Essays on Financial Markets and the Macroeconomy

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Dr. rer. pol.)
im Fach Volkswirtschaftslehre

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

von
Herr Dipl.-Vw. Emanuel Mönch
geboren am 11.02.1977 in Halle/Saale

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Christoph Markschies

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Oliver Günther, Ph.D.

Gutachter:

1. Prof. Harald Uhlig, Ph.D.
2. Prof. Bartosz Maćkowiak, Ph.D.

eingereicht am:	28. Juni 2006
Tag der mündlichen Prüfung:	20. Oktober 2006

Abstract

This thesis consists of four essays of independent interest which make empirical and methodological contributions to the fields of financial economics and macroeconomics. The first essay deals with the proper specification of investors' information set in tests of conditional asset pricing models. In particular, it advances the use of dynamic factors as conditioning variables. By construction, dynamic factors summarize the information in a large number of variables and are therefore intuitively appealing proxies for the information set available to investors. The essay demonstrates that this approach substantially reduces the pricing errors implied by conditional models with respect to traditional approaches that use individual indicators as instruments. Following previous evidence that the central bank uses a large set of conditioning information when setting short-term interest rates, the second essay employs a similar insight in a model of the term structure of interest rates. Precisely, the dynamics of the short-term interest rate are modelled using a Factor-Augmented Vector-Autoregression. Based on this dynamic characterization of monetary policy, the term structure of interest rates is derived under the assumption of no-arbitrage. The resulting model is shown to provide superior out-of-sample forecasts of US government bond yields with respect to a number of benchmark models. The third essay analyzes the predictive information carried by the yield curve components level, slope, and curvature within a joint dynamic factor model of macroeconomic and interest rate data. The model is estimated using a Metropolis-within-Gibbs sampling approach and unexpected changes of the yield curve components are identified employing a combination of zero and sign restrictions. The analysis reveals that the curvature factor is more informative about the future evolution of the yield curve and of economic activity than has previously been acknowledged. The fourth essay provides a monthly business cycle chronology for the Euro area. A monthly series of Euro area real GDP is constructed using an interpolation routine that nests previously suggested approaches as special cases. Then, a dating routine is applied to the interpolated series which excludes business cycle phases that are short and flat.

Keywords:
Financial economics, macroeconomics, applied econometrics, asset pricing, term structure of interest rates, dynamic factor models, business cycle dating

Zusammenfassung

Diese Arbeit besteht aus vier Essays, die empirische und methodische Beiträge zu den Gebieten der Finanzmarktökonomik und der Makroökonomik liefern. Der erste Essay beschäftigt sich mit der Spezifikation der Investoren verfügbaren Informationsmenge in Tests bedingter Kapitalmarktmodelle. Im Speziellen schlägt es die Verwendung dynamischer Faktoren als Instrumente vor. Diese fassen per Konstruktion die Information in einer Vielzahl von Variablen zusammen und stellen daher intuitive Maße für die Investoren zur Verfügung stehenden Informationen dar. Es wird gezeigt, dass so die Schätzfehler bedingter Modelle im Vergleich zu traditionellen, auf einzelnen Indikatoren beruhenden Modellvarianten substantiell verringert werden. Ausgehend von Ergebnissen, dass die Zentralbank zur Festlegung des kurzfristigen Zinssatzes eine große Menge an Informationen berücksichtigt, wird im zweiten Essay im Rahmen eines affinen Zinsstrukturmodells eine ähnliche Idee verwandt. Speziell wird die Dynamik des kurzfristigen Zinses im Rahmen einer Faktor-Vektorautoregression modelliert. Aufbauend auf dieser dynamischen Charakterisierung der Geldpolitik wird dann die Zinsstruktur unter der Annahme fehlender Arbitragemöglichkeiten hergeleitet. Das resultierende Modell liefert bessere Vorhersagen US-amerikanischer Anleihenzinsen als eine Reihe von Vergleichsmodellen. Der dritte Essay analysiert die Vorhersagekraft der Zinsstrukturkomponenten “level”, “slope” und “curvature” im Rahmen eines dynamischen Faktormodells für makroökonomische und Zinsdaten. Das Modell wird mit einem Metropolis-within-Gibbs Sampling Verfahren geschätzt, und Überraschungsänderungen der drei Komponenten werden mit Hilfe von Null- und Vorzeichenrestriktionen identifiziert. Die Analyse offenbart, dass der "curvatureFaktor informativer in Bezug auf die zukünftige Entwicklung der Zinsstruktur und der gesamtwirtschaftlichen Aktivität ist als bislang vermutet. Der vierte Essay legt eine monatliche Chronologie der Konjunkturzyklen im Euro-Raum vor. Zunächst wird mit Hilfe einer verallgemeinerten Interpolationsmethode eine monatliche Zeitreihe des europäischen BIP konstruiert. Anschließend wird auf diese Zeitreihe ein Datierungsverfahren angewandt, das kurze und flache Konjunkturphasen ausschließt.

Schlagwörter:
Finanzmarktökonomik, Makroökonomik, angewandte Ökonometrie, Aktienbewertungsmodelle, Zinstrukturmodelle, dynamische Faktormodelle, Datierung von Konjunkturzyklen

Acknowledgements

This thesis is the result of my work at the Institute of Economic Policy at Humboldt University Berlin over the last three and a half years. During this time, I have enjoyed support of many different people to all of whom I am much obliged. Some stand out, however, and their contribution deserves individual mention.

First of all, I would like to express great gratitude to my thesis supervisor Professor Harald Uhlig. He significantly stimulated my interest in statistical methods and their application to macroeconomics and finance. Not only has he directly contributed to this thesis as the co-author of "Towards a Monthly Business Cycle Chronology for the Euro Area" (Chapter 5, published in *Journal of Business Cycle Measurement and Analysis* **2(1)**, May 2005). He also had an impact on all other chapters via numerous comments and suggestions that he pointed out to me in seminars and personal discussions. Apart from the academic advice, I further thank him for his support in finding external funding and for employing me in the Sonderforschungsbereich 649.

Professor Bartosz Maćkowiak has been co-supervising this thesis. I was extremely lucky to find with him someone at our department who shares my interest in factor models. I benefitted enormously from discussing with him all different kinds of questions and I still enjoy working on a joint project.

I will remember the time of my thesis as an enjoyable and fruitful stage in my life. Much of this is owed to my office mates and colleagues with whom I shared a lot of fun and many interesting discussions. In writing the final draft of this thesis, I have benefitted from helpful comments and careful proofreading by Martin Kliem, Stefan Ried, Samad Sarferaz, Almuth Scholl, Andreas Schrimpf from ZEW Mannheim, and in particular Christian Stoltenberg. My research requires good computational facilities. I am therefore grateful to Georg Man who did a lot to make our UNIX machines run. A similar thanks goes to Holger Gerhardt who moreover helped me with layout issues. For the most part, however, the professional appearance of my thesis is the merit of Jenny Simon who helped me turning four papers into one document. I would further like to thank our secretary Susann Roethke for silently yet so efficiently doing all the important administrative work behind the scenes.

I did most of the work on Chapter 3 as a visitor to the European Central Bank. Being at the ECB offered me the opportunity to discuss my work with outstanding researchers, and sharpened my understanding of which economic questions are of interest to practitioners. The publication of an earlier version of chapter 3 in the ECB Working Paper Series has considerably enhanced the reach of this study. Among the many nice people I met at the ECB, two have been particularly important to me: PETER HÖRDAHL and MANFRED KREMER. I thank both of them for their support and advice.

Much of my thesis deals with dynamic factor analysis for large cross-sections and therefore requires the availability of big datasets. I am highly indebted to ROBERT BLISS, LUCREZIA REICHLIN, and MARK WATSON for generously sharing their data with me or for providing them on the internet.

The chapters of this thesis and earlier drafts have been presented at several occasions. In addition to various presentations in the Macro Brownbag Seminar at Humboldt University, I have benefitted from comments received at the doctoral workshops of the European Finance Association in Glasgow and the German Finance Association in Mainz, the Association Française de Finance Meeting in Paris, the "Macroeconomic Issues in the EMU" workshop in Brescia, the Spring Meeting of Young Economists in Warsaw, the European Meeting of the Financial Management Association in Zurich, the Econometric Society European Meeting in Madrid, research seminars at the ECB and the Bundesbank, the Econometric Society World Congress in London, the European Finance Association Meeting in Moscow, and at Barclays Global Investors in London.

Man does not live by science alone. I am therefore highly indebted to the STUDIENSTIFTUNG DES DEUTSCHEN VOLKES for funding my doctoral studies for as long as three years. I am further grateful to the DEUTSCHE FORSCHUNGSGEMEINSCHAFT for complementary funding through the SFB 649.

All the direct support mentioned so far could not have developed its value without the mental backing by my friends and family. Above all, MY PARENTS and OMA ILSE are a constant source of energy through their unconditional trust in me and my abilities. Finally, a special word of thanks is reserved for TINA who shared all my ups and downs along the way. Her cheerfulness makes everything so much easier.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Scope of the Study

Ever since I started studying economics, I wanted to understand how financial markets work and how they interact with the macroeconomy. Over the last three and a half years, I have had a first chance to explore this topic. This thesis is a progress report of my exploration.

Financial and macroeconomic theory have developed independently of one another for a long time. As a consequence, the benchmark models put forth by the two strands of economic research are strikingly disconnected. In macroeconomics, this is mainly reflected by the fact that models which successfully replicate the dynamics of key economic aggregates largely fail to explain observed patterns of asset prices. A popular example for such a failure is the equity premium puzzle which denotes the problem of standard macroeconomic models to explain the high average returns of stocks as compared to bonds. In financial economics, the separation from macroeconomic theory is most obviously notable from the fact that traditional asset pricing models completely ignore the information carried by non-financial variables. Famous witnesses of this disregard still build the core of most finance textbooks: the Capital Asset Pricing Model, the Arbitrage Pricing Theory, and various latent factor models of the term structure of interest rates.

Lately, the strict demarcation between pure finance and pure macro models has somewhat broken up. In particular, a number of recent studies have successfully employed macroeconomic variables in models of the cross-section of stock returns and models of the term structure of interest rates. Macroeconomic information enters both types of models via different channels. On the one hand, state of the art cross-sectional asset pricing models assume a time-varying relationship between the stochastic discount factor and the pricing factors that summarize the fundamental sources of risk faced by investors. The time-variation in the discount factor specification is inherently linked to investors' conditional expectations about future returns. As these expectations are unobservable, they are commonly modeled to depend on macroeconomic variables which are assumed to proxy for the information set available to investors. The recent term structure literature, on the other hand, mainly incorporates macroeconomic information via a formulation of monetary policy. In particular, affine term structure models use no-arbitrage arguments to develop the yield curve starting from a specification of the short-term interest rate. The latter is adopted from benchmark macroeconomic models which map the short rate to output and inflation via some monetary policy reaction function. Accordingly, the macroeconomic indicators that are used to summarize the central bank's information set become state variables in the term structure model. In sum, the assumptions made about the information employed by investors or the central bank define which macroeconomic variables enter the two types of models.

The particular structure of state of the art models of the cross-section of returns and the term structure of interest rates allows to include only few indicators as proxies for the information set available to investors or the central bank. Finding sensible instruments therefore represents a common theme in much of the current macro-finance literature. Yet, both investors and the central bank have access to huge amounts of information. As a consequence, asset pricing studies that employ only few instruments neglect potentially important information and are thus prone to the risk of misspecification. In my dissertation, I therefore go beyond this common practice by suggesting ways to explicitly incorporate broad macroeconomic information sets in the two types of asset pricing models. My approach is based on dynamic factor analysis for large datasets which attracts a good deal of interest in contemporary empirical macroeconomics. Recent research on dynamic factor models has shown that the information in a large number of economic time series can efficiently be summarized by a few factors. Accordingly, these factors are natural and intuitively appealing proxies for the information set available to investors or the central bank. In Chapters 2 and 3 of my thesis, I exploit this feature and show that the use of large macroeconomic datasets can substantially improve the performance of asset pricing models.

The information linkages between macroeconomic and financial variables are not unidirectional. While macroeconomic variables have only recently found their way into asset pricing, financial variables are traditionally employed as forward-looking indicators of economic activity. For example, spreads between long-term and short-term interest rates are often used to predict recessions. However, the yield curve carries more information than what is captured by interest rate spreads. Indeed, the financial literature commonly decomposes the term structure of interest rates into three factors - “level”, “slope”, and “curvature” - which together summarize almost all of the variation across yields. Although this factor decomposition is extensively used in asset pricing applications, there is only scattered evidence on the predictive content of the three factors for future economic activity. Therefore, Chapter 4 of my thesis analyzes the information carried by the three yield curve components. In particular, I study whether unexpected changes of the factors capture important news about the future evolution of output and other economic variables.

On a more general basis, forecasting the beginning and end of recessions is one of the most prominent tasks of economists in applied research and practice. Other than the use of interest rate spreads, there exists a variety of different methods to predict economic turning points. Compared to the comprehensive repertory of recession prediction models, however, surprisingly little research effort has been devoted to the proper definition of business cycle phases and what marks their beginning and end. In the US, an official chronology of business cycle turning points is maintained by the National Bureau of Economic Research, NBER. The dating decision is reached by a group of economists in a judgmental process based on the evolution of different economic indicators. The introduction of the Euro has created a similar need for the Euro area economy. The Centre for Economic Policy Research, CEPR, has therefore recently started to publish a chronology of business cycle turning points, employing a procedure akin to the one used by the NBER. Different from the NBER, however, the CEPR announces the beginning and end of recessions only on a quarterly basis. Yet, since a wide range of empirical analyses employ the state of the economy as an indicator, it is useful to have a monthly chronology of business cycle turning points available. Chapter 5 of my thesis derives such a chronology for the Euro area.

Altogether, much of my thesis ties in with three popular themes in the contemporary financial and macroeconomic literature: the cross-section of stock returns, the term-structure of interest rates, and dynamic factor models for large datasets. In the following, I provide a brief review of the literature on these three issues. I then summarize the results of my thesis in relation to the previous work.

1.2 Literature Review

This section reviews the existing literature on the cross-section of stock returns, the term structure of interest rates, and dynamic factor models for large datasets. As there are almost no intersections between the three fields of research, I treat them separately. For ease of exposition, I report on the main contributions to each theme in a chronological order.

1.2.1 The Cross-Section of Stock Returns

It is a well-documented fact that returns differ across stocks and across portfolios. A vast literature tries to identify the systematic patterns behind these cross-sectional return differences. In the following, I provide a brief chronological review of the main contributions to this literature.

The Capital Asset Pricing Model (CAPM) of Sharpe [1964], Lintner [1965] and Black [1972] was the first model that provided a quantitative explanation of cross-sectional return differences. The main result of the CAPM is that if all investors hold mean-variance efficient portfolios in the sense of Markowitz [1952], then the market portfolio is also mean-variance efficient. A consequence of this finding is that expected returns of individual stocks or portfolios depend on their exposure to the market risk. The latter is commonly measured as the coefficient in a linear regression of the individual stock return on the market return and is mostly labeled the market *beta*. According to the CAPM, the expected return of a stock therefore depends on the expected return of the market portfolio and the stock's beta.

The CAPM has been derived under the assumption that investors choose their mean-variance efficient portfolios within a static framework where asset returns have constant means and variances. In its original form, the CAPM thus abstracts from changes in the investment opportunity set. This critique was first pointed out by Merton [1973] who in response developed a dynamic model of asset allocation, commonly labeled the Intertemporal CAPM (ICAPM). Solving his model, Merton derived an asset pricing formula which implies that the expected return of an asset depends on two components: its exposure to the market risk and its comovement with a portfolio that hedges against future changes of the set of investment opportunities. Merton assumes the latter to be captured by some unspecified state variables. Therefore, Fama [1991] referred to the ICAPM as a "fishing license". Indeed, Merton's theory has given rise to various ad-hoc pricing models that are based on empirical risk factors which are assumed to proxy for changes in investment opportunities. One early example for such an ad-hoc approach was the article by Chen, Roll, and Ross [1986]. These authors showed that macroeconomic indicators such as the term spread, the default spread, unexpected inflation, and production growth have incremental explanatory power over the market return in cross-sectional regressions.

Breeden [1979] provided an important simplification of Merton's (1973) model. He showed that the pricing formula of the ICAPM can be transformed into a single-beta representation where the beta captures the asset's comovement with aggregate consumption. Breeden's model is therefore known as the Consumption Capital Asset Pricing Model (CCAPM). Yet, while the CCAPM represents an important milestone in asset pricing theory, it has largely failed empirical tests (see e.g. Lettau and Ludvigson 2001).

Another strand of research on the cross-section of stock returns has taken a statistical approach. The Arbitrage Pricing Theory (APT) of Ross [1976] represents the seminal contribution in this field. The central idea of the APT is that in equilibrium, idiosyncratic risk will not be priced since investors can diversify it away by holding portfolios. Therefore, only systematic risk shared by all assets in the market will carry premia. If there are such systematic sources of risk, asset returns will have a factor structure. As a consequence, expected returns of individual assets are approximately linear functions of factor loadings. Chamberlain and Rothschild [1983] showed that Ross' (1976) result prevails when returns have an approximate factor structure, i.e. when their idiosyncratic components are mildly cross-correlated. Connor and Korajczyk [1986] extended these results and developed a procedure for factor extraction based on principal components that accommodates approximate factor structures. More recently, Jones [2001] provides an extension of Connor and Korajczyk's method which is robust to heteroskedasticity of returns. While potentially delivering good in-sample estimates of the common factors underlying asset returns, the APT does not provide economic explanations for the observed cross-sectional return differences. It has therefore received little attention in recent years.

The empirical failure of the CCAPM, the purely statistical nature of the APT, and the adhocness of empirical specifications of the ICAPM have bestowed a long life upon the standard CAPM. In the 1990s, however, the model as been seriously challenged by the work of Fama and French. First, Fama and French [1992] showed that a firm's market capitalization and the ratio of a firm's book to its market value together capture an important share of the cross-sectional variation in average stock returns. Their analysis was inspired by previously documented "CAPM anomalies", i.e. failures of the CAPM when applied to portfolios sorted by firm characteristics. For example, Basu [1977] reported that firms with low price-earnings ratios tend to have higher returns than their market betas would suggest. Furthermore, Banz [1981] documented that small capitalization stocks on average earn returns that are higher than what the CAPM would predict. After noting that size and book-to-market play the most prominent role in explaining these anomalies, Fama and French [1993] constructed a set of portfolios designed to mimic the two factors. Together with the market return, these mimicking portfolios exhibit a strikingly good ability to explain average stock returns. Consequently, the Fama-French three-factor model has lately succeeded the CAPM as the benchmark asset pricing model.

Yet, the model is only empirically motivated and provides no insight about the economic foundations of the size and book-to-market factors. Recent research has therefore tried to give the Fama-French factors an economic interpretation. Liew and Vassalou [2000], for example, show that the size and book-to-market factors help forecast GDP growth. Vassalou [2003] finds that much of the information in the two factors is news related to future GDP growth. More recently, Petkova [2005] provides empirical evidence that the Fama-French factors are correlated with innovations of a set of variables that describe investment opportunities, e.g. the dividend yield, the term spread, and the default spread. Altogether, though, the two Fama-French factors still lack a clear-cut economic interpretation.

As a consequence, the CAPM has lately attracted renewed interest. In particular, many authors have focused on conditional formulations of the model. In this framework, an asset's expected return in period t depends on the expected return on the market portfolio and its expected market beta in period $t+1$, where the expectations are formed based on all information available up to period t. Jagannathan and Wang [1996] were the first to suggest a conditional version of the CAPM. They showed that their conditional model implies an unconditional model with an additional risk factor. Cochrane [1996] more generally demonstrated how conditioning information can be included in asset pricing models and how such models translate into unconditional models that can be estimated using standard statistical methods. Employing Cochrane's setup, Lettau and Ludvigson [2001] more recently find strikingly strong support for conditional versions of the CAPM and the CCAPM. Based on prior evidence that the log consumption-wealth ratio forecasts excess returns, they use this variable as a proxy for investors' information set. Subsequent to Lettau and Ludvigson's influential study, a number of different conditioning variables have been proposed. Two examples are Santos and Veronesi [2006] who show that the labor income to consumption ratio is a useful instrument and Lustig and Van Nieuwerburgh [2005] who suggest the housing wealth to human wealth ratio as conditioning variable.

Despite the recent popularity of conditional pricing models in the spirit of Lettau and Ludvigson [2001], the approach suffers from the fundamental problem of employing individual variables as proxies for the information set available to investors. This is obviously a strong assumption. In Chapter 2 of my thesis, I therefore suggest a new strategy of testing conditional asset pricing models on the basis of large sets of conditioning information.

1.2.2 The Term-Structure of Interest Rates

The term structure of interest rates commonly denotes the cross-section of zero-coupon government bond yields ordered by maturity. Plotting combinations of maturities and yields at a given point in time usually reveals a curve-shaped functional link. Accordingly, "yield curve" is often used as a synonym for term structure of interest rates. The following section provides a brief review of the main cornerstones of term structure research over the last three decades.

A central result of asset pricing theory is that the price of a financial asset must equal its discounted expected future payoff. This general principle translates into a specific relationship between the prices of bonds of different maturity. In particular, it implies that the price of an n-periods to maturity zero-coupon bond in period t equals the discounted expected price of an $(n-1)$-periods to maturity bond in $t+1$. Hence, only two ingredients are needed to derive the prices of bonds of all maturities: the formulation of a time-series process for the stochastic discount factor, and a specification of the one-month interest rate.

Vasicek [1977] was the first who suggested a model that has these two features. In the Vasicek model, the stochastic discount factor is driven by the short-term interest rate as the only state variable. Accordingly, it belongs to the class of one-factor models of the term structure. The short-rate itself is modeled as a diffusion with parameters chosen such that closed-form solutions for bond prices of all maturities can be derived.

Cox, Ingersoll, and Ross [1985] suggested two important extensions to the Vasicek model. First, by inserting square root terms in the volatility of the short rate, they accounted for the observation that higher interest rates are more volatile. Second, their specification keeps the short rate from falling below zero, a feature that is obviously desirable since nominal interest rates cannot be negative.

Although providing a good starting point, the models by Vasicek [1977] and Cox et al. [1985] turned out to be inconsistent with a number of empirical facts. For example, if the model parameters are calibrated so as to match the auto-correlation of the short rate, then the implied average yield curve is considerably less concave than in the data. As one strategy to better fit observed yield curves, Ho and Lee [1986] introduced additional time-dependent adjustment terms in the discount factor specification. Heath, Jarrow, and Morton [1992] further extended this approach, allowing for state variables with time-varying volatility. This feature made their model particularly useful for the evaluation of interest rate sensitive contingent claims such as bond options.

Despite these generalizations, one-factor models appear to be unable to replicate the rich dynamic patterns of the yield curve, see Backus, Foresi, and Telmer [1998] for a nice review. Therefore, term structure research has recently focussed on multi-factor models of the yield curve. Duffie and Kan [1996] developed a generalized class of models which nests multi-factor versions of the models by Vasicek [1977] and Cox et al. [1985] as special cases. In this model class, yields are linear functions of a vector of state variables and hence its members are commonly referred to as "affine" models of the term structure. Dai and Singleton [2000] and Duffee [2002] complement Duffie and Kan's (1996) framework by providing useful results on the specification and estimation of multi-factor affine models.

Apart from the modeling approaches based on no-arbitrage principles, another strand of research has taken a statistical approach to term structure analysis. In particular, a number of authors have employed factor analysis techniques to decompose the variation across yields into a few components. Nelson and Siegel [1987], for example, approximated the forward curve using a three-factor decomposition with loadings given by exponential functions of the time-to-maturity and a shape parameter. Litterman and Scheinkman [1991] and Knez, Litterman, and Scheinkman [1994] performed classical factor analysis of yields and also found that three factors explain almost all of their cross-sectional variation. According to the shape of the factor loadings, Litterman and Scheinkman [1991] identified the three factors as the "level", "steepness", and "curvature" of the yield curve. This nomenclature has largely survived, with "steepness" nowadays mostly being replaced by "slope".

The preceding summary reveals that traditional term structure analysis completely neglects the information carried by observable macroeconomic variables. On the one hand, affine models commonly assume exogenous time series processes for the state variables. On the other hand, statistical factor decompositions focus on the estimation of unobserved components that explain as much as possible of the cross-sectional variation of yields. Yet, since the short-rate is set by the central bank in response to economic fluctuations, there must be some link between the term structure and the macroeconomy. In a widely recognized paper, Ang and Piazzesi [2003] explicitly model this transmission channel by specifying the short-term interest rate in the spirit of a Taylor-type monetary policy reaction function. Accordingly, output and inflation become observable states which complement the three latent factors of the standard affine model. Quite strikingly, Ang and Piazzesi find that the two macroeconomic variables account for a large share in the cross-sectional variation of yields.

This result has triggered a lot of ongoing research in the macro-finance term structure literature. So far, only few papers have been published, but the number of studies is constantly growing. Two prominent contributions are due to Hördahl, Tristani, and Vestin [2006] and Diebold, Rudebusch, and Aruoba [2006]. Hördahl et al. [2006] build a small structural model that describes the joint evolution of output, inflation, and the short rate, and add the term structure using no-arbitrage restrictions. In contrast to the study by Ang and Piazzesi [2003], their model comprises only one latent state variable which they structurally interpret as the inflation target pursued by the central bank. Hördahl et al. [2006] provide evidence that their model delivers better out-of-sample forecasts of yields than Ang and Piazzesi's approach. A common feature of the models by Ang and Piazzesi [2003] and Hördahl et al. [2006] is the unidirectional link between macroeconomic variables and the yield curve. Diebold et al. [2006] therefore study a model which allows for interaction in both directions. Adopting the decomposition of yields suggested in Diebold and Li [2006], they find evidence for a bidirectional feedback between the yield curve and a set of macroeconomic variables.

This result is in line with a strand of the literature that highlights the usefulness of interest rate spreads as predictors of economic activity. Estrella and Hardouvelis [1991], Estrella and Mishkin [1998] and a number of other authors provide regression-based evidence that yield spreads can successfully be employed to forecast recessions. Yet, the previously documented factor structure of the term structure suggests that the yield curve carries more information than what is captured by its slope alone. So far, however, little effort has been devoted to investigating the predictive information contained in the other components, in particular curvature. In Chapter 4, I therefore perform a systematic analysis of the informational content of each of the three yield curve factors.

Similar to factor models of the cross-section of stock returns, degrees of freedom problems restrict the number of factors that can be incorporated in state-of-the-art term structure models. Accordingly, much of the recent term structure literature deals with model specification issues such as the selection of appropriate macroeconomic state variables. As discussed above, the choice of states is determined by the specification of the short-rate. A recent literature finds evidence that central banks set interest rates in response to the information contained in many variables. Accordingly, term structure models built upon short-rate equations that are based on only few macroeconomic indicators are potentially misspecified. In Chapter 3 of my thesis, I therefore suggest a term structure model which parsimoniously incorporates a large macroeconomic information set. My approach builds on recent advances in dynamic factor analysis. The next section provides a brief review of the main contributions in this domain of macroeconomic research.

1.2.3 Dynamic Factor Models for Large Datasets

Factor models are statistical tools that can be used to decompose a set of observed variables into unobserved common and idiosyncratic components. The term "dynamic factor model" has different usages. It mostly denotes a factor model for time series variables. However, it is also sometimes used to highlight the inclusion of lagged factors in the observation equation in contrast to "static" factor models which only have contemporary factor values in the observation equation.

While factor models have a long tradition in other social sciences, they have only recently become a popular tool among macroeconomists. Yet, a few early contributions have inspired much of the subsequent work, in particular the studies by Geweke [1977] and Sargent and Sims [1977] who analyze the common dynamics among a small number of economic time series using frequency domain approaches. Engle and Watson [1981], Sargent [1989], and Stock and Watson [1991] estimate dynamic factor models for small sets of variables in the time domain using maximum likelihood methods. Based on the EM algorithm, Quah and Sargent [1993] extend this approach in order to accommodate as much as 60 time series. A common feature of these studies is the assumption of an "exact" factor structure meaning that the idiosyncratic components are modeled to be uncorrelated across variables.

As has been mentioned above in the context of the Arbitrage Pricing Theory (APT), a branch of financial research has developed techniques to extract factors from a cross-section of returns under the assumption of an "approximate" factor structure, i.e. allowing for some mild cross-correlation of the idiosyncratic components. The main contribution in this field is Connor and Korajczyk [1986] who show that the factors in an approximate factor model are consistently estimated by the first principal components of the data when the number of cross-sectional elements N goes to infinity and when the number of time series observations T is fixed.

Recently, Stock and Watson [1998] have attracted a great deal of attention among empirical macroeconomists with a working paper that has subsequently been published in the twin articles Stock and Watson [2002a,b]. Their contribution is twofold. First, they provide a number of methodological refinements of the principal components approach of Connor and Korajczyk. In particular, they show consistency of the estimated factors when both the number N of cross-sectional observations and the number T of time series observations tend to infinity, assuming $T/N \to 0$ and allowing for time-varying factor loadings. As a second main contribution, Stock and Watson use factors extracted from a large dataset to forecast various economic time series. Strikingly, they find that factor-based forecasts deliver significantly better out-of-sample predictions of measures of output and inflation than various benchmark forecasting models. Note that the principal components estimates are similar to cross-sectional averages of many variables with weights chosen so as to minimize the sum of squared idiosyncratic components. Accordingly, Stock and Watson label them "diffusion indexes", a term that remains a common synonym for factors extracted from large cross-sections of time series.

The strong forecast performance of diffusion indexes has raised a lot of interest in dynamic factor modeling. Besides the many empirical applications of Stock and Watson's estimation approach, there are a number of recent methodological advances. Most importantly, Bai and Ng [2002] show consistency of the estimated factors without imposing a restriction on the relation between N and T. They further propose some panel information criteria to estimate the number of common factors in a large dataset. More recently, Bai [2004] extends the methodology to accommodate nonstationary data, and Bai and Ng [2006] derive asymptotic confidence intervals for diffusion index forecasts.

While Stock and Watson analyze comovement of economic variables in the time domain, a parallel literature has taken a frequency domain approach. Forni and Reichlin [1998] show that the number of common factors in a panel of economic time series can be determined by applying principal components analysis to the spectral density matrix of the cross-sectional averages of all variables in the panel. Yet, they estimate the common shocks using a structural VAR technique. Forni, Hallin, Lippi, and Reichlin [2000] demonstrate that the common factors in an approximate dynamic factor model can be consistently estimated by the first principal components of the spectral density matrix of the data. An important drawback of their method is that the estimated factors are infeasible for forecasting purposes since they are obtained using two-sided filters. To overcome this problem, Forni, Hallin, Lippi, and Reichlin [2005] develop an estimation technique that is based on one-sided filtering. However, due to its computational simplicity the method suggested by Stock and Watson remains more widely used in practice.

Factor models allow to predict individual economic time series using the information contained in a large cross-section of data. Bernanke and Boivin [2003] first note that this qualifies dynamic factors for application in monetary policy analysis. In particular, they argue that central banks actively monitor a large number of economic variables, and therefore likely set the short-term interest rate based on information beyond output and inflation. In empirical tests, Bernanke and Boivin find evidence that monetary policy reaction functions based on estimated factors outperform standard Taylor-rule specifications. This result is confirmed by Favero, Marcellino, and Neglia [2005] who employ different datasets and factor extraction methods. Moreover, Giannone, Reichlin, and Sala [2004] show that forecasts of the federal funds rate are significantly improved by using a factor model approach.

Dynamic factor models can in principle be estimated using the Kalman filter and maximum likelihood. Yet, this approach becomes infeasible when the number of variables in the dataset is large. As a consequence, Stock and Watson and other authors apply principal components techniques. However, this approach does not allow joint determination of the common factors and the parameters governing their dynamics. Recently, some authors have therefore suggested to use Markov Chain Monte Carlo (MCMC) techniques - a branch of Bayesian statistics - to estimate dynamic factor models. The seminal contributions in this field are Otrok and Whiteman [1998] and Eliasz [2002]. While Otrok and Whiteman consider a model with only one factor but serially correlated idiosyncratic components, Eliasz studies a model with multiple factors and no autocorrelation in the error terms. Both authors estimate their factor models using Gibbs sampling techniques. The Gibbs sampler allows approximation of the unknown joint posterior distribution of a set of parameters by alternately sampling from their conditional posteriors. As dynamic factor models can be written in state-space form, derivation of conditional distributions of the latent factors and the parameters governing their dynamics is straightforward. Lately, a vast literature has emerged that applies Bayesian factor model techniques to a variety of economic and financial questions, but so far only a few papers have been published. Kose, Otrok, and Whiteman [2003], for example, use a generalization of the model in Otrok and Whiteman [1998] to identify the common factors among macroeconomic time series on a global, regional and country level. Bernanke, Boivin, and Eliasz [2005] extend the approach of Eliasz [2002] to accommodate observable variables as additional factors, a technique which they label "Factor-Augmented Vector-Autoregression" (FAVAR).

In Chapter 3 of my thesis, I employ the FAVAR approach by Bernanke et al. [2005] as the state equation of an affine term structure model which I estimate using classical statistical methods. In Chapter 4, I set up a dynamic factor model of macroeconomic and interest rate data to study the informational content of yield curve news about the future evolution of key macroeconomic variables. This model is estimated using a Gibbs sampling approach.

1.3 Outline of the Thesis

My thesis consists of four main chapters each of which represents a study of independent interest. In Chapter 2, I propose a strategy to incorporate a broad set of conditioning information in tests of conditional asset pricing models. I document that my approach substantially reduces the pricing errors implied by such models. In Chapter 3, I suggest an affine term structure model that parsimoniously exploits a large macroeconomic information set and show that this model significantly outperforms a number of benchmarks in out-of-sample forecasts of the yield curve. In Chapter 4, I then study the informational content of yield curve surprises within a Bayesian dynamic factor model. The main finding of this analysis is that the curvature factor carries important predictive information about the future course of the economy. Finally, Chapter 5 contains an exercise in dating the Euro area business cycle. The following paragraphs provide a fast guide to the main contributions made in each of the four chapters. Details on the employed estimation approaches and on implementation issues are given in the Technical Appendix.

Conditional asset pricing models relate the observed prices of financial assets to investors' conditional expectations about future payoffs. Accordingly, they require the specification of investors' expectations and hence of their information set. Since this information set is intrinsically unobservable, it is common to use proxies that are known to predict returns, for example the dividend yield, the term spread or the log consumption-wealth ratio. Conditional pricing models are thus traditionally tested on the basis of very little conditioning information and are therefore potentially misspecified.

In Chapter 2, I suggest a solution to this problem which explicitly exploits a broad macroeconomic information set. My approach, outlined in Section 2.2, is based on dynamic factor analysis for large datasets which allows to efficiently summarize the information in many variables by a few estimated factors. Sections 2.3 and 2.4 summarize the methodology and the data that I use to empirically assess my approach. Section 2.5 documents the outcomes of various tests which show that the pricing errors implied by the conditional CAPM are substantially reduced when dynamic factors instead of commonly used conditioning variables are employed as instruments. This is a strong result which casts doubt upon the common practice of testing conditional models on the basis of individual instruments. The findings summarized in Table 2.4 in Section 2.5.3 underscore this conclusion by showing that diffusion indexes exhibit incremental explanatory power over the best-performing benchmark instruments while the latter are found to be insignificant when added to diffusion index based specifications.

Some benchmark instruments are only available at the quarterly frequency, most importantly the log consumption-wealth ratio suggested by Lettau and Ludvigson [2001]. Comparing dynamic factors extracted from a quarterly panel with this popular conditioning variable in Section 2.5.4, I also find that diffusion indexes imply substantially smaller pricing errors. This indicates that the log consumption-wealth ratio - similar to all other studied benchmark instruments - misses important variation in conditional moments of returns. Finally, Section 2.5.5 documents that my approach withstands a number of robustness tests. Altogether, the results of Chapter 2 show that dynamic factors are better proxies for investors' information set than previously suggested individual instruments. This finding carries important implications for the specification of conditional pricing models in applied research and practice.

In Chapter 3, I apply dynamic factor model techniques in a different asset pricing context. In particular, I incorporate a large macroeconomic information set in an affine term structure model which I use to forecast the yield curve. The background of this analysis is the following. On the one hand, a recent literature includes individual macroeconomic indicators, in particular measures of output and inflation, as state variables in term structure models. These variables enter the models via a Taylor-rule formulation of monetary policy. On the other hand, specifications of monetary policy that exploit large macroeconomic information sets have recently been shown to empirically outperform reaction functions based on output and inflation alone. Quite intuitively, this result has a bearing on term structure analysis. In Chapter 3, I elaborate on this insight and show how a large macroeconomic information set can be incorporated in a yield curve model. Specifically, I use a Factor-Augmented Vector-Autoregression (FAVAR) suggested in Bernanke et al. [2005] as the state equation in an affine model. The FAVAR model describes the evolution of the short-term interest rate conditional on a large information set. Based on this dynamic short rate specification, my model builds up the yield curve using parameter restrictions implied by no arbitrage. Accordingly, I label my approach a "No-Arbitrage Factor-Augmented VAR".

An earlier version of Chapter 3 has appeared as No. 544 in the *ECB Working Paper Series*. The chapter is organized as follows. In Section 3.2, I outline the details of my term structure model and show how it is estimated in Section 3.3. The estimation results are summarized in Section 3.4. Some preliminary evidence that diffusion indexes contain useful information about the federal funds rate and yields of higher maturity is provided in Section 3.4.4. I then show in Section 3.4.5 that my model captures the cross-sectional variation of interest rates well in-sample.The main focus of the chapter is on yield curve prediction, however. In Section 3.5, I document the results of an out-of-sample forecast exercise for yields of various maturities at horizons from one to twelve months ahead. Strikingly, my approach produces significantly better out-of-sample forecasts of government bond yields than a number of successful competitor models which are summarized in Section 3.5.1. This important result becomes most apparent from Tables 3.4 and 3.5. As a partial explanation for the strong predictive power of my model, Section 3.5.3 documents a close link between the yield curve components level and slope and the macroeconomic factors of my model.

Chapter 4 studies the linkages between the term structure and the macroeconomy from a reverse angle by analyzing the predictive content of the yield curve for key macroeconomic variables. The motivation for this analysis is as follows. While interest rate spreads are often used as predictors of recessions, the traditional decomposition of yields into the three factors level, slope, and curvature suggests that the term structure carries more information than what is captured by yield differentials. However, there is yet little evidence about the predictive content of each of the three yield curve components. Chapter 4 therefore seeks to provide a systematic analysis of the information carried by unexpected changes of the three factors. To this end, I set up a dynamic factor model of macroeconomic and interest rate data that is outlined in Section 4.2. I adopt the decomposition of yields into level, slope, and curvature recently suggested by Diebold and Li [2006]. The three yield components share unrestricted dynamics with a set of factors that summarize the comovement in a number of macroeconomic variables. Hence my model setup allows to study the interaction between the yield curve and the macroeconomy within a rich parametric framework. This represents a major improvement over the existing literature on the joint dynamics of macroeconomic variables and interest rates, in particular the study by Diebold et al. [2006].

Estimation of my model is via a Metropolis-within-Gibbs sampling algorithm. A Metropolis step needs to be added to the standard Gibbs sampling procedure in order to draw from the nonstandard distribution of the shape parameter in the yield factor loadings. Section 4.3 briefly discusses the general operating mode of my approach. A detailed treatment of the individual steps carried out in the estimation process is provided in Appendix 6.3. For reasons broadly discussed in Chapter 4, a recursive identification of the yield curve shocks is inappropriate for the problem at hand. To identify unexpected changes of the individual term structure components, I therefore employ sign restriction techniques similar to those suggested by Uhlig [2005] and Mountford and Uhlig [2005]. My approach, outlined in detail in Section 4.4, allows to identify surprise changes of each of the three yield factors that are not accompanied by simultaneous responses of the remaining two factors. Accordingly, term structure movements are dissected into unexpected changes of the three components level, slope, and curvature. Since the yield and the macro factors share common dynamics, the future evolution of key economic indicators subsequent to yield curve surprises can be studied.
The results of my impulse response analysis are summarized in Section 4.5. The

most important finding is that surprise changes of the curvature factor carry important news about the future evolution of the term structure and of output. In particular, unexpected increases of the curvature factor announce a strong and persistent response of the yield curve slope and a significant decline of the yield curve level. Together, these two features imply a successive flattening of the yield curve which is commonly associated with an upcoming recession. Consistent with this interpretation, output growth exhibits a pronounced hump-shaped response following an unexpected increase of curvature, eventually falling below zero about one year after the shock occurs (see Figures 4.9 and 4.18). This result is surprising since curvature has previously been documented to be unrelated to macroeconomic variables (e.g. Diebold et al. 2006, Dewachter and Lyrio 2006).

While the curvature factor displays strong predictive power, I find the yield curve slope to be less informative than the regular use of interest rate spreads as predictors of recessions would suggest. In particular, an unexpected increase of the slope factor - tantamount to diminishing yield spreads - is followed by an almost immediate but not very pronounced decline of output (see Figures 4.8 and 4.17). According to these results, a rising slope factor is therefore associated with a decline of output, but appears to be announced by the curvature factor. This might qualify curvature as a forward-looking indicator. Finally note that consistent with conventional wisdom, surprise surges of the level factor are followed by a pronounced subsequent rise of inflation. Altogether, Chapter 4 provides a systematic analysis of the predictive content of the level, slope, and curvature of the term structure. My results indicate that the yield curve carries more predictive information than what is captured by interest rate spreads alone.

The paper underlying Chapter 5 is joint work with Harald Uhlig and has been published in the *Journal of Business Cycle Measurement and Analysis*, Volume 2 No. 1 in May 2005. It contains an exercise in dating the Euro area business cycle on a monthly basis. Much of the chapter deals with the construction of monthly time series of real GDP. We suggest an interpolation routine that nests some popular temporal disaggregation models which have been proposed in previous studies. Our model takes into account information from related monthly indicators and is cast in state-space form. Appendix 6.5.1 provides a detailed discussion of the approach. Based on our constructed monthly series of real GDP, we obtain business cycle turning points using a modified version of the nonparametric dating routine suggested by Bry and Boschan [1971] which we discuss in Section 5.2. The variant of the Bry-Boschan algorithm that we consider adds a combined amplitude/phase-length criterion to the original procedure so as to rule out business cycle phases that are short and flat. Applied to our constructed series of monthly real GDP for the US and the Euro area, we show in Section 5.3 that the algorithm closely replicates the dating decisions of the NBER and the CEPR.

2 Conditional Asset Pricing with a Large Information Set

Dynamic factors summarize the information in a large number of variables and are therefore intuitively appealing proxies for the information set available to investors. This chapter demonstrates that conditioning on dynamic factors instead of commonly used instruments substantially reduces the pricing errors implied by conditional models. Dynamic factors are further shown to exhibit incremental explanatory power over benchmark conditioning variables. The results withstand a number of robustness tests and carry important implications for the specification of conditional asset pricing models in applied research and practice.

2.1 Introduction

While the Capital Asset Pricing Model (CAPM) of Sharpe [1964], Lintner [1965], and Black [1972] has long been the workhorse asset pricing model, it is now a widely accepted fact that it fails to explain the cross-section of portfolio returns ordered by firm characteristics such as size and book-to-market. The empirical failure of the CAPM has given rise to different interpretations, however. Some authors have argued that the comovement with the market portfolio is not the only source of risk faced by investors, and hence factors that capture additional hedging concerns need to be added in order to explain cross-sectional return differences (e.g. Merton 1973, Chen et al. 1986, Campbell 1996). In contrast, a number of studies have emerged recently which state that the model holds in a conditional sense. Accordingly, some authors have suggested to "resurrect" the CAPM by taking into account time-variation in assets' exposure to the market risk and the associated risk premium. Indeed, various conditional versions of the CAPM have been shown to explain the cross-section of returns better than the unconditional CAPM and not worse than the Fama-French three-factor model (e.g. Jagannathan and Wang 1996, Lettau and Ludvigson 2001, Santos and Veronesi 2006).[1]

This chapter also adopts the view that the CAPM holds conditionally. In contrast to previous studies, however, I test the model using a broad set of conditioning information rather than individual instruments. To do so, I build on recent research in dynamic factor analysis for large datasets which has shown that the information contained in many time series can parsimoniously be summarized by a few factors. In particular, I extract factors from a large panel of macroeconomic time series using the methodology of Stock and Watson [2002a,b] and employ them as conditioning variables in tests of the CAPM. I show that this approach substantially reduces the pricing errors with respect to specifications based on individual benchmark instruments. Therefore, the main conclusion of this chapter is that dynamic factors are better proxies for the information set available to investors than commonly used individual instruments.

More precisely, the results of my study are the following. I first show in tests of the conditional CAPM based on a single instrument that dynamic factors imply substantially smaller pricing errors than commonly used conditioning variables such as the term spread, the dividend yield or the log consumption-wealth-ratio. Second, specifications of the conditional CAPM using two dynamic factors as instruments strongly outperform specifications based on two benchmark conditioning variables. They are also shown to price the 25 Fama-French size and book-to-market sorted stock portfolios more precisely than the Fama-French three-factor model. Third, dynamic factors have strong incremental explanatory power over commonly used instruments when they are jointly used as conditioning variables. In contrast, no benchmark instrument carries useful conditioning information in addition to the two best-performing diffusion indexes. The results are robust to the change of test assets and variations of the sample period. Moreover, they prevail in tests of the conditional Consumption-CAPM. Altogether, the empirical evidence strongly supports the hypothesis that factors extracted from a large macroeconomic data panel are more useful conditioning variables than individual instruments commonly employed in the literature. This result carries

[1] Note that in a recent paper, Lewellen and Nagel [2005] generally question the ability of the conditional CAPM to explain cross-sectional asset pricing anomalies.

important implications for the specification of conditional asset pricing models in applied research and practice.

My analysis bears some relationship with a recent study by Ludvigson and Ng [2005]. They show that adding dynamic factors to a set of commonly used instruments, one can significantly improve out-of-sample forecasts of the time-varying mean and volatility of excess stock returns. Since the time-varying parameters in conditional pricing models are inherently linked to the conditional moments of returns, their finding provides another argument for using dynamic factors as instruments in conditional asset pricing models. As the results in this chapter document, dynamic factors have explanatory power for returns not only in the time-series but also in the cross-sectional dimension.

The remainder of this chapter is organized as follows. In Section 2.2, I show how conditional asset pricing models can be tested using a large conditioning information set. Section 2.3 briefly summarizes the estimation methodology that is used to assess the performance of the different model specifications. A more detailed treatment of the estimation and test methodology as well as some implementation issues is provided in Appendix 6.1.2. Section 2.4 documents the data used and in Section 2.5 I summarize the empirical results of my study. Section 2.6 concludes the chapter.

2.2 Conditioning on a Large Information Set

It has become common practice to formulate asset pricing models in the particularly tractable stochastic discount factor language. I follow this convention and start with the basic pricing equation which states that in the absence of arbitrage opportunities, there exists at least one pricing kernel m that prices all assets in the payoff space correctly, i.e.

$$E_t[R_{i,t+1}m_{t+1}] = 1 \qquad \forall \;\; i = 1 \ldots N \tag{2.1}$$

where E_t denotes the expectation conditional on all information available in period t, m_{t+1} denotes the stochastic discount factor, and $R_{i,t+1}$ the return on the i-th asset. A large class of empirical asset pricing models assume a linear relationship between the discount factor and the pricing factors f which capture the fundamental sources of risk faced by investors, i.e.

$$m_{t+1} = a_t + b_t' f_{t+1}, \tag{2.2}$$

where a_t is a scalar and b_t a $k \times 1$ vector of pricing parameters. Unconditional pricing models assume constant parameters a and b. However, one can show that the parameters in any multi-factor pricing model depend on the conditional means and variances of the pricing factors and therefore must be time-varying (see Cochrane 2001 or Wang [2004]). Since conditional moments are not observed, the time-dependence is commonly modeled by assuming a_t and b_t to be linear functions of conditioning variables Z_t that are in investors' information set in period t , i.e.

$$a_t = a_0 + a_1' Z_t \quad \text{and} \quad b_t = b_0 + b_1' Z_t,$$

where a_0, b_0, a_1, and b_1 are constant coefficients and where Z_t is a $M \times 1$ vector of the conditioning variables. As Cochrane [2001] shows, this specification of a conditional pricing model implies an unconditional model with pricing factors given by the lagged instruments Z_t, the risk factors f_{t+1}, and all products $Z_t f_{t+1}$ of factors and lagged instruments. Obviously, this results in a large number of factors. For example, a two-factor pricing model scaled with five instruments yields a total of 17(!) factors that must in principle be included in the unconditional model.

Consequently, in order to avoid degrees of freedom problems, most conditional pricing models are tested using only one instrument at the same time. The particular choice of conditioning variable is commonly motivated either by its capacity to forecast returns or by its relevance as a cyclical indicator since expected returns have been shown to vary over the business cycle. For example, based on its previously documented predictive power for excess returns, Lettau and Ludvigson [2001] advocate the use of the log consumption-wealth ratio as an instrument in conditional asset pricing tests. Jagannathan and Wang [1996] suggest to use the default spread as a proxy for the time-varying market risk premium. Other authors use the cyclical component of industrial production or GDP (Hodrick and Zhang 2001), or the term spread, a short-term interest rate, and the dividend yield (Ferson and Harvey 1999) as conditioning variables. Overall, there is yet little agreement on which variables should be used as instruments in tests of conditional asset pricing models. Moreover, taking into account that market participants have access to a huge amount of information, single indicators more principally appear to be inadequate proxies for investors' information sets. Ideally, one should therefore use as much information as possible when testing conditional pricing models. However, as shown above, simply adding instruments does not do the trick as one quickly runs into degrees of freedom problems.

The solution to this problem which I consider in this chapter employs dynamic factor analysis for large datasets. Research on dynamic factor models has shown that the information in many time series can effectively be summarized by a few common factors (Stock and Watson 2002a,b, Forni et al. 2005). Under the assumption that investors process a large amount of information about the state of the economy, dynamic factor analysis quite intuitively should prove beneficial also in asset pricing applications. Indeed, there is recent empirical evidence supporting this conjecture. In Mönch (2005, Chapter 3) I show that combining a Factor-Augmented VAR (Bernanke et al. 2005) with no-arbitrage restrictions significantly improves out-of-sample forecasts of government bond yields. Moreover, Ludvigson and Ng [2005] employ factors extracted from two large macroeconomic and financial data panels to forecast stock returns. Adding these factors to a set of commonly used instruments, they show that forecasts of the conditional mean and volatility of the market return are significantly improved.

In this chapter, I test whether dynamic factors represent useful conditioning variables. In order to keep the analysis tractable, I focus on the conditional CAPM, i.e. I test models of the form

$$m_{t+1} = a_t + b_t R_{m,t+1},$$

where the return on the market portfolio R_m is the only relevant pricing factor. As before, let the time-varying parameters a_t and b_t depend on some conditioning variables available in period t. I assume that there is a large number of conditioning variables $\{x_1 \ldots x_M\}$ which have the following factor structure:

$$x_{it} = \lambda_i' F_t + \epsilon_{it}, \tag{2.3}$$

where x_{it} denotes the time-t observation of the i-th instrument, F_t is the $q \times 1$ vector of common factors, λ_i denotes the corresponding vector of factor loadings and ϵ_{it} an idiosyncratic component.[2] Hence, the common variation of the M

[2] Note that the estimation approach of Stock and Watson [2002a,b] that is employed here allows for some mild serial and cross-correlation of the idiosyncratic components. The setup thus describes an approximate factor model.

variables in the panel is summarized by a small number q of factors where it is assumed that $q \ll M$. As a consequence, dynamic factor analysis allows to substantially reduce the dimensionality of the problem that is posed by choosing a set of conditioning variables. I therefore study the empirical performance of conditional versions of the CAPM where the pricing parameters a_t and b_t are assumed to be linear functions of time-t observations of the factors F extracted from a large set of macroeconomic variables,

$$a_t = a_0 + a_1' F_t \quad \text{and} \quad b_t = b_0 + b_1' F_t,$$

This specification of the time-varying pricing coefficients implies an unconditional factor model of the form

$$\mathbf{1} = E\left[R_{t+1}\left(a_0 + a_1' F_t + b_0 R_{m,t+1} + b_1' F_t R_{m,t+1}\right)\right]$$

with pricing factors F_t, $R_{m,t+1}$, and $F_t R_{m,t+1}$ and constant coefficients a_0, a_1, b_0, and b_1. Stacking the coefficients and the scaled pricing factors into the vectors $b = (a_0, a_1, b_0, b_1)$ and $\bar{\mathbf{f}}_{t+1} = (F_t', R_{m,t+1}, F_t' R_{m,t+1})'$ and letting $f_{t+1} = (1, \bar{\mathbf{f}}_{t+1})$, the model becomes

$$\mathbf{1} = E\left[b' f_{t+1} R_{t+1}\right]. \tag{2.4}$$

Based on this formulation, I carry out tests of the conditional CAPM using unconditional moments.

2.3 Test Methodology

Before specifying the empirical strategy used to test the conditional factor pricing models, I briefly describe the estimation approach used to extract the common factors from a large panel of economic time series variables. As Ludvigson and Ng [2005], I employ the method popularized by Stock and Watson [2002a,b]. Using the standardization $F'F/T = I_q$, Stock and Watson show that for large M and T the space spanned by the common factors is consistently estimated by the first q principal components of the cross-sectional variance-covariance matrix XX'. I provide more details on how the Stock-Watson procedure is implemented in appendix 6.1.1. Since the true number of common components is not known, I employ the panel information criteria developed by Bai and Ng [2002] to determine the number of factors that summarize the common variation among the variables X. Note that Stock and Watson label the factors extracted from a large panel of macro time series "diffusion indexes". In order to clearly keep apart the notion "factor" in the sense of risk factor in an asset pricing model and "factor" denoting the common component of a large number of macro variables, I adopt Stock and Watson's terminology throughout the chapter and refer to the factors extracted from the large panel of macro time series as diffusion indexes.

2.3.1 The SDF Method

Factor pricing models formulated in the stochastic discount factor language give rise to a set of moment conditions that can be used for estimation via the Generalized Method of Moments (GMM). In particular, any model of the form (2.4) implies a vector of pricing errors

$$g(b) = E\left[b' f_t R_t - \mathbf{1}\right] \tag{2.5}$$

If the model is valid, $g(b)$ must be zero. The GMM procedure uses this condition to choose parameter estimates b which minimize a weighted sum of squared pricing errors. Different weighting matrices are commonly employed in the literature. Hansen [1982] suggested to weight the pricing errors with the inverse of their sample variances in order to obtain efficient estimates of the coefficients b. Moreover, he proposed a J-statistic to assess whether the pricing errors implied by the model are jointly zero. As the optimal GMM weights are model-dependent, Hansen and Jagannathan [1997] have suggested to instead use the inverse of the second moment matrix of returns $E[RR']^{-1}$. This approach allows direct comparison of different models by assessing their pricing errors. In the empirical results below, I denote J_{HJ} the J-statistic obtained using the Hansen-Jagannathan weighting matrix.

Hansen and Jagannathan [1997] provide another test statistic that is directly suitable for model comparisons. Noting that any true stochastic discount factor prices all assets in the payoff space correctly, they argue that a false asset pricing model will give rise to a strictly positive minimum distance between the pricing kernel implied by the model and the set of true stochastic discount factors. In my empirical results summarized below, I provide estimates of the Hansen-Jagannathan distance measure (HJ-distance) for all compared models. Jagannathan and Wang [1996] derive the asymptotic sampling distribution of this statistic which equals a weighted sum of $(N-k)$ $\chi^2(1)$-distributed random variables.

To obtain the p-value for the HJ-distance, I simulate this weighted sum 100,000 times.

A recurrent theme throughout this chapter is to study how well different specifications of conditional pricing models compare to each other. The tests depicted above allow to investigate these questions. It is also of interest, however, to directly assess whether some conditioning variables drive out others when they simultaneously enter a model. More precisely, I am interested in studying whether instruments commonly used in asset pricing tests describe the time-variation in the pricing relationship of the CAPM sufficiently well or whether adding factors extracted from a large panel of macro data improves the fit of the model. Conversely, it is important to assess whether conditioning on a large information set by using diffusion indexes as instruments is sufficient or whether individual instruments capture additional useful pricing information. These questions can easily be tested in the GMM framework. In particular, let the time-varying coefficients of the conditional CAPM be given by

$$\begin{aligned} a_t &= a_0 + a_F' F_t + a_Z' Z_t \\ b_t &= b_0 + b_F' F_t + b_Z' Z_t, \end{aligned} \tag{2.6}$$

where F denotes the vector of factors summarizing the variation in a large macroeconomic dataset and Z some additional conditioning variables. Now stack the pricing factors and constants into the vectors $\bar{\mathbf{f}}_{t+1} = (F_t', F_t' R_{m,t+1})'$, $\bar{\mathbf{z}}_{t+1} = (Z_t', Z_t' R_{m,t+1})'$, $b_f = (a_F, b_F)$, and $b_z = (a_Z, b_Z)$, respectively. Moreover, let $f_{t+1} = (1, R_{m,t+1}, \bar{\mathbf{f}}_{t+1}, \bar{\mathbf{z}}_{t+1})$ and $b = (a_0, b_0, b_f, b_z)$. Then, the model can again be written in the form

$$\mathbf{1} = E_t[b' f_{t+1} R_{t+1}], \tag{2.7}$$

and testing $b_f = 0$ or $b_z = 0$ allows to assess whether one set of conditioning variables drives out the other. Following Cochrane [2001], I employ two different strategies to do so. The first is a simple Chi-square test of the form

$$\hat{b}_j' \; Cov(\hat{b}_j)^{-1} \; \hat{b}_j \sim \chi^2(k_j), \tag{2.8}$$

where k_j denotes the number of elements in b_j and where $j = \{f, z\}$. Second, a χ^2 difference test can be applied by estimating both the unrestricted model (2.7) and a restricted version setting e.g. $b_f = 0$. Computing for both models the J-statistic defined above using the same weighting matrix, the statistic

$$\Delta J = T \; J(\hat{b}_f, \hat{b}_z) - T \; J(\hat{b}_z) \sim \chi^2(k_F) \tag{2.9}$$

can be used to assess whether imposing the restriction leads to a significant increase of the sum of squared pricing errors or not.

Ghysels [1998] has noted that the out-of-sample performance of conditional pricing models may be impaired by parameter instability in the relationship between the pricing factors and the returns on the test assets. In order to assess for each model specification whether this is a concern, I report as another diagnostic Andrews' (1993) supLM test for structural breaks. To be more precise, I compute the LM-statistic at increments of 0.05 over the interval [0.15 ; 0.85] and report as supLM the supremum of these statistics. The judgment whether a model fails or passes the supLM test is based on the distribution tables provided in Andrews [1993].

2.3.2 The Beta Method

Pricing models of the form (2.4) imply an unconditional multifactor beta representation for returns given by

$$E[R_i] = E[R_0] + \beta_{i,1}\gamma_1 + \ldots + \beta_{i,k}\gamma_k, \tag{2.10}$$

where $E[R_0]$ is the average return of a zero-beta portfolio that is uncorrelated with the pricing kernel and where $\beta_{i,j}$ denotes the exposure of return R_i to variation in the pricing factor f_j and γ_j the associated price of risk. Models of this form can be consistently estimated using the cross-sectional regression methodology of Fama and MacBeth [1973]. Since it is intuitively appealing and easy to implement, this estimation approach remains a widely used tool in empirical tests of asset pricing models.[3] The Fama-MacBeth procedure works in two steps. First, estimates of the betas are obtained by regressing individual returns on the pricing factors. Second, the market prices of risk are estimated by running cross-sectional regressions of returns on the betas. As the betas are estimated, the Fama-MacBeth method suffers from an errors in variables problem. A popular adjustment method for the resulting bias has been proposed by Shanken [1992]. I report Fama-MacBeth and Shanken-corrected t-statistics for the estimated factor risk premia below. To evaluate the overall model performance in the Fama-MacBeth setup, I report the cross-sectional R-square suggested by Jagannathan and Wang [1996]. Moreover, I test

[3] Some recent applications of the method are Lettau and Ludvigson 2001, Li, Vassalou, and Xing 2004, and Lustig and Van Nieuwerburgh 2005.

whether the average pricing errors are jointly zero using the Chi-Square statistic[4]

$$J_\alpha = \bar{\alpha}' \; Cov(\bar{\alpha})^{-1} \; \bar{\alpha} \sim \chi^2(N-1), \tag{2.11}$$

where $\bar{\alpha}$ denotes the vector of model-implied average pricing errors.

2.4 Data

Since the influential paper by Fama and French [1992], it is common practice to evaluate asset pricing models based on their ability to explain the returns on portfolios sorted by size and book-to-market ratio. I follow this practice and use as test assets the returns on Fama-French's 25 size and book-to-market sorted US stock portfolios.[5]

I extract factors from a large panel of monthly macroeconomic time series for the US which has been compiled by Stock and Watson [2005]. The dataset comprises 132 variables from various economic categories: real output and income, employment and hours, real retail, manufacturing and trade sales, consumption, housing starts and sales, real inventories, orders, stock prices, exchange rates, interest rates and spreads, money and credit quantity aggregates, price indexes, average hourly earnings, and miscellaneous. As in Stock and Watson [2005], the series are subjected to some preliminary transformations in order to achieve stationarity.[6] Moreover, prior to extracting the factors, all series have been standardized to have zero mean and unit variance.

I compare the usefulness of the extracted factors as instruments in conditional asset pricing models with five benchmark conditioning variables that have been employed in previous studies. These are the term spread between a ten-year US government bond yield and the three-month Treasury bill yield, "TERM", the default spread between Moody's Baa and Aaa corporate bond yields, "DEF", the

---

[4] See Cochrane [2001], p. 246.

[5] I thank Kenneth French for making the return data available on his website.

[6] I am grateful to Mark Watson for making these data available on his website. For details on the exact composition of the panel and the data transformations, the reader is referred to Stock and Watson's (2005) paper. Notice that I slightly modify their preadjustment approach by computing annual inflation rates instead of monthly growth rates of inflation for the price series in the panel. As documented in Mönch (2005, Chapter 3), this increases the persistence of the estimated factors and possibly enhances their ability to explain risk premia.

one-month Treasury bill yield, "TB1", the spread between the three-month and the one-month Treasury bill yield, "TB31", and the dividend yield of the S&P 500 index, "DIV". These are exactly the same instruments that have been selected by Ferson and Harvey [1999] on the basis of their previously documented ability to forecast returns or to proxy for time-variation in risk premia. Note that Wang [2004] includes the lagged market return as a sixth instrument whereas Petkova [2005] only considers TERM, DEF, TB1, and DIV.

Recently, some conditioning variables have received attention which by construction capture investors' expectations about future excess returns. Two examples for such instruments are the log consumption wealth ratio, suggested by Lettau and Ludvigson [2001] who denote this variable cay, and the labor income to consumption ratio, s_w, proposed by Santos and Veronesi [2006]. Both variables have been shown to be valuable instruments in cross-sectional tests of asset pricing studies. I therefore use them as two additional benchmarks to assess the relative usefulness of diffusion indexes as conditioning variables. Since cay and s_w are only available on a quarterly basis, I transform all series in the monthly panel by Stock and Watson [2005] into the quarterly frequency and then extract factors from the resulting quarterly dataset.[7]

I estimate the different CAPM specifications over the time period 1963:01-2003:12. The sample thus covers 41 years of data and a total of 492 observations.

2.5 Empirical Results

In this section, I present the results obtained from estimating different specifications of the conditional CAPM based on individual instruments, dynamic factors or both. As stated above, I extract factors from the large panel of macroeconomic time series using the methodology of Stock and Watson [2002a,b]. The panel information criteria by Bai and Ng [2002] indicate that a total of 8 factors captures the bulk of common variation among the 132 time series in the dataset.[8]

[7] I thank Sydney Ludvigson and Martin Lettau for providing the cay series and its components on their website. Notice further that I compute s_w using the same consumption and labor income series that have been employed to construct cay.

[8] Note that Stock and Watson [2005] find an optimal number of 7 factors to summarize the same dataset. The small difference either owes to the fact that I use a shorter sample period or that I

I start with GMM tests of the conditional CAPM where only a single conditioning variable is employed. This gives a first impression of whether the factors extracted from the large panel of macro series capture useful pricing information and how they compare to instruments commonly used in conditional asset pricing tests. I then confront specifications of the conditional CAPM where two benchmark instruments are used as conditioning variables with specifications based on two diffusion indexes. To facilitate comparison with previous studies, I relate the best-performing specification using two diffusion indexes to the unconditional CAPM and the Fama-French three-factor model on the basis of Fama-MacBeth regressions. In a next step, I use the GMM framework to test whether the diffusion indexes have additional explanatory power over the benchmark instruments. This is important to check before one can confidently argue that the factors extracted from a large dataset incorporate useful pricing information that is not captured by the benchmark instruments. As an additional test of how the diffusion indexes relate to individual conditioning variables, I then extract factors from a quarterly dataset and compare their usefulness with the log consumption wealth ratio suggested by Lettau and Ludvigson [2001] and the labor income to consumption ratio recently proposed by Santos and Veronesi [2006]. Finally, I carry out a set of robustness tests to demonstrate that the strong relative performance of diffusion indexes as conditioning variables is not sensitive to the choice of test assets, variations of the sample period, or the choice of pricing factors.

2.5.1 Conditional CAPM with One Instrument

I first estimate specifications of the conditional CAPM using the diffusion indexes one by one as conditioning variables. I compare their performance with the variables TERM, DEF, TB1, TB31, and DIV described above, which are often used as instruments in asset pricing studies. Table 2.1 summarizes the results of GMM estimations obtained for the eight diffusion indexes and the five benchmark conditioning variables. Several conclusions can be drawn from these results. First, all benchmark instruments imply a Hansen-Jagannathan distance of 0.13 and thus perform about equally well. Second, five out of the eight diffusion indexes also give rise to a HJ-distance of 0.13. However, the fourth, sixth, and eighth index imply values of 0.11 and 0.12, respectively, and thus perform better than all benchmark instruments. This is reflected also by the J_{HJ}-statistics, summarized in the

apply a slightly different set of preadjustments to the individual series.

third column of Table 2.1. Four out of the eight diffusion indexes imply a value of J_{HJ} smaller than the one obtained for the term spread which performs best among the benchmark instruments. Hence, most diffusion indexes imply smaller pricing errors than the commonly used conditioning variables. The results obtained from optimal GMM estimations confirm this finding. The J-statistic is not significantly different from zero for five out of the eight diffusion indexes. In contrast, among the benchmark instruments this only holds true for the default spread. Overall, these results indicate that it is beneficial to exploit a large information set when testing the conditional CAPM. Finally notice that the $supLM$ statistics indicate parameter instability for some model specifications. In particular, this is the case for the diffusion indexes F_5, F_6, and F_8, and the dividend yield DIV. The latter result is somewhat consistent with the instability of the dividend-price ratio as a predictor of returns that has been documented by e.g. Lettau, Ludvigson, and Wachter [2006].

2.5.2 Conditional CAPM with Two Instruments

The results documented above show that some diffusion indexes outperform benchmark instruments in tests of the conditional CAPM based on a single conditioning variable. I now study whether this result prevails when more than one instrument is used. In particular, I compare specifications of the conditional CAPM based on two diffusion indexes or benchmark instruments, respectively. In order to restrict the total number of different specifications to estimate, I focus on combinations which comprise the single best performing diffusion index, F_4, and the best benchmark instrument, TERM, according to the results above. Moreover, I only consider the first five diffusion indexes which together explain about 60% of the total variance of the 132 series in the macro dataset. Overall, eight different specifications of the conditional CAPM are compared, four using pairs of diffusion indexes as conditioning variables and four using pairs of benchmark instruments.

Table 2.2 summarizes the GMM estimation results obtained for these eight specifications. Several remarks need to be made. First and most importantly, the best performing model using two individual instruments, which combines the term spread and the dividend yield, prices the 25 Fama-French portfolios less correctly

Table 2.1: **GMM Tests of the Conditional CAPM with One Instrument**

This table summarizes GMM estimation results for different specifications of the conditional CAPM based on a single conditioning variable. The 25 Fama-French portfolios are used as test assets. The estimation period is from 1963:01 to 2003:12. HJ-dist denotes the Hansen and Jagannathan [1997] distance measure, J is Hansen's (1982) test on the overidentifying restrictions of the model, J_{HJ} the equivalent test statistic based on the Hansen-Jagannathan weighting matrix $E[RR']^{-1}$, and $supLM$ denotes Andrews' (1993) test for structural breaks with unknown change point. p-values are provided in parentheses below the estimates. I indicate that a model does not pass Andrews' stability test at the 10%, 5%, and 1% level of significance with one, two, and three asterisks.

Z_{t-1}	HJ-dist	J	J_{HJ}	supLM
F_1	0.13	19.05	70.06	2.52
	(.81)	(.58)	(.00)	
F_2	0.13	17.58	47.75	3.08
	(.66)	(.68)	(.00)	
F_3	0.13	59.19	69.90	6.87
	(.15)	(.00)	(.00)	
F_4	0.11	21.60	38.85	5.50
	(.90)	(.42)	(.01)	
F_5	0.13	16.83	82.82	**18.05
	(.69)	(.72)	(.00)	
F_6	0.12	57.62	40.03	***71.39
	(.08)	(.00)	(.01)	
F_7	0.13	38.10	76.36	4.72
	(.49)	(.01)	(.00)	
F_8	0.12	20.01	42.48	***22.98
	(.77)	(.52)	(.00)	
TERM	0.13	48.57	57.39	6.57
	(.24)	(.00)	(.00)	
DEF	0.13	28.57	77.66	7.88
	(.75)	(.12)	(.00)	
TB1	0.13	51.08	70.96	8.61
	(.30)	(.00)	(.00)	
TB31	0.13	38.44	70.25	9.64
	(.60)	(.01)	(.00)	
DIV	0.13	64.05	68.76	**18.63
	(.14)	(.00)	(.00)	

than the worst performing specification based on two diffusion indexes. While the point estimates of the HJ-distance and the J_{HJ}-statistic are 0.13 and 39.07 for the (TERM, DIV)-specification, all model versions based on pairs of diffusion indexes give rise to values of the HJ-distance measure of 0.11 and 0.10, respectively, and to values of the J_{HJ}-statistic ranging from about 17.4 to 38.4. Hence, diffusion index-based models imply considerably smaller pricing errors than model specifications using individual conditioning variables. According to the J-statistics and the corresponding p-values reported in the second column of Table 2.2, the optimal GMM estimates for none of the diffusion index specifications imply pricing errors that are significantly different from zero. This is only true for two of the models using benchmark instruments. Moreover, the estimates obtained using the Hansen-Jagannathan weighting matrix give rise to pricing errors that are significantly different from zero for all specifications based on individual variables. In contrast, two combinations of diffusion indexes yield values of the J_{HJ}-statistic that are not significantly different from zero.

Finally note that all diffusion index-based specifications of the conditional CAPM price the 25 size and book-to-market sorted portfolios considerably better the unconditional CAPM and the Fama-French three-factor model. This is remarkable since the Fama-French model has become an important benchmark in recent years.[9] Altogether, these results provide pervasive evidence that conditioning on a large information set by using diffusion indexes as instruments considerably enhances the fit of the CAPM.

Fama-MacBeth Regressions

Traditionally, asset pricing models have often been tested using the cross-sectional regression methodology of Fama and MacBeth [1973]. To facilitate comparison with previous studies, I also report results obtained from applying this estimation strategy. To conserve space, I restrict the analysis to the conditional CAPM specifications which have been shown to perform best in the GMM tests documented above. Recall that these are the combination of the term spread and the dividend yield, and the combination of the fourth and fifth diffusion index. Table 2.3 summarizes the outcomes of cross-sectional regressions of these two CAPM specifications, the static CAPM and the Fama-French three-factor model. For each model, the risk premia estimates for the individual pricing factors, the corresponding t-values and their Shanken-adjusted counterparts are reported. Further, two model diagnostics are provided: the cross-sectional R^2 and the J_α-statistic which allows

[9] Notice, however, that the specifications using two instruments imply a total number of five unconditional factors whereas the Fama-French model only comprises three factors.

Table 2.2: **GMM Tests of the Conditional CAPM with Two Instruments**

This table summarizes GMM estimation results for different specifications of the conditional CAPM based on two conditioning variables. The 25 Fama-French portfolios are used as test assets. The estimation period is from 1963:01 to 2003:12.

Z_{t-1}	HJ-dist	J	J_{HJ}	supLM
F_1, F_4	0.11	22.61	38.38	13.19
	(.84)	(.25)	(.01)	
F_2, F_4	0.10	18.23	25.92	14.05
	(.85)	(.51)	(.13)	
F_3, F_4	0.11	19.14	34.35	7.37
	(.86)	(.45)	(.02)	
F_4, F_5	0.10	10.19	17.73	8.79
	(.98)	(.95)	(.54)	
TERM, DEF	0.13	33.87	50.01	*18.30
	(.37)	(.02)	(.00)	
TERM, TB1	0.12	22.66	40.79	17.06
	(.46)	(.25)	(.00)	
TERM, TB31	0.13	44.26	56.97	15.33
	(.29)	(.00)	(.00)	
TERM, DIV	0.12	21.57	39.07	17.27
	(.58)	(.31)	(.00)	
CAPM	0.14	61.76	70.89	2.78
	(.00)	(.00)	(.00)	
FF	0.12	57.23	56.13	10.92
	(.00)	(.00)	(.00)	

to test whether the model-implied average pricing errors are jointly equal to zero. Consistent with the evidence in many previous studies, the CAPM with constant coefficients explains only a small share, 16%, of the variation of average returns across the 25 size and book-to-market sorted stock portfolios. Likewise consistent with previous studies, the Fama-French three-factor model performs considerably better than the static CAPM, explaining about 80% of the cross-sectional variation of average returns. According to the J_α-statistic, however, the Fama-French model implies average pricing errors significantly different from zero at the 5% confidence level. As one would expect from the GMM estimation results documented above, the conditional versions of the CAPM perform better than the unconditional Fama-French model. The specification employing TERM and DIV as instruments explains about 85% of the cross-sectional variation of the test assets and implies average pricing errors which are not statistically different from zero. Strikingly, the fit of the conditional CAPM is even better using the diffusion

Table 2.3: **Fama-MacBeth Tests of the Conditional CAPM with Two Instruments**

This table summarizes results from Fama-MacBeth regressions for the unconditional CAPM, the Fama-French three-factor model, and different specifications of the conditional CAPM based on two benchmark instruments or diffusion indexes, respectively. The models are of the form

$$E[R_i] = E[R_0] + \beta_{i,f_1}\gamma_1 + \ldots + \beta_{i,f_k}\gamma_k.$$

The second and third row of each panel provide the Fama-MacBeth and Shanken-adjusted t-statistics for these estimates. R^2 and J_α refer to the cross-sectional R-square and the test for zero average pricing errors defined in (2.11). The 25 Fama-French portfolios are used as test assets. The estimation period is from 1963:01 to 2003:12.

Model	Pricing Factors						$R^2(\bar{R}^2)$	J_α
CAPM	cst	R_m						
$\hat{\gamma}$	1.81	-0.57					0.16	56.58
t-value	4.73	-1.31					0.12	0.00
Shanken-t	4.69	-1.18						
FF	cst	R_m	SMB	HML				
$\hat{\gamma}$	1.84	-0.86	0.21	0.47			0.79	45.01
t-value	5.97	-2.31	1.42	3.45			0.76	0.01
Shanken-t	5.78	-1.97	1.00	2.42				
cond CAPM (Z)	cst	R_m	TERM	DIV	$R_m\times$TERM	$R_m\times$DIV		
$\hat{\gamma}$	1.45	-0.10	1.88	-2.68	0.57	-5.56	0.85	31.51
t-value	5.29	-0.27	5.02	-3.47	0.41	-1.99	0.81	0.14
Shanken-t	2.71	-0.13	2.55	-1.75	0.21	-0.99		
cond CAPM (F)	cst	R_m	F_4	F_5	$R_m\times F_4$	$R_m\times F_5$		
$\hat{\gamma}$	1.97	-1.04	0.21	0.60	4.07	-2.00	0.91	25.04
t-value	4.29	-2.15	0.53	2.92	3.78	-3.13	0.88	0.40
Shanken-t	2.70	-1.31	0.33	1.82	2.36	-1.93		

indexes F_4 and F_5 as instruments. In particular, this specification implies a cross-sectional R^2 of 91% and even smaller pricing errors than the conditional CAPM based on individual instruments.

Figure 2.1 visualizes these results, providing plots of average realized returns against fitted expected returns for the 25 Fama-French portfolios. These are denoted by two digits where the first corresponds to the size quintile and the second to the book-to-market quintile. Visibly, the static CAPM fails to explain the cross-section of 25 portfolios. The other three models price the test assets much more precisely although with some differences. Most importantly, the Fama-French three-factor model implies relatively large pricing errors for portfolios with a very low or very high book-to-market ratio (second digit equal to 1 or 5, respectively).

The (TERM, DIV)-specification of the conditional CAPM appears to explain average returns of low value firms better than the Fama-French model, but still implies relatively large pricing errors for high value firms. In contrast, the CAPM conditional on two diffusion indexes prices both groups of assets more precisely than the competitor models. The performance is particularly remarkable for the small growth portfolio that previous studies (e.g. Lettau and Ludvigson 2001) have documented to be difficult to price. Overall, the outcomes of the Fama-MacBeth regressions confirm that diffusion indexes capture conditioning information which enhances the fit of the CAPM.

2.5.3 Do Factors Have Incremental Explanatory Power?

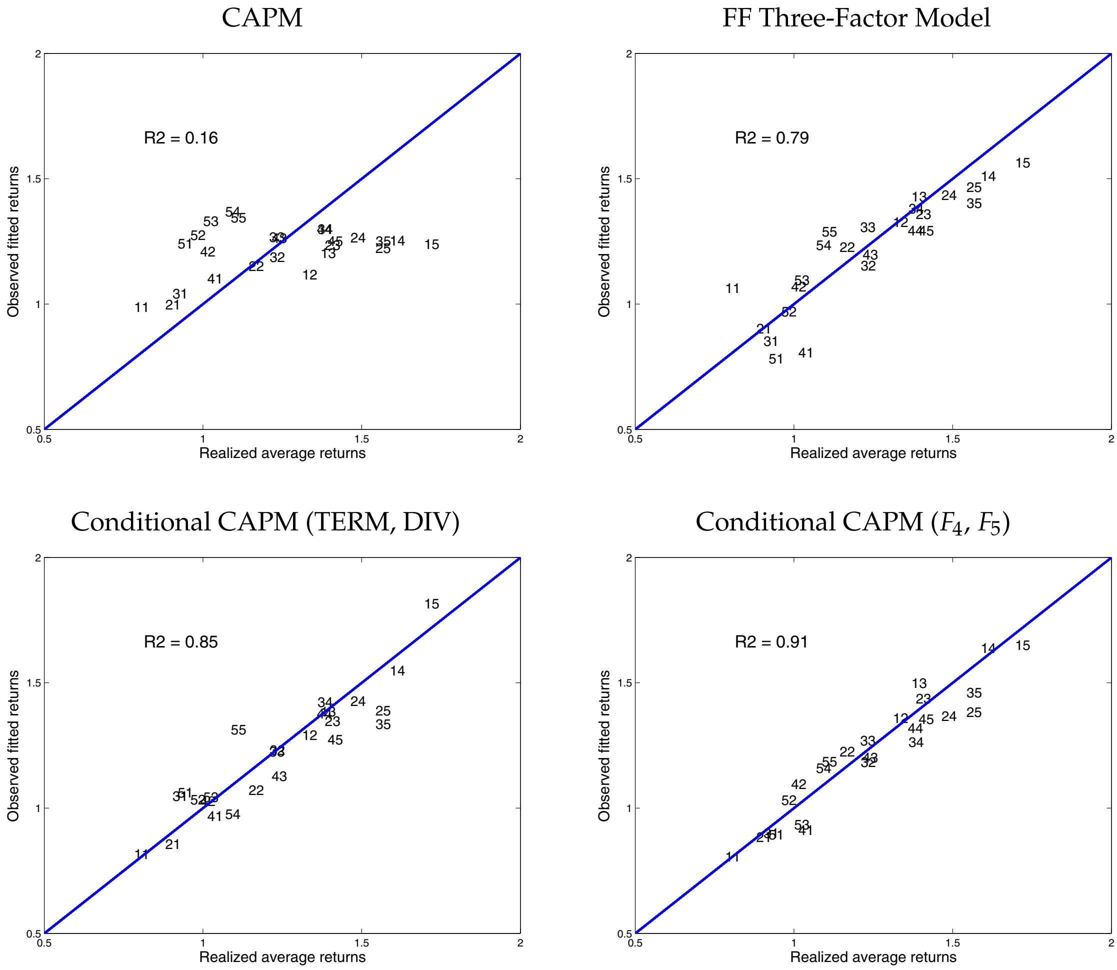
The tests documented above show that specifications of the conditional CAPM using diffusion indexes as instruments outperform specifications based on traditionally used conditioning variables. While these results indicate that it is advantageous to employ a large information set, they do not in general allow to conclude that diffusion indexes have incremental explanatory power over benchmark instruments. As discussed in Section 2.3, the hypothesis that some added instrument is useless can explicitly be tested in the GMM framework. In this section, I employ the two tests stated in (2.8) and (2.9) to evaluate whether the best performing pairs of benchmark conditioning variables and diffusion indexes capture the relevant pricing information or whether adding instruments significantly improves the model fit.

Table 2.4 summarizes the outcomes of these tests. The upper panel shows estimation results of conditional CAPM specifications based on the term spread and the dividend yield, to which the first five diffusion indexes are added one by one. The lower panel reports the estimation outcomes of conditional CAPM versions using the fourth and fifth diffusion index as conditioning variables to which the five benchmark instruments are added individually. The last two columns of Table 2.4 provide the statistics ΔJ and $\chi^2(bj = 0)$ on the basis of which I assess whether the added instruments are useless or not.

The results of these tests can be summarized as follows. First, despite the strong explanatory power of the instruments TERM and DIV documented above, the hypotheses $b_{F_j} = 0$ can be rejected at the 5% significance level for three out of

Figure 2.1: **Plots of Model-Fit Based on Fama-MacBeth Regressions**

This figure plots average realized returns of the 25 Fama-French portfolios against the fitted returns implied by four different models: the unconditional CAPM, the Fama-French three-factor model, and two specifications of the conditional CAPM using two benchmark instruments or diffusion indexes as conditioning variables. The portfolios are denoted by two digits where the first corresponds to the size quintile (1=small, 5=large) and the second to the book-to-market quintile (1=low, 5=high).



the first five diffusion indexes. This implies that these three factors incorporate useful pricing information that is not already captured by the two benchmark instruments. Not surprisingly, these useful conditioning variables are the diffusion indexes F_2, F_4, and F_5 which have been documented particularly powerful in tests of the conditional CAPM using only one instrument. Second, as the results in the lower panel of Table 2.4 show, the hypotheses $b_{Z_j} = 0$ cannot be rejected for any of the added instruments. Hence, none of the five benchmark conditioning variables adds useful information to the best-performing model specification based on diffusion indexes. One can therefore conclude that factors extracted from a large panel of macroeconomic data have incremental explanatory power over commonly used instruments, while the reverse is not true. This underlines that the information set available to investors is better captured by factors extracted from a large macroeconomic data panel than by individual variables.

2.5.4 Comparison with Quarterly Conditioning Variables

The results above show that conditioning on a large information set leads to a better fit of the conditional CAPM. In this section, I provide additional evidence for this finding by comparing diffusion indexes with two other popular instruments. These are the log consumption wealth ratio *cay* suggested by Lettau and Ludvigson [2001] and the labor income to consumption ratio s_w proposed by Santos and Veronesi [2006]. Both variables have been shown to have strong explanatory power for the cross-section of returns. Since they are only available on a quarterly basis, I transform all monthly series in the panel of macro times series into the quarterly frequency and extract diffusion indexes from the resulting quarterly dataset.[10]

Table 2.5 provides GMM estimation results obtained for different specifications of the conditional CAPM using quarterly diffusion indexes and the two quarterly benchmark conditioning variables as instruments. These results can be summarized as follows. First, *cay* and s_w both give rise to a Hansen-Jagannathan distance of 0.65 and thus seem to perform about equally well. According to the J_{HJ}-statistic, however, the log consumption-wealth ratio implies slightly smaller pricing errors ($J_{HJ} = 123.28$) than the labor income to consumption ratio ($J_{HJ} = 154.25$). Second and most importantly, the first two factors extracted from the large panel of quarterly time series strongly outperform *cay* and s_w. The first diffusion index implies a HJ-distance of 0.62 and a J_{HJ}-statistic of 67.28, which shows that this factor produces considerably smaller pricing errors than the two bench-

[10] Note that different strategies of temporal aggregation have been employed for various groups of time series. In particular, for all interest rate series, exchange rates, monetary aggregates, stock market indices, and business outlook indices I have defined the last monthly observation in a quarter as the quarterly figure. For all other series, the average of three consecutive monthly observations has been used as the quarterly figure.

Table 2.4: **GMM Tests of Incremental Explanatory Power**

This table summarizes GMM estimation results for tests of incremental explanatory power of single instruments. In the first panel, diffusion indexes are added to a conditional CAPM specification using the term spread and the dividend yield as conditioning variables. In the second panel, benchmark instruments are added to a specification based on the fourth and fifth diffusion index. For each specification, estimates of the pricing coefficients are provided in the first row and the corresponding t-values are given in parentheses below. $\chi^2(b_j = 0)$ denotes the Chi-square test for the hypothesis $b_j = 0$ for $j = \{F, Z\}$ defined in (2.8), ΔJ refers to the χ^2 difference test defined in (2.9). The 25 Fama-French portfolios are used as test assets. The sample period is 1963:01-2003:12.

cst	$R_{m,t}$	Z_{t-1}		$R_{m,t} \cdot Z_{t-1}$		F_{t-1}	$R_{m,t} \cdot F_{t-1}$	ΔJ	$\chi^2(b_F = 0)$
		TERM	DIV	TERM	DIV	F_1			
0.39	-0.03	-0.18	0.04	0.01	0.01	0.17	0.00	1.30	2.05
(1.11)	(-.30)	(-1.69)	(.72)	(.43)	(.73)	(1.15)	(.07)	(.52)	(.36)
		TERM	DIV	TERM	DIV	F_2			
0.24	-0.19	-0.25	0.09	0.02	0.03	0.27	0.06	10.78	6.09
(.75)	(-1.81)	(-2.02)	(1.77)	(.82)	(2.06)	(1.07)	(1.78)	**(.00)**	**(.05)**
		TERM	DIV	TERM	DIV	F_3			
0.63	-0.18	-0.33	0.07	0.04	0.02	0.15	-0.02	4.13	4.05
(2.07)	(-1.72)	(-2.92)	(1.54)	(1.68)	(1.83)	(1.38)	(-.70)	(.13)	(.13)
		TERM	DIV	TERM	DIV	F_4			
0.13	0.04	-0.01	0.02	-0.00	-0.00	0.01	-0.07	8.78	8.41
(.47)	(.53)	(-.07)	(.68)	(-.06)	(-.03)	(.07)	(-2.65)	**(.01)**	**(.01)**
		TERM	DIV	TERM	DIV	F_5			
0.78	-0.23	-0.29	0.04	0.05	0.03	-0.28	0.01	8.05	6.85
(2.39)	(-2.17)	(-2.99)	(.87)	(2.11)	(2.25)	(-2.04)	(.31)	**(.02)**	**(.03)**

cst	$R_{m,t}$	F_{t-1}		$R_{m,t} \cdot F_{t-1}$		Z_{t-1}	$R_{m,t} \cdot Z_{t-1}$	ΔJ	$\chi^2(b_Z = 0)$
		F_4	F_5	F_4	F_5	TERM			
0.38	0.03	-0.19	-0.29	-0.06	-0.01	-0.03	-0.00	4.58	0.11
(2.28)	(1.15)	(-.95)	(-2.08)	(-1.75)	(-.23)	(-.29)	(-.04)	(.10)	(.95)
		F_4	F_5	F_4	F_5	DEF			
0.26	-0.00	-0.13	-0.23	-0.07	0.01	0.10	0.03	4.62	0.34
(.90)	(-.02)	(-.65)	(-1.62)	(-2.38)	(.22)	(.37)	(.38)	(.10)	(.84)
		F_4	F_5	F_4	F_5	TB1			
0.37	-0.05	-0.21	-0.33	-0.07	0.01	0.07	0.13	4.53	1.14
(1.22)	(-.66)	(-1.00)	(-1.97)	(-2.15)	(.54)	(.11)	(1.05)	(.10)	(.57)
		F_4	F_5	F_4	F_5	TB31			
0.94	-0.01	-0.12	-0.45	-0.08	0.02	-0.18	0.01	4.31	2.29
(2.22)	(-.16)	(-.54)	(-2.34)	(-2.33)	(.73)	(-1.46)	(.51)	(.12)	(.32)
		F_4	F_5	F_4	F_5	DIV			
0.40	-0.06	-0.21	-0.32	-0.07	0.02	0.01	0.01	4.48	1.44
(1.15)	(-.79)	(-.95)	(-1.80)	(-2.09)	(.66)	(.09)	(1.14)	(.11)	(.49)

Table 2.5: **Conditional CAPM with One Instrument, Quarterly Data**

This table summarizes GMM estimation results for different specifications of the conditional CAPM based on one conditioning variable using quarterly data. The 25 Fama-French portfolios are used as test assets. The estimation period is from 1963:Q1 to 2003:Q4.

Z_{t-1}	HJ-dist	J	J_{HJ}	supLM
F_1	0.62	35.12	67.28	***18.28
	(.67)	(.04)	(.00)	
F_2	0.49	19.02	48.49	9.31
	(.95)	(.64)	(.00)	
F_3	0.64	44.89	145.48	3.70
	(.87)	(.00)	(.00)	
F_4	0.64	9.99	173.25	***21.19
	(.96)	(.99)	(.00)	
F_5	0.64	20.93	189.81	***20.80
	(.76)	(.53)	(.00)	
F_6	0.63	103.37	152.66	***25.71
	(.71)	(.00)	(.00)	
F_7	0.64	47.52	134.23	*15.39
	(.50)	(.00)	(.00)	
cay	0.65	60.30	123.28	***24.28
	(.07)	(.00)	(.00)	
s_w	0.65	78.53	154.25	11.68
	(.31)	(.00)	(.00)	

mark instruments. Strikingly, the second diffusion index implies even smaller pricing errors, as it gives rise to a HJ-distance of 0.49 and a J_{HJ}-statistic of 49.49. The relative performance of the remaining factors in comparison with the two benchmark instruments is somewhat less pronounced. The diffusion indexes $F_3, \ldots, F_7$ imply smaller Hansen-Jagannathan distance measures than the two variables *cay* and s_w. Yet, they give rise to J_{HJ}-statistics that are larger than the one implied by *cay* and smaller than the corresponding value of s_w only for F_3, F_6, and F_7. Overall, however, the diffusion indexes strongly withstand the comparison with the two quarterly benchmark conditioning variables. The results obtained using quarterly data thus underscore the main conclusion drawn before: conditioning on a large information set significantly improves the fit of the conditional CAPM with respect to specifications based on individual conditioning variables.

2.5.5 Robustness Tests

In this section, some of the results documented above are subjected to three types of robustness tests. The first one regards the choice of test assets, the second the stability of the results across different sample periods, and the third investigates whether the strong relative pricing performance of diffusion indexes is limited to the CAPM or whether this result extends to other conditional models.

Scaled Returns

Following Cochrane [1996], one can test whether an asset pricing model is able to price a different set of assets by considering managed portfolios, i.e. portfolios that are reweighed period by period according to new information about expected returns. To construct such managed portfolios, he proposes to scale the returns of the test assets with conditioning variables that captures news about future returns. I follow this suggestion and scale the 25 Fama French portfolios with the term spread, the default spread, the short rate, and the dividend yield, respectively. All variables have previously been used to construct managed portfolios (see e.g. Li et al. 2004). Table 2.7 summarizes the results of GMM tests of the conditional CAPM with two instruments using the scaled portfolio returns as test assets. A comparison of these results with Table 2.2 shows that the strong relative performance of the conditional CAPM using diffusion indexes as instruments prevails when the model is called to price alternative sets of assets. Indeed, both in terms of the HJ-distance and the J_{HJ}-statistic, the worst performing specification using diffusion indexes as conditioning variables outperforms the best performing conditional CAPM based on benchmark instruments for *all* four sets of scaled portfolios. This is a striking result since the employed scaling variables should by construction capture the variation of expected returns across the scaled test assets pretty well.

Subsample Analysis

The second robustness check aims at detecting whether the strong ability of diffusion indexes to proxy for investors' information sets is specific to the sample used in the previous tests. I thus repeat the tests documented in Tables 2.1 and 2.2 using different sample periods. To keep things simple, I split the estimation period in two subsamples of about the same length, 1963:01-1983:12 and 1984:01-2003:12.

The outcomes of GMM tests for these two subsamples are reported in Tables 2.8 and 2.9. Several remarks are in order. First, in the early subsample 1963-1983, the panel information criteria by Bai and Ng [2002] indicate that seven factors are sufficient to summarize the common variation in the panel of macro time series. Hence, I extract seven diffusion indexes using the method by Stock and Watson. When only one conditioning variable is employed, five out of these seven diffusion indexes imply smaller pricing errors for the conditional CAPM than all benchmark instruments. The same holds true for the subsample 1984-2003 for which eight factors capture the common variation in the macroeconomic dataset. Hence, the conclusion drawn before that individual diffusion indexes represent better proxies for investors' information sets than individual instruments is not sample-specific. Note that different conditioning variables may in principle be most useful in different subsamples. According to the results in Table 2.8, this is the case here. While the fourth common factor is the best performing single diffusion index in the early sample period, the fifth diffusion index clearly stands out in the subsample 1984-2003. Equivalently, the default spread implies smaller pricing errors over the second subsample than the term spread which performs best over the period 1963-1983. Combining the single best conditioning variables with each of the four remaining instruments, one can see from the results in Table 2.9 that pairs of diffusion indexes strongly outperform pairs of individual benchmark indicators in both subsamples.

Tests of the Conditional Consumption-CAPM

The results presented so far show that diffusion indexes outperform commonly used conditioning variables in tests of the conditional CAPM. A potential concern with respect to these results might point at the usefulness of dynamic factors as instruments in other conditional pricing models. Indeed, as the results in Hodrick and Zhang [2001] and Wang [2005] show, specific conditioning variables must not be equally useful in combination with different pricing factors. The Consumption-CAPM (CCAPM henceforth) originally due to Breeden [1979] has recently been revived by the work of Lettau and Ludvigson [2001] who have shown that conditional versions of the model explain the cross-section of returns well. To underscore my previous results, I therefore compare the fit of the conditional CCAPM using diffusion indexes as instruments with specifications based on benchmark conditioning variables. Hence, I compare models of the form

$$m_{t+1} = a_t + b_t \Delta c_{t+1},$$

where Δc_{t+1} denotes the growth rate of aggregate consumption and where a_t and b_t are linear functions of the conditioning variables as before.

Table 2.10 documents the outcomes of GMM estimations of the conditional CCAPM based on different instruments. The left and right panel summarize the results obtained using monthly and quarterly data, respectively. According to the estimates of the HJ-distance and the J_{HJ}-statistic, the conditional CCAPM implies slightly larger pricing errors than the corresponding specifications of the conditional CAPM, reported in Tables 2.1 and 2.5. More importantly, however, the overall pattern found before that most diffusion indexes perform better than benchmark instruments extends to the CCAPM. This holds true both for monthly and quarterly instruments. Hence, the strong relative performance of dynamic factors as conditioning variables is not specific to tests of the conditional CAPM. Altogether, the robustness checks carried out in this section provide strong evidence that the usefulness of dynamic factors as instruments in tests of conditional asset pricing models is independent from the set of test assets, the sample-period considered, and the particular choice of pricing factors.

2.5.6 What Information Do the Factors Summarize?

Table 2.6: Correlation of Diffusion Indexes and Benchmark Instruments

This table reports sample correlations between the diffusion indexes extracted from the panel of 132 monthly time series for the US and the five monthly benchmark conditioning variables introduced in Section 2.4. The sample period is 1963:01-2003:12.

	TERM	DEF	TB1	TB31	DIV
F_1	0.09	-0.52	-0.36	-0.34	-0.33
F_2	0.58	-0.06	-0.59	-0.56	-0.64
F_3	0.41	0.53	0.09	0.34	0.13
F_4	0.16	0.18	-0.09	0.10	-0.10
F_5	-0.02	-0.00	-0.05	-0.32	-0.04
F_6	0.43	-0.04	-0.19	0.12	-0.18
F_7	0.08	0.25	0.33	0.08	0.31
F_8	-0.19	0.13	0.18	0.16	0.17

So far, this chapter has shown that factors which by construction summarize the information in a large number of macroeconomic time series imply significantly smaller pricing errors than commonly used information variables when they are employed as instruments. Since not all diffusion indexes are equally useful, though, it is interesting to investigate which dimensions of conditioning information are more important than others. The factors extracted via the method by Stock and Watson [2002a,b] are only identifiable up to a nonsingular $q \times q$ rotation. Hence, a detailed interpretation is unwarranted. Nevertheless, it may be instructive to briefly characterize those diffusion indexes which have been found to be particularly useful instruments. Consider first the sample correlations between the extracted factors and the benchmark conditioning variables, summarized in Table 2.6. There is some substantial correlation between individual factors and benchmark conditioning variables. For example, the second diffusion index is strongly correlated with the term spread, the short-term interest rate, and the dividend yield. However, the fourth and fifth diffusion index which have also been shown to be useful conditioning variables are largely uncorrelated with the benchmark instruments. This indicates that both sets of instruments only partly capture similar conditioning information.

For a more detailed investigation of the economic underpinnings of the extracted factors, Table 2.11 lists those variables which are most highly correlated with the first five diffusion indexes. According to the results from the various tests of the conditional CAPM documented above, the diffusion indexes F_2, F_4, and F_5 stand out. I therefore focus on discussing these factors. The second factor is clearly an inflation-related index since it is most highly correlated with various price indices.[11] This is interesting since inflation has not previously been used as a conditioning variable.[12] Next consider the fourth factor which has individually implied the smallest pricing errors, strongly outperforming all benchmark conditioning variables. According to Table 2.11, this factor is mainly driven by variables related to the housing industry. This is an interesting result because a recent literature identifies housing risk as a factor that influences agents' investment decisions (see e.g. Lustig and Van Nieuwerburgh 2005, and Piazzesi, Schneider, and

[11] This is consistent with the results in Stock and Watson [2002a] who also find the second diffusion index extracted from a similarly constructed dataset to be an inflation factor.

[12] Notice, however, that some link between unexpected inflation and stock returns has previously been documented. Using different model setups and empirical approaches, Chen et al. [1986], Kelly [2003], and more recently Aretz, Bartram, and Pope [2006] find (weak) evidence that unexpected inflation is related to the Fama-French factor SMB or is itself a factor in the cross-section of stock returns.

Tuzel 2006). Both studies show that valuation ratios based on housing wealth or housing consumption forecast stock returns. My result that a housing factor is a highly useful conditioning variable thus provides additional evidence for the important role of housing-related information in stock returns. Notice finally that the fifth diffusion index is mainly driven by interest rate series and thus bears a natural relation with benchmark instruments such as the term spread or the default spread.

2.6 Conclusion

Tests of conditional asset pricing models require the specification of proxies for the information set available to investors. Degrees of freedom problems restrict to very few the number of conditioning variables that can be used at the same time. It is therefore a common practice to employ individual variables as instruments. Consequently, most tests of conditional pricing models are based on very limited amounts of conditioning information.

This chapter suggests to exploit a broad conditioning information set by using dynamic factors as instruments. Dynamic factors by construction summarize the information in many economic variables and are therefore intuitively appealing proxies for the information set available to investors. I show that dynamic factors imply substantially smaller pricing errors than commonly employed conditioning variables when they are used as instruments in tests of conditional models. Moreover, dynamic factors exhibit significant incremental explanatory power over benchmark instruments. The obtained results strongly support the hypothesis that dynamic factors represent better proxies for investors' information set than individual indicators. The results withstand a number of robustness tests and carry important implications for the specification of conditional asset pricing models in applied research and practice.

A.2 Additional Tables and Figures

Table 2.7: **GMM Tests Using Scaled Returns**

This table summarizes GMM estimation results for different specifications of the conditional CAPM based on two conditioning variables, respectively. The test assets are the 25 Fama-French portfolios scaled by different conditioning variables. The upper left panel provides results for portfolios scaled by the term spread, the upper right panel by the default spread. the lower left panel by the 1-month Treasury Bill and the lower right panel by the dividend yield.

	Returns Scaled by TERM				Returns Scaled by DEF			
Z_{t-1}	HJ-dist	J	J_{HJ}	supLM	HJ-dist	J	J_{HJ}	supLM
F_1, F_4	0.11	22.61	38.38	13.19	0.11	32.66	41.41	13.97
	(0.84)	(0.25)	(0.01)		(.63)	(.03)	(.00)	
F_2, F_4	0.10	18.23	25.92	14.05	0.10	22.35	29.02	*18.95
	(0.85)	(0.51)	(0.13)		(.64)	(.27)	(.07)	
F_3, F_4	0.11	19.14	34.35	7.37	0.11	27.73	41.73	12.87
	(0.86)	(0.45)	(0.02)		(.59)	(.09)	(.00)	
F_4, F_5	0.10	10.19	17.73	8.79	0.10	16.31	23.44	10.14
	(0.98)	(0.95)	(0.54)		(.95)	(.64)	(.22)	
TERM, DEF	0.13	33.87	50.01	*18.30	0.12	34.96	49.57	16.85
	(0.37)	(0.02)	(0.00)		(.36)	(.01)	(.00)	
TERM, TB1	0.12	22.66	40.79	17.06	0.12	24.24	46.20	**24.08
	(0.47)	(0.25)	(0.00)		(.50)	(.19)	(.00)	
TERM, TB31	0.13	44.26	56.97	15.33	0.13	41.74	58.41	***33.16
	(0.29)	(0.00)	(0.00)		(.29)	(.00)	(.00)	
TERM, DIV	0.12	21.57	39.07	17.27	0.12	22.04	45.55	***30.12
	(0.59)	(0.31)	(0.00)		(.59)	(.28)	(.00)	

	Returns Scaled by TB1				Returns Scaled by DIV			
Z_{t-1}	HJ-dist	J	J_{HJ}	supLM	HJ-dist	J	J_{HJ}	supLM
F_1, F_4	0.22	24.69	36.38	12.95	0.04	35.29	46.84	14.29
	(.68)	(.17)	(.01)		(.52)	(.01)	(.00)	
F_2, F_4	0.16	17.60	21.07	10.31	0.03	27.48	33.70	8.14
	(.96)	(.55)	(.33)		(.56)	(.09)	(.02)	
F_3, F_4	0.23	27.27	40.93	10.88	0.04	20.30	36.69	9.35
	(.57)	(.10)	(.00)		(.74)	(.38)	(.01)	
F_4, F_5	0.21	27.27	32.96	13.73	0.03	18.10	27.66	11.37
	(.68)	(.10)	(.02)		(.87)	(.52)	(.09)	
TERM, DEF	0.25	44.19	77.58	8.63	0.04	34.69	64.98	13.42
	(.25)	(.00)	(.00)		(.42)	(.02)	(.00)	
TERM, TB1	0.25	36.81	67.64	7.12	0.04	32.46	59.33	16.61
	(.39)	(.01)	(.00)		(.43)	(.03)	(.00)	
TERM, TB31	0.25	45.09	70.86	16.26	0.04	35.60	68.09	11.30
	(.26)	(.00)	(.00)		(.24)	(.01)	(.00)	
TERM, DIV	0.25	37.19	64.61	7.56	0.04	31.51	58.71	16.88
	(.39)	(.01)	(.00)		(.51)	(.04)	(.00)	

Table 2.8: **Subsample Analysis - Conditional CAPM with One Instrument**

This table summarizes GMM estimation results for different specifications of the conditional CAPM based on one conditioning variable, respectively. The left-hand panel provides results for the sample period 1963:01-1983:12, and the right-hand panel for the period 1984:01-2003:12. The 25 Fama-French portfolios are used as test assets.

	1963:01 - 1983:12				1984:01 - 2003:12			
Z_{t-1}	HJ-dist	J	J_{HJ}	supLM	HJ-dist	J	J_{HJ}	supLM
F_1	0.21	374.62	230.41	***79.64	0.15	66.86	96.15	9.11
	(.00)	(.00)	(.00)		(.12)	(.00)	(.00)	
F_2	0.19	44.87	286.47	8.42	0.15	44.57	96.03	**20.68
	(.63)	(.00)	(.00)		(.23)	(.00)	(.00)	
F_3	0.20	28.40	144.58	**20.69	0.15	61.29	***94.48	3.17
	(.49)	(.13)	(.00)		(.43)	(.00)	(.00)	
F_4	0.19	39.84	100.61	8.95	0.16	98.64	140.74	***78.32
	(.40)	(.01)	(.00)		(.01)	(.00)	(.00)	
F_5	0.20	95.13	168.22	***20.96	0.13	36.86	69.07	***23.54
	(.17)	(.00)	(.00)		(.63)	(.02)	(.00)	
F_6	0.21	62.17	159.73	13.42	0.15	65.93	130.02	11.24
	(.31)	(.00)	(.00)		(.07)	(.00)	(.00)	
F_7	0.20	66.56	126.33	***26.07	0.16	84.85	116.50	13.18
	(.07)	(.00)	(.00)		(.05)	(.00)	(.00)	
F_8	-	-	-	-	0.15	40.71	69.64	8.93
					(.20)	(.01)	(.00)	
TERM	0.18	36.09	172.79	9.58	0.15	95.83	121.46	***58.27
	(.82)	(.02)	(.00)		(.09)	(.00)	(.00)	
DEF	0.19	43.56	198.94	***71.32	0.15	72.30	101.09	***67.96
	(.50)	(.00)	(.00)		(.23)	(.00)	(.00)	
TB1	0.21	83.88	225.46	8.65	0.15	46.58	106.61	***27.36
	(.17)	(.00)	(.00)		(.10)	(.00)	(.00)	
TB31	0.21	37.46	206.11	***24.24	0.16	114.91	137.92	***38.22
	(.24)	(.01)	(.00)		(.01)	(.00)	(.00)	
DIV	0.21	84.62	225.24	12.78	0.15	47.60	113.06	***22.35
	(.03)	(.00)	(.00)		(.10)	(.00)	(.00)	

Table 2.9: **Subsample Analysis - Conditional CAPM with Two Instruments**

This table summarizes GMM estimation results for different specifications of the conditional CAPM based on two conditioning variables, respectively. The left-hand panel provides results for the sample period 1963:01-1983:12, and the right-hand panel for the period 1984:01-2003:12. Notice that due to their superior individual performance over the second subsample, the instruments F_5 and DEF in parentheses are used instead of F_4 and $TERM$ in tests using data for the period 1984-2003. The 25 Fama-French portfolios are used as test assets. For comparison, the last two panels provide the estimation results for the unconditional CAPM and the Fama-French (1993) three-factor model, respectively.

	1963:01 - 1983:12				1984:01 - 2003:12			
Z_{t-1}	HJ-dist	J	J_{HJ}	supLM	HJ-dist	J	J_{HJ}	supLM
$F_1, F_4(F_5)$	0.19	25.74	***77.02	6.62	0.11	16.84	33.69	**21.08
	(.59)	(.14)	(.00)		(.96)	(.60)	(.02)	
$F_2, F_4(F_5)$	0.17	54.40	122.00	11.5	0.13	27.29	61.38	***45.24
	(.40)	(.00)	(.00)		(.58)	(.10)	(.00)	
$F_3, F_4(F_5)$	0.18	26.00	56.25	11.2	0.13	61.18	61.22	15.49
	(.44)	(.13)	(.00)		(.48)	(.00)	(.00)	
F_4, F_5	0.18	37.23	67.78	7.27	0.13	30.97	62.66	***49.39
	(.39)	(.01)	(.00)		(.59)	(.04)	(.00)	
TERM, DEF	0.18	35.49	177.51	***42.88	0.15	65.55	110.86	***70.46
	(.71)	(.01)	(.00)		(.00)	(.00)	(.00)	
TERM (DEF), TB1	0.18	29.86	153.94	**23.36	0.15	53.00	101.90	***34.49
	(.60)	(.05)	(.00)		(.00)	(.00)	(.00)	
TERM (DEF), TB31	0.18	37.69	163.12	***25.7	0.15	44.42	72.17	***25.38
	(.57)	(.01)	(.00)		(.00)	(.00)	(.00)	
TERM (DEF), DIV	0.17	25.94	140.85	**21.94	0.15	48.73	101.27	***26.14
	(.74)	(.13)	(.00)		(.00)	(.00)	(.00)	
CAPM	0.21	310.78	215.44	3.99	0.16	56.34	126.60	1.58
	(.00)	(.00)	(.00)		(.11)	(.00)	(.00)	
FF	0.16	78.26	72.42	4.37	0.14	86.72	83.24	17.70
	(.00)	(.00)	(.00)		(.00)	(.00)	(.00)	

Table 2.10: **Conditional CCAPM with One Instrument**

This table summarizes GMM estimation results for different specifications of the conditional Consumption-CAPM based on a single conditioning variable. The 25 Fama-French portfolios are used as test assets. The estimation period is from 1963:01 to 2003:12.

	Monthly Data				Quarterly Data		
Z_{t-1}	HJ-dist	J	J_{HJ}	Z_{t-1}	HJ-dist	J	J_{HJ}
F_1	0.18	26.41	80.90	F_1	0.67	26.70	256.25
	(.75)	(.19)	(.00)		(.80)	(.22)	(.00)
F_2	0.16	19.24	45.26	F_2	0.53	24.07	59.84
	(.71)	(.57)	(.00)		(.89)	(.34)	(.00)
F_3	0.17	20.60	64.02	F_3	0.68	13.48	329.74
	(.82)	(.48)	(.00)		(1.00)	(.92)	(.00)
F_4	0.17	18.85	50.73	F_4	0.63	33.74	100.75
	(.86)	(.59)	(.00)		(.71)	(.05)	(.00)
F_5	0.15	20.00	50.04	F_5	0.67	26.88	369.62
	(.68)	(.52)	(.00)		(.80)	(.22)	(.00)
F_6	0.17	25.14	82.15	F_6	0.67	43.35	325.74
	(.73)	(.24)	(.00)		(.29)	(.00)	(.00)
F_7	0.15	10.90	30.65	F_7	0.64	43.58	167.82
	(.90)	(.96)	(.08)		(.62)	(.00)	(.00)
F_8	0.16	31.87	67.93				
	(.50)	(.06)	(.00)				
TERM	0.17	10.74	58.92	cay	0.67	14.65	306.60
	(.97)	(.97)	(.00)		(.97)	(.88)	(.00)
DEF	0.17	35.02	79.39	s_w	0.67	5.75	347.98
	(.31)	(.03)	(.00)		(.99)	(.99)	(.00)
TB1	0.18	15.06	86.72				
	(.84)	(.82)	(.00)				
TB31	0.17	25.46	75.25				
	(.75)	(.23)	(.00)				
DIV	0.18	12.74	89.51				
	(.89)	(.92)	(.00)				

Table 2.11: **Factor Variance Decomposition**

This table summarizes R-squares of univariate regressions of the first five factors extracted from the macroeconomic dataset on all individual variables. I report the variables that are most highly correlated with the individual factors, respectively. Notice that the series have been transformed to be stationary prior to extraction of the factors, i.e. for most variables the regressions correspond to regressions on growth rates. The five diffusion indexes together explain more than 60% of the total variation in the panel.

Mnemonic	Description	R^2
	Factor 1 (27.3% of total variance)	
IPS43	INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)	0.80
IPS10	INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX	0.80
PMP	NAPM PRODUCTION INDEX	0.77
CES003	EMPLOYEES ON NONFARM PAYROLLS - GOODS-PRODUCING	0.77
CES002	EMPLOYEES ON NONFARM PAYROLLS - TOTAL PRIVATE	0.73
	Factor 2 (15.9% of total variance)	
PUC	CPI-U: COMMODITIES	0.67
GMDCN	PCE,IMPL PR DEFL:PCE; NONDURABLES	0.67
PWFCSA	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS	0.61
PWFSA	PRODUCER PRICE INDEX: FINISHED GOODS	0.60
PUXHS	CPI-U: ALL ITEMS LESS SHELTER	0.60
	Factor 3 (7.3% of total variance)	
A0M077	Ratio, mfg. and trade inventories to sales	0.35
PU85	CPI-U: MEDICAL CARE	0.30
A0M070	Manufacturing and trade inventories	0.29
LHU680	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS	0.26
aom001	Average weekly hours, mfg.	0.26
	Factor 4 (5.7% of total variance)	
HSFR	HOUSING STARTS:TOTAL FARM&NONFARM(	0.45
HSBR	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS	0.41
HSBWST	HOUSES AUTHORIZED BY BUILD. PERMITS:WEST	0.40
HSSOU	HOUSING STARTS:SOUTH	0.39
HSWST	HOUSING STARTS:WEST	0.36
	Factor 5 (4% of total variance)	
HSBSOU	HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH	0.26
FYGT1	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.	0.23
FYGM6	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.	0.22
FYGT5	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.	0.22
FYGM3	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.	0.21

3 Forecasting the Yield Curve in a Data-Rich Environment

This chapter suggests a term structure model which parsimoniously exploits a broad macroeconomic information set. The model does not incorporate latent yield curve factors, but instead uses the common components of a large number of macroeconomic variables and the short rate as explanatory factors. Precisely, an affine term structure model with parameter restrictions implied by no-arbitrage is added to a Factor-Augmented Vector Autoregression (FAVAR). The model is found to strongly outperform different benchmark models in out-of-sample yield forecasts, reducing root mean squared forecast errors relative to the random walk up to 50% for short and around 20% for long maturities.

3.1 Introduction

Traditional models of the term structure of interest rates are built upon decompositions of yields into latent factors using one or another statistical method (e.g. Nelson and Siegel 1987, Knez et al. 1994, Duffie and Kan 1996). While the fit of these models is usually rather good, their economic meaning is somewhat limited since they have relatively little to say about the relationship between observable economic variables and interest rates of different maturities. To explore this issue, one therefore needs to construct models which jointly describe macro and term structure dynamics.

In a seminal paper, Ang and Piazzesi [2003] augment a standard latent factor affine term structure model with two macroeconomic factors. They find that the included macroeconomic variables improve yield forecasts, accounting for up to 85 % of the variation in interest rates. Inspired by this finding, a vivid literature has emerged lately which explores different approaches of jointly modelling the term structure and the macroeconomy. Hördahl et al. [2006], for example, build a small structural model that describes the joint evolution of output, inflation, and the monetary policy instrument, and add the term structure using no-arbitrage restrictions. Diebold et al. [2006] estimate a model which allows for correlated latent and observed macroeconomic factors and find that macroeconomic variables have strong effects on future movements of the yield curve, while latent interest rate factors have a relatively small impact on macroeconomic variables. Further examples of recent papers which jointly model term structure and macro dynamics are e.g. Dewachter and Lyrio [2006], Wu [2002], Rudebusch and Wu [2003], and Dai and Philippon [2005].

Based on different model setups and methodologies, these papers consistently argue that macroeconomic variables are useful for explaining and/or forecasting government bond yields. However, only very small macroeconomic information sets are commonly exploited in these studies. The main reason for this informational limitation is that state-of-the-art affine term structure models imply the estimation of a large number of parameters, thereby considerably restricting the number of explanatory variables one can include. Yet, by limiting the analysis to only a few variables, other potentially useful macroeconomic information is being neglected.

A recent strand of the macroeconomic literature advances the use of dynamic factor models in order to exploit large information sets in economic analysis (e.g. Stock and Watson 2002a,b, Forni et al. 2005). Such models break down the cross-sectional information contained in large panels of time series into common and series-specific components, and thereby enable the researcher to separate out aggregate and idiosyncratic shocks. A number of studies have found that dynamic factor models are particularly powerful in forecasting economic time series, especially measures of output and inflation (e.g. Stock and Watson 2002a, Giannone et al. 2004).

In this chapter, I examine the usefulness of factors extracted from a large macroeconomic dataset for explaining and forecasting the term structure of interest rates. The exercise is motivated by three previously documented results. First, it has been argued by some authors that central banks actively monitor a large number of macroeconomic time series, and that monetary policy decisions are thus based on the information contained in not only a few key aggregates but many economic variables. Loosely speaking, the central bank sets interest rates in a "data-rich environment" (Bernanke and Boivin 2003). Accordingly, dynamic factors which effectively summarize the information contained in a large number of time series should prove useful in explaining interest rates set by central banks. Bernanke and Boivin [2003], Favero et al. [2005] and Belviso and Milani [2005] consistently provide empirical evidence supporting this claim. By comparing standard Taylor rules with specifications based on dynamic factors, these papers show that the latter exhibit information beyond output and inflation that helps to explain monetary policy. Moreover, Giannone et al. [2004] show that factor-based forecasts of the federal funds rate perform as good as market-based forecasts. Overall, short-term interest rates are thus well explained by dynamic factors. Second, factor models have been shown to perform well in forecasting measures of output and inflation (see, e.g. Stock and Watson 2002a). Since both expected output and expected inflation are likely to have an impact on bond yields, this delivers another argument for using them in a term structure model. Finally, in Mönch (2006, Chapter 2) I show that dynamic factors are highly useful instruments in tests of conditional pricing models which implies that they capture well the time-variation of risk premia. Altogether, since the prices and yields of non-defaultable government bonds are driven by expectations about future short-term interest rates, expected future inflation and risk premia, the evidence pointed to above suggests that factors extracted from large panels have explanatory power also for the yield curve.

What is the appropriate model setup for incorporating a broad macroeconomic information set into term structure analysis through the use of dynamic factors? In a recent paper, Bernanke et al. [2005] suggest to combine the advantages of factor modeling and structural VAR analysis by estimating a joint vector-autoregression of factors extracted from a large cross-section of time series and the short-term interest rate, an approach which they label "factor-augmented VAR" (FAVAR). The FAVAR model provides a dynamic characterization of short-term interest rates set by the central bank in response to the main economic shocks which are summarized by a few common factors. As a by-product, it delivers a path of expected future short rates conditional on a broad macroeconomic information set. On the other hand, given a short rate equation, affine term structure models provide a tool to build up the entire yield curve subject to no-arbitrage restrictions. It is thus an obvious next step to combine a factor-augmented VAR model with the standard affine setup by using the FAVAR as the state equation in an essentially affine term structure model. This is done in the present chapter.

Estimation of my model is in two steps. First, I extract common factors from a large macroeconomic dataset using principal components and estimate the parameters governing their joint dynamics with the monetary policy instrument in a VAR. Second, I estimate a no-arbitrage vector autoregression of yields on the exogenous pricing factors. Specifically, I obtain the price of risk parameters by minimizing the sum of squared fitting errors of the model following the nonlinear least squares approach suggested by Ang, Piazzesi, and Wei [2006]. Since my model does not include latent yield curve factors, the parameters governing the dynamics of the state variables can be estimated separately by standard OLS. Hence, estimation is fast which makes the model particularly useful for recursive out-of-sample forecasts which are the main focus of this study.

The results of the chapter can be summarized as follows. A model including as factors the short rate and four common components which together explain the bulk of variation in a large panel of monthly macroeconomic time series variables for the US, provides a good in-sample fit of the term structure of interest rates. Preliminary regressions show that factors extracted from a large macroeconomic dataset contain information useful for explaining the federal funds rate beyond output and inflation. Moreover, the model factors are highly significant explanatory variables for yields. Compared to a model which incorporates the short rate and four individual measures of output and inflation as factors, there is a clear advantage in using the larger macroeconomic information set. The results from out-of-sample forecasts of yields underpin this finding. The term structure model based on common factors clearly outperforms the model based on individual variables for all maturities at all horizons. More importantly, the No-Arbitrage FAVAR model shows a striking superiority with respect to a number of benchmark models in out-of-sample yield forecasts. At forecast horizons beyond one month ahead, the model outperforms the random walk, a standard three-factor affine model and the model recently suggested by Diebold and Li [2006], reducing root mean squared forecast errors relative to the random walk up to 50% for short and around 20% for long maturities

The remainder of this chapter is summarized as follows. In Section 3.2, the affine term structure model based on common dynamic macro factors is motivated and its exact parametrization discussed. Section 3.3 briefly sketches the method used to estimate the model. Additional details on the estimation and on the implementation of a test of relative forecast performance are given in Appendix 6.2. In Section 3.4, I first provide some preliminary evidence on the usefulness of dynamic macro factors to explain yields and then discuss the results of the out-of-sample forecasts in Section 3.5. Section 3.6 concludes the chapter.

3.2 The Model

Monetary policy decisions are likely based on the information contained in not only a few key aggregates but many economic variables. Yet, it is infeasible to empirically model the short-term interest rate set by the central bank as a function of a large number of individual variables. Economists therefore customarily map the monetary policy instrument to a few variables, including mostly a measure of the output gap and a measure of inflation. A convenient way of keeping track of a plethora of information without including too many variables into a model, is to think of all macroeconomic variables as being driven by a few common factors and an idiosyncratic component. In such a setup, the reaction of the monetary policy maker to shocks affecting different categories of economic variables can be modeled by relating the short-term interest rate to factors which by construction capture the common response of a large number of individual variables to the economy-wide shocks. This framework thus allows to considerably reduce the dimensionality of the policy problem in a "data-rich" environment (Bernanke and Boivin 2003).

3.2.1 State Dynamics and Short Rate Equation

More formally, assume there is a large number of macroeconomic time series that are each driven by the monetary policy instrument r, a small number of unobserved common factors F and an idiosyncratic component e, i.e.

$$X_t = \Lambda_f F_t + \Lambda_r r_t + e_t, \tag{3.1}$$

where X_t is a $M \times 1$ vector of period-t observations of the variables in the panel, Λ_f and Λ_r are the $M \times k$ and $M \times 1$ matrices of factor loadings, r_t is the short-term interest rate, F_t is the $k \times 1$ vector of period-t observations of the common factors, and e_t is an $M \times 1$ vector of idiosyncratic components.[13] Note that equation (3.1) can also be written in a way that allows X_t to depend on current and lagged values of the fundamental factors. Stock and Watson [2002b] show, however, that the static formulation is not restrictive since F_t can be interpreted as including an arbitrary number of lags of the fundamental factors. Accordingly they refer to the model above - without the observable r_t - as a dynamic factor model.

Economists typically think of the economy as being affected by monetary policy through the short term interest rate, r_t. On the other hand, the central bank is assumed to set interest rates in response to the overall state of the economy, characterized e.g. by the deviations of inflation and output from their desired levels. As has been discussed by Bernanke et al. [2005], theoretical macroeconomic aggregates as inflation and output might not be perfectly observable neither to the policy-maker nor to the econometrician. More realistically, the macroeconomic time series observed by the central bank or the econometrician will in general be noisy measures of broad economic concepts such as output and inflation. Accordingly, these variables should be treated as unobservable in empirical work so as to avoid confounding measurement error or idiosyncratic dynamics with fundamental economic shocks. Bernanke et al. [2005] therefore suggest to extract a few common factors from a large number of macroeconomic time series variables and to study the mutual dynamics of monetary policy and the key economic aggregates by estimating a joint VAR of the factors and the policy instrument, an approach which they label "Factor-Augmented VAR" (FAVAR).

The term structure model suggested here is built upon the assumption that yields are driven by movements of short term interest rates as well as the main shocks hitting the economy. The latter are proxied for by the factors which capture the bulk of common variation in a large number of macroeconomic time series variables. The joint dynamics of these factors and the monetary policy instrument are modelled in a vector autoregression. I thus employ the FAVAR model suggested by Bernanke et al. [2005] as a central building block for my term structure model. In addition, restrictions are imposed on the parameters governing the impact of the state variables on the yields in order to ensure no-arbitrage. Accordingly, I

[13] The idiosyncratic components may display some slight cross- and serial correlation, see Stock and Watson (2002a,b) for a detailed discussion of this issue.

will term the approach pursued here a "No-Arbitrage Factor-Augmented Vector Autoregression". Overall, the dynamics of the economy are described by

$$\begin{pmatrix} F_t \\ r_t \end{pmatrix} = \tilde{\mu} + \tilde{\Phi}(L) \begin{pmatrix} F_{t-1} \\ r_{t-1} \end{pmatrix} + \tilde{\omega}_t, \tag{3.2}$$

where $\tilde{\mu} = (\tilde{\mu}_f', \tilde{\mu}_r)'$ is a $(k+1) \times 1$ vector of constants, $\tilde{\Phi}(L)$ is a $(k+1) \times (k+1)$ matrix of order-p lag polynomials and $\tilde{\omega}_t$ is a $(k+1) \times 1$ vector of reduced form shocks with variance covariance matrix $\tilde{\Omega}$. To summarize, equation (3.2) says that the factors capturing the common variation in many economic time series variables are driven partly by their own dynamics, partly by monetary policy through the short term rate, and partly by exogenous shocks.

Let us have a closer look at the policy reaction function implied by this model. Since the short term interest rate is included in the state vector, the dynamics of the policy instrument are completely characterized by the last equation in the VAR above, i.e.

$$r_t = \tilde{\mu}_r + \tilde{\phi}_f(L)F_{t-1} + \tilde{\phi}_r(L)r_{t-1} + \tilde{\omega}_t^r. \tag{3.3}$$

Hence, in the FAVAR model the short-term interest rate set by the central bank is characterized by a response to the lagged observations of the main economic driving forces, $\tilde{\phi}_f(L)F_{t-1}$, by some interest rate smoothing element $\tilde{\phi}_r(L)r_{t-1}$, and by a monetary policy shock orthogonal to the former two components. The policy reaction function is thus purely backward looking. Yet, since the evolution of r and the main economic driving forces are jointly characterized by a Factor-Augmented VAR model, the implied dynamics of the short term interest rate are potentially much richer than in standard affine term structure models where the short rate is an affine function of contemporaneous observations of the factors whose dynamics are described independently of changes in monetary policy. Hence, the No-Arbitrage FAVAR model studied here explicitly allows for feedback from monetary policy to the macroeconomy, a feature missing e.g. from the model in Ang and Piazzesi [2003] who assume macroeconomic and term structure factors to be orthogonal. The approach pursued in this chapter is thus closer in spirit to the work by Hördahl et al. [2006] who describe the joint evolution of output, inflation, and short-term interest rates within a structural economic model. As in their paper, I expect the richer dynamic structure of the FAVAR model to improve forecast performance.

To facilitate notation in the sequel, I rewrite the VAR in equation (3.2) in companion form as

$$Z_t = \mu + \Phi Z_{t-1} + \omega_t, \tag{3.4}$$

where $Z_t = (F_t', r_t, F_{t-1}', r_{t-1}, \ldots, F_{t-p+1}', r_{t-p+1})'$, and where μ denotes a vector of constants and zeros, Φ the respective companion form matrix of VAR coefficients, and Ω the companion form variance covariance matrix of the reduced form innovations ω. Accordingly, the short rate r_t can be expressed in terms of Z_t as $r_t = \delta' Z_t$ where $\delta' = (0_{1\times k}, 1, 0_{1\times(k+1)(p-1)})$.

In the present model, the vector of state variables Z only comprises the macro driving factors, F, and the short term rate, r. Notice that this assumption could in principle be relaxed by augmenting the state vector with latent yield factors as in Ang and Piazzesi [2003]. In this case, however, the two-step estimation method would no longer be feasible, and one would have to resort to standard maximum-likelihood techniques that are commonly employed in the affine term structure literature. Moreover, the number of parameters to estimate jointly would be considerably higher and thus estimation speed lower. After all, the results below show that no latent factors are needed to obtain a satisfactory in-sample fit of the model.

3.2.2 Pricing Kernel

To model the dynamics of the pricing kernel, I follow the arbitrage-free term structure literature initiated by Duffie and Kan [1996] which has also been applied, among others, by Ang and Piazzesi [2003] and Hördahl et al. [2006]. These authors define the nominal pricing kernel as $m_{t+1} = \exp(-r_t)\frac{\psi_{t+1}}{\psi_t}$, where ψ_t denotes the Radon-Nikodym derivative which converts the risk-neutral into the true data-generating distribution. ψ is assumed to follow the lognormal process $\psi_{t+1} = \psi_t \exp(-\frac{1}{2}\,\lambda_t'\Omega\lambda_t - \lambda_t'\omega_{t+1})$ and is thus driven by the innovations ω of the state variables. Accordingly, the nominal pricing kernel m is given by

$$\begin{aligned} m_{t+1} &= \exp(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t - \lambda_t'\omega_{t+1}), \\ &= \exp(-\delta' Z_t - \frac{1}{2}\lambda_t'\Omega\lambda_t - \lambda_t'\omega_{t+1}). \end{aligned} \tag{3.5}$$

The vector λ_t denotes the market prices of risk. Following Duffee [2002], these are commonly assumed to be affine in the underlying state variables Z, i.e.

$$\lambda_t = \lambda_0 + \lambda_1 Z_t. \tag{3.6}$$

In order to keep the model parsimonious, I restrict the prices of risk to depend only on current observations of the model factors. Obviously, there is some arbitrariness in this restriction. In principle, one can also think of theoretical models that give rise to market prices of risk which depend on lagged state variables. However, since the dimensionality of the problem requires to make some identification restrictions, assuming that market prices of risk depend only on current observations of the states seems to be a plausible compromise.[14] In an arbitrage-free market, the price of a n-months to maturity zero-coupon bond in period t must equal the expected discounted value of the price of an (n-1)-months to maturity bond in period $t+1$:

$$P_t^{(n)} = E_t[m_{t+1} \, P_{t+1}^{(n-1)}].$$

Assuming that yields are affine in the state variables, bond prices $P_t^{(n)}$ are exponential linear functions of the state vector:

$$P_t^{(n)} = \exp\left(A_n + B_n' Z_t\right),$$

where the scalar A_n and the coefficient vector B_n depend on the time to maturity n. Closely following Ang and Piazzesi [2003], I show in appendix 6.2.1 that no-arbitrage is guaranteed by computing coefficients A_n and B_n according to the following recursive equations:

$$A_n = A_{n-1} + B_{n-1}' \left(\mu - \Omega\lambda_0\right) + \frac{1}{2} B_{n-1}' \Omega B_{n-1}, \tag{3.7}$$

$$B_n = B_{n-1}' \left(\Phi - \Omega\lambda_1\right) - \delta'. \tag{3.8}$$

Given the price of an n-months to maturity zero-coupon bond, the corresponding yield is thus obtained as

$$\begin{aligned} y_t^{(n)} &= -\frac{\log P_t^{(n)}}{n} \\ &= a_n + b_n' Z_t, \end{aligned} \tag{3.9}$$

where $a_n = -A_n/n$ and $b_n' = -B_n'/n$.

[14] Note that since the state vector Z_t includes current and lagged observations of the macro factors and the short rate, this choice implies a set of obvious zero restrictions on the parameters λ_0 and λ_1. In particular, $\lambda_0 = (\tilde{\lambda}_0', 0_{1\times(k+1)(p-1)})'$ where $\tilde{\lambda}_0$ is a vector of dimension $(k+1)$ and $\lambda_1 = \begin{pmatrix} \tilde{\lambda}_1 & 0_{(k+1)\times(k+1)(p-1)} \\ 0_{(k+1)(p-1)\times(k+1)} & 0_{(k+1)(p-1)\times(k+1)(p-1)} \end{pmatrix}$ where $\tilde{\lambda}_1$ is a $(k+1)\times(k+1)$ matrix.

3.3 Estimation of the Term Structure Model

Prior to estimating the term-structure model, the common factors have to be extracted from the panel of macro data. This is achieved using standard static principal components following the approach suggested by Stock and Watson (2002a,b). Precisely, let V denote the eigenvectors corresponding to the q largest eigenvalues of the $T \times T$ cross-sectional variance-covariance matrix XX' of the data. Then, subject to the normalization $F'F/T = I_q$, estimates $\hat{F}$ of the factors and $\hat{\Lambda}$ the factor loadings are given by[15]

$$\begin{aligned} \hat{F} &= \sqrt{T}\,V \quad \text{and} \\ \hat{\Lambda} &= \sqrt{T}\,X'V, \end{aligned}$$

i.e. the common factors are estimated as the q largest eigenvalues of the variance-covariance matrix XX'. Given the factor estimates, estimation of the term structure model is performed using a consistent two-step approach following Ang et al. [2006]. First, estimates of the parameters (μ, Φ, Σ) governing the dynamics of the model factors are obtained by running a VAR(p) on the estimated factors and the short term interest rate. Second, given the estimates from the first step, the parameters $\tilde{\lambda}_0$ and $\tilde{\lambda}_1$ which drive the evolution of the state prices of risk, are estimated by minimizing the sum of squared fitting errors of the model. That is, for a given set of parameters the model-implied yields $\hat{y}_t^{(n)} = \hat{a}_n + \hat{b}_n' Z_t$ are computed and then the sum

$$S = \sum_{t=1}^{T} \sum_{n=1}^{N} (\hat{y}_t^{(n)} - y_t^{(n)})^2 \tag{3.10}$$

is minimized with respect to $\tilde{\lambda}_0$ and $\tilde{\lambda}_1$ given the estimates of the VAR parameters μ, Φ, and Ω. Although being possibly less efficient than a joint estimation of all model parameters in a one-step maximum likelihood procedure, the two-step approach has the clear advantage that it is fast and thus much better suited for

[15] To account for the fact that r is an observed factor which is assumed unconditionally orthogonal to the unobserved factors F in the model (3.1), its effect on the variables in X has to be concentrated out prior to estimating F. Here, this is achieved by simply regressing all variables in X onto r and extracting principal components from the variance-covariance matrix of residuals of these regressions. Note that Bernanke et al. [2005] use a slightly a slightly more elaborate approach in order to identify monetary policy shocks within their FAVAR model.

the recursive out-of-sample forecast exercise carried out in this chapter.[16]

Due to the recursive formulation of the bond pricing parameters, the sum of squared fitting errors is highly nonlinear in the underlying model parameters. It is thus helpful to find good starting values to achieve fast convergence. This is done in the following way. I first estimate the parameters $\tilde{\lambda}_0$ under the assumption that risk premia are constant but nonzero, i.e. I set to zero all elements of the matrix $\tilde{\lambda}_1$ which governs the time-varying component of the market prices of risk. Then, I take these estimates as starting values in an estimation step that allows for variation in the market prices of risk, i.e. I let all elements of $\tilde{\lambda}_0$ and $\tilde{\lambda}_1$ be estimated freely. Finally, to enhance tractability of the model, I follow the common practice in the affine term structure literature and re-estimate the model after setting to zero those elements of $\tilde{\lambda}_1$ which are insignificant. Standard errors of the prices of risk parameters, reported in section 3.4 are computed via the numerical gradient of the sum of squared fitting errors function S. The standard errors of the state equation parameters are unadjusted OLS standard errors.[17]

3.4 Empirical Results

3.4.1 Data

I estimate the model using the following data. The macroeconomic factors are extracted from a dataset which contains about 160 monthly time series of various economic categories for the US. Among others, it includes a large number of time series related to industrial production, more than 30 employment-related variables, around 30 price indices and various monetary aggregates. It further

[16] Nonetheless, it would be interesting to estimate the latent macro factors and the parameters characterizing their impact on yields jointly within a one-step estimation procedure. The cross-equation restrictions of the yield curve model would then put additional structure on the estimation of the factors, thereby potentially sharpening up our understanding of the macroeconomic driving forces behind the yield curve. In a recent paper, Law [2004] uses a similar idea to study the extent to which variation in bond yields can be explained by macroeconomic fundamentals.

[17] Notice that Ang et al. [2006] compute standard errors using GMM to adjust for the two-stage estimation process. However, since the No-Arbitrage FAVAR model involves estimation of a VAR of lag order higher than 1, a large number of moment conditions would be needed to identify the state equation parameters via GMM and thus computation would be burdensome. Hence, since the focus here is on forecast performance rather than in-sample fit, I do not follow the approach of Ang et al. [2006].

contains different kinds of survey data, stock indices, exchange rates etc. This dataset has been compiled by Giannone et al. [2004] to forecast US output, inflation, and short term interest rates.[18] Stock and Watson's (2002a,b) principal components estimation of the common factors in large panels of time series requires stationarity. I therefore follow Giannone et al. [2004] in applying different preadjustments to the time series in the dataset.[19] Finally, I standardize all series to have mean zero and unit variance.

I use data on zero-coupon bond yields of maturities 1, 3, 6, and 9 months, as well as 1, 2, 3, 4, 5, 7, and 10 years. All interest rates are continuously-compounded smoothed Fama-Bliss yields and have been constructed from US treasury bonds using the method outlined in Bliss [1997].[20] I estimate and forecast the model over the post-Volcker disinflation period, i.e. from 1983:01 to the last available observation of the macro dataset, 2003:09.

3.4.2 Model Specification

In the first step of the estimation procedure, I extract common factors from the large panel of macroeconomic time series using static principal components following Stock and Watson (2002a,b). Together, the first 10 factors explain about 70% of the total variance of all variables in the dataset. The largest contribution is accounted for by the first four factors, however, which together explain more than 50% of the total variation in the panel. Interestingly, a look at the correlation patterns of all 10 factors with yields of all maturities and their lags, reveals that it is the first four factors that are most highly correlated with yields.

The number of factors I can include in my term-structure model is limited due to parameterization constraints imposed by the market prices of risk specification. If

---

[18] I am grateful to Lucrezia Reichlin who generously provided me with this dataset. Note that I exclude all interest rate related series from the original panel and instead include the zero-coupon yields used in the term structure model. For a detailed description of the data, the reader is referred to the paper by Giannone et al. [2004].

[19] Though with a slight difference as regards the treatment of price series: instead of computing first differences of quarterly growth rates, I follow Ang and Piazzesi [2003] and compute annual inflation rates. The resulting increase in persistence of the estimated factors appears to help explain the persistence of yields.

[20] I am grateful also to Robert Bliss who provided me with the programs and raw data to construct the Fama-Bliss yields.

no additional restrictions are imposed on the market prices of risk, the number of parameters to estimate in the second step of the estimation procedure increases quadratically with the number of factors. For the sake of parsimony I thus restrict the number of factors to the first four principal components extracted from the large panel of monthly time series and the short rate. Unreported results with smaller and larger number of factors have shown that this specification seems to provide the best tradeoff between estimability and model fit. A similar choice has to be made regarding the number of lags to include in the factor-augmented VAR which represents the state equation of my term structure model. Standard information criteria indicate an optimal number of four lags for the joint VAR of factors and the short rate. Therefore, I employ this particular specification of the state equation.

3.4.3 Factor Estimates

As stated before, I extract factors from the large panel of macro time series using the method of Stock and Watson (2002a,b). According to their approach, factors are estimated as $\sqrt{T}$ times the eigenvectors corresponding to the q largest eigenvalues of XX' in descending order. This identifies the common factors against any rotations. Implicitly, one can think of the factors as cross-sectional averages of many time series with weights chosen such that the sum of squared idiosyncratic components in equation (3.1) is minimized. In order to get some understanding of what type of economic information the estimated factors capture, it is instructive to regress them onto the individual variables in the panel. Table 3.7 lists for each of the four factors those five series with which it exhibits the strongest correlation. It turns out from these results that the first factor is closely linked to business cycle variables such as measures of employment and industrial production. In contrast, the second factor is most strongly correlated with different measures of consumer price inflation. Hence, there is a clear dichotomy between a real and a nominal factor as the two main driving forces behind a large number of various economic time series.[21] The third factor loads most strongly on leading indicators of the business cycle such as M1, inventories and loans and securities series. Finally, the fourth factor is most strongly correlated with measures of money supply and producer prices. A plot of the factor time-series

[21] Using the same dataset, Giannone et al. [2004] find that the dynamic dimension of the US economy is two, i.e. they identify a real and a nominal shock which explain the bulk of variation in all time series contained in the panel.

together with some important real and nominal variables is provided in Figure 3.5 in Appendix A.3.

3.4.4 Preliminary Evidence

Before estimating the term structure model subject to no-arbitrage restrictions, I run a set of preliminary regressions to check whether the extracted macro factors are potentially useful explanatory variables in a term structure model. First, I use a simple encompassing test to assess whether a factor-based policy reaction function provides a better explanation of monetary policy decisions than a standard Taylor-rule based on individual measures of output and inflation. I then perform unrestricted regressions of yields on the model factors.

Test of "Excess Policy Response"

The use of dynamic factors instead of individual macroeconomic variables to forecast yields has been justified with the argument that central banks react to larger information sets than individual measures of output and inflation. Whether this conjecture holds true empirically can be tested by comparing the fit of a standard Taylor-rule policy reaction function with that of a policy reaction function based on dynamic factors. Bernanke and Boivin [2003] present evidence for an "excess policy reaction" of the Fed by showing that the fitted value of the federal funds rate from a factor-based policy reaction function is a significant additional regressor in an otherwise standard Taylor-rule equation. Alternatively, one can separately estimate the two competing policy reaction functions and then perform an encompassing test *à* la Davidson and MacKinnon [1993]. This is the strategy adopted by Belviso and Milani [2005]. I follow these authors and compare a standard Taylor rule with partial adjustment,[22]

$$r_t = \rho r_{t-1} + (1-\rho)(\phi_\pi \pi_t + \phi_y y_t),$$

[22] Inflation π is defined as the annual growth rate of the GDP implicit price deflator (GDPDEF). The output gap is measured as the percentage deviation of log GDP (GDPC96) from its trend (computed using the Hodrick-Prescott filter and a smoothing parameter of 14400). Both quarterly series have been obtained from the St. Louis Fed website and interpolated to the monthly frequency using the method described in Mönch and Uhlig (2005, Chapter 5). For the interpolation of GDP, I have used industrial production (INDPRO), total civilian employment (CE16OV) and real disposable income (DSPIC96) as related monthly series. CPI and PPI finished goods have been employed as related monthly series for interpolating the GDP deflator.

with a policy reaction function based on the four factors which I use as state variables in my term structure model,

$$r_t = \rho r_{t-1} + (1-\rho)\phi_F' F_t.$$

The results from both regressions are summarized in Tables 3.8 and 3.9 in Appendix A.3. As indicated by the regression R^2s of 0.967 and 0.970, the factor-based policy rule seems to fit the data slightly better than the standard Taylor rule. The Davidson-MacKinnon (1993) encompassing test can now be used in order to asses whether this improvement in model fit is statistically significant. I thus regress the federal funds rate onto the fitted values from both alternative specifications. This yields the following result:

$$\begin{aligned} r_t &= \alpha\, \hat{r}_t^{Taylor} + (1-\alpha)\, \hat{r}_t^{Factors} + \epsilon_t \\ &= 0.119\, \hat{r}_t^{Taylor} + 0.881\, \hat{r}_t^{Factors} \\ &= (0.173) \qquad\quad (0.173) \end{aligned}$$

Hence, the coefficient on the standard Taylor rule is insignificant whereas the coefficient on the factor-based fitted federal funds rate is highly significant.[23] I interpret this result as evidence supporting the hypothesis that the Fed reacts to a broad macroeconomic information set.

Unrestricted Estimation

To obtain a first impression whether the factors extracted from the panel of macro variables also capture predictive information about yields of higher maturities, Table 3.10 summarizes the mutual correlations between the yields and various lags of the four factors used for estimating the model. As one can see in this table, the short-term interest rate ($y^{(1)}$) shows strongest contemporaneous correlation with yields of any other maturity. Yet, all four macro factors extracted from the panel of monthly US time series, are also strongly correlated with yields of different maturities. The first factor, which closely tracks the business cycle (see also Table 3.7), is positively correlated with yields. The second factor, which clearly captures inflation movements, is also strongly positively correlated with yields of all maturities. The third factor which is most closely related to leading indicators,

[23] Unreported results have shown that this is robust to alternative specifications of both reaction functions using a larger number of lags of the policy instrument and the macro variables or factors.

Table 3.1: **Unrestricted VAR of Yields on Factors**

This table summarizes the results of an unrestricted VAR of yields of different maturities on the four macro factors extracted from the panel of economic time series, and the short term interest rate. The estimation period is 1983:01 to 2003:09. t-values are in brackets.

	$y^{(3)}$	$y^{(6)}$	$y^{(12)}$	$y^{(24)}$	$y^{(36)}$	$y^{(48)}$	$y^{(60)}$	$y^{(84)}$	$y^{(120)}$
cst	1.084	1.697	2.458	3.735	4.683	5.348	5.825	6.452	6.985
	[12.081]	[14.331]	[16.281]	[19.931]	[22.204]	[23.459]	[24.178]	[24.818]	[24.955]
F1	0.253	0.429	0.614	0.792	0.885	0.947	0.992	1.055	1.113
	[13.252]	[17.038]	[19.097]	[19.853]	[19.728]	[19.520]	[19.353]	[19.073]	[18.680]
F2	0.314	0.470	0.626	0.824	0.957	1.052	1.124	1.225	1.319
	[10.966]	[12.444]	[12.974]	[13.762]	[14.210]	[14.455]	[14.610]	[14.758]	[14.759]
F3	0.026	0.045	0.108	0.285	0.435	0.540	0.615	0.710	0.787
	[1.806]	[2.399]	[4.470]	[9.505]	[12.878]	[14.811]	[15.945]	[17.056]	[17.577]
F4	0.091	0.149	0.217	0.309	0.369	0.409	0.438	0.476	0.510
	[5.189]	[6.418]	[7.312]	[8.389]	[8.905]	[9.137]	[9.251]	[9.324]	[9.278]
$y^{(1)}$	0.861	0.795	0.718	0.574	0.459	0.376	0.315	0.235	0.166
	[58.071]	[40.613]	[28.766]	[18.529]	[13.164]	[9.965]	[7.909]	[5.460]	[3.592]
$\bar{R}^2$	0.99	0.98	0.98	0.96	0.95	0.94	0.93	0.92	0.91

is uncorrelated with yields of shorter maturities, but positively correlated with longer maturity yields. Finally, the fourth factor is also positively correlated with yields of all maturities. Correlating lagged factors with yields, one can see that the strong impact of the short rate on yields of all maturities decreases for the benefit of the macro factors. In particular, the correlation between yields and the lagged observations of the business cycle related first and third factor increases with the lag length. This gives a first indication that the macro factors should be good predictors of yields.

To explore further the question whether the models' factors have explanatory power for yields, Table 3.1 provides estimates of an unrestricted VAR of yields of different maturities onto a constant, the four macro factors and the federal funds rate, i.e. it estimates the pricing equation for yields,

$$Y_t = A + BZ_t + u_t,$$

where no cross-equation restrictions are imposed on the coefficients A and B. The first observation to make is that the R^2 of these regressions are all very high. Together with the short rate, the four factors explain more than 95% of the variation in short yields, and still about 90% of the variation in longer yields. Not surprisingly, the federal funds rate is the most highly significant explanatory variable for short maturity yields. However, in the presence of the macro factors its impact

Table 3.2: **In-sample Fit: Observed and Model-Implied Yields and Returns**

This table summarizes empirical means and standard deviations of observed and fitted yields and model-implied 1-year holding period returns. Yields are reported in percentage terms and holding period returns are stated in basis points. The first and second row in each panel report the respective moment of observed yields and fitted values implied by the No-Arbitrage FAVAR model. The third and fourth row in each panel report the respective moment of observed and model-implied 1-year holding period returns.

	$y^{(1)}$	$y^{(3)}$	$y^{(6)}$	$y^{(9)}$	$y^{(12)}$	$y^{(24)}$	$y^{(36)}$	$y^{(48)}$	$y^{(60)}$	$y^{(84)}$	$y^{(120)}$
						Mean					
$y^{(n)}$	5.22	5.44	5.62	5.77	5.90	6.31	6.58	6.76	6.89	7.04	7.17
$\hat{y}^{(n)}$	5.22	5.45	5.61	5.75	5.90	6.33	6.57	6.76	6.90	7.04	7.17
$rx^{(n)}$	-	-	-	-	-	6.91	7.75	8.35	8.85	17.00	11.08
$\hat{rx}^{(n)}$	-	-	-	-	-	6.92	7.67	8.37	8.95	16.84	10.78
						Standard Deviation					
$y^{(n)}$	2.12	2.18	2.26	2.30	2.33	2.31	2.28	2.25	2.24	2.23	2.24
$\hat{y}^{(n)}$	2.12	2.12	2.18	2.24	2.28	2.31	2.26	2.21	2.18	2.16	2.17
$rx^{(n)}$	-	-	-	-	-	2.79	3.93	5.08	6.23	8.73	12.45
$\hat{rx}^{(n)}$	-	-	-	-	-	2.67	3.44	4.11	4.76	6.80	8.62

decreases strongly towards the long end of the maturity spectrum. This shows that the factors extracted from the large panel of macro variables exhibit strong explanatory power for longer yields and thus represent potentially useful states in a term structure model.

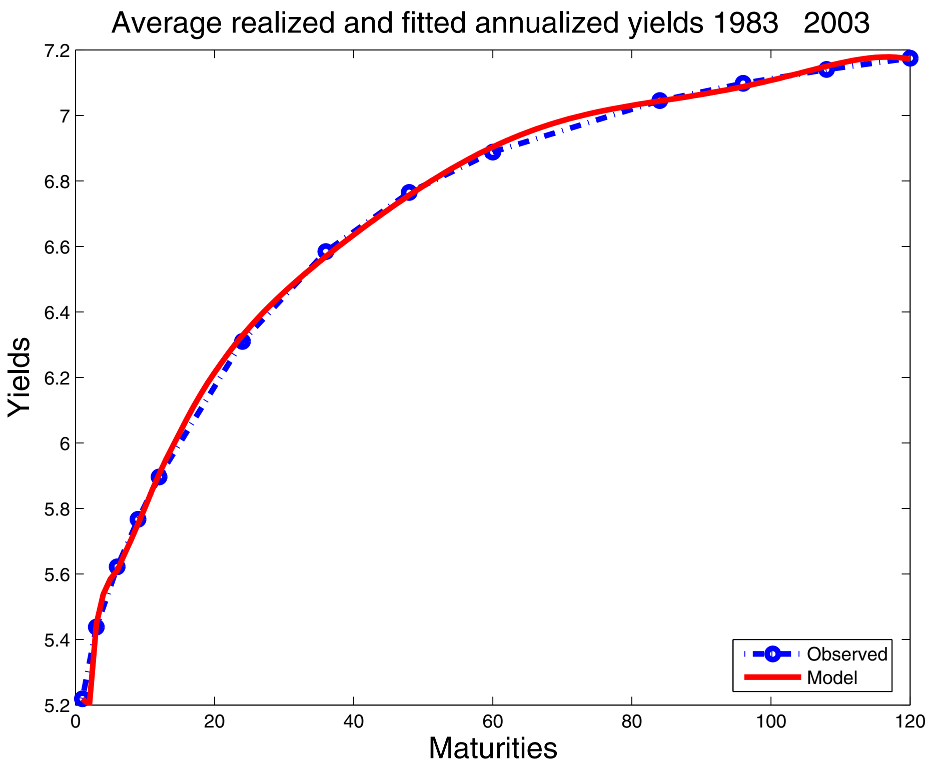
3.4.5 Estimating the Term Structure Model

In-Sample Fit

In this section, I report results obtained from estimating the FAVAR model subject to the cross-equation restrictions (3.7) and (3.8) implied by the no-arbitrage assumption as outlined in Section 3.2. The model fits the data surprisingly well, given that it does not make use of latent yield curve factors. Table 3.2 reports the first and second moments of observed and model-implied yields and one-year holding period returns, respectively. These numbers indicate that on average the No-Arbitrage FAVAR model fits the yield curve almost exactly. Figure 3.1 provides a visualization of this result by showing average observed and model-implied yields across the entire maturity spectrum. Notice that the model seems to be missing some of the variation in longer maturities since the standard deviations of fitted interest rates are slightly lower than the standard deviations of the

Figure 3.1: **Observed and Model Implied Average Yield Curve**

This figure plots average observed yields against those implied by the No-Arbitrage FAVAR model.



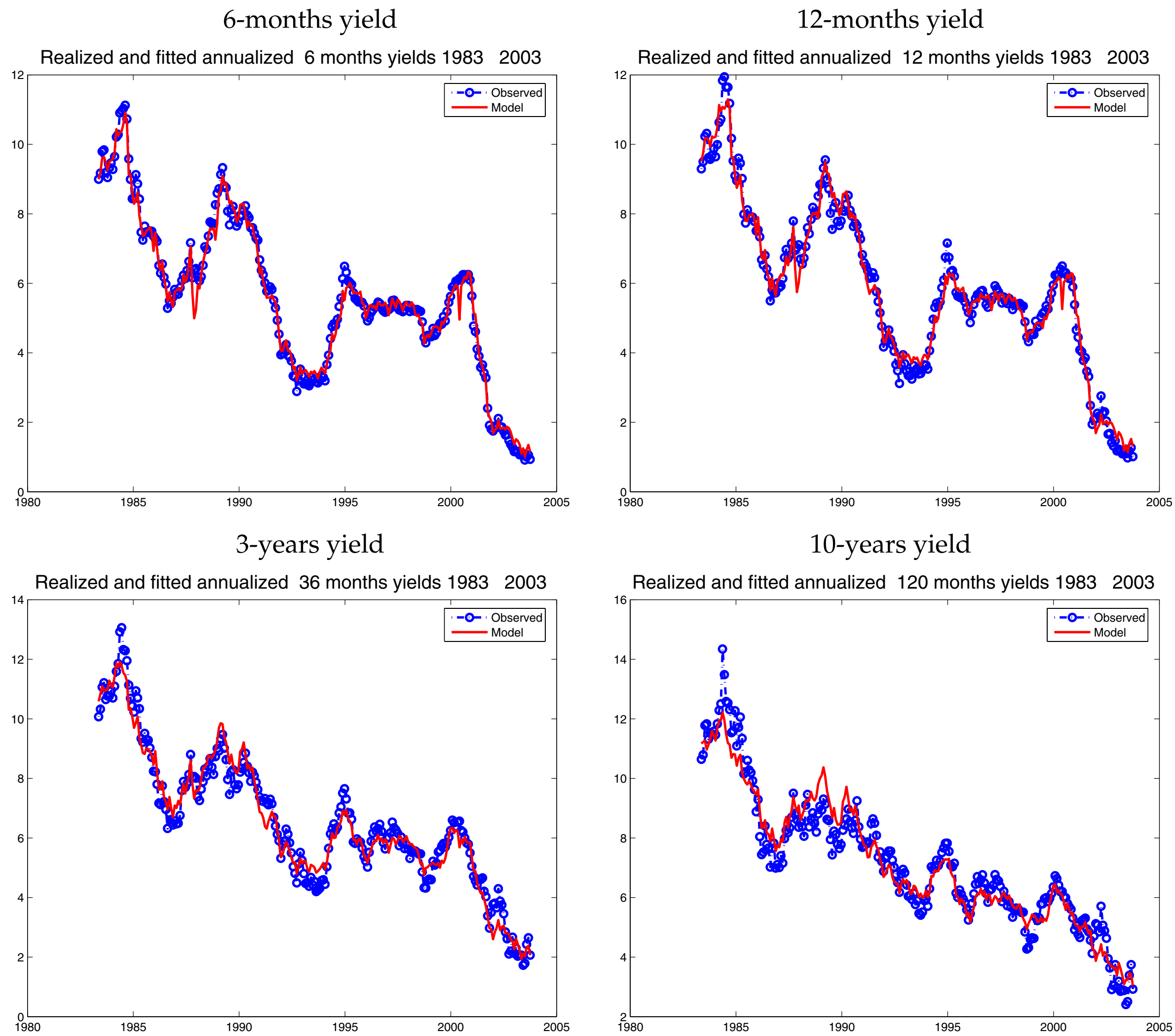
observed yields, especially at the long end of the curve. This can also be seen in Figure 3.2 which plots the time series for a selection of observed and model-implied yields.

While the fit is very good at the short end of the yield curve, the model does not perfectly capture all the variation at the long end of the maturity spectrum. Looking at the lower panel of Table 3.2, one can see that observed and model-implied holding period returns are almost identical on average whereas the fitted returns exhibit standard deviations slightly smaller than the observed returns. Yet, the difference amounts to only a few basis points and is thus fairly small.

Overall, the No-Arbitrage FAVAR model is able to capture the cross-sectional variation of government bond yields quite well, with a slightly better in-sample fit at the short end of the curve. As we will see further below, this has an impact also

Figure 3.2: **Observed and Model-Implied Yields**

This table provides plots of the observed and model-implied time series for four selected interest rates, the 6-months yield, the 12-months yield and the 3-and 10-years yields.



on the forecast results obtained from the model. Indeed, the improvement over latent-factor based term structure models is more pronounced at the short than at the long end of the yield curve. Yet, as has been discussed above, estimating a TSM without latent yield factors considerably facilitates estimation of the model and thus makes recursive out-of-sample forecasts feasible.

Parameter Estimates

Table 3.11 in Appendix A.3 reports the parameter estimates and associated standard errors of the No-Arbitrage FAVAR model. The upper panel shows parameter estimates of the Factor-Augmented VAR that represents the state equation of the model, the second panel provides the estimates of the state prices of risk which constitute the remaining components of the recursive bond pricing parameters A and B.

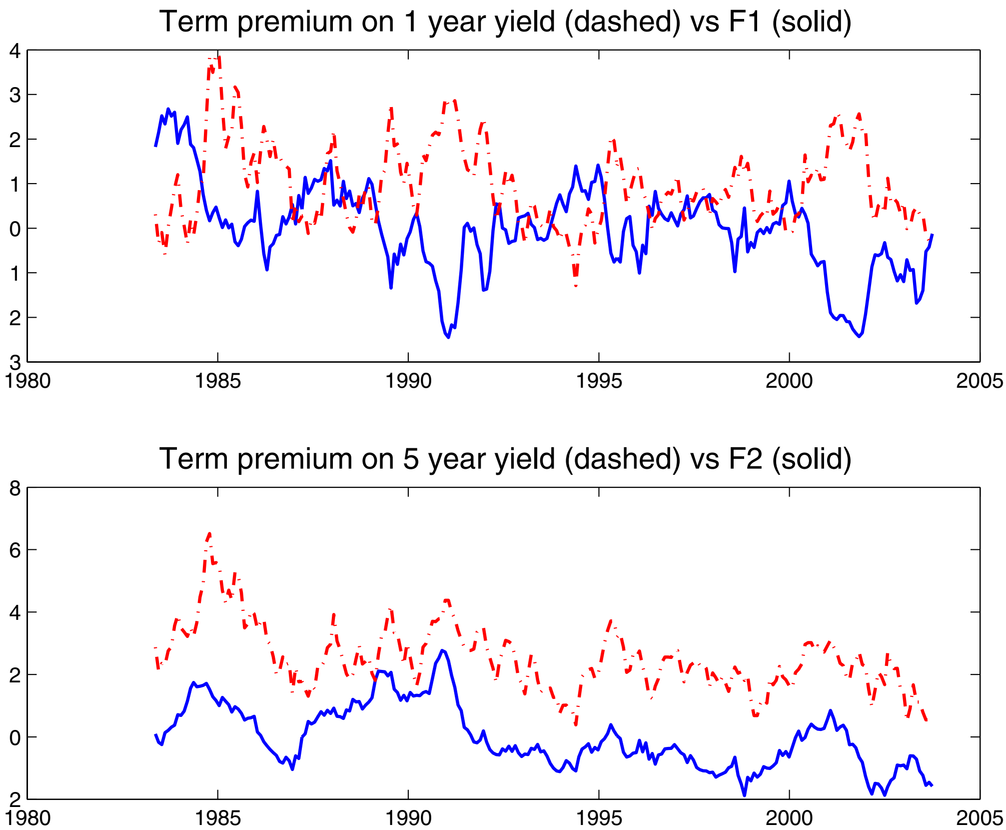
As the diagonal elements of the first lag's coefficient matrix indicate, all five model factors are relatively persistent, a feature that seems to be needed to explain time-variation in yields which are themselves highly persistent time series processes. Since the model factors are by construction unconditionally uncorrelated only few of the off-diagonal elements of the autoregression coefficients in Φ are significant.

As the second panel of Table 3.11 shows, all elements of the vector $\tilde{\lambda}_0$ governing the unconditional mean of the market prices of risk are large and highly significant. This suggests that risk premia are characterized by a large constant component. As indicated by the size and significance of the estimates $\tilde{\lambda}_1$, there is also some significant amount of time variation in risk premia over the sample period considered. It is difficult to interpret individual elements in the estimated prices of risk matrix, however. Indeed, unreported results from alternative model specifications varying e.g. the number of factors, the number of lags in the state equation or the sample period, have shown that the price of risk estimates are quite sensitive to changes in model specification. Hence, economic reasoning based on the significance of individual parameters governing the state prices of risk is unwarranted. Instead, in order to visualize the relation between risk premia and the model factors, Figure 3.3 provides a plot of model-implied term premia for the 1-year and 5-year yield. As indicated by these plots, term premia at the short end of the yield curve are more closely related to the business cycle as proxied by the first macro factor whereas premia for longer yields seem to track inflation which is represented by the second factor.

Figure 3.4 shows a plot of the loadings b_n of the yields onto the contemporaneous observations of the model factors. The signs of these loadings are consistent

Figure 3.3: **Risk Premia Dynamics**

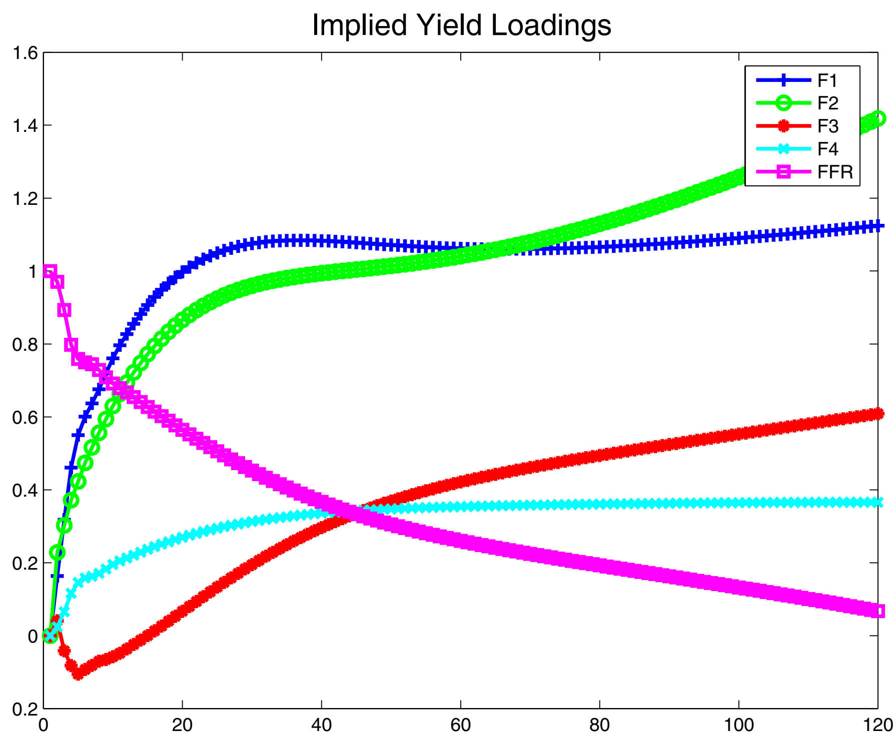
This figure provides a plot of the term premia for 2-year and 5-year yields as implied by the No-Arbitrage FAVAR model. For comparison, they are related to the first and second model factor, respectively.



with those obtained from regressing yields onto the model factors without imposing no-arbitrage restrictions, summarized in Table 3.1. By construction of my arbitrage-free model, the loading of the 1-month yield onto the short rate factor equals unity and those for the macro factors are zero. However, the impact of the short rate on longer yields strongly decreases with maturity and is close to zero at the very long end of the maturity spectrum. Hence, movements in the short-term interest rate only have a marginal direct effect on long-term interest rates. These are almost entirely driven by macroeconomic factors. Most importantly, the inflation-related second factor has a strongly increasing impact on yields going up the maturity spectrum. In contrast, the business cycle related first factor has an equally strong impact on yields of medium and longer maturities. The third factor which is leading the business cycle with a reversed sign

Figure 3.4: **Implied Yield Loadings**

This figure provides a plot of the yield loadings b_n implied by the No-Arbitrage FAVAR model. The coefficients can be interpreted as the response of the n-month yield to a contemporary shock to the respective factor.



has an increasingly positive impact on yields of longer maturities and a negative but small impact on very short maturities. This result is consistent with the well-documented procyclicality of the slope of the yield curve.

3.5 Out-of-Sample Forecasts

In this section, I compare the out-of-sample forecast performance of the No-Arbitrage FAVAR with that of the no-arbitrage VAR model, a VAR(1) on yield levels, the Diebold-Li (2006) version of the Nelson-Siegel (1987) three-factor model, an essentially affine latent factor only model ($A_0(3)$), and a simple random walk. The latter three models are expected to be the most challenging competitors. Diebold and Li [2006] have shown their model to outperform a variety of yield forecasting models including different specifications of forward regressions, AR and VAR models for yields and the random walk. Moreover, Duffee [2002] has shown that the essentially affine latent factor only model has strong out-of-sample forecast performance. Finally, the random walk is often reported to be difficult to beat in out-of-sample forecasts of interest rates.

3.5.1 The Competitor Models

Precisely, the forecasts for the different competitor models are computed as follows.

1. **No-Arbitrage FAVAR model**:

$$\hat{y}^{(n)}_{t+h|t} = \hat{a}_n + \hat{b}_n \hat{Z}^{FAVAR}_{t+h|t}$$

   where Z^{FAVAR} contains the four factors explaining the bulk of variation in the panel of monthly time series for the US, and the 1-month yield. The coefficients $\hat{a}_n$ and $\hat{b}_n$ are obtained recursively according to equations (3.7) and (3.8), using as input the estimates $\hat{\mu}$, $\hat{\Phi}$, and $\hat{\Sigma}$ obtained by running a VAR(1) on the states, as well as the estimates $\hat{\lambda}_0$ and $\hat{\lambda}_1$ obtained by minimizing the sum of squared fitting errors of the model. Forecasts $\hat{Z}^{FAVAR}_{t+h|t}$ are obtained from a VAR(1) fitted to the companion form state vector, i.e.

$$\hat{Z}^{FAVAR}_{t+h|t} = \hat{\Phi}^h Z^{FAVAR}_t + \sum_{i=0}^{h-1} \hat{\Phi}^i \hat{\mu}$$

2. **No-Arbitrage VAR**:

$$\hat{y}^{(n)}_{t+h|t} = \hat{a}_n + \hat{b}_n \hat{Z}^{VAR}_{t+h|t}$$

   where Z^{VAR} contains the quarterly growth rate of IP, the help-wanted index, the annual growth rates of CPI and PPI, and the 1-month yield. The coefficients $\hat{a}_n$ and $\hat{b}_n$ are obtained recursively according to equations (3.7) and (3.8) and guarantee the absence of arbitrage opportunities. The specification and estimation of the model is the same as for the No-Arbitrage FAVAR model.

3. **VAR(1) on Yield Levels**:

$$\hat{y}_{t+h|t} = \hat{c} + \hat{\Gamma} y_t$$

   where $y_t = \{y^{(1)}, y^{(3)}, \ldots, y^{(120)}\}$ and $\hat{c}$ and $\hat{\Gamma}$ are obtained by regressing the vector y_t onto a constant and its h-months lag.

4. **Diebold-Li (2006)**:

$$\hat{y}^{(n)}_{t+h|t} = \hat{\beta}_{1,t+h|t} + \hat{\beta}_{2,t+h|t}\left(\frac{1-e^{-\tau n}}{\tau n}\right) + \hat{\beta}_{3,t+h|t}\left(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)$$

where

$$\hat{\beta}_{t+h|t} = \hat{c} + \hat{\Gamma}\hat{\beta}_t$$

Diebold and Li [2006] obtain estimates of the factors β by fixing τ to 0.0609 and then simply regressing yields onto the factor loadings 1, $(\frac{1-e^{-\tau n}}{\tau n})$, and $(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n})$.[24] Note that Diebold and Li find better forecasting performance of their model when the factor dynamics are estimated by fitting simple AR(1) processes instead of a VAR(1). With the data and sample period used here, however, I find that their model performs better when the latent factor dynamics are estimated using a VAR as specified above.

5. **Essentially Affine Latent Factor Only Model ($A_0(3)$)**:

$$\hat{y}^{(n)}_{t+h|t} = \hat{a}_n + \hat{b}_n \hat{Z}^{A_0(3)}_{t+h|t}$$

where $Z^{A_0(3)}$ is composed of three latent yield factors, backed out from the yields using the method by Chen and Scott [1993]. In particular, it is assumed that the 1-month, 1-year and 10-year yield are observed without error. Otherwise the model setup is the same as for the No-Arbitrage FAVAR model, but only one lag of the state vector enters the state equation. Moreover, the transition matrix Φ in the state equation is assumed to be lower-triangular and the variance-covariance matrix Ω to be an identity matrix so as to ensure exact identification of the model (see Dai and Singleton [2000] for a discussion of the identification issue in affine TSM). Following Duffee [2002], prices of risk are affine in the state variables $Z^{A_0(3)}$ and not assumed to be driven by the factor volatility. Duffee [2002] provides evidence that this "essentially affine" model yields the best out-of-sample forecast results among a set of different affine term structure model specifications. Moreover, Dai and Singleton [2002] show that risk premia are best captured by the essentially affine model. Notice that since estimating the model involves backing out the latent factors from the yields, estimation is tedious and takes considerably longer than estimation of the No-Arbitrage FAVAR

[24] The particular value of τ chosen by Diebold and Li maximizes the curvature loading for a maturity of 30 months. For more details on this choice, the reader is referred to Diebold and Li's paper.

and VAR models where the parameters of the state equation are estimated in a first stage of the estimation via OLS.

6. **Random Walk**:

$$\hat{y}_{t+h|t}^{(n)} = y_t^{(n)}$$

Assuming a random walk model for interest rates implies a simple "no-change" forecast of individual yields. Hence, in this model the h-months ahead prediction of an n-maturity bond yield in period t is simply given by its time t observation.

3.5.2 Forecast Results

The forecasts are carried out over the time period 2000:01-2003:09. The affine models are first estimated over the period 1983:01 - 1999:12 to obtain starting values for the parameters. All models are then estimated recursively using data from 1983:01 to the time that the forecast is made, beginning in 2000:01. Table 3.3 summarizes the root mean squared errors obtained from these forecasts. Three main observations can be made. First, the No-Arbitrage FAVAR model clearly outperforms the no-arbitrage VAR model for all maturities at all forecast horizons. This implies strong support for the use of a broad macroeconomic information set when forecasting the yield curve based on macroeconomic variables only. Second, at the one month horizon, the essentially affine latent factor only model and the random walk outperform the macro-based FAVAR and VAR models for yields of all maturities, with the random walk being slightly superior for medium and longer maturities and the $A_0(3)$ model performing best for short maturities. Third and most importantly, however, the No-Arbitrage FAVAR model exhibits a strikingly strong superiority with respect to all considered benchmark models in yield forecasts six months and twelve months ahead. As the first column of panels B and C of Table 3.3 document, the FAVAR model outperforms the benchmark models across the entire maturity spectrum. This indicates that the combination of a large information set, the rich dynamics of a fourth-order VAR, and the parameter restrictions implied by no-arbitrage delivers a model which is particularly powerful in out-of-sample predictions.

Table 3.4 reports RMSEs of all considered models relative to the random walk forecast. These results show that the improvement in terms of root mean squared

Table 3.3: **Out-of-sample Yield Forecasts : RMSEs**

This table summarizes the root mean squared errors obtained from out-of-sample yield forecasts. The models have been estimated using data from 1983:01 until 1999:12. The forecasting period is 2000:01-2003:09. "FAVAR" refers to the essentially affine term structure model using as states four factors extracted from a large macro panel and the short rate; "VAR" refers to an essentially affine model with IP growth, the index of help-wanted adds in newspapers, CPI growth, PPI growth and the short rate in the state vector. "VAR yields" refers to a VAR(1) on yield levels, "Diebold-Li" denotes the Diebold-Li (2006) version of the three-factor Nelson-Siegel model, "$A_0(3)$" the essentially affine three latent factor only model, and "RW" refers to a simple random walk forecast.

	FAVAR	VAR	VAR yields	Diebold-Li	$A_0(3)$	RW
			Panel A: 1-month ahead forecasts			
$y^{(1)}$	0.759	0.784	0.340	0.363	**0.336**	0.412
$y^{(3)}$	0.650	0.607	0.223	0.298	**0.218**	0.267
$y^{(6)}$	0.654	0.667	0.231	0.353	**0.207**	0.255
$y^{(9)}$	0.619	0.659	0.263	0.410	**0.237**	0.268
$y^{(12)}$	0.624	0.669	0.289	0.436	**0.270**	0.282
$y^{(24)}$	0.612	0.844	0.332	0.394	0.351	**0.313**
$y^{(36)}$	0.587	0.963	0.351	0.367	0.434	**0.331**
$y^{(48)}$	0.596	0.957	0.367	0.371	0.460	**0.347**
$y^{(60)}$	0.609	0.952	0.383	0.385	0.451	**0.361**
$y^{(84)}$	0.564	0.907	0.410	0.419	0.422	**0.384**
$y^{(120)}$	0.532	0.895	0.441	0.464	0.407	**0.407**
			Panel B: 6-month ahead forecasts			
$y^{(1)}$	**0.561**	0.699	1.065	1.213	0.789	1.202
$y^{(3)}$	**0.493**	0.698	1.123	1.240	0.851	1.147
$y^{(6)}$	**0.565**	0.777	1.219	1.316	0.916	1.127
$y^{(9)}$	**0.645**	0.884	1.288	1.369	0.973	1.112
$y^{(12)}$	**0.692**	0.989	1.322	1.383	1.001	1.095
$y^{(24)}$	**0.711**	1.116	1.262	1.262	0.930	1.012
$y^{(36)}$	**0.721**	1.195	1.164	1.144	0.856	0.955
$y^{(48)}$	**0.736**	1.236	1.105	1.091	0.841	0.929
$y^{(60)}$	**0.735**	1.251	1.075	1.073	0.848	0.921
$y^{(84)}$	**0.740**	1.252	1.051	1.073	0.848	0.924
$y^{(120)}$	**0.716**	1.203	1.045	1.088	0.950	0.937
			Panel C: 12-month ahead forecasts			
$y^{(1)}$	**0.995**	1.343	2.116	2.095	1.626	2.093
$y^{(3)}$	**1.056**	1.508	2.301	2.141	1.730	2.122
$y^{(6)}$	**1.185**	1.587	2.476	2.267	1.816	2.140
$y^{(9)}$	**1.321**	1.735	2.561	2.346	1.867	2.120
$y^{(12)}$	**1.345**	1.850	2.572	2.366	1.873	2.069
$y^{(24)}$	**1.226**	1.860	2.321	2.178	1.641	1.787
$y^{(36)}$	**1.181**	1.802	2.054	1.976	1.419	1.584
$y^{(48)}$	**1.139**	1.804	1.887	1.859	1.324	1.478
$y^{(60)}$	**1.086**	1.810	1.788	1.796	1.302	1.425
$y^{(84)}$	**1.120**	1.808	1.688	1.742	1.280	1.386
$y^{(120)}$	**1.098**	1.780	1.627	1.715	1.417	1.376

forecast errors implied by the FAVAR model is particularly pronounced for short and medium term maturities. At the one-month forecast horizon, all yield-based models outperform the affine models based on macro variables. However, at forecast horizons beyond one month, the No-Arbitrage FAVAR model strongly outperforms all other models across the entire spectrum of maturities. Relative to the random walk, the No-Arbitrage FAVAR model reduces root mean squared forecast errors up to 50% at the short end of the yield curve and still improves forecast performance of long yields about 20%. Compared to the best performing competitor model, the essentially affine latent factor model $A_0(3)$, the improvement is still of a remarkable order of 15%.[25] One can formally assess whether the improvement of the FAVAR model over the benchmark models in terms of forecast error is significant by applying White's (2000) "reality check" test. This test can be used to evaluate superior predictive ability of a model with respect to one or more benchmark models. Here, I test whether the No-Arbitrage FAVAR model has superior predictive accuracy with respect to the five considered benchmark models. The test statistics are reported in Table 3.5. Negative figures indicate that the average squared forecast loss of the No-Arbitrage FAVAR model is smaller than that of the respective competitor model while positive test statistics indicate the opposite. White [2000] shows how to derive the empirical distribution of the test statistic by means of a block bootstrap of the forecast error series. I perform 1,000 block-bootstrap resamples from the prediction error series to compute the significance of the forecast improvement.

We have seen above that the FAVAR model outperforms the VAR model at the one-month forecast horizon. As the first column of panel A in Table 3.5 shows, the improvement over the VAR model is significant at almost all maturities. Yet, the FAVAR model is outperformed by all yield-based predictions at the one-month horizon. In sharp contrast, the No-Arbitrage FAVAR model beats all benchmark models at all maturities in forecasts 6-months ahead. As the results in panel B of Table 3.5 show, the improvement in terms of root mean squared forecast errors is significant at the 5% level for all maturities with respect to all benchmark mod-

[25] Note that unreported results from a version of the No-Arbitrage FAVAR model including only one lag in the transition equation have shown a slightly worse performance. In particular, this model specification has been outperformed by the random walk and the latent factor affine model at the very long end of the yield curve. Hence, allowing for a relatively rich specification of the joint dynamics of macro factors and the short rate appears to considerably enhance forecast accuracy.

Table 3.4: **RMSEs Relative to Random Walk**

This table summarizes the root mean squared errors of out-of-sample yield forecasts relative to the simple random walk forecasts. The models have been estimated using data from 1983:01 until 1999:12. The forecasting period is 2000:01-2003:09. "FAVAR" refers to the essentially affine term structure model using as states four factors extracted from a large macro panel and the short rate; "VAR" refers to an essentially affine model with IP growth, the index of help-wanted adds in newspapers, CPI growth, PPI growth and the short rate in the state vector. "VAR yields" refers to a VAR(1) on yield levels, "Diebold-Li" denotes the Diebold-Li (2006) version of the three-factor Nelson-Siegel model and "$A_0(3)$" the essentially affine three latent factor only model.

	FAVAR	VAR	VAR yields	Diebold-Li	$A_0(3)$
			Panel A: 1-month ahead forecasts		
$y^{(1)}$	1.842	1.904	0.827	0.881	**0.816**
$y^{(3)}$	2.437	2.277	0.837	1.117	**0.818**
$y^{(6)}$	2.559	2.611	0.904	1.381	**0.811**
$y^{(9)}$	2.307	2.456	0.978	1.526	**0.883**
$y^{(12)}$	2.210	2.369	1.024	1.544	**0.957**
$y^{(24)}$	1.959	2.698	**1.063**	1.260	1.123
$y^{(36)}$	1.773	2.910	**1.061**	1.108	1.310
$y^{(48)}$	1.717	2.758	**1.057**	1.069	1.326
$y^{(60)}$	1.685	2.637	**1.059**	1.064	1.247
$y^{(84)}$	1.467	2.361	**1.067**	1.090	1.097
$y^{(120)}$	1.309	2.202	1.084	1.142	**1.000**
			Panel B: 6-month ahead forecasts		
$y^{(1)}$	**0.467**	0.582	0.886	1.009	0.656
$y^{(3)}$	**0.430**	0.608	0.979	1.082	0.742
$y^{(6)}$	**0.501**	0.689	1.081	1.168	0.812
$y^{(9)}$	**0.579**	0.795	1.158	1.230	0.874
$y^{(12)}$	**0.632**	0.904	1.208	1.264	0.914
$y^{(24)}$	**0.702**	1.103	1.248	1.247	0.919
$y^{(36)}$	**0.755**	1.252	1.219	1.199	0.897
$y^{(48)}$	**0.792**	1.330	1.189	1.174	0.905
$y^{(60)}$	**0.798**	1.359	1.167	1.165	0.920
$y^{(84)}$	**0.801**	1.354	1.137	1.161	0.918
$y^{(120)}$	**0.764**	1.284	1.115	1.161	1.014
			Panel C: 12-month ahead forecasts		
$y^{(1)}$	**0.476**	0.642	1.011	1.001	0.777
$y^{(3)}$	**0.498**	0.711	1.085	1.009	0.816
$y^{(6)}$	**0.554**	0.742	1.157	1.059	0.848
$y^{(9)}$	**0.623**	0.818	1.208	1.107	0.881
$y^{(12)}$	**0.650**	0.894	1.243	1.143	0.905
$y^{(24)}$	**0.686**	1.041	1.299	1.219	0.918
$y^{(36)}$	**0.746**	1.138	1.297	1.247	0.896
$y^{(48)}$	**0.771**	1.221	1.277	1.258	0.896
$y^{(60)}$	**0.762**	1.270	1.255	1.261	0.914
$y^{(84)}$	**0.808**	1.305	1.218	1.257	0.923
$y^{(120)}$	**0.799**	1.294	1.183	1.247	1.030

Table 3.5: **White's Reality Check Test**

This table summarizes "White's Reality Check" test statistics based on a squared forecast error loss function. I choose the no-arbitrage FAVAR model as the benchmark model and compare it bilaterally with the competitor models. Negative test statistics indicate that the average squared forecast loss of the FAVAR model is smaller than that of the respective competitor model. Bold figures indicate significance at the 5% interval. Significance is checked by comparing the average forecast loss differential with the 5% percentile of the empirical distribution of the loss differential series approximated by applying a block bootstrap with 1,000 resamples and a smoothing parameter of 1/12. Bold figures highlight significance at the 5% level.

	VAR	VARylds	DL	$A_0(3)$	RW
			Panel A: 1-month ahead forecasts		
$y^{(1)}$	**-0.218**	3.064	2.967	3.088	2.708
$y^{(3)}$	0.401	2.537	2.276	2.554	2.392
$y^{(6)}$	-0.091	2.543	2.063	2.611	2.462
$y^{(9)}$	**-0.292**	2.177	1.515	2.259	2.154
$y^{(12)}$	**-0.355**	2.108	1.397	2.175	2.130
$y^{(24)}$	**-2.244**	1.812	1.515	1.728	1.892
$y^{(36)}$	**-3.862**	1.553	1.478	1.134	1.640
$y^{(48)}$	**-3.710**	1.554	1.534	1.057	1.645
$y^{(60)}$	**-3.563**	1.566	1.554	1.204	1.667
$y^{(84)}$	**-3.340**	1.091	1.038	1.028	1.224
$y^{(120)}$	**-3.417**	0.671	0.520	0.856	0.859
			Panel B: 6-month ahead forecasts		
$y^{(1)}$	**-1.064**	**-4.996**	**-7.109**	**-1.881**	**-6.938**
$y^{(3)}$	**-1.491**	**-6.198**	**-7.971**	**-2.944**	**-6.570**
$y^{(6)}$	**-1.747**	**-7.102**	**-8.705**	**-3.184**	**-5.825**
$y^{(9)}$	**-2.245**	**-7.578**	**-8.969**	**-3.253**	**-5.022**
$y^{(12)}$	**-3.070**	**-7.725**	**-8.818**	**-3.209**	**-4.395**
$y^{(24)}$	**-4.574**	**-6.621**	**-6.639**	**-2.194**	**-3.156**
$y^{(36)}$	**-5.606**	**-5.066**	**-4.798**	**-1.285**	**-2.369**
$y^{(48)}$	**-6.079**	**-4.121**	**-3.930**	**-0.984**	**-1.939**
$y^{(60)}$	**-6.324**	**-3.729**	**-3.701**	**-1.071**	**-1.863**
$y^{(84)}$	**-6.279**	**-3.386**	**-3.656**	**-1.031**	**-1.858**
$y^{(120)}$	**-5.748**	**-3.537**	**-4.081**	**-2.371**	**-2.229**
			Panel C: 12-month ahead forecasts		
$y^{(1)}$	**-4.491**	**-19.400**	**-19.110**	**-9.214**	**-18.973**
$y^{(3)}$	**-6.448**	**-23.301**	**-19.471**	**-10.487**	**-18.904**
$y^{(6)}$	**-6.236**	**-26.380**	**-20.876**	**-10.553**	**-17.702**
$y^{(9)}$	**-7.098**	**-26.890**	**-20.971**	**-9.717**	**-15.335**
$y^{(12)}$	**-9.065**	**-26.910**	**-21.130**	**-9.487**	**-13.841**
$y^{(24)}$	**-11.105**	**-21.837**	**-18.076**	**-6.705**	**-9.563**
$y^{(36)}$	**-10.617**	**-15.928**	**-13.999**	**-3.519**	**-6.357**
$y^{(48)}$	**-11.224**	**-12.776**	**-12.047**	**-2.614**	**-5.073**
$y^{(60)}$	**-12.038**	**-11.404**	**-11.446**	**-2.961**	**-4.884**
$y^{(84)}$	**-11.619**	**-9.019**	**-9.949**	**-2.191**	**-3.840**
$y^{(120)}$	**-11.323**	**-8.159**	**-9.711**	**-4.536**	**-3.954**

els. This underlines the observation made above that the model predicts yields considerably better than the best performing competitor, the essentially affine latent factor model, $A_0(3)$. The same pattern is found for the 12-months ahead predictions for which the forecast loss of the No-Arbitrage FAVAR model is also significantly smaller than those of the considered benchmarks at all maturities. Altogether, the evidence shows that the No-Arbitrage FAVAR model is particularly useful in predicting yields at forecast horizons beyond one month, the improvement over benchmark models being particularly strong at the short end of the curve. Thus, augmenting a Factor-Augmented VAR model with tight parameter restrictions implied by no-arbitrage may lead to significantly improved yield forecasts. The fact that the No-Arbitrage FAVAR model outperforms a model based on a VAR of four individual macro variables plus the short rate which is otherwise identically specified, further underscores the usefulness of incorporating a broad macroeconomic information set into term structure analysis.

To summarize, the No-Arbitrage FAVAR model exhibits strong relative advantages over a variety of benchmark models which have been documented powerful tools in forecasting the yield curve. The improvement is particularly pronounced for short and medium term maturity yields. Notice that I have not compared the model to alternative affine term structure models which incorporate macro factors such as the models by Ang and Piazzesi [2003] or Hördahl et al. [2006]. Simultaneously including macro and latent yield curve factors, these models are considerably more cumbersome to estimate than the model presented in this chapter and thus a comparison based on recursive out-of-sample forecasts is infeasible. As has already been discussed above, the No-Arbitrage FAVAR model has the advantage that the state equation parameters are obtained in a separate step of the estimation procedure, a feature that considerably enhances estimation speed and thus also might make the approach more suitable for application in practice.

3.5.3 How are the Macro Factors Related to Latent Yield Factors?

In order to better understand the source of the strong forecast performance of the No-Arbitrage FAVAR model, it is interesting to relate the macro factors to the traditional latent decomposition of yields into level, slope, and curvature. In this section, I thus regress estimates of latent factors onto the macro factors and the

Table 3.6: **Regression of Latent Yield Factors on the Model Factors**

This table summarizes the results obtained from a regression of level, slope, and curvature yield factors onto the factors of the FAVAR model. Level, slope, and curvature are computed as the first three principal components extracted from the yields used to estimate the term structure model. They explain 90.8%, 6.4% and 1.6% of the total variance of all yields, respectively. The sample period is 1984:01-2003:9. t-statistics are in brackets.

	Level	Slope	Curvature
cst	0,040	0,244	-0,145
	[22.769]	[18.712]	[-7.103]
F1	0,007	0,032	-0,058
	[20.133]	[11.828]	[-13.783]
F2	0,008	0,038	-0,049
	[14.400]	[9.215]	[-7.491]
F3	0,004	0,037	0,009
	[13.880]	[16.737]	[2.431]
F4	0,003	0,016	-0,017
	[8.832]	[6.309]	[-4.245]
$y^{(1)}$	0,005	-0,041	0,024
	[15.617]	[-18.643]	[7.011]
$\bar{R}^2$	0,959	0,786	0,481

short rate. The latent yield factors are computed as the first three principal components of the yields used to estimate the term structure model. Similar to results from previous studies, the first three principal components explain about 90.8%, 6.4% and 1.6% of the total variance of the panel. Following the conventional notation, I label them "level", "slope", and "curvature". The first three columns of Table 3.6 summarize the results of these regressions. The four macro factors and the short-term interest rate explain almost all of the variation in the yield level. The main contribution comes from the short rate, the business cycle related first factor and the inflation-related second factor, but the remaining macro factors are also significant explanatory variables for the yield level. Almost 80% of the variation in the slope of the yield curve is explained by the macro factors. Both the business cycle related first and the inflation-related second factor are positively linked with the slope of the yield curve. This is consistent with the fact that short-term interest rates are expected to rise relative to long-term interest rates in an inflationary environment. Moreover, the short rate has a strongly significant negative coefficient in the slope equation which is consistent with the intuition that rises in the short rate lead to a decreasing yield curve slope. Finally note that only about 48% of the variation in the curvature of the yield curve are explained by the macro factors. Hence, variations in the relative size of short, medium and long-term yields seem to be the least related to macroeconomic news.

3.6 Conclusion

This chapter presents a model of the term structure of interest rates which is entirely built upon observable macroeconomic information. Instead of relying on a latent factor-based decomposition of interest rates, yields are modelled as affine functions of the short rate and a few factors which capture the bulk of variation in a large number of macroeconomic time series variables. This particular modelling approach which I label a "No-Arbitrage Factor-Augmented Vector Autoregression" is motivated by recent evidence which suggests that factors extracted from large macro panels are powerful predictors of short-term interest rates and measures of output and inflation. Moreover, since monetary policy decisions are likely based on the developments in a variety of economic time series, it is straightforward to model interest rates as a function of the factors which by construction summarize the main sources of economic fluctuation.

The model is estimated in two steps. First, the factors are extracted from a large panel of macroeconomic time series using the principal components-based approach suggested by Stock and Watson (2002a,b) and the parameters governing their joint dynamics with the short-term interest rate are estimated in a VAR. In a second step, the price of risk parameters of the affine term structure model specification are obtained by minimizing the sum of squared fitting errors of the model. This consistent two-step approach makes estimation fast and allows to carry out a recursive out-of-sample forecasting exercise.

Preliminary regressions show that the factors of the model contain information for explaining the monetary policy instrument which is not captured by individual measures of output and inflation. Moreover, unrestricted regressions of yields on the model factors show that common components extracted from the large panel of macroeconomic time series are highly significant explanatory variables for yields. Accordingly, an affine term structure model built upon these factors and the short rate provides a good in-sample fit of the term structure of interest rates. Compared to a model which incorporates the short rate and four individual measures of output and inflation as factors, there is an advantage in using the larger macroeconomic information set. The results from out-of-sample forecasts of yields underpin this finding. The term structure model based on common factors clearly outperforms the model based on individual variables for all maturities at all horizons. Moreover, in forecasts beyond one month ahead the model strongly outperforms a set of yield-based forecast models including the model recently suggested by Diebold and Li [2006], a standard three latent factor essentially affine model, and the random walk. At forecast horizons of six and twelve months ahead, the reduction in terms of root mean squared forecast errors relative to the random walk amounts up to 50% for short yields and still is about 20% for very long yields. The improvement in forecast accuracy is shown to be statistically significant for all maturities.

A number of potential extensions to the work carried out in this chapter are conceivable. First, since financial markets are assumed to respond quickly to macroeconomic news, the forecast exercise could be done using real-time data. Unfortunately, however, real-time macroeconomic datasets of the size necessary for the use of large-scale factor models are still scarce. Second, to improve on the interpretability of the model, a more structural factor model approach could be applied. Instead of extracting factors from a large cross-section of macroeconomic time series, Belviso and Milani [2005] have recently suggested to extract factors from groups of variables of the same economic category and to use this structural factor-augmented FAVAR model to assess the effect of monetary policy. In such a framework, particular emphasis could be given to factors summarizing agents' expectations of inflation and output developments which have been documented important determinants of long-term yields (see e.g. Dewachter and Lyrio 2006). Finally, the model setup employed in this chapter can in principle also be used as a tool to disentangle the effects of specific economic shocks on risk premia and on the risk-adjusted future path of expected short-term rates.

A.3 Additional Tables and Figures

Table 3.7: **Factor Loadings**

This table summarizes R-squares of univariate regressions of the factors extracted from the panel of macro variables on all individual variables. For each factor, I list the five variables that are most highly correlated with it. Notice that the series have been transformed to be stationary prior to extraction of the factors, i.e. for most variables the regressions correspond to regressions on growth rates. The four factors together explain more than 50% of the total variation in the large panel of macroeconomic time series.

Factor 1 - 24.9 % of total variance	R^2
Employment on nonag payrolls: Manufacturing	0.79
Employment on nonag payrolls: Goods-producing	0.77
Capacity Utilization: Total (NAICS)	0.76
Index of IP: Non-energy excl CCS and MVP (NAICS)	0.76
Index of IP: Total	0.76
Factor 2 - 13.3 % of total variance	
CPI: all items (urban)	0.79
CPI: all items less medical care	0.76
CPI: all items less food	0.74
CPI: all items less shelter	0.69
PCE chain weight price index: Total	0.69
Factor 3 - 7.6 % of total variance	
M1 (in mil of current $)	0.49
CPI: medical care	0.47
Inventories: Mfg and Trade: Mfg, durables (mil of chained 96$)	0.41
Loans and Securities @ all comm banks: Securities, U.S. govt (in mil of $)	0.36
Inventories: Mfg and Trade: Mfg (mil of chained 96$)	0.36
Factor 4 - 5.4 % of total variance	
Employment on nonag payrolls: Financial activities	0.33
PPI: finished goods excl food	0.27
PPI: finished consumer goods	0.24
CPI: transportation	0.23
M3 (in mil of current $)	0.23

Table 3.8: **Policy Rule Based on Individual Variables**

This table reports estimates for a policy rule with partial adjustment based on individual measures of output and inflation, i.e.

$$r_t = c + \rho r_{t-1} + (1-\rho)(\phi_y y_t + \phi_\pi \pi_t),$$

where r denotes the federal funds rate, y the deviation of log GDP from its trend, and π the annual rate of GDP inflation. The sample period is 1983:01 to 2003:09. Standard errors are in parentheses. The R^2 of this regression is 0.967.

c	ρ	ϕ_y	ϕ_π
-0.011	0.955	1.332	2.592
(0.078)	(0.017)	(0.627)	(0.850)

Table 3.9: **Policy Rule Based on Factors**

This table reports estimates for a policy rule with partial adjustment based on the four factors extracted from a large panel of macroeconomic variables, i.e.

$$r_t = c + \rho r_{t-1} + (1-\rho)(\phi_{F1} F1_t + \phi_{F2} F2_t + \phi_{F3} F3_t + \phi_{F4} F4_t),$$

where r again denotes the federal funds rate and $F1$ to $F4$ the four macro factors extracted from a panel of about 160 monthly time series for the US. The sample period is 1983:01 to 2003:09. Standard errors are in parentheses. The R^2 of this regression is 0.97.

c	ρ	ϕ_{F1}	ϕ_{F2}	ϕ_{F3}	ϕ_{F4}
0.564	0.902	0.174	0.160	-0.004	0.050
(0.152)	(0.025)	(0.031)	(0.049)	(0.025)	(0.030)

Table 3.10: **Correlation of Macro Factors and Yields**

This table summarizes the mutual correlation patterns between the yields and factors used for estimating the term structure model. $F1, F2, F3$ and $F4$ denote the macro factors extracted form the large panel of monthly economic time series for the US, $y^{(1)}$ to $y^{(120)}$ denote the yields of maturities 1-month to 10-years, respectively.

	$y^{(1)}$	$y^{(6)}$	$y^{(12)}$	$y^{(36)}$	$y^{(60)}$	$y^{(120)}$
Correlation of observable factors and yields						
F1	0.392	0.478	0.514	0.545	0.546	0.541
F2	0.723	0.725	0.712	0.688	0.671	0.649
F3	0.025	0.014	0.031	0.151	0.223	0.289
F4	0.296	0.272	0.266	0.254	0.241	0.223
Correlation of 1-month lagged observable factors and yields						
F1(-1)	0.441	0.520	0.550	0.567	0.562	0.551
F2(-1)	0.706	0.701	0.688	0.668	0.654	0.634
F3(-1)	0.004	0.001	0.020	0.145	0.220	0.288
F4(-1)	0.272	0.250	0.248	0.242	0.231	0.215
Correlation of 3-months lagged observable factors and yields						
F1(-3)	0.515	0.577	0.596	0.589	0.573	0.552
F2(-3)	0.661	0.651	0.638	0.629	0.623	0.611
F3(-3)	-0.024	-0.015	0.008	0.139	0.216	0.283
F4(-3)	0.244	0.228	0.228	0.227	0.216	0.198
Correlation of 6-months lagged observable factors and yields						
F1(-6)	0.576	0.627	0.638	0.616	0.589	0.556
F2(-6)	0.591	0.567	0.555	0.566	0.575	0.580
F3(-6)	-0.057	-0.035	-0.008	0.125	0.201	0.267
F4(-6)	0.221	0.218	0.217	0.209	0.195	0.175
Correlation of 9-months lagged observable factors and yields						
F1(-9)	0.638	0.675	0.679	0.641	0.606	0.568
F2(-9)	0.514	0.473	0.460	0.493	0.517	0.536
F3(-9)	-0.066	-0.019	0.014	0.140	0.209	0.271
F4(-9)	0.177	0.181	0.182	0.175	0.157	0.127
Correlation of 12-months lagged observable factors and yields						
F1(-12)	0.656	0.676	0.671	0.621	0.583	0.540
F2(-12)	0.431	0.384	0.375	0.436	0.475	0.502
F3(-12)	-0.073	-0.009	0.024	0.129	0.192	0.255
F4(-12)	0.169	0.178	0.189	0.191	0.173	0.146

Table 3.11: **Parameter Estimates for No-Arbitrage FAVAR Model**

State dynamics : $Z_t = \tilde{\mu} + \Phi_1 Z_{t-1} + \ldots \Phi_4 Z_{t-4} + \tilde{\omega}_t, \quad E[\tilde{\omega}_t \tilde{\omega}_t'] = \tilde{\Omega}$

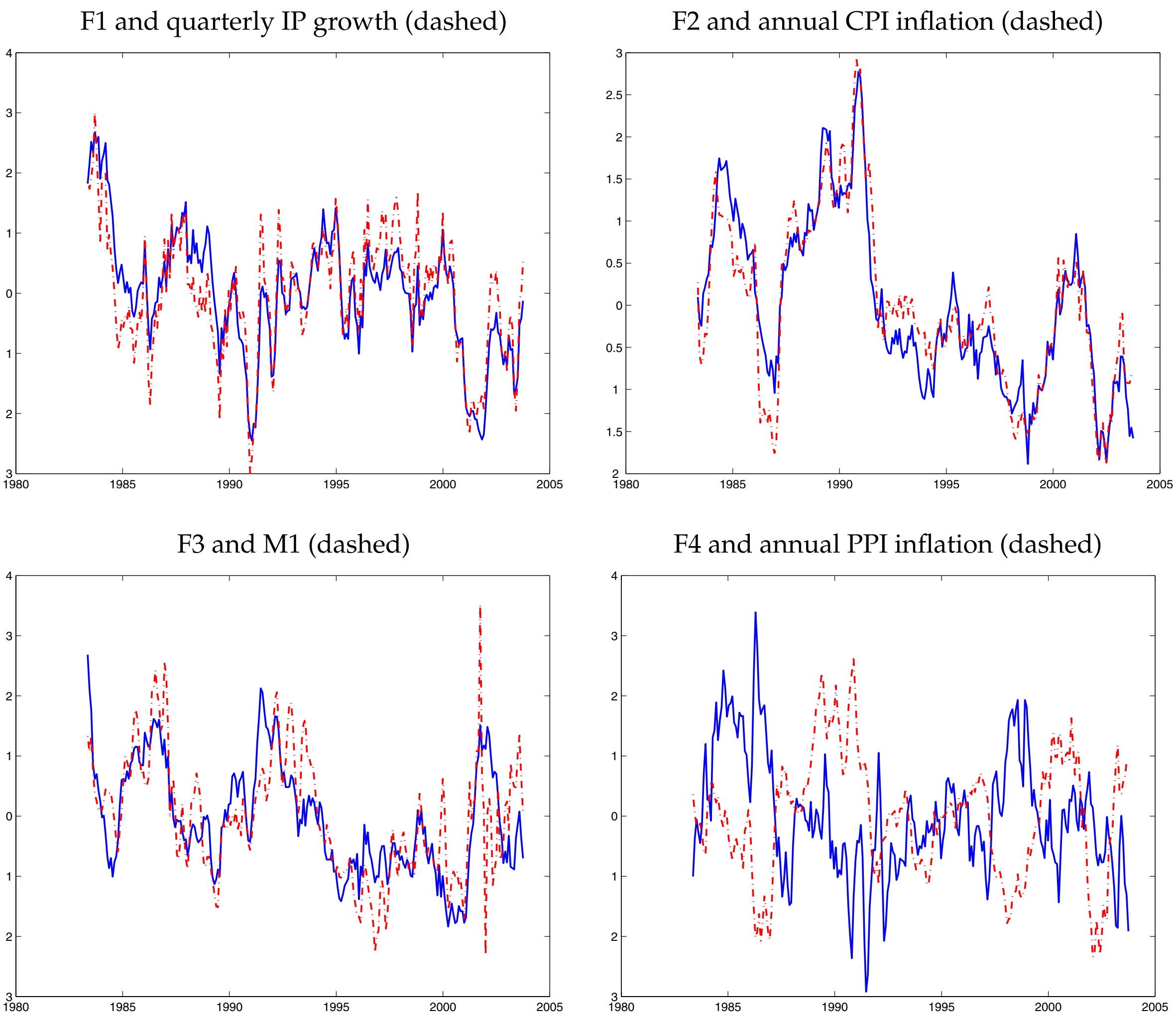
	Φ_1					Φ_2				
F1	1.149	0.211	0.025	0.039	-0.007	0.132	-0.271	0.034	0.148	0.078
	(0.108)	(0.153)	(0.114)	(0.062)	(0.053)	(0.165)	(0.237)	(0.145)	(0.083)	(0.072)
F2	0.179	1.200	0.007	-0.057	0.006	-0.235	-0.238	-0.053	0.025	0.023
	(0.070)	(0.099)	(0.074)	(0.040)	(0.035)	(0.107)	(0.154)	(0.095)	(0.054)	(0.047)
F3	-0.213	-0.056	0.900	0.023	-0.054	0.040	-0.098	0.158	0.017	-0.023
	(0.079)	(0.113)	(0.084)	(0.045)	(0.039)	(0.122)	(0.174)	(0.107)	(0.061)	(0.053)
F4	-0.384	-0.792	-0.142	0.893	0.041	0.058	0.650	0.057	-0.268	-0.139
	(0.138)	(0.197)	(0.146)	(0.079)	(0.069)	(0.212)	(0.304)	(0.187)	(0.107)	(0.093)
$y^{(1)}$	0.341	0.451	0.075	0.045	0.929	-0.094	-0.581	-0.361	0.057	-0.120
	(0.125)	(0.177)	(0.132)	(0.071)	(0.062)	(0.192)	(0.274)	(0.169)	(0.096)	(0.084)
	Φ_3					Φ_4				
F1	-0.621	0.113	-0.055	-0.119	0.035	0.251	-0.046	0.062	-0.018	-0.120
	(0.163)	(0.235)	(0.146)	(0.084)	(0.072)	(0.122)	(0.158)	(0.103)	(0.059)	(0.052)
F2	0.142	-0.018	0.128	-0.033	-0.047	-0.016	-0.000	-0.102	0.027	0.037
	(0.106)	(0.153)	(0.095)	(0.054)	(0.047)	(0.079)	(0.103)	(0.067)	(0.038)	(0.034)
F3	0.217	0.235	-0.432	0.053	0.066	-0.120	0.034	0.299	-0.014	-0.018
	(0.120)	(0.173)	(0.108)	(0.062)	(0.053)	(0.090)	(0.116)	(0.076)	(0.044)	(0.039)
F4	0.283	-0.309	0.139	-0.129	-0.020	0.206	0.367	-0.067	0.329	0.153
	(0.210)	(0.302)	(0.187)	(0.108)	(0.093)	(0.156)	(0.203)	(0.132)	(0.076)	(0.067)
$y^{(1)}$	0.038	0.368	0.246	-0.007	-0.130	-0.117	-0.097	-0.024	-0.049	0.233
	(0.189)	(0.272)	(0.169)	(0.097)	(0.084)	(0.141)	(0.183)	(0.119)	(0.069)	(0.061)
	$\tilde{\Omega}$					$\tilde{\mu}$				
F1	0.086					0.084				
	(0.008)					(0.128)				
F2	-0.036	0.036				-0.104				
	(0.004)	(0.003)				(0.083)				
F3	0.036	-0.027	0.047			0.132				
	(0.005)	(0.003)	(0.004)			(0.094)				
F4	-0.062	0.009	-0.013	0.142		-0.216				
	(0.008)	(0.005)	(0.005)	(0.013)		(0.164)				
$y^{(1)}$	0.005	-0.000	-0.002	-0.003	0.116	0.428				
	(0.006)	(0.004)	(0.005)	(0.008)	(0.011)	(0.148)				

Market prices of risk : $\lambda_t = \lambda_0 + \lambda_1 Z_t$

$\tilde{\lambda}_0$	$\tilde{\lambda}_1$				
-29.535	1.536	-1.241	-1.701	-	-3.701
(0.038)	(0.724)	(0.172)	(0.624)	-	(1.550)
-290.060	-1.420	-4.239	-1.202	-0.347	-1.076
(0.034)	(0.266)	(0.044)	(0.113)	(0.076)	(0.061)
-141.987	-2.407	-	1.217	-	3.649
(0.018)	(1.078)	-	(0.266)	-	(0.964)
-52.033	-	-	-1.821	1.090	-5.523
(0.013)	-	-	(0.146)	(0.712)	(0.010)
-3.113	-	-	-	-	-
(0.081)	-	-	-	-	-

Figure 3.5: **Plot of Model Factors**

This figure provides plots of the factors used in the No-Arbitrage FAVAR model. Each factor is confronted with an individual macroeconomic variable in order to show the close correspondence to the real and the nominal side of the economy.



4 Term Structure Surprises: The Predictive Content of Curvature, Level, and Slope

This chapter analyzes the predictive content of the term structure components level, slope, and curvature within a dynamic factor model of macroeconomic and interest rate data. Surprise changes of the three components are identified using sign restrictions, and their macroeconomic underpinnings are studied via impulse response analysis. The curvature factor is found to carry predictive information both about the future evolution of the yield curve and of output. In particular, unexpected increases of the curvature precede a flattening of the yield curve and announce a significant decline of output about one year ahead. Surprise surges of the yield curve level anticipate large persistent increases in inflation and a hump-shaped response of output growth. Somewhat contrary to conventional wisdom, positive slope surprises are followed by an immediate though not very pronounced decline in output.

4.1 Introduction

It is widely accepted that the yield curve carries information about the prospective evolution of economic activity, inflation, and monetary policy. Interest rate spreads, for example, are often used as predictors of recessions and inflation. Since the yield curve assumes similar shapes over time, it is common to think of it in terms of the three factors level, slope, and curvature which together explain almost all of the cross-sectional variation of interest rates. However, despite the factor structure of the yield curve and its informational richness, there is only scattered evidence about the predictive content of each of its components. Therefore, this chapter provides a systematic analysis of the economic underpinnings of level, slope, and curvature by studying the evolution of key macroeconomic variables subsequent to surprise changes of the three components. While partly confirming conventional wisdom, the most important result is that unexpected changes of the curvature factor are more informative about the future evolution of the yield curve and of output than has previously been acknowledged.

To carry out this exercise I use a Bayesian factor model of macroeconomic and interest rate data that has the following properties. Yields are decomposed into three factors as recently suggested by Diebold and Li [2006]. Their approach is a variant of the Nelson and Siegel [1987] functional form and allows a straightforward interpretation of the latent factors as level, slope, and curvature. Macroeconomic variables are also assumed to have a factor structure. This has two advantages. On the one hand, it allows to study the dynamic effects of yield curve shocks on various macroeconomic variables. On the other hand, it represents a remedy to the problem caused by the use of revised data in studies of the macro-finance link. Finally, both sets of factors - macro and term structure - share common dynamics within a VAR. Together, these features of the model allow an unrestricted set of interactions between the term structure and the real economy. In this respect, the model goes beyond previous macro-finance models of the term structure such as Ang and Piazzesi [2003] or Hördahl et al. [2006] who only offer a unidirectional linkage from macroeconomic variables to the yield curve.

My model is similar in spirit to the one studied in Diebold, Rudebusch, and Aruoba (2006, DRA henceforth). These authors also allow for a bidirectional linkage between macroeconomic variables and use the same factor decomposition of yields. Although related, my model differs in a number of important dimensions from the one studied in DRA. First, DRA include only three individual economic variables in their macro-finance model of the term structure. Employing a factor structure, my model in contrast incorporates a broad macroeconomic information set. Second, while the joint dynamics of macro variables and term structure factors in DRA is limited to a VAR of order one, my model includes more lags. Overall, the model in this chapter exhibits a richer structure and therefore allows a more comprehensive analysis of macro-term structure dynamics than previous studies.

The additional generality comes at the cost of computational complexity. To estimate the model, I therefore build on recent advantages in Bayesian dynamic factor model analysis. As has been recognized by Kim and Nelson [1999] and others, Gibbs sampling algorithms are well-suited to approximate the joint posterior distribution of parameters and unobserved factors in state-space models. In the application of such methods to my model, a complication arises due to the non-standard distribution of the exponential decay parameter in the Nelson-Siegel spline functions. This is solved by adding a Metropolis step to the Gibbs sampler.

The third crucial difference with respect to the study by DRA regards the identification of shocks. In order to dissect the informational content of yield curve innovations into level, slope, and curvature, it is important to properly identify the surprise changes of the three components. This is an intricate issue. As has been pointed out by Sarno and Thornton [2004], zero restrictions on impulse responses of financial variables to contemporaneous macroeconomic shocks are inconsistent with the efficient market hypothesis. Hence, an appropriate identification scheme must allow the yield curve factors to contemporaneously react to all macroeconomic shocks. This is not the case in the recursive identification scheme employed by DRA who order the yield curve factors first. It could be achieved by ordering the yield factors last. Yet, this identification would preclude macroeconomic variables from contemporaneously responding to yield curve surprises, an assumption that is quite restrictive. I solve this problem by using identification schemes which make use of sign restriction techniques as recently suggested by Uhlig [2005] and Mountford and Uhlig [2005]. Precisely, to identify a positive surprise change of, say, the curvature factor, the impulse responses of the level and slope are restricted to be zero on impact while the response of the curvature is required to be positive over some periods after the shock occurs. At the same time, the impulse responses of the macroeconomic variables remain unrestricted. Imposing zero restrictions on individual yield curve factors might be an unrealistically strong restriction. Therefore, I also study a variant of the above identification scheme in which initial responses to the two remaining factors are required to be small but not necessarily zero. Technically, this is achieved using a penalty function approach.

The results of this chapter can be summarized as follows. Most importantly, I find unexpected changes of the curvature factor to be more informative about the future evolution of the yield curve and macroeconomic variables than has previously been recognized. In particular, positive surprise changes of the curvature factor announce a strongly significant and very persistent hump-shaped movement of the yield curve slope and a significant decline of the yield curve level. Together, these two features imply a successive flattening of the yield curve which is commonly associated with an upcoming recession. This is paralleled by a pronounced hump-shaped response of output. In particular, the growth rate of industrial production increases sharply for about three months, then slowly declines, and eventually falls below zero one year after the initial surprise. Hence, unexpected rises of the curvature factor - not accompanied by simultaneous changes of the yield curve level or slope - appear to announce economic slowdowns. This result is surprising since the curvature factor has previously been documented to be unrelated to macroeconomic variables. DRA, for example, report negligible responses of macro variables to shocks in the curvature factor, an observation restated in Diebold, Piazzesi, and Rudebusch [2005]. Yet, their results are based on a recursive identification inconsistent with the efficient market hypothesis. Dewachter and Lyrio [2006] estimate latent yield factors within an essentially affine term structure model and find evidence suggesting that the curvature factor is related to real interest rate movements that are uncorrelated with macroeconomic variables. However, they base their results on regression analysis and do not investigate the information carried by yield curve innovations. Evans and Marshall [2004] study the responses of yield curve factors to specific macroeconomic shocks which they obtain based on estimations of theoretical models. They conclude that the curvature is largely unaffected by macroeconomic shocks. However, their approach does not allow to study the evolution of macroeconomic variables subsequent to surprise changes of the yield curve factors which is the focus of this chapter.

Somewhat more consistent with conventional wisdom, surprise surges of the level factor - not paralleled by simultaneous changes of the slope or curvature - announce strong and persistent movements in inflation. They also anticipate real effects: positive level surprises are followed by a significant hump-shaped response of output growth. In contrast, slope surprises are followed by an immediate decline of output.[26] Yet, the responses are of moderate size and not statistically significant across all different identification schemes applied. Alto-

[26] Note that according to the Diebold-Li formulation of the factor loadings, a positive slope shock is associated with a strong increase of short-term rates and a small increase of long maturities, i.e. a flattening of the yield curve.

gether, unexpected rises of the slope factor seem to announce economic downturns shortly before they start, but the connection is surprisingly decent given that yield spreads are popular predictors of recessions.

The chapter is organized as follows. I present the empirical model that has been used to study the joint dynamics of the yield curve and the macroeconomy in Section 4.2. Section 4.3 briefly discusses identification of the model and its estimation via a Metropolis-within-Gibbs sampling algorithm. In Section 4.4, I present the different approaches used to identify surprise changes of the yield curve components and summarize the main empirical results obtained in Section 4.5. Section 4.6 concludes the chapter. Details on the Metropolis-within-Gibbs sampling algorithm used to estimate the model are provided in Appendix 6.3.2.

4.2 The Model

Assume that yields of different maturities are driven by three common factors and an idiosyncratic component,

$$Y_t = \Lambda_y F_t^y + e_t^y, \tag{4.1}$$

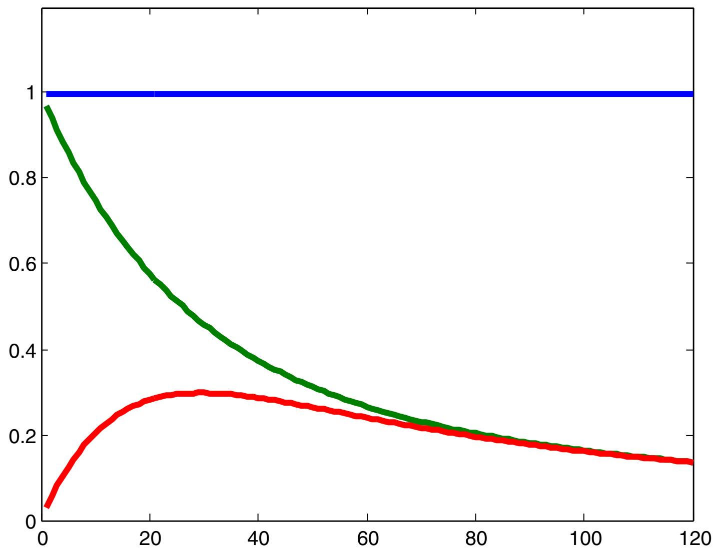
where Y_t is a $N_y \times 1$ vector of yields of different maturity, Λ_y is a $N_y \times 3$ matrix of factor loadings, F_t^y is a 3×1 vector of factors, and e_t^y is a $N_y \times 1$ vector of idiosyncratic components or pricing errors. The common factors explicitly represent the level, slope, and curvature of the yield curve. Recently, Diebold and Li [2006] have suggested the following variant of the well-known Nelson-Siegel (1987) decomposition of yields. In their model, the factor loadings are given by

$$\lambda_n^y = \left[1 \quad \left(\frac{1-e^{-\tau n}}{\tau n}\right) \quad \left(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)\right], \tag{4.2}$$

where τ denotes a shape parameter and n maturity. Hence, the loading on the first factor equals 1 for all yields. A shock to this factor therefore results in a simultaneous upward or downward shift of yields of all maturities. Accordingly, it has a clear-cut interpretation as a level factor. The loadings on the second factor are given by the functions $(\frac{1-e^{-\tau n}}{\tau n})$. Independently of the value of τ, this function assumes its maximum at $n = 0$ and then decays towards zero as n increases. The exponential decay parameter τ governs the speed of convergence. Shocks to the second factor thus affect short yields much stronger than long-term interest rates

Figure 4.1: **Diebold-Li Loadings**

This figure plots the Diebold and Li [2006] loadings corresponding to the level, slope, and curvature factor. The exponential decay parameter has been set to $\tau = 0.0609$.



and it therefore has a straightforward interpretation as a slope factor. Finally, loadings of yields on the third factor are given by functions of the form $(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n})$. For $n = 0$, this function has a value of zero. As Figure 4.1 shows, the function value increases with maturity and eventually reverts towards zero. Accordingly, the third factor mainly captures movements in medium-term maturities and can be interpreted as a curvature factor.[27] Applying the Diebold-Li setup has several advantages. First, it provides a very parsimonious way of decomposing yields into few common factors. Indeed, Diebold and Li [2006] have shown that their model reproduces different yield curve shapes and fits the term structure well over time. Second, as movements of the term structure of interest rates are often stated in terms of the level, slope, and curvature, these factors have become economic concepts of independent interest. It is thus appealing to separately study their predictive content. Note that the no-arbitrage assumption is not explicitly taken into account in this setup. However, as it explains yields of all maturities very precisely, the model should approximately capture no-arbitrage to the extent that it is satisfied in the data.

[27] Note that the shape parameter τ determines for which maturity the function assumes its maximum. Diebold and Li [2006] set $\tau = 0.0609$ which implies a maximum at $n = 30$ months that the authors choose as a reference maturity for the "medium term".

Assume further that macroeconomic variables in the model are driven by a few common factors and an idiosyncratic component, i.e.

$$X_t = \Lambda_x F_t^x + e_t^x, \tag{4.3}$$

where X_t is a $N_x \times 1$ vector of period-t observations of the variables in the panel, Λ_x is a $N_x \times k_x$ matrix of factor loadings, F_t^x is the $k_x \times 1$ vector of period-t observations of the common factors, and e_t^x is an $N_x \times 1$ vector of idiosyncratic components. By construction, the factors F^x capture the common variation in a large number of economic times series. Accordingly, impulse responses of different macroeconomic time series to surprise changes of the yield curve factors can be studied. Another reason for employing a factor model approach instead of using individual macro variables relates to the problem of data revisions. In fact, a common objection against empirical macro-finance models is that data revisions imply that the information set available to the econometrician is different from the information set available to investors. Hence, estimates of the parameters governing the mutual interactions between macroeconomic and financial variables may be biased. Studying the interaction of financial variables with the common components of many macro variables represents one way to address this critique. Indeed, assuming that data-revision errors are series-specific (see e.g. Bernanke and Boivin 2003, Giannone et al. 2004), the common factors extracted from different vintages of the same macroeconomic dataset will be the same. Thus, factor estimates obtained from revised data span the space of information available to investors in real-time.

Obviously, equations (4.1) and (4.3) have the same structure. I therefore consider the unified framework

$$\begin{aligned} \begin{pmatrix} X_t \\ Y_t \end{pmatrix} &= \begin{bmatrix} \Lambda_x & 0 \\ 0 & \Lambda_y \end{bmatrix} \begin{pmatrix} F_t^x \\ F_t^y \end{pmatrix} + \begin{pmatrix} e_t^x \\ e_t^y \end{pmatrix} \\ \text{or} \qquad Z_t &= \Lambda \quad F_t \quad + \quad e_t. \end{aligned} \tag{4.4}$$

I assume that the idiosyncratic disturbances are mutually orthogonal and not auto-correlated, i.e. $E[ee'] = R$ is a diagonal $N \times N$ matrix where $N = N_x + N_y$.[28] A central feature of my model is the assumption that a few structural shocks cause

[28] The orthogonality assumption is traditionally made in exact factor models. Relaxing it in order to encompass approximate factor structures is possible in the framework studied here, but would introduce additional complexity to the model.

the common variation in both sets of variables. By construction, the comovement is captured by the two sets of factors, F^x and F^y. Their common dynamics are modeled within a VAR, i.e.

$$\begin{aligned} \begin{pmatrix} F_t^x \\ F_t^y \end{pmatrix} &= \Phi(L) \begin{pmatrix} F_{t-1}^x \\ F_{t-1}^y \end{pmatrix} + \begin{pmatrix} \omega_t^x \\ \omega_t^y \end{pmatrix} \\ \text{or} \qquad F_t &= \Phi(L) \quad F_{t-1} \quad + \quad \omega_t. \end{aligned} \tag{4.5}$$

The reduced-form errors $\omega_t = (\omega_t^{x\prime} \omega_t^{y\prime})'$ are assumed to have non-diagonal variance-covariance matrix $E[\omega\omega'] = \Omega$. Furthermore, the idiosyncratic disturbances of individual variables and the shocks driving the common factors are assumed to be mutually independent.

4.3 Estimation of the Model

Before estimating the model, it needs to be ensured that its parameters and latent factors are uniquely identified. Exact identification is crucial since observationally equivalent sets of factors and parameters may give rise to the same likelihood but lead to different economic conclusions. Hence, assumptions need to be made which exclude such indeterminacies. A standard identification approach in factor models of the form (4.4)-(4.5) is due to Geweke and Zhou [1996]. These authors show that restricting the upper $k \times k$ block of Λ to be lower-triangular with positive diagonal elements uniquely determines the factors and loadings. A variant of this "hierarchical" identification approach is employed here. A complication arises due to the fact that two separate groups of factors drive the common dynamics of the variables Z. However, as shown in appendix 6.3.1, it is sufficient to restrict the upper $k_x \times k_x$ block of the submatrix Λ^x to be lower-triangular in order to ensure unique identification of the model.

Estimation of the model via maximum likelihood techniques is infeasible due to the large number of model parameters. An alternative would be to independently extract factors from both sets of variables via e.g. principal components and to study their joint dynamics within a VAR. This two-step estimation approach is inefficient, however, as it does not permit to jointly estimate the factors and model parameters. Recently, some authors have started estimating large-scale dynamic factor models via likelihood-based Markov Chain Monte Carlo (MCMC) methods (see e.g. Eliasz 2002, Bernanke et al. 2005, Kose et al. 2003). In particular the Gibbs sampling algorithm has is an useful estimation device. The Gibbs sampler is based upon iterative draws from the conditional posterior distributions of the individual model parameters given the data and the remaining parameters. Gibbs sampling for dynamic factor models involves an additional step for drawing the unobserved factors conditional on the model parameters and the data. Kim and Nelson [1999] provide a nice introduction to the estimation of dynamic factor models using the Gibbs sampler. In this chapter, I take advantage of the recent methodological developments and set up an MCMC algorithm to estimate the model presented above. As will be discussed in more detail below, posterior conditional distributions cannot be derived for all parameters of the model and thus a Metropolis-within-Gibbs algorithm is employed.

Estimation of dynamic factor models via Gibbs sampling requires a state-space formulation of the model. For notational simplicity, I thus rewrite the model (4.4)-(4.5) in companion form as

$$Z_t = \bar{\Lambda}\,\bar{F}_t + \bar{e}_t \tag{4.6}$$

$$\bar{F}_t = \bar{\Phi}\,\bar{F}_{t-1} + \bar{\omega}_t \tag{4.7}$$

where $\bar{F}_t = (F_t, \ldots, F_{t-p+1})$ and where $\bar{\Lambda}, \bar{e}_t, \bar{\Phi}$, and $\bar{\omega}_t$ denote the companion form equivalents of Λ, e_t, Φ, and ω_t, respectively, and $\bar{R}$ and $\bar{\Omega}$ the corresponding variance covariance matrices.

Let $\theta = (\Lambda^x, \Lambda^y, R, \Phi, \Omega)$ denote the set of model parameters. Moreover, let $\tilde{X}_T = \{X_1, \ldots, X_T\}$ and $\tilde{Y}_T = \{Y_1, \ldots, Y_T\}$ be all T observations on yields and macro variables and let $\tilde{Z}_T = \{\tilde{X}_T, \tilde{Y}_T\}$. Analogously, let $\tilde{F}_T = \{\bar{F}_1, \ldots, \bar{F}_t\}$ denote all observations of the factors F. The objective is to generate samples from the joint posterior distribution $p(\theta, \tilde{F}_T|\tilde{Z}_T)$ of model parameters and unobserved factors. If this distribution is not given or is not standard so that drawing from it is infeasible, the Gibbs sampler allows to approximate it by the empirical distributions of simulated values from the conditional posteriors $p(\theta|\tilde{Z}_T, \tilde{F}_T)$ and $p(\tilde{F}_T|\tilde{Z}_T, \theta)$. After finding starting values θ^0, any iteration of the Gibbs sampler involves the following two steps:

Step 1: Draw $\tilde{F}_T^{(i)}$ from $p(\tilde{F}_T|\tilde{Z}_T, \theta^{(i-1)})$.

Step 2: Draw $\theta^{(i)}$ from $p(\theta|\tilde{Z}_T, \tilde{F}_T^{(i)})$.

The exact procedures to sample from the conditional distributions of factors and individual parameters are described in appendix 6.3.2. The crucial result employed in the Gibbs sampler is that the empirical distribution of draws from the conditional posterior densities converges to the joint marginal posterior distribution as the number of iterations goes to infinity. Accordingly, after discarding an initial number of draws (the "burn-in"), sampling from the known conditional posterior densities of factors and parameters is equivalent to sampling from their unknown joint posterior distribution.

4.4 Identification of Shocks

The main focus of this chapter is to analyze the information that surprise changes of level, slope, and curvature convey about the future evolution of key macroeconomic variables. To answer this question, it is crucial to properly disentangle shocks to the three components. As has been pointed out by Sarno and Thornton [2004], there is a fundamental problem related to the identification of shocks to financial variables in structural VARs. In particular, they argue that zero restrictions on impulse responses of financial variables to macroeconomic shocks are inappropriate under the assumption of efficient markets. Assume that US government bond yields are efficient market variables, i.e. variables which reflect all information relevant for their determination. Then, the critique by Sarno and Thornton implies that an identification scheme needs to be found which allows the yield curve components to contemporaneously react to all macroeconomic shocks. In principle, this could be achieved by ordering the term structure factors last in recursive structural VARs. However, the yield curve shocks so identified would not have contemporaneous effects on the macro factors. Having in mind a structural interpretation of shocks, this could be an appropriate assumption. In this chapter, though, I seek to analyze the predictive content of surprise changes of the yield curve. Unexpected movements of interest rates are likely associated with simultaneous changes of macroeconomic variables. I therefore allow the latter to contemporaneously move when a yield curve surprise occurs.

I employ two different approaches to carry out this identification. Both impose restrictions on the sign of impulse responses of one of the three yield factors and the initial impact of the other two factors and draw on previous work in Uhlig [2005] and Mountford and Uhlig [2005]. In the following, I briefly describe the general operating mode of identification via sign restrictions. I then discuss the two main identification procedures used to generate the results in this chapter.

4.4.1 Identifying Surprise Changes with Sign Restrictions

Identifying shocks via sign restrictions is based upon prior assumptions about the impact of a certain type of shock on different economic variables. Uhlig [2005], for example, uses prior restrictions on the responses of prices, the federal funds rate and nonborrowed reserves to identify contractionary monetary policy shocks. The economic reasoning behind this approach is that a monetary policy shock should be characterized by a rise of the federal funds rate and a decline of prices and nonborrowed reserves. In this chapter, I apply similar methods in order to identify surprise changes of the latent yield curve factors, without however attributing a structural interpretation to these surprises. In contrast, I investigate their informational content by studying the subsequent dynamic responses of the macro variables stacked in X. I start by introducing some useful notation and then turn to explaining the different identification strategies in detail.

Notation

For convenience, I restate the VAR in (4.5) which represents the state equation of my model:

$$F_t = \Phi_1 F_{t-1} + \Phi_2 F_{t-2} + \ldots + \Phi_p F_{t-p} + \omega_t,$$

where $F_t = (F_t^{x\prime}, F_t^{y\prime})'$ is a $k = (k_x+3)$ vector of common factors and where $\Omega = E[\omega_t \omega_t']$ is the constant unconditional variance-covariance matrix of the one-step ahead prediction errors. The aim is to identify structural shocks ν that are mutually uncorrelated and standardized to have unit variance, i.e. $E[\nu_t \nu_t'] = I_k$. In order to trace impulse responses of the factors F to the structural shocks ν, one needs to find a matrix A that satisfies $\omega_t = A\nu_t$. Obviously, this implies $\Omega = AA'$. Following Uhlig [2005], I define an impulse vector as the column of some matrix A that has this property. As he shows, any such impulse vector a can be obtained by performing a Cholesky decomposition $\Omega = \tilde{A}\tilde{A}'$ and multiplying $\tilde{A}$ with some $k \times 1$ vector q of unit length, i.e. $a = \tilde{A}q$.

To compute impulse responses of the factors F, stack (4.5) as

$$\mathbf{F} = \bar{\mathbf{F}}\boldsymbol{\Phi} + \boldsymbol{\omega}, \tag{4.8}$$

where $\mathbf{F} = [F_1, \ldots, F_T]'$, $\bar{F}_t = [F'_{t-1}, \ldots, F'_{t-p}]'$, $\bar{\mathbf{F}} = [\bar{F}_1, \ldots, \bar{F}_T]'$, $\boldsymbol{\Phi} = [\Phi_1, \ldots, \Phi_p]'$, and $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_T]'$. For any $k \times 1$ impulse vector a, let $\mathbf{a} = [a', 0_{1,k(p-1)}]'$ and

$$\Gamma = \begin{bmatrix} \multicolumn{2}{c}{\boldsymbol{\Phi}'} \\ I_{k(p-1)} & 0_{k(p-1),k} \end{bmatrix}.$$

Then, the impulse response $r_{a,i}(h)$ of factor i to an impulse a at horizon h can be computed as

$$r_{a,i}(h) = (\Gamma^h\, \mathbf{a})_i \quad \text{for} \quad h = 0, \ldots, H \tag{4.9}$$

In this chapter, a joint dynamic factor model of macro and yield data is set up. This modeling framework allows to trace responses to the impulses driving the factors F to the individual variables in Z. Hence, the economic interpretation of a given shock can be based on a much broader information set than in usual VAR studies. Denoting $r_a(h)$ the vector of impulse responses of the factors F to an impulse a at horizon h, the response of the n-th variable in Z to that impulse a, denoted $r^n_a(h)$, can be computed as

$$r^n_a(h) = \lambda'_n r_a(h) \quad \text{for} \quad h = 0, \ldots, H, \tag{4.10}$$

where λ_n is the n-th row of the factor loading matrix Λ.

Combination of Zero and Sign Restrictions

I identify shocks that have a positive impact on the yield curve level, slope, or curvature.[29] This is achieved by imposing a positive response of one of the three factors over a given interval after the shock occurs. In order to separate out the initial effects on level, slope, and curvature, additional zero restrictions are introduced which ensure that only one of the three components moves on impact. These restrictions are summarized in the following definition:

Definition 1 *A "pure" level impulse vector is an impulse vector a such that the response of the level factor is positive at horizons $h = 0, \ldots, H$ and such that the responses of the slope and the curvature factor are zero on impact.*

Similar definitions apply to the slope and curvature factor. How are the zero restrictions imposed in practice? Recall from (4.9) that the response on impact to

[29] The positivity assumption is made for normalization reasons and is not restrictive. Since the model is linear, all reported results equivalently apply to negative surprise changes after flipping signs of the impulse responses.

an impulse vector is given by the impulse vector itself. Hence, in order to identify shocks which produce a pre-specified zero response on impact of e.g. factors i and j, one needs to find vectors q of unit length such that

$$\begin{pmatrix} a_i \\ a_j \end{pmatrix} = \begin{bmatrix} \tilde{A}_{(i)} \\ \tilde{A}_{(j)} \end{bmatrix} q \equiv \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{4.11}$$

where $\tilde{A}_{(i)}$ and $\tilde{A}_{(j)}$ denote the i-th and j-th row of the lower-triangular Cholesky factor $\tilde{A}$, respectively. Finding such vectors q is a straightforward numerical exercise that can be eased by parameterizing the space of k-dimensional vectors of unit length. In the application studied in this chapter, the total number of factors is $k = 7$ and thus the parameterization

$$q = \begin{bmatrix} \cos(\alpha_1)\cos(\alpha_2)\cos(\alpha_3)\cos(\alpha_4) \\ \cos(\alpha_1)\cos(\alpha_2)\cos(\alpha_3)\sin(\alpha_4) \\ \cos(\alpha_1)\cos(\alpha_2)\sin(\alpha_3) \\ \cos(\alpha_1)\sin(\alpha_2) \\ \sin(\alpha_1)\cos(\alpha_5)\cos(\alpha_6) \\ \sin(\alpha_1)\cos(\alpha_5)\sin(\alpha_6) \\ \sin(\alpha_1)\sin(\alpha_5) \end{bmatrix} \tag{4.12}$$

can be used. According to this parameterization, any 7×1 vector of unit length is characterized by a set of six angles $\{\alpha_1, \ldots, \alpha_6\}$ defined over the interval $[0, 2\pi]$.

In the first step of the identification procedure, I use numerical optimization routines to find vectors q which - for a given draw of the parameters (Φ, Ω) - fulfil the zero restrictions in (4.11). In a second step, I discard those impulse vectors a that do not satisfy the sign restriction imposed on the impulse responses of the factor of interest. From the draws that are retained, I then compute median impulse responses and the corresponding confidence intervals.[30]

Imposing Orthogonality

Structural shocks are commonly assumed to be exogenous orthogonal disturbances hitting the economy. Yields are prices of financial assets that mainly reflect market expectations about future monetary policy and inflation. Therefore,

[30] To enhance the speed of the algorithm, I keep candidate impulse vectors which satisfy the restriction for the opposite sign and multiply them by minus one.

in this chapter I take the stand that surprise changes of the yield curve summarize term structure reactions to structural shocks rather than representing exogenous disturbances themselves. Accordingly, they need not necessarily be orthogonal. In order to analyze whether the results obtained above are robust to the assumption that the surprise changes to the three components are independent, one can additionally impose orthogonality restrictions. Using the parameterization (4.12), this amounts to finding a vector q_1 that satisfies the zero and sign restrictions for, say, the level factor. In a subsequent step, one then has to find a vector q_2 that fulfils the zero restriction for, say, the slope factor. In addition, it is required that q_1 and q_2 be orthogonal, i.e. $q_1'q_2 = 0$. Similar to the approach in Mountford and Uhlig [2005], this is achieved by stacking the zero and orthogonality restrictions,

$$\begin{bmatrix} \tilde{A}_{(i)} \\ \tilde{A}_{(j)} \\ q_1' \end{bmatrix} q_2 \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \tag{4.13}$$

where again $\tilde{A}_{(i)}$ and $\tilde{A}_{(j)}$ denote the i-th and j-th row of the lower-triangular Cholesky factor $\tilde{A}$. All candidates q_2 that fulfil these restrictions are then subjected to the sign restriction test and discarded if they do not meet them. In that case, the entire procedure is restarted by drawing a pair (Φ, Ω) from the joint posterior of the model parameters, finding a vector q_1 that satisfies the zero and sign restrictions for the level factor and so on. Once two vectors q_1 and q_2 have been found that fulfil all requirements, the third shock can be identified by finding vectors q_3 that meet the according zero restrictions and which are orthogonal to both q_1 and q_2. These candidates are then kept or discarded depending on whether they pass the sign restriction test.

Penalty Function Approach

The identification approach discussed above allows to clearly separate out shocks which on impact move only one of the three yield curve factors. Yet, surprise changes of the yield curve that translate into unilateral movements of one of the three factors may be rare events in reality and thus the imposed restrictions not supported by the data. Still, a thorough analysis of term structure dynamics reveals that many yield curve changes are largely driven by movements of one of its three components. In order to accommodate this behavior, I employ a second identification routine which eases the strong restriction of zero contemporaneous impact. In particular, using sign restrictions as above, I identify positive surprises of the three yield curve factors which exert as little as possible an effect on the remaining two components. This is achieved by applying the following penalty function identification approach which draws on the approach in Uhlig [2005].

I identify a surprise as, say, a "strong" level surprise if it implies a strong response of the level factor, but has as little as possible an impact on the slope and the curvature factor over a certain period of time after the initial impact. That is, for each draw of parameters (Φ, Ω) from equation (4.5), I find impulse vectors a that imply a large response of one yield curve factor while the responses of the other two factors are required to be close to zero. This is achieved by minimizing a function that accordingly rewards and penalizes the impulse responses of the three yield factors. An intuitive penalty function satisfying these requirements is

$$\Psi_i(a) = \sum_{h=0}^{H} \left[\gamma_i \frac{r_{a,i}(h)}{\sigma_i} \right] + \sum_{h=0}^{H_2} \left[\sum_{j \neq i} \gamma_j \left(\frac{r_{a,j}(h)}{\sigma_j} \right)^2 \right] \tag{4.14}$$

where i defines "level", "slope", or "curvature" and j the remaining two yield factors. Further, σ_i and σ_j denote the standard deviations of the first difference of the respective factors and are included in order to normalize the impulse responses. The weights γ_i and γ_j can be chosen such that particular emphasis is put on one of the potentially conflicting objectives.With this functional form, positive (negative) responses of one yield curve factor are rewarded and at the same time responses of the remaining two factors different from zero over a fixed number of periods after the shock occurs are penalized.[31] More formally, with the penalty function approach I have the following definition of a level shock:

Definition 2 *A "strong" level impulse vector is an impulse vector a that maximizes the penalty function* Ψ.

Again, similar definitions apply to the slope and curvature factor.

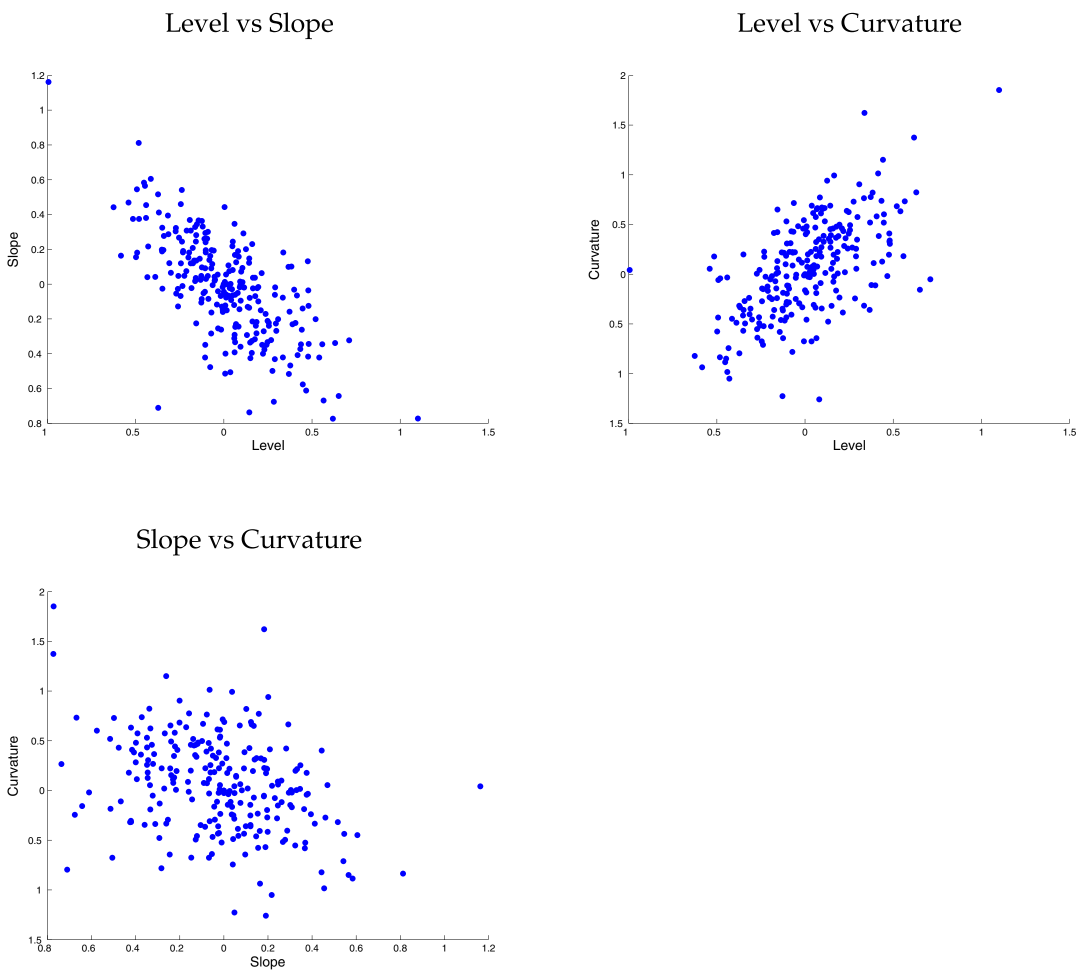
4.4.2 Impulse Responses to "Typical" Yield Curve Shocks

The identification schemes discussed above allow to study the economic implications of shocks to the yield curve which mainly or exclusively can be attributed

[31] Obviously, the choice of weights γ attributed to the responses of the three factors is arbitrary. As will be discussed further below, I report results for $\gamma = [-1, 100, 100]$. Hence, positive responses of the factor of interest are rewarded while any responses of the remaining two factors that deviate from zero are penalized with a much stronger weight.

Figure 4.2: **Innovations of Level, Slope, and Curvature**

This figure provides scatter plots of the model-implied one-step ahead forecast errors of the three yield curve factors level, slope, and curvature. The estimation period is 1983:01-2003:09.



to one of its three components level, slope, and curvature. While this helps to dissect the information in the term structure, it does not allow to study the economic implications of yield curve surprises that affect all three components. Yet, this might be a relevant scenario. Figure 4.2 relates the model-implied one-step ahead forecast errors of the level, slope, and curvature factors. As these scatter plots show, the forecast errors of level and slope, and those of level and curvature are indeed noticeably correlated. This indicates that there is a strong common component in the innovations to the three yield curve factors. By identifying this common component one can thus, roughly speaking, study the impulse responses of macroeconomic variables to a "typical" yield curve shock.

How can this be done? One needs to perform an eigen decomposition of the variance-covariance matrix of forecast errors of the yield factors. Then, by construction, the eigenvector corresponding to the largest eigenvalue explains most of the common variation in the forecast errors.[32] In my model, the yield factors constitute a subgroup of the state variables. Hence, the variance-covariance matrix of all forecast errors needs to be decomposed into blocks. In particular, let

$$\Omega = \begin{pmatrix} \Omega_{xx} & \Omega_{xy} \\ \Omega'_{xy} & \Omega_{yy} \end{pmatrix} \tag{4.15}$$

where Ω_{xx}, Ω_{xy}, and Ω_{yy} are of dimension $k_x \times k_x, k_x \times 3$, and 3×3. Define D as the 3×3 diagonal matrix with entries given by the eigenvalues of Ω_{yy} in descending order and V as the matrix of corresponding eigenvectors. Then, Ω_{yy} can be written as

$$\begin{aligned} \Omega_{yy} &= A_{yy} A'_{yy} \\ \text{where } A_{yy} &= V D^{1/2}. \end{aligned} \tag{4.16}$$

Now the task at hand is to find a matrix A that has A_{yy} as the lower-right block and which further satisfies $\Omega = AA'$. One possible decomposition fulfilling this condition is given by

$$A = \begin{pmatrix} A_{xx} & A_{xy} \\ 0 & A_{yy} \end{pmatrix} \tag{4.17}$$

where A_{xx} and A_{xy} are of dimension $k_x \times k_x$ and $k_x \times 3$, respectively. Then, together with the condition $\Omega = AA'$, (4.15) and (4.17) imply

$$\begin{aligned} \Omega_{xy} &= A_{xy} A'_{yy} \\ \text{and thus } A_{xy} &= \Omega_{xy} (A'_{yy})^{-1}. \end{aligned} \tag{4.18}$$

Moreover, we have

$$\Omega_{xx} = A_{xx} A'_{xx} + A_{xy} A'_{xy}.$$

Hence,
$$A_{xx} A'_{xx} = \Omega_{xx} - \Omega_{xy} (A'_{yy})^{-1} (A_{yy})^{-1} \Omega_{yx}. \tag{4.19}$$

One solution to this equation is given by the Cholesky decomposition of the term $\Omega_{xx} - \Omega_{xy}(A'_{yy})^{-1}(A_{yy})^{-1}\Omega_{yx}$.[33] Altogether, the (k_x+1)-st column of any decom-

[32] Notice that similar approaches of shock identification have been employed in Uhlig [2004] and Giannone et al. 2004.

[33] Notice that an infinite number of orthogonal rotations of the Cholesky factor A_{xx} can be found which equally solve equation (4.19). Being only interested in shocks to the yield curve factors corresponding to the last three columns of A, this does not, however, affect the results of my analysis.

position of the form (4.17), with A_{xx}, A_{xy}, and A_{yy} as given above, represents the impulse vector which explains the largest share of the common variation of one-step ahead forecast errors of the three yield curve factors. It can thus be interpreted as a "typical" yield curve shock. Note that this decomposition also allows to study the share of variance of individual macroeconomic variables explained by those shocks which together capture all of the variation in the three yield curve factors. This may serve as a rough measure of how much of the variation in macroeconomic variables is anticipated by movements in the yield curve.

4.5 Empirical Results

This section summarizes the results of the chapter. First, I describe the data used and discuss the empirical specification. Then, I document the fit of the joint factor model of macro and yield data before I finally turn to the results of the impulse response analysis.

4.5.1 Data and Model Specification

I estimate the model using monthly data for the US from 1983:01 until 2003:09. This time span covers the post-Volcker disinflation period and can thus be seen as a consistent monetary policy regime. The interest rate data used in this study are unsmoothed Fama-Bliss yields for maturities 1, 3, 6, 9, 12, 15, 18, 21, 24, 30 months and 3, 4, 5, 6, 7, 8, 9, and 10 years.[34] Hence, the term structure information is extracted from a wide range of maturities. I have further selected 25 variables of different economic categories in order to exploit the information in a variety of different macroeconomic time series. Table 4.1 lists these variables. Note that they are taken from a more comprehensive macroeconomic data panel for the US that has been compiled by Giannone et al. [2004].[35]

As discussed above, the model decomposes yields into three factors which have a straightforward interpretation as level, slope, and curvature. The true number of factors driving the macroeconomic variables in the panel X is not known, however. Bai and Ng [2002]have developed formal tests for the optimal number of

[34] I am grateful to Robert Bliss for sharing these data with me.

[35] I thank Lucrezia Reichlin for letting me use these data. For further details on the origin of these data, the reader is referred to the paper by Giannone et al. [2004].

Table 4.1: **Macro Variables and Share of Variance Explained by Factors**

This table lists the 25 macro variables that have been used to estimate the dynamic factor model for macro and yield data. As is common in the literature, most series have been subjected to some transformation prior to the estimation, provided in the second column of the table. The transformation codes are: 0 = no transformation, 1 = logarithm, 2 = monthly differences, 3 = quarterly growth rate, 4 = annual growth rate. The third column provides for each of the 25 variables the share of variance explained by the four estimated factors.

Series	Transf.	R^2
Index of IP: Total	3	0.724
Capacity Utilization: Total	2	0.805
Purchasing Managers Index	0	0.901
Index of help-wanted advertising	3	0.618
Employment on nonag payrolls: Total	3	0.573
Avg weekly hrs. of production or nonsupervisory workers	3	0.158
Personal Cons. Expenditure: Total	3	0.084
Construction put in place: Total	3	0.344
Inventories: Mfg and Trade: Total	3	0.203
ISM mfg index: new orders	0	0.831
NYSE composite index	3	0.989
S&P composite	3	0.990
Nominal effective exchange rate	3	0.036
M1	3	0.085
M2	3	0.981
M3	3	0.656
Loans and Securities @ all commercial banks: Total	3	0.191
CPI: all items (urban)	4	0.940
PPI: finished goods	4	0.608
PCE chain weight price index: Total	4	0.960
Avg hourly earnings: Total nonagricultural	4	0.235
Philadelphia Fed Business Outlook: General activity	0	0.746
Outlook: Prices paid	0	0.556
Outlook: Prices received	0	0.531
Federal govt deficit or surplus	3	0.013

factors in large-scale factor models estimated using static principal components analysis. Model choice is a much more intricate issue in Bayesian factor analysis, at least from a computational point of view. Justiniano [2004] discusses methods for selecting the optimal number of factors in Bayesian factor models that allow for lags in the observation equation. I present results for a model which extracts four factors from the set of 25 macroeconomic time series.[36]

[36] Unreported results have shown that model specifications using different numbers of factors yielded qualitatively very similar results.

Another choice to make regards the number of lags in the joint VAR of macro and yield factors which represents the state equation of my model. As in standard VAR studies, there exists a tradeoff between parsimony and richness of the dynamic structure. I document results for a model with $p = 6$ lags which seems to be a good compromise between the two conflicting objectives.

4.5.2 Convergence of the Metropolis-within-Gibbs Sampler

The results reported below are based on 8,000 simulations of the Metropolis-within-Gibbs sampler summarized in appendix 6.3.2. In order to ensure convergence to the ergodic distribution, the first 3,000 iterations have been discarded as a burn-in. The remaining 5,000 draws have been used to compute median estimates and standard errors of the model factors and parameters. Figures 4.3 and 4.4 provide plots of factor estimates and their 95% confidence bands. As one can see from these plots, all factors are sharply estimated.

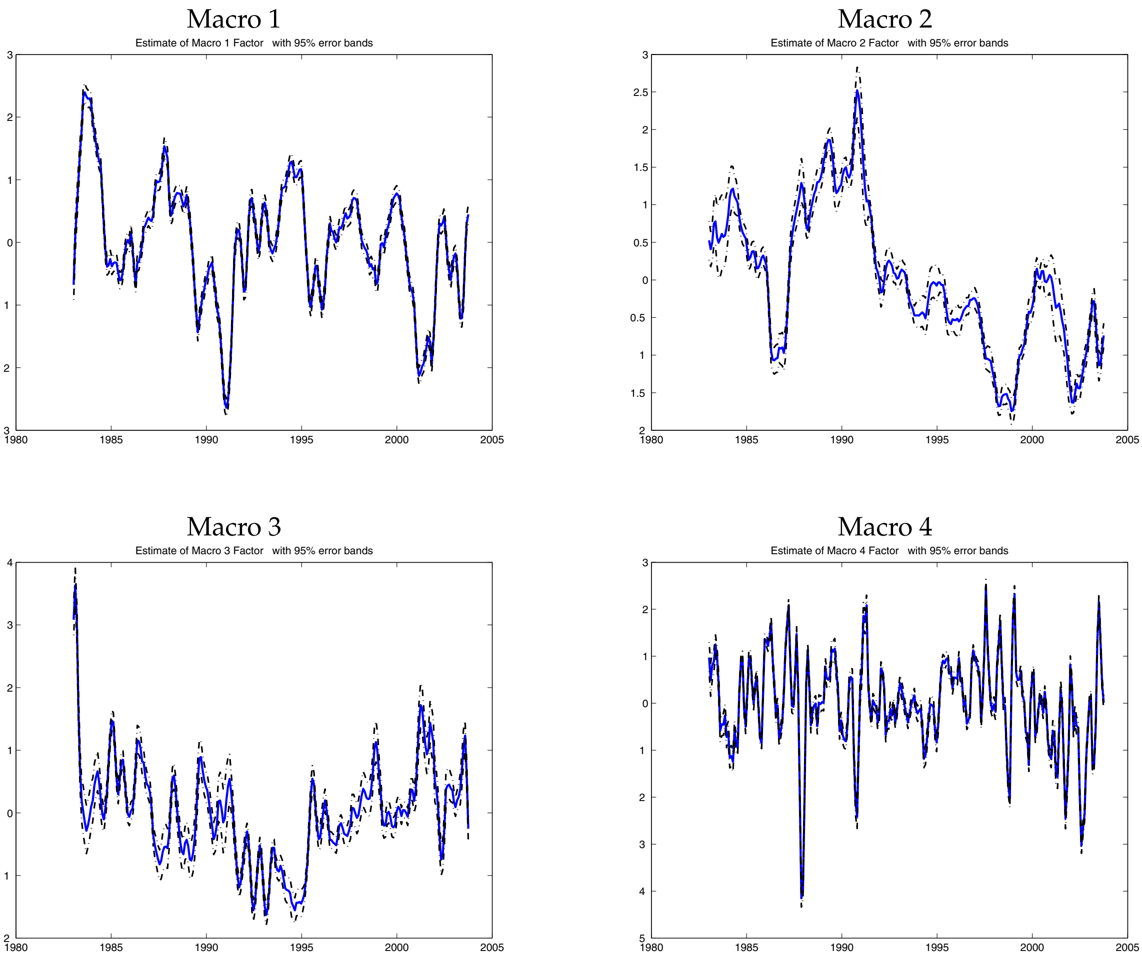
It has been discussed above that specific restrictions on the elements of Λ^x are imposed to ensure exact identification of the model. As a cross-check on whether the model is uniquely identified and as another test of convergence, I executed the sampler several times using different sets of randomly generated starting values. Figures 4.12 and 4.13 in Appendix A.4 show that the factor estimates are highly similar across the different initializations of the sampler. I interpret this as additional evidence for convergence of the algorithm. Finally note that to compute impulse responses, I randomly selected pairs (Φ, Ω) from the stored draws of their posterior conditional distributions. The random selection is important since MCMC draws are usually autocorrelated and therefore error bands might be understated using consecutive draws. To address this problem, once could alternatively employ e.g. every 10 draws (see also Law 2004).

4.5.3 Assessing the Model Fit

In this section, I provide a set of results that allow to assess how well the factor model fits the data. Table 4.1 lists for all macro variables in the panel the shares of variance explained by the four estimated factors. According to these figures, most variables exhibit a rather large common component. In particular, variables related to output and inflation are well explained by the four factors. About 74%

Figure 4.3: **Estimated Macro Factors and 95%-Confidence Intervals**

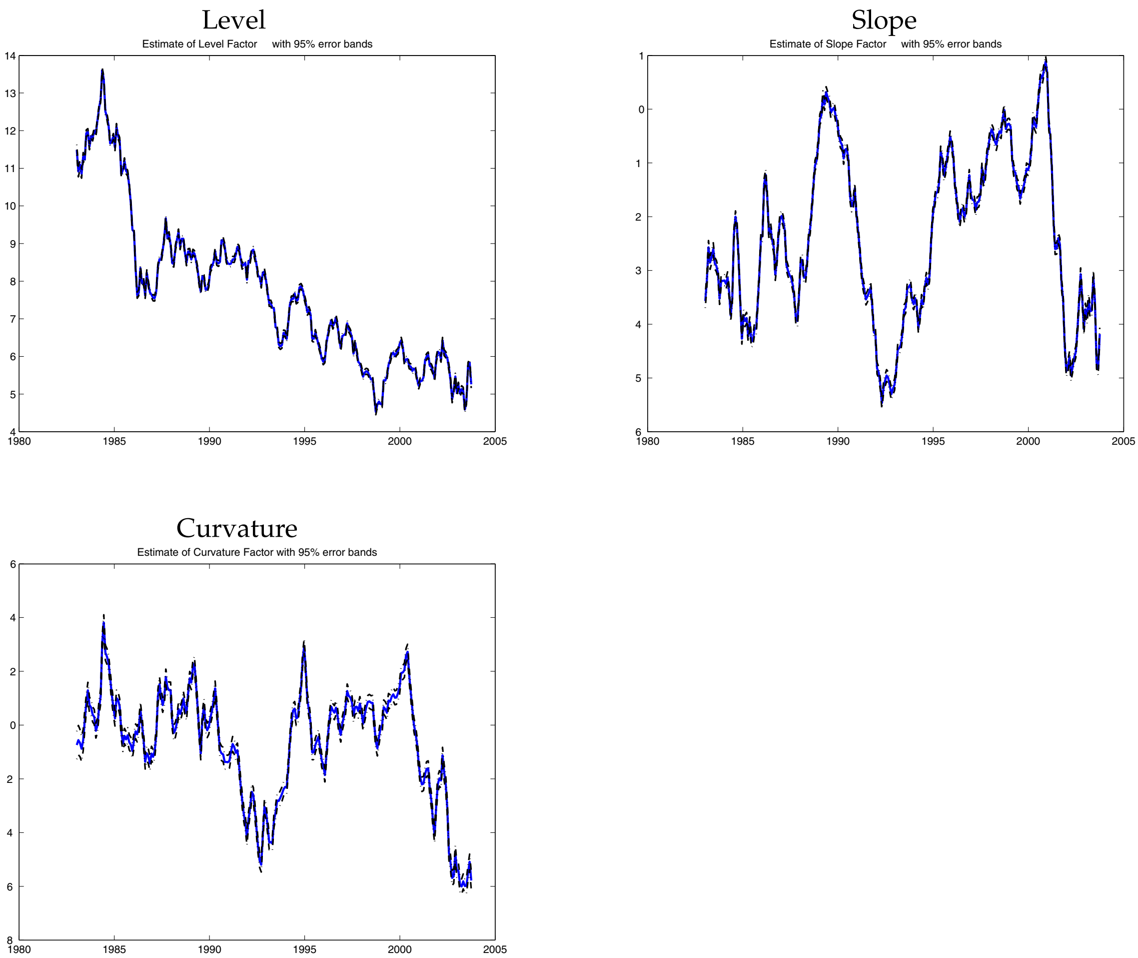
This figure provides plots of the estimated macro factors together with their 95%-confidence intervals. The estimates have been obtained as the median of the factor draws kept after the initial burn-in period of the Metropolis-within-Gibbs sampler. The estimation period is 1983:01-2003:09.



of the growth rate of industrial production, for example, is captured by the macro factors. Moreover, about 90% of the variation of annual CPI inflation is explained by the common components. Monetary aggregates, stock indices, and other variables are equally well explained by the four factors. However, the variation of some time series such as personal consumption expenditure, the nominal effective exchange rate, and the federal government deficit is largely attributable to variation not captured by the common factors. Obviously, impulse responses of these variables to shocks driving the common factors would be more or less meaningless. Hence, the results shown below are exclusively based only on variables

Figure 4.4: **Estimated Yield Factors and 95%-Confidence Intervals**

This figure provides plots of the estimated yield factors together with their 95%-confidence intervals. The estimates have been obtained as the median of the factor draws kept after the initial burn-in period of the Metropolis-within-Gibbs sampler. The estimation period is 1983:01-2003:09.

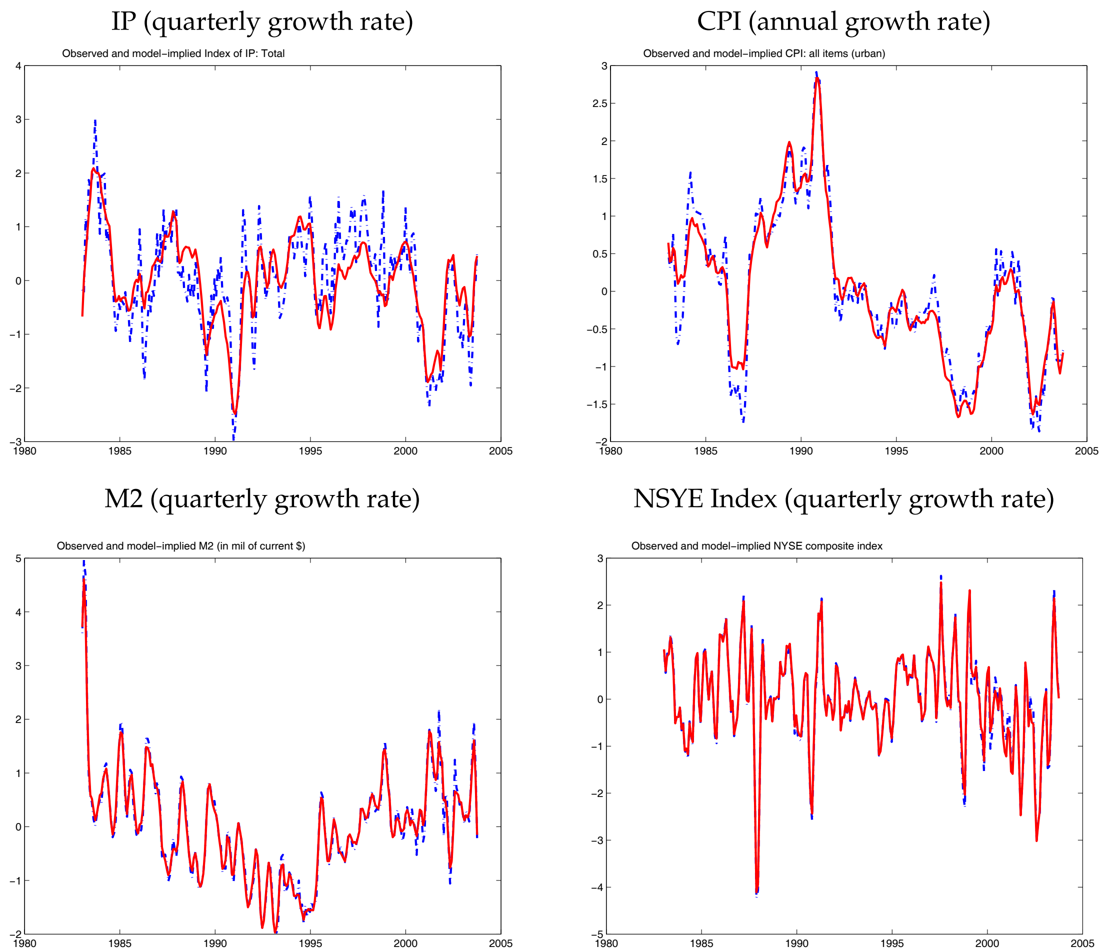


that have an important common component. Figure 4.5 provides plots of some selected macro variables and how well they are explained by the common factors.

As has already been documented in previous studies (e.g. Diebold and Li 2006, Diebold et al. 2006), the Nelson-Siegel decomposition of yields almost perfectly explains the cross-sectional variation of interest rates of different maturities over time. Table 4.2 shows that this result is confirmed by the analysis conducted here. Indeed, except for a few maturities at the very short and the very long end of the

Figure 4.5: **Model Fit - Macro Variables**

This figure provides plots of observed (dashed) and fitted (solid) values for a selection of macro variables.



curve, all yields are nearly perfectly matched by the three factors. This implies that very little information about the yield curve dynamics is lost by restricting the analysis to surprise changes of its three components. Figure 4.6 highlights this result visually by providing plots of observed and fitted yields for selected maturities.

4.5.4 Impulse Response Analysis

The main focus of this chapter is to shed light on the question what information surprise changes of the term structure level, slope, and curvature convey

Table 4.2: **Yields and Share of Variance Explained by Factors**

This table lists the 18 maturities of unsmoothed Fama-Bliss government bond yields that have been used for the estimation. The second to fourth column provides, for each maturity, the cumulative shares of variance explained by the level (L), slope (S), and curvature (C) factor.

Maturity	L	L,S	L,S,C
1-month	0.499	0.986	0.986
3-months	0.542	0.993	0.998
6-months	0.589	0.980	0.999
9-months	0.622	0.966	0.999
12-months	0.636	0.951	0.999
15-months	0.665	0.942	0.999
18-months	0.694	0.937	0.999
21-months	0.720	0.934	0.999
24-months	0.742	0.933	0.999
30-months	0.776	0.933	0.999
3-years	0.806	0.937	0.999
4-years	0.858	0.946	0.999
5-years	0.889	0.954	0.999
6-years	0.918	0.964	0.999
7-years	0.933	0.970	0.999
8-years	0.948	0.977	0.998
9-years	0.957	0.981	0.998
10-years	0.964	0.984	0.997

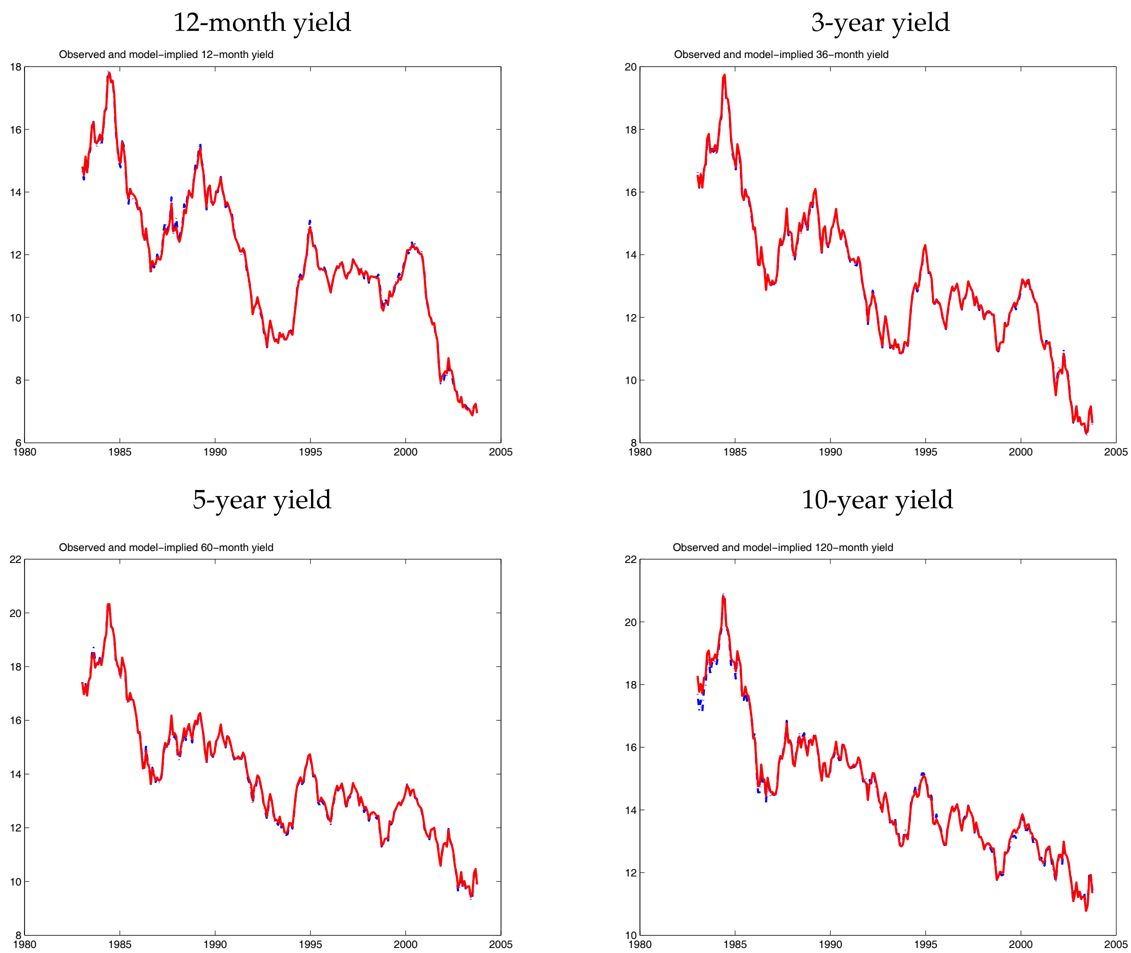
about the future evolution of the economy. As extensively discussed above, this is achieved by studying the impulse responses of macroeconomic variables to shocks which on impact exclusively (mainly) move one of the three components. In this section, I summarize the results obtained from these exercises. I start by presenting impulse responses to "pure"yield curve surprises which *exclusively* move one of the three components on impact. I then turn to the discussion of results obtained from the penalty function approach which identifies surprise changes that *strongly* move one of the three yield curve factors while the remaining two are restricted to move only little on impact. For the sake of brevity, I only report impulse responses for industrial production, CPI inflation, and M2. These three variables represent important economic categories and are all well explained by the common factors.

Combination of Zero and Sign Restrictions

The results documented in this section are based on a value of $H = 5$. Hence, responses of the yield curve factors to the impulses defined above are required to be

Figure 4.6: **Model Fit - Yields**

This figure provides plots of observed (dashed) and fitted (solid) yields for different maturities.



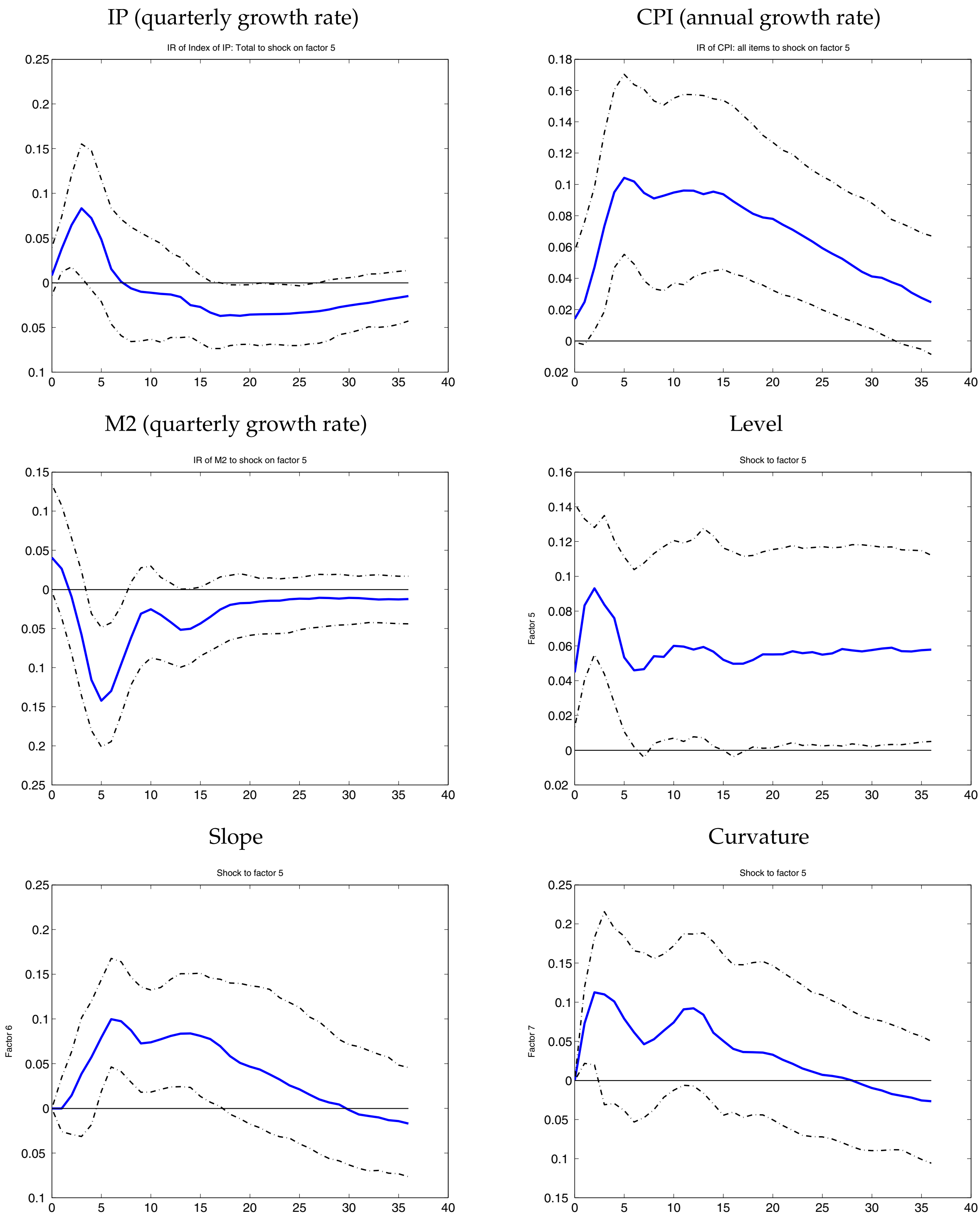
positive for half a year after the initial surprise. Since all three yield curve components are persistent time series processes, this restriction is not very strong. Unreported results have shown that the conclusions remain qualitatively unchanged when different sign restriction lengths are imposed.

Surprise Changes of the Level

Figure 4.7 plots impulse responses to a positive surprise change of the level factor for the three yield curve components and the three macro variables referred to above. Several remarks are in order. First, consistent with conventional wisdom, positive surprise changes of the yield curve level indicate a significant subsequent

Figure 4.7: **Impulse Responses - "Pure" Level Surprise**

This figure plots impulse responses of some selected macro variables and the three yield curve components after a "pure" level surprise as defined in Section 4.4.1. Hence, the response of the level factor is restricted to be positive over the first six periods after the shock occurs while the responses of the slope and curvature factors are restricted to be zero on impact. Dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of impulse responses.



increase of inflation. As the upper right panel of Figure 4.7 reveals, inflation rates rise for about five months after a level surprise and then slowly revert towards their initial level. Second, unexpected surges of the yield curve level also announce significant changes in output growth. According to the upper left panel of Figure 4.7, the growth rate of industrial production increases for about three months and then quickly reverts towards zero, eventually turning negative about seven months after the initial level surprise. While the response is not highly significant, it still indicates that positive surprise changes of the level factor announce a decline of output about six months to one year ahead.[37] Third, as the right middle panel of Figure 4.7 shows, a positive surprise change of the level factor is followed by a persistent upward shift of the yield curve level. In contrast, both slope and curvature show hump-shaped but not very pronounced responses to an unexpected rise of the level factor.

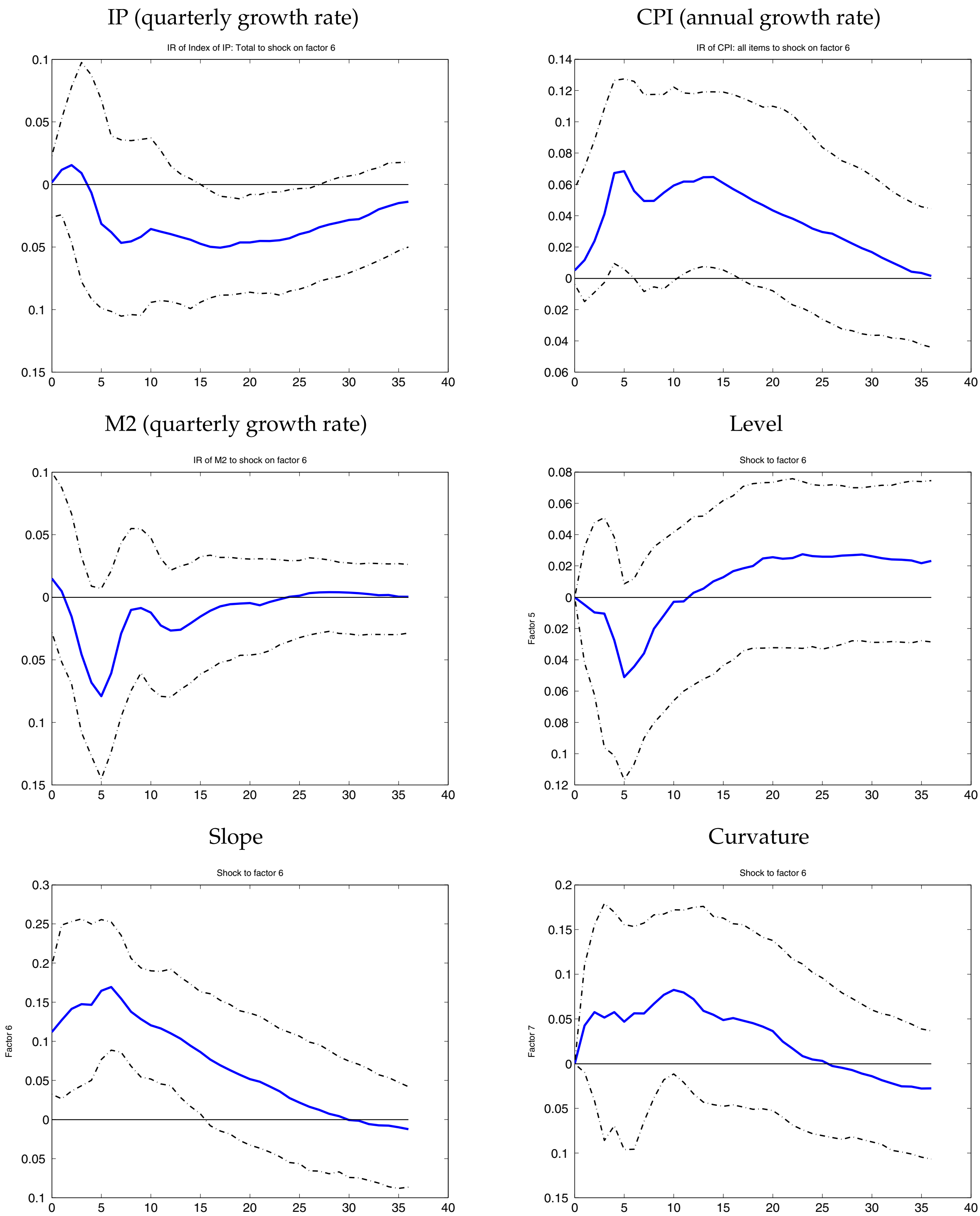
Surprise Changes of the Slope

Before turning to the impulse responses following a surprise change of the yield curve slope, it is worth recalling that term spreads are considered to be strong predictors of recessions. Indeed, as has been documented by Estrella and Hardouvelis [1991], Estrella and Mishkin [1998] and many others, spreads between long-term and short-term interest rates forecast recessions better than most other leading indicator variables. There are different mechanisms which may explain this fact. One is monetary policy. If the central bank raises short-term interest rates in order to avoid an overheating of the economy, long-term yields will typically increase by a smaller amount since inflation is assumed to be lower in the future. Hence, the yield curve becomes flatter or even inverted. A decreasing yield curve slope therefore announces an economic slowdown. A second explanation why the term spread should anticipate recessions is related to the expectations hypothesis of the term structure. If long-term interest rates equal average expected future short rates, then increasing (declining) long rates must indicate a monetary policy tightening (easing) which in turn is associated with a decrease (increase) of output growth. Finally, agents' savings decisions represent another channel to explain the link between the term spread and economic activity. In fact, if agents expect a recession to occur, they might decide to save more and invest in

[37] Note that somewhat consistent with this result, Ang et al. [2006] find that the short rate forecasts GDP growth better than term spreads.

Figure 4.8: **Impulse Responses - "Pure" Slope Surprise**

This figure plots impulse responses of some selected macro variables and the three yield curve components to a "pure" slope surprise as defined in Section 4.4.1. Hence, the response of the slope factor is restricted to be positive over the first six periods after the shock occurs while the responses of the level and curvature factors are restricted to be zero on impact. Dash-dotted lines indicate the 16% and 84% quantiles of the posterior distribution of impulse responses.



long-term bonds. Accordingly, the demand for long-term debt would increase, long-term yields would fall, and thus the term spread would decline. Altogether, one would expect that positive changes of the slope factor announce a decline of economic activity.[38]

Studying the responses of different economic variables to surprise changes of the yield curve slope, the exercises conducted in this chapter complement the mostly regression-based evidence on the link between the term spread and real activity. Figure 4.8 summarizes impulse responses to an unexpected increase of the slope factor when both level and curvature do not initially move. The upper left panel shows the impulse response of IP growth. At first sight, the conventional view that a rising slope factor (a decreasing spread between long-term and short-term bonds) announces an economic downturn seems to be confirmed since the growth rate of industrial production - after a brief initial rise - falls below zero. Yet, the link is surprisingly weak. Indeed, the downturn is little pronounced and the response of output is almost insignificant. This is a striking result given that the term spread is known to be a strong predictor of output growth. Moreover, the almost immediate decline of the growth rate of industrial production does not match the common view that declining spreads announce the beginning of a recession four to six quarters ahead. When discussing the effects of unexpected curvature shocks further below, I provide evidence that partly reconciles these findings with previous results. The reaction of inflation to an unexpected increase of the slope factor looks qualitatively similar to the response to a level shock. Also note that an unexpected increase of the slope factor is followed by an initial decline of the yield curve level for about five months and a subsequent rise above its initial level. The curvature factor instead increases for about ten months before it slowly reverts towards zero. Yet, neither of the two responses are significant implying that these patterns can hardly be regarded as systematic.

Surprise Changes of the Curvature

By contrast, the impulse responses following a surprise change of the curvature factor are strikingly pronounced. As the middle right and lower left panel of

[38] Recall that an increase of the slope factor in my model corresponds to a lower spread between long-term yields and short-term yields. This is because the loading of the slope factor decreases with maturity.

Figure 4.9: **Impulse Responses - "Pure" Curvature Surprise**

This figure plots impulse responses of some selected macro variables and the three yield curve components to a "pure" curvature surprise as defined in Section 4.4.1. Hence, the response of the curvature factor is restricted to be positive over the first six periods after the shock occurs while the responses of the level and slope factors are restricted to be zero on impact. Dash-dotted lines indicate the 16% and 84% quantiles of the posterior distribution of impulse responses.

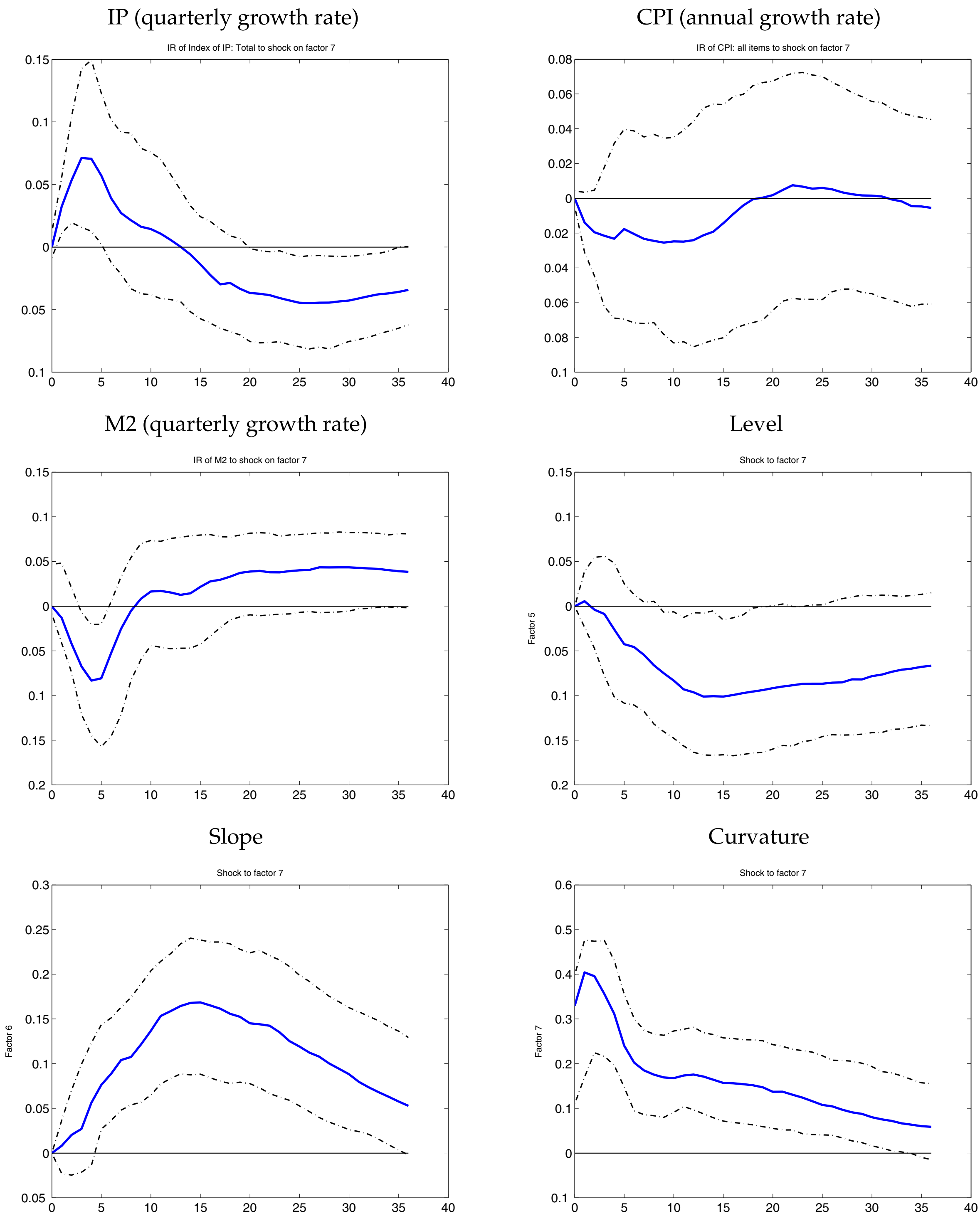


Figure 4.9 show, unexpected increases of the curvature factor are followed by a strong hump-shaped response of the yield curve slope. More precisely, the slope factor increases for about 15 months subsequent to a curvature surprise, and then gradually declines towards its initial level. At the same time, the level factor features a persistent decrease for about one year after the shock, before it slowly reverts towards its initial level. Hence, following a surprise surge of the curvature factor, long rates fall and short rates strongly increase which implies a flattening of the yield curve. As has been argued above, this is commonly associated with an upcoming recession. This interpretation is clearly mirrored in the response of output growth. In particular, subsequent to a curvature surge, the growth rate of industrial production rises sharply for about five months and then slowly reverts, eventually falling below zero about one year after the shock. Hence, an unexpected increase of the curvature factor anticipates a pronounced increase of the yield curve slope (a fall of the term spread) and at the same time announces a significant decline of output about one year ahead. This might call for the curvature instead of the slope as an early predictor of recessions.

Altogether, these results challenge conventional wisdom by providing evidence that unexpected changes of the yield curve slope do not anticipate significant declines of economic activity whereas unexpected changes of the curvature factor are more informative about the future evolution of output than has previously been acknowledged. Note that these results are not qualitatively altered by imposing the restriction that surprise changes of the three yield curve factors be orthogonal. As Figures 4.14 and 4.15 in Appendix A.4 show, the impulse responses follow identical patterns as for the case of non-orthogonal shocks, but exhibit slightly wider error bands.

Penalty Function Approach

I have argued above that the zero restrictions on the initial responses of the yield curve factors may be implausibly restrictive. I therefore also study scenarios where one of the three components strongly moves on impact while the initial responses of the other two are small. This is achieved with the penalty function approach described in Section 4.4.1 and the surprise changes so identified are labeled “strong” level, slope, and curvature shocks, respectively.[39] Figures 4.16 to

[39] The results reported refer to $H = 5$ and $H_2 = 0$, i.e. the response of the factor of interest is maximized over the first six months and the response of the remaining two components is restricted

4.18 in Appendix A.4 summarize the results obtained using the penalty function approach. Three remarks are in order. First, the impulse responses are qualitatively very similar to those obtained using the identification approach with zero restrictions. Second, individual impulse responses show a more clear-cut behavior. In particular, the hump-shaped pattern of output growth following a "strong" curvature surprise, and the decline of output subsequent to a strong slope surge are more pronounced than before. Third, the median impulse responses are estimated with higher precision than using the pure sign restriction approach as the relatively tight confidence bands indicate. This is not surprising since the penalty function algorithm selects impulse vectors from a subset of those captured with the pure sign restriction approach if one abstracts from the small nonzero initial responses that it allows.

4.5.5 Comparison to Recursive Identification

It has been discussed above that recursive identification schemes common in the SVAR literature are not appropriate to study the question pursued in this chapter. Nevertheless, it might be instructive to compare the impulse responses obtained using the sign restriction approaches with those from recursive identification schemes. Figure 4.19 in Appendix A.4 plots responses of IP growth and CPI inflation after shocks to the yield curve level, slope, and curvature, respectively. Solid lines represent the median estimate resulting from the pure sign restriction approach and dashed lines show estimates from the recursive identification with the factor of interest ordered last. The confidence intervals are those of the sign restriction approach. According to these plots, both approaches yield rather similar results. The most notable differences are the following. First, with the recursive identification IP growth turns negative earlier than what is implied by the sign restriction identification subsequent to shocks on all three yield curve factors. Second, the negative response of IP growth subsequent to a slope shock is more pronounced in the recursive identification scheme. Yet, the timing is similar to what has been found using the sign restriction approach which contradicts previous regression-based evidence that a decrease of the yield spread announces the beginning of a recession four to six quarters ahead. Finally, subsequent to a positive curvature shock the recursive method implies a less pronounced response of output growth and a persistent decline of inflation rates below zero.

to be close to zero on impact.

4.5.6 Impulse Responses to "Typical" Yield Curve Shocks

The last set of results relates to the identification of shocks discussed in Section 4.4.2 above. These shocks have been labeled "typical" in the sense that they explain as much as possible of the one-step ahead forecast error variance of the three yield curve components. Figure 4.2 shows that the one-step ahead forecast errors of level and slope and level and curvature are indeed visibly correlated. Accordingly, the first eigenvector explains about 70% of the variance of the model-implied yield curve innovations captured in Ω_{yy}. Figure 4.10 provides plots of impulse responses following the shock that is represented by this eigenvector. On impact, both level and curvature respond positively while the slope factor is driven downwards. Hence, longer and medium-term maturities are shifted up while short maturities fall. Together, these responses imply an initial steepening of the yield curve. As one would expect, this steepening is paralleled by a strongly significant increase of output growth. Rising for about four months after the shock occurs, the growth rate of industrial production then slowly reverts towards zero and eventually turns negative about one year later. This behavior is consistent with the subsequent dynamics of the yield curve components. In fact, after their strong initial reactions, level and the curvature decrease sharply while the slope factor rises. The yield curve thus becomes successively flatter. Interestingly, inflation does not significantly change subsequent to a typical yield curve surprise. This indicates that news about the future evolution of output might be more important for the dynamics of the yield curve than inflationary concerns.

Figure 4.11 reports for some macroeconomic variables and the 5-year yield the shares of variance explained by the three shocks corresponding to the eigenvectors of Ω_{yy}. By construction, the one-step ahead prediction error variance of the 5-year yield is entirely captured by these three shocks. The macroeconomic variables exhibit similar variance decompositions. Only small fractions of the forecast error variance are explained a few months ahead, but the shares increase strongly with the forecast horizon, attaining median levels of about 40% to 50% three years ahead. Altogether, about half of the long-run variation in different macroeconomic variables is explained by the shocks which summarize the variation of one-step ahead yield curve surprises. This shows that term structure news capture important information about the dynamics of macroeconomic aggregates in the long-run.

Figure 4.10: **Impulse Responses - "Typical" Yield Curve Shock**

This figure plots impulse responses of some selected macro variables and the three yield curve components to a "typical" curvature shock as defined in Section 4.4.2. Hence, by construction the shock explains the largest share of the one-step ahead forecast error variance of the three yield curve factors. Dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of impulse responses.

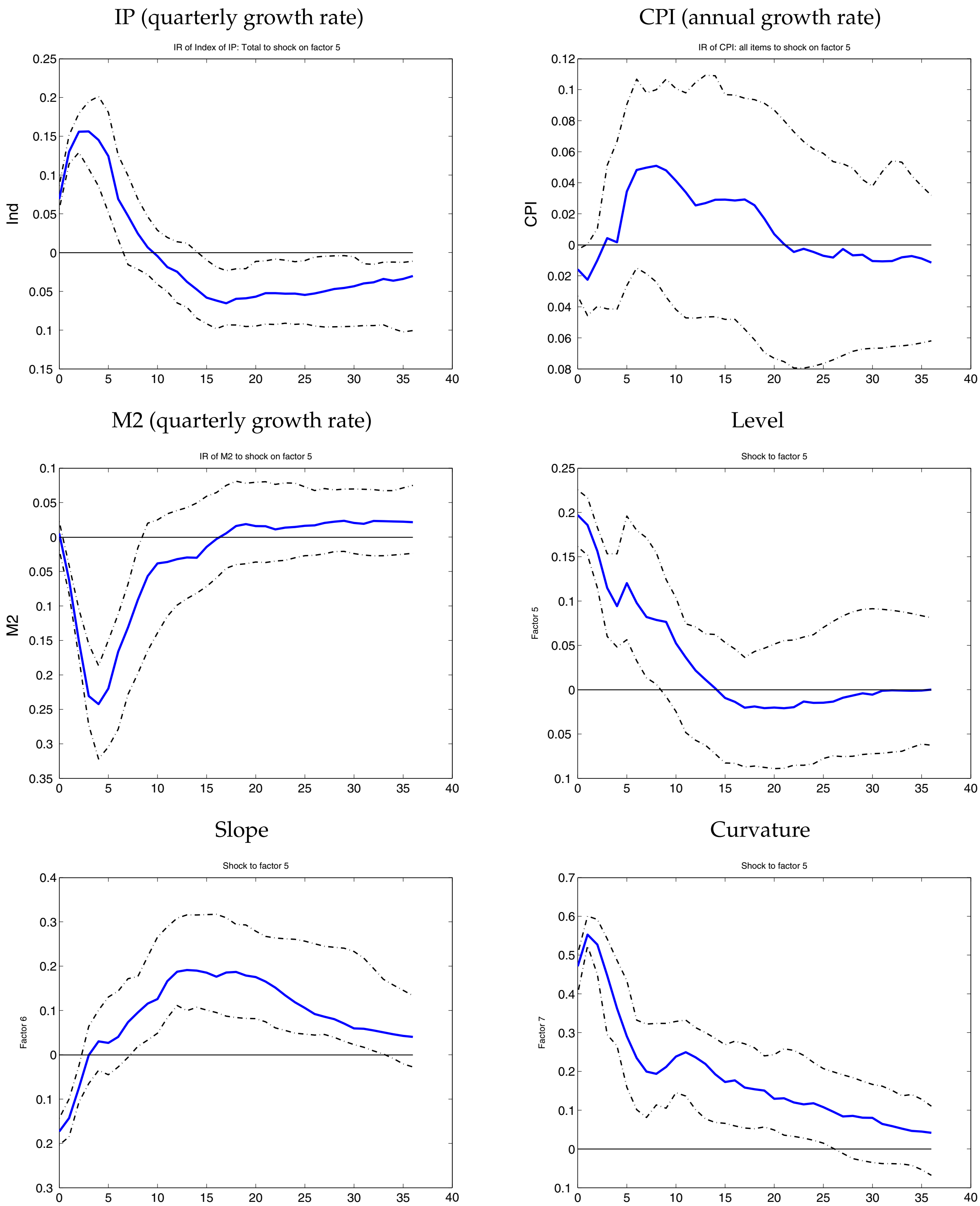
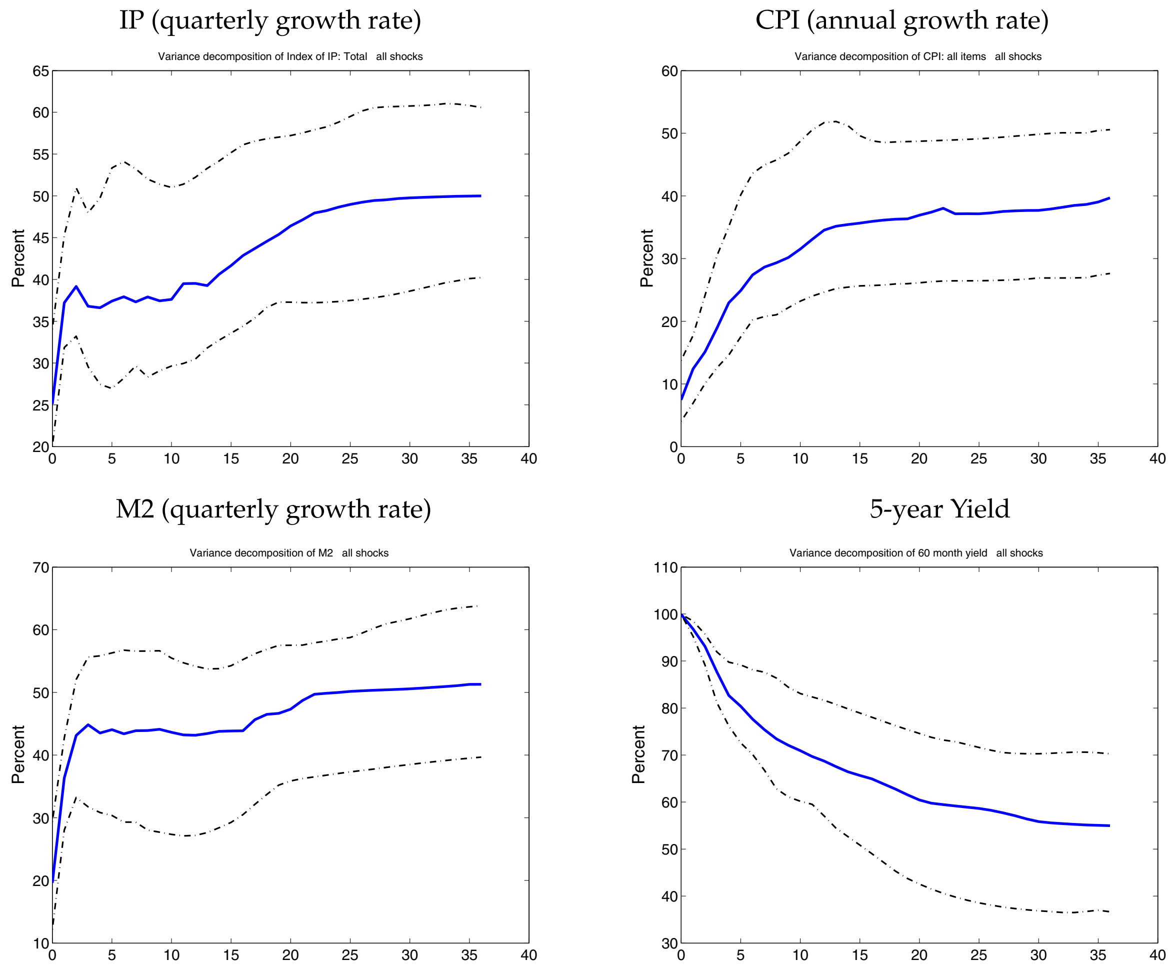


Figure 4.11: **Forecast Error Variance Decompositions**

This figure plots for some selected macro variables the shares of forecast error variance explained by the three shocks corresponding to the eigenvectors of Ω_{yy}. These shares can be interpreted as measures of how much information about the future evolution of macroeconomic variables is embodied in the yield curve. The solid line denotes the median estimate, while dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of variance decompositions.



4.6 Conclusion

The yield curve is known to convey information about the future course of the economy. Moreover, most of the yield curve variation can be described by the three factors level, slope, and curvature. Although interest rate spreads are commonly used as predictors of recessions, the informational content of innovations of the three yield curve components is largely unstudied. In this chapter, I have therefore systematically analyzed the predictive information carried by surprise changes of the yield curve level, slope, and curvature.

The main result of my study is surprising: unexpected changes of the curvature are quite informative about the future evolution of the term structure and the prospective dynamics of output. In particular, surprise surges of the curvature factor precede a pronounced flattening of the yield curve. At the same time, output growth follows a hump-shaped pattern, significantly falling below zero about one year after the shock. In contrast, unexpected increases of the slope factor - tantamount to diminishing yield spreads - are followed by an immediate but not very pronounced decline of output. According to these findings, a rising slope factor is associated with a future decline of output, but is announced by changes of the curvature factor. This qualifies curvature as a forward-looking indicator. The results obtained for the level factor are more consistent with conventional wisdom. In particular, surprise surges of the yield curve level are followed by a strong and persistent increase of inflation rates. All results remain qualitatively unchanged when different identification schemes are applied.

My findings suggest a few promising lines of future research. First and foremost, a regression analysis should be carried out to assess whether the curvature factor can indeed be used as a forward-looking indicator and how it compares to term spreads. Second, the model can be employed to forecast yields making use of the common dynamics of term structure and macroeconomic factors. In particular, it would be interesting to compare its forecast performance with that of models which exclusively employ term structure (Diebold and Li 2006) or macroeconomic dynamics (Mönch 2005, Chapter 3). Third, including the federal funds rate as an additional factor, one could disentangle expected and unexpected monetary policy actions and study their effects on macroeconomic variables.

A.4 Additional Tables and Figures

Figure 4.12: **Macro Factors - Different Starting Values**

This figure provides plots of the estimated macro factors obtained using five different sets of randomly generated starting values.

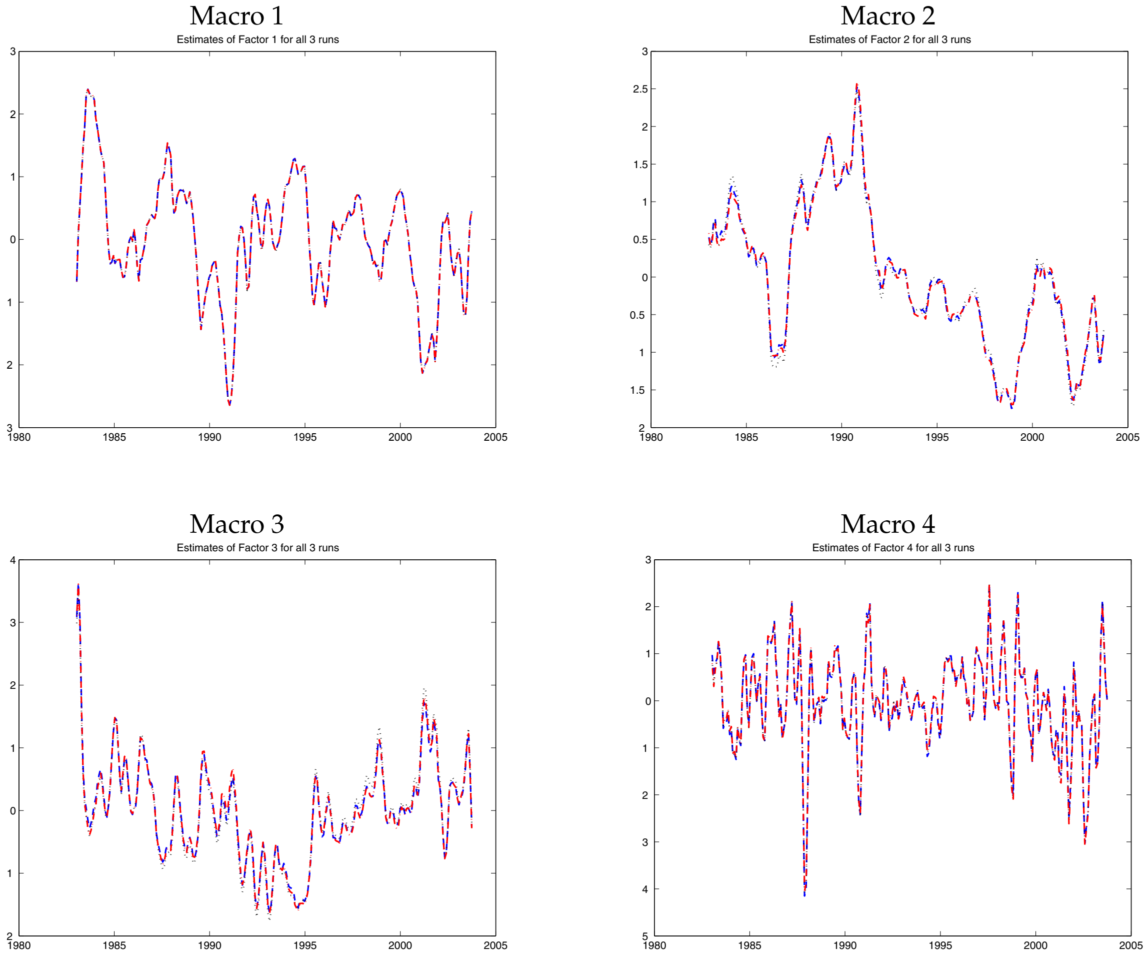


Figure 4.13: **Yield Factors - Different Starting Values**

This figure provides plots of the estimated yield factors obtained using five different sets of randomly generated starting values.

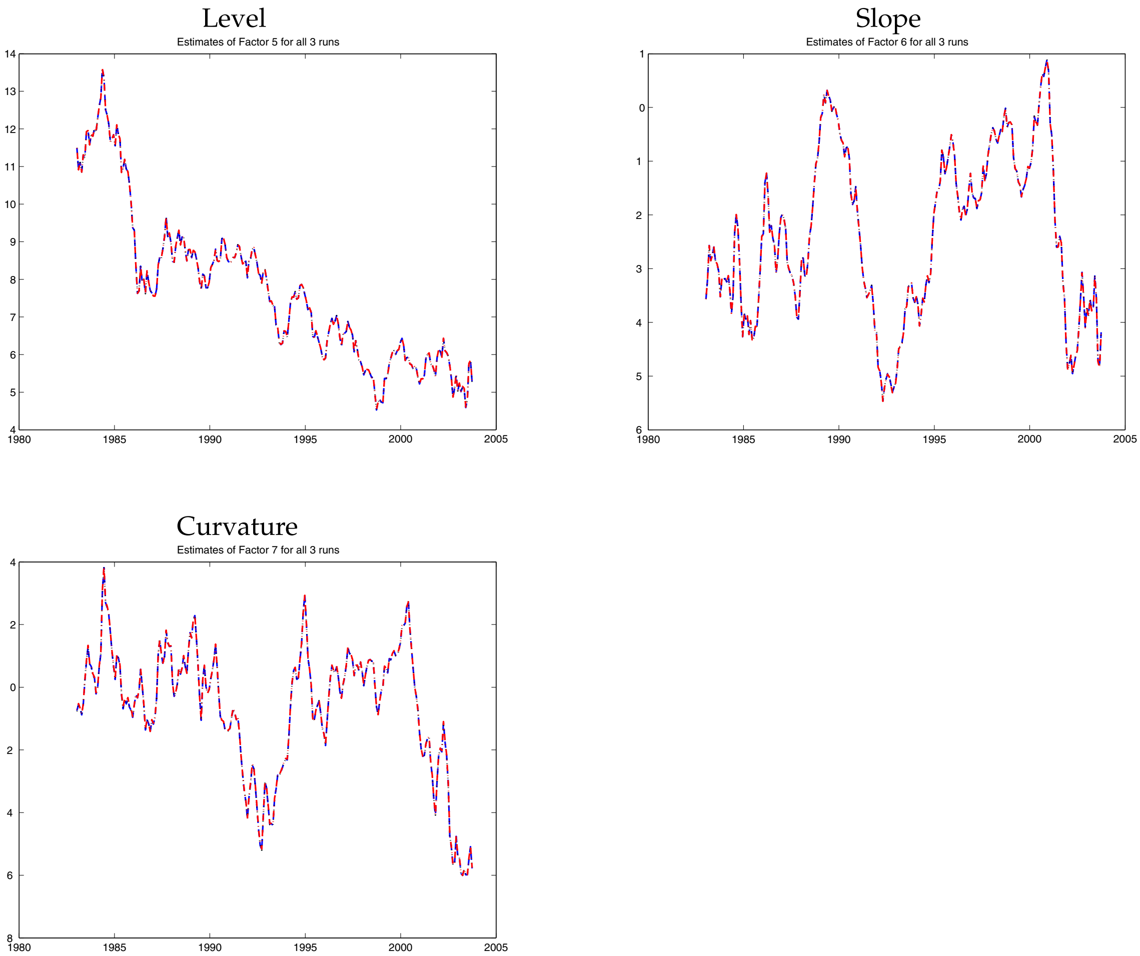


Figure 4.14: **Orthogonal Yield Surprises and Subsequent Yield Dynamics**

This figure plots impulse responses of the yield factors level (upper row), slope (middle row), and curvature (lower row) to orthogonal surprises identified with the "pure" sign restriction approach. Hence, in addition to imposing the sign and zero restrictions discussed above, the three identified shocks are required to be orthogonal. Dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of impulse responses.

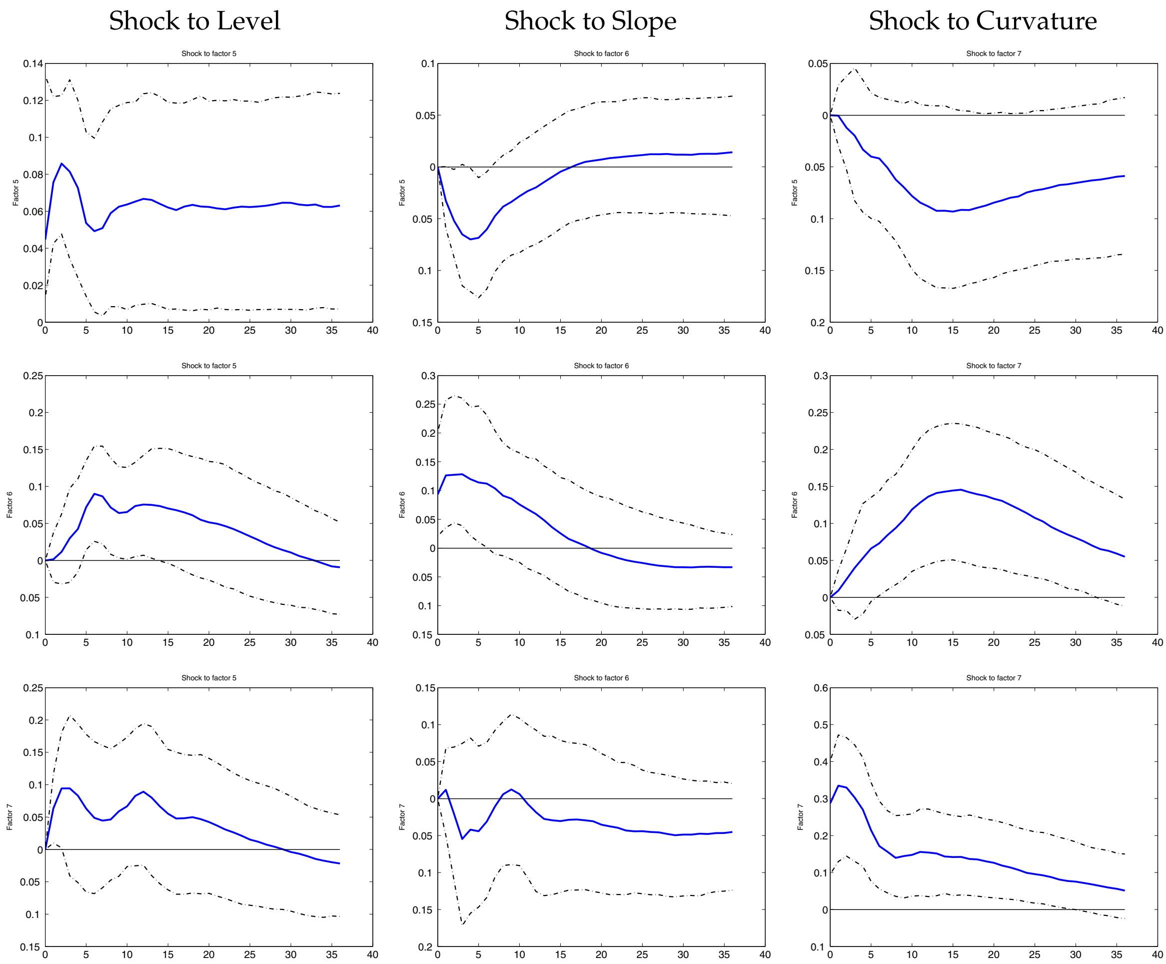


Figure 4.15: **Orthogonal Yield Surprises and Subsequent Macro Dynamics**

This figure plots impulse responses of the quarterly growth rate of industrial production (upper row), annual CPI inflation (middle row), and the quarterly growth rate of M2 (lower row) to orthogonal yield curve surprises. Dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of impulse responses.

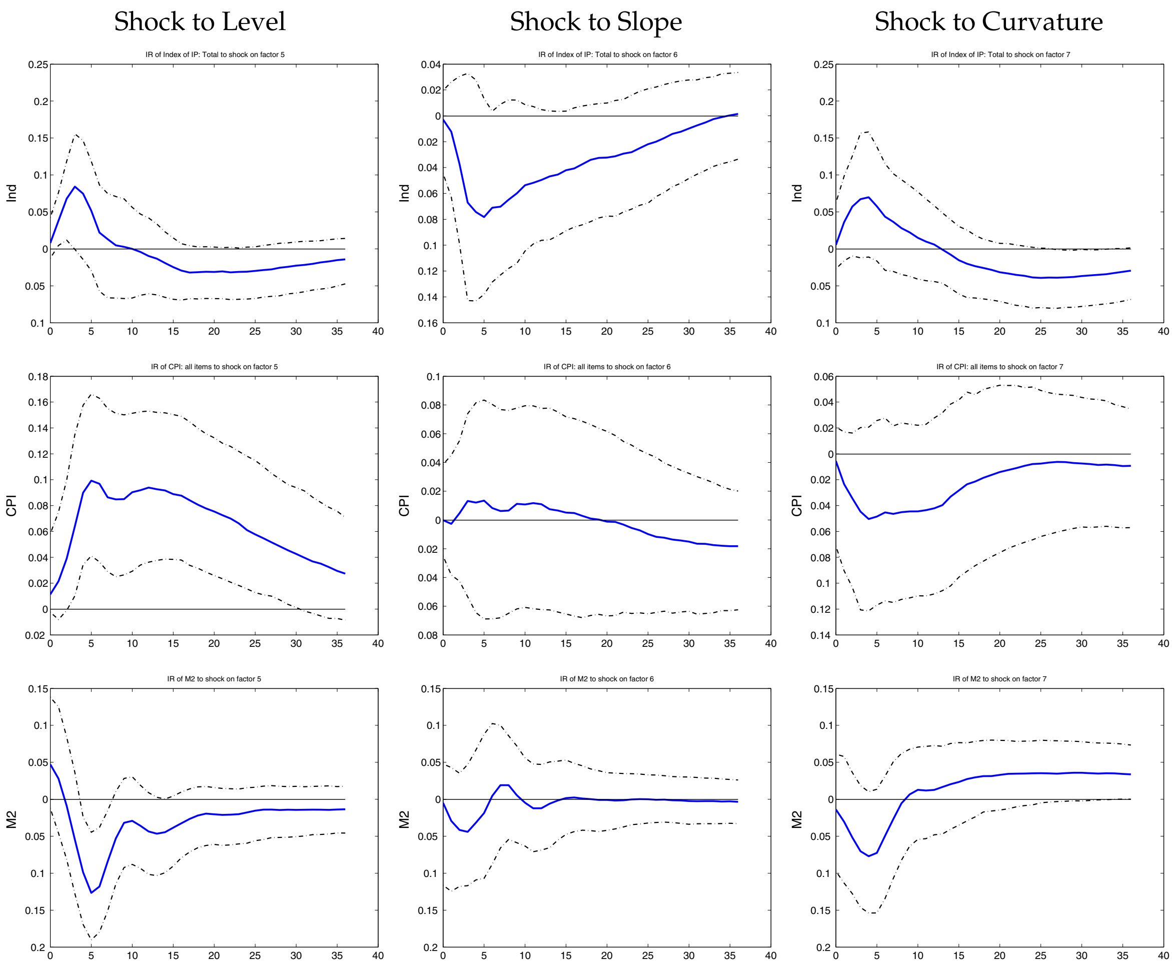


Figure 4.16: **Impulse Responses - "Strong" Level Surprise**

This figure plots impulse responses of some selected macro variables and the three yield curve components to a "strong" level surprise as defined in Section 4.4.1. Hence, the response of the level factor is restricted to be positive over the first six periods after the shock occurs while the responses of the slope and curvature factors are restricted to be small on impact. Dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of impulse responses.

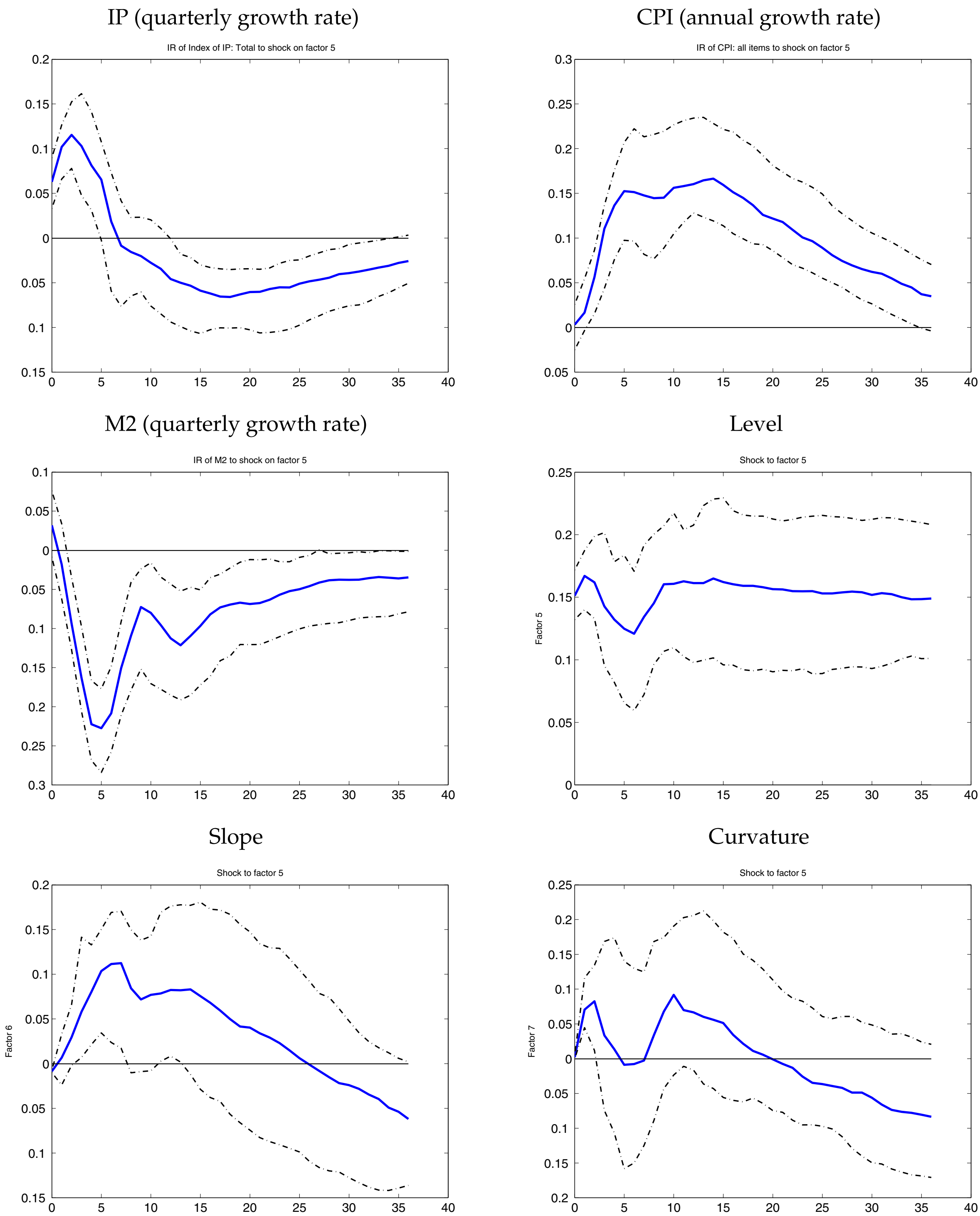


Figure 4.17: **Impulse Responses - "Strong" Slope Surprise**

This figure plots impulse responses of some selected macro variables and the three yield curve components to a "strong" slope surprise as defined in Section 4.4.1. Hence, the response of the slope factor is restricted to be positive over the first six periods after the shock occurs while the responses of the level and curvature factors are restricted to be small on impact. Dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of impulse responses.

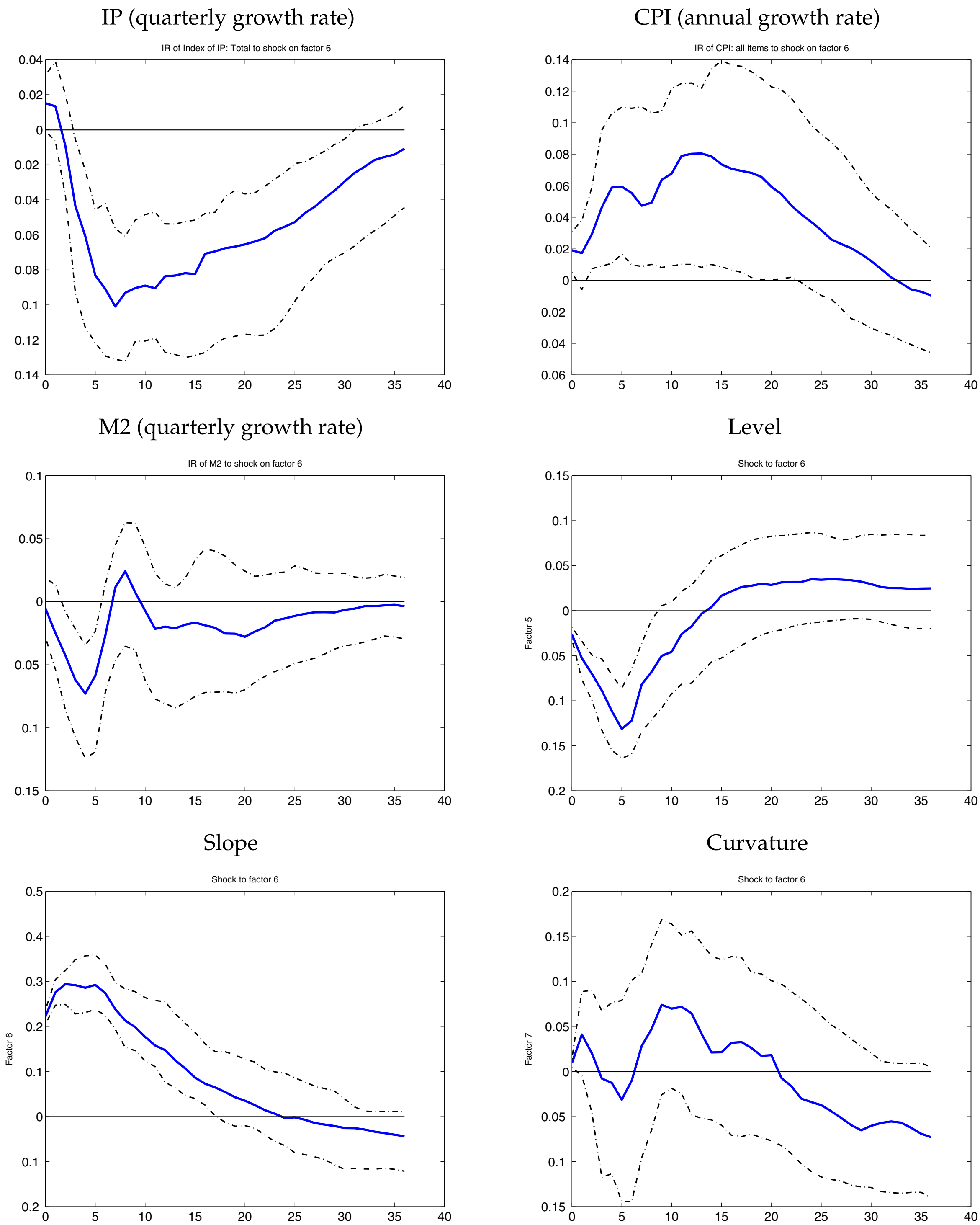


Figure 4.18: **Impulse Responses - “Strong” Curvature Surprise**

This figure plots impulse responses of some selected macro variables and the three yield curve components to a “strong” curvature surprise as defined in Section 4.4.1. Hence, the response of the curvature factor is restricted to be positive over the first six periods after the shock occurs while the responses of the level and slope factors are restricted to be small on impact. Dash-dotted lines indicate the 16% and the 84% quantiles of the posterior distribution of impulse responses.

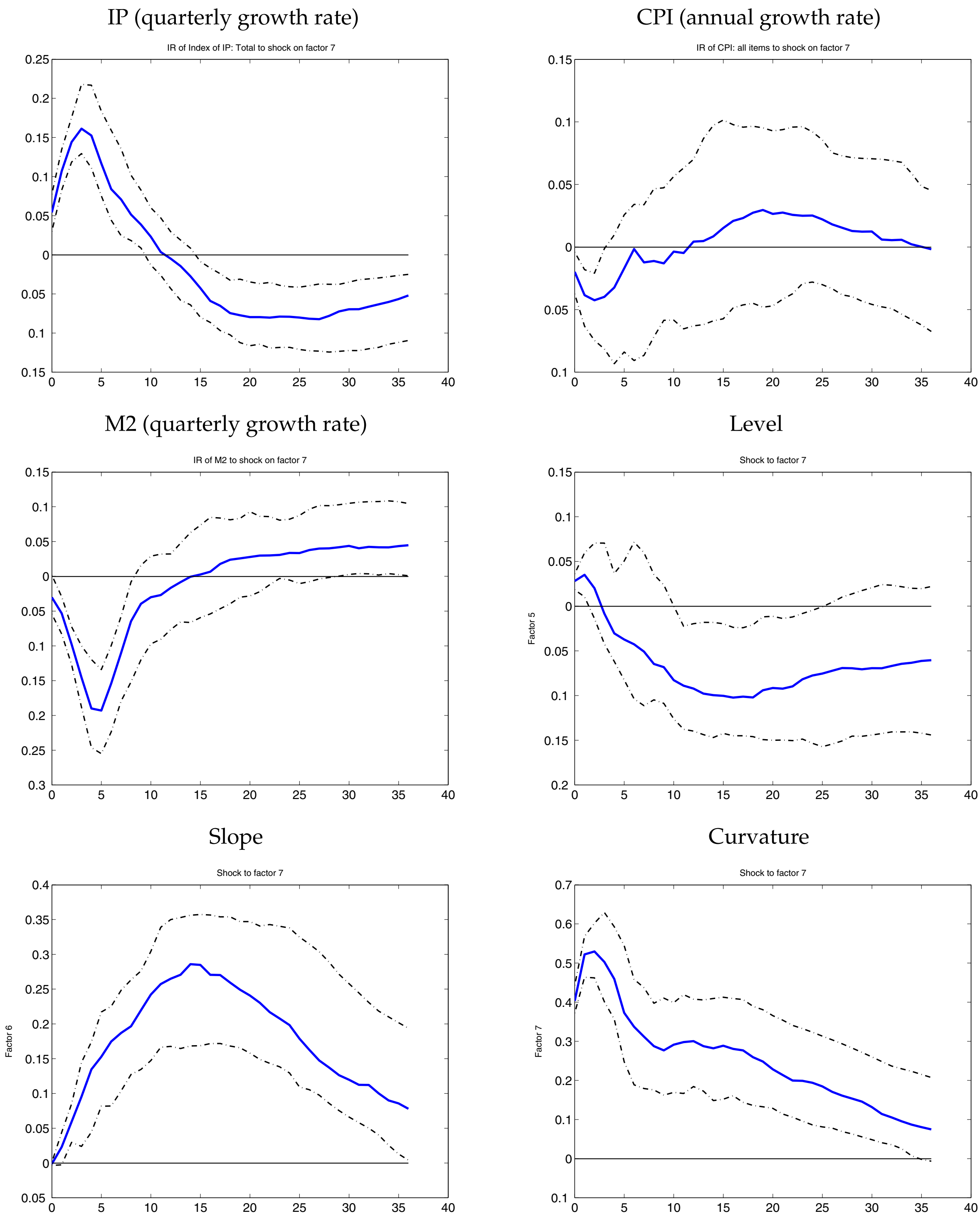
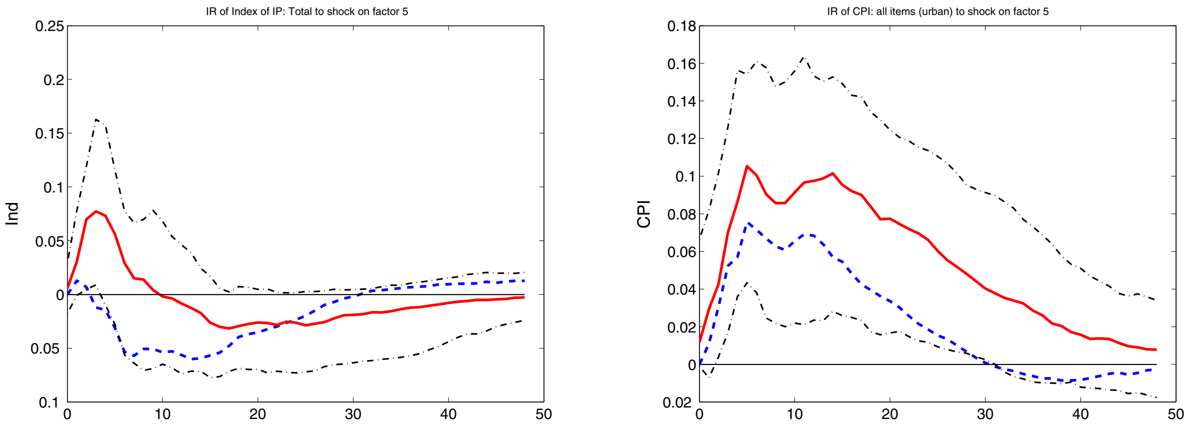


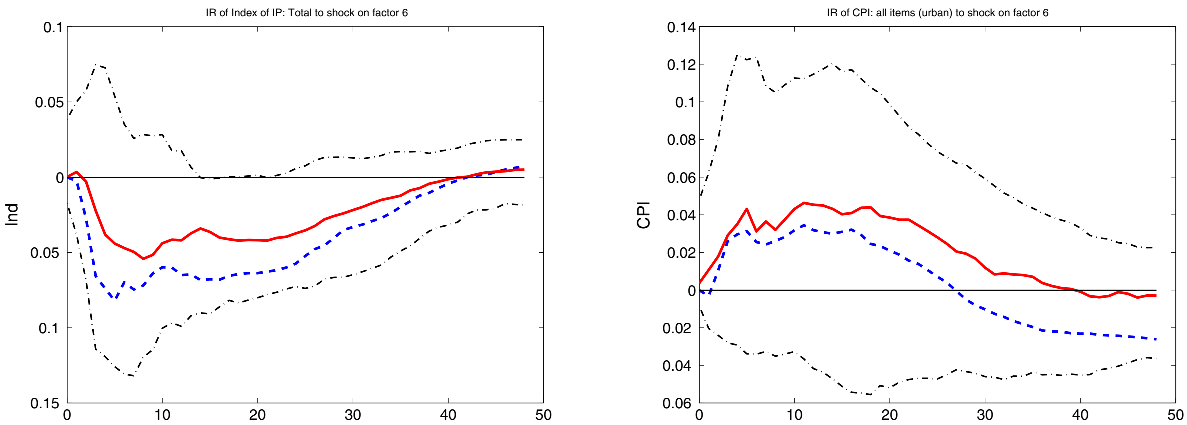
Figure 4.19: **Sign Restriction (solid) vs. Recursive Identification (dashed)**

This figure contrasts impulse responses of selected macro variables and the three yield curve components to shocks identified using different approaches. Red solid lines indicate median responses to "pure" yield curve surprises as defined above whereas blue (dashed) lines show median impulse responses obtained using a recursive identification scheme with the respective yield factor ordered last. Dash-dotted lines refer to the 16% and 84% quantiles of the posterior distribution of the sign restriction approach.

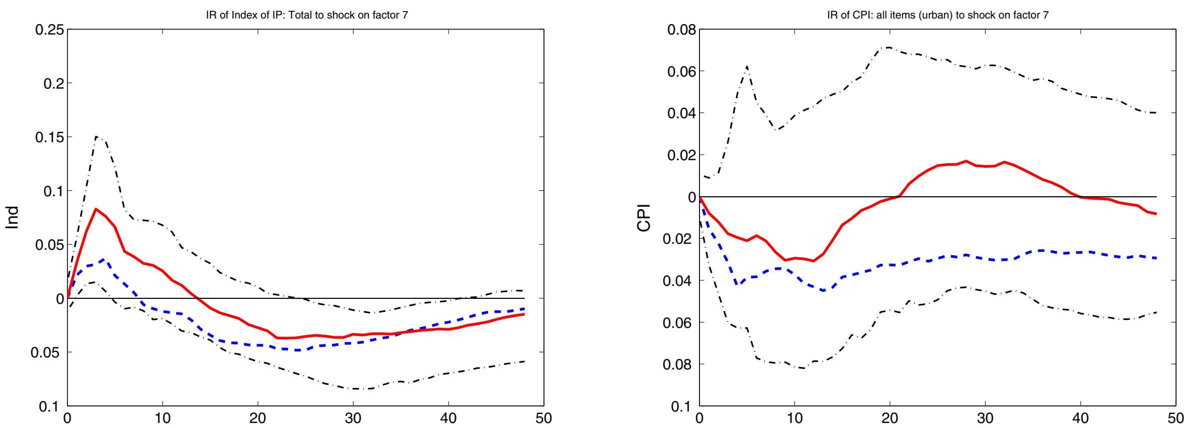
IP Level Shock CPI



IP Slope Shock CPI



IP Curvature Shock CPI



5 Towards a Monthly Business Cycle Chronology for the Euro Area

with Harald Uhlig

This chapter is an exercise in dating the Euro area business cycle on a monthly basis. Using a quite flexible interpolation routine, we construct several monthly series of Euro area real GDP, and then apply the Bry-Boschan (1971) procedure. To account for the asymmetry in growth regimes and duration across business cycle phases, we propose to extend this method with a combined amplitude/phase-length criterion ruling out expansionary phases that are short and flat. Applying the extended procedure to US and European data, we are able to replicate approximately the dating decisions of the NBER and the CEPR.

5.1 Introduction

Official dating of business cycles has a long tradition in the United States. The dates of peaks and troughs in the US economy's activity are officially announced by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER). According to the committee, a peak in activity determines the beginning of a recession which is defined as "*a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales*", see "The NBER's Business-Cycle Dating Procedure", Business Cycle Dating Committee, National Bureau of Economic Research, October 2003. In accordance with this definition, the Business Cycle Dating Committee is predominantly basing its judgment on the behavior of four monthly observable economic time series: total employment, real personal income less transfer payments, price-adjusted total sales of the manufacturing and wholesale-retail sectors, and industrial production. Since real GDP is only measured quarterly, it plays a minor role in the judgment of the Business Cycle Dating Committee. Information from other economic time series may also influence the decision of the committee, albeit with less weight.

Although the Business Cycle Dating Committee does not specify in more detail the method employed to date peaks and troughs, it seems to be following the traditional NBER view of business cycle behavior as described in Burns and Mitchell [1946]. Their approach of measuring business cycles consisted in first identifying turning points in a variety of individual economic time series which usually tend to cluster around certain dates. In a second step, reference cycle dates for aggregate economic activity were selected from within these clusters on the basis of different criteria as, for example, bounds on the length and amplitude of business cycles.

With the creation of the Euro area on January 1st, 1999 and a single currency in circulation as of January 1st, 2002, it has become of greater urgency to establish such an official tradition in Europe as well. Therefore, the Centre for Economic Policy Research (CEPR) has recently formed a committee to set the dates of the Euro area business cycle in a manner similar to the NBER. Taking into account the particular features of the Euro area as a group of national economies, the Committee defines a recession *as a significant decline in the level of economic activity, spread across the economy of the euro area, usually visible in two or more consecutive quarters of negative growth in GDP, employment and other measures of aggregate economic activity for the euro area as a whole, and reflecting similar developments in most countries*, see "Business Cycle Dating Committee of the Centre for Economic Policy Research", CEPR, September 2003. To make sure that expansions or recessions are widespread over the countries of the area, the CEPR bases its judgment on euro area aggregate statistics as well as country statistics. Further, the committee has decided to date the Euro area business cycle in terms of quarters rather than months, arguing that the most reliable European data for dating purposes are available only on a quarterly basis. However, being well aware of the scarcity of appropriate historical monthly time series for most of the European countries, we think that it is nevertheless useful to establish a monthly business cycle chronology also for the Euro area. In fact, if the figures in some quarterly time series are viewed as the average or sum of the three consecutive months in a quarter, then dating on the quarterly level amounts to identifying turning points in a filtered monthly series. Monthly and quarterly dating of the same underlying monthly series might therefore lead to different results. Hence, dating business cycles at the monthly level is likely to provide a more precise information about the exact turning points than quarterly dating. Furthermore, since the state of the economy is an important variable in empirical models, applications are conceivable which would require knowledge about the business cycle turning points of the Euro area on a monthly basis. Thus, applying the highest diligence in interpreting the available data, this chapter aims at filling the gap of a monthly business cycle chronology for the Euro area.

To arrive at such a chronology, two difficulties must be overcome. First, rather than examining a plethora of data for each of the months of the last 30 years, an econometric methodology needs to be found which successfully finds the NBER dates, and then apply that methodology to European data. Second, appropriate Euro area data needs to be found. For solving both of these difficulties, we can build on existing research.

For an econometric methodology, we build on the research which has tried to reverse-engineer a time-series based methodology replicating the dates chosen by the NBER. The methodology by Bry and Boschan [1971] is generally considered to be quite successful at that. We will show that this is indeed the case in Section 5.3.1, although with some caveats: the Bry-Boschan procedure sometimes finds the exact NBER date, but sometimes only comes close to the official dates within a few months. Furthermore, the Bry-Boschan treats business cycle expansions and contractions symmetrically, thus not taking into account differences in terms of growth and duration across regimes. As a consequence, the procedure may identify business cycle phases that are implausibly flat. To avoid this, we therefore propose to augment the Bry-Boschan procedure with a suitable amplitude/phase-length criterion in Section 5.2.1, ruling out business cycle expansions that are both short and flat.

To use the Bry-Boschan procedure, one needs a monthly time series for real GDP. Even for the US, such a time series is not officially available, although one can construct a pretty good time series with the help of an interpolation procedure which is described in detail in appendix 6.5.1. We have done so for the exercise in Section 5.3.1 and discuss the resulting series in Appendix 6.5.2.

For the Euro area, building a good monthly real GDP time series is more difficult than for the US for a number of reasons. First, quarterly real GDP for the Euro area has only been recorded officially as of January 1991. Since our aim is to determine the Euro area business cycle turning points for at least the last 30 years, we have proceeded to construct a Euro area monthly real GDP series by interpolating and then aggregating appropriate country time series.[40] Even there, data availability is a serious problem. The details on available data and our construction are provided in appendix 6.5.3. To check the dating results obtained using our series we have additionally determined the turning points of two different monthly interpolations of the Euro area quarterly real GDP series constructed by Fagan, Henry, and Mestre [2001]. For all three series, the results are very similar, see Section 5.3.2 for a comparison.

Section 5.4 finally provides a summary of the challenges in improving on this exercise, discusses limitations and provides the key conclusions.

5.2 Bry-Boschan's Method and an Extension

Bry and Boschan [1971] provide a nonparametric, intuitive and easily implementable algorithm to determine peaks and troughs in individual time series. Although the method is quite commonly used in the literature, we briefly sketch its main constituents here. For a detailed description, the reader is referred to Bry and Boschan's paper. The procedure consists of six sequential steps. First, on the basis of some well-specified criterion, extreme observations are identified and replaced by corrected values. Second, troughs (peaks) are determined for a 12-month moving average of the original series as observations whose values are lower (higher) than those of the five preceding and the five following months. In case two or more consecutive troughs (peaks) are found, only the lowest (highest) is retained. Third, after computing some weighted moving average, the highest and lowest points on this curve in the ± 5 months-neighborhood of the before determined peaks and troughs are selected. If they verify some phase length criteria and the alternation of peaks and troughs, these are chosen as the intermedi-

[40] Eurostat has recently launched a project whose aim is the construction of a historical monthly time series for Euro area real GDP. However, this series has not yet been made officially available. Moreover, since their approach seems to differ somewhat from ours in terms of methodology, it would be interesting to compare the time series propoerties of both series.

ate turning points. Fourth, the same procedure is repeated using an unweighted short-term moving average of the original series. Finally, in the neighborhood of these intermediate turning points, troughs and peaks are determined in the unsmoothed time series. If these pass a set of duration and amplitude restrictions, they are selected as the final turning points.

5.2.1 A Simple Combined Amplitude/Phase-Length Criterion for the Bry-Boschan Procedure

Obviously, as a univariate procedure the Bry-Boschan turning point selection method is unsuited to take into account information from more than one time series as is done by the business cycle dating committees of the NBER and the CEPR. Despite this shortcoming, we would like to stick to the Bry-Boschan algorithm instead of using a multivariate methodology since it is both intuitive and transparent. In its original form, the method incorporates a minimum cycle and phase length criterion, restricting business cycles and phases to last at least 15 and 5 months, respectively. Turning points corresponding to cycles or phases that do not fulfil these criteria are simply deleted. As we will see further below, with the minimum phase length criterion switched off, the Bry-Boschan procedure identifies two recessions in Euro area real GDP in the early 1980s. On the contrary, with the minimum phase length criterion switched on it censors the shorter of the two downturns without taking into account that there has been only a brief period (19 months) of very moderate growth (annualized growth rate of less than 1.4%) in between the two phases of declining GDP. Considering also information from other economic indicators, the CEPR has defined the period starting in the first quarter of 1980 and ending in the third quarter of 1982 as a long recession, see 5.3.2. During the same time, US monthly real GDP has shown a similar behavior, first falling shortly from January to June 1980, then rising until August 1981 and declining again until February 1982. Yet, both recessions and the intermediate upturn were more pronounced in the US than in the Euro area. For example, US real GDP grew at an annual rate of 3.3 % in between the trough in June 1980 and the peak in August 1981. Accordingly, the NBER has dated two distinct recessions interrupted by a short intermediate upturn.

It is our view that the different patterns that real GDP followed in the US and in the Euro area in the early 1980s suffice to explain the dating decisions of the

NBER and the CEPR, without having to take into account further measures of activity. We therefore would like to augment the univariate Bry-Boschan procedure with a combined amplitude/phase-length criterion that embraces both types of pattern. Such a rule should ensure that business cycle phases that are both short and flat are suppressed while phases that are short but pronounced are retained. Hence, it remains to appropriately define what is "short" and what is "flat" in the present context, and whether the criterion shall apply to business cycle expansions or contractions.[41]

As has already been noted above, the original Bry-Boschan procedure provides for a minimum phase-length criterion of five months, i.e. once the turning points in the time series to be dated are determined, business cycle expansions or contractions that are shorter than five months are deleted, independently of their amplitude. Notice, however, that due to the widely documented asymmetry of business cycles that is associated with much longer booms than recessions, this criterion in practice exclusively applies to business cycle contractions. Having studied the time series behavior of GDP for different countries, it is our view that episodes shorter than five months occur which can be classified as business cycle contractions without any doubt. In contrast, the length of expansions seems to be a more distinctive feature of business cycles at least in the postwar period. In fact, there is a comprehensive literature on the stabilization of business cycles in the US and other industrialized countries in the postwar period (see, e.g., Diebold and Rudebusch 1992 and Romer 1994). Our reading of this literature is that there is widespread agreement that business cycle expansions have been significantly longer in the postwar than in the prewar period, while it is not so clear that business cycle contractions have become shorter over time. We therefore base our combined amplitude/phase-length criterion that is designed to date postwar data on the growth rate and length of expansions rather than contractions.

Given this decision, it appears intuitive to designate a "short" business cycle ex-

[41] Artis, Kontolemis, and Osborn [1997] suggest a turning point selection procedure similar to the Bry-Boschan algorithm which incorporates a minimum amplitude criterion. According to their criterion, phases (peak to trough or trough to peak) are excluded that have an amplitude of less than one standard deviation of log changes of the series to be dated. This rule is obviously aimed at use for rather volatile series such as industrial production which Artis et al. [1997] employed for their dating exercise. However, applied to our (comparatively smooth) monthly real GDP series for the US and the Euro area, it did not yield the desired exclusion of flat expansions.

pansion as one that is significantly shorter than the average expansion. We therefore define a short expansion as one whose length is outside the one-standard deviation interval around the average expansion length. Based on the official NBER business cycle dates, the average length of expansions in the US has been 57 months in the postwar period, with a standard deviation of 36 months. Given these numbers, the threshold below which a business cycle upturn would be defined as short according to the above criterion is thus 21 months or 7 quarters.[42]

By similar reasoning we define a "flat" expansion as an upturn in which the annualized growth rate is significantly lower than the average positive annual growth rate, i.e. which is outside the one-standard deviation interval around the average positive annual growth rate.[43] Computing this indicator for the US, we find a value of 2.1 %, whereas for the Euro area it amounts to 1.5 %. In order not to make our rule excessively restrictive, we take the lower of both values as our threshold for minimum annual growth in a short business cycle upturn. Altogether, our combined amplitude/phase-length criterion thus excludes expansions that are not longer than 21 months *and* during which the annualized growth rate is lower than 1.5 %. In practice, applying this criterion amounts to deleting the trough and peak which mark the beginning and the end of a short and flat expansion, respectively, in the ultimate step of the Bry-Boschan procedure.

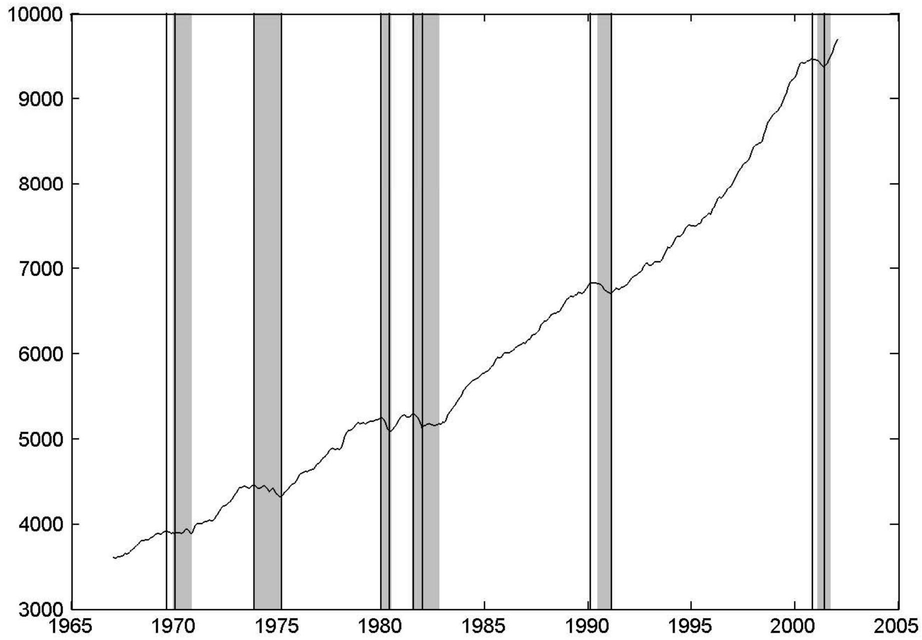
It might be worth noting that Artis, Marcellino, and Proietti [2003] make a similar point. These authors discuss the usefulness of amplitude restrictions as a censoring device for dating algorithms. Analogously to our reasoning, they conclude that since expansions are usually longer and characterized by a lower average drift rate than recessions, different threshold values for amplitude restrictions should be used for booms and contractions. However, they do not investigate this issue further and do not provide such a phase-dependent amplitude rule.

[42] Obviously, there is some arbitrariness in this choice. Using European data or a longer time span of US data might have led to a slightly different threshold. Yet, given the well-documented business cycle stabilization after world war II and the close correspondence between US and Euro area business cycle characteristics (see Agresti and Mojon 2001), this choice appears by all means appropriate.

[43] We restrict this indicator to positive annual growth rates since including contractions would obviously result in a biased threshold for low growth during expansions.

Figure 5.1: **Official NBER Dates and Bry-Boschan Dates for the US**

This plot allows a visual comparison of official NBER dates and the turning points found by the Bry-Boschan algorithm. The recessions identified by the NBER are indicated by shaded areas, the peaks and troughs determined by the Bry-Boschan procedure by vertical bold lines.



5.3 Monthly Business Cycle Chronologies for the US and the Euro Area

In this section, we apply the original and augmented Bry-Boschan algorithm to our monthly real GDP series for the US and the Euro area and compare the results.

5.3.1 The US Dates

As a first check on our procedure and for comparison, we apply the programmed turning point selection algorithm to US data. To that end, we construct a monthly time series for real US GDP for the period 1967:1 to 2002:09 (see appendix 6.5.1 for details on the interpolation method) to which we then apply the Bry-Boschan procedure as well as our augmented version of it.

The results can be seen in a "birds eye view" in Figure 5.1. The NBER recessions

Table 5.1: **NBER and Bry-Boschan Dates for the US**

This table compares business cycle turning points defined by the NBER with those obtained using the Bry-Boschan algorithm based on our series of monthly real GDP for the US.

Peaks:						
Bry-Boschan	69M8	73M11	80M1	81M8	90M3	00M12
Augmented Bry-Boschan	69M8	73M11	80M1	81M8	90M3	00M12
NBER	69M12	73M11	80M1	81M7	90M7	01M3
Troughs:						
Bry-Boschan	70M1	75M3	80M6	82M1	91M3	01M7
Augmented Bry-Boschan	70M1	75M3	80M6	82M1	91M3	01M7
NBER	70M11	75M3	80M7	82M11	91M3	01M11

have been indicated by shaded areas, whereas the peaks and troughs determined by the Bry-Boschan procedure are shown as vertical bold lines. As expected, the dating results for the US do not change by augmenting the Bry-Boschan procedure with our amplitude/phase-length criterion since the short recovery in between the two recessions in the early 1980s was rather pronounced.

A comparison of the dates is given in Table 5.1. Note that the Bry-Boschan procedure sometimes finds the exact NBER date, but sometimes only comes close to the official dates within a few months. Further, for those dates that do not coincide, the Bry-Boschan dates tend to lead the NBER dates slightly, the only exception being the peak in July 1981. Employment is one of the main four monthly time series the Business Cycle Dating Committee of the NBER bases its judgement on. Since employment is known to lag output, this might partly explain the slight lead in the Bry-Boschan dates. Notice further that when the moving average window parameter in the first step of the procedure is set to twelve months as in Bry and Boschan [1971], the procedure misses two complete business cycles towards the beginning and the end of the sample. However, since our objective has been to come as close as possible to the NBER dates, we have set the window parameter for the pre-smoothing to eight months. The business cycle dates we propose for the Euro area have been obtained using the same setting.

5.3.2 A Monthly Business Cycle Chronology for the Euro Area

The term "Euro area" in this chapter refers to the area of the 12 member countries of the European monetary union as of January 1st, 2002, including in particular Greece and Eastern Germany. As already noted above, we perform our business cycle dating exercise on different monthly time series for Euro area real GDP. The construction of these series is briefly sketched in Section 5.3.2, and in more detail in appendix 6.5.3. In Section 5.3.2 we present the dating results obtained by applying the Bry-Boschan turning point selection procedure to these series, and compare them with the quarterly dates obtained by other authors and those recently published by the CEPR. We discuss the monthly business cycle dates taking into consideration further aggregate measures of Euro area business activity in Section 5.3.2 and finally apply the Bry-Boschan procedure augmented with our amplitude/phase-length criterion in Section 5.3.2.

Monthly GDP Series for the Euro area

For our business cycle dating experiment, we use three different time series for monthly Euro area real GDP. Our benchmark series is our own series for the period 1970:1 to 2002:12. Although the details about the construction of this series are provided in appendix 6.5.3, we shall briefly outline the main steps here. First, we have constructed monthly time series for GDP volume for all twelve Euro area member countries from interpolating appropriate quarterly and annual time series. For each country separately, we choose the "best" interpolation procedure among a set of possible specifications of a general model which nests some of the most commonly used interpolation methods such as the ones suggested by e.g. Chow and Lin [1971], Fernandez [1981], or Mitchell, Smith, Weale, Wright, and Salazar [2005]. The general model treats monthly figures of real GDP as the unobserved component in a state-space model, employing the observation equation to ensure that quarterly (annual) figures are the averages of three (twelve) consecutive monthly observations. We use the Kalman filter to estimate the model and maximum likelihood ratio tests to select the best specification. As related variables, we employ monthly series for industrial production, real retail sales, employment or exports, depending on availability, see Table 6.2. Finally, we aggregate these series to obtain a measure of Euro area real GDP using the same aggregation method and weights as Fagan et al. [2001], FHM in short, in their latest update of the ECB's area wide model dataset.

Table 5.2: **Turning Points for Different Monthly Series of Euro Area Real GDP**

This table compares turning points identified by the Bry-Boschan algorithm when applied to our monthly series of Euro area GDP, a linear interpolation of the quarterly FHM series, and a monthly interpolation of the FHM series constructed using a chained volume index of aggregate Euro area industrial production as related series.

Peaks:				
Our series	74M8(QIII)	80M3(QI)	82M4(QII)	92M2(QI)
FHM IP	74M8(QIII)	80M3(QI)	82M4(QII)	92M2(QI)
FHM lin	74M8(QIII)	80M2(QI)	82M5(QII)	92M2(QI)
CEPR	74QIII	80QI		92QI
Troughs:				
Our series	75M4(QII)	80M9(QIII)	82M7(QIII)	93M1(QI)
FHM IP	75M1(QI)	80M9(QIII)	82M8(QIII)	93M4(QII)
FHM lin	75M2(QI)	80M8(QIII)	82M8(QIII)	93M2(QI)
CEPR	75QI		82QIII	93QIII

The other two series are based on interpolations of the quarterly real GDP series constructed by FHM which has recently been updated. The first is a linear interpolation, viewing the quarterly data as referring to the middle of the three months in a quarter. The second has been constructed by interpolating the quarterly FHM series employing the interpolation method described in appendix 6.5.1. As related series, we have used an aggregate monthly chained volume index series for Euro area industrial production, which we have constructed using the same weights and aggregation method as FHM for their area-wide model dataset.[44]

The Euro Area Dates

Applying the original Bry-Boschan procedure, we obtain the results listed in Table 5.2. As can be seen from the dating results, the original programmed turning point selection procedure finds four business cycles for all three series.[45] A

[44] Since there is no such series for Ireland covering the entire sample period, we have omitted Ireland from this aggregate.

[45] Notice that the minimum phase length criterion included in the original Bry-Boschan procedure has been set off here. In case this criterion is put on, the third cycle is censored for all three series since the corresponding recession is always shorter than 5 months.

visual "birds-eye" view of the dates obtained for our own monthly time series is provided in Figure 5.2. Concerning the exact dates of the identified turning points, there is a surprising agreement between the three series: three out of four peaks found in our series coincide exactly with those obtained from the monthly interpolation of the FHM series using aggregate industrial production as related variable, and the dates of the third peak only differ by one month. Further, in terms of quarterly business cycle peaks, those dates are fully consistent with the ones obtained using the linear interpolation of the FHM series. There are only slight differences when the identified business cycle troughs are concerned, the maximum deviation between our series and the instrumental variable interpolation of the FHM series being three months. For the first and the fourth trough, however, this deviation results in a different quarterly turning point.

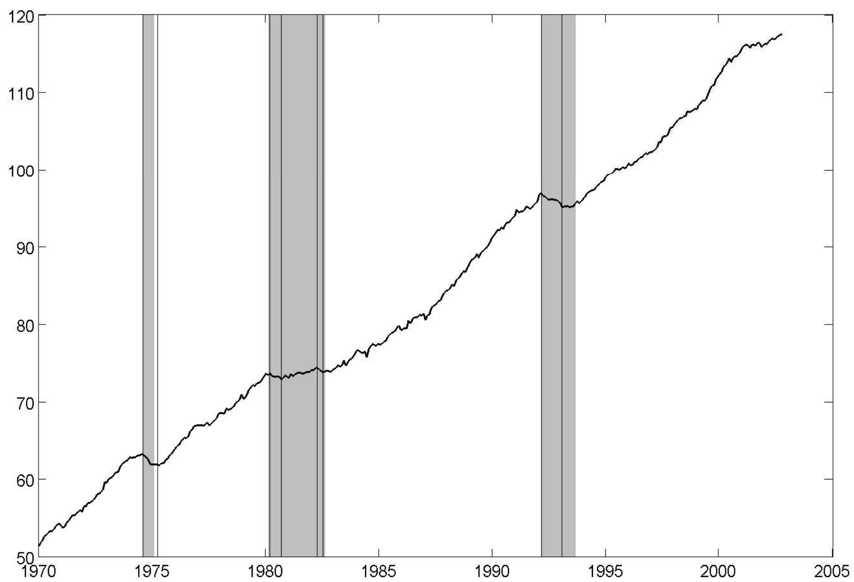
The quarterly turning points can be compared with the dating results obtained by other authors and the turning points recently provided by the CEPR. Let us begin the comparison by considering first the findings of other authors. Krolzig [2001] employs a univariate Markov-switching model for Euro area quarterly GDP growth using the time series constructed by Beyer, Doornik, and Hendry [2001] which covers the post-1979 period. Over that time span, he identifies two business cycles with peaks in 1980QI and 1992QI, and troughs in 1981QI and 1993QI, respectively. Hence, for the identified cycles, there is a close correspondence with our results, the only difference being the trough in 1980. Interestingly, however, Krolzig's (2001) procedure indicates that the Euro area has experienced only one complete cycle in the 1980s.[46] As will be discussed in Section 5.3.2 below, we come to the same conclusion by studying the time series behavior of further business-cycle related variables.

Employing a business cycle dating method called "ABCD approach", Anas and Ferrara [2004] determine business cycle turning points for the Euro area using Eurostat's aggregate GDP series starting in 1995 and own backward calculations before. They find their method to deliver similar results as e.g. Krolzig [2001] and Anas, Billio, Ferrara, and Duca [2003]. The latter paper, using the same method-

[46] Note that Krolzig [2001] finds very similar results using a multivariate Markov-switching model for GDP growth rates of eight Euro area member countries. Notice further that Anas and Ferrara [2002] find a very recent business cycle peak in 2001QI by extending the univariate analysis in Krolzig [2001] up to 2002QII.

Figure 5.2: **Real GDP and Business Cycle Turning Points for the Euro Area**

This figure shows turning points for the Euro area business cycle based on our monthly series for real Euro area GDP. The recessions identified by the CEPR are indicated by shaded areas, the peaks and troughs determined by the original Bry-Boschan procedure by vertical bold lines. The quarterly CEPR dates have been interpreted as monthly turning points by taking the middle month of the respective quarter as the monthly date. Notice further that the five month minimum phase length rule in the original Bry-Boschan algorithm has been set off here.



ology and time series as Anas and Ferrara [2004], identifies four business cycles over the 1970-2003 period. Although they do not correspond exactly, the quarterly turning points found by Anas et al. [2003] are rather similar to the ones identified in this chapter, differing by one quarter at the most.

Applying a quarterly version of the Bry-Boschan procedure to the previous release of the quarterly FHM series, Harding and Pagan (2001b) and Artis et al. [2003] both obtain slightly different results as we do using interpolations of the latest update of the FHM series.[47] However, applying the Bry-Boschan algorithm to the linear interpolation of the previous version of the quarterly FHM series,

[47] The latest update of the ECB's area-wide model database has been made available in November 2003 and differs from the previous one in a number of respects : the inclusion of Greece, new availability of data including ESA95 data, revisions to historical data and the interpolation of quarterly historical data using a methodology similar to the one employed here.

we obtain exactly the same dates as Harding and Pagan (2001b) and Artis et al. [2003]. In fact, dating the old version of the FHM series results in business cycle troughs in 1981Q1 instead of 1980QIII and 1982QIV instead of 1982QIII.[48] This difference emphasizes the importance which exerts the construction method of the employed time series on the dating result. Moreover, the fact that the latest update of the FHM series exhibits turning points which are much more similar to the ones obtained using our series than those of the previous FHM release, clearly underscores the usefulness of our series as a measure of monthly Euro area real GDP.

Examining the Individual Dates

According to the dating results discussed so far, all measures of Euro area GDP that are available to us seem to support the view that the Euro area has experienced four cycles since 1970. Interestingly, however, the CEPR has only identified three business cycles over the same period, considering the short cycle in the early 1980s as a long recession, see Table 5.2. In its inaugural release, the business cycle dating committee of the CEPR notes:

The third recession, in the 1980s, exhibits different and specific characteristics. The recession in terms of aggregate output is milder but longer. GDP does not decline sharply but rather stagnates for almost three years. Our dating is thus based on the behaviour of employment and investment which, unlike GDP, declined sharply during the period. In this episode, we also observe more heterogeneity in output dynamics across the three large economies than in the other two recessions. That affects our designation of the date of the trough, in particular.

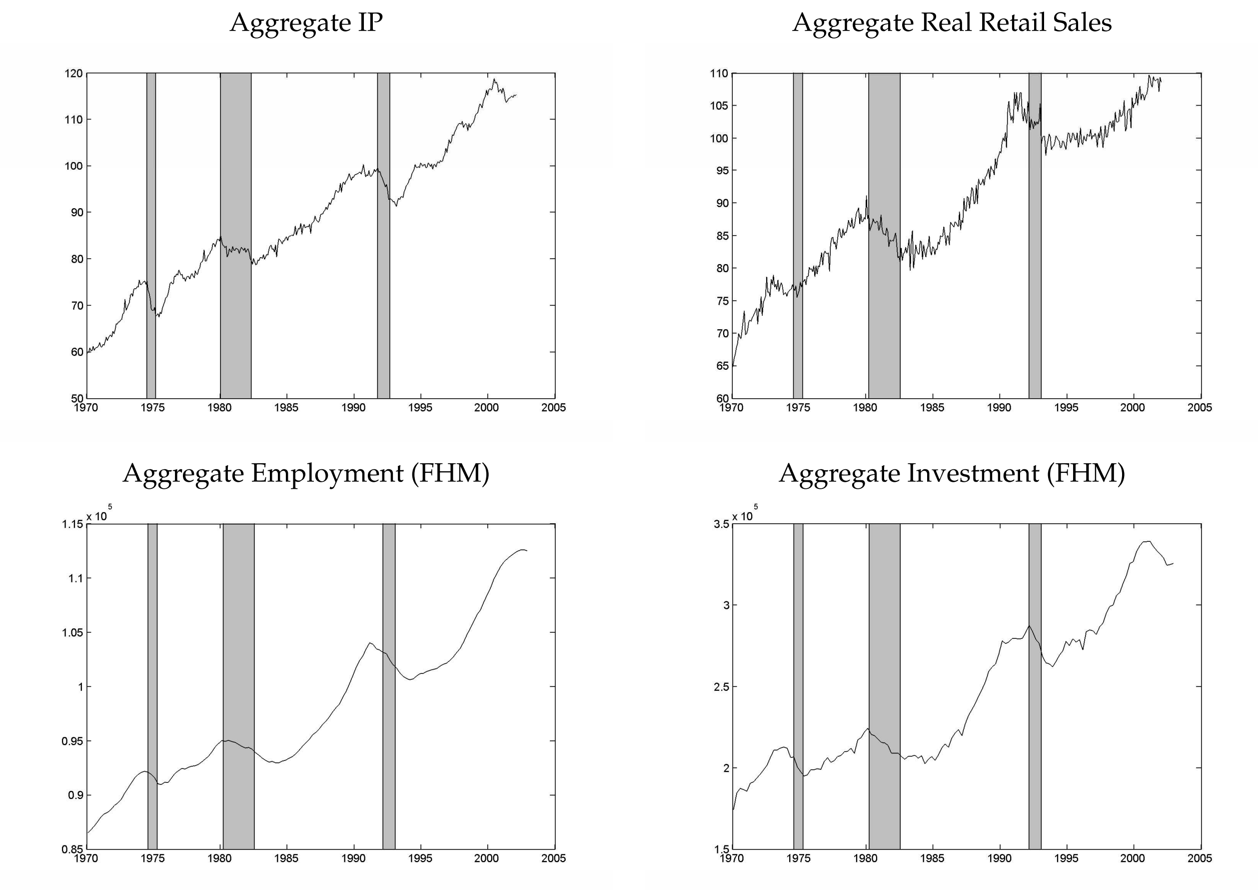
To assess whether there have been one or two cycles in the 1980s, we therefore follow the business cycle dating committee by examining further relevant time series. Figure 5.3 provides plots of Euro area aggregates for industrial production, real retail sales, employment, and investment, the latter two being linear interpolations of the quarterly series constructed by Fagan et al. [2001] for the ECB's area wide model.[49] Eye-ball checking is sufficient to see that Euro area

[48] It may be worth noting that Artis et al. [2003] find the same turning points for the post-1979 period, employing the quarterly real GDP series for the Euro area provided by Beyer et al. [2001] which only covers the post-EMU period.

[49] Notice that due to the limited data availability, the aggregate IP series does not include Ireland. The aggregate retail sales series is constructed using data from Belgium, Finland, Germany, Greece, Ireland, and the Netherlands. The aggregation method is the same as the one that has

Figure 5.3: **Related Monthly Indicators for the Euro Area**

This figure plots the aggregate related monthly series for the Euro area. The shaded areas indicate the recession periods identified by the Bry-Boschan procedure based on our monthly series of aggregate GDP. The IP series is obtained from aggregating monthly IP series for all Euro area member countries except Ireland. The aggregate retail sales series is constructed using data from Belgium, Finland, Germany, Greece, Ireland, and the Netherlands. The aggregation method is the same as the one that has been used to construct the GDP series.



employment, investment, and retail sales clearly have exhibited one pronounced cycle in the 1980s instead of two short cycles. The aggregate IP series shows a slightly less clear-cut behavior, declining sharply from March 1980 to September 1980, remaining almost constant until April 1982, and then falling again sharply. A central feature of business cycles is the common movement of different measures of economic activity. Given that three such variables in the Euro area clearly exhibit only one cycle in the 1980s instead of two, and that industrial production

been used to construct the GDP series.

does not regain its pre-March 1980 level until 1985, it thus appears appropriate to consider the period between early 1980 and mid 1982 as a long recession even though GDP has recovered slightly in between these dates.[50]

Interestingly, although Euro area industrial production and investment clearly have experienced a peak towards the end of 2000 (see Table 5.3), the Bry-Boschan procedure does not identify a business cycle peak in real GDP around that time. Indeed, all three monthly time series of Euro area real GDP remain more or less constant in 2001 and start rising again in early 2002. Accordingly, the short-term moving averages of the respective series are rather flat, hence explaining why the Bry-Boschan procedure does not identify a turning point. Thus further data observations will have to be awaited before it can doubtlessly be decided whether there has been a business cycle peak in Euro area real GDP around 2001.

Applying the Augmented Bry-Boschan Method

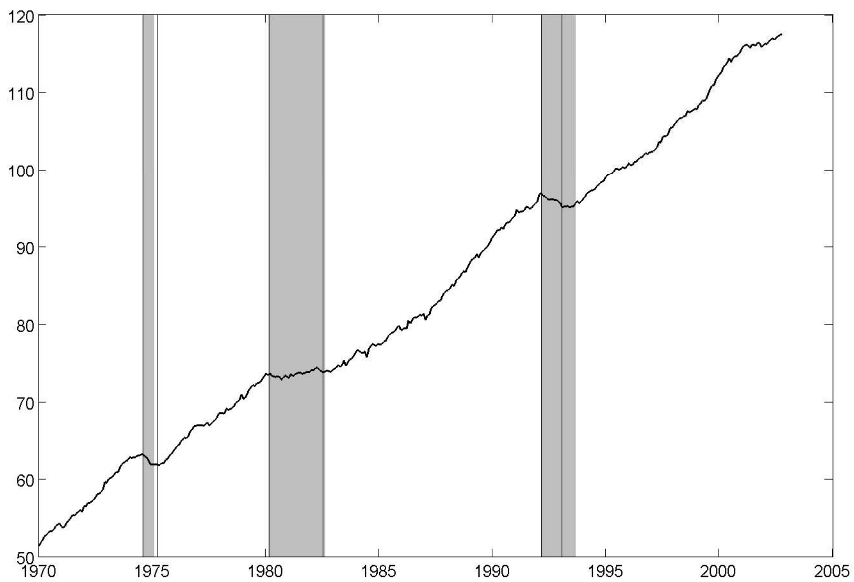
To assess whether our combined amplitude/phase-length criterion is able to identify this feature of the data, we now apply the augmented Bry-Boschan procedure to the three monthly time series of Euro area real GDP. A visual representation of the outcome of this exercise is provided in Figure 5.4, while the corresponding business cycle chronologies are stated in Table 5.3. As can be seen from these results, the extended algorithm identifies the two short recessions in the 1980s connected by a very brief and moderate upturn as a long recession, and thus matches very closely the dating decision of the CEPR.

Rather than dating the Euro area business cycle, some recent studies have focussed on the European business cycle, thus also considering countries that are not member of the European Monetary Union, as for example the UK. Applying a multivariate Markov-Switching model to quarterly GDP growth rates of six European countries including the UK, Krolzig and Toro [2002] identify three cycles over the 1970-1996 period, with business cycle peaks in 1974QI, 1980QI, and 1992QII, and troughs in 1975QII, 1982QIV and 1993QII, respectively. These dates are rather similar to our findings when we apply the amplitude/phase-

[50] According to our measure of monthly GDP for the Euro area, output grew only about 2.2 % in between the two peaks identified by the Bry-Boschan procedure in September 1980 and April 1982. This corresponds to an annual rate of less than 1.4 % which appears unusually low for a business cycle upturn. During the same period, the quarterly FHM series grew about 1.45 % corresponding to an annual rate of less than 1%.

Figure 5.4: **Monthly Real GDP and Turning Points for the Euro Area**

This figure shows the turning points of the Euro area business cycle based on our monthly series for Euro area real GDP. The Bry-Boschan algorithm has been augmented with the combined amplitude/phase-length criterion discussed above. The recessions identified by the CEPR are indicated by shaded areas, the peaks and troughs determined by the Bry-Boschan procedure by vertical bold lines. The quarterly CEPR dates have been interpreted as monthly turning points by taking the middle month of the respective quarter as the monthly date.



length criterion. This might indicate that UK GDP has experienced a more pronounced downturn in the early 1980s whereas Euro area member countries such as Belgium and the Netherlands that are excluded from Krolzig and Toro's (2002) dataset, contributed to the long and flat expansion distinctive for Euro area GDP in the early 1980s.

5.4 Conclusion

We have performed an exercise in dating the business cycle in the Euro area from 1970 to 2002 on a monthly basis. We construct several monthly European real GDP series, and then apply the Bry-Boschan (1971) procedure. The Bry-Boschan procedure comprises a censoring rule which treats business cycle expansions and contractions symmetrically without taking into account the differences in average drift rate and duration across regimes. To overcome this shortcoming, we

Table 5.3: **CEPR Dates and Modified Bry-Boschan Dates for the Euro Area**

This table summarizes the turning points identified by the augmented Bry-Boschan algorithm when applied to our monthly series of Euro area GDP. Further, the quarterly turning points determined by the CEPR are provided.

Peaks:			
Our series	74M8(QIII)	80M3(QI)	92M2(QI)
FHM IP	74M8(QIII)	80M3(QI)	92M2(QI)
FHM lin	74M8(QIII)	80M2(QI)	92M2(QI)
CEPR	74QIII	80QI	92QI
Troughs:			
Our series	75M4(QII)	82M7(QIII)	93M1(QI)
FHM IP	75M1(QI)	82M8(QIII)	93M4(QII)
FHM lin	75M2(QI)	82M8(QIII)	93M2(QI)
CEPR	75QI	82QIII	93QIII

propose a combined amplitude/phase-length criterion for the Bry-Boschan procedure that rules out expansionary phases which are short and flat.

For US data, we show that this procedure comes close to replicating the official NBER dates. For European data, a number of additional issues needed to be resolved. In particular, a monthly real GDP series had to be constructed, to which to apply the Bry-Boschan procedure. We have constructed such a series by first interpolating quarterly and annual GDP series for individual countries, using different monthly available variables as instruments. In a second step, we have aggregated the individual interpolated series to obtain a monthly real GDP series for the Euro area.

As a cross-check on the dating results obtained using our series of monthly Euro area real GDP, we have constructed two alternative series. We find a surprising agreement between the dating results obtained from the three different series. However, since our benchmark series has been constructed using information contained in a number of different monthly available instruments, we think this series reflects the monthly variation of business activity in the Euro area most appropriately.

The original Bry-Boschan dating procedure has identified four peaks and four troughs over the period 1970 to 2002, see Table 5.2. Yet the two contractions and the interjacent expansion identified in the early 1980s are not very pronounced. We have therefore examined other measures of business activity in order to assess whether the Euro area has experienced one or two cycles in that period. As all these series do exhibit only one complete cycle during that time, we consider the period of very low GDP growth in the early 1980s as a long recession. Applying the Bry-Boschan procedure augmented with our combined amplitude/phase-length criterion to the different monthly GDP series for Euro area, we are able to replicate this feature and match the turning point decision of the CEPR quite closely.

It is important to keep in mind that the Bry-Boschan procedure cannot detect peaks and troughs very close to the beginning or the end of the sample. In particular, the procedure may have missed turning points in the Euro area since mid 2002. However, the methodology applied in this chapter can be easily used to determine more recent turning points of the Euro area business cycle when new data becomes available. Moreover, the flexibility of our approach to construct a monthly time series of real GDP for the Euro area makes it readily applicable in case of future enlargements of the European Monetary Union.

Bibliography

Anna-Maria Agresti and Benoît Mojon. Some stylised facts on the euro area business cycle. Working Paper Series 95, European Central Bank, 2001.

Omar Aguilar and Mike West. Bayesian dynamic factor models and portfolio allocation. *Journal of Business & Economic Statistics*, 18(3):338–357, July 2000.

Jaques Anas and Laurent Ferrara. A comparative assessment of parametric and non parametric turning points detection methods: The case of the euro-zone economy. Discussion paper, Centre d'Observation Economique, Paris, 2002.

Jaques Anas and Laurent Ferrara. Detecting cyclical turning points: The abcd approach and two probabilistic indicators. *Journal of Business Cycle Measurement and Analysis*, 1(2):193–225, 2004.

Jaques Anas, Monica Billio, Laurent Ferrara, and Marco Lo Duca. A turning point chronology for the euro-zone classical and growth cycle. Eurostat working paper, Paper presented at the 4th Colloquium on Modern Tools for Business Cycle Analysis, Luxembourg, November 2003.

Donald W.K. Andrews. Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4):821–856, July 1993.

Andrew Ang and Monika Piazzesi. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–787, May 2003.

Andrew Ang, Monika Piazzesi, and Min Wei. What does the yield curve tell us about gdp growth? *Journal of Econometrics*, 127(1-2):359–403, 2006.

Kevin Aretz, Söhnke M. Bartram, and Peter F. Pope. Macroeconomic risks and the fama and french/carhart model. mimeo, Lancaster University, 2006.

Michael J. Artis, Zenon G. Kontolemis, and Denise R. Osborn. Business cycles for g7 and european countries. *Journal of Business*, 70(2):249–279, April 1997.

Michael J. Artis, Massimiliano Marcellino, and Tommaso Proietti. Dating the euro area business cycle. CEPR Discussion Paper 3696, Centre for Economic Policy Research, January 2003.

David Backus, Silverio Foresi, and Chris Telmer. Discrete-time models of bond pricing. NBER Working Paper 6736, National Bureau of Economic Research, 1998.

Jushan Bai. Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics*, 122(1):137–183, September 2004.

Jushan Bai and Serena Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, January 2002.

Jushan Bai and Serena Ng. Confidence intervals for diffusion index forecasts and inference for factor augmented regressions. *Econometrica*, 74(4):1133–1150, 2006.

Rolf W. Banz. The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9:3–18, 1981.

S. Basu. Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *Journal of Finance*, 32 (3):663–682, June 1977.

Luc Bauwens, Michel Lubrano, and Jean-François Richard. *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, 1999.

Francesco Belviso and Fabio Milani. Structural factor-augmented var (sfavar) and the effects of monetary policy. mimeo, Princeton University, March 2005.

Ben Bernanke, Mark Gertler, and Mark Watson. Systematic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity*, 1:91–157, 1997.

Ben Bernanke, Jean Boivin, and Piotr S. Eliasz. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. *The Quarterly Journal of Economics*, 120(1):387–422, January 2005.

Ben S. Bernanke and Jean Boivin. Monetary policy in a data-rich environment. *Journal of Monetary Economics*, 50(3):525–546, April 2003.

Andreas Beyer, Jurgen A. Doornik, and David F. Hendry. Constructing historical euro-zone data. *Economic Journal*, 111(469):F102–F121, February 2001.

Fischer Black. Capital market equilibrium with restricted borrowing. *Journal of Business*, 45(3):444–455, July 1972.

Robert R. Bliss. Testing term structure estimation methods. *Advances in Futures and Options Research*, 9:197–231, 1997.

Douglas T. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3): 265–296, September 1979.

Gerhard Bry and Charlotte Boschan. Cyclical analysis of economic time series: Selected procedures and computer programs. NBER Technical Working Paper 20, National Bureau of Economic Research, 1971.

Arthur F. Burns and Wesley C. Mitchell. *Measuring Business Cycles*. National Bureau of Economic Research, New York, 1946.

John Y. Campbell. Understanding risk and return. *Journal of Political Economy*, 104 (2):298–345, April 1996.

C.K. Carter and Robert Kohn. On gibbs sampling for state space models. *Biometrika*, 81:541–553, 1994.

Gary Chamberlain and Michael Rothschild. Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, 51(5):1281–1304, September 1983.

Nai-Fu Chen, Richard Roll, and Stephen A. Ross. Economic forces and the stock market. *Journal of Business*, 59(3):383–403, July 1986.

Ren-Raw Chen and Louis Scott. Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates. *Journal of Fixed Income*, 3:14–31, 1993.

Gregory C. Chow and An-loh Lin. Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4):372–375, November 1971.

John H. Cochrane. *Asset Pricing*. Princeton University Press, NJ, June 2001.

John H. Cochrane. A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy*, 104(3):572–621, June 1996.

Timothy Cogley, Sergei Morozov, and Thomas J. Sargent. Bayesian fan charts for u.k. inflation: Forecasting and sources of uncertainty in an evolving monetary system. CFS Working Paper Series 2003/44, Center for Financial Studies, February 2003.

Gregory Connor and Robert A. Korajczyk. Performance measurement with the arbitrage pricing theory. *Journal of Financial Economics*, 15:373–394, 1986.

John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–408, 1985.

Qiang Dai and Thomas Philippon. Fiscal policy and the term structure of interest rates. NBER Working Paper 11574, National Bureau of Economic Research, August 2005.

Qiang Dai and Kenneth J. Singleton. Specification analysis of affine term structure models. *Journal of Finance*, 55(5):1943–1978, October 2000.

Qiang Dai and Kenneth J. Singleton. Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics*, 63 (3):415–441, March 2002.

Russell Davidson and James G. MacKinnon. *Estimation and Inference in Econometrics*. Oxford University Press, New York, 1993.

Hans Dewachter and Marco Lyrio. Macro factors and the term structure of interest rates. *Journal of Money, Credit, and Banking*, 38(1):119–140, 2006.

Tommaso Di Fonzo. Temporal disaggregation of economic time series: Towards a dynamic extension. Working papers and studies, European Commision, 2003.

Francis X. Diebold and Canlin Li. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 127(1-2):337–364, February 2006.

Francis X. Diebold and Glenn D. Rudebusch. Have postwar economic fluctuations been stabilized? *American Economic Review*, 82(4):993–1005, September 1992.

Francis X. Diebold, Monika Piazzesi, and Glenn Rudebusch. Modeling bond yields in finance and macroeconomics. *American Economic Review Papers and Proceedings*, 95(2):415–420, May 2005.

Francis X. Diebold, Glenn D. Rudebusch, and Boragan S. Aruoba. The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, 127(1-2):309–338, 2006.

Gregory R. Duffee. Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57(1):405–443, 02 2002.

Donald Duffie and Rui Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.

Piotr Eliasz. Likelihood based inference in large dynamic factor models using gibbs-sampling. unpublished manuscript, Princeton University, 2002.

Robert F. Engle and Mark W. Watson. A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association*, 76(376): 774–781, 1981.

Arturo Estrella and Gikas A. Hardouvelis. The term structure as a predictor of real economic activity. *Journal of Finance*, 46(2):555–576, June 1991.

Arturo Estrella and Frederic S. Mishkin. Financial variables as leading indicators predicting u.s. recessions:. *The Review of Economics and Statistics*, 80(1):45–61, February 1998.

Charles L. Evans and David Marshall. Economic determinants of the nominal treasury yield curve. mimeo, Federal Reserve Bank of Chicago, 2004.

Gabriel Fagan and Jérôme Henry. Long run money demand in the eu: Evidence for area-wide aggregates. *Empirical Economics*, 23(3):483–506, 1998.

Gabriel Fagan, Jérôme Henry, and Ricardo Mestre. An area-wide model (awm) for the euro area. Working Paper Series 42, European Central Bank, January 2001.

Eugene F. Fama. Efficient capital markets: Ii. *Journal of Finance*, 46(5):1575–1617, December 1991.

Eugene F. Fama and Kenneth R. French. The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465, June 1992.

Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, February 1993.

Eugene F. Fama and James D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, May-June 1973.

Carlo A. Favero, Massimiliano Marcellino, and Francesca Neglia. Principal components at work: The empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics*, 20(5):603–620, 2005.

Roque B. Fernandez. A methodological note on the estimation of time series. *The Review of Economics and Statistics*, 63(3):471–476, August 1981.

Wayne E. Ferson and Campbell R. Harvey. Conditioning variables and the cross section of stock returns. *Journal of Finance*, 54(4):1325–1360, 08 1999.

Mario Forni and Lucrezia Reichlin. Let's get real: A factor analytical approach to disaggregated business cycle dynamics. *Review of Economic Studies*, 65(3): 453–473, July 1998.

Mario Forni, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics*, 82(4):540–554, November 2000.

Mario Forni, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association*, 100:830–840, 2005.

John Geweke. The dynamic factor analysis of economic time series. In D.J. Aigner and A.S. Goldberger, editors, *Latent Variables in Socio-Economic Models*. North-Holland, Amsterdam, 1977.

John Geweke and Guofu Zhou. Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies*, 9(2):557–587, 1996.

Eric Ghysels. On stable factor structures in the pricing of risk: Do time-varying betas help or hurt? *Journal of Finance*, 53(2):549–573, 1998.

Domenico Giannone, Lucrezia Reichlin, and Luca Sala. Monetary policy in real time. In Mark Gertler and Kenneth Rogoff, editors, *NBER Macroeconomics Annual 2004*, NBER Macroeconomics Annual, pages 161–225. The MIT Press, 2004.

Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054, July 1982.

Lars Peter Hansen and Ravi Jagannathan. Assessing specification errors in stochastic discount factor models. *Journal of Finance*, 52(2):557–590, June 1997.

David Heath, Robert Jarrow, and Andrew Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105, January 1992.

Thomas S.Y. Ho and Sang-bin Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41(5):1011–1029, December 1986.

Robert J. Hodrick and Xiaoyan Zhang. Evaluating the specification errors of asset pricing models. *Journal of Financial Economics*, 62(2):327–376, November 2001.

Peter Hördahl, Oreste Tristani, and David Vestin. A joint econometric model of macroeconomic and term-structure dynamics. *Journal of Econometrics*, 127(1-2): 405–444, 2006.

Ravi Jagannathan and Zhenyu Wang. The conditional capm and the cross-section of expected returns. *Journal of Finance*, 51(1):3–53, March 1996.

Christopher S. Jones. Extracting factors from heteroskedastic asset returns. *Journal of Financial Economics*, 62:293–325, 2001.

Alessandro Justiniano. Estimation and model selection in dynamic factor analysis. mimeo, Princeton University, 2004.

K. Rao Kadiyala and Sune Karlsson. Numerical methods for estimation and inference in bayesian var-models. *Journal of Applied Econometrics*, 12(2):99–132, March-Apr 1997.

Patrick J. Kelly. Real and inflationary macroeconomic risk in the fama and french size and book-to-market portfolios. mimeo, Arizona State University, 2003.

Chang-Jin Kim and Charles R. Nelson. *State Space Models with Regime Switching*. The MIT Press, Cambridge, Massachusetts, 1999.

Peter J. Knez, Robert Litterman, and José A. Scheinkman. Explorations into factors explaining money market returns. *Journal of Finance*, 49(5):1861–1882, December 1994.

M. Ayhan Kose, Christopher Otrok, and Charles H. Whiteman. International business cycles: World, region, and country-specific factors. *American Economic Review*, 93(4):1216–1239, September 2003.

Hans-Martin Krolzig. Markov-switching procedures for dating the euro-zone business cycle. *Vierteljahreshefte zur Wirtschaftsforschung*, 70:339–351, 2001.

Hans-Martin Krolzig and Juan Toro. Classical and modern business cycle measurement: The european case. Discussion Paper in Economics 60, University of Oxford, 2002.

Peyron Law. Macro factors and the yield curve. mimeo, Stanford University, 2004.

Martin Lettau and Sydney Ludvigson. Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, 109(6):1238–1287, December 2001.

Martin Lettau, Sydney Ludvigson, and Jessica A. Wachter. The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies*, 2006. forthcoming.

Jonathan Lewellen and Stefan Nagel. The conditional capm does not explain asset-pricing anomalies. *Journal of Finance*, 2005. forthcoming.

Qing Li, Maria Vassalou, and Yuhang Xing. Sector investments growth rates and the cross-section of equity returns. *Journal of Business*, 2004. forthcoming.

Jimmy Liew and Maria Vassalou. Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics*, 57(2): 221–245, August 2000.

John Lintner. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47: 13–37, February 1965.

Robert Litterman and José A. Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, 1:54–61, June 1991.

Kirsten Lommatzsch and Sabine Stephan. Seasonal adjustment methods and the determination of turning points of the emu business cycle. *Vierteljahreshefte zur Wirtschaftsforschung*, 70(3):399–415, 2001.

Sydney C. Ludvigson and Serena Ng. The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics*, 2005. forthcoming.

Hanno N. Lustig and Stijn G. Van Nieuwerburgh. Housing collateral, consumption insurance, and risk premia: An empirical perspective. *Journal of Finance*, 60(3):1167–1219, 06 2005.

Harry Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–99, September 1952.

Robert C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41 (5):867–887, September 1973.

James Mitchell, Richard J. Smith, Martin R. Weale, Stephen Wright, and Eduardo L. Salazar. An indicator of monthly gdp and an early estimate of quarterly gdp growth. *Economic Journal*, 115(501):F108–F129, 02 2005.

Emanuel Mönch. Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented var approach. Working Paper Series 544, European Central Bank, November 2005.

Emanuel Mönch. Conditional asset pricing with a large information set. mimeo, Humboldt University Berlin, 2006.

Emanuel Mönch and Harald Uhlig. Towards a monthly business cycle chronology for the euro area. *Journal of Business Cycle Measurement and Analysis*, 2(1): 43–69, 2005.

Andrew Mountford and Harald Uhlig. What are the effects of fiscal policy shocks? SFB 649 Discussion Papers 2005-039, Humboldt University Berlin, July 2005.

Charles R. Nelson and Andrew F. Siegel. Parsimonious modeling of yield curves. *Journal of Business*, 60(4):473–489, October 1987.

OECD. A complete and consistent macro-economic data set of the euro area: Methodological issues and results. mimeo, OECD, Paris, 2002.

Christopher Otrok and Charles H Whiteman. Bayesian leading indicators: Measuring and predicting economic conditions in iowa. *International Economic Review*, 39(4):997–1014, November 1998.

Ralitsa Petkova. Do the fama-french factors proxy for innovations in predictive variables? *Journal of Finance*, 2005. forthcoming.

Monika Piazzesi, Martin Schneider, and Selale Tuzel. Housing, consumption, and asset pricing. *Journal of Financial Economics*, 2006. forthcoming.

Dimitris N. Politis and Joseph P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313, 1994a.

Dimitris N. Politis and Joseph P. Romano. Limit theorems for weakly dependent hilbert space valued random variables with applications to the stationary bootstrap. *Statistica Sinica*, 4(2):461–476, 1994b.

Tommaso Proietti. Temporal disaggregation by state space methods: Dynamic regression methods revisited. mimeo, Università di Udine, November 2004.

Danny Quah and Thomas J. Sargent. A dynamic index model for large cross sections. In J.H. Stock and M.W. Watson, editors, *Business Cycles, Indicators, and Forecasting*, pages 285–306. University of Chicago Press, 1993.

Christina D. Romer. Remeasuring business cycles. *Journal of Economic History*, 54 (3):573–609, 1994.

Stephen A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13:341–360, 1976.

Glenn Rudebusch and Tao Wu. A macro-finance model of the term structure, monetary policy, and the economy. Federal Reserve Bank of San Francisco Working Paper 03-17, Federal Reserve Bank of San Francisco, 2003.

Tano Santos and Pietro Veronesi. Labor income and predictable stock returns. *Review of Financial Studies*, 19(1):1–44, 2006.

Thomas J. Sargent. Two models of measurements and the investment accelerator. *Journal of Political Economy*, 97(2):251–287, April 1989.

Thomas J. Sargent and Christopher A. Sims. Business cycle modeling without pretending to have too much a priori economic theory. In Christopher A. Sims, editor, *New Methods in Business Cycle Research*. Federal Reserve Bank of Minneapolis, 1977.

Lucio Sarno and Daniel L. Thornton. The efficient market hypothesis and identification in structural vars. *Federal Reserve Bank of St. Louis Review*, 86(1):49–60, 2004.

Paul Schreyer. Some observations on international area aggregates. OECD Statistics Working Papers 2001/2, OECD, Paris, December 2001.

Jay Shanken. On the estimation of beta-pricing models. *Review of Financial Studies*, 5(1):1–33, 1992.

William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442, 1964.

James H. Stock and Mark W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162, April 2002a.

James H. Stock and Mark W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97:1167–1179, 2002b.

James H. Stock and Mark W. Watson. Implications of dynamic factor models for var analysis. NBER Working Paper 11467, National Bureau of Economic Research, July 2005.

James H. Stock and Mark W. Watson. A probability model of the coincident economic indicators. In K. Lahiri and G.H. Moore, editors, *Leading Economic Indicators: New Methods and Forecasting Records*. Cambridge University Press, New York, 1991.

James H. Stock and Mark W. Watson. Diffusion indexes. NBER Working Paper 6702, National Bureau of Economic Research, 1998.

Harald Uhlig. What moves gnp? mimeo, Humboldt University Berlin, August 2004.

Harald Uhlig. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419, March 2005.

Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

Maria Vassalou. News related to future gdp growth as a risk factor in equity returns. *Journal of Financial Economics*, 68(1):47–73, April 2003.

Kevin Q. Wang. Conditioning information, out-of-sample validation, and the cross-section of stock returns. mimeo, University of Toronto, 2004.

Yong Wang. A comparison of factor models for explaining the cross section of stock returns. mimeo, Northwestern University, 2005.

Halbert White. A reality check for data snooping. *Econometrica*, 68(5):1097–1126, September 2000.

Tao Wu. Macro factors and the affine term structure of interest rates. Federal Reserve Bank of San Francisco Working Paper 02-06, Federal Reserve Bank of San Francisco, 2002.

6 Technical Appendix

6.1 Technical Appendix to Chapter 2

6.1.1 Estimation of the Common Factors

To extract the common factors from the panel of 132 macroeconomic time series for the US, I employ the method popularized by Stock and Watson [2002a,b]. Under the normalization $F'F/T = I_q$, Stock and Watson show that for large M and T the space spanned by the common factors can be consistently estimated by the principal components of the cross-sectional variance-covariance matrix of the data. Precisely, let V denote the eigenvectors corresponding to the q largest eigenvalues of the $T \times T$ cross-sectional variance-covariance matrix XX'. Then, estimates $\hat{F}$ of the factors and $\hat{\Lambda}$ of the factor loadings are given by

$$\begin{aligned} \hat{F} &= \sqrt{T}\,V \\ \text{and} \quad \hat{\Lambda} &= \sqrt{T}\,X'V, \end{aligned}$$

i.e. the common factors are estimated as $\sqrt{T}$ times the q largest eigenvalues of the variance-covariance matrix XX'.

In practice, the true number q of common components is not known and therefore needs to be estimated. This is done using the panel information criteria developed by Bai and Ng [2002]. Using the MATLAB code provided by Serena Ng on her website,[51] I obtain an estimate of the true number of factors. Precisely, I employ the BIC_3 criterion considered by Bai and Ng [2002] which the authors note to be useful in practice since it has particularly good properties in the presence of cross-correlations. This criterion delivers an estimate of $\hat{q} = 8$ for the panel of monthly time series and of $\hat{q} = 7$ for the quarterly panel.

6.1.2 Estimation of the Pricing Model and Diagnosis Tests

After the factors are extracted from the panel of macro time series, they are individually used as instruments in tests of the conditional CAPM. In this section, I provide details on the estimation of the model and on the test statistics based on which the performance of different model specifications has been compared. The models are estimated using Hansen's (1982) Generalized Method of Moments

[51] http://www-personal.umich.edu/∼ngse/research.html

(GMM) which I describe in some detail in Section 6.1.2. To facilitate comparison with previous studies, I also provide results from cross-sectional regressions. These are briefly sketched in Section 6.1.2.

GMM Estimation

Factor pricing models formulated using the SDF terminology give rise to a set of moment conditions that can be used for estimation via the Generalized Method of Moments (GMM). To obtain parameter estimates of my conditional pricing models, I closely follow Cochrane (2001, chapter 13) who provides a nice overview of how the GMM procedure can be applied in practice. Any model of the form (2.4) implies a vector of pricing errors

$$g(b) = E\left[b' f_t R_t - \mathbf{1}\right]$$

with sample analog

$$g_T(b) = \sum_{t=1}^{T} \left(b' f_t R_t - \mathbf{1}\right).$$

If the model is valid, the pricing errors must be zero. The GMM procedure uses this condition to choose parameter estimates b which minimize the weighted sum of squared pricing errors

$$J_T = g_T(b)' W_T g_T(b),$$

where W_T is some weighting matrix. Hansen [1982] shows that estimates $\hat{b}$ are efficient if the weighting matrix $W_T = S_T^{-1}$ is used, where S_T is a consistent estimator of $S = [T\ Cov(g_T)]$. As S_T is a function of g_T which depends on b, estimation of the parameters in optimal GMM proceeds in two steps. In the first stage, the weighting matrix $W_T = I$ is used and the second stage has $W_T = S^{-1}$. The corresponding parameter estimates are given by

$$\begin{aligned}
\text{First stage}: \hat{b}_1 &= (d'd)^{-1} d' E_T(p) && (6.1)\\
\text{Second stage}: \hat{b}_2 &= (d'S^{-1}d)^{-1} d' S^{-1} E_T(p), && (6.2)
\end{aligned}$$

where $E_T(p) = \mathbf{1_N}$ is a $N \times 1$ vector of ones and where $d' = \frac{1}{T}(f'R)$ is the sample mean of the product of returns R and the pricing factors f of the unconditional model.[52] To obtain an estimate of S, I follow Cochrane (2001, p. 221) and use the

[52] Recall that the expected discounted value of a gross return is always equal to one, hence $E_T(p) = \mathbf{1_N}$. In case excess returns are used to estimate the models, $E_T(p)$ would equal a vector of zeros. This in turn would result in an identification problem and the procedure would have to be slightly adjusted. See Cochrane (2001, pp. 256-258) for details.

Bartlett estimate

$$S_T = \sum_{j=-k}^{k} \left(\frac{k-|j|}{k} \right) \frac{1}{T} \sum_{t=1}^{T} (u_t u'_{t-j}), \tag{6.3}$$

with a window size of k=12. Here, u denotes the vector of pricing errors implied by the first stage parameter estimates $\hat{b}_1$.

Based on the second-stage estimates, a test for the null hypothesis that the pricing errors are jointly zero is given by the J-statistic:

$$TJ(\hat{b}) = T\, g_T(\hat{b})'\, S_T^{-1}\, g_T(\hat{b}) \quad \sim \ \chi^2(N-k), \tag{6.4}$$

where N denotes the number of moment conditions, equal to the number of test assets, and k the number of parameters to estimate.

The weighting matrix W_T=S_T^{-1} is model-dependent since it assigns bigger weights to assets with small variances in their pricing errors and smaller weights to assets with large variances in their pricing errors. Hence, comparisons of model fit based on the statistic J_T are infeasible. Hansen and Jagannathan [1997] have therefore suggested to employ a weighting matrix that is identical across different models. In particular, they propose to use as weighting matrix the inverse of the second moments of asset returns $W_T = E[RR']^{-1}$ which implies bigger weights for assets with small variance and smaller weights for assets with large variance. Using the HJ weighting matrix, I obtain estimates $\hat{b}_{HJ}$ of the model coefficients from

$$\hat{b}_{HJ} = \left(d'E[RR']^{-1}d \right)^{-1} d'E[RR']^{-1}E_T(p). \tag{6.5}$$

The J-test then becomes

$$J_{HJ} = g_T(\hat{b}_{HJ})'\, Cov(g_T(\hat{b}_{HJ}))^{-1}\, g_T(\hat{b}_{HJ}) \quad \sim \ \chi^2(N-k), \tag{6.6}$$

where $g_T(\hat{b}_{HJ})$ denotes the vector of average pricing errors implied by $\hat{b}_{HJ}$ and $Cov(g_T(\hat{b}_{HJ}))$ their sample covariance. The latter is given by

$$Cov(g_T(\hat{b}_{HJ})) = \frac{1}{T}(\mathbf{I_N} - d(d'W_T d)^{-1} d'W_T) S_T (\mathbf{I_N} - d(d'W_T d)^{-1} d'W_T)', \tag{6.7}$$

where S_T now denotes the Bartlett estimate of the spectral density matrix of the pricing errors implied by $\hat{b}_{HJ}$ (see e.g. Cochrane 2001, p. 255 or Hodrick and Zhang 2001, p. 335).

Another test statistic suitable for model comparisons is the Hansen-Jagannathan distance measure which denotes the minimum distance between the pricing kernel implied by a model and the set of true stochastic discount factors. The HJ-distance is given by

$$\delta = \left(g_T(\hat{b}_{HJ})' \; E[RR']^{-1} \; g_T(\hat{b}_{HJ}) \right)^{1/2}. \tag{6.8}$$

Jagannathan and Wang [1996] show that the asymptotic sampling distribution of the HJ-distance equals a weighted sum of $(N-k)$ $\chi^2(1)$-distributed random variables. The weights are given by the $N-k$ nonzero eigenvalues of the matrix

$$A = S_T^{1/2} W_T^{1/2'} \left(\mathbf{1_N} - W_T^{1/2} d (d' W_T d)^{-1} d' W_T^{1/2'} \right) W_T^{1/2} S_T^{1/2'}, \tag{6.9}$$

where $W_T^{1/2}$ and $S_T^{1/2}$ denote the upper-triangular Choleski factors of the weighting matrix W_T and the spectral density matrix of the model-implied pricing errors S_T. To obtain the p-value for the HJ-distance, I simulate this weighted sum 100,000 times.

To test for parameter instability in the relationship between the pricing factors and the returns on the test assets, I apply Andrews' (1993) supLM test for structural breaks. More precisely, I compute Andrews' LM-statistic at increments of 0.05 over the interval [0.15 ; 0.85] and report as supLM the supremum of these statistics. The LM-statistic is given by

$$LM(\pi) = \frac{T}{\pi(1-\pi)} g_T(\pi)' S_T^{-1} d (d' S_T^{-1} d)^{-1} d' S_T^{-1} g_T(\pi), \tag{6.10}$$

where $g_T(\pi)$ denotes the pricing errors implied by the optimal GMM estimator averaged over the period $t = \{1, \ldots, \pi T\}$ and where S_T denotes the corresponding spectral density matrix. The judgment whether a model fails or passes the supLM test is based on the distribution tables provided in Andrews [1993].

Cross-Sectional Regressions

Pricing models of the form (2.4) imply an unconditional multifactor beta representation for returns as given in (2.10). Based on this formulation, the model can be consistently estimated using the cross-sectional regression methodology of Fama and MacBeth [1973]. The Fama-MacBeth procedure works in two steps. First, estimates of the risk exposures $\hat{\beta}_{i,f_j}$ are obtained from time-series regressions of individual returns on the pricing factors, i.e.

$$R_i = \beta_{i,f_1} f_1 + \ldots + \beta_{i,f_k} f_k + \alpha_i. \tag{6.11}$$

In a second step, the price of risk parameters $\gamma_{f_j,t}$ are estimated by running for each time period a cross-sectional regression of the vector of returns on the betas, i.e.

$$R_t = \gamma_{f_1,t}\hat{\beta}_{i,f_1} + \ldots + \gamma_{f_k,t}\hat{\beta}_{i,f_k} + \epsilon_t. \tag{6.12}$$

Estimates and standard errors of the unconditional factor risk premia $\gamma = (\gamma_{f_1}, \ldots, \gamma_{f_k})$ are obtained by computing time series averages and the associated sample variances of the $\hat{\gamma}_t$, i.e.

$$\hat{\gamma} = \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_t \tag{6.13}$$

$$\hat{\Sigma}_\gamma = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\gamma}_t - \hat{\gamma}). \tag{6.14}$$

Since the regressors $\hat{\beta}_{i,f_j}$ are estimated in the first step, the Fama-MacBeth method suffers from an errors in variables problem. Shanken [1992] derives a correction term for the variance estimate that is designed to adjust for the bias. In particular, he shows that

$$\sqrt{T}(\hat{\gamma} - \gamma) \sim N\left(0, (1 + \hat{\gamma}'\hat{\Sigma}_f^{-1}\hat{\gamma}) \cdot \hat{\Sigma}_\gamma + \Sigma_f\right), \tag{6.15}$$

where $\hat{\Sigma}_f$ denotes the sample covariance of the pricing factors.

To evaluate the overall model performance in the Fama-MacBeth setup, one can compute the cross-sectional R-square following Jagannathan and Wang [1996] as

$$R^2 = \frac{Var_c(\bar{R}_i) - Var_c(\bar{\alpha}_i)}{Var_c(\bar{R}_i)}, \tag{6.16}$$

where Var_c denotes the cross-sectional variance, $\bar{\alpha}_i$ is the time-series average of the pricing error for asset i, and $\bar{R}_i$ is the time series average of the return on asset i. I report the R^2 statistic as well as its degrees of freedom-adjusted counterpart for each of the tested models.

I further test whether the average pricing errors are jointly zero. Following Cochrane (2001, p. 246), let

$$\bar{\alpha} = \frac{1}{T}\sum_{t=1}^{T}\alpha_t \quad \text{and} \quad Cov(\bar{\alpha}) = \frac{1}{T^2}\sum_{t=1}^{T}(\alpha_t - \bar{\alpha})(\alpha_t - \bar{\alpha})'$$

denote the vector of model-implied average pricing errors and its sample covariance matrix. Then, the Chi-square statistic

$$J_\alpha = \bar{\alpha}'\ Cov(\bar{\alpha})^{-1}\ \bar{\alpha} \sim \chi^2(N-1) \tag{6.17}$$

can be used to assess whether the average pricing errors are jointly equal to zero. I report J_α-statistics and the corresponding p-values as a complementary model diagnosis to the cross-sectional R^2.

6.2 Technical Appendix to Chapter 3

6.2.1 Derivation of the Bond Pricing Parameters

The absence of arbitrage between bonds of different maturity implies the existence of the stochastic discount factor m such that

$$P_t^{(n)} = E_t[m_{t+1}\, P_{t+1}^{(n-1)}],$$

i.e. the price of a n-months to maturity bond in month t must equal the expected discounted price of an $(n-1)$-months to maturity bond in month $(t+1)$. Following Ang and Piazzesi [2003], the derivation of the recursive bond pricing parameters starts with assuming that the nominal pricing kernel m takes the form

$$m_{t+1} = \exp(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t - \lambda_t'\omega_{t+1})$$

and by guessing that bond prices P are exponentially affine in the state variables Z, i.e.

$$P_t^{(n)} = \exp(A_n + B_n' Z_t).$$

Plugging the above expressions for P and m into the first relation, one obtains

$$\begin{aligned}
P_t^{(n)} &= E_t[m_{t+1}\, P_{t+1}^{(n-1)}] \\
&= E_t\left[\exp(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t - \lambda_t'\omega_{t+1})\exp(A_{n-1} + B_{n-1}' Z_{t+1})\right] \\
&= \exp(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t + A_{n-1})\; E_t\left[\exp(-\lambda_t'\omega_{t+1} + B_{n-1}'(\mu + \Phi Z_t + \omega_{t+1}))\right] \\
&= \exp(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t + A_{n-1} + B_{n-1}'\mu + B_{n-1}'\Phi Z_t) E_t\left[\exp((-\lambda_t' + B_{n-1}')\omega_{t+1})\right]
\end{aligned}$$

Since the innovations ω of the state variable process are assumed Gaussian with variance-covariance matrix Ω, it is obvious that

$$\begin{aligned}
\ln E_t\left[\exp((-\lambda_t' + B_{n-1}')\omega_{t+1})\right] &= E_t\left[\ln(\exp((-\lambda_t' + B_{n-1}')\omega_{t+1}))\right] + \\
&\quad \frac{1}{2} Var_t\left(\ln(\exp((-\lambda_t' + B_{n-1}')\omega_{t+1}))\right) \\
&= \frac{1}{2}\left[\lambda_t'\Omega\lambda_t - 2B_{n-1}'\Omega\lambda_t + B_{n-1}'\Omega B_{n-1}\right] \\
&= \frac{1}{2}\lambda_t'\Omega\lambda_t - B_{n-1}'\Omega\lambda_t + \frac{1}{2}B_{n-1}'\Omega B_{n-1}.
\end{aligned}$$

Hence, $E_t\left[\exp((-\lambda_t' + B_{n-1}')\omega_{t+1})\right] = \exp(\frac{1}{2}\lambda_t'\Omega\lambda_t - B_{n-1}'\Omega\lambda_t + \frac{1}{2}B_{n-1}'\Omega B_{n-1})$ and thus

$$\begin{aligned} P_t^{(n)} &= \exp(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t + A_{n-1} + B_{n-1}'\mu + B_{n-1}'\Phi Z_t + \ldots \\ &\quad + \frac{1}{2}\lambda_t'\Omega\lambda_t - B_{n-1}'\Omega\lambda_t + \frac{1}{2}B_{n-1}'\Omega B_{n-1}). \end{aligned}$$

Using the relations $r_t = \delta' Z_t$ and $\lambda_t = \lambda_0 + \lambda_1 Z_t$, and matching coefficients finally yields

$$P_t^{(n)} = \exp(A_n + B_n' Z_t),$$

where

$$\begin{aligned} A_n &= A_{n-1} + B_{n-1}'(\mu - \Omega\lambda_0) + \frac{1}{2}B_{n-1}'\Omega B_{n-1}, \\ \text{and} \quad B_n &= B_{n-1}'(\Phi - \Omega\lambda_1) - \delta'. \end{aligned}$$

These are the recursive equations of the pricing parameters stated in (3.7)-(3.8).

6.2.2 Estimation of the No-Arbitrage FAVAR Model

Estimation of the No-Arbitrage Factor-Augmented VAR model proceeds in three consecutive steps. First, I extract the common factors from the panel of 160 monthly time series using the method of Stock and Watson [2002a,b] which has been described in detail in Section 6.1 above. The extracted factors are then treated as if they were data, and estimation of the term structure model is performed using the consistent two-step approach suggested by Ang et al. [2006]. This method involves first the estimation of the parameters (μ, Φ, Σ) by running OLS regressions individually on each of the equations of the VAR(p) in (3.4). In a second step, then, the market price of risk parameters (λ_0, λ_1) are estimated by minimizing the sum of squared pricing errors implied by the model. Precisely, I compute the model-implied yields $\hat{y}_t^{(n)} = \hat{a}_n + \hat{b}_n' Z_t$ based on the recursive formulae in (3.7) and (3.8) holding the FAVAR parameters (μ, Φ, Σ) fixed, and minimize the sum of squared pricing errors stated in (3.10) with respect to λ_0 and λ_1. Computationally, this is achieved using the MATLAB function *lsqnonlin* which implements a subspace trust region method that is based on the interior-reflective Newton algorithm.[53]

[53] See the MATLAB documentation for more details.

Due to the recursive formulation of the parameters of the affine model, the minimization problem is highly nonlinear. It is therefore helpful to find sensible starting values to initialize the algorithm. I achieve this using the following strategy. First, I set to zero all elements of the matrix λ_1 to obtain initial estimates of the parameters in λ_0. Then, I take these estimates as starting values in a second estimation step, i.e. I let all elements of λ_0 and λ_1 be estimated freely. Finally, to enhance tractability of the model, I follow an approach common in the affine term structure literature, and re-estimate the model after setting to zero those elements of λ_1 which are insignificant. To assess significance, I compute standard errors of the prices of risk parameters via the numerical gradient of the sum of squared fitting errors that is delivered by the function *lsqnonlin*.

6.2.3 Implementation of White's "Reality Check"

White's (2000) "Reality Check" test is based on a bootstrap resampling from the forecast errors implied by different forecast models. It can be used to evaluate whether some model has superior predictive ability with respect to one or more benchmark models. In my application, I use the No-Arbitrage FAVAR model as the benchmark and compare it with each of the competitor models. White's algorithm is based on the stationary block bootstrap of Politis and Romano [1994a,b]. Given a smoothing parameter q between 0 and 1 and a forecast period $\{T_1, \ldots, T_2\}$, the bootstrap procedure of Politis and Romano involves the following steps.

(i) Set $t = T_1$. Draw $\theta(T_1)$ independently and uniformly from $\{T_1, \ldots, T_2\}$.

(ii) Increment t. If $t > T$, stop. Otherwise, draw a random variable U from the uniform distribution with support $[0, 1]$. If $U < q$, draw $\theta(t)$ independently and uniformly from $\{T_1, \ldots, T_2\}$. If $U \geq q$, set $\theta(t) = \theta(t-1) + 1$; if $\theta(t) > T$, reset to $\theta(t) = T_1$.

(iii) Repeat (ii).

This algorithm delivers blocks of indexes between T_1 and T_2 which are of random length and distributed according to the Geometric distribution with mean block length equal to $1/q$. I follow Hördahl et al. [2006] and use a smoothing parameter of $q = 1/12$.

White's "Reality Check" algorithm proceeds in the following steps:

1. Generate M sets of random indexes using the block bootstrap algorithm described above. Denote them with $\{\theta_i(t), t = T_1, \ldots, T_2\}, i = 1, \ldots, M$.

2. Compute the series of squared forecast errors of the benchmark model as $\hat{h}_{0,t} = -(\hat{y}_t - y_t)^2, t = T_1, \ldots, T_2$. Do the same for the competitor model and denote the resulting series of squared forecast errors as $\hat{h}_1$.

3. Form the difference of forecast losses $\hat{f}_t = \hat{h}_{1,t} - \hat{h}_{0,t}, t = T_1, \ldots, T_2$ and compute their average $\bar{f} = \frac{1}{m}\sum_{t=T_1}^{T_2} \hat{f}_t$ where $m = T_2 - T_1 + 1$.

4. Perform the same for each of the M sets of bootstrapped indexes. For each $i = \{1, \ldots, M\}$, denote the average of the forecast loss difference series as $\bar{f}_i^*$. Furthermore, set $\bar{V} = m^{1/2}\bar{f}$ and compute $\bar{V}_i^* = m^{1/2}(\bar{f}_i^* - \bar{f})$.

5. Compare the sample value of $\bar{V}$ to the percentiles of $\bar{V}_i^*$.

I repeat this procedure for each of the competitor models and report the corresponding test statistics $\bar{V}$ in Table 3.5. Negative figures obviously indicate that the average squared forecast loss of the benchmark model is smaller than that of the respective competitor model while positive test statistics indicate the opposite. I perform $M = 1,000$ block-bootstrap resamples from the prediction error series to compute the significance of the forecast improvement.

6.3 Technical Appendix to Chapter 4

This appendix discusses in detail identification and estimation of the Bayesian factor model suggested in Chapter 4, as well as the methods used to identify the shocks to the yield curve factors.

For convenience, I restate the main equations of the joint factor model of macroeconomic and interest rate data.

$$\begin{aligned}
\begin{pmatrix} X_t \\ Y_t \end{pmatrix} &= \begin{bmatrix} \Lambda_x & 0 \\ 0 & \Lambda_y \end{bmatrix} \begin{pmatrix} F_t^x \\ F_t^y \end{pmatrix} + \begin{pmatrix} e_t^x \\ e_t^y \end{pmatrix} \\
\text{or} \qquad Z_t &= \Lambda \quad F_t \quad + \quad e_t. \qquad (6.18)
\end{aligned}$$

$$\begin{aligned}
\begin{pmatrix} F_t^x \\ F_t^y \end{pmatrix} &= \Phi(L) \begin{pmatrix} F_{t-1}^x \\ F_{t-1}^y \end{pmatrix} + \begin{pmatrix} \omega_t^x \\ \omega_t^y \end{pmatrix}. \\
\text{or} \qquad F_t &= \Phi(L) \quad F_{t-1} \quad + \quad \omega_t. \qquad (6.19)
\end{aligned}$$

In a nutshell, a set of N_x macro time series stacked in the vector X and a set of N_y yields stacked in the vector Y are driven by two different groups of factors, F^x and F^y of dimension $k_x \times 1$ and 3×1, respectively. Furthermore, all variables exhibit idiosyncratic components e^x and e^y that are assumed to be mutually and serially uncorrelated, i.e. $R = E[ee']$ is a diagonal $N \times N$ matrix where $N = N_x + N_y$. The two groups of factors are modeled to share common dynamics within a VAR. The innovations of this VAR are assumed to be contemporaneously correlated, i.e. $\Omega = E[\omega\omega']$ is a symmetric matrix of dimension $(k_x+3) \times (k_x+3)$.

6.3.1 Identification of the Factor Model

Before discussing how the model (6.18)-(6.19) can be estimated, the restrictions needed to ensure unique identification of the model parameters and the unobserved factors have to be stated. Factor models suffer from the well-known problem of rotational indeterminacy meaning that different rotations of the factors and model parameters may be observationally equivalent. To illustrate this point, rewrite model (6.18)-(6.19) as

$$\begin{aligned}
Z_t &= \Lambda P^{-1} P F_t + e_t \\
P F_t &= P \Phi(L) P^{-1} P F_{t-1} + P \omega_t
\end{aligned}$$

where P is some non-singular $k \times k$ matrix.

Denoting $\tilde{\Lambda} = \Lambda P^{-1}, \tilde{\Phi} = P\Phi(L)P^{-1}, \tilde{F}_t = PF_t$, and $\tilde{\omega}_t = P\omega_t$, the model can be rewritten as

$$Z_t = \tilde{\Lambda}\tilde{F}_t + e_t \tag{6.20}$$

$$\tilde{F}_t = \tilde{\Phi}(L)\tilde{F}_{t-1} + \tilde{\omega}_t. \tag{6.21}$$

This model gives rise to the same likelihood as model (6.18)-(6.19), but implies different parameter and factor estimates. To ensure that one can uniquely identify the model parameters and the latent factors, one therefore has to impose restrictions such that no nonsingular rotation P is feasible. This can be done in the following way. First, fixing the unconditional variance of the latent factors to be an identity matrix implies

$$Var(F_t) = I_k \equiv Var(\tilde{F}_t) = PVar(F_t)P' = PP'.$$

Hence, P has to be orthogonal. Further restrictions on P can be derived from the structure of the factor loading matrix Λ. As we have seen above,

$$\Lambda = \begin{pmatrix} \Lambda_x & 0 \\ 0 & \Lambda_y \end{pmatrix}$$

which implies a set of zero restrictions. Hence, any rotation $\tilde{\Lambda} = \Lambda P'$ must be of the same form

$$\tilde{\Lambda} = \begin{pmatrix} \tilde{\Lambda}_x & 0 \\ 0 & \tilde{\Lambda}_y \end{pmatrix} \tag{6.22}$$

where $\tilde{\Lambda}_x$ and $\tilde{\Lambda}_y$ are of dimension $k_x \times k_x$ and 3×3, respectively. This particular structure implies that the orthogonal rotation matrix P must also be block-diagonal. Accordingly, identification of the joint factor model reduces to separate identification problems for each subset of factors. For convenience, denote the upper-left $k_x \times k_x$ block of P with P_{xx} and the lower-right 3×3 block P_{yy}. Consider first the lower N_y equations of (6.18) corresponding to the observation equations for the yields. We know from equation (4.2) that the yield loadings are given by $\lambda_n^y = \left[1 \quad \left(\frac{1-e^{-\tau n}}{\tau n}\right) \quad \left(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)\right]$. Any rotation P_{yy} of the yield factors must preserve this particular structure. It is straightforward to show that this only holds true for an identity matrix P_{yy}. Hence, the tight parametric structure imposed by the Nelson-Siegel decomposition of yields ensures unique identification of the level, slope, and curvature factors. It thus remains to fix Λ_x such that a unique orthogonal rotation matrix P_{xx} of the macro factors is implied. Here, I can built on existing results by Geweke and Zhou [1996] who show that in case the variance covariance matrix of the factors is identity, exact identification is ensured by restricting the upper left $k_x \times k_x$ block of Λ_x to be lower-triangular with positive diagonal elements. Aguilar and West [2000] extend this "hierarchical" identification scheme to the case where the factor variances are left unrestricted but instead the diagonal elements of the the upper left $k_x \times k_x$ block of Λ_x are set to unity. In the sequel, I stick to that latter identification scheme, i.e. I estimate the factor variances and restrict the diagonal elements of the upper left $k_x \times k_x$ block of Λ_x to be unity.

6.3.2 Estimation using a Metropolis-within-Gibbs Sampler

The Metropolis-within-Gibbs Sampling Algorithm used to estimate the joint dynamic factor model of macroeconomic and interest rate data of Chapter 4 involves several consecutive steps that are briefly sketched in Section 4.3. The objective of the present section is to describe in greater detail the implementation of the different steps.

For convenience, I restate the notation introduced in Section 4.3. Let $\theta = (\Lambda^x, \Lambda^y, R, \Phi, \Omega)$ denote the set of model parameters. Furthermore, stack all T observations on yields and macro variables in the vectors $\tilde{X}_T = \{X_1, \ldots, X_T\}$ and $\tilde{Y}_T = \{Y_1, \ldots, Y_T\}$, and let $\tilde{Z}_T = \{\tilde{X}_T, \tilde{Y}_T\}$. Analogously, let $\tilde{F}_T = \{\bar{F}_1, \ldots, \bar{F}_t\}$ denote all observations of the factors F.

The Gibbs sampling algorithm approximates the joint posterior distribution $p(\theta, \tilde{F}_T|\tilde{Z}_T)$ of the model parameters and the unobserved factors by sampling from the conditional posteriors $p(\theta|\tilde{Z}_T, \tilde{F}_T)$ and $p(\tilde{F}_T|\tilde{Z}_T, \theta)$. The state-space form of dynamic factor models allows straightforward derivation of the latter. In the model studied here, a complication arises due to the nonstandard distribution of the exponential decay parameter τ. To overcome this problem, I introduce a Metropolis algorithm. The details are discussed further below.

The Gibbs sampler proceeds as follows.

Step 0: Find starting values θ^0.

Step 1: Draw $\tilde{F}_T^{(i)}$ from $p(\tilde{F}_T|\tilde{Z}_T, \theta^{(i-1)})$.

Step 2: Draw $\theta^{(i)}$ from $p(\theta|\tilde{Z}_T, \tilde{F}_T^{(i)})$.

Steps 1 and 2 constitute one iteration of the sampler, and are repeated until the empirical distributions of $\tilde{F}_T^{(i)}$ and $\theta^{(i)}$ converge. The crucial result employed in the Gibbs sampler is that these empirical distributions converge to the joint marginal posterior distribution as the number of iterations goes to infinity. Accordingly, after discarding an initial number of draws (the "burn-in"), sampling from the known conditional posterior densities of factors and parameters is equivalent to sampling from their unknown joint posterior distribution. I provide more details on each of the three steps below.

Find Starting Values θ^0

If the model is exactly identified, the algorithm should converge to the ergodic distribution of the model parameters independently of the choice of initial parameter values. The estimation results reported in Section 4.5.3 confirm that the factor estimates are identical for different sets of randomly chosen starting values. However, to achieve fast convergence of the sampler, it is advisable to choose a meaningful set of initial parameter values. In the following, I document in detail the individual steps carried out to find starting values for the different model parameters.

1. I use principal components estimates as starting values for the factors F^x driving the macro variables. That is, I use the method depicted in Section 6.1.1 above to extract the first $k_x = 4$ principal components from the panel of macro time series stacked in X.

2. I obtain starting values for the latent yield factors F^y by setting $\tau = 0.0609$ which is the value chosen by Diebold and Li [2006]. The entries of Λ_y are completely determined by τ and the maturities n of the yields stacked in Y. Given the loadings, estimates of the factors can therefore be obtained

by regressing period by period the vector of yield observations onto the loadings Λ_y, i.e. $\hat{F}_t^y = \left(\Lambda_y{}'\Lambda_y\right)^{-1}\Lambda_y{}'Y_t \quad \forall\, t = (1,\ldots,T)$.

3. I get starting values of the parameters Φ and Ω in the joint VAR of macro and yield factors by performing OLS estimation of the VAR equation by equation.

4. I obtain starting values of the parameters Λ_x by running OLS regressions of the variables in X onto the principal components estimates of the factors F^x equation by equation. Moreover, I use the sample variances of the regression residuals and the residuals of the yield regressions as starting values for the entries of R.

Sample from $p(\tilde{F}_T|\tilde{Z}_T,\theta^{(i-1)})$

This section provides additional details on the algorithm that I employ to sample from the posterior distribution of the latent factors. My approach closely follows the exposition in Kim and Nelson [1999]. Based on a result of Carter and Kohn [1994], Kim and Nelson show that draws from the conditional posterior distribution of the latent factors in a state-space model can be obtained by performing the following steps. First run the Kalman filter forward to obtain estimates $\bar{F}_{T|T}$ of the factors in period T and their variance covariance matrix $P_{T|T}$ based on all available sample information. For convenience, I restate the Kalman filter formulae here:

$$
\begin{aligned}
\bar{F}_{t|t-1} &= \bar{\Phi}\,\bar{F}_{t-1|t-1} \\
P_{t|t-1} &= \bar{\Phi}\,P_{t-1|t-1}\,\bar{\Phi}' + \bar{\Omega} \\
e_{t|t-1} &= Z_t - Z_{t|t-1} = Z_t - \bar{\Lambda}\bar{F}_{t-1|t-1} \\
v_{t|t-1} &= \bar{\Lambda}P_{t-1|t-1}\bar{\Lambda}' + R \\
K &= P_{t|t-1}\bar{\Lambda}'v_{t|t-1}^{-1} \\
\bar{F}_{t|t} &= \bar{F}_{t|t-1} + K\,e_{t|t-1} \\
P_{t|t} &= P_{t|t-1} - K\,\bar{\Lambda}P_{t|t-1},
\end{aligned}
$$

where K denotes the Kalman gain. The Kalman filter needs to be initialized with starting values $\bar{F}_0$ and P_0 of the unconditional mean and variance of the latent factors. I follow a common practice in Bayesian dynamic factor models and let $\bar{F}_0$ be a $k(p+1)\times 1$ vector of zeros and P_0 an identity matrix of the corresponding dimension.

Running the Kalman filter forward delivers estimates $\bar{F}_{T|T}$ of the latent factors and their variance covariance matrix $P_{T|T}$ in the last sample period. Given these estimates, Kim and Nelson [1999] show that draws of the latent factors $\bar{F}$ for all previous observations can be obtained by performing the following procedure. For $t = T-1, \ldots, 1$ proceed backwards to generate draws $\bar{F}_{t|T}$ from

$$\begin{aligned} \bar{F}_{t|T}|\bar{F}_{t+1|T}, \tilde{Z}_T, \theta &\sim N(\bar{F}_{t|t,\bar{F}_{t+1|T}}, P_{t|t,\bar{F}_{t+1|T}}) \\ \text{where} \quad \bar{F}_{t|t,\, \bar{F}_{t+1|T}} &= \bar{F}_{t|t} + P_{t|t}\bar{\Phi}' \, (\bar{\Phi}\, \bar{P}_{t|t}\, \bar{\Phi}' + \bar{\Omega})^{-1} \, (\bar{F}_{t+1} - \bar{\Phi}\, \bar{F}_{t|t}) \\ \text{and} \quad P_{t|t,\, P_{t+1|T}} &= P_{t|t} - P_{t|t}\, \bar{\Phi}' \, (\bar{\Phi}\, \bar{P}_{t|t}\, \bar{\Phi}' + \bar{\Omega})^{-1} \, \bar{\Phi}\, P_{t|t}. \end{aligned} \tag{6.23}$$

As Kim and Nelson show, this algorithm needs to be slightly modified if $\bar{\Omega}$ is singular. This is the case here since the state equation includes more than one lag and the model is written in companion form. Denote $\bar{\Omega}^*$ the upper left $k \times k$ block of $\bar{\Omega}$ which is positive-definite and let $\bar{F}_t^*$ and $\bar{\Phi}^*$ be the first k rows of $\bar{F}_t$ and $\bar{\Phi}_t$. As before, run the Kalman filter forward to obtain $\bar{F}_{T|T}$ and $P_{T|T}$. Then, for $t = T-1, \ldots, 1$ proceed backwards to generate draws $\bar{F}_{t|T}$ from

$$\begin{aligned} \bar{F}_{t|T}|\bar{F}^*_{t+1|T}, \tilde{Z}_T, \theta &\sim N(\bar{F}^*_{t|t,\bar{F}_{t+1|T}}, P^*_{t|t,\bar{F}_{t+1|T}}) \\ \text{where} \quad \bar{F}^*_{t|t,\, \bar{F}_{t+1|T}} &= \bar{F}_{t|t} + P_{t|t}\bar{\Phi}^{*'}(\bar{\Phi}^*\, \bar{P}_{t|t}\, \bar{\Phi}^{*'} + \bar{\Omega}^*)^{-1} \, (\bar{F}^*_{t+1} - \bar{\Phi}^*\, \bar{F}_{t|t}) \\ \text{and} \quad P^*_{t|t,\, P_{t+1|T}} &= P_{t|t} - P_{t|t}\bar{\Phi}^{*'}(\bar{\Phi}^*\, \bar{P}_{t|t}\, \bar{\Phi}^{*'} + \bar{\Omega}^*)^{-1} \, \bar{\Phi}^*\, P_{t|t}. \end{aligned}$$

Sample from $p(\theta^{(i)}|\tilde{Z}_T, \tilde{F}_T^{(i)})$

Conditional on the data and draws of the unobserved factors, the observation equation (6.18) of the state-space model amounts to a set of $N_x + N_y$ regressions. Since the errors are assumed to be mutually orthogonal, one can sample each equation's parameters independently using standard results from Bayesian regression analysis. Moreover, conditional on the factor draws, equation (6.19) is just a VAR(p) in the factors. Hence, the distributional theory from Bayesian VARs can be applied to estimate the parameters of the state equation. I start with a description of the algorithm to sample from the parameters of the observation equation and then move to the procedure used to draw the parameters of the state equation.

Sampling Λ and R

Consider first the set of equations characterizing the decomposition of the macroeconomic variables X into common components F^x and idiosyncratic components e^x. Since the idiosyncratic components are assumed to be independent (R is diagonal), the parameters of the macro observation equations can be obtained by estimating N_x independent univariate linear regression models. Using natural conjugate priors

$$\begin{aligned} p\left(\lambda_i^x | R_{ii}\right) &= N(\lambda_{i0}^x, R_{ii} V_{i0}^{-1}) \\ \text{and} \quad p\left(R_{ii}\right) &= iG(\nu_{i0}/2, \nu_{i0}\, \sigma_{i0}^2/2), \end{aligned}$$

standard Bayesian results (see e.g. Bauwens, Lubrano, and Richard 1999) show that the conditional posterior distributions of λ_i^x and R_{ii} are given by

$$p\left(\lambda_i^x | R_{ii}, \tilde{X}_T, \tilde{F}_T, \theta_{-\lambda_i^x}\right) = N(\bar{\lambda}_i^x, R_{ii} V_i^{-1}) \tag{6.24}$$

$$\text{and} \quad p\left(R_{ii} | \tilde{X}_T, \tilde{F}_T, \theta_{-R_{ii}}\right) = iG(\nu_i/2, \nu_i \sigma_i^2/2) \tag{6.25}$$

$$\begin{aligned} \text{where} \quad \bar{\lambda}_i^x &= V_i^{-1} \left(V_{i0}^{-1}\, \lambda_{i0}^x + F^{x\prime} F^x\, \hat{\lambda}_i^x \right), \\ V_i &= V_{i0} + F^{x\prime} F^x, \\ \nu_i &= \nu_{i0} + T, \quad \text{and} \\ \nu_i\, \sigma_i^2 &= \nu_{i0} \sigma_{i0}^2 + (T - k_x)\, S_i^2 + \\ &\quad + (\hat{\lambda}_i^x - \lambda_{i0})' \left[V_{i0}^{-1} + (F^{x\prime} F^x)^{-1} \right]^{-1} (\hat{\lambda}_i^x - \lambda_{i0}) \end{aligned}$$

where $S_i^2 = \frac{1}{T-k_x}(x_i - F^x \hat{\lambda}_i^{x\prime})'(x_i - F^x \hat{\lambda}_i^{x\prime})$ is the sum of squared fitting errors of the i-th equation estimated via OLS. Moreover, $\hat{\lambda}_i^x$ denotes the estimated coefficient from that regression. I use uninformative priors ($\lambda_{i0}^x = 0_{k_x \times 1}, \nu_{i0} = 0, V_{i0}^{-1} = I$) for the estimation. Computationally, I draw from the normal and inverse Gamma distributions using MATLAB's *mvnrnd* and *gamrnd* functions.[54] Finally, recall that the upper left $k_x \times k_x$ block of Λ_x is restricted to be lower-triangular with ones on the diagonal in order to ensure exact identification of the factors. I impose this restriction by setting to zero or to one the respective elements of each draw from the posterior of Λ_x.

[54] Regarding the latter, note that one has to hand over the *inverse* of the scale parameter of the gamma distribution, i.e. I sample from the inverse gamma using the command "1/gamrnd($\nu_i/2, 2/(\nu_i \sigma_i^2)$)".

Matters are more complicated when it comes to estimating Λ_y. Recall that the Diebold-Li loadings of the yield of maturity n on the three factors F^y are given by 1, $\left(\frac{1-e^{-\tau n}}{\tau n}\right)$, and $\left(\frac{1-e^{-\tau n}}{\tau n} - e^{-\tau n}\right)$. Hence, the exponential decay parameter τ is the only unknown parameter in Λ_y. The observation equations related to the N_y yields therefore amount to a nonlinear least squares problem that has to be solved for τ. As a consequence, the distribution of τ is non-standard and cannot be sampled from directly. To solve this problem, I set up the following random walk Metropolis algorithm to draw from τ:

1. Conditional on τ^{j-1}, generate a proposal from $\tau^* = \tau^{(j-1)} + c\epsilon$, where $\epsilon \sim N(0,1)$, i.e. draw τ^* from $N(\tau^{(j-1)}, c^2)$.

2. Accept τ^* with probability

$$\begin{aligned} \alpha &= \min\left\{1, \frac{p\ (\tau^*|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})\ q(\tau^*|\tau^{(i-1)})}{p\ (\tau^{(i-1)}|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})\ q(\tau^{(i-1)}|\tau^*)}\right\} \\ &= \min\left\{1, \frac{p\ (\tau^*|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})}{p\ (\tau^{(i-1)}|\tilde{Z}_T, \tilde{F}_T, \theta_{-\tau})}\right\} \\ &= \min\left\{1, \frac{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^*, \theta_{-\tau})\ p(\tau^*)}{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^{(i-1)}, \theta_{-\tau})\ p(\tau^{(i-1)})}\right\}, \end{aligned} \tag{6.26}$$

   where $\theta_{-\tau}$ denotes the vector of all parameters except for τ.

Note that the q-terms drop since the proposal density is symmetric around $\tau^{(j-1)}$. Hence, with a flat prior, drawing τ amounts to generating a proposal value from a normal distribution centered around the last iteration's value, and to accept that proposal with probability given by the ratio of likelihoods implied by the two candidates, $\frac{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^*, \theta_{-\tau})}{p\ (\tilde{Z}_T|\tilde{F}_T, \tau^{(i-1)}, \theta_{-\tau})}$. This ratio can easily be computed using factor loadings $\Lambda_y(\tau^*)$ and $\Lambda_y(\tau^{(i-1)})$. The scaling parameter c is calibrated so that acceptance ratios between 0.2 and 0.5 are obtained.[55]. Further, as mentioned above, I initialize the sampler using the value of $\tau = 0.0609$.

Conditional on a draw of τ and the latent yield factors F^y, the variances R^y_{ii} of the pricing errors represent mutually independent regression residuals that can be drawn individually. Specifying a flat ($\nu_{i0} = 0$) natural conjugate inverse-Gamma prior distribution, the posterior is given by a conjugate inverse Gamma distribution as in (6.25).

[55] The reported results are based on a value of $c = 0.02$

Sampling Φ and Ω

The state equation of the model (6.18)-(6.19) is a VAR(p) in the factors F. To estimate the parameters of this VAR, I follow Bernanke et al. [2005] by imposing a diffuse conjugate Normal-Wishart prior,

$$\begin{aligned} p\ (vec(\Phi)|\Omega) &= N(0, \Omega \otimes Q_0) \\ p\ (\Omega) &= iW(\Omega_0, \nu_0), \end{aligned}$$

where diagonal elements of Ω_0 are set so as to equal the residual variances of the corresponding p-lag univariate autoregressions, σ_i^2. Further, in the spirit of the Minnesota prior, diagonal elements of Q_0 are set so that the prior variance of the l-lagged jth variable in equation i equals $\sigma_i^2/(l\sigma_j^2)$. Using a result from Kadiyala and Karlsson [1997], the conditional posterior distributions are then given by

$$p\ (vec(\Phi)|\ \Omega, \tilde{Z}_T, \tilde{F}_T, \theta_{-\Phi}) = N\left(vec(\bar{\Phi}), \Omega \otimes \bar{Q}\right) \tag{6.27}$$

$$p\ (\Omega|\ \tilde{Z}_T, \tilde{F}_T, \theta_{-\Omega}) = iW\left(\bar{\Omega}, T + \nu_0\right) \tag{6.28}$$

where

$$\begin{aligned} \bar{\Omega} &= \Omega_0 + \hat{\omega}'\hat{\omega} + \hat{\Phi}'(\tilde{F}_T'\tilde{F}_T)\hat{\Phi} - \bar{\Phi}'\left[Q_0^{-1} + (\tilde{F}_T'\tilde{F}_T)\right]\bar{\Phi}, \\ \bar{\Phi} &= \bar{Q}\ (\tilde{F}_T'\tilde{F}_T)\ \hat{\Phi}, \\ \text{and}\quad \bar{Q} &= \left[Q_0^{-1} + (\tilde{F}_T'\tilde{F}_T)\right]^{-1}, \end{aligned}$$

and where $\hat{\omega}$ denotes the matrix of OLS residuals. Stationarity of the VAR parameters is enforced by discarding draws of Φ that have eigenvalues greater than 1.001 in absolute value. To draw from the posterior inverse Wishart $iW\left(\bar{\Omega}, T + \nu_0\right)$, I follow Cogley, Morozov, and Sargent [2003] and use the following shortcut. First, I draw $(T + \nu_0)k$ random numbers from the standard normal distribution and arrange them in a matrix u of dimension $(T + \nu_0) \times k$. I then set $\Omega = \left(\bar{\Omega}^{-1/2} u u' \bar{\Omega}^{-1/2'}\right)^{-1}$ where $\bar{\Omega}^{-1/2} = chol(\bar{\Omega}^{-1})$ is the Choleski factor of the inverse of $\bar{\Omega}$. This is equivalent to drawing Ω from the inverse Wishart with parameters $\bar{\Omega}$ and $T + \nu_0$.

Identification of Shocks with Sign Restrictions

The objective of Chapter 4 is to identify surprise changes of the yield curve components level, slope, and curvature and to study the subsequent evolution of key macroeconomic aggregates. As stated before, the identification is achieved using sign restriction techniques as in Uhlig [2005] and Mountford and Uhlig [2005]. While the approach has been generally discussed in Section 4.4, I provide some details on the implementation of the different variants in this section.

Combination of Zero and Sign Restrictions

I define as a "pure" surprise change any rise of one of the three yield curve components that is not accompanied by simultaneous responses of the remaining two factors. Computationally, these surprise changes are identified using a combination of zero and sign restrictions. Precisely, I adopt the following strategy.

(i) Randomly and uniformly select an index n from $\{1, \ldots, N_I\}$, where N_I is the number of saved draws from the posterior distribution of the model parameters.

(ii) Compute the Choleski decomposition of $\Omega^{(n)}$. Denote the lower-triangular Choleski factor with $\tilde{A}^{(n)}$ and let $\tilde{A}_i^{(n)}$ and $\tilde{A}_j^{(n)}$ be the rows corresponding to the two remaining yield factors.

(iii) Randomly select a vector of starting values $\alpha = (\alpha_1, \ldots, \alpha_6)$ from the uniform distribution with support $[0, \pi]$. Using the MATLAB function *fsolve* and the parameterization (4.12), find a vector $q(\alpha)$ of unit length which implies

$$\begin{bmatrix} \tilde{A}_{(i)} \\ \tilde{A}_{(j)} \end{bmatrix} q \equiv \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

(iv) Compute the corresponding impulse vector $a = \tilde{A}q$. Check whether it satisfies the sign restrictions for the factor of interest. If it does, compute impulse responses for all factors and save them. If it doesn't, discard q and restart the procedure at step (i). If a fulfils the sign restriction with reversed sign, multiply it with minus one and save it.

This algorithm is repeated until a total of $n = 250$ combinations of parameter draws and vectors q are found that satisfy the sign restrictions. Using the corresponding draw $\Lambda^{(n)}$ of the factor loading matrix, I compute impulse responses for individual macro variables according to (4.10). I finally compute the median estimate and the 16% and 84% quantiles from these impulse responses and report the results in Section 4.5.

I also consider orthogonal impulse vectors for the three yield factors. To obtain these, I extend the above algorithm to incorporate an extra loop over the three factors. Precisely, the algorithm is first carried out for the level factor. Denote the resulting vector of unit length q_1. Second, I search a vector q_2 of unit length for the slope factor that is orthogonal to q_1. This additional restriction can easily be incorporated replacing the zero restriction in step (iii) with

$$\begin{bmatrix} \tilde{A}_{(i)} \\ \tilde{A}_{(j)} \\ q_1' \end{bmatrix} q_2 \equiv \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

If the resulting impulse vector $a(q_2)$ satisfies the sign restriction for the slope factor, it is retained and the procedure is carried out for the curvature factor.[56] Otherwise, the entire procedure is started again for the level factor.

Penalty Function Approach

In addition to the "pure" yield curve surprises identified with the procedure described above, I also study surprise changes which imply a strong reaction of one of the three yield factors and at the same time allow for a small contemporaneous reaction of the remaining two components. This is achieved using a penalty function approach that draws on previous work by Uhlig [2005] and Mountford and Uhlig [2005]. Precisely, I identify a "strong" level surprise as an impulse vector a that minimizes the function (4.14) which I restate here for convenience:

$$\Psi_i(a) = \sum_{h=0}^{H} \left[\gamma_i \frac{r_{a,i}(h)}{\sigma_i} \right] + \sum_{h=0}^{H_2} \left[\sum_{j \neq i} \gamma_j \left(\frac{r_{a,j}(h)}{\sigma_j} \right)^2 \right]$$

where i defines "level", j the remaining two yield factors, γ the weights attributed to the different factors, and H and H_2 the time horizons over which the responses shall be restricted. Computationally, I perform the minimization using the MATLAB function *fminsearch* to which I hand over the value of Ψ_i for a given a, where $a = \tilde{A}q$ as above and where q is a 7×1 vector of unit length, computed using the

[56] Orthogonality of the curvature shock to the level and slope shock then obviously implies two additional zero restrictions.

parameterization (4.12). That is, I search for the set of angles $(\alpha_1, \ldots, \alpha_6)$ which minimize the function Ψ_i. To initialize the procedure, I randomly draw starting values from a uniform distribution with support $[0, \pi]$. Moreover, as discussed in the text, I set $\gamma = [-1, 100, 100]$. Hence, positive responses of the factor of interest are rewarded while any responses of the remaining two factors that deviate from zero are penalized with a much stronger weight. I further set $H = 5$ and $H_2 = 0$, i.e. the response of the factor of interest is maximized over the first six months after the shock and the response of the remaining two components is restricted to be close to zero on impact.

"Typical" Yield Curve Shocks

As a final exercise, I study impulse responses of key economic variables subsequent to a "typical" yield curve shock, i.e. a shock that explains most of the variance of the one-step ahead forecast errors of the three yield curve factors. My approach to identifying this shock has been described in detail in Section 4.4.2. In the following, I discuss some implementational questions related to this method.

As stated above, my identification method requires a decomposition of the variance-covariance matrix Ω into blocks corresponding to the macro and the yield factors, i.e.

$$\Omega = \begin{pmatrix} \Omega_{xx} & \Omega_{xy} \\ \Omega'_{xy} & \Omega_{yy} \end{pmatrix}$$

where Ω_{xx}, Ω_{xy}, and Ω_{yy} are of dimension $k_x \times k_x, k_x \times 3$, and 3×3, respectively. My approach implies finding a matrix

$$A = \begin{pmatrix} A_{xx} & A_{xy} \\ 0 & A_{yy} \end{pmatrix}$$

with A_{yy} and A_{xy} given by

$$\begin{aligned} A_{yy} &= VD^{1/2} \\ \text{and} \quad A_{xy} &= \Omega_{xy}(A'_{yy})^{-1}. \end{aligned}$$

D and V denote the diagonal matrix of eigenvalues and the matrix of corresponding eigenvectors of Ω_{yy}, respectively. The "typical" yield curve shock is given by the column of A that corresponds to the largest eigenvalue of Ω_{yy}. To obtain estimates of impulse responses subsequent to this shock, I perform this decomposition for 250 randomly selected draws from the saved posteriors of Φ and Ω.

6.4 Technical Appendix to Chapter 5

6.5 Constructing monthly time series for real GDP

6.5.1 Interpolation

A variety of different interpolation methods has been suggested in the literature. While some methods estimate higher frequency representations of a low frequency time series on the basis of a time series model, others explicitly take into account the information in related series and thus perform interpolation on the basis of a regression model. Here, we focus on this second class of models since we would like to derive monthly estimates of real GDP using the information in economically related time series which are available at the monthly frequency.

A very prominent and often applied interpolation model that makes use of related high frequency information is the method suggested by Chow and Lin [1971]. These authors assume the high frequency observations of the series to be interpolated as being generated by a linear regression model in the related series with first-order autocorrelated residuals. Depending on the time series properties of both the interpoland and interpolator variables, however, different specifications might be more appropriate. For example, Fernandez [1981] suggests a regression model in first differences to take account of potential non-stationarity of the data. A somewhat more general formulation has been used by e.g. Gregoir (1995) and Mitchell et al. [2005] who suggest to perform the interpolation using dynamic regression models, i.e. they incorporate lagged observations of the interpoland in the regression equation.[57]

Since there is no a priori criterion to decide upon the superiority of any of these approaches, we derive here a unified framework which nests a few of the prominent interpolation methods. This allows us to gauge the relative performance of different models for a given set of series and we will choose the one that is most appropriate on the basis of likelihood ratio tests. Following the work by Bernanke,

[57] For a more exhaustive overview on different interpolation methods, the reader is referred to the nice reviews provided in Di Fonzo [2003] and Proietti [2004], respectively.

Gertler, and Watson [1997] and Proietti [2004], we cast the models in a state-space setup in which it is particularly straightforward to handle the aggregation restrictions implied by the interpolation problem. We will now describe the most general interpolation method considered and then show how different approaches suggested in the literature can be obtained by imposing simple restrictions on individual parameters of the model. Consider the following dynamic regression model

$$\begin{aligned}(1-\phi(L))y_t &= x_t'\beta + u_t,\\ u_t &= \rho u_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0,\sigma^2),\end{aligned}$$

where y_t is the high frequency observation of the variable to interpolated, $\phi(L)$ is a lag polynomial of order p,[58] x_t are the time t observations of a set of related series, and u_t is the regression residual which is assumed to follow an AR(1) process.

We assume here that the quarterly GDP figures are the average of the unobserved three consecutive monthly observations. Hence, defining the quarterly indicator variable y^+ as

$$y^+ = (0 \;\; 0 \;\; y_3^+ \;\; 0 \;\; 0 \;\; y_6^+ \;\; 0 \;\; 0 \;\; y_9^+ \ldots)',$$

we obtain the following measurement equation:

$$\begin{aligned}y_t^+ &= \frac{1}{3}\sum_{i=0}^{2} y_{t-i}, \quad t = 3,6,9,\ldots\\ y_t^+ &= 0 \quad \text{otherwise.}\end{aligned}$$

Notice that there is no error term in this equation since the mean of three consecutive months shall *exactly* equal the quarterly observation. Moreover, the Kalman filter proves particularly useful in such a setup since it can easily handle missing observations by letting the Kalman update be zero in the periods where no new

[58] For simplicity, we assume $p = 1$ since otherwise the number of different models to compare to one another would be quite large. Unreported results based on higher order values of p showed that in most cases, higher order autoregressive parameters were insignificant.

information becomes available. Cast in state-space form, the model is

$$y_t^+ = H_t' \ \xi_t \tag{6.29}$$

$$\xi_t = \begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ u_t \end{pmatrix} = \begin{pmatrix} \phi & 0 & 0 & \rho \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \rho \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ u_{t-1} \end{pmatrix} + \begin{pmatrix} x_t'\beta \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ 0 \\ 0 \\ \epsilon_t \end{pmatrix} \tag{6.30}$$

where

$$H_t' = \left\{ \begin{array}{ll} [\frac{1}{3}\ \frac{1}{3}\ \frac{1}{3}\ 0] & \text{for } \ t = 3, 6, 9, \ldots \\ [0\ 0\ 0\ 0] & \text{otherwise} \end{array} \right\}$$

Notice that due to the limited data availability in the Euro area, for our purpose of constructing monthly GDP series for all member countries of the Euro area, the interpolation method needed to be generalized to incorporate the possibility of using also annual data for interpolation. This is easily done by letting the indicator variable contain observations in annual frequency. As a matter of course, the measurement equation has to be adapted accordingly.[59]

We now briefly show how different interpolation models suggested in the literature can be obtained by fixing either the ϕ or the ρ parameter in the above model. Chow-Lin (1971) suggest a regression model without lagged dependent variables, but autoregressive errors, hence the Chow-Lin model obtains by fixing ϕ to be zero and by letting ρ be estimated freely. Fernandez [1981] suggests a model in first differences to take account of nonstationarity in the data. As the reader will easily notice, this model is obtained by letting the regression residuals be a random walk, i.e. $\rho = 1$ and $\phi = 0$. Note also that one can generalize this model to become a dynamic regression model in first differences by allowing ϕ to be nonzero. As noted above, Mitchell et al. [2005] suggest a dynamic regression model with IID errors, i.e. they let ϕ be nonzero, but have $\rho = 0$. Again, this model can be generalized to have autocorrelated residuals.[60] Table 6.1 sum-

[59] The countries for which the adapted algorithm had to be used were Belgium, Greece, Ireland, Luxembourg, and Portugal. Since those five countries only have a total weight of 6.7 % in our series of Euro area GDP, the uncertainty introduced by performing annual to monthly interpolations is rather small.

[60] It is important to note that the model we refer to here effectively is only a simplified version of the model suggested by Mitchell et al. [2005] who in addition allow the dependent variable to be stated in logarithms and also allow lagged observations of the related series to enter the regression.

Table 6.1: **Parameter Restrictions in the Interpolation Model**

This table summarizes the parameter restrictions that have to be imposed on the model (6.29)-(6.30) in order to obtain a particular type of interpolation model.

Model	ϕ	ρ
Static model in levels with IID residuals	0	0
Static model in levels with AR(1) residuals (Chow-Lin)	0	free
Static model in 1st differences with IID residuals (Fernandez)	0	1
Dynamic model in levels with IID residuals (Mitchell et al)	free	0
Dynamic model in 1st differences with IID residuals	free	1
Dynamic model in levels with AR(1) residuals	free	free

marizes the different interpolation methods and their corresponding parameter restrictions.

In our interpolation exercise, we estimate for all countries the six models summarized above via Maximum Likelihood using the Kalman filter. We then perform country by country a set of bilateral likelihood ratio tests to discover whether the imposed restrictions are borne by the data or not and select the most appropriate model accordingly. We then aggregate the corresponding series using the method described below to obtain our benchmark series of monthly real GDP for the Euro area.

To assess the quality of interpolation, we follow Bernanke et al. [1997] by using R^2 measures of fit. Denoting $y_{t|T}$ the expected value of monthly GDP in period t conditional on the estimated model parameters and the full information set, this measure of fit is given by

$$R^2_{levels} = \frac{Var(y_{t|T})}{Var(y_{t|T}) + Var(u_{t|T})}.$$

As we will see below, when both the interpoland and interpolator variables are upward trending, this measure of fit will be very close to unity in most cases. Hence, it appears more informative to report the R^2 in first differences:

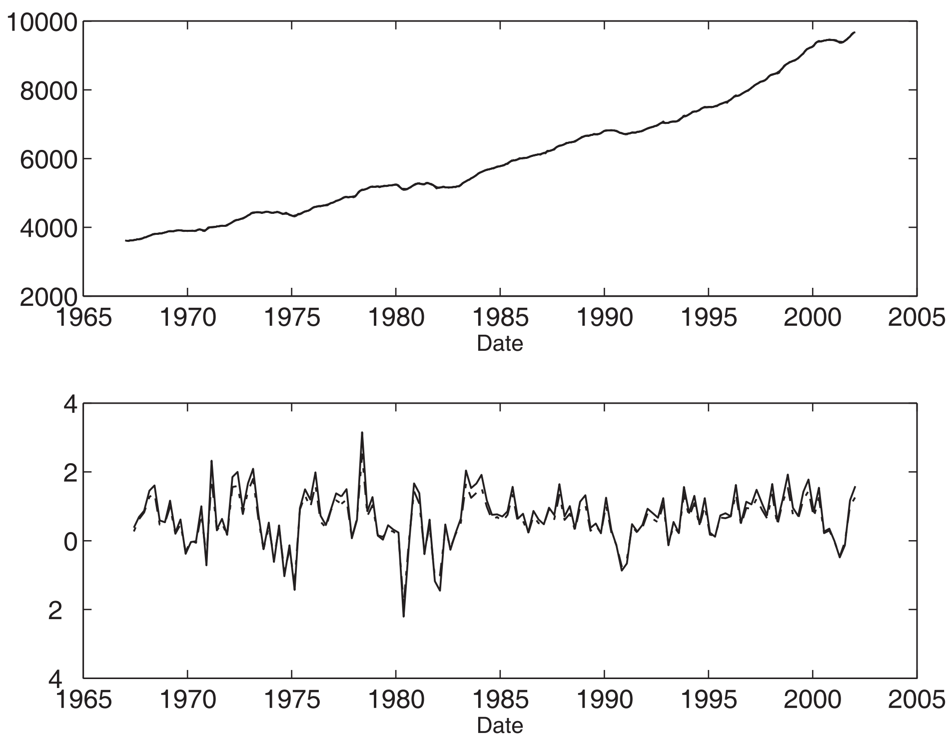
$$R^2_{diffs} = \frac{Var(\Delta y_{t|T})}{Var(\Delta y_{t|T}) + Var(\Delta u_{t|T})}.$$

6.5.2 A monthly time series for real US GDP

This appendix documents the monthly real GDP series from 1967:1 to 2002:09, used for the US in Section 5.3.1. It is based on time series for quarterly real GDP, industrial production, the total civilian employment, and real disposable personal income, which have all been obtained from the Federal Reserve Bank of St. Louis web site. The interpolation is done using the procedure described above. According to the results of bilateral likelihood ratio tests, the best interpolation method for the set of series used is the Chow-Lin model, i.e. the static model in levels with autocorrelated residuals. Values of 0.99 and 0.58 for R^2_{levels} and R^2_{diffs} indicate a good overall interpolation quality. Figure 6.1 provides a plot of the resulting series and a comparison of the implied quarterly growth rates with the actual quarterly growth rates.

Figure 6.1: **Constructed Monthly Series for US Real GDP**

This figure plots our monthly series of US real GDP, based on the four time series GDP96, INDPRO, CE16OV, and DSPIC96, which have all been obtained from the Federal Reserve Bank of St. Louis web site. The interpolation is done using the procedure described above. The upper panel shows the level, and the lower panel the the quarterly growth rates of the original and the interpolated series, respectively.



6.5.3 A monthly time series for real Euro area GDP

Difficulties

Official data covering the Euro area as a whole only exist from 1991 on. Hence, to obtain Euro area aggregates that cover a longer time span, one has to perform some aggregation of individual countries' real GDP series. However, since exchange rate changes have to be taken into account in the pre-Euro period, aggregation of real GDP series across the Euro area member countries is not a trivial task. Competing methods with different merits and shortcomings have been proposed in the literature, and the choice of an appropriate aggregation method seems largely to depend on the requirements one wants it to fulfil. Two often cited references for constructing Euro area aggregates from the individual countries' series are Fagan et al. [2001] and Beyer et al. [2001]. To be used for estimation in the area-wide econometric model of the ECB, Fagan et al. [2001] have constructed a dataset of quarterly Euro area aggregates covering the period from 1970q1 to 2000q4 and including a series of real GDP.[61] They adopt an aggregation method with fixed weights that are computed as the countries' respective shares in total GDP at market prices in 1995. Beyer et al. [2001] propose an aggregation method with time-varying weights that are computed on the basis of exchange rates for converting into a common currency (i.e. the ECU in applications to the Euro area). The authors claim their method to be more general than the one adopted by Fagan et al. [2001]. However, it only delivers estimates of euro area aggregates from 1979 onwards when the European monetary system was constituted. Beyer et al. [2001] provide aggregated Euro area time series for real GDP, nominal GDP, and M3 over the post-1979 period.

The availability of historical quarterly real GDP time series varies considerably across the Euro area member countries. While, for example, real GDP data for Italy is available from 1960 onwards, the equivalent time series for Ireland only covers the post 1997 period. For most of the Euro area countries, chained indices of GDP volume are available over longer time spans than real GDP series. Since volume indices can be directly transformed into 'constant price' level data, we

[61] The authors note that to construct the dataset they have used data from different sources, some of which are not publicly available. Further, when only annual data were available, quarterly time series were constructed by means of some interpolation method similar to the one used in Chapter 5.

use these to generate our aggregate monthly time series of real GDP for the Euro area on the basis of which we then perform the business cycle dating exercise. Yet, since such series are not available for all Euro area countries on a quarterly basis from 1970 onwards, annual series had to be used for some countries, namely Belgium, Greece, Ireland, Luxembourg, and Portugal.

The availability of related monthly series that can be used for interpolating quarterly real GDP is also very limited. While monthly series for industrial production and are available from 1970 onwards for all Euro area member countries except for Ireland, additional variables that are potentially useful for interpolating real GDP are rather scarce. For some countries, a chained index of real retail sales is available from the OECD. For others, if available, monthly employment or export series have been used as additional related series.

The Approach Employed

This section describes our approach to constructing a monthly real GDP series for the Euro area subject to the requirements and limitations mentioned above, especially the problem of data availability for the individual member countries.

1. Interpolation of the individual countries' GDP volume series via the method described above, using industrial production, and, if available, real retail sales, and/or employment as related series. The instrumental variables have been obtained from the OECD and the IMF database, respectively, the chained indices of GDP volume are from the OECD database. All country data is seasonally adjusted before aggregating.[62] Table 6.2 summarizes for all countries the set of related series that have been used for interpolation, the method that has been found to perform best, as well as R^2 statistics as measures of interpolation quality.

2. Next, we compute a weighted average of the interpolated GDP volume series using the so-called "index method" for aggregation (see Fagan and

[62] There clearly is a potential sensitivity of the dating outcome with respect to the seasonal adjustment method employed. Lommatzsch and Stephan [2001] study this issue in detail and find that for quarterly Euro area real GDP series, the dating of the classical cycle is almost completely unaffected by the choice of seasonal adjustment method. Although we expect this issue to be more relevant for the monthly dating exercise that we perform, it is not the focus of Chapter 5 to study the sensitivity of our results to different seasonal adjustment methods. We instead rely on the seasonal adjusted data from official sources to make our procedure as transparent as possible.

Table 6.2: **Interpolation Specifications of Individual GDP Series**

This table summarizes for each country the particular specification used for the interpolation of quarterly and annual GDP volume series into monthly series. The third to sixth column report the related series used for interpolation, the goodness of fit, and the weights in the aggregate series corresponding to the countries' shares in total Euro area GDP in 1995.

Country	Interpolation Method	Rel. Series	R^2_{levels}	R^2_{diffs}	w_i(%)
Austria	Dynamic, levels, AR(1)	IP, Empl	0.99	0.49	3.0
Belgium	Static, 1st diffs, IID	IP, Rsal	0.99	0.89	3.6
Finland	Static, levels, AR(1)	IP, Rsal	0.99	0.72	1.7
France	Static, levels, AR(1)	IP	0.99	0.69	20.1
Germany	Static, levels, AR(1)	IP, Rsal	0.99	0.81	28.3
Greece	Static, levels, AR(1)	IP, Rsal	0.98	0.89	2.5
Ireland	Dynamic, levels, IID	Rsal, Expts	0.99	0.24	1.5
Italy	Static, levels, AR(1)	IP	0.99	0.57	19.5
Luxembourg	Dynamic, levels, IID	IP, Empl	0.99	0.13	0.3
Netherlands	Static, levels, AR(1)	IP, Rsal	0.99	0.70	6.0
Portugal	Static, levels, AR(1)	IP	0.99	0.69	2.4
Spain	Static, levels, AR(1)	IP	0.99	0.55	11.1
					100.0

Henry 1998). According to this method, the log level index for aggregate monthly GDP is given by

$$log(Y) = \sum_{i=1}^{N} w_i log(Y_i).$$

We use the weights provided by Fagan et al. [2001] in their latest update of the ECB's area-wide model dataset for the aggregation.[63]

Since the OECD's GDP volume series for unified Germany only starts in 1991, we have used the West-German series as the historical German series, rescaled to

[63] To see whether the weighting scheme used for aggregation has an impact on the business cycle dating results, we have also constructed an aggregate series using time-varying weights computed as linear interpolations of the annual shares of total GDP at market prices. This series has a peak in 1975:5 instead of 1975:4, all other turning points being equal. Moreover, it exhibits an additional peak in 2001:5. However, since this method of computing time-varying weights is unusual in the literature, we do not rely on this series for the dating exercise. Notice that the OECD's methodology of constructing international area aggregates with time-varying weights for volume indices requires data on the corresponding value series (see OECD 2002 and Schreyer 2001). However, as Schreyer [2001] notes, in case such information is missing, value-added shares at exchange rates or PPPs of a fixed base-year should be used. This is exactly the approach adopted here. As already note above, we could not adopt the aggregation method suggested by Beyer et al. [2001] since this approach can only be used for constructing aggregates in the post 1979-period.

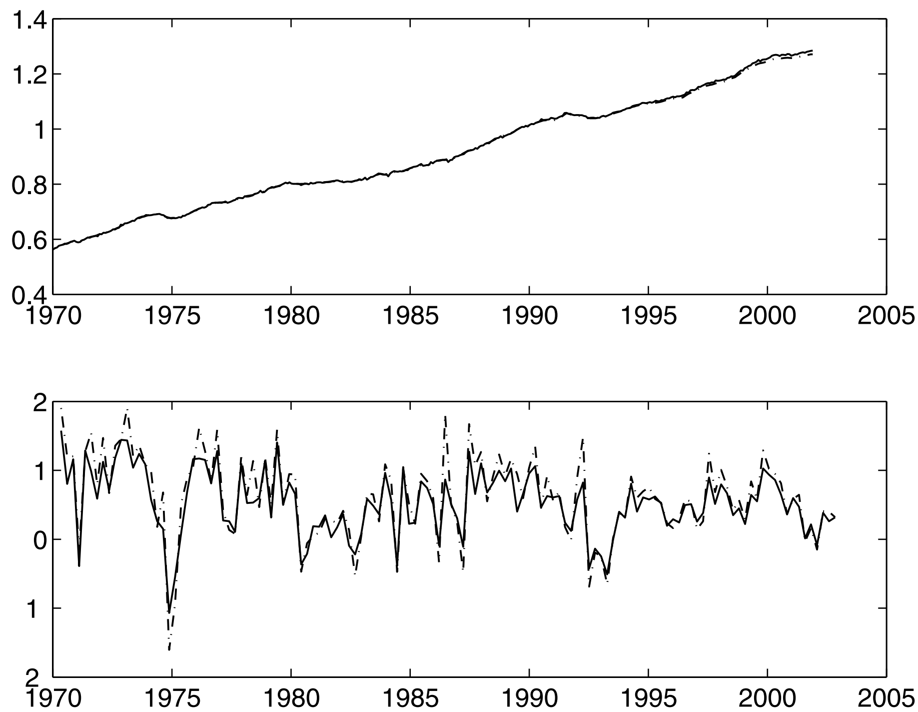
the whole German series by multiplying it with the ratio of the two series in the first quarter of 1991. This is the approach that has also been used by FHM for the construction of their area-wide model dataset.

Result

The resulting time series is available from the authors upon request. A visual comparison to the latest update of the time series by Fagan et al. [2001] is given in Figure 6.2. The upper panel plots our series (solid) against the FHM series (dash-dotted) in levels, whereas the lower panel plots the quarterly growth rates of both series. Obviously, our monthly series is close to the quarterly series, with a slightly more jagged appearance (as desired) due to the interpolation using related series. On the other hand, the FHM series exhibits slightly more volatile quarterly growth rates.

Figure 6.2: **Our Monthly Series vs the Series by FHM**

This figure compares our monthly series of Euro area real GDP (solid) to the quarterly series of Euro area real GDP constructed by Fagan et al. [2001] (dash-dotted). The upper panel shows the levels and the lower panel the quarterly growth rates of the two series, respectively.



Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfasst und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Emanuel Mönch

28. Juni 2006