Multiple Threshold Regimes and Macroeconomic Predictors for Analyzing and Pricing Interest Rate-Dependent Instruments

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

submitted by

Kameliya Filipova

from

Bulgaria

Approved on the application of

Prof. Francesco Audrino, PhD
and
Prof. George Tauchen, PhD

Dissertation no. 4062

Adag Copy AG, Zürich, 2012

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

St. Gallen, March 6, 2012

The President:

Prof. Dr. Thomas Bieger

To my Family

Acknowledgement

I am indebt to many people who have contributed to this dissertation. First of all, I would like to express my deep and sincere gratitude to my supervisor Professor Francesco Audrino. It has been a privilege and a pleasure to learn about computational statistics and financial econometrics by working with him. He has always been very patient and has always found time to discuss various research issues and ideas with me. Without his guidance and support this dissertation would not have been possible.

I am also deeply indebt with my mentor at Duke University, Professor George Tauchen, for being my co-referee and sponsoring my visit at Duke. Financial support of the Swiss National Science Foundation is gratefully acknowledged.

I would like to thank my colleagues at the Institute: (names in random order) Pirmin Meier, Anselm Ivanovas, Tatiana Dvinyaninova, Yujia Hu, Simon Knaus, Markus Harlacher, Sukma Dewi Erawan, Prof. Daniel Buncic, Fulvio Corsi, Mathias Hasler, Prof. Matthias Fengler, Alexander Melnikov and Prof. Enrico De Giorgi for the interesting and fruitful discussions as well as for the very nice time spend at the Institute and outside it. Many special thanks go to Fadrina Denoth. For me our coffee breaks were the sunshine in those foggy and rainy St. Gallen days. I am blessed to have such good friends.

Finally, I am still looking for words to express my gratitude to my family for their unconditional love and support. I am especially grateful to my sister, the big tiger David, and Sophcho for keeping me down to earth and constantly reminding me that there is nothing more important than the family tree. I owe you a lot.

Contents

List of Tables

List of Figures

Preface

My dissertation is motivated by the limited number of multivariate macro–finance asset pricing models which allow for (multiple threshold) regime shifts. Indeed, after more than 40 years of research on asset pricing, one of the central unresolved problems in the financial literature is the relation between the state of the economy and the prices of financial assets.

The literature exhibits numerous attempts to model and explain the nature of fluctuations in the term structure of interest rates. In light of comprehensive academic work on identifying the sources driving the yield curve, the notion of nonlinear, regime-switching relation appears to be a natural, yet non-trivial, extension. It implies that the impact of different fundamentals on bond dynamics changes over time and goes beyond the simple inflation fluctuations and business cycle variation.

The areas so far unexplored in the literature, pertain to our understanding of the sources of those regime switches, the application of appropriate multivariate stochastic processes, and the pricing of bonds and options, to mention a few. Answering these questions is important for solving practical issues faced by the financial industry, monetary policy regulation and as a contribution to the academic literature.

My contribution to the asset pricing literature evolve around two main themes: *(i)* estimating and forecasting term structure models of interest rates with regime shifts; and *(ii)* modeling and explaining the relationships between interest rates and various macroeconomic and financial fundamentals.

The dissertation consists of three essays. The goal of the first chapter is to contribute new empirical evidence to the various economic sources driving the US yield curve. To this end, Francesco Audrino and I present a methodology to build and estimate a discrete-time regime–switching model for the term structure dynamics over time. We allow the conditional dynamics of the yield at different maturities to change in reaction to past information coming from several relevant predictor variables. We consider both endogenous, yield curve factors and exogenous, macroeconomic factors as predictors in our model, letting the data themselves choose the most important variables. We find clear, different economic patterns in the local dynamics and regime specification of the yields depending on the maturity. Moreover, we present strong empirical evidence for the accuracy of the model in reproducing various stylized facts and predicting out-of-sample the yield curve in comparison to several alternative approaches.

The choice of the suboptimal (linear) term structure modeling framework might be just one of the aspects why a direct link between government bonds (or bond risk premia) and macroeconomic fundamentals is hard to detect. Indeed, there are many other possible explanations for this phenomena. Most of them are related to macroeconomic data. First, nowadays econometricians observe hundreds or even thousands of different macroeconomic measures of inflation, consumption, labor markets, housing, etc. Those variables are typically highly correlated within each individual group and they are usually driven by just a small number of latent common factors, impossible to summarize with a few observable series. A second possible explanation is that macroeconomic variables are usually very noisy and imperfectly measured, especially in comparison to the financial fundamentals. Therefore, it is not a surprise that most of the term structure and risk premia variation is captured by financial factors. Last but not least, models themselves are just an imprecise description of the reality. For example, most of the models found in the economic literature, do not take into account the heteroskedastic nature of the macroeconomic fundamentals.

In the second chapter, Francesco Audrino, Fulvio Corsi and I address all of the above mentioned issues by proposing a simple but effective estimation procedure to extract the level and the volatility dynamics of a latent macroeconomic factor from a panel of observable indicators. Our approach is based on a multivariate conditionally heteroskedastic exact factor model that can take into account the heteroskedasticity feature shown by most macroeconomic variables and relies on an iterated Kalman filter procedure. In simulations we show the unbiasedness of the proposed estimator and its superiority to different approaches introduced in the literature. Simulation results are confirmed in applications to real inflation data with the goal of forecasting long-term bond risk premia. Moreover, we find that the extracted level and conditional variance of the latent factor for inflation are strongly related to NBER business cycles.

Guided by the empirical findings presented in the previous two chapters, in the third chapter I develop a new discrete time multivariate regime–switching asset pricing framework, which takes into account the time–varying relation between the short rate and the state of the economy. My approach combines the no–arbitrage restrictions on the cross–section of bonds together with macroeconomic factors that drive bond yields. In contrast to the classical term structure literature, where nonlinearities in the short rate are captured by increasing the number of latent state variables, or by latent regime shifts, in my framework the regimes are governed by thresholds and they are directly linked to different economic fundamentals. Specifically, starting from a simple monetary policy model for the short rate, I introduce a model for the yield curve, which takes into account not only the possibility of regime switches in the behavior of the Federal Reserve, but also agents' beliefs around these changes. In the empirical part, I show the merit of our approach along four dimensions: *(i)* interpretable bond dynamics; *(ii)* superior out-of-sample performance; *(iii)* design of no-arbitrage dynamic term structure model; and *(iv)* accurate short end yield curve pricing. In this way, without resorting to purely latent factors, I am able to capture into a coherent framework the short–term risk that drive bond dynamics. Finally, I take the approach a step further and discuss how it can be successfully applied in modeling stock and bond return comovements.

Chapter 1

Yield Curve Predictability, Regimes, and Macroeconomic Information: A Data-Driven Approach

1.1 Introduction

Beginning with Ang and Piazzesi (2003) the idea of incorporating macroeconomic variables on the top of yield curve factors for modeling bond yields plays a major role in today's term structure literature, giving raise to a new so called macro-finance modeling framework.[1] Despite the various macro-finance modeling strategies proposed in the last years for the U.S. term structure of interest rates dynamics, several questions and controversies are still open. Most of the open issues in the recent macro-finance literature evolve around the central theme of how yields are associated with macro variables. In this chapter we propose an empirical approach to determine the various economic sources driving the U.S. yield curve. We consider both endogenous, yield curve factors and exogenous, macroeconomic factors as predictors in our model, letting the data themselves choose the most important variables.

A common approach in the macro-finance field is to model the short rate dynamics as a function of latent and macroeconomic factors. Yields of other maturities are then derived as risk-adjusted averages of expected future short rates. Thus, the factors driving the short rate contain all the relevant information needed for building and estimating term structure models.[2] Factor analysis of the unconditional variance-covariance matrix of

[1]The macro-finance literature is vast. Important contributions in that area include for example Dewachter, Lyrio, and Maes (2006), Dewachter and Lyrio (2006), Hoerdahl, Tristani, and Vestin (2006), Moench (2008), Joslin, Priebsch, and Singleton (2009), De Pooter, Ravazzolo, and Van Dijk (2007) and Rudebusch and Wu (2008).

[2]This statement is only true under the convention that the market price of risk is also a function of the same state and/or macroeconomic variables driving the short rate dynamics.

yields commonly suggests the number of latent factors needed to explain the cross-sectional dynamics. In addition, standard macroeconomic intuition, a Taylor-style policy rule, is typically used to determine the macro factors entering the short rate equation. Consequently, based on this modeling framework, the same latent and macro variables should help explain not only the short rate but also the entire yield curve dynamics.

However, empirical observations cast some doubt on this view. Short and long maturities are known to react quite differently in shocks hitting the economy. Whereas the central bank (U.S. Federal Reserve) is actively targeting the short rate in order to achieve economic stability (to promote their national economic goals), the long rates tend to be based mainly on real activity, forecasts of inflation and judgements regarding the gap between long-term interest rates and inflation. Many forces are at work in driving the term structure dynamics, and identifying these forces and understanding their impact is of crucial importance.

Almost all the above-mentioned models treat the whole post-war period as a homogeneous sample and do not take into account the possibility of structural breaks in the economy documented in the macroeconomic literature. An exception to this practice is the regime-switching models of interest rates introduced by Hamilton (1988) and - followed for example by Sola and Driffill (1994), Evans and Lewis (1995), Garcia and Perron (1996), and Gray (1996). These papers attempt to build a model that captures the stochastic behavior of the interest rate within a stationary model. Extensive empirical literature (see, for example, Aït-Sahalia (1996b), Stanton (1997), and Ang and Bekaert (2002)) reveals that the regime-switching models better describe the nonlinearities in the yields' drift and the volatility found in the historical interest rate data. More recent works, for example Ang and Bekaert (2002), Bansal and Zhou (2002), Dai, Singleton, and Yang (2007), Bansal, Tauchen, and Zhou (2004), and Audrino and De Giorgi (2007), have managed informally to link the succession of alternating regimes to business cycles and interest rate policies. Rudebusch and Wu (2007) suggest a link between the shift in the interest rate behavior and the dynamics of the central bank's inflation target. Ang, Bekaert, and Wei (2008) develop a regime-switching model to study real interest rates and inflation risk premia by combining latent and macroeconomic factors.

In this chapter we build a regime-switching multifactor model for the term structure dynamics over time in which for every maturity we are able to identify or infer the most important macroeconomic and latent variables driving both the local dynamics and the regime shifts. Our basic framework for the yield curve is a macro-factor model, yet not the usual no-arbitrage factor representation typically used in the macro-finance literature. The methodology adopted in this chapter is mainly motivated by Audrino's (2006) tree-structured model for the short rate. Similarly to Audrino (2006) we employ a multiple threshold model that is able to take into account regime-shifts in the yield curve's dynamics and to exploit both macroeconomic and term structure information. However, in this chapter we do not restrict the local dynamics to follow Cox, Ingersoll, and Ross (1985) process, as in Audrino (2006), but allow for a more flexible data-driven structure selected by a given decision rule. Moreover, we extended the data-sample and the macroeconomic factors used in Audrino (2006).

The technique we propose, have several advantages over the existing modeling strategies. First, it allows us to select the most relevant macroeconomic and latent predictors driving the yields dynamics for each maturity entirely based on the data. This is a huge advantage in comparison to the other techniques, where the number of latent and macroeconomic factors has to be determined a-priori and has to be the same for every maturity. Second, our tree-structured threshold model enables us to identify the possible structural breaks in the yields dynamics in a purely data-driven way. In our estimation technique the regimes are linked to particular macroeconomic and/or monetary policy variables, allowing clear interpretation and separability between monetary policy and macroeconomic changes. Moreover, in contrast to the other term structure models with regime shifts, we do not rely on the assumption that structural changes in the short rate cause also structural changes in the whole yield curve. Instead, we let the data themselves choose the optimal regime structure for each maturity. Third, our approach remains highly competitive in terms of in- and out-of-sample forecasting performance.

We apply our modeling framework to U.S. data. Based on the observed patterns the results can be summarized by three groups: short-, mid- and long-term maturities. Like the monetary policy rules found in the macroeconomic literature,[3] the short rate local dynamics is mainly driven by inflation, real activity, and an autoregressive component. The regimes for the short rate are linked to the level of inflation. The mid-term maturities follow an autoregressive process (AR(1)-GARCH(1,1)), whose behavior is determined by the term structure slope and the level of real activity. In addition, we also find some correspondence between NBER business cycles and our limiting regimes. The long rates capture strong macroeconomic effects. Here the volatility of inflation plays a major role in the threshold structure as well as in the piecewise linear dynamics.

In order to improve the prediction accuracy of our model, we use bagging (short for bootstrap aggregating). In essence, bagging is a variance reduction technique aimed at improving the predictive performance of unstable estimators, especially trees. We compare the out-of-sample forecasting ability of our model to that of several strong competitor models. Using the superior predictive ability (SPA) test of Hansen (2005), we find that such improvements are in most cases statistically significant.

[3]See for example Clarida, Gali, and Gertler (2000) or Taylor (1993), among others.

The remainder of this chapter is organized as follows: Section 1.2.1 and Section 1.2.2 present the modeling framework we use for fitting and forecasting the term structure. Section 1.2.3 describes the techniques we employ for model estimation. The role of bagging is discussed in Section 1.2.4. In Section 1.3 we present the empirical application to U.S. yield data, test our model's ability to reproduce the most important stylized facts, and discuss the results of the out-of-sample forecast. Section 1.4 concludes.

1.2 The Model

This section introduces the modeling framework we use for fitting and forecasting the yield dynamics. To infer the yield curve behavior, we use a model with four distinctive features. First, to capture the cross-sectional dynamics of the yield curve, we employ two latent term structure factors often used in the finance literature, interpreted as level and slope. The two factors usually account for about 95% of the cross-sectional variation of yields.[4] Second, we allow heteroscedasticity in the error term. Since our goal is to build a realistic model for the term structure dynamics over time, this feature is crucial. Third, motivated by the interpretability and the improved forecasting performance of the macro-finance literature in comparison to the pure finance approach, we incorporate macroeconomic variables (such as macroeconomic indicators for real activity and inflation). Fourth, our model accommodates regime-switching behavior but still allows interpretation and clear endogenous regime specification.

1.2.1 The yield-macro model: specification

Let $Y_t = (y(t,n_1),\ldots,y(t,n_T))'$ be a T-dimensional vector of yields with maturities $n_1,\ldots,n_T$ observed at time t and let $\Delta y(t,n_\tau) \equiv y(t,n_\tau) - y(t-1,n_\tau)$ denote the first difference of yields at time t with maturity n_τ. Further, let us assume the following model for the term structure dynamics

$$\Delta y(t,n_\tau) = \mu_{t,n_\tau} + \varepsilon_{t,n_\tau}, \quad \tau = 1,\ldots,T, \tag{1.1}$$

where $\mu_{t,n_\tau} \equiv \mu(\Phi_{t-1,n_\tau};\psi_{n_\tau})$ is a parametric function representing the conditional mean and ε_{t,n_τ} is the error term of the yields' returns with maturity n_τ.[5] More formally, ε_{t,n_τ}

[4]For an extensive survey see for example Litterman and Scheinkman (1991) and Dai and Singleton (2000).

[5]Here, similar to the CIR process for the short rate, we model yields first differences. This choice is mainly motivated by the non-stationarity of the yield process. However, without loss of generality, our model could be easily rewritten and interpreted in terms of yields levels, allowing comparisons with other studies in the literature like, for example, Nelson-Siegel term structure models.

can be decomposed as $\varepsilon_{t,n_\tau} = \sqrt{h(\Phi_{t-1,n_\tau};\psi_{n_\tau})}z_t$, where $(z_t)_{t\in\mathbb{Z}}$ is a sequence of independent identically distributed random variables with zero mean and unit variance, and where $h(\Phi_{t-1,n_\tau};\psi_{n_\tau})$ is the time-varying conditional variance. Above we denoted by Φ_{t,n_τ} all the relevant conditional information up to time t for maturity n_τ. In our application (see Section 3), Φ_{t,n_τ} corresponds to a large number of term structure and macroeconomic variables.

1.2.2 The yield-macro model with regime shifts: specification

In practice, changes in business cycle conditions or monetary policy may affect real rates, expected inflation, as well as other macroeconomic indices and cause interest rates with different maturities to behave quite differently in different time periods, in terms of both level and volatility. An adequate characterization of this stylized fact requires building a term structure model with regime shifts (see for example Ang and Bekaert (2002), Bansal and Zhou (2002), Dai, Singleton, and Yang (2007), Rudebusch and Wu (2007), Bansal, Tauchen, and Zhou (2004), Audrino (2006), and Audrino and De Giorgi (2007)). Rather than following the common Markovian regime-switching approach of specifying the distribution of the regime-switching variable conditionally on the future regime, here, following Audrino (2006) and Audrino and Trojani (2006), the regimes are determined endogenously and represent thresholds partitioning[6] the predictor space into a set of disjoint regions. This approach enables us to determine the current regime based solely on the realization of the state variables, macroeconomic variables, and the threshold structure. This is a major advantage in comparison with the other regime-switching models proposed in the literature, where information about the whole yield curve is needed. In particular, the regime-switching dynamics for the conditional mean and the conditional variance can be written as:

$$\begin{aligned}\mu_{t,n_\tau} &= \sum_{j=1}^{K_{n_\tau}}(\alpha^j_{0,n_\tau} + \alpha^j_{1,n_\tau}\Delta y(t-1,n_\tau) + (\beta^{\mathbf{j}}_{\mathbf{n}_\tau})'\mathbf{x_{t-1}} + (\gamma^{\mathbf{j}}_{\mathbf{n}_\tau})'\mathbf{x^{ex}_{t-1}})I_{[\Phi_{t-1,n_\tau}\in\mathcal{R}^j_{n_\tau}]},\\ h_{t,n_\tau} &= \sum_{j=1}^{K_{n_\tau}}(\omega^j_{n_\tau} + a^j_{n_\tau}\varepsilon^2_{t-1,n_\tau} + b^j_{n_\tau}h_{t-1,n_\tau})I_{[\Phi_{t-1,n_\tau}\in\mathcal{R}^j_{n_\tau}]},\end{aligned}$$

where $\psi_{\mathbf{n}_\tau} = (\alpha^j_{0,n_\tau}, \alpha^j_{1,n_\tau}, (\beta^{\mathbf{j}}_{\mathbf{n}_\tau})', (\gamma^{\mathbf{j}}_{\mathbf{n}_\tau})', \omega^j_{n_\tau}, a^j_{n_\tau}, b^j_{n_\tau}, j = 1,\ldots,K_{n_\tau})$ is a $((m+4)\times K_{n_\tau})$-dimensional vector of the unknown (true) parameters $\tau = 1,\ldots,T$. $I(\cdot)$ is the indicator function and $\mathcal{R}^j_{n_\tau}$ represents a region of the partition $\mathcal{P}_{n_\tau} = \{\mathcal{R}^1_{n_\tau},\ldots,\mathcal{R}^{K_{n_\tau}}_{n_\tau}\}$ of the state

[6] Here we restrict attention to recursive binary partitions. The problem with the multiple splits is that it usually fragments the data too quickly, leaving an insufficient number of observations at the next level down. Moreover, this assumption is not a drawback since multiple splits can easily be achieved by a series of binary splits.

space G_{n_τ} of $\Phi_{t,n_\tau} = \{(\Delta y(t,n_\tau), \mathbf{x}_t', \mathbf{x}_t^{\mathsf{ex}\prime})' \in \mathbb{R}^1 \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}\}$ such that

$$\mathcal{P}_{n_\tau} = \{\mathcal{R}_{n_\tau}^1, \ldots, \mathcal{R}_{n_\tau}^{K_{n_\tau}}\},\ G_{n_\tau} = \cup_{j=1}^{K_{n_\tau}} \mathcal{R}_{n_\tau}^j,\ \mathcal{R}_{n_\tau}^i \bigcap_{(i \neq j)} \mathcal{R}_{n_\tau}^j = \varnothing \quad \tau = 1, \ldots, T.$$

Above we denoted by $(\Delta y(t,n_\tau), \mathbf{x}_t')$ and by $\mathbf{x}_t^{\mathsf{ex}\prime}$ all the endogenous and all the exogenous (macroeconomic) information, respectively, available at time t.

1.2.3 Model estimation

A common approach in the term structure literature to estimating a macro-finance model is to assume that the term structure factors are latent and then to use one-step maximum likelihood estimation. However, this procedure typically requires some additional restrictions due to the multiple likelihood maxima with close-to-identical likelihood values but very different yield decompositions.[7] Consequently, this approach leads to severe estimation difficulties in implementation. Instead, in order to obtain an estimate for the unknown (true) parameters ψ we employ a two-step procedure. As in Ang, Piazzesi, and Wei (2006), the key assumption here is that all factors are observable.

Step 1: Best subset selection

One of the main questions in the term structure literature is how many yield curve factors and/or macro variables should be included in the model. Studies such as Litterman and Scheinkman (1991) and Dai and Singleton (2000) find that, at monthly frequency, the first three principal components account for more than 99% of the cross sectional variation of yields. Applying principal component analysis to our data, we find that the first principal component explains 96.7% of the yield curve variation. Adding the second principal component brings the percentage of yield curve variation to 99.8%.

While just a small number of factors (two or three) are sufficient to model the cross sectional variation of yields, a few questions still remains open. How many factors are needed to build a good model for the time series dynamics? Is there any predictability of macro variables on top of latent factors? If so, how many and which macroeconomic variables should be included in the model? Do these variables always have the same impact on the yields with different maturities? A simple way to answer these questions is to perform best subset selection. Although this statistical dimensionality reduction technique does not impose any economic structure, it helps us identify the most relevant predictors for each maturity.

The main idea behind best subset selection is to retain only a subset of the most informative variables and to eliminate the noise variables from the model. This is achieved

[7] See for example Kim and Orphanides (2005) and Duffee (2002) for discussion of this.

by finding for each number of variables $p \in \{0, 1, 2, \ldots, m\}$ the subset of size p that gives the smallest residual sum of squares. The optimal number of predictors p is usually chosen according to some information criteria. In this chapter we use the Bayesian Schwarz Information Criterion (BIC) since it does not suffer from convergence problems and it is known to provide accurate results in a time series framework.[8]

There are at least four reasons why we favor employing a dimensionality reduction technique rather than including all the possible predictors in the yield curve's local dynamics. *(i)* The first reason is interpretability. With a large number of predictors we would like to identify a smaller subset that contains the most relevant information. *(ii)* The second reason is prediction accuracy. In general, including all possible prediction variables often leads to poor forecasts, due to the increased variance of the estimates in an overly complex model. Therefore, it is crucial to identify the most informative (relevant) predictors and to separate them from the noise variables. By doing so, we reduce the variance of the predicted values: the result is a parsimonious model with better prediction accuracy. *(iii)* Besides the improved forecasting ability, a parsimonious model often helps avoid data-mining problems. *(iv)* Since only a few sources of systematic risk drive the yield curve dynamics, nearly all bond information can be summarized with just a few variables. Therefore, just a small set of variables is needed in order to obtain a close fit to the entire yield curve at any point in time.

Step 2: Regime specification

The second step of our estimation procedure involves regime specification. As stated earlier, the regimes are built as multiple tree-structured thresholds partitioning the predictor space G into relevant disjoint regions. In particular, the partition $\mathcal{P}_{n_\tau}$ for maturity n_τ, $\tau = 1, \ldots, T$, is constructed on a binary tree, where every terminal node represents a partition region $\mathcal{R}^j_{n_\tau}$ whose edges are determined by thresholds. In the general case, the regime classification at time t is based on all the endogenous information $(\Delta y(t-1, n_\tau), \mathbf{x}'_{t-1})$ and the exogenous macroeconomic variables $\mathbf{x}^{\mathbf{ex}}_{t-1}$ up to time $t-1$. As noted above, in contrast to the Hamilton-Markovian framework, here the number of regimes as well as the threshold structure are derived purely from the data.

In this chapter we will mention only the main steps of the binary tree construction and estimation. However, an exact description, illustrative examples, further applications of the algorithm, and discussions about the consistency and reliability of the parameter estimation can be found for example in Audrino and Bühlmann (2001), Audrino (2006), and Audrino and Trojani (2006).

[8]Other possibilities include other information criteria AIC or C_p as well as cross validation.

In short, the estimation procedure involves the following three steps:

(i) Growing a large tree (a tree with a large number of nodes). The threshold selection is based on minimizing the conditional negative pseudo log-likelihood.

The maximal binary tree constructed in *(i)* can be too large and easily lead to overfitting. In order to overcome this problem we proceed by

(ii) Combining some of the branches of this large tree to generate a series of sub-trees of different sizes (varying numbers of nodes);

(iii) Selecting an optimal tree via the application of measures of accuracy of the tree. Analogously to the best subset selection, we chose BIC.

1.2.4 Improving the forecasting ability: Bagging

Bagging, introduced by Breiman (1996), is a variance reduction technique aimed at improving the predictive performance of various estimators as for example classification and regression trees. In general, bagging involves the following steps: *(i)* generate a large number of bootstrap resamples from the data; *(ii)* apply the decision rule to each of the resamples; *(iii)* and average the forecasts from the models selected by the decision rule for each bootstrap sample. Initially bagging was developed for i.i.d. data (see for example Breiman (1996)) and later extended to the time series framework (see, for example, Inoue and Kilian (2004), Audrino and Medeiros (2011)).

The dramatic reduction of the prediction error for a wide range of models with a similar structure has motivated us to use bagging to improve the forecasting performance of our model. In particular, for every maturity, we use the following three-step procedure:

(i) Build a $(n-1)\times(m+1)$ matrix, where the first column corresponds to our response variable Δy_t and the next m columns include all the potential predictors.

$$\{\Delta y(t, n_\tau), \Delta y(t-1, n_\tau), \mathbf{x}'_{t-1,n_\tau}, \mathbf{x}^{\mathbf{ex}\prime}_{t-1,n_\tau}\}, \quad t = 2, \ldots, n.$$

Construct B bootstrap samples denoted by

$$\{\Delta y^*_{(i)}(j+1, n_\tau), \Delta y^*_{(i)}(j, n_\tau), \mathbf{x}^{*\prime}_{(i),j,n_\tau}, \mathbf{x}^{\mathbf{ex}*\prime}_{(i),j,n_\tau}\}, \quad j = 1, \ldots, n-1,$$

where $i = 1, \ldots, B$ by randomly drawing with replacement blocks of rows of length q from the matrix constructed above, where the block size q is chosen in such a way that it captures the dependence in the error term.

(ii) For each bootstrap sample apply the two-step procedure proposed in Section 2.3.1 and Section 2.3.2. Since our two-step approach is purely data-driven, each bootstrap tree will typically involve features different from the original. Note that for every bootstrap sample, the number of predictors, the optimal selection for the local dynamics, the number of terminal nodes, as well as the splitting points may be different. Using the optimal parameters estimated from the i-th bootstrap sample, for $t = 1, \ldots, T_{out}$ compute the conditional mean of the yield process denoted by $\mu^*_{(i)t,n_\tau}$.

(iii) For $t = 1, \ldots, T_{out}$ average the forecasts of the conditional mean

$$\hat{\mu}_{t,n_\tau} = \frac{1}{B} \sum_{i=1}^{B} \mu^*_{(i)t,n_\tau}.$$

1.3 Empirical Results

We start this section with a brief description of the data we use for the empirical part of the chapter. Afterwards, we give an interpretation of the estimated results and test the flexibility of the resulting model. Finally, we compare the forecasting performance of our model to that of several strong competitors.

1.3.1 Data

The *term structure data* consist of monthly data of U.S. Treasury bills with eight different maturities: 3 and 6 months and 1, 2, 3, 5, 7 and 10 years taken from the Fama-Bliss files in the CRSP database. The data cover the time period from January 1960 until June 2005. This is quite a standard data set, a part of which has already been used for example by Audrino (2006), Audrino and De Giorgi (2007), Bansal and Zhou (2002) and Dai, Singleton, and Yang (2007). Table 1.1 provides a fairly detailed description of the data.

Since almost all the cross-sectional term structure information can be summarized in just a few variables associated with the empirical proxies of level, slope, and curvature, we build the endogenous predictors in the following way: we define the level as the 10-year yield and the slope as the difference between the longest (10-year) and the shortest (3-month) maturity in our data set. There are two reasons why we do not build an empirical proxy for the curvature component. First, studies like Litterman and Scheinkman (1991) find that the third principal component accounts for about 2% of the yield curve variation, whereas in our data set it explains less than 0.2% of the variation. Second, in the term structure models the third factor is usually related to heteroskedasticity. Since we model the heteroskedasticity of the error term explicitly, adding a third factor may easily lead to overparametrization. The curvature component also seems unimportant in a broad range

Summary Statistics of Data

	Central moments				Autocorrelations		
	Mean	Stdev	Skew	Kurt	Lag 1	Lag 2	Lag 3
Δ Yield 3M	-0.0020	0.5230	-2.1023	18.0171	0.1517	-0.0661	-0.0291
Δ Yield 6M	-0.0029	0.5156	-1.6226	17.4492	0.1661	-0.0622	-0.0712
Δ Yield 1Y	-0.0028	0.5038	-1.0525	16.0674	0.1630	-0.0986	-0.0863
Δ Yield 2Y	-0.0024	0.4587	-0.6168	10.9402	0.1395	-0.0970	-0.0740
Δ Yield 3Y	-0.0022	0.4199	-0.4246	7.5918	0.1305	0.0884	-0.0748
Δ Yield 5Y	-0.0019	0.3709	-0.2641	4.8015	0.1068	-0.0863	-0.0676
Δ Yield 7Y	-0.0018	0.3426	-0.1923	3.5260	0.0856	-0.0852	-0.0596
Δ Yield 10Y	-0.0014	0.3177	-0.1267	2.7397	0.0642	-0.0771	-0.0533
CPI	4.1503	2.7383	1.4282	1.6165	0.9914	0.9784	0.9639
PPI	3.5834	4.4352	1.0159	1.5395	0.9759	0.9451	0.9153
HELP	82.4983	25.8153	-0.1730	-1.1146	0.9892	0.9787	0.9658
IP	3.1122	4.3763	-0.8378	1.0030	0.9642	0.9093	0.8426
UE	1.2577	15.6301	1.1064	1.2066	0.9560	0.9132	0.8564
CPI.sq	24.7100	34.7598	2.4530	5.8395	0.9930	0.9811	0.9644
PPI.sq	32.4777	60.4715	3.2759	12.5513	0.9614	0.9265	0.8893
HELP.sq	7471.2510	4189.7830	0.2397	-1.0312	0.9886	0.9787	0.9660
IP.sq	28.8038	31.1617	1.6884	3.2806	0.9316	0.8390	0.7311
UE.sq	245.4443	463.4325	3.8516	18.7554	0.9265	0.8375	0.7377
vol.CPI	0.8168	0.5930	1.3497	1.1439	0.9937	0.9774	0.9527
vol.PPI	1.9207	1.3050	1.1334	0.4714	0.9900	0.9656	0.9295
vol.HELP	7.2742	4.3801	0.6408	-0.7335	0.9902	0.9639	0.9228
vol.IP	2.7813	1.8615	1.2099	0.8497	0.9890	0.9657	0.9321
vol.UE	9.9326	6.0700	0.9794	0.0586	0.9889	0.9670	0.9357
slope	1.3401	1.3334	-0.3714	0.1274	0.9438	0.8799	0.8264
Yield 10Y (level)	7.0158	2.4334	0.8696	0.3816	0.9891	0.9770	0.9662

Table 1.1: Descriptive statistics for monthly yields at eight different maturities, and for the yield curve level and slope, where we define the level as the 10-year yield and the slope as the difference between the 10-year and 3-month yields. The inflation measures CPI and PPI refer to CPI inflation and PPI (finished goods) inflation, respectively. We calculate the inflation measure at time t using $log(P_t/P_{t-12})$ where P_t is the (seasonally adjusted) inflation index. The real activity measures HELP, IP, and UE refer to the index of help wanted advertising in newspapers, the (seasonally adjusted) growth rate in industrial production, and the unemployment rate, respectively. The growth rate in industrial production is calculated using $log(I_t/I_{t-12})$ where I_t is the (seasonally adjusted) industrial production index. The conditional volatility measures vol.CPI, vol.PPI, vol.HELP, vol.IP, vol.UE are constructed by using a simple 24-month rolling window approach. By CPI.sq, PPI.sq, HELP.sq, IP.sq, UE.sq we denote the square of the macroeconomic indices CPI, PPI, HELP, IP, UE, respectively. The last three columns contain sample autocorrelations at displacements of 1, 2, and 3 months. The sample period is January 1960 to June 2005.

of macro-finance papers including for example the macro Nelson-Siegel framework studied by Diebold, Rudebusch, and Aruoba (2006).

Macroeconomic data (from January 1960 onward) including some of the leading U.S. indicators of inflation (consumer price index of finished goods (CPI), producer price index of finished goods (PPI)), and real activity (the index of Help Wanted Advertising in Newspapers (HELP), unemployment (UE), the growth rate of industrial production (IP)) are available from the *Datastream* International. In order to ensure stationarity, we transform the monthly macro time series by using annual log differences. We follow Ang and Piazzesi (2003), Audrino (2006) and Diebold, Rudebusch, and Aruoba (2006) in computing the annual growth rates. The caption for Table 1 lists the applied transformations.

An important stylized fact is that shocks in the economy have a significant impact on the dynamics of the yield curve. Therefore, it is intuitive that the term structure dynamics may not only be linked to the level but also to the volatility of the different macroeconomic indicators. In order to exploit this additional macroeconomic information, we construct our measures of conditional volatility of the macro indices by using a simple 24-month rolling window approach. The size of the rolling window is mainly motivated by the degree of smoothness as well as the magnitude of correlation between the yields of different maturities and the conditional volatility of the macroeconomic data. Finally, we also include in our pool of predictors the empirical proxies of the variance of the macroeconomic data just by squaring the different indices. Cross–correlations reveal that there is only a weak correlation among the different macroeconomic variables.

We divide our data set into two parts. We use the data between January 1961 and December 2001 as the in-sample period, whereas the remaining data from January 2002 to June 2005 are left to evaluate the out-of-sample forecasts of the different models.

1.3.2 What is driving the Yield Curve Predictability?

Level dynamics

As discussed in the previous section, using best subset selection we are able to infer the most important variables determining the level dynamics of the yields for every maturity. Although the methodology itself has no economic structure, the consistency between the selected variables via best subset selection and the economic literature is striking. The results are presented in Table 1.2.

Judging from the results presented in Table 1.2 Panel A, we can draw a number of conclusions. Based on the clear pattern the results can be summarized by 3 groups: short, mid-term, and long maturities. Whereas the behavior of the short- and long-term maturities is linked to both endogenous and exogenous variables, the mid-term maturities exploit only

Panel A: Best Subset Selection

Maturity (n_τ)	Δy_{n_τ}	slope	level	PPI	HELP	HELP.sq	vol.PPI	vol.CPI
3M	⋆	⋆	⋆	⋆		⋆	⋆	
6M	⋆	⋆	⋆	⋆		⋆	⋆	
1Y	⋆							
2Y	⋆							
3Y	⋆							
5Y			⋆		⋆		⋆	⋆
7Y			⋆		⋆		⋆	⋆
10Y			⋆		⋆		⋆	⋆

Panel B: Optimal Regime Structure

Maturity	Optimal Regime Structure	# Regimes
3M	$CPI_{t-1} \leq 3.5316$ $CPI_{t-1} > 3.5316$	2
6M	$CPI_{t-1} \leq 3.5316$ $CPI_{t-1} > 3.5316$	2
1Y	$HELP_{t-1} \leq 61.82$ $HELP_{t-1} > 61.82$ and $slope_{t-1} \leq -0.0662$ $HELP_{t-1} > 61.82$ and $slope_{t-1} > -0.0662$	3
2Y	$HELP_{t-1} \leq 61.82$ $HELP_{t-1} > 61.82$ and $slope_{t-1} \leq -0.0662$ $HELP_{t-1} > 61.82$ and $slope_{t-1} > -0.0662$	3
3Y	$HELP_{t-1} \leq 61.82$ $HELP_{t-1} > 61.82$ and $slope_{t-1} \leq -0.0662$ $HELP_{t-1} > 61.82$ and $slope_{t-1} > -0.0662$	3
5Y	$volatilityPPI_{t-1} \leq 0.5935$ $volatilityPPI_{t-1} > 0.5935$	2
7Y	no regimes	1
10Y	$volatilityPPI_{t-1} \leq 0.5935$ $volatilityPPI_{t-1} > 0.5935$	2

Table 1.2: Best subset selection results (Panel A) and optimal regime structure (Panel B) found for every maturity. The variables we take into consideration are the following: the yield's first difference for maturity n_τ, $\tau = 1, \ldots, 8$ denoted by Δy_{n_τ}, yield curve's level, defined as the yield with the longest maturity in our sample (10 years), the yield curve's slope (the longest (10 years) minus the shortest maturity (3 months) in our sample) the macroeconomic indices CPI, PPI, HELP, IP, UE, the square of the macroeconomic indices CPI.sq, PPI.sq, HELP.sq, IP.sq, UE.sq, and the conditional volatility of the above-mentioned macroeconomic indices vol.CPI, vol.PPI, vol.HELP, vol.IP, vol.UE. See text for more details about the model setup and the estimation procedure.

endogenous information.

The linear dynamics for the three- and six-month yields' returns found in our model is very similar to those implied by the standard macroeconomic models. In particular, similar to the Clarida, Gali and Gertler's (2000) framework, which encompasses Taylor's (1993) rule as a special case, the central bank determines the short nominal interest rate (r_{t+1}) depending on the difference between the expected inflation ($\mathbb{E}_t[\pi_{t+1}]$) and the inflation target ($\pi_t^\star$) set by the central bank (which is allowed to be time-varying), on the output gap $\mathbb{E}_t(z_{t+1})$ as well as on the lagged short-term interest rate r_{t-1}. Precisely,

$$r_t = \beta(\mathbb{E}_t[\pi_{t+1}] - \pi_t^\star) + \gamma \mathbb{E}_t(z_{t+1}) + \rho r_{t-1}. \tag{1.2}$$

For the linear dynamics of our resulting model, the combination of the yield curve's level and the inflation variables (level and conditional volatility of inflation (vol.PPI)) might be seen as a measure for the difference between the expected and the target inflation. However, the exact behavior of the two variables, expected inflation and inflation target, is rather difficult to disentangle. The reason is that both are in general unobservable. In addition, the linear combination of the square of the real activity component (HELP), and the slope of the yield curve may be considered as a measure for the expected output gap. The above-mentioned conclusions about the level and the slope of the yield curve are fully in line with the existing macro-finance literature. Examining the correlations between Nelson-Siegel yield factors and a large set of macroeconomic variables, Diebold, Rudebusch, and Aruoba (2006) find that the level factor is highly correlated with inflation, and the slope factor is highly correlated with real activity. Rudebusch and Wu (2008) provide a similar interpretation. They find that the level factor reflects market participants views about the underlying or medium term inflation target of the central bank, whereas the slope factor captures the cyclical response of the central bank aimed at stabilizing the real economy and keeping inflation close to target. Finally, the autoregressive term in our resulting model corresponds to the last term in (1.2), reflecting the Federal Reserve policy to smooth changes in interest rates.

For the mid-term maturities (one-, two- and three-year yields' returns), we find that the linear dynamics is driven only by endogenous information. More precisely, the mid-term yield returns follow an AR(1)-GARCH(1,1) process.

Perfectly in line with the empirical observations, the long-term maturities (five-, seven- and ten-year yields) capture a strong macroeconomic effect. They are linked to the level of the yield curve, the level of real activity (HELP), and the conditional volatility of the two inflation indices CPI and PPI.

Regimes

Similar to the previous subsection, based on the threshold structure, the results could be split into three parts: short-, middle- and long-term maturities. As mentioned above, the regimes for every maturity are determined endogenously, based on our in-sample period between January 1961 and December 2001.

Short-term maturities

For the short-term maturities we find two limiting regimes, characterized by the level of inflation or more precisely, CPI. The results are given in Table 1.3.[9]

SHORT-TERM MATURITIES' PARAMETER ESTIMATES

		3 Months		6 Months	
Optimal Regime Structure	Variable	Coefficient	t-statistic	Coefficient	t-statistic
$CPI_{t-1} \leq 3.5316$	`const`	0.1729	29.5631	0.2101	31.1557
	Δ`y`	0.1826	33.6940	0.2463	29.4939
	`slope`	0.0406	18.9258	0.1205	6.8212
	`level`	0.0262	15.8349	-0.0557	-14.2697
	`PPI`	0.0084	19.0392	0.0092	27.4288
	`HELP.sq`	8e-06	2.2432	0.0000	0.0000
	`vol.PPI`	-0.0580	-31.1357	-0.0909	-69.8259
	ω_{n_τ}	0.0385	27.0689	0.0462	27.0864
	ϵ^2_{t-1,n_τ}	0.1347	4.2985	0.1068	3.5909
	h_{t-1,n_τ}	0.0000	0.0046	0.0006	0.8040
$CPI_{t-1} > 3.5316$	`const`	0.2131	9.5540	0.2541	46.8190
	Δ`y`	0.1202	10.4486	0.1010	19.8693
	`slope`	0.1064	40.2900	0.0849	7.2371
	`level`	-0.0458	-11.6777	-0.0361	-7.3638
	`PPI`	-0.0012	-0.3901	0.0059	0.1370
	`HELP.sq`	4e-06	0.3462	0.0000	0.0000
	`vol.PPI`	-0.0152	-5.8731	-0.0519	-3.3263
	ω_{n_τ}	0.1800	22.7663	0.1598	10.3497
	ϵ^2_{t-1,n_τ}	0.8275	51.1376	0.5685	44.9099
	h_{t-1,n_τ}	0.0077	0.2294	0.2093	11.5955
LB^2_5		7.2069	(0.2057)	6.1919	(0.2880)
LB^2_{10}		15.033	(0.1309)	14.0398	(0.1712)
LB^2_{15}		16.6322	(0.3413)	15.9012	(0.3886)

Table 1.3: Local parameter estimates, optimal threshold structure and related statistics for 3- and 6-month yields from the macro-tree regime-switching model. The sample period is January 1961 - December 2001, for a total of 492 monthly observations. t-statistics are based on heteroskedastic-consistent standard errors. LB^2_i denotes the Ljung-Box statistic for serial correlation of the squared residuals out to i lags. p-values are in parentheses.

[9]Note that the coefficients reported in Table 1.3, Table 1.4, and Table 1.5 are for the yields first differences.

The threshold structure is fully in line with the Federal Reserve's monetary policy, where the short rate is used as an instrument to promote national economic goals. A well-known fact (general monetary policy rule) is that in times of high inflation, the Federal Reserve tends to raise the short end of the yield curve in order to provide economic stability. Therefore, it is not a surprise that the regimes are linked to the level of the leading inflation index CPI. Though our in-sample period encompasses several Fed monetary policy changes with substantial differences in the short rate response to the expected inflation,[10] our resulting model is still valid. The reason for this is that in our model the inflation threshold has an impact mainly on the level of the short rate, whereas the conditional piecewise linear dynamics - especially the linear combination of the yield curve's level, slope, the macroeconomic level of inflation PPI, and the conditional volatility of inflation vol.PPI - captures the fluctuations in the short-term maturities. In other words, the main difference between the conditional means for the two limiting regimes lies in the magnitude of the resulting yield values. This finding is perfectly in line with the existing macro-finance literature. For example, examining the structural impulse responses of their macro-factor model for joint dynamics of the yields, Ang and Piazzesi (2003) document that inflation surprises have large effects on the level of the entire yield curve.

Another interesting finding is that in both regimes, shocks in the economy have an immediate impact on the short-term yields' returns. In periods of moderate to low inflation ($CPI \leq 3.5316$), shocks in the economy have a small but significant impact on the yield dynamics. In the second limiting regime, characterized by moderate to high inflation ($CPI > 3.5316$), the impact of individual shocks is much higher than in the first regime. Note also that in the second regime, the individual impact of shocks in the economy decreases (from 0.8275 for 3-month to 0.5685 for 6-month yield returns), whereas the persistence of the shocks increases significantly (from 0.0077 for 3-month to 0.2093 for 6-month yield returns) with time to maturity.

Based on our in-sample data, we find that the short maturities exploit two regimes. At first glance, this finding seems a little bit intact with the study of Audrino (2006) and Audrino and Medeiros (2011), where using almost the same in-sample period the authors found that the short rate is subject to four and three limiting regimes, respectively. There are at least two possible reasons why we find only two limiting regimes in our study. One reason could be that both studies focus on one-month U.S. Treasury bill rates, whereas the shortest maturity in our data set is three months. A stylized fact is that the yield persistence increases with maturity. Therefore it seems quite natural that the short end

[10]For a discussion of the Federal Reserve policy rules in the different subperiods, see Clarida, Gali, and Gertler (2000). Although the results are not reported here, we have also tested for structural breaks in the economy.

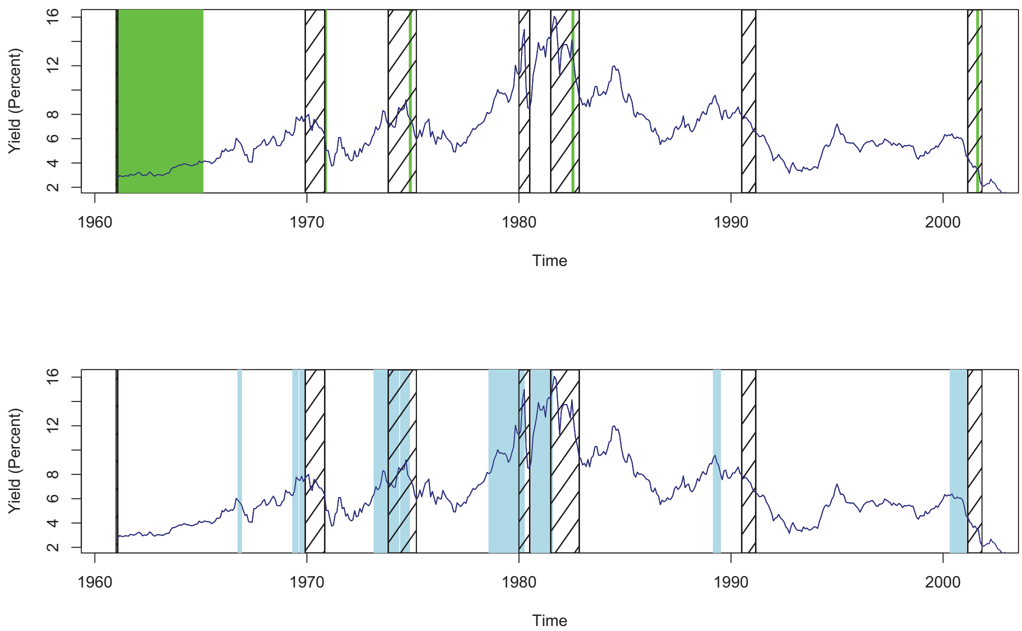


Figure 1.1: The top and the bottom panels plot the one-year yield time series for the period January 1961-December 2001. The gray bars in the top panel overlay periods with low real activity HELP $\leq$ 61.82 as found in Regime 1. The gray bars in the bottom panel overlay periods with medium and high real activity HELP $>$ 61.82 and yield curve slope $\leq$ -0.0662 as found in Regime 2. NBER recessions are indicated by shaded bars. See text for more details.

of the yield curve is subject to more regimes that the long end. A second reason could be that in both Audrino (2006) and Audrino and Medeiros (2011) the local dynamics of the short rate is fixed and follows the classical CIR process. In this chapter, however, we allow for more flexible linear structure using best subset selection on first place and afterwards determining the regimes. This essentially implies that some of the hierarchial structure found in Audrino (2006) and Audrino and Medeiros (2011) enters in a linear way in our modeling framework. In the next section we show that this feature is not a drawback since our model outperforms significantly the model of Audrino (2006).

Mid-term maturities

The threshold structure with three limiting regimes found for the mid-term maturities mainly reflects the yield curve behavior across business cycles. The dependence of the regimes on the real activity index HELP confirms Ang and Piazzesi's (2003) finding that output shocks have a significant impact on intermediate yields. The regime structure and the estimated coefficients are presented in Table 1.4.

Mid-term Maturities' Parameter Estimates

		1 Year		2 Years		3 Years	
Regime Structure	Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
$HELP_{t-1} \leq$ 61.82	const	0.0372	1.6952	0.0240	1.3363	0.0189	1.5250
	Δ y	-0.1259	-0.8400	-0.1606	-1.2175	-0.0027	-0.0243
	ω_{n_τ}	0.0071	1.4674	0.0061	2.2538	0.0032	2.2138
	ϵ^2_{t-1,n_τ}	0.5131	1.1935	0.3814	1.5724	0.2993	1.7977
	h_{t-1,n_τ}	0.1387	0.8706	0.2040	1.0613	0.3263	2.1271
$HELP_{t-1} >$ 61.82 $slope_{t-1} \leq$ -0.0662	const	-0.0568	-0.8287	-0.0468	-0.6951	-0.0159	-0.2608
	Δ y	-0.2591	-1.2252	-0.2654	-1.7572	-0.2788	-1.9412
	ω_{n_τ}	0.0317	1.0229	0.0000	0.0000	0.0000	0.0000
	ϵ^2_{t-1,n_τ}	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	h_{t-1,n_τ}	1.0709	20.5100	1.1434	26.8699	1.1303	32.5502
$HELP_{t-1} >$ 61.82 $slope_{t-1} >$ -0.0662	const	0.0206	1.2551	0.0131	0.6677	0.0131	0.6852
	Δ y	0.0635	1.3306	0.0613	1.2087	0.0405	0.8011
	ω_{n_τ}	0.0040	0.9031	0.0164	1.9815	0.0183	2.2763
	ϵ^2_{t-1,n_τ}	0.0338	2.6079	0.0721	2.9940	0.0607	2.6512
	h_{t-1,n_τ}	0.9161	9.7915	0.7935	11.7520	0.7852	10.9538
LB^2_5		5.3243	(0.3776)	4.5076	(0.4789)	4.1067	(0.5342)
LB^2_{10}		5.8277	(0.8295)	10.7420	(0.3780)	6.4213	(0.7787)
LB^2_{15}		15.8527	(0.3919)	14.1564	(0.5137)	9.3152	(0.8605)

Table 1.4: Local parameter estimates, optimal threshold structure and related statistics for 1-, 2- and 3-year yields from the macro-tree regime-switching model. The sample period is January 1961 - December 2001, for a total of 492 monthly observations. t-statistics are based on heteroskedastic-consistent standard errors. LB^2_i denotes the Ljung-Box statistic for serial correlation of the squared residuals out to i lags. p-values are in parentheses.

The first regime ($HELP \leq 61.82$) essentially encompasses short periods towards or right after the end of recessions with particularly low mid-term yields. The upper panel of Figure 1.1 illustrates this finding.

The second limiting regime is characterized by both a negative slope of the yield curve ($slope \leq -0.0662$) and moderate to high real activity ($HELP > 61.82$). The dependence on the slope is not a surprise, since in general the slope of the yield curve is considered one of the most important forecasters of the short- and mid-term economic growth.[11] This regime structure mainly describes the mid-term yield behavior right before or in the very beginning of recession periods. The bottom panel of Figure 1.1 confirms this finding. The resulting GARCH dynamics for this limiting regime clearly shows that individual shocks have no immediate impact. The estimated coefficient for the autoregressive term in the GARCH dynamics for each of the mid-term maturities in this regime (Regime 2) exceeds one. This non-stationarity in the GARCH model indicates not only high persistence of the individual shocks but also reflects the uncertainty in the economy.

The third regime with moderate to high real activity ($HELP > 61.82$) and in general positive yield curve slope ($slope > -0.0662$) spans more than 70 percent of the in-sample period and reflects the standard mid-term yield curve behavior. In this regime individual shocks in the economy have a small but significant impact. They are also strongly persistent, although less so than those found in the second regime. Here, it is also important to note that the shock persistence in this regime decreases with time to maturity (from 0.9161 for the one-year yield to 0.7852 for the three-year yield).

Long-term maturities

Finally, for the long maturities we find that the regimes are characterized by the conditional volatility of inflation (vol.PPI). Results are reported in Table 1.5.

This threshold structure is fully in line with the macro-finance literature, where the behavior of the long-end of the yield curve is strongly related to inflation (inflation level, volatility of inflation, expected inflation, inflation target, inflation gap, inflation risk premium, etc.). For the first regime we find that it is characterized by low conditional volatility of inflation ($vol.PPI \leq 0.5935$). In this regime the resulting yields are low, reflecting the stability in the economy. Individual shocks have moderate (for the five-year yield) to negligible (for the ten-year yield) impact on the yields' returns, whereas their persistence increases with maturity. The other limiting regime is characterized by moderate to high conditional volatility of inflation ($vol.PPI > 0.5935$). Here the levels of the long-term yields are significantly higher than those found in the other limiting regime. The persistence of

[11] The rule of thumb is that an inverted yield curve (short rates above long rates) indicates a recession in about a year.

Long-term Maturities' Parameter Estimates

Optimal Regime Structure	Variable	5 Years Coefficient	5 Years t-statistic	10 Years Coefficient	10 Years t-statistic
$volPPI_{t-1} \leq 0.5935$	`const`	-0.1540	-0.8477	0.2117	1.1637
	`level`	0.0654	1.0950	-0.0374	-0.7929
	`HELP`	-0.0017	-0.7444	-0.0004	-0.2575
	`vol.CPI`	-0.0643	-0.4163	0.1795	1.1959
	`vol.PPI`	0.0024	0.0131	-0.1450	-0.9841
	ω_{n_τ}	0.0005	0.2769	0.0003	1.9806
	ϵ^2_{t-1,n_τ}	0.6736	2.4245	0.0001	0.0011
	h_{t-1,n_τ}	0.3216	1.7743	0.7778	16.6318
$volPPI_{t-1} > 0.5935$	`const`	0.1331	1.4707	0.1131	1.3419
	`level`	-0.0527	-3.9959	-0.0445	-4.0519
	`HELP`	0.0024	2.3462	0.0020	2.0917
	`vol.CPI`	0.1886	3.2182	0.1551	3.6091
	`vol.PPI`	-0.0678	-2.9253	-0.0520	-3.1801
	ω_{n_τ}	0.0094	1.1868	0.0056	1.6716
	ϵ^2_{t-1,n_τ}	0.1036	2.9457	0.0930	2.3062
	h_{t-1,n_τ}	0.8334	11.0224	0.8543	13.8319
LB^2_5		4.6216	(0.4638)	4.8898	(0.4295)
LB^2_{10}		9.7054	(0.4667)	10.4846	(0.3991)
LB^2_{15}		10.8031	(0.7664)	13.5046	(0.5634)

Optimal Regime Structure	Variable	7 Years Coefficient	7 Years t-statistic
no regimes	`const`	0.0474	1.0659
	`level`	-0.0462	-2.6178
	`HELP`	0.0029	2.5623
	`vol.CPI`	0.1869	2.7582
	`vol.PPI`	-0.0641	-2.6711
	ω_{n_τ}	0.0002	0.0602
	ϵ^2_{t-1,n_τ}	0.1856	2.5730
	h_{t-1,n_τ}	0.8496	13.5133
LB^2_5		21.4143	(0.0007)
LB^2_{10}		39.9147	(0.0000)
LB^2_{15}		57.6575	(0.0000)

Table 1.5: Local parameter estimates, optimal threshold structure and related statistics for 5-, 7- and 10-year yields (in the upper table) from the macro-tree regime-switching model. The optimal resulting structure for the 7-year yield is the global model (without regime shifts). The sample period is January 1961 - December 2001, for a total of 492 monthly observations. t-statistics are based on heteroskedastic-consistent standard errors. LB^2_i denotes the Ljung-Box statistic for serial correlation of the squared residuals out to i lags. p-values are in parentheses.

individual shocks is very high, whereas their immediate impact is comparatively small. For the seven-year yield we were not able to find any optimal threshold structure.

Based on the threshold structure found for each maturity, one may easily conclude that overall the entire yield curve is potentially subject to twelve (two for the short-term, three for the mid-term, and up to two for the long-term maturities) regime shifts. However, due to the mutual dependence among the different thresholds, in reality, the number of regimes is much smaller, since the resulting thresholds (level of CPI, volatility of PPI, slope of the yield curve, and level of HELP) are correlated.

Finally, analogously to Audrino (2006), we analyze the correspondence between NBER business cycles and the regime structure found for each maturity. In particular, we compute the frequency of the regimes in the recessions versus expansions. The results are reported in Table 1.6.

In addition, as in Bansal, Tauchen, and Zhou (2004) and Audrino (2006), we compute correlations between the yield curve's slope, HELP, CPI and NBER business cycles. The absolute correlations between yield curve slope, HELP, CPI, and the NBER indicator are 0.1248, 0.1654, and 0.4452, respectively. Thus, we can once again conclude that the optimal threshold structure we find for each maturity is quite natural.

Stylized Facts

Since we focus explicitly on modeling the yield curve dynamics over time, an important model diagnostics would be to see how well our resulting model is able to fit also the cross-section of yields. In this section we test our model's ability to replicate the most important stylized facts found in the the term structure literature. Here we focus on the following stylized facts, summarized by Diebold and Li (2006): *(i)* the average yield curve is upward-sloping and concave; *(ii)* the fitted model is able to reproduce the variety of yield curve shapes observed through time: upward-sloping, downward-sloping, humped, and inverted-humped; *(iii)* short rates are more volatile than long rates; *(iv)* long rates are more persistent than short rates.

Figure 1.2 and Figure 1.3 provide a graphical representation of the above-mentioned facts.

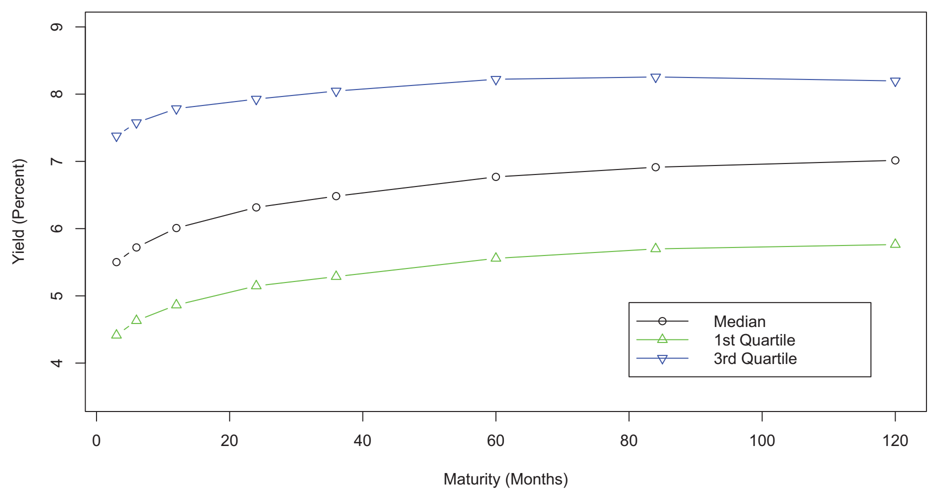
The upper panel of Figure 1.2 shows the average (median) fitted yield curve together with its interquartile ranges. The average upward-sloping form, the concavity, as well as the fact that short rates are more volatile than long rates are apparent. The short end of the yield curve is obviously steeper and flattens with maturity. Based on Figure 1.2, we can easily draw one more conclusion - the distribution of yields around their median is asymmetric with a longer right tail.

Next, Figure 1.3 presents four fitted yield curves for some selected dates. Apparently, Regime Frequency in Recession and Expansions

Maturity	Regime	Frequency		
		total	recession	expansion
3 and 6 months	$CPI \leq 3.516$	$246/492 = 0.500$	$11/73 = 0.152$	$235/419 = 0.561$
	$CPI > 3.516$	$246/492 = 0.500$	$62/73 = 0.850$	$184/419 = 0.439$
1, 2 and 3 years	$HELP \leq 61.82$	$65/492 = 0.132$	$10/73 = 0.137$	$55/419 = 0.132$
	$HELP > 61.82$ & $slope \leq -0.066$	$62/492 = 0.126$	$17/73 = 0.233$	$45/419 = 0.107$
	$HELP > 61.82$ & $slope > -0.066$	$365/492 = 0.742$	$46/73 = 0.630$	$319/419 = 0.761$
5 and 10 years	$vol.PPI \leq 0.5935$	$62/492 = 0.126$	$0/73 = 0.000$	$62/419 = 0.148$
	$vol.PPI > 0.5935$	$430/492 = 0.874$	$73/73 = 1.000$	$357/419 = 0.852$

Table 1.6: Frequency of the different regimes in NBER recessions and expansions for the in-sample period January 1961 - December 2001, for a total of 492 observations. Total (#observations in the regime/total number of observations), recession (# recession observations in the regime/ total number of recession observations) and expansion (# expansion observations in the regime/ total number of expansion observations) frequencies for our model are reported.

Panel A: Median Yield Curve



Panel B: Boxplots Yield Curve

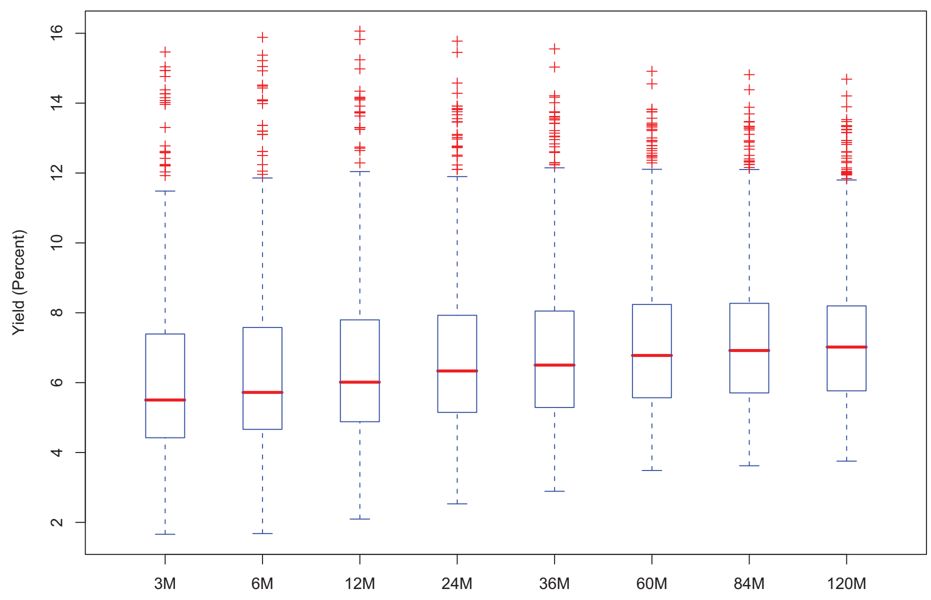


Figure 1.2: Panel A shows the median fitted (data-based) yield curve with interquartile range (25th and 75th percentiles). Panel B presents Boxplots for the fitted (model-based) yields for every maturity. The data span the time period January 1961-December 2001, for a total of 492 observations.

Selected Fitted Yield Curves

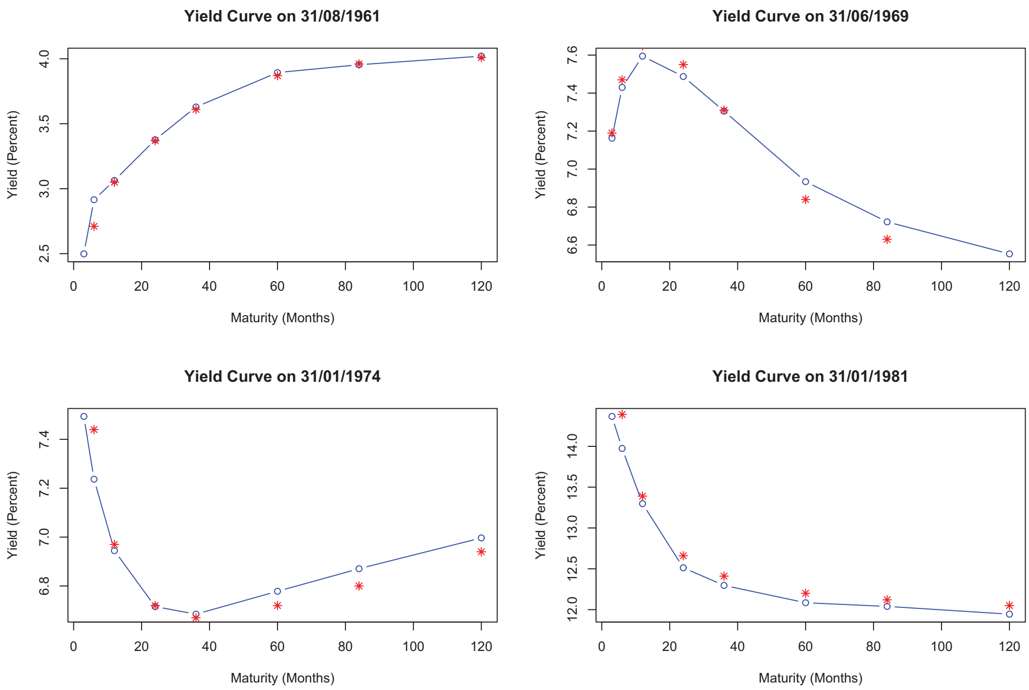


Figure 1.3: Fitted (model-based) yield curves for selected dates (dotted lines), together with actual yields (stars). See text for details.

our model is able to capture the broad variety of shapes the actual yield curve assumes through time: upward-sloping, downward-sloping, humped, and inverted-humped. The model does not provide a perfect fit at any point in time, but its overall match is quite good.

The boxplots presented in the bottom panel of Figure 1.2 show that our model is perfectly in line with the stylized fact that short rates are more volatile than long rates.

The clear linear pattern presented in Table 1.2 Panel A as well as the threshold structure given in Table 1.2 Panel B reflect one additional stylized fact: yields of near maturities are highly correlated, and therefore it is quite natural that the forces moving the short, middle, and long part of the yield curve are one and the same within the three groups, but quite different among them.

1.3.3 Out-of-Sample Forecasting

Apart from the economic linkage and the ability to replicate at least the most important stylized facts, a good term structure model should also be able to provide a good out-of-sample fit. In this section we compare the out-of-sample performance of our model to those of several strong competitors for maturities of 3 and 6 months and 1, 2, 3, 5, 7 and 10 years. In particular, we focus on the following 6 models: *(i)* Random walk; *(ii)* VAR(1) on yields levels; *(iii)* two dynamic specifications of Nelson-Siegel for the yields levels proposed by Diebold and Li (2006); *(iv)* Markovian regime switching model of Gray (1996); *(v)* tree structured regime switching model of Audrino (2006); and *(vi)* the one regime version of our model.

In this chapter, we assess the prediction accuracy of the different models by means of two different measures. In particular, we focus on the mean squared errors (MSE). The measure is given by:

$$\texttt{MSE-mean} = \frac{1}{n}\sum_{t=1}^{n}\left(\Delta y(t, n_\tau) - \hat{\mu}_{t,n_\tau}\right)^2.$$

To improve the prediction accuracy of our model, we use bagging. As stated above, bagging is a machine learning technique aimed at reducing the variance and thus improving the forecasting performance of various estimators such as trees. Applied to our data set, for building the bootstrap samples we use block bootstrapping of Künsch (1989), where we set the block size value q to be equal to 20 and the number of iterations B to be equal to 50.

For completeness, we also apply bagging to all the competitors' models. Apart from Audrino's (2006) model we do not find any significant improvement in the out-of-sample performance of the other models. The reason for this lies in the structure of the modeling framework.[12] The results are presented in Table 1.7.

Comparing the one-month-ahead out-of-sample results of the different models (see Table 1.7), without considering bagging, we find that our model has overall good performance at all eight maturities. Matters improve dramatically, once we apply bagging. The SPA p-values, based on all fifteen model specifications, presented in Table 1.8 reveal that the forecasts yield from the bagged versions of our model are significantly better than almost all of the alternative approaches. Based on the multiple comparison test, we cannot conclude that our model significantly outperforms the random walk.[13] However, a direct comparison between the bagged version of our model and those of the random walk via Diebold and Mariano (1995) test indicates that we are able to beat the random walk at least for the short- and the long-term maturities. For completeness, in the Appendix we provide in–sample results for our models.

1.4 Conclusion

In this chapter we present a methodology to build and estimate a discrete-time regime-switching model of interest rates that incorporates latent and macroeconomic factors and

[12]Bühlmann and Yu (2002) have conducted extensive research on this topic.

[13]Several studies (see, for example, Duffee (2002) and Ang and Piazzesi (2003)) have documented that beating the random walk is indeed a challenging task, especially over short horizons.

Out-of-Sample Results

	Out-of-sample MSE							
Maturity	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.0124	11.5538	0.2777	0.3280	0.0147	0.0150	0.0174	0.0293
6M	0.0216	12.9953	0.1844	0.2254	0.0171	0.0248	0.0363	0.0504
1Y	0.0352	0.0322	0.0853	0.1100	0.0346	0.0448	0.0395	0.5491
2Y	0.0949	0.0901	0.1280	0.0934	0.0901	0.0961	0.0910	0.1428
3Y	0.1179	0.1178	0.1648	0.1366	0.1236	0.1261	0.1166	0.1481
5Y	0.1346	0.2294	0.2153	0.2316	0.1345	0.1313	0.1269	0.1217
7Y	0.1286	0.1286	0.2796	0.4171	0.1219	0.1185	0.1261	0.1441
10Y	0.1014	0.1151	0.0996	0.1108	0.1035	0.1024	0.0974	0.1205

	Out-of-sample MSE for the bagged models						
Maturity	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	VAR(1)	Audrino Tree	Gray's RS
3M	0.0068	0.0820	0.5781	0.6626	0.1315	0.0128	0.1440
6M	0.0099	0.0368	0.4329	0.5083	0.1135	0.0196	0.0798
1Y	0.0284	0.0653	0.2420	0.3014	0.1486	0.0357	0.3754
2Y	0.0824	0.0905	0.0845	0.1253	0.2045	0.0887	0.3112
3Y	0.1149	0.1449	0.1550	0.1439	0.2295	0.1142	0.2941
5Y	0.1242	0.1434	0.1679	0.1323	0.1905	0.1230	0.2607
7Y	0.1155	0.1155	0.4108	0.6082	0.1510	0.1116	0.2707
10Y	0.0918	0.0995	0.1230	0.1731	0.1204	0.0951	0.2093

Table 1.7: Results of out-of-sample 1-month-ahead forecasting using eight models: Our macro-tree model (`Macro Tree`), Random walk (`RW`), VAR(1) on yields levels (`VAR(1)`) two dynamic specifications of Nelson-Siegel for the yields levels proposed by Diebold and Li (2006) (`NS AR(1)` and `NS VAR(1)`), Markovian regime switching model of Gray (1996) (`Gray's RS`), tree structured regime switching model of Audrino (2006) (`Audrino Tree`) and the one regime version of our model (`Best Subset`). The results are based on the out-of-sample period January 2002 - June 2006.

Superior Predictive Ability Test

Maturity	Out-of-Sample SPA test for the MSE							
	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.1006	0.0000	0.0000	0.0000	0.2474	0.0687	0.0000	0.0068
6M	0.0944	0.0000	0.0000	0.0000	0.4006	0.0880	0.0000	0.0000
1Y	0.5676	0.5671	0.1427	0.0000	0.4637	0.2033	0.0836	0.0000
2Y	0.1071	0.3433	0.3635	0.3652	0.5876	0.5519	0.4364	0.0187
3Y	0.4202	0.4341	0.4520	0.6570	0.4399	0.4442	0.5854	0.0371
5Y	0.2370	0.0124	0.1720	0.4418	0.5135	0.6211	0.6414	0.6219
7Y	0.4120	0.4120	0.0000	0.0000	0.5557	0.5475	0.3730	0.1071
10Y	0.4441	0.0446	0.5197	0.0565	0.3647	0.2809	0.6092	0.0000

Maturity	Out-of-Sample SPA test for the MSE for the bagged models						
	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	VAR(1)	Audrino Tree	Gray's RS
3M	0.5952	0.0000	0.0000	0.0000	0.0000	0.1265	0.0000
6M	0.5260	0.0123	0.0000	0.0000	0.0000	0.0231	0.0039
1Y	0.5371	0.0000	0.0000	0.0000	0.0000	0.5118	0.0000
2Y	0.6423	0.2842	0.5915	0.0973	0.0000	0.4496	0.0000
3Y	0.6677	0.4843	0.3679	0.3726	0.0063	0.6521	0.0000
5Y	0.6578	0.0740	0.0278	0.6332	0.0000	0.6947	0.0000
7Y	0.6845	0.6845	0.0000	0.0000	0.0521	0.6721	0.1050
10Y	0.6290	0.2346	0.0473	0.0000	0.0226	0.4608	0.0000

Table 1.8: p-values of superior predictive ability (SPA) test of Hansen (2005) for all eight models and their bagged versions. The results are based on the out-of-sample period January 2002 - June 2005.

takes into account the heteroskedastic nature of the interest rates.

In contrast to the existing models, the proposed model is purely data-driven and is able to identify, for every maturity, the most relevant latent and macroeconomic factors both for the local dynamics as well as for the regime structure. As such, it offers a clear interpretation and regime specification while remaining highly competitive in terms of out-of-sample forecasting.

Applying our model to US interest rate data we draw a number of conclusions. First, we find one and the same clear pattern both for the resulting local dynamics and for the regime structure. Based on the pattern, we split the results into three groups: short-, mid- and long-term maturities. For the short maturities we find correspondence between the resulting local structure and the monetary policy models described in the macroeconomic literature. More precisely, the local dynamics of the short end of the yield curve is driven by macroeconomic (inflation, real activity) and term structure (level, slope, and autoregressive term) information. Not surprisingly, we find two limiting regimes linked to the level of inflation (CPI). The optimal threshold structure for the mid-term maturities is determined by the sign of the term structure slope coefficient and the leading real activity indicator HELP. Here, the local dynamics follows a pure AR(1)+GARCH(1,1) process. For the long-term maturities we find that they are subject to up to two regime shifts determined by the conditional volatility of inflation. The local structure of the long end of the yield curve captures the strong macroeconomic impact related to the level of the real activity (HELP) and the inflation's conditional volatility (CPI and PPI).

Second, we conclude that our framework is consistent with the key stylized facts of the yield curve behavior. Finally, we compare the out-of-sample accuracy of our model to those of several strong competitors and find that the bagged version of our model significantly outperforms the other approaches most of the time.

Chapter 2

Bond Risk Premia Forecasting: A Simple Approach for Extracting Macroeconomic Information from a Panel of Indicators

2.1 Introduction

In their highly influential paper, using a reduced form no–arbitrage framework with time–varying risk premia, Ang and Piazzesi (2003) conclude that macroeconomic variables have an important explanatory power for yields and that the inclusion of such variables in term structure models can improve their forecasting performances significantly. More recently, many other studies (see, among others, Ludvigson and Ng (2009b), Joslin, Priebsch, and Singleton (2009), Duffee (2011) for the U.S. or Wright (2009) in an international context) have documented that macroeconomic variables capture significant predictive power for bond excess returns over and above the standard financial factors. In order to avoid relying on specific macro series, Ang and Piazzesi (2003) and Ludvigson and Ng (2009a), measure different macroeconomic fundamentals as the first principal components of blocks of large numbers of macroeconomic series.

In this chapter we propose considering macroeconomic variables as possible relevant factors for modeling the dynamics of the bond risk premia process (and therefore the whole term–structure). We take into account not only the level of a macroeconomic variable, but also its volatility. Moreover, we also propose a different method for reconstructing the level and volatility dynamics of the latent macro–factor from a bunch of observable indicators. Our approach is considerably simpler from a computational perspective than the classical ones introduced in the literature and at the same time performs better in simulations as well as in a real data applications.

In macroeconomics, it is common to have a large set of indexes that measure or are highly dependent on a latent macroeconomic variable. Given the pervasiveness of heteroskedasticity in macroeconomic variables, we model the observable set of proxies using a multivariate conditionally heteroskedastic exact factor model, i.e. a linear factor model where the heteroskedastic conditional variance is a function of the past values of the latent factor (see for instance, Diebold and Nerlove (1989)). In such a type of model, the conditional density, depending on unobservable variables, is generally unknown. As a consequence, the log-likelihood function cannot be obtained explicitly and hence standard maximum likelihood estimators cannot be employed (Harvey, Ruiz, and Sentana (1992)). To overcome this problem, alternative estimation procedures have been proposed in the literature: the Bayesian Markov chain Monte Carlo (MCMC) estimation methods introduced by Fiorentini, Sentana, and Shephard (2004) and the indirect inference estimators introduced by Sentana, Calzolari, and Fiorentini (2008).

However, following the direction proposed by Diebold and Nerlove (1989) and Sentana (2004), in this study we introduce a (computationally) simple estimation approach that relies on filtering the latent factor from a panel of data via an iterated Kalman filter procedure. This approach hinges on recent results about efficient estimation of the macro-parameters in dynamic panel data models with a common factor. In particular, Gagliardini and Gourieroux (2009) showed that substituting the true factor values by their cross-sectional approximations does not lead to any asymptotic efficiency loss. For the cross–sectional reconstruction of the latent factor we propose an iterated process in which we estimate the volatility dynamics of the factor from the time series of a first (time–invariant) Kalman filter approximation of the factor and use it in a new cross–sectional conditional (time–varying) Kalman filter estimation. New volatility dynamics can be estimated from the dynamics of the new estimated factor and the procedure can be iterated until convergence.

Simulation results based on different data–generating processes and the same amount of data that are available in the empirical application show the unbiasedness of the proposed estimator for the conditional variance parameters and its superiority to other simple alternative methods, in particular, to the principal component approach used by Ludvigson and Ng (2009a).

The superiority of our approach is also confirmed by a real data application. Using a panel of 21 monthly inflation time series, we filter the level and the volatility of inflation via several different techniques. We test the ability of the estimated factors in forecasting long–term bond risk premia and find that both the level and the volatility of inflation obtained via an iterated Kalman filter significantly outperform the other competitors. Moreover, by analyzing the correspondence between the different factors and National Bureau of Economic Research (NBER) business cycles, we show that our inflation estimates are not only statistically but also economically significant.

The reminder of the chapter is organized as follows. Section 2.2.1 describes in detail the procedure of reconstructing the level and volatility dynamics of a latent factor. Section 2.2.2 shows the performance of the latent macroeconomic variable and its volatility in a simulation study. In Section 2.3 we apply our estimation technique on real macroeconomic data. Section 2.4 concludes.

2.2 Reconstructing the dynamics and volatility of the latent factor

Our purpose in this section is to reconstruct the underlying time series dynamics of a latent macroeconomic variable and its volatility process from the observations of a certain number of proxies. We propose a simple estimation approach that exploits the possibility of filtering the latent factor from cross-sectional information via an iterated Kalman filter procedure.

2.2.1 Model and estimation procedure

We model the latent factor dynamics at time t through a factor model for the N-dimensional vector of the observables $r_t = (r_{t,i})_{i=1}^N$

$$r_t = Bf_t + e_t, \qquad \text{for} \qquad t = 1...T \tag{2.1}$$

with B the $N \times k$ matrix of factor loadings, e_t the $N \times 1$ vector of idiosyncratic noises, and the latent factor f_t being the variable of interest. In our empirical study of Section 3 we consider a univariate factor representing the latent inflation (i.e., using the standard notation, $f_t = \pi_t$) and, as observables, a number of index proxies for inflation such as different types of Producer and Consumer price indices.

The main assumptions of the model can be expressed in the following form:

$$\begin{pmatrix} f_t \\ u_t \end{pmatrix} |\mathbb{I}_{t-1} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Delta_t & 0 \\ 0 & \Phi \end{pmatrix}\right]. \tag{2.2}$$

The latent factor f_t is assumed to follow a general GARCH type dynamic with (for simplicity) mean zero and (for identifiability) unconditional unit variance i.e. $f_t|\mathbb{I}_{t-1} \sim N(0, \Delta_t)$ with $\mathbb{E}[\Delta_t] = \Delta = I_k$ the identity matrix of order k.[1] The information set $\mathbb{I}_t$ contains current and past values of r and f, i.e. $\mathbb{I}_t = \{r_t, f_t, r_{t-1}, f_{t-1}, \cdots,\}$. As in standard factor models, the vector of idiosyncratic noises e_t is conditionally orthogonal to f_t and has a positive semidefinite diagonal variance matrix Φ, then the conditional distribution of r_t is $r_t|\mathbb{I}_{t-1} \sim N(0, \Sigma_t)$ where $\Sigma_t = B\Delta_t B' + \Phi$ has the usual exact factor structure.

[1] Although possible in principle to extend the model to include dynamics in the conditional mean of the factor, this would certainly complicate both the reconstruction of the latent factor and the estimation of the dynamics of Δ_t. Since our purpose in this chapter is to propose an unbiased estimation method which is as simple as possible, we leave this extension of the model for future research.

In the literature this type of model is called a multivariate conditionally heteroskedastic exact factor model and nests several models widely used in empirical finance (for instance, Diebold and Nerlove (1989)). When the variance of the factor is a function of lagged values of f_t, as in the GARCH case, the exact form of the conditional density of r_t given its past is generally unknown and, hence, the log-likelihood function cannot be explicitly obtained (Harvey, Ruiz, and Sentana (1992)). To overcome this problem, Bayesian Markov chain Monte Carlo (MCMC) estimation methods, simulated EM algorithm Fiorentini, Sentana, and Shephard (2004) and indirect inference estimators Sentana, Calzolari, and Fiorentini (2008) have been proposed in the literature.

Here, instead, we propose a simpler approach in which we iterate between filtering the factor with a Kalman filter in the cross–sectional dimension and estimating its variance dynamics in the time series dimension. This approach hinges on the idea contained in the recent literature on estimators of the macro-parameters in dynamic panel data models with a common factor where the macro-parameter is estimated by means of cross-sectional approximations (Forni and Reichlin (1998), Forni, Hallin, Lippi, and Reichlin (2004), Gagliardini and Gourieroux (2009)). These studies show that, under certain speed of convergence assumptions,[2] estimating the macro-parameter on the cross-sectional approximations of the factors is root–T consistent, asymptotically normal and achieves the same asymptotic efficiency bound as the one obtained with an observable factor (i.e. the Cramer-Rao bound in linear Gaussian models). Therefore, the estimators built on the approximated factor are asymptotically equivalent to the unfeasible estimator that uses the true factor values. These efficiency results are obtained under certain asymptotic schemes which are not expected to necessary hold in our setting. Therefore, whenever these asymptotic conditions are not satisfied, estimators based on more complex simulated estimation methods (as the ones in Fiorentini, Sentana, and Shephard (2004) and Sentana, Calzolari, and Fiorentini (2008)) are expected to be asymptotically more efficient. However, the big advantage of the proposed estimator is to be computationally much simpler. This advantage is due to the way the proposed estimator effectively exploit the cross–sectional dimension (to reconstruct the factor) in combination with the time–series information (used to filter the variance of the factor).

Different approaches can be used to approximate f_t: simple cross sectional averaging, principal component analysis (PCA) or factor analysis (FA). In this study we propose a reconstruction of the f_t factor by an iterative procedure in which the factor is first estimated with a Kalman filter using the cross-section of the observable indicators at our disposal. From the time series of this first approximation of the factor, the variance dynamics are estimated in a classical GARCH framework. The estimated GARCH dynamics of the factor

[2]When $N, T \to \infty$ and $T/N \to c > 0$ the fixed effects estimator is consistent, while if $N, T \to \infty$ such that $T^b/N = O(1), b > 1$ the estimator is efficient.

conditional variance are then used in a conditional Kalman filter estimation to obtain new factor estimates. This iterative procedure is run until convergence. Although we apply this approach to a case where a one factor model arises naturally, this procedure could be directly extended to the case of multiple factors provided that one is not interested in the exact identification of the different factors (because of the indeterminacy induced by factor rotation).

Before starting the procedure, we need an estimate of the factor loading matrix B. Given that in these types of models the factor loadings are assumed to be constant over time, they can be conveniently estimated from unconditional quantities. Moreover, conditionally heteroskedastic factor models also imply unconditional covariance matrices that have an exact k factor structure as in the traditional factor models. Hence, recalling that $\Delta = I_k$, the unconditional covariance matrix Σ can be written as

$$\Sigma = BB' + \Phi. \tag{2.3}$$

Given the different scale of the indices (which have different units of measures), it is desirable to standardize the variable to avoid the problem of having one variable with a large variance unduly influencing the determination of the factor loadings. Standardizing by the individual volatility and working with the correlation matrix is then a customary choice. Clearly, the correlation matrix $R = D^{-1}\Sigma D'^{-1}$ with $D = \mathrm{diag}(\Sigma)$ will also have the same factor structure

$$R = B^*B^{*'} + \Phi^* \tag{2.4}$$

with $B^* = D^{-1}B$ and $\Phi^* = D^{-1}\Phi D'^{-1}$.

Since in our case all the observed indexes are mainly driven by a single latent macroeconomic variable they are supposed to measure, we assume a factor structure with only one common factor (i.e. $k = 1$). Then, the correlation matrix takes the following simple structure.

$$R = \begin{bmatrix} 1 & b_1^*b_2^* & \dots & b_1^*b_N^* \\ b_2^*b_1^* & 1 & \dots & b_2^*b_N^* \\ \vdots & & \ddots & \vdots \\ b_N^*b_1^* & b_N^*b_2^* & \dots & 1 \end{bmatrix}$$

where $[B^*]_i = b_i^*$ is the generic element of the $N \times 1$ vector B^*. This structure, together with the fact that the factor loadings of the proxy are assumed to be all positive, suggests the possibility to estimate the vector of standardized factor loadings B^* by simply minimizing the difference between any generic off diagonal element of the matrix $B^*B^{*'}$ with the corresponding element of the sample unconditional correlation matrix $[S^*]_{ij} = s_{ij}^*$, that is

$$\hat{b}^* = \underset{b^*}{\mathrm{argmin}} \sum_{i=1}^{N} \sum_{j \neq i} (b_i^* b_j^* - s_{i,j}^*)^2. \qquad s.t. \quad 0 < b_i^* < 1 \qquad \forall i \tag{2.5}$$

The minimization algorithm in (2.5) projects the sample correlation matrix into the space spanned by single factor models.

Having the estimated standardized factor loadings $\hat{B}^*$'s, we can estimate the elements of the diagonal matrix Φ^* as $[\hat{\Phi}^*]_{ii} = 1 - (\hat{b}_i^*)^2$. Then the original idiosyncratic variance matrix and factor loadings are simply obtained as $\hat{\Phi} = \hat{D}\hat{\Phi}^*\hat{D}'$ and $\hat{B} = \hat{D}\hat{B}^*$ respectively.

With $\hat{B}$ and $\hat{\Phi}$ at hand, we can now start the Kalman filter iteration. If the joint conditional distribution of r_t and f_t given $\mathbb{I}_{t-1}$ is normal, the model (2.1) has a natural time–series state–space representation. In fact, considering the common factor f_t as state variable, equation (2.1) could be seen as a standard measurement equation. When Δ_t is considered as a given observable, the Kalman filter would coincide with the conditional expectation of f_t given r_t and Δ_t, i.e. $E[f_t|r_t, \Delta_t]$, which is optimal in the conditional mean squared error sense.[3] Thus, the conditional Kalman filter estimate of the common factor would be given by the (unfeasible) updating equation of the filter

$$f_t^{CK} = \Delta_t B' \Sigma_t^{-1} r_t = \Delta_t B'(B\Delta_t B' + \Phi)^{-1} r_t. \tag{2.6}$$

This estimator can be seen as a Bayesian approach for the cross–sectional estimation of the factor. More precisely, the unfeasible estimator in (2.6) corresponds to the mean of the posterior distribution of f_t given the data r_t in a Bayesian approach that considers f_t as a random variable with prior distribution $f_t \sim N(0, \Delta_t)$.

In order to have a feasible conditional Kalman filter, we propose to start the iterative procedure from the following filter with time–invariant weights

$$\hat{f}_t^{(0)} = \hat{B}'\hat{\Sigma}^{-1} r_t = \hat{B}'(\hat{B}\hat{B}' + \hat{\Phi})^{-1} r_t \tag{2.7}$$

using the estimates $\hat{B}$ and $\hat{\Phi}$ obtained from the unconditional information.

Having this first reconstruction of the dynamics of the latent macro–variable, we then get an estimate of the dynamics of its volatility by estimating a GARCH model on $\hat{f}_t$. In this way we obtain a first estimate of the dynamics of the conditional variance of the factor i.e. $\hat{\Delta}_t^{(0)}$ which is then used in the conditional Kalman filter estimation of the factor

$$\hat{f}_t^{(1)} = \hat{\Delta}_t^{(0)} \hat{B}'\hat{\Sigma}_t^{-1} r_t = \hat{\Delta}_t^{(0)} \hat{B}' \left(\hat{B}\hat{\Delta}_t^{(0)} \hat{B}' + \hat{\Phi} \right)^{-1} r_t \tag{2.8}$$

from which a new reconstruction of the latent factor can be computed and a new conditional variance dynamics $\hat{\Delta}_t^{(1)}$ estimated. Iterating this procedure provides our proposed estimator for the dynamics of the latent factor and its conditional variance. Note that in practice, only a small number of iterations is necessary to reach converge and the algorithm is very fast.

[3]Actually, the optimality of the Kalman filter extraction of the factor holds under the more general assumption that f_t and r_t follow a conditional joint distribution that is elliptically symmetric (Sentana (1991)).

2.2.2 Simulations

We first judge the performance of the proposed approach on the accuracy in the reconstruction of the time series of the latent factor f_t. The first employed data generating process (DGP) is a one factor model with the latent factor following a GARCH type dynamics with zero mean and unconditional unit variance. We simulate 1000 paths and for each path we assume 49 years of monthly observations ($T = 588$). Similarly to our real data application, we assume to have 20 observable indicators for the latent macroeconomic variable ($N = 20$). The true βs in the DGP are randomly chosen within a range of values analogous to that estimated on the empirical data.

For comparison purposes we also include the result obtained with a simple cross–sectional average of the indexes, the factor score obtained with cross-sectional OLS regression, the PCA and the FA with one factor. When N is large enough (so that idiosyncratic errors are diversified away) we have that the simple cross-sectional average is $\bar{r_t} \simeq \left(\frac{1}{N}\sum_{i=1}^{N} b_i\right) f_t$; thus the factor values are recovered up to a scaling constant. We account for this scaling constant by simply dividing the series of the cross-sectional averages $\bar{r_t}$ by $\frac{1}{N}\sum_{i=1}^{N} \hat{b}_i$. The cross–sectional OLS regression is another common method to generate factor scores. Contrary to our Kalman filter approach which consider f_t as a random variable, this approach assume the factor to be an unknown parameter and estimate it by the cross–sectional regression $f_t^{OLS} = (\hat{B}'\hat{B})^{-1}\hat{B}'r_t$. The FA is performed using the Matlab command "factoran" which performs maximum likelihood estimate of the factor loadings and computes factor scores using the weighted least–square (or Barlett) method (which also treats the factor scores as fix parameters).

Given that the OLS regression approach completely discards the information contained in Δ_t and Φ, while FA neglects the information contained in the dynamics of Δ_t, we expect them to be less efficient than the Iterated Kalman filter method who optimally exploits the information in both Φ and Δ_t.

To judge the accuracy in reconstructing the f_t series with the various approaches, we compute the Root Mean Square Error (RMSE) for each simulated path between the true path of the latent factor and the estimated one. For each simulation path we also compute the correlation coefficient between the two series. Results are reported in the first two rows of Table 2.1 Panel A.

According to both metrics, our proposed procedure for the latent f_t process turns out to be the most precise; it is the one with, on average, the smallest RMSE and the highest correlation coefficient.

We then evaluate the ability of the different approaches to reconstruct the volatility dynamics of the true factor by computing the RMSE and correlation coefficient between the true series of simulated volatilities and the reconstructed ones obtained by fitting a

Performance Comparison - Simulations

	Simple Average	cross-section OLS	Factor Analysis	Principal Component	Iterated Kalman
Panel A					
Avg corr on f_t	0.9640	0.9526	0.9898	0.9433	0.9899
Avg RMSE on f_t	0.2742	0.3161	0.1467	0.3337	0.1391
Avg corr on σ_t	0.9477	0.9379	0.9677	0.9289	0.9691
Avg RMSE on σ_t	0.1404	0.1456	0.1293	0.1460	0.0548
Panel B					
Avg corr on f_t	0.9636	0.9517	0.9900	0.9416	0.9902
Avg RMSE on f_t	0.2756	0.3193	0.1455	0.3382	0.1381
Avg corr on σ_t	0.9454	0.9340	0.9670	0.9244	0.9684
Avg RMSE on σ_t	0.1397	0.1449	0.1285	0.1451	0.0549
Panel C					
Avg corr on f_t	0.9635	0.9521	0.9895	0.9426	0.9899
Avg RMSE on f_t	0.2757	0.3181	0.1468	0.3361	0.1396
Avg corr on σ_t	0.6176	0.6077	0.6397	0.5993	0.6444
Avg RMSE on σ_t	0.2260	0.2291	0.2195	0.2271	0.2032

Table 2.1: Performance comparison of different filtering methods for the factor dynamics and its conditional volatility over 1000 simulation paths. The methods are: simple cross–sectional averages, cross–sectional OLS regression, Factor Analysis, Principal Component, and Iterated Kalman filter. The performance measures are the average correlation and the average Root Mean Square Error (RMSE). The different DGP in the three panels are given by the one factor model with: GARCH dynamics in Δ_t (Panel A), GARCH dynamics in Δ_t and Φ_t (Panel B), GARCH dynamics with two regimes in Δ_t (Panel C).

GARCH(1,1) process to the estimated f_t series. Again, the Iterated Kalman filter provides the reconstruction of the latent factor volatility with, on average, the lowest RMSE and the highest correlation coefficient, as shown in the last two rows of Table 2.1 Panel A.

Finally, in Figure 2.1, Panel A, we plot the distributions of the estimated parameters of the GARCH process for the volatility.

The figure clearly shows that the estimates of the true parameters α and β of the GARCH process in the factor DGP are both unbiased and reasonably accurate.

We also test the procedure on two more challenging volatility DGP processes (i.e. with a purposely misspecified DGP). In the first one the diagonal variance matrix of the idiosyncratic noise Φ, which was kept constant over time in the previous set up, is now also time–varying, with each idiosyncratic component following a different GARCH process. The objective of this simulation exercise is to test the robustness of our procedure in a misspecified set up featuring GARCH dynamics in both the factor and idiosyncratic conditional variances, i.e. with time–varying Δ_t and Φ_t.

The second one consists of a two-regime process with lagged return as the threshold variable where the local conditional variance evolves according to a FIGARCH(1,d,1) model (see Baillie, Bollerslev, and Mikkelsen (1996)) in one regime and a model that is not of a GARCH type in the second regime.

The results are similar to those previously obtained in the correctly specified set up, confirming in both cases the more accurate reconstruction of the latent process by the proposed iterated Kalman filter method.

Finally, as in Figure 2.1, Panel A, in Panel B we plot the distribution of the α and β parameter estimates in the case of DGP process with time varying (GARCH type) idiosyncratic noise Φ_t. GARCH parameter estimates seem to remain unbiased even in this misspecified context.

[Figure 2.1 about here.]

2.3 Real data application: bond risk premia forecasting

Economic theory suggests that (a great portion of) bond term premia variation is driven by macroeconomic fundamentals. Yet, the link between macroeconomic activity and risk premia might be hard to detect. Using different modeling setups, many recent studies (see, among others, Ludvigson and Ng (2009b), Joslin, Priebsch, and Singleton (2009), or Duffee (2011)) document that macroeconomic variables capture significant predictive power for excess returns over and above the standard financial factors. In this section we assess the performance of our iterated Kalman filter technique in forecasting long–term bond excess

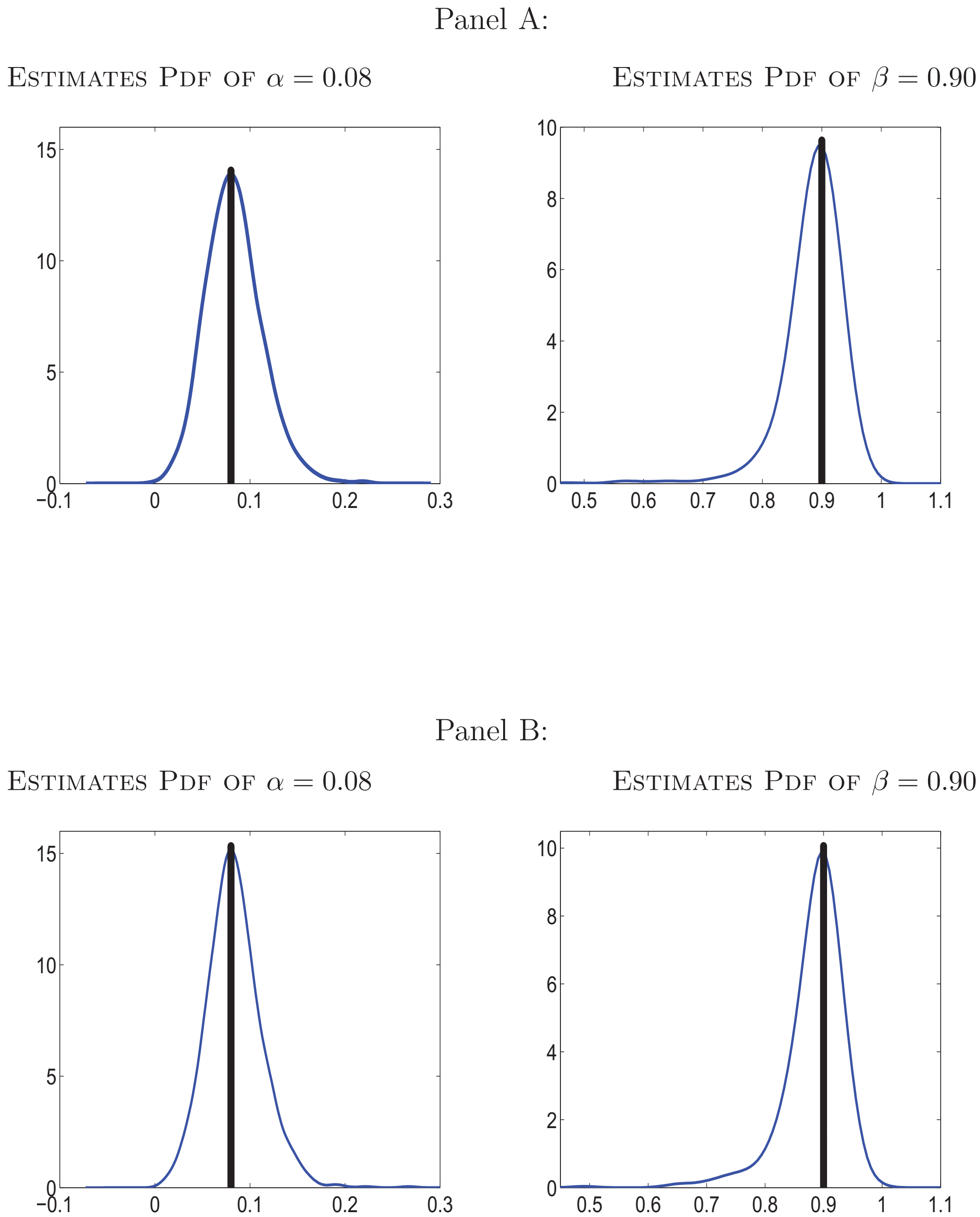


Figure 2.1: Probability distribution function of the estimation error over 1000 simulation paths of the parameters of the GARCH(1,1) process for the factor conditional variance $\sigma_t^2 = c + \alpha f_{t-1}^2 + \beta \sigma_{t-1}^2$ in a DGP with constant (Panel A) and time varying (GARCH type, Panel B) idiosyncratic noise Φ_t.

returns in comparison with principal components analysis and factor analysis.[4] Results obtained applying a principal components analysis are not reported in Section 3.3 given that they are qualitatively the same, but slightly worse, than those obtained using factor analysis.

2.3.1 Data and estimated inflation levels and variances

In our empirical study two different datasets are used.

Bond Data

We use monthly data (June 1961 onward) from the Federal Reserve Board constructed as in Gürkaynak, Sack, and Wright (2006).[5] Following Cochrane and Piazzesi's (2005) procedure, bond excess returns are calculated in the classical way as 1-year holding period returns in excess of the one–year risk–free rate. Furthermore, we construct our tent–shape bond–return forecasting factor described in Cochrane and Piazzesi (2005) (hereafter CP factor) as a linear combination of forward rates. The inclusion of the CP factor is motivated simply by the fact that it has high explanatory power for bond excess returns.

Macroeconomic Data

The second dataset consists of monthly observations for 21 U.S. inflation time series. Exact description of the data is given in Appendix 2.5. The panel spans the period January 1959 – December 2007 and has already been used as a part of other studies: see, among others, Stock and Watson (2005), Ludvigson and Ng (2009b) and Ludvigson and Ng (2009a). We build two alternative pairs of estimates for inflation levels and variances. First, similar to Ludvigson and Ng (2009a), we extract both the first principal component (PC) and the first factor (FA) as measures for inflation's level. PC and FA volatility are computed from fitting a GARCH(1,1) to the estimated principal component and the estimated factor, respectively. Our second approach for reconstructing the level and the variance of inflation is based on the iterated Kalman filter procedure described in Section 2.2.1.

For our analysis we take the largest common period of the two datasets and split it into two parts. We consider June 1961 to December 2003 as in-sample period. The rest of the data (January 2004 - December 2008) has been left to evaluate the out-of-sample forecasting performance of the different predictors. Summary statistics of the data are reported in Table 2.2.

[4]Another possible alternative procedure is the one proposed by Harvey, Ruiz, and Sentana (1992). Given the dimensionality of the problem, however, that approach is too computational expensive and is not implemented.

[5]The data are available under http://www.federalreserve.gov/econresdata/researchdata.htm.

Summary Statistics of Data

	$rx^{(5)}$	$rx^{(10)}$	$rx^{(20)}$	$rx^{(30)}$	CP	π^{IK}	$\mathtt{vol}\pi^{IK}$	π^{PC}	$\mathtt{vol}\pi^{PC}$	π^{FA}	$\mathtt{vol}\pi^{FA}$
Panel A:											
Mean	0.011	0.013	0.011	0.009	0.006	0.838	1.034	0.034	0.876	0.497	0.499
Std	0.056	0.104	0.198	0.325	0.019	0.579	2.109	1.014	0.694	0.401	0.370
AC1	0.931	0.921	0.880	0.798	0.916	0.989	0.993	0.973	0.953	0.991	0.990
Panel B:											
$rx^{(5)}$	1.00										
$rx^{(10)}$	0.96	1.00									
$rx^{(20)}$	0.82	0.92	1.00								
$rx^{(30)}$	0.62	0.72	0.90	1.00							
CP	0.43	0.48	0.49	0.44	1.00						
π^{IK}	-0.22	-0.25	-0.26	-0.26	-0.06	1.00					
$\mathtt{vol}\pi^{IK}$	-0.31	-0.36	-0.31	-0.29	-0.15	0.89	1.00				
π^{PC}	-0.30	-0.29	-0.30	-0.29	-0.41	0.55	0.42	1.00			
$\mathtt{vol}\pi^{PC}$	-0.23	-0.24	-0.26	-0.26	-0.38	0.49	0.49	0.69	1.00		
π^{FA}	-0.27	-0.29	-0.30	-0.30	-0.13	0.97	0.85	0.68	0.61	1.00	
$\mathtt{vol}\pi^{FA}$	-0.26	-0.30	-0.30	-0.30	-0.11	0.97	0.89	0.53	0.55	0.98	1.00
NBER					0.04	0.46	0.45	0.17	0.24	0.43	0.41

Table 2.2: Panel A reports summary statistics for the following variables: 5, 10, 20, 30 year bond excess returns (denoted by $rx^{(5)}$, $rx^{(10)}$, $rx^{(20)}$, $rx^{(30)}$, respectively), Cochrane and Piazzesi (2005) factor (denoted by CP), inflation level and inflation volatility factors estimated by iterated Kalman filter (denoted by π_t^{IK} and $\mathtt{vol}\pi_t^{IK}$),by factor analysis (denoted by π_t^{FA} and $\mathtt{vol}\pi_t^{FA}$), and by principal components analysis (denoted by π_t^{PC} and $\mathtt{vol}\pi_t^{PC}$). NBER is a binary variable, where one indicates month designated as recessions by the National Bureau of Economic Research. AC1 denotes the first autocorrelation coefficient. Panel B reports cross–correlations.

The adequacy of the one factor structure may be questionable. In fact, the assumption of the one factor structure is primarily given by the economic consideration that all the variable in the data set are all proxy of the same underlying macroeconomic variable i.e. inflation. From a statistical point of view we can observe that the first principal component explain about 53% of the total variance of the dataset while all the other components are below 10%. The presence of a highly persistent heteroscedasticity in the series, which justifies the use of a GARCH(1,1) model, is given by the highly significant results of the Engle ARCH test with a large number of lags of 50.

Figure 2.2 illustrates the difference between the inflation's level and the inflation's volatility factors obtained using the three different techniques.

As Figure 2.2 clearly shows, the estimated inflation's level and volatility factors obtained from the three competing approaches are significantly different.

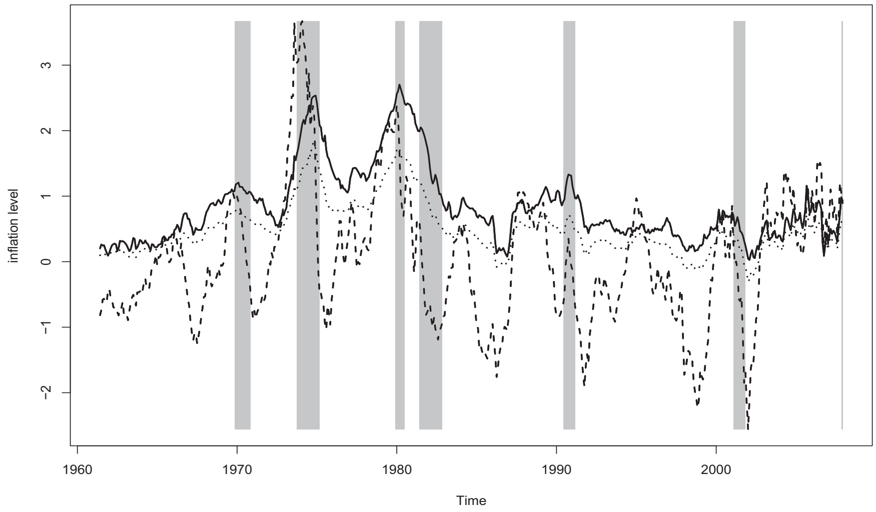
2.3.2 Financial variables, inflation measures, and business cycles

To begin with, we analyze the correspondence between the NBER business cycles and the different financial and inflation measures. The last row of Table 2.2 reports the results. The weak correlation (around 0.04) between the NBER recession and CP factor confirms Ludvigson and Ng (2009b) finding that, without macro factors, bond risk premia appear virtually acyclical. Yet, theory says that risk premia have a marked counter–cyclical behavior, compensating the investors for macroeconomic risks. The almost two times higher correlation between the NBER business cycles indicator and the iterated Kalman filter inflation variables as well as the ones obtained from FA in comparison to those estimated with the PC approach assures more pronounced cyclical fluctuations in bond risk premia. By its iterated nature, our measures for inflation seem to better capture perceptions of risks looming on the investors horizon. Thus, they convey valid and timely information over and above that contained in other financial and PC inflation fundamentals. These findings make the inflation factors obtained by the iterated Kalman filter approach highly economically significant. No particularly significant difference in the correlations with NBER business cycles can be seen between our procedure and a classical FA.

2.3.3 Long–term bond risk premia forecasting results

First of all, following the term structure literature to assess the in-sample performance of our procedure in comparison with the alternative approaches we investigate the impact of the different pairs of inflation factors (i.e. level and volatility) as predictors for bond excess returns at different maturities. To this goal we run the following regressions:

Inflation Level



Inflation Volatility

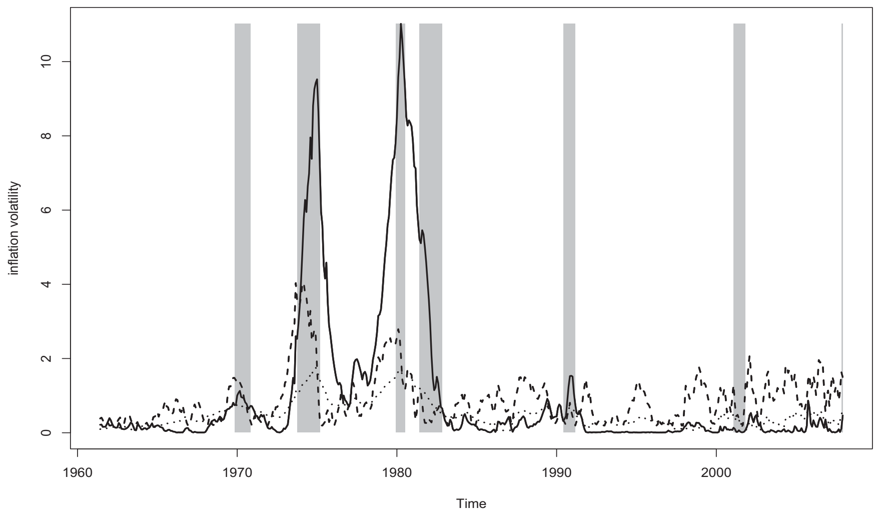


Figure 2.2: The upper panel plots the three estimates of inflation level: iterated Kalman filter (solid line), factor analysis (dotted line), and principal components (dashed line) based on a panel of 21 inflation time series, as described in the text. The lower panel plots the inflation volatility filtered by the three techniques. Once again the solid line indicates the iterated Kalman filter estimate, the dotted line the estimate got applying factor analysis, whereas the dashed line represents the dynamics of the principal components volatility. The shaded bars denote months designated as recessions by NBER.

$$
\begin{aligned}
\text{Model } \mathcal{M}_1 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \varepsilon^{(n)}_{t+12} \\
\text{Model } \mathcal{M}_2 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \pi^{IK}_t + \varepsilon^{(n)}_{t+12} \\
\text{Model } \mathcal{M}_3 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \texttt{vol}\pi^{IK}_t + \varepsilon^{(n)}_{t+12} \\
\text{Model } \mathcal{M}_4 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \pi^{FA}_t + \varepsilon^{(n)}_{t+12} \\
\text{Model } \mathcal{M}_5 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \texttt{vol}\pi^{FA}_t + \varepsilon^{(n)}_{t+12} \\
\text{Model } \mathcal{M}_6 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \pi^{IK}_t + \gamma_3 \texttt{vol}\pi^{IK}_t + \varepsilon^{(n)}_{t+12} \\
\text{Model } \mathcal{M}_7 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \pi^{PC}_t + \gamma_3 \texttt{vol}\pi^{FA}_t + \varepsilon^{(n)}_{t+12}, \\
\text{Model } \mathcal{M}_8 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \pi^{IK}_t + \gamma_3 \pi^{FA}_t + \varepsilon^{(n)}_{t+12} \\
\text{Model } \mathcal{M}_9 : \quad & rx^{(n)}_{t+12} = \gamma_0 + \gamma_1 CP_t + \gamma_2 \texttt{vol}\pi^{IK}_t + \gamma_3 \texttt{vol}\pi^{FA}_t + \varepsilon^{(n)}_{t+12},
\end{aligned}
$$

where $rx^{(n)}_{t+12}$ are the excess returns on an n year nominal bond ($n = 5, 10, 20, 30$) at time $t + 12$. CP_t represents the CP factor, π_t and $\texttt{vol}\pi_t$ denote the inflation level and inflation volatility factors, estimated by the two different approaches: iterated Kalman filter (denoted by π^{IK}_t and $\texttt{vol}\pi^{IK}_t$) and factor analysis (denoted by π^{FA}_t and $\texttt{vol}\pi^{FA}_t$), respectively. To this end, we estimate nine different models. First, we regress the excess returns only on CP factor (Model $\mathcal{M}_1$). This regression should serve as a benchmark model. Then, in Model $\mathcal{M}_2$ and Model $\mathcal{M}_3$ we add one more predictor, the level and the volatility of inflation, each estimated by the iterative Kalman filter approach. We repeat the same procedure for the next two models (Model $\mathcal{M}_4$ and Model $\mathcal{M}_5$), where we add once again the level and the volatility of inflation, this time estimated by the FA technique. In Model $\mathcal{M}_6$ and Model $\mathcal{M}_7$ we take into consideration all three predictors: CP factor, level and volatility of inflation. The only difference between Model $\mathcal{M}_6$ and Model $\mathcal{M}_7$ is in the way the inflation variables are measured. In particular, in Model $\mathcal{M}_6$ the inflation variables are derived by the iterated Kalman filter procedure, whereas in $\mathcal{M}_7$ FA has been used. In contrast to the previous models, where the main idea is to assess performance, the individual filtering techniques, the last two models (Model $\mathcal{M}_8$ and Model $\mathcal{M}_9$) provide a direct comparison between the two level (Model $\mathcal{M}_8$) and the two volatility (Model $\mathcal{M}_9$) factors. All coefficients are estimated with ordinary least squares, and standard errors are corrected for autocorrelation and heteroskedasticity. [6] Table 2.3, Table 2.4, Table 2.5, and Table 2.6 present the results.

The estimated coefficients for the CP factor are positive and highly significant for predicting bond risk premia at all maturities. Fully in line with the literature, the CP factor accounts for around 28% of the excess returns variation. The strength of the predictive power of the inflation factors changes with time to maturity of a bond, explaining up to 6% of the variation in addition to the CP factor. The estimated coefficients for level and volatility of inflation are negative, and they are significant most of the time. The negative

[6] In particular, we follow Ludvigson and Ng (2009a) and Cochrane and Piazzesi (2005) and compute t-statics using the Newey-West adjustment with 18 lags.

PREDICTIVE REGRESSION ANALYSIS 5 YEAR EXCESS RETURNS

	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$	$\mathcal{M}_6$	$\mathcal{M}_7$	$\mathcal{M}_8$	$\mathcal{M}_9$
Intercept	−0.002 (−0.310)	0.013 (1.168)	0.005 (0.678)	0.013 (−0.173)	0.013 (0.185)	0.004 (0.339)	0.011 (1.118)	0.003 (0.724)	0.008 (0.824)
CP Factor	1.533*** (4.749)	1.474*** (4.555)	1.410*** (4.219)	1.430*** (3.982)	1.444*** (4.333)	1.408*** (3.992)	1.420*** (4.294)	1.3200*** (4.016)	1.418*** (4.080)
Inflation Level (Iterated Kalman)		−0.027* (−1.756)				-0.060 (-1.131)		0.068 (1.599)	
Inflation Vol (Iterated Kalman)			−0.028* (−1.650)			−0.005 (−0.826)			−0.003 (−0.474)
Inflation Level (FA)				−0.009 (−1.430)			−0.060 (−1.312)	−0.124* (−2.028)	
Inflation Vol (FA)					0.035 (0.602)		0.006 (0.615)		-0.012 (-0.398)
R^2	0.257	0.285	0.294	0.287	0.293	0.294	0.296	0.316	0.312
Adjusted R^2	0.256	0.283	0.292	0.285	0.290	0.290	0.291	0.296	0.292

Table 2.3: Results for ordinary least squares regressions for nine different models (labeled as $\mathcal{M}_1$, $\mathcal{M}_2$,…,$\mathcal{M}_9$) utilizing annual returns on 5-year Treasury bonds. Standard errors are corrected for autocorrelation and heteroskedasticity. t-statistics are reported in parenthesis. Asterisks * ,** ,*** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The data span the period June 1962 to December 2003. See text for more details.

Predictive Regression Analysis 10 Year Excess Returns

	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$	$\mathcal{M}_6$	$\mathcal{M}_7$	$\mathcal{M}_8$	$\mathcal{M}_9$
Intercept	−0.013	0.018	0.001	0.017	0.019	0.002	0.017	0.005	0.011
	(−1.164)	(0.991)	(0.088)	(1.052)	(1.097)	(0.121)	(0.983)	(0.781)	(0.556)
CP Factor	3.025***	2.899***	2.772***	2.817***	2.840***	2.778***	2.817***	2.841***	2.792***
	(4.649)	(4.521)	(4.251)	(4.379)	(4.457)	(4.203)	(4.336)	(4.177)	(4.243)
Inflation Level (Iterated Kalman)		−0.035*				−0.002		0.087	
		(−1.751)				(−0.078)		(1.186)	
Inflation Vol (Iterated Kalman)			−0.011*			−0.010			−0.006
			(−1.917)			(−0.960)			(−0.562)
Inflation Level (FA)				−0.056*			−0.057	−0.179*	
				(−1.990)			(−.581)	(−1.709)	
Inflation Vol (FA)					−0.059*		0.001		−0.029
					(−1.941)		(0.011)		(−0.572)
R^2	0.283	0.320	0.328	0.327	0.327	0.328	0.328	0.338	0.330
Adjusted R^2	0.282	0.317	0.325	0.324	0.325	0.324	0.324	0.334	0.326

Table 2.4: Results for ordinary least squares regressions for nine different models (labeled as $\mathcal{M}_1$, $\mathcal{M}_2$,…,$\mathcal{M}_9$) utilizing annual returns on 10-year Treasury bonds. Standard errors are corrected for autocorrelation and heteroskedasticity. t-statistics are reported in parenthesis. Asterisks * ,** ,*** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The data span the period June 1962 to December 2003. See text for more details.

Predictive Regression Analysis: 20 Year Excess Returns

	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$	$\mathcal{M}_6$	$\mathcal{M}_7$	$\mathcal{M}_8$	$\mathcal{M}_9$
Intercept	-0.042^{**}	0.017	-0.016	0.014	0.020	-0.005	0.019	-0.005	0.013
	(-1.976)	(0.634)	(-0.972)	(0.639)	(0.784)	(-0.157)	(0.730)	(-0.159)	(0.391)
CP Factor	5.901^{***}	5.659^{***}	5.438^{***}	5.507^{***}	5.540^{***}	5.482^{***}	5.533^{***}	5.270^{***}	5.498^{***}
	(4.366)	(4.439)	(4.300)	(4.368)	(4.425)	(4.298)	(4.354)	(4.216)	(4.287)
Inflation Level (Iterated Kalman)		-0.066^{*}				-0.019		0.147	
		(-1.964)				(-0.362)		(1.186)	
Inflation Vol (Iterated Kalman)			-0.019^{**}			-0.015			-0.005^{**}
			(-2.348)			(-1.169)			(-0.378)
Inflation Level (FA)				-0.106^{*}			-0.018	-0.314^{*}	
				(-2.268)			(-0.139)	(-1.841)	
Inflation Vol (FA)					-0.116^{*}		-0.097		-0.090
					(-2.248)		(-0.637)		(-1.043)
R^2	0.299	0.336	0.340	0.341	0.342	0.342	0.343	0.351	0.346
Adjusted R^2	0.297	0.334	0.338	0.339	0.339	0.337	0.339	0.348	0.342

Table 2.5: Results for ordinary least squares regressions for nine different models (labeled as $\mathcal{M}_1$, $\mathcal{M}_2$,…,$\mathcal{M}_9$) utilizing annual returns on 20-year Treasury bonds. Standard errors are corrected for autocorrelation and heteroskedasticity. t-statistics are reported in parenthesis. Asterisks * ,** ,*** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The data span the period June 1962 to December 2003. See text for more details.

Predictive Regression Analysis: 30 Year Excess Returns

	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$	$\mathcal{M}_6$	$\mathcal{M}_7$	$\mathcal{M}_8$	$\mathcal{M}_9$
a Intercept	-0.072^{**}	0.029	-0.031	0.024	0.034	0.008	0.035	-0.007	-0.040
	(-2.027)	(0.798)	(-1.215)	(0.730)	(0.792)	(1.016)	(0.946)	(-0.193)	(-0.368)
CP Factor	8.779^{***}	8.372^{***}	8.057^{***}	8.118^{***}	7.874^{***}	8.213^{***}	8.179^{***}	7.727^{***}	8.203^{***}
	(3.830)	(4.013)	(3.911)	(3.974)	(4.023)	(3.886)	(3.966)	(3.815)	(3.928)
Inflation Level (Iterated Kalman)		-0.112^{**}				-0.070		0.243	
		(-2.292)				(-0.808)		(1.370)	
Inflation Vol (Iterated Kalman)			-0.030^{***}			-0.013			0.004
			(-2.602)			(-0.751)			(0.233)
Inflation Level (FA)				-0.178^{***}			0.030	-0.522	
				(-2.604)			(0.169)	(-2.059)	
Inflation Vol (FA)					-0.197^{***}		-0.023		-0.220
					(-2.628)		(-1.145)		(-1.577)
R^2	0.246	0.285	0.283	0.289	0.287	0.286	0.290	0.298	0.296
Adjusted R^2	0.244	0.282	0.280	0.287	0.284	0.282	0.287	0.295	0.291

Table 2.6: Results for ordinary least squares regressions for nine different models (labeled as $\mathcal{M}_1$, $\mathcal{M}_2$,...,$\mathcal{M}_9$) utilizing annual returns on 30-year Treasury bonds. Standard errors are corrected for autocorrelation and heteroskedasticity. t-statistics are reported in parenthesis. Asterisks * ,** ,*** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The data span the period June 1962 to December 2003. See text for more details.

correlation between the different inflation measures and excess returns is quite intuitive, as higher inflation decreases the value of the nominal bond. Including both level and volatility of the inflation factor (see Models $\mathcal{M}_6$ and $\mathcal{M}_7$) in the regression does not seem to improve the accuracy, and both predictors become statistically not significant.[7]

Although, at first glance, both filtering techniques seem to perform equally well in this in-sample investigation, the ability of our approach to reconstruct in a more accurate way both the level and the volatility of inflation has empirical merits out-of-sample. To support this, we run a genuine out-of-sample experiment for the remaining period in our sample. Forecasting results covering the period January 2004 to December 2008 are shown in Table 2.7.

The superior predictive ability test of Hansen (2005) (see Table 2.7) reveals that our inflation's level and volatility measures on top of the CP factor matter for forecasting bond risk premia, significantly outperforming other alternatives. Importantly, however, their impact can differ, depending on the time to maturity of a bond.

We also test the performance of the two filtering techniques in a more challenging framework. Without making any additional assumptions, we create a pool of predictors, including the two different pairs of inflation measures and the CP factor, and let the data themselves choose the most informative variables. This is achieved by finding for each possible number of predictors the subset of the corresponding size that gives the smallest residual sum of squares.[8] Then, we use the Bayesian Schwarz Information Criterion (BIC) to select the best model. We find that regressing the excess returns on the CP factor and the volatility of inflation obtained by the iterated Kalman filter i.e. Model $\mathcal{M}_3$ leads to optimal results.

Finally, we discuss the overall impact of the individual inflation factors in forecasting bond risk premia. Based on the in–sample fit, out–of–sample forecasting, and economic significance, we document that the most important macroeconomic variable for bond excess returns represents the volatility of inflation estimated via the iterated Kalman filter technique. Yet, our inflation volatility measure is no longer a statistically significant predictor of long–term bond risk premia once the level of inflation is in the same regression. The reason for this is the high correlation between the two iterated Kalman filter factors. However, their impact varies with the time to maturity of a bond. In general, we may conclude that the iterated Kalman filter technique allows us to extract in a more accurate way the investors' perceptions of inflation risk in comparison with alternative approaches.

[7]This result is a consequence of the high correlation between the two variables (and both series are very persistent) together with the necessary Newey-West correction that substantially lowers the t-statistics.

[8]This procedure is known in the literature as best subset selection. See Hastie, Tibshirani, and Friedman (2001) for more details.

Panel A: Out-of-Sample Mean Squared Errors

	5Y Bond Exret	10Y Bond Exret	20Y Bond Exret	30Y Bond Exret
$\mathcal{M}_1$	0.0033 (0.0679)	0.0091 (0.0163)	0.0320 (0.0000)	0.0873 (0.0000)
$\mathcal{M}_2$	0.0029 (0.3114)	0.0074 (0.1292)	0.0256 (0.0818)	0.0706 (0.0860)
$\mathcal{M}_3$	0.0028 (0.7174)	0.0070 (0.4381)	0.0241 (0.4289)	0.0674 (0.3816)
$\mathcal{M}_4$	0.0029 (0.1424)	0.0076 (0.0060)	0.0265 (0.0007)	0.0731 (0.0000)
$\mathcal{M}_5$	0.0029 (0.3086)	0.0075 (0.0067)	0.0259 (0.0006)	0.0709 (0.0090)
$\mathcal{M}_6$	0.0028 (0.5228)	0.0070 (0.6444)	0.0242 (0.5927)	0.0683 (0.6727)
$\mathcal{M}_7$	0.0030 (0.3163)	0.0076 (0.0265)	0.0260 (0.0007)	0.0707 (0.0186)

Panel B: Out–of–Sample Mean Absolute Errors

	5Y Bond Exret	10Y Bond Exret	20Y Bond Exret	30Y Bond Exret
$\mathcal{M}_1$	0.0418 (0.0954)	0.0786 (0.0000)	0.1576 (0.0000)	0.2473 (0.0000)
$\mathcal{M}_2$	0.0400 (0.4728)	0.0697 (0.1435)	0.1381 (0.0681)	0.2146 (0.1380)
$\mathcal{M}_3$	0.0391 (0.7175)	0.0676 (0.3838)	0.1349 (0.4187)	0.2120 (0.4242)
$\mathcal{M}_4$	0.0439 (0.1362)	0.0709 (0.015)	0.1412 (0.0006)	0.2200 (0.0000)
$\mathcal{M}_5$	0.0422 (0.2092)	0.0704 (0.0257)	0.1396 (0.0010)	0.2169 (0.0018)
$\mathcal{M}_6$	0.0390 (0.5643)	0.0676 (0.6722)	0.1348 (0.6992)	0.2118 (0.7385)
$\mathcal{M}_7$	0.0440 (0.4819)	0.0709 (0.0066)	0.1398 (0.0030)	0.2165 (0.0016)

Table 2.7: Results (mean squared errors (Panel A) and mean absolute errors (Panel B)) of out–of–sample forecasting performance of seven different models for 5-, 10-, 20- and 30-year Treasury Bond excess returns, as described in detail in the text. p-values of the superior predictive ability (SPA) test of Hansen (2005) are reported in parenthesis. The results are based on the out-of-sample period, January 2004 - December 2008.

2.4 Conclusions

In this chapter we propose a new, computationally simple approach for reconstructing the level and volatility dynamics of a latent macroeconomic factor from a large panel of macroeconomic indices. Our estimation procedure is based on the iterated Kalman filter technique in which we iterate between filtering the unobservable factor with a Kalman filter in the cross–sectional dimension and estimating its variance dynamics in the time series dimension.

We assess the performance of our iterated Kalman filter approach on a set of empirical studies. Extensive simulation results reveal the accuracy of our latent factor volatility estimates and its superiority in comparison with other alternative approaches. Encouraged by those results, we test the ability of our approach to reconstruct in a more accurate way the unobservable macroeconomic driver and its volatility on a real data application – bond risk premia forecasting. Using a panel of a large number of inflation time series, we filter the level and the volatility of inflation via different techniques. We find that in predicting long–term bond risk premia, our inflation estimates significantly outperform the other competitors. In addition, looking at the correspondence between NBER business cycles and inflation fundamentals, we conclude that our estimates are not only statistically but also economically significant.

Our analysis could be taken a step further by studying the performance of bond risk premia in a term structure modeling framework. The iterated Kalman technique could also be to used obtain more accurate estimates for other important macroeconomic predictors such as real activity. However, those extensions are left for future research.

2.5 Data Appendix

This appendix presents U.S. inflation data used in our real data analysis. The first column lists the short name of the inflation variable, followed by its mnemonic in column 2, and a brief data description in column 4. All data series are from Global Insights Basic Economic Database. The third column shows the transformations used to assure stationarity of the individual time series. In particular, $\Delta \ln$ and lv denote the first difference of the logarithm and the level of the series, respectively. These data span the period January 1959 - December 2007 for a total of 588 monthly observations.

Short Name	Mnemonic	Tran	Description
PPI: fin gds	pwfsa	Δ ln	Producer Price Index: Finished Goods (82=100,Sa)
PPI: cons gds	pwfcsa	Δ ln	Producer Price Index: Finished Consumer Goods (82=100,Sa)
PPI: int materials	pwimsa	Δ ln	Producer Price Index:Intermed Mat.Supplies & Components(82==100,Sa)
PPI: crude matls	pwcmsa	Δ ln	Producer Price Index: Crude Materials (82=100,Sa)
Spot market price	psccom	Δ ln	Spot market price index: bls & crb: all commodities(1967=100)
PPI: nonferrous materials	pw102	Δ ln	Producer Price Index: Nonferrous Materials (1982=100, Nsa)
NAPM com price	pmcp	lv	Napm Commodity Prices Index (Percent)
CPI-U: all	punew	Δ ln	Cpi-U: All Items (82-84=100,Sa)
CPI-U: apparel	pu83	Δ ln	Cpi-U: Apparel & Upkeep (82-84=100,Sa)
CPI-U: transp	pu84	Δ ln	Cpi-U: Transportation (82-84=100,Sa)
CPI-U: medical	pu85	Δ ln	Cpi-U: Medical Care (82-84=100,Sa)
CPI-U: comm.	puc	Δ ln	Cpi-U: Commodities (82-84=100,Sa)
CPI-U: dbles	pucd	Δ ln	Cpi-U: Durables (82-84=100,Sa)
CPI-U: services	pus	Δ ln	Cpi-U: Services (82-84=100,Sa)
CPI-U: ex food	puxf	Δ ln	Cpi-U: All Items Less Food (82-84=100,Sa)
CPI-U: ex shelter	puxhs	Δ ln	Cpi-U: All Items Less Shelter (82-84=100,Sa)
CPI-U: ex med	puxm	Δ ln	Cpi-U: All Items Less Medical Care (82-84=100,Sa)
PCE defl	gmdc	Δ ln	Pce, Impl Pr Defl:Pce (2000=100) (AC) (BEA)
PCE defl: dlbes	gmdcd	Δ ln	Pce, Impl Pr Defl:Pce; Durables (2000=100) (AC) (BEA)
PCE defl: nondble	gmdcn	Δ ln	Pce, Impl Pr Defl:Pce; Nondurables (2000=100) (AC) (BEA)
PCE defl: service	gmdcs	Δ ln	Pce, Impl Pr Defl:Pce; Services (2000=100) (AC) (BEA)

Chapter 3

Monetary Policy Regime Shifts and Asset Prices

3.1 Introduction

Over the last 40 years researchers have actively studied the dynamics of the term structure of interest rates and its relation with the states of the economy. Most of the recent attention has been devoted to jointly modeling yield and macroeconomic variables within a no–arbitrage framework. Motivated by the fact that the unconditional cross-sectional variation of yields can be almost perfectly summarized by just a small number of latent factors, the so called "macro–finance" models provide a parsimonious and tractable way to capture the yield curve dynamics. The link between nominal interest rates and macroeconomy is then typically made by augmenting the latent factors with relevant (for the term structure) macroeconomic variables.

One often mentioned but seldom taken into account fact in the macro-finance literature is the time–varying relation between the short rate and the state of the economy. Indeed, the literature so far has focused mainly on studying the long end of the yield curve (and the risk premia associated with it), ignoring the nonlinear behavior of short term maturities. In general, those nonlinearities can be easily captured by increasing the number of latent state variables, but this leaves the link between the state of the economy and interest rates a puzzle.

Taking into account the time–varying relation between the short rate and the state of the economy is crucial for at least two reasons. First, ignoring nonlinearities in the short rate dynamics may give misleading results about the impact of the different macroeconomic fundamentals on the term structure. To gain some insight into this problem, let us consider a simple Taylor rule monetary policy model for the short rate over two different twenty-year periods: 1965 - 1984 and 1985 - 2004. Table 3.1 juxtaposes the parameter estimates over the two subsamples.

Relying only on parameter estimates from the second subsample, one can easily conclude

Taylor Rule Monetary Policy Model

Global Taylor Rule $r_{t+1} = \gamma_0 + \rho r_t + \gamma_\pi \pi_t + \gamma_g g_t + \varepsilon_{t+1}$			
Coefficient	1965:01–1984:12	1985:01–2004:12	1965:01–2004:12
γ_0	0.003^{*}	0.001	0.001
	(2.576)	(1.497)	(2.130)
ρ	0.908^{***}	0.974^{***}	0.930^{***}
	(2.946)	(109.230)	(55.536)
γ_π	0.072^{**}	0.011	0.065^{***}
	(2.857)	(0.597)	(3.568)
γ_g	-0.007^{**}	-0.008^{***}	-0.008^{***}
	(-3.267)	(-5.068)	(-5.172)

Table 3.1: The table reports parameter estimates for the Taylor rule for three different sample periods. The short rate (r_t) is represented by the three month zero coupon T-Bills. Inflation (π_t) is computed as year-on-year log differences of CPI all items. g_t denotes unemployment growth rate. The first sample (1964:01–1984:12) includes the OPEC crisis in 1973–1975 and the Fed experiment in 1979–1982, a period typically referred to as the "Great Inflation". From the mid-80s until the recent financial crisis the economy has been characterized by remarkable economic stability. This period is often called the "Great Moderation". The last column presents the parameter estimates over the whole sample. All t-statistics (in parentheses) are obtained using Newey-West adjustment with 18 lags.

that inflation plays almost no role for the short rate dynamics. Looking at the first period, however, we can see that inflation is one of the main monetary policy drivers. This simple illustration indicates the importance of including regime shifts in the Taylor rule monetary policy model.

Second, macroeconomic variables such as real activity, unemployment and inflation are known to contain predictive power for future yields beyond the one that is captured by the latent cross–sectional yield curve factors. Thus, relying entirely on latent factors and ignoring the nonlinear response of yields to changing economic fundamentals blurs our understanding of the bond risk premia. Furthermore, it also biases our forecasts of future economic conditions based on interest rates.

In this chapter we propose a new approach to model and analyze yields and their relation to the economy. We argue that the linkage between the term structure of interest rates and the state of the economy is indeed quite complex and goes beyond the simple inflation fluctuations and business cycle variation. A simple but crucial assumption is that the short end of the yield curve is subject to regime shifts. More specifically, starting from a common Taylor rule type short rate representation, we introduce a model for the yield curve, which takes into account not only the possibility of regime switches in the behavior of the Federal Reserve, but also agents' beliefs around these changes. We study the merit of our approach along three dimensions: (i) interpretable short rate dynamics, (ii) out-of-sample performance, and (iii) design and implications of no-arbitrage dynamic term structure models.

We start our empirical analysis with a simple Taylor rule policy model for the short rate. In the first step, we document the impact of non-linearities in the policy function itself. Depending on the underlying model dynamics, we find two/three limiting regimes for the short rate linked to inflation and unemployment. We document that the estimated transition probabilities show asymmetric responses to good and bad news on the short rate in different states.

Two interesting consequences emerging of our model are worth emphasizing. First, due to the flexible econometric technique for selecting the optimal number of regimes and the corresponding threshold structure, we contribute new evidence to the debate about the changes in the U.S. monetary policy over the last 50 years. In fact, we show that the idea that U.S. monetary policy can be described in terms of pre– and post–Volcker "era" (as in Table 3.1) proves to be misleading. The behavior of the Fed is instead closely related to those of the level of inflation and unemployment. Second, we demonstrate the ability of our modeling framework to forecast future economic conditions. More precisely, the regime shifts uncover three clear business cycle patterns (counter cyclical, cyclical and acyclical).

Motivated by the results for the short rate, we take our analysis a step further and design a modeling framework suitable for pricing bonds and other financial derivatives.

Our approach combines the no–arbitrage restrictions on the cross–section of bonds together with macroeconomic factors that drive bond yields. Additionally, it accommodates nonlinearities (regime shifts) in the conditional mean of the short rate and bond risk premia and therefore carries over to the entire term structure. Moreover, we are able to obtain iterative closed-form solutions for the yield. In this way, without relying on purely latent factors, the model establishes a tight link between the term structure of interest rates and directly interpretable macroeconomic quantities. To estimate the model we use a nonlinear filtering technique, unscented Kalman filter, proposed by Julier and Uhlmann (1997), and recently used in finance by Campbell, Sunderam, and Viceira (2010).

We assess the goodness of fit of our regime–switching modeling framework studying a series of model implications. We show that our model leads to reduced pricing errors in the cross-section of yields for maturities below one year. In addition, in a slightly different setting, our model is also able to reproduce the Cochrane and Piazzesi's (2005) tent–shape bond risk premia factor. Finally, we show our model's ability to accommodate unspanning and to generate the various yield curve shapes.

The content of this chapter can be summarized as follows. Section 3.2 gives a brief summary of the related literature. A description of the model for the short rate is given in Section 3.3. Section 3.4 discusses the design of a pricing framework and study its implications. Section 3.5 describes the estimation strategy and the empirical results for the short rate. Finally, Section 3.6 studies several generalizations and possible applications.

3.2 Related Literature

This chapter is closely related to two strands of the term structure literature: macro–finance models and models with regime shifts.

Macro–finance models: An important part of the term structure literature has focused on studying the relation between the term structure of interest rates and the economy. Ang and Piazzesi (2003) have drawn attention to this question by showing that the inclusion of macroeconomic factors on the top of latent components improves the predictability and shed some light into the economic nature of the underlying forces that drive changes in interest rates. Other closely related studies, which focus on the linkage between macro variables and the yield curve using little or no macroeconomic structure are, for example, Kozicki and Tinsley (2005), Ang, Piazzesi, and Wei (2006), Dewachter and Lyrio (2006), Joslin, Priebsch, and Singleton (2009). Alternative studies such as Hoerdahl, Tristani, and Vestin (2006), and Rudebusch and Wu (2008), embed the yield factors within a macroeconomic structure. This additional structure eases the interpretation of a bidirectional relation between the term structure factors and macro variables. To uncover the relationship between the yield curve and macro factors Moench (2008) uses a large panel of macro and financial variables. Diebold, Rudebusch, and Aruoba (2006) provide a macroeconomic interpretation of the dynamic Nelson-Siegel representation by combining it with a vector autoregression (VAR) representation for the macroeconomy.

Models with regime shifts. A well–known stylized fact in the term structure literature is that the yields are subject to regime shifts. Indeed, extensive empirical literature (see, for example Aït-Sahalia (1996a), Stanton (1997)) reveals that the regime switching models better describe the nonlinearities in the yields' drift (particularly in the short rate) and the volatility found in the historical interest rate data. More recent works, for example Ang and Bekaert (2002), Bansal and Zhou (2002), Dai, Singleton, and Yang (2007), Bansal, Tauchen, and Zhou (2004), and Audrino and De Giorgi (2007), have managed to link informally the succession of alternating regimes to business cycles and interest rate policies. Rudebusch and Wu (2007) suggest a link between the shift in the interest rate behavior and the dynamics of the central bank's inflation target. Ang, Bekaert, and Wei (2008) develop a regime-switching model to study real interest rates and inflation risk premia by combining latent and macroeconomic factors.

Evidences of non-linearities have been discussed in the Taylor-rule literature (see, e.g., Clarida, Gali, and Gertler (2000)) and in the structural VAR literature (see, e.g., Sims and Zha (2006), Cogley and Sargent (2005)). These studies typically focus on the structural monetary policy's response to inflation dynamics in the transition between Feds Chairman Burn, Volcker and Greenspan. In a no–arbitrage setting Ang, Boivin, Dong, and Loo-Kung (2010) and Bibikov and Chernov (2008) study the effects of monetary policy shifts on interest rates. Although these papers allow for state–dependence, the regimes remain latent and hard to interpret.

This chapter is also connected to threshold type models introduced by Tong (1978) and Tong and Lim (1980). The different smooth version of these threshold models, known in the econometrics literature also as smooth transition models, has been extensively reviewed by van Dijk, Tersvirta, and Franses (2002). Recently, da Rosa, Veiga, and Medeiros (2008) and Audrino and Medeiros (2011) combine smooth transition regressions with regression trees.

From a methodological perspective, our study can be seen as a multivariate extension of Audrino and De Giorgi's (2007) model for the yield curve. Similar smooth transition tree technique has been used by Audrino and Medeiros (2011). However, their study uses a CIR process and focuses only on the short rate. To the best of our knowledge, we are the first who develop multivariate (smooth) threshold model for the term structure of interest rates and study its implications.

3.3 Model Development: Short Rate Dynamics

Our main goal is to introduce a discrete–time model for the term structure of interest rates which incorporates both macroeconomic factors and regime shifts. To this end, we first introduce a simple model for the monetary policy. Then, we focus on a regime-switching specification, where the conditional mean of the short rate dynamics is subject to regime shifts.

3.3.1 Basic Model for the Monetary Policy

In most country economies the nominal short rate is the main instrument available to central banks for conducting monetary policy. One of the simplest representations of the monetary policy is by means of the Taylor rule. According to that rule the central bank sets the nominal short term interest rate, $(r_t)_{t\in\mathbb{Z}}$, based on the following equation:

$$r_{t+1} = \gamma_0 + \rho r_t + \gamma_\pi \pi_t + \gamma_g g_t + \varepsilon^r_{t+1}. \tag{3.1}$$

In our notation π_t denotes inflation, g_t is output gap and ε^r_t is a sequence of independent and normally distributed innovations with mean zero and variance σ_r^2. As usual, we assume that ε^r_t is independent from r_s, π_s and g_s for $s < t$.

The lagged short-term interest rate on the right-hand side of Equation (3.1) can be seen as either a smoother for the future short rate dynamics or as a proxy for additional macroeconomic, monetary policy or even financial variables. It is important to point out that in this reduced–form specification of the Taylor rule, to determine r_{t+1}, we rely only on lagged information, available at time $t + 1$. This makes our analysis suitable for forecasting and avoids issues that may arise from relying on particular parametric model for the economic fundamental and simultaneously solving equation systems.

3.3.2 Taylor Rule and Monetary Policy Regimes

Monetary policy has changed substantially over the last five decades. The very high and volatile U.S. nominal interest rates in the mid 70s and early 80s, followed by 20 years of remarkably stable bond dynamics is one of the most popular illustrations of Fed's changing respond to macroeconomic fundamentals. Taking into account this stylized fact, the aim of this subsection is to generalize the Taylor rule by allowing for shifting response of monetary policy to macroeconomic fundamentals.

One constructive way of thinking when modeling the regime shifts is to ask what kind of economic mechanism drives the changes in the monetary policy. For example, how does Fed response to inflation and unemployment changes under different economic conditions? To shed some light,borrowing the idea from the threshold models, we conjecture that the regimes are governed by thresholds partitioning the predictor space into disjoint regions. The fact that regimes are directly linked to some relevant macroeconomic and/or monetary policy variables such as inflation, unemployment growth rate or business cycles makes our model easy to interpret. Each of those regions is then characterized by a Taylor rule model with different policy response parameters. The agents observe the state of the economy today and based only on the current relevant macroeconomic information and the optimal threshold structure, they infer the future state. The transition between the individual states changes over time and is directly related to economic fundamentals.

additional term in higher order moments, as for example variance, skewness, kurtosis. Intuitively, the possibility of switch to a new regime introduces an additional source of risk. Moreover, differences in means help us generate persistence in levels as well as in the volatility of the short rate. This feature is especially important for modeling adequately the realized nominal interest rates.

Employing the idea of the tree structured smooth autoregressive threshold models, similar to da Rosa, Veiga, and Medeiros (2008) and Audrino and Medeiros (2011), we construct the regime shifts using a (smooth threshold type) binary tree structure. More formally, collecting all predictor variables into vector $x_t = (\pi_t, g_t, r_t)'$, we can design the short rate regime-switching dynamics as follows:

$$r_{t+1} = \sum_{i \in \mathbb{T}} (\gamma_{0,i} + \gamma_{\pi,i}\pi_t + \gamma_{g,i} g_t + \rho_i r_t) B_{\mathbb{J}i}(x_t; \theta_i) + \varepsilon^r_{t+1}. \tag{3.2}$$

The functions $B_{\mathbb{J}i}$, $0 < B_{\mathbb{J}i} < 1$, are known as membership functions and express the tree–structured transition probabilities. In the representation below, we follow the notation of Audrino and Medeiros (2011). The structure of every binary tree is characterized by the set of pairs (J, T), where $\mathbb{J}$ is the set of indexes of the parent (nonterminal) nodes and $\mathbb{T}$ the set of terminal nodes. The root of the tree, i.e. the node of the tree that is not a successor to any other node, is indexed by 0. Every parent node has two successors - a left and a right child. All left nodes carry even numbers and all right nodes are associated with odd numbers. More precisely, a parent node at position j generates left- and right-child nodes at positions $2j + 1$ and $2j + 2$, respectively. In addition, every non terminal node has an associated split variable $x_{s_j,t} \in x_t$, where $s_j \in \mathbb{S} = \{1, 2, 3\}$. In Equation (3.2)

$$B_{\mathbb{J}i}(x_t; \theta_i) = \prod_{j \in \mathbb{J}} G(x_{s_j,t}; \alpha_j, c_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}} \left[1 - G(x_{s_j,t}; \alpha_j, c_j)\right]^{(1-n_{i,j})(1+n_{i,j})}, \tag{3.3}$$

and

$$n_{i,j} = \begin{cases} -1 & \text{if the path to leaf } i \text{ does not include the parent node } j; \\ 0 & \text{if the path to leaf } i \text{ includes the right-child node of the parent node } j; \\ 1 & \text{if the path to leaf } i \text{ includes the left-child node of the parent node } j, \end{cases} \tag{3.4}$$

where $\mathbb{J}_i$ denotes the subset of $\mathbb{J}$ containing the indexes of the parent nodes that form the path to leaf i. Then, θ_i is the vector containing all corresponding tree structured transition probability parameters. Two examples of a binary tree are given in Figure 3.1, Panel A and Panel B.

To complete the model, the only remaining structure needed to be specified is the form of the transition function $G(x_{s_j,t}; \alpha_j, c_j)$. We focus on one of simplest, yet flexible enough choices for the transition probability G, the linear logistic function:

$$G(x_{s_j,t}; \alpha_j, c_j) = \frac{1}{1 + e^{-\alpha_j \left(x_{s_j,t} - c_j\right)}}, \qquad \alpha_j \geq 0. \tag{3.5}$$

The transition function is a bounded function of $x_{s_j,t}$, taking values between zero and one. In the equation above, the parameter c_j is the threshold value between two regimes, associated with $x_{s_j,t} \leq c_j$ and $x_{s_j,t} > c_j$, and smooth transitions between them.[1] In other words the regime decision boundary for the short rate at time $t+1$ is given by the set of points for which $\{x_{s_j,t} | x_{s_j,t} - c_j = 0\}$, $j \in \mathbb{J}_i$, $i \in \mathbb{T}$. When the relevant variable is close to its threshold value, the probability of staying in the same regime is close to 0.5, and thus giving considerable weight to the other states. On the contrary, if the variable is far away above (below) the threshold, the probability of staying in the same regime is close to one (zero), putting almost zero (100 percent) probability to the alternative state.

The second parameter of the logistic function G is the smoother α_j. As the name suggests, α_j determines the amount of smoothing across regimes. In particular, when α_j approaches infinity, the logistic function turns into an indicator function, equal one, if the relevant threshold variable is above its threshold value and zero otherwise. A smoothing variable close to zero implies equal probabilities. In addition, for identification reasons, we require the parameter α_j to be greater or equal zero. In the empirical part of this chapter we show some logistic function's examples (see Figure 2) .

Although the logistic function is symmetric around the threshold value, note that it is particularly useful for modeling asymmetric behavior. As an example, assume that the $x_{s_j,t}$ represents unemployment growth, i.e., $j = 2$ (or $x_{s_2,t} = g_t$). Then the resulting regime-switching model can be used to describe different monetary policy processes during periods

[1]An alternative is to think of the model as a constantly changing monetary policy, where the regimes are associated with each individual value of the transition function. In this chapter we focus on the first interpretation.

of positive and negative unemployment growth with a smooth transition between the two limiting regimes.

Note that in the representation above, all potential transition variables are also present in the local monetary policy Taylor rule dynamics. In fact, the vector of explanatory variables x_t, can be easily extended by including other relevant observable (e.g., macroeconomic) and/or latent (e.g., monetary policy, yield curve variables). In more general case, the transition variable can also be a (linear) combination of several variables.

3.4 Bond Pricing

A large number of both macroeconomic and finance studies uses the short rate as a building block for modeling the whole yield curve. Motivated by the more flexible, yet interpretable smooth threshold regime–switching short rate dynamics, in this section we take our analysis a step further and introduce a no-arbitrage term structure framework suitable for pricing interest rates and other derivatives.

Our study shows how to incorporate successfully short rate nonlinearities into the term structure. By imposing no–arbitrage conditions, we are able to derive iterative closed–form solutions for bond prices.

3.4.1 Model Specification

In the previous section we obtained results for the short rate without relying on any particular parametric assumptions about the distribution of the macroeconomic fundamentals. Here, we extend our approach and specify stochastic processes also for inflation and unemployment. This allows us to take our analysis a step further and design a quadratic bond pricing framework.

Short Rate Dynamics Revised

Consider again the regime-switching Taylor rule models given in Equation (3.2). We look at two economically meaningful models with different degrees of complexity (see, e.g., Ang, Boivin, Dong, and Loo-Kung (2010)). In particular, we focus on a Taylor rule specification, where the Fed's response to macroeconomic environment is changing over time. The reaction to the lagged short rate remains unaltered. We refer to this model as restricted regime-switching short rate model:

$$r_{t+1} = \sum_{i \in \mathbb{T}} (\gamma_{0,i} + \gamma_{\pi,i}\pi_t + \gamma_{g,i}g_t + \rho r_t) B_{\mathbb{J}i}(x_t; \theta_i) + \varepsilon^r_{t+1}. \tag{3.6}$$

In the second specification, we relax the assumption of constant lagged short rate response. This feature makes all conditional mean coefficients regime dependent. We refer to this model as unrestricted regime-switching short rate model:

$$r_{t+1} = \sum_{i\in\mathbb{T}} (\gamma_{0,i} + \gamma_{\pi,i}\pi_t + \gamma_{g,i}g_t + \rho_i r_t) B_{\mathbb{J}^i}(x_t; \theta_i) + \varepsilon^r_{t+1}. \tag{3.7}$$

Both representations of the regime–switching short rate considered above are quite general and cannot be used directly for bond pricing. The reason for that is the complex form of the transition probability function. To overcome this issue we use a first order approximation for the transition probabilities. This approximation leads to iterative closed-form bond pricing solutions whether restricted or unrestricted short rate specification is used. Below we formalize the procedure.[2]

Let G denotes the logistic function of Equation (3.5). Using first order Taylor expansion around the corresponding threshold, we obtain the following approximation T_1 for G:

$$\begin{aligned} T_1(x_{s_j,t}; \alpha_j, c_j) &= G(c_j; \alpha_j, c_j) + \left.\frac{\partial G(x_{s_j,t}; \alpha_j, c_j)}{\partial x_{s_j,t}}\right|_{x_t = c} (x_{s_j,t} - c_j) \\ &= \frac{1}{2} + \alpha_j \frac{1}{4}(x_{s_j,t} - c_j). \end{aligned}$$

The result gives rise to the following auxiliary model for the short rate:[3]

$$r_{t+1} = \sum_{i\in\mathbb{T}} (\gamma_{0,i} + \gamma_{\pi,i}\pi_t + \gamma_{g,i}g_t + \rho_i r_t) b_{\mathbb{J}^i}(x_t; \theta_i) + \varepsilon^r_{t+1}, \tag{3.8}$$

[2]For simplicity, until the end of this section we focus on the unrestricted model representation. The results for the restricted Taylor rule specification are straightforward. When deriving the closed-form solutions for bond prices, we will discuss the theoretical differences between the two specifications.

[3]An interesting parallel can be made between the logistic function and the uniform distribution. In fact, the same regime–switching representation as in Equation (3.8) can be obtained directly, without relying on any approximation. The idea is to replace the transition logistic function(s) assumption with a (multivariate) uniform distribution, u. Below, we discuss briefly this approach. First, since the uniform distribution of any fixed size is independent of the location itself (but it dependents on the interval size, area, volume, etc.), to pin the location down, analogously to the logistic case, we adapt the convention that the transition function is symmetrically distributed around the threshold value. Thus, the threshold plays the role of a separating hyperplane between the two individual regimes. The resulting transition probabilities are of the form $F_u = \frac{(x_{s_j,t} - c_j + l_j/2)}{l_j}$, where $F_u(\cdot)$ denotes the cumulative uniform distribution function and l_j is the length of the range. The link between the length of the interval and the volatility is given by $\sqrt{\frac{1}{12}} l_j$. In practice, due to the less flexible constant increment assumption implied by the uniform distribution, this approach leads to slightly worse empirical fit. The optimal threshold structure, however, found in both cases is quite similar.

$$
\begin{aligned}
b_{\mathbb{J}i}(x_t;\theta_i) &= \prod_{j\in\mathbb{J}} T_1(x_{s_j,t};\alpha_j,c_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}}\left[1-T_1(x_{s_j,t};\alpha_j,c_j)\right]^{(1-n_{i,j})(1+n_{i,j})} \\
&= \prod_{j\in\mathbb{J}} (b_j x_{s_j,t}+a_j)^{\frac{n_{i,j}(1+n_{i,j})}{2}}\left[1-(b_j x_{s_j,t}+a_j)\right]^{(1-n_{i,j})(1+n_{i,j})}, \qquad (3.9)
\end{aligned}
$$

where $n_{i,j}$ is given in Equation 3.4 and

$$
a_j = \frac{1}{2} - \frac{1}{4}\alpha_j c_j, \qquad b_j = \frac{1}{4}\alpha_j.
$$

To provide some intuition, suppose that the relevant threshold variable is inflation. In case of just two regimes, equation (3.8) can be conveniently rewritten as quadratic short rate model of the form:

$$
\begin{aligned}
r_{t+1} &= (\gamma_{0,1}+\gamma_{\pi,1}\pi_t+\gamma_{g,1}g_t+\rho_1 r_t)(1-a_1-b_1\pi_t) \\
&\quad + (\gamma_{0,2}+\gamma_{\pi,2}\pi_t+\gamma_{g,2}g_t+\rho_2 r_t)(a_1+b_1\pi_t)+\varepsilon^r_{t+1}
\end{aligned} \qquad (3.10)
$$

If we interpret the model as a linear model with time varying coefficients, Equation (3.10) can be conveniently rewritten in a more general form as

$$
\begin{aligned}
r_{t+1} &= \beta_0+\beta_\pi\pi_t+\beta_g g_t+\beta_r r_t+\beta_{\pi\pi}\pi_t^2+\beta_{gg}g_t^2 \\
&\quad + \beta_{rr}r_t^2+\beta_{\pi g}\pi_t g_t+\beta_{\pi r}\pi_t r_t+\beta_{gr}g_t r_t+\varepsilon^r_{t+1}, \qquad (3.11)
\end{aligned}
$$

where the parameters $\beta = (\beta_0, \beta_\pi, \beta_g, \beta_r, \beta_{\pi\pi}, \beta_{gg}, \beta_{rr}, \beta_{\pi g}, \beta_{\pi r}, \beta_{gr})'$ are functions of the short rate regression parameters, the threshold value(s) and the smoothing logistic parameter(s). The explicit representation of β is straightforward and is skipped here for brevity. In Section 5 we provide explicit solutions for β for our two optimal specifications.

Note that even though we write Equation (3.8) as a quadratic model for the short rate, in presence of more than two regimes, higher order terms could still be present. We ignore third and higher order terms (if present) for several reasons. First, from an econometrics point of view, in the context of the U.S. economy, in the last 60 years the short rate dynamics is comparatively smooth. Indeed, in our empirical study we find that the quadratic approximation is much more closer (in a mean squared sense) to the true short rate dynamics than its higher order model counterpart.[4]This feature makes the empirical impact of any higher order terms on the short rate negligible (see Table 3.2 for summary

[4]The mean squared error between our 3 regime restricted short rate model Equation (3.8) and the true short rate dynamics is 4.11e-05, whereas the mean squared distance between Equation (3.11) and the short rate is 5.40e-06.

statistics). Second, while first and second order terms are typically associated with level and volatility (correlation), the economic interpretation of higher order cross products is not very clear. Since one of the final goals is interpretable term structure dynamics, we resign of including those additional terms. Third, various empirical studies for the U.S. nominal bonds (see, e.g., Sims and Zha (2006), Cogley and Sargent (2005)) indicate the presence of just a small number (typically two or three) of regime shifts. In fact, consistent with the macroeconomics literature, the maximum number of regimes found in our empirical study is three.

3.4.2 Term Structure of Interest Rates

The previous section establishes a tight link between the nominal short rate and macroeconomic fundamentals. While the results were limited to the short rate, they serve as a fundament for the dynamics of the whole yield curve. We now embed the results for the regime–switching (quadratic) short rate into a no-arbitrage framework. In particular, a discrete–time dynamic term structure model incorporating nonlinearities in the short rate is developed. An important feature of our modeling framework is that shifts in regimes impact not only the short end of the yield curve, but also allow for possible significant impact on the entire term structure. No–arbitrage restrictions are imposed in order to ensure the consistent pricing of the cross–section of bonds with different maturities.

Following the recent literature on bond pricing (see, e.g., Dai and Singleton (2000), Dai and Singleton (2002), Ang and Piazzesi (2003)), below we specify the main assumptions characterizing the dynamic term structure model.

Short rate:

The form of the regime–switching short rate is given in Equation (3.11). The representation is general and encompasses both the restricted and unrestricted versions of the nonlinear short rate specification.

State dynamics:

Borrowing the intuition from general equilibrium models, we assume that under the physical measure, $\mathbb{P}$, in our reduced–form representation the joint behavior of the short rate, inflation and unemployment is of the form:

$$\pi_{t+1} = \mu_\pi^{\mathbb{P}} + \phi_{\pi\pi}^{\mathbb{P}} \pi_t + \phi_{\pi g}^{\mathbb{P}} g_t + \varepsilon_{t+1}^{\pi}, \tag{3.12}$$

$$g_{t+1} = \mu_g^{\mathbb{P}} + \phi_{g\pi}^{\mathbb{P}} \pi_t + \phi_{gg}^{\mathbb{P}} g_t + \varepsilon_{t+1}^{g}. \tag{3.13}$$

Again, the dynamics of our third state variable, the short rate, is given in Equation (3.11). Collecting all innovations in a vector $\varepsilon_{t+1} = (\varepsilon_{t+1}^{\pi}, \varepsilon_{t+1}^{g}, \varepsilon_{t+1}^{r})' \sim N(0, \Sigma\Sigma')$, we adopt the simplest form of the variance-covariance matrix $\Sigma\Sigma'$

$$\Sigma = \begin{pmatrix} \sigma_\pi & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & \sigma_r \end{pmatrix}.$$

The generalization to a more complex volatility matrix representation is straightforward (see, e.g., Dai and Singleton (2002), Ang, Boivin, Dong, and Loo-Kung (2010)).

While it might be economically meaningful to include the short rate as an additional predictor for inflation and unemployment dynamics (Equation (3.12) and Equation (3.13), respectively), empirically we find almost no difference in the resulting macro and yield dynamics. Once included in the macro dynamics, however, the short rate parameter estimates are only marginally significant. In addition, this specification complicates considerably the recursive solution for bond pricing and leads to a worse out-of-sample performance. The latter result is probably due to overparametrization.

Stochastic discount factor:

Our representation of the economy is complete by formulating the stochastic discount factor $M_{t,t+1}$ (SDF) between the date t and $t+1$. We assume that the SDF has the standard exponential affine form

$$M_{t,t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}\right).$$

Following the contemporaneous term structure literature (see, e.g., Duffee (2002)) we conjecture an essentially affine structure for $\lambda_t = (\lambda_{\pi,t}, \lambda_{g,t}, \lambda_{r,t})'$ with the following representation:

$$\begin{aligned} \lambda_{\pi,t+1} &= \sigma_\pi^{-1}(\lambda_{0\pi} + \lambda_{\pi\pi}\pi_t + \lambda_{\pi g}g_t) \\ \lambda_{g,t+1} &= \sigma_g^{-1}(\lambda_{0g} + \lambda_{g\pi}\pi_t + \lambda_{gg}g_t) \\ \lambda_{r,t+1} &= \sigma_r^{-1}\sum_{i\in\mathbb{T}}(\lambda_{0ri} + \lambda_{r\pi i}\pi_t + \lambda_{rgi}g_t + \lambda_{rri}r_t)b_{\mathbb{J}i}(x_t;\theta_i) \end{aligned}$$

Note that similar to Bansal and Zhou (2002) and Audrino and De Giorgi (2007), we use the same regime structure for the market price in our state dynamics. Apart from the improved empirical fit (additional flexibility), the assumption of regime–switching market price of risk for the short rate has a direct economic motivation: In the classical general equilibrium context, the stochastic discount factor (or intertemporal marginal rate of substitution) depends on consumption and inflation processes. While risk is typically attributed either to variation of consumption (in the real economy context) or inflation (in nominal economy context), recent studies (see, e.g., Campbell, Sunderam, and Viceira (2010), Hasseltoft (2008), David and Veronesi (2009)) show the importance of time variation in consumption - inflation correlation dynamics.

The economic intuition of the consumption–inflation relation is straightforward: In periods when inflation and consumption are negatively correlated (and the level of inflation is low or moderate), nominal bonds do well in bad times and hedge against consumption risk. But, bonds are risky investments in times when inflation is high, especially in times when the correlation between inflation and consumption becomes positive. In the former case, the agents are willing to accept low rates of returns, whereas in the latter case bonds are avoided unless the term premium is high. It is important to note that in contrast to the other models introduced in the term structure literature, in our approach the consumption (unemployment)–inflation interaction term is not modeled explicitly (determined a-priori), but it is a by–product of the presence of regime shifts.

Bond Pricing

Having specified all three necessary ingredients, we now derive the bond pricing equation. Let $P(t,n)$ be the price at date t of a nominal bond with n periods to maturity. Bond prices satisfy the law of one price:

$$P(t,n) = \mathbb{E}_t^{\mathbb{P}}\left(M_{t,t+1}P(t+1,n-1)\right) \tag{3.14}$$

Assuming $M_{t,t+1}P(t+1,n-1)$ are jointly lognormally distributed under $\mathbb{P}$, the price of a single period bond at time t can be written as:

$$\begin{aligned} P(t,1) &= \mathbb{E}_t^{\mathbb{P}}\left(\exp\left(-r_{t+1} - \frac{1}{2}\lambda\prime_t\lambda_t + \lambda_t\varepsilon_{t+1}\right)\right) \\ &= \exp(-r_t - \frac{1}{2}\lambda_t'\lambda_t + \frac{1}{2}\lambda_t'\lambda_t) \\ &= \exp(-r_t). \end{aligned}$$

Equivalently, under the risk neutral measure, $\mathbb{Q}$, the price of a bond with n periods to maturity can be solved as:

$$P(t,n) = \mathbb{E}_t^{\mathbb{Q}}\left(\exp\left(-\sum_{i=0}^{n-1} r_{t+i}\right)\right). \tag{3.15}$$

The dynamics of x_t under the risk neutral distribution follows from the standard drift adjustment. Details are provided in Appendix .1 and Appendix .2.

The regime–switching short rate dynamics, or more precisely its quadratic representation, motivates the following exponential quadratic bond price for maturity $n \geq 2$:[5]

$$P(t,n) = \exp\left(A(n) + B(n)'x_t + x_t'C(n)x_t\right). \tag{3.16}$$

[5]Quadratic term structured models has been studied by Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2003) and are related to a bigger class of models, whose domain is not an intersection of half-plans. Wishart term structure models, for example, use a process of stochastic positive definite matrices Gourieroux and Sufana (2003) apply Wishart autoregressive process to the term structure of interest rates.

Since the price of a bond at time t with n periods to maturity is modeled as a quadratic function of the underlying latent factors, the model-implied yield $y(t,n)$, on an n-period zero coupon bond is given by

$$y(t,n) = -\frac{\log P(t,n)}{n} = -\frac{A(n)}{n} - \frac{B(n)'}{n}x_t - x_t'\frac{C(n)}{n}x_t. \tag{3.17}$$

The coefficients $A(n)$, $B(n) = (B_\pi(n), B_g(n), B_r(n))'$ and

$$C(n) = \begin{pmatrix} C_{\pi\pi}(n) & \frac{C_{\pi g}(n)}{2} & \frac{C_{\pi r}(n)}{2} \\ \frac{C_{\pi g}(n)}{2} & C_{gg}(n) & \frac{C_{gr}(n)}{2} \\ \frac{C_{\pi r}(n)}{2} & \frac{C_{gr}(n)}{2} & C_{rr}(n) \end{pmatrix}$$

are determined recursively by the risk neutral parameters (including the threshold values and their corresponding smoothing parameters), and the variance–covariance matrix Σ of the innovations $\varepsilon = (\varepsilon_t^\pi, \varepsilon_t^g, \varepsilon_t^r)'$. For consistency, we impose the initial conditions $A(0) = 0$, $B(0) = (0,0,0)'$ and $C(0) = \mathbf{0}_{3\times3}$. In addition, we require that $A(1) = 0$, $B(1) = (0,0,-1)'$ and $C(1) = \mathbf{0}_{3\times3}$, so that $r_t = y(t,1)$. The explicit derivation for both the restricted and the unrestricted models is provided in Appendix .1 and Appendix .2, respectively. The two specifications lead to different yield curve pricing formulas.

In general, to assure positivity of yields, we can impose the following sufficient conditions: $C(n)$ is a negative semidefinite matrix and $A(n) \leq \frac{1}{4}B(n)'C(n)^{-1}B(n)$. However, since the quadratic specification arises as a direct consequence of the regime–switching short rate dynamics, we leave these restrictions as an empirical issue.

There are two key differences between our bond pricing framework and the classical quadratic term structure models. First, rather than taking a stance and modeling the quadratic terms apriori, the quadratic structure is determined endogenously and is due to the regime–switching Taylor rule dynamics. Thus, we are able to add more flexibility without specifying a full model, which typically leads to overparametrization. Second, our quadratic term structure model is asymmetric around zero and nests the linear single regime specification.[6] If there are no regimes, all quadratic terms in Equation (3.17) will be equal zero, yielding a linear relationship between yields and the economy. This property reveals an important economic interpretation of our quadratic model representation: While the volatility of the state variables is the key determinant of both global and regime–switching models, there is an additional mechanism that drives the yield dynamics, reflected by the regime shifts. This mechanism is related to the agents' valuation of the various economic states, and is determined (expressed) by (i) the threshold structure; (ii) the empirical distribution of the corresponding threshold variables within each individual regime; (iii) the smoothing parameter.

[6]Note that this is not possible neither in the classical latent quadratic nor in the Wishart term structure framework, where due to identification restrictions linear terms are not present in the term structure dynamics.

Intuitively, the possibility of changing to a new regime with different mean introduces an additional source of risk. Suppose, for example, that at time t there are only two regimes for which the transition probability function G is close to zero (or one). In that case, the agents are sure that in the next period the economy will stay in same regime and therefore the variation in bond prices is almost entirely due to the volatility of the state variables. In contrast, if G is close to 0.5 even a milder uncertainty in the macro fundamentals may generate a high uncertainty in the model (the latter is due to the fact that the difference between both local means enter the higher moments such as variance, skewness, kurtosis).[7] This increase in uncertainty can be attributed to the agents' valuation of the economy. In fact, if the agents expect a possible change in the economic policy soon (i.e. the relevant threshold variable is close to its threshold value), the variance of the bond prices is high, even though the volatility of macroeconomic variables is low. In a nutshell, using macroeconomic volatility and regime shifts, we open up two distinct channels through which uncertainty is reflected in bond prices. In that way, our framework allows us a clear differentiation between objective macroeconomic risk (risk coming from the volatility of the macroeconomic variables) and its perception by the agents (regime shifts).

3.5 Estimation Procedure and Empirical Results

We start this section with a description of the data used in our empirical analysis. Then, we discuss briefly the estimation procedure, and finally we focus on the empirical results for the short rate.

3.5.1 Data

Yield data

We use monthly U.S. zero-coupon bonds with time to maturity 3, 6, 12, 24 and 60 months. The data are obtained from the Federal Reserve Board and are constructed as in Gürkaynak, Sack, and Wright (2006). We consider the sample period January 1964 – November 2011 for a total of 575 observations. All yields are at monthly frequency.

Macroeconomic Data

The two macroeconomic variables used in our empirical study are taken from the Bureau of Labour Statistics database. We use monthly seasonally adjusted Consumer Price Index for All Urban Consumers (CPI) and Unemployment Rate (UNEMPL) as proxies for inflation and unemployment rate, respectively. We compute year-on-year log differences in CPI

[7] Let us denote by μ_1 and μ_2 the two short rate's local mean dynamics. The variance of short rate is then equal to $G(1-G)(\mu_1-\mu_2)^2+\sigma_r^2$.

to construct our measure for inflation. We define unemployment as the year-on-year log growth in the unemployment rate. Similar to the yield data, the sample period is we take into consideration is 1964:01 – 2011:11.

We chose to use unemployment growth instead of the classical output gap for several reasons. First, next to inflation, unemployment regulation is one of the main monetary policy goals. Therefore, its dynamics plays an important role in short rate variation. Second, output gap is typically hard to measure in real time. Since one of our final goals is forecasting, we resign relying on any future information and use only data available in real time.

Table 3.2 provides summary statistics of the data. As well documented in the literature, all yields and macro variables are very persistent, positively skewed and leptokurtic. Not surprisingly, the correlation between the yields and inflation is high, decreasing with time to maturity of the yields. Unemployment and yields as well unemployment and inflation are unconditionally weakly negatively (positively) correlated.

3.5.2 Finding the Optimal Short Rate Structure

We find and estimate the optimal threshold structure and, thus, the optimal number of limiting regimes using a binary tree. The main difference between the regime specification used here and the one introduced Audrino (2006) lies in the way we construct the splits. Specifically, here the splits are smooth probabilistic functions of the predictor variables. In contrast, in Audrino (2006) they are hard splits, assigning values either zero or one, depending on whether the relevant variable was below or above the corresponding threshold value.

The procedure we use here to find the optimal tree is the same as the one introduced in Audrino and Medeiros (2011). Specifically, first we grow a large tree. The most relevant variable, optimal threshold value, transition probability and local regression parameters found at each step are chosen in such a way that the conditional negative quasi log-likelihood

$$-\ell(\psi; \{x_t\}_{t=2}^{n}) = -\sum_{t=2}^{n} \log\left(\frac{1}{\sqrt{h(\psi)}} p_Z\left(\frac{(r_t - \mu_t(\psi))}{\sqrt{h(\psi)}}\right)\right) \tag{3.18}$$

is minimized. In Equation (3.18) p_Z is the standardized gaussian density, μ and h are the time-varying conditional mean and variance of the short rate and ψ is a parameter vector, consisting of all relevant short rate's transition probability, local mean and variance parameters.

To overcome the problem with overparametrization present in large tree structures, we prune the tree model. We use the Bayesian information criteria (BIC) to select the optimal tree specification. Specifically, as described by Audrino (2006), we search for a best subtree

Summary Statistics of Data

	Yield 3M	Yield 6M	Yield 1Y	Yield 2Y	Yield 5Y	CPI	UNEMPL
Panel A:							
Mean	0.056	0.057	0.058	0.061	0.065	0.042	0.010
Std	0.031	0.031	0.030	0.029	0.027	0.027	0.168
Skew	0.818	0.732	0.639	0.611	0.603	1.318	1.152
Kurt	1.433	1.087	0.818	0.50	0.449	1.633	0.982
AC(1)	0.974	0.981	0.982	0.983	0.986	0.989	0.969
AC(2)	0.946	0.957	0.959	0.963	0.971	0.970	0.939
AC(3)	0.922	0.936	0.939	0.945	0.957	0.950	0.894
Panel B:							
Yield 3M	1.000						
Yield 6M	0.991	1.000					
Yield 1Y	0.978	0.995	1.000				
Yield 2Y	0.961	0.982	0.994	1.000			
Yield 5Y	0.918	0.942	0.962	0.984	1.000		
CPI	0.708	0.701	0.679	0.654	0.613	1.000	
UNEMPL	-0.103	-0.104	-0.102	-0.088	-0.046	0.178	1.000
NBER	0.160	0.159	0.151	0.147	0.123	0.360	0.534

Table 3.2: Panel A contains summary statistics of the following yield and macro variables: yields with time to maturity 3, 6, 12, 24 and 60 months (denoted by Yield 3M, Yield 6M, Yield 1Y, Yield 2Y, Yield 5Y respectively), inflation (denoted by CPI) measured as year-on-year differences of CPI, and unemployment growth rate (denoted by UNEMPL). AC(1), AC(2), AC(3) report the first three autocorrelations of the corresponding variables. Panel B shows cross–correlations. NBER is a business cycle indicator variable, equal to one, when the economy is in recession.

of $\mathcal{P}_{opt}^{(M)}$, so that

$$\text{BIC}(\mathcal{P}_i) = -2 \cdot \ell\big(\hat{\psi}^{\mathcal{P}_i}; \{x_t\}_{t=2}^{n}\big) + \dim(\hat{\psi}^{\mathcal{P}_i}) \cdot \log(n-1) \tag{3.19}$$

is minimized. Above $\hat{\psi}^{\mathcal{P}_i}$ denotes the quasi maximum likelihood estimate for the subtree $\mathcal{P}_i$ implied by our model. Note that by using BIC, the best subtree is chosen in a purely data–driven way. In this way we overcome two of the major regime–switching issues present in the literature, namely, selecting the optimal number of regimes and determining the particular regime–switching structure. For detail treatment and asymptotic results, we refer to Audrino (2006) and Audrino and Medeiros (2011).

3.5.3 Empirical Results for the Short Rate

Depending on the underlying model specification restricted or unrestricted regime-switching Taylor rule we find that the optimal model has three or two limiting regimes, respectively. Table 3.3 and Table 3.4 present the parameter estimates of the two different specifications.

The endogenous way in which we determine the number of regimes, the best–fitting threshold structure for the short rate dynamics and the resulting local structure deserves a more detailed attention. In the remainder of this subsection we study the implications of the regime–switching Taylor rule along the following two dimensions: (i) realistic description of the monetary policy changes in U.S. economy in the last 50 years, (ii) empirical fit.

Before turning to our discussion about the regime–switching Taylor rule and the monetary policy, we make several remarks. First, we model the short–term interest rate as a reduced–form time series and not in a structural model context. Therefore, the results uncover only the unidirectional, purely data–driven relation between the short rate and macroeconomic variables. The advantage of this technique is that we do not have to rely on any parametric assumptions about the dynamics of the economic fundamentals. Second, this is a backward–looking model, which is build and estimated using information available at real time (i.e., relying only on past information). An alternative approach is to use, for example, a New Keynesian framework. However, to estimate the latter model, information about the future state of the economy is needed. Without taking a stance about the dynamics of the macroeconomic fundamentals, the forward–looking models are not suitable for forecasting purposes.

Taylor Rule - 2 Regime Specification

Based on BIC, the unrestricted short rate model has two limiting regimes characterized by the level of inflation. Figure 3.1, Panel A plots the resulting structure.

At first glance, the comparatively low value of the inflation threshold (0.0335) found here might seems at odds with the typical high inflation regime separation found in the literature.

Panel A: Restricted Taylor Rule Monetary Policy Model - 3 Regimes

Taylor Rule - 3 Regimes $r_{t+1} = \gamma_{0i} + \gamma_{\pi i}\pi_t + \gamma_{gi}g_t + \rho r_t + \varepsilon_{t+1}, \quad i = 1,2,3$			
	Regime 1	Regime 2	Regime 3
Coefficient	$CPI_t \leq 0.033$	$CPI_t > 0.033$ and $UNEMPL_t \leq 0.089$	$CPI_t > 0.033$ and $UNEMPL_t > 0.089$
γ_0	0.001^{**}	$-3.02e-5^{***}$	0.003^{***}
	(2.321)	(3.693)	(2.499)
ρ	0.960^{***}	0.960^{***}	0.960^{***}
	(34.384)	(34.384)	(34.384)
γ_π	-0.086^{**}	0.061	0.140^{**}
	(-2.552)	(0.981)	(5.404)
γ_g	-0.002^{***}	-0.039^{***}	-0.017^{***}
	(-3.422)	(-2.532)	(-3.797)
Log Likelihood			-2235.491
BIC			-4382.021
AIC			-4442.982

Panel B: Transition Probability Parameters

Transition probability parameters	
α_1	α_2
4.44^{***}	6.21^{***}
(13.125)	(14.472)

Table 3.3: The table reports parameter estimates for the restricted regime–switching Taylor rule for the sample period 1964:01 – 2011:11. The short rate (r_t) is represented by the three month zero coupon T-Bills. Inflation (π_t) is computed as year-on-year log differences of CPI all items. g_t denotes unemployment growth rate. α denotes the transition probability parameter as described in the text. t -statistics (in parenthesis) are based on heteroscedastic-consistent standard errors. Asterisks * ,** ,*** denote significance at the 10%, 5% and 1% level, respectively.

UNRESTRICTED TAYLOR RULE MONETARY POLICY MODEL - 2 REGIMES

Taylor Rule - 2 Regimes		
$r_{t+1} = \gamma_{0i} + \gamma_{\pi i}\pi_t + \gamma_{gi}g_t + \rho_i r_t + \varepsilon_{t+1}$		$i = 1, 2$
	Regime 1	Regime 2
Coefficient	$CPI_t \leq 0.033$	$CPI_t > 0.033$
γ_0	0.001	0.002
	(0.418)	(6.286)
ρ	0.913***	0.996***
	(85.872)	(36.544)
γ_π	-0.017	0.093***
	(-1.142)	(10.242)
γ_g	-0.002***	-0.001
	(-4.199)	(-0.955)
α		4.52
Log Likelihood		-2226.082
BIC		-4432.164
AIC		-4432.164

Table 3.4: The table reports parameter estimates for the unrestricted regime–switching Taylor rule for the sample period 1964:01 – 2011:11. The short rate (r_t) is represented by the three month zero coupon T-Bills. Inflation (π_t) is computed as year-on-year log differences of CPI all items. g_t denotes unemployment growth rate. α denotes the transition probability parameter as described in the text. t -statistics (in parenthesis) are based on heteroscedastic-consistent standard errors. Asterisks * ,** ,*** denote significance at the 10%, 5%, 1% level, respectively.

Best–Fitting Tree Threshold Structure

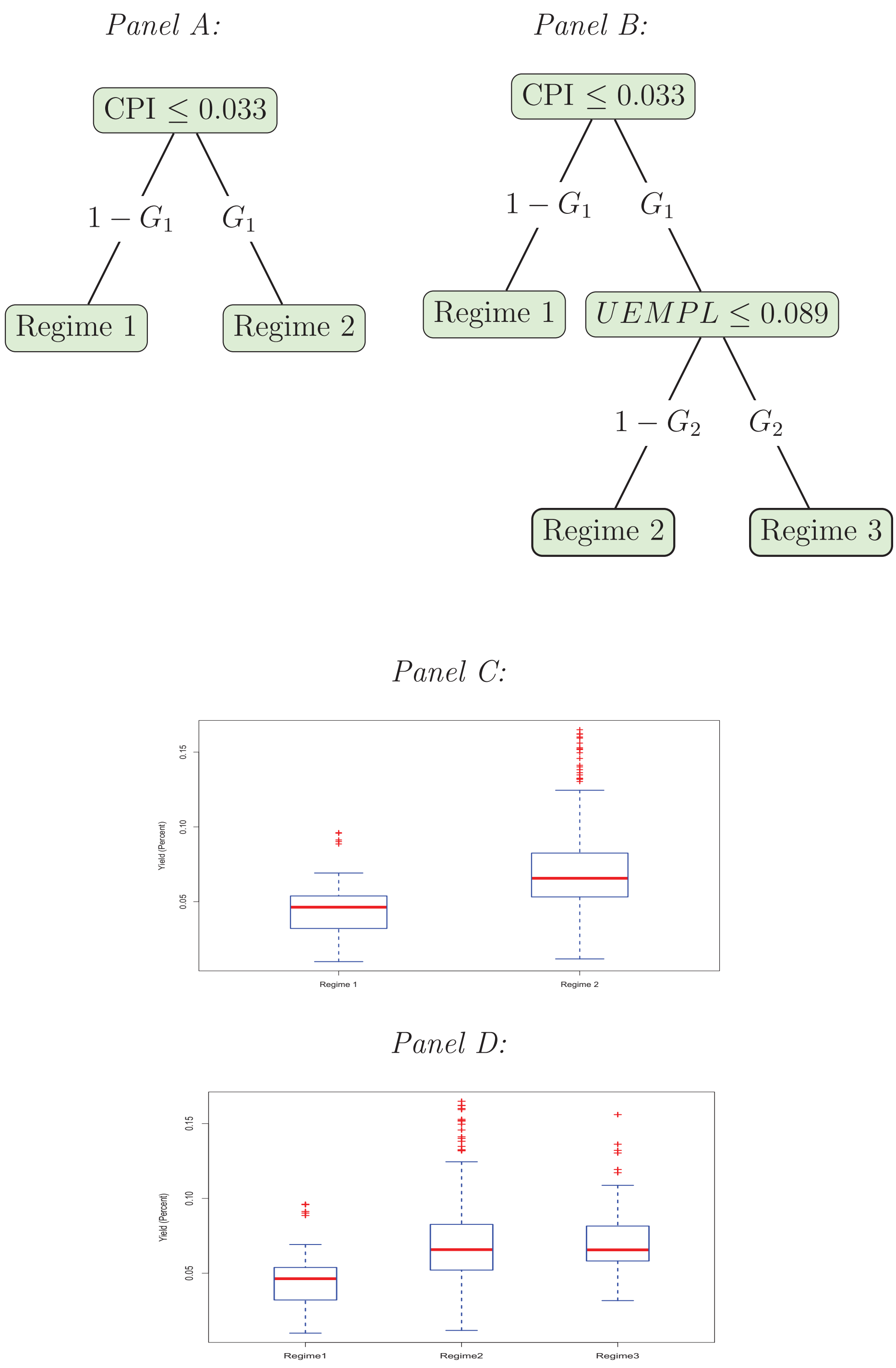


Figure 3.1: Panel A and Panel B provide graphical representation of the best–fitting unrestricted and restricted models' regime structure, respectively. Panel C shows boxplots for the short rate under the unrestricted two regime specification. Panel D presents boxplots for the short rate under the restricted three regime specification. The results are based on the sample period 1964:01 – 2011:11.

A more detailed inspection of our modeling approach, however, reveals the accuracy of our result. At the threshold value the probability of the short rate to be in one of the two regimes is one half. Therefore, the natural state of the short rate is modeled as a convex combination of the two regimes with probabilities hovering around 0.5. The more unusual (extreme) states are still described by linear combinations of the two local dynamics, but one of the microfounded models dominates significantly. In fact, the threshold value of 0.0335 almost coincide with the median of the empirical inflation distribution. Figure 3.1, Panel C presents boxplots for the empirical short rate dynamics under both regimes. Not surprisingly, the low inflation regime is associated with low short rates. High inflation periods, on the other hand, cause high short rate level.

Concerning the parameters of the Taylor rule, presented in Table 3.4, a direct comparison between the two local dynamics is not possible, since the data has different distribution in the two regions. The relation between both macro variables in the two subspaces changes as well. The correlation between CPI and UNEMPL under the first regime is negative (-0.102), whereas in the second regime it becomes comparatively high and positive (0.390). At this point, we want to emphasize that despite the fact that the dynamics of the short rate under the second regime is (close to) non-stationary, the global model is still stationary, since it is always a mixture of the two local dynamics.

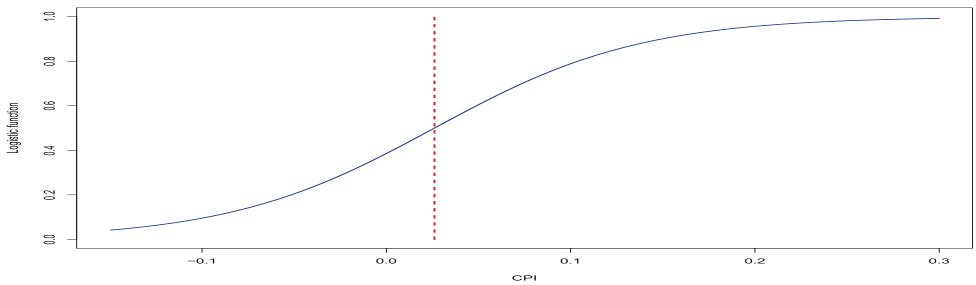
Taylor Rule - 3 Regime Specification

The BIC chooses a 3–regime model specification as optimal description for the restricted short rate dynamics. The best–fitting tree structure is presented in Figure 3.1, Panel B. The regimes here are linked to the level of inflation and unemployment rate. Similar to the two regime case the first regime is characterized by low inflation. Once inflation is above its mode value, unemployment rate starts to play major role in the regime structure. The parameter estimates for the local Taylor rule dynamics are given in Table 3.3. However, the interpretation in this case becomes quite complex. To gain some intuition, Figure 3.1, Panel D superimposes boxplots, describing the empirical distribution of the short rate conditional on each one of the three regimes. Regime 1 describes periods of low interest rate level and volatility. Regimes two and three are both associated with high level of the short rate, but differ in terms of volatility.

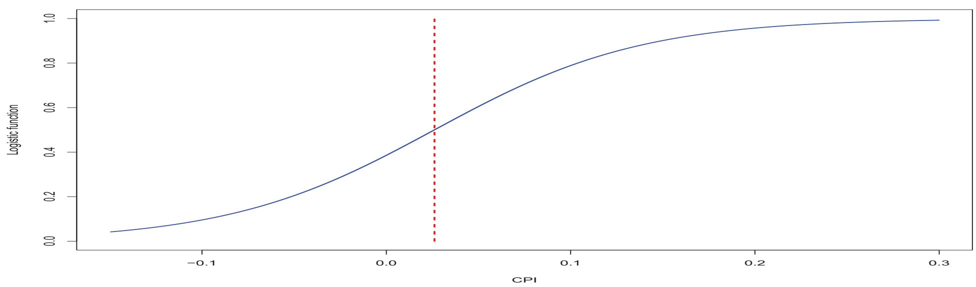
Figure 3.2 displays the resulting transition probabilities for the best–fitting restricted and unrestricted regime specifications.

Note that the figure in the upper panel (the transition probability function in the unrestricted specification) almost coincides with the transition probability in the first regime of the restricted regime–switching model. The bottom left plot displays the probability function for the second regime. Here the transition to the other two regimes is smooth, asymmetric, more sensitive to changes of unemployment rate. Finally, the bottom right

Panel A: Unrestricted Model



Panel B: Restricted Model



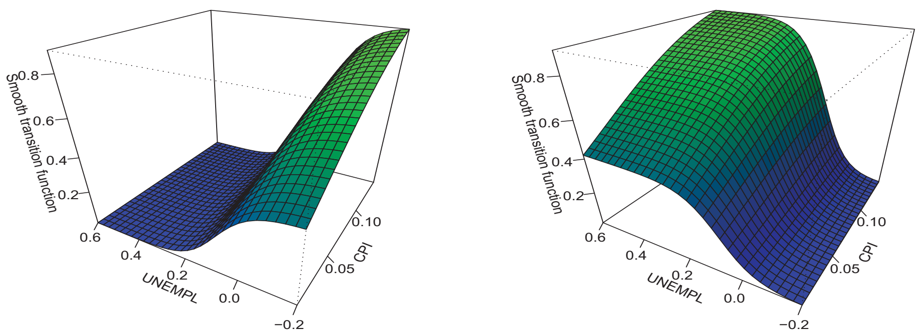
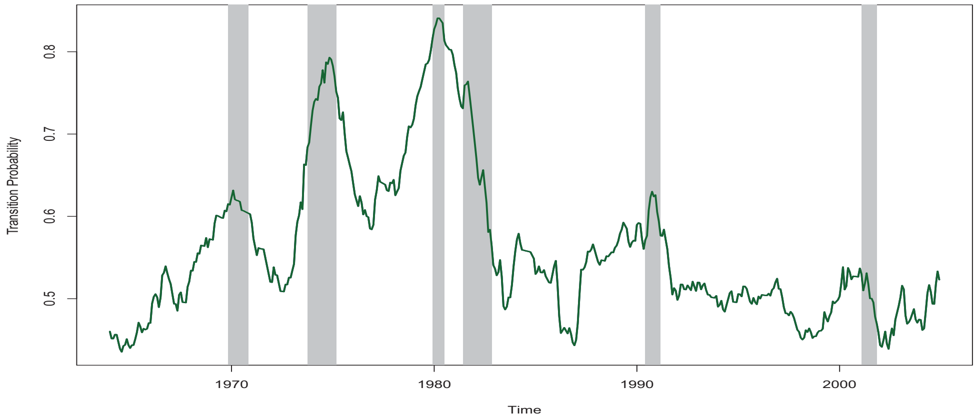


Figure 3.2: The figure shows the estimated transition probabilities associated with the best–fitting limiting regimes. Panel A displays the transition probability of the unrestricted short rate model. Panel B presents the transition probabilities of the restricted short rate representation. The upper plot of Panel B presents the transition probability model associated with regime one, the bottom left figure presents the transition probability function associated with regime two. Finally, the bottom right figure shows the results for the third regime. The probabilities are estimated over the sample period 1964:01–2011:11.

Transition Probability Dynamics over Time

Panel A: Unrestricted Model



Panel B: Restricted Model

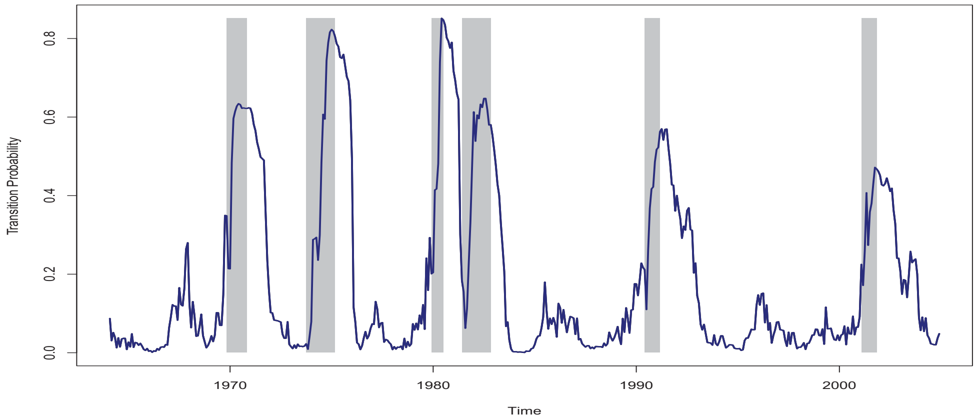


Figure 3.3: The two graphs report the estimated transition probabilities from the optimal unrestricted (Panel A) and restricted (Panel B) three–regime Taylor rule specification.

panel presents the transition probability function for the third regime. Due to the unemployment rate the shape of the probability surface is much more rougher compared to the second regime. The reaction to unemployment rate in this case is close to those of the classical tree structure with hard splits. Figure 3.3 shows the estimated transition probability dynamics over time.

To assess the ability of our model–implied regimes to describe different economic conditions, we relate the regimes with NBER business cycles. Looking at the correlation between the individual regimes and the NBER expansion/recession indicator function, we find 3 clear business cycle patterns. The first regime ($CPI \leq 0.0335$) carries clear counter cyclical component (the correlation with NBER indicator equals to -0.28). It captures mainly contraction periods, characterized by decrease in inflation. With correlation value exceeding 0.53, the third regime ($CPI > 0.0335$ and $UNEMP > 0.089$) uncovers most of the cyclical variation. This regime describes periods of expansion, typically right after recessions. The correlation between the second regime ($CPI > 0.0335$ and $UNEMP \leq 0.089$) is relative weak (equal to -0.12), showing acyclical pattern. At this point we want to emphasize that with typical correlation values of 0.3 found in the literature (see, for example, Audrino and De Giorgi (2007), Bansal, Tauchen, and Zhou (2004), Ang, Bekaert, and Wei (2008)), the results here deserve special attention.

It is important to stress that in contrast to the mid/high inflation state ($CPI > 0.0335$), where the response of unemployment growth rate is nonlinear, under the low inflation regime ($CPI \leq 0.0335$), an additional low/high unemployment growth rate split is not needed. In fact, both states - deep recession (characterized by low (negative) inflation and high unemployment growth rate) and post recession expansion (characterized by low inflation and low (negative) unemployment growth rate) can be adequately captured by a simple linear Taylor rule type dynamics.

In addition, we assess the out-of-sample forecasting performance of our two- and three–regime smooth threshold Taylor rule models. To this end, we compare the mean squared errors (MSE) of our models with those of other strong competitors. In particular, the concurrent models taken into consideration are (i) global Taylor rule model; (ii) a Taylor rule model with one estimated structural break over time; (iii) random walk; (iv) two-regime markovian Taylor rule model estimated as in Gray (1996), where we allow the regimes to depend also on inflation and unemployment rate; and (v) optimal tree threshold model estimated as in Audrino (2006), where the local dynamics follows a Taylor rule instead of CIR dynamics.

The in-sample period 1964:01 - 2004:12 is used to estimate the optimal parameters for each individual model. Given the estimated parameters, we compute the one-month ahead forecast for the out-of-sample period 2005:01 - 2011:11. The superior predictive ability tests of Hansen (2005) show that our smooth threshold Taylor rule models are able to produce accurate forecasts. The unrestricted version of the model yields the best performance overall.

The reasons why our smooth threshold model's forecast is better than forecasts based on the other models may be the following. First, from econometric perspective, our model can be seen as a time-varying weighted average of different autoregressive processes (conditional mixture model) The merit of model averaging has been shown in the term premia context in Jardet, Monfort, and Pegoraro (2011). Second, the randomness coming from the volatilities of the macroeconomic variables in not the only source through which uncertainty comes into my model. Using time–varying regime shifts, we effectively open up a second channel of uncertainty, so that macroeconomic information impacts monetary policy also in a nonlinear way. We allow this uncertainty to be associated with both level and the volatility of the macroeconomic and monetary policy variables. Thus, our model can

Out–of–Sample Results: Short Rate

Model	MSE		MAE	
Global Taylor rule	0.0095	(0.0080)	0.2215	(0.0075)
Structural Break Taylor rule	0.0119	(0.0001)	0.2521	(0.0232)
Random Walk	0.0062	(0.4742)	0.1488	(0.5097)
Markovian Two Regime Taylor rule	0.0165	(0.1052)	0.3039	(0.1264)
Smooth Threshold Two Regime Taylor rule	0.0065	(0.5126)	0.1688	(0.3432)
Smooth Threshold Three Regime Taylor rule	0.0097	(0.0266)	0.2271	(0.0102)

Table 3.5: Mean squared errors (MSE) presented in column 2 and mean absolute errors (MAE) shown in column 4 of out–of–sample forecasting performance of seven different models for the short rate, as described in detail in the text. p-values of the superior predictive ability (SPA) tests of Hansen (2005) are reported in parenthesis. The results are based on out-of-sample period, January 2005 - November 2011, for a total of 83 observations.

generate high uncertainty even if the macroeconomic volatility is low. By contrast, in the single regime VAR models used in the macroeconomic literature, the variation in yields is entirely expressed by the volatility of relevant economic fundamentals.

It is not a surprise that our smooth threshold tree structure model performs better than the classical tree model. From purely statistical point of view, when the smoothing parameter approaches infinity, the soft split becomes a hard one. Therefore, our model nests the tree model as a special case.[8] But it offers more. From economic perspective the time–varying transition probabilities reflect the agents' perception of the future economic conditions. Therefore, just the notion of possible unusual (extreme) events may have an impact on the agents' behavior. In fact, those (extreme) regimes might never occur in reality (in our sample) or be visited just very few times but they still could have important effects on the low of motion of the short rate through the agents' beliefs. In this way, under some regularity conditions, our modeling framework can capture rare events. Note that this is not possible in the classical threshold tree structure, where all thresholds should be realized values, present in the sample.

Finally, it is important to point out that our framework is very general and is also for applicable for latent factor models as well as models with economic factors only. Of course in the latent approach, in order to be able to identify the number regimes and the optimal threshold structure, we should first take a stance on the dynamics of those factors. This goes beyond the scope of the chapter and is considered as future research.

[8]Note, however that the optimal classical tree structure found above is different from the smooth one.

3.5.4 Estimation Procedure for the Whole Term Structure

Following Audrino and De Giorgi (2007) to estimate the model for the whole term structure, we employ a two–step procedure. On the first step, we focus on the short rate, compute the optimal number of limiting regimes, identify the relevant regime–switching variables and the corresponding threshold values. This has already been done in section 3.5.3. Given the optimal structure we found on step one, in step two we estimate the term structure model via maximum likelihood combined with unscented Kalman filter.

We choose to use a filtering approach for two reasons. First, macroeconomic information is in general not precisely measured. One possibility to uncover the true, unobservable economic quantities is to employ a filtering procedure. Using a version of Kalman filter enables us to estimate the dynamics of the corresponding macroeconomic fundamentals i.e. inflation and unemployment in a more accurate way. Second, without loss of generality, our modeling framework is also applicable in the presence of latent factors (e.g. stochastic volatility, long-run inflation expectation). The unscented Kalman filter procedure presented below, makes the estimation procedure for these type of models straightforward.

Unscented Kalman Filter

Nonlinearities in the transition and measurement equation makes the filtering approach introduced by Kalman (1960) not directly applicable in our setting. One way to overcome this problem is to use a particle filter. This kind of filtering, however, increases considerably the computational burden. Here instead, we use a version of a Kalman filter, known as unscented Kalman filter (UKF). It has been introduced in the literature by Julier and Uhlmann (1997) and recently applied in finance by Carr and Wu (2007) and Campbell, Sunderam, and Viceira (2010), among others. Instead of drawing a large number of point, as it will be the case if we use a particle filter, the UKF works through deterministic sampling of points in the distribution of the innovations. Those "sigma" points are chosen in such a way that they can capture at least the conditional mean and variance-covariance matrix of the state variables accurately.[9]

As mentioned earlier, the second step of the estimation procedure is based on a UKF estimation for the data of the whole term structure. In this second step we keep the structure of the regimes (i.e. the best–fitting thresholds linked to the corresponding variables and

[9] Another popular choice in the literature to cope with this kind of nonlinearities is the Extended Kalman Filter (EKF). The EKF approach employs the idea of approximating the nonlinearities analytically through first-order linearization. Moreover, the EKF requires explicit computation of Jacobians and Hessians, making the procedure difficult to tune. UKF is conceptually different. As pointed by Julier and Uhlmann (1997) it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function or transformation. Indeed, in their work Wan and van der Merwe (2001) show the superiority of the UKF in a series of experiments.

the smoothing values) estimated in Section 3.5.3 fixed. For completeness, we consider both unrestricted and restricted model specification from the previous section. We estimate the model over the sample period 1964:01-2011:11, using monthly zero coupon yields with maturities 3, 6, 12, 24 and 60 months.

3.5.5 State Space Dynamic

This subsection gives a brief description of the UKF procedure. All details can be found in the Appendix .3.

Transition Equations

To estimate the model parameters on the observed yields, inflation and unemployment, we cast the model into a state-space form and infer the three unobserved state variables. The transition equations are given by:

$$\pi_{t+1} = \mu_\pi^{\mathbb{P}} + \phi_{\pi\pi}^{\mathbb{P}}\pi_t + \phi_{\pi g}^{\mathbb{P}} g_t + \varepsilon_{t+1}^{\pi}, \tag{3.20}$$

$$g_{t+1} = \mu_g^{\mathbb{P}} + \phi_{g\pi}^{\mathbb{P}}\pi_t + \phi_{gg}^{\mathbb{P}} g_t + \varepsilon_{t+1}^{g}, \tag{3.21}$$

$$r_{t+1} = \sum_{i\in\mathbb{T}}(\gamma_{0,i} + \gamma_{\pi,i}\pi_t + \gamma_{g,i}g_t + \rho_i r_t) b_{\mathbb{J}i}(x_t;\theta_i) + \varepsilon_{t+1}^{g}. \tag{3.22}$$

The system can be written in a more parsimonious form as follows:

$$x_{t+1} = F(x_t, u) + \varepsilon_{t+1}, \tag{3.23}$$

where the state variable $x_t = (\pi_t, g_t, r_t)'$ represents the unobserved inflation, unemployment and short rate dynamics.

The covariance matrix of the transition equation is diagonal and $\varepsilon_{t+1} = (\varepsilon_{t+1}^{\pi}, \varepsilon_{t+1}^{g}, \varepsilon_{t+1}^{r})' \sim N(0, \Sigma\Sigma')$, where

$$\Sigma = \begin{pmatrix} \sigma_\pi & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & \sigma_r \end{pmatrix}$$

Measurement equations

We introduce two types of measure equations based on macroeconomic variables and yields. The measurement equations are of the form:

$$\tilde{\pi}_t = \pi_t + \eta_{1,t},$$
$$\tilde{g}_t = g_t + \eta_{2,t}.$$
$$\tilde{y}(t,n) = -\frac{A(n)}{n} - \frac{B(n)'}{n}x_t - x_t'\frac{C(n)}{n}x_t + \xi_t,$$

or in a matrix notation

$$m_t = H(x_t, u) + \vartheta_t, \tag{3.24}$$

where $m_t = (\tilde{\pi}_t, \tilde{g}_t, \tilde{y}(t,3), \tilde{y}(t,6), \tilde{y}(t,12), \tilde{y}(t,24), \tilde{y}(t,60))'$ is the collection of observed inflation, observed unemployment, and five observed yields with time to maturity 3, 6, 12, 24 and 60 months, respectively. Similar to the transition equation, the measurement shocks $\vartheta_t = (\eta_{1,t}, \eta_{2,t}, \xi_t')'$ are normally distributed and are uncorrelated with each other. The variance–covariance matrix of the measurement equation is of the form

$$Cov(\vartheta) = \begin{pmatrix} \widetilde{\sigma}_\pi^2 & 0 & \mathbf{0}_{5\times5} \\ 0 & \widetilde{\sigma}_g^2 & \mathbf{0}_{5\times5} \\ 0 & 0 & \widetilde{\sigma}_y^2\mathbf{I}_{5\times5} \end{pmatrix}.$$

Furthermore, we assume that the shocks of the measurement equation are uncorrelated with the transition equation's innovations.

Quasi-maximum likelihood estimation

Let $\hat{m}_{t+1}^-$ and $\hat{P}_{t+1}^-$ be one period ahead predictions of m_t and of their conditional volatility matrix $\hat{P}_{t+1}$, respectively, as estimated by the filter. Assuming normality of measurement errors, we can compute the quasi–log–likelihood value for each time point in our sample as follows:

$$\ell_{t+1}(\Theta) = -\frac{1}{2}\ln|\hat{P}_{t+1}^-| - \frac{1}{2}(\hat{m}_{t+1}^- - m_{t+1})'(\hat{P}_{t+1}^-)^{-1}(\hat{m}_{t+1}^- - m_{t+1}) - \tag{3.25}$$

We obtain the parameter estimates by solving

$$\hat{\Theta} = arg\max_{\Theta}\sum_{t=0}^{T-1}\ell_{t+1}(\Theta), \tag{3.26}$$

where T is the length of the in–sample period expressed in months. The starting values of the log-likelihood correspond to the the unconditional moments of the state vector. Detailed description of the algorithm is provided in Appendix .3.

To ease the computational burden and prevent invariance, we adopt several conventions. In particular, in a preliminary step we estimate the optimal parameters for inflation and unemployment in the transition equation by maximum likelihood. Given also the optimal estimates for the regime–switching process for the short rate, obtained in Section 3.5.3, we keep all the parameters in the transition equation fixed. Then, in a first step we filter the parameters for the market price of risk. Given the dynamics of the market price of risk, in a second step we estimate the optimal parameters for inflation, unemployment and the short rate. We iterate the procedure several times until convergence.

3.5.6 Model Implications

Parameter Estimates

Table 3.6 and Table 3.7 present parameter estimates for the best–fitting 3–regime restricted and 2–regime unrestricted Taylor rule models, respectively.

The estimated inflation and unemployment dynamics across both specifications is almost identical, matching closely their realized dynamics. Consistent with the macro–finance literature (see, e.g., Ang and Piazzesi (2003), Bibikov and Chernov (2010)) inflation is highly persistent with an autoregressive coefficient close to unity. Unemployment growth rate is persistent, shows however somewhat faster decay with an autoregressive parameter of 0.96. The estimated VAR representation also suggests that increase in unemployment rate leads to decrease in future inflation. This finding reflects the inverse Phillips curve type relationship between inflation and unemployment.

The second panel of Table 3.6 and 3.7 show parameter estimates for the short rate for the optimal restricted and unrestricted monetary policy models. The model estimates found here match closely the one obtained in our short rate analysis (see Section 3.5.3). In practice, however, it is not possible to identify uniquely all the short rate parameters only from the second step. This is due to the fact that the quadratic short rate model representation used for pricing, has several parameters less than the affine regime–switching counterpart.[10] To recover the local linear dynamics under each regime, we exploit, where necessary, the ratio(s) between the individual variable estimates among the regimes found in Section 3.5.3.

The above–mentioned identification issue persist in the context of the market price of risk parameters. While the quadratic specification is uniquely determined, some of the market price of risk parameters under the local regimes cannot be recovered. Therefore, we prefer to report the parameter estimates from the quadratic representation instead of making additional assumptions about the market price of risk. Keeping that in mind, we

[10] In the two–regime unrestricted specification, for example, the number of parameters is equal nine, whereas the quadratic short rate model has one parameter less. We face the same issue in the three regime restricted representation, where the number of parameters is 11. Our quadratic short rate representation has on the other hand only 8 parameters.

Parameter Estimates: Restricted Model

Macroeconomic Fundamentals (π_t, g_t):

	μ	π	g	π	g
π	2.8e-5	0.995	-0.005	0.004	0
	(0.001)	(0.006)	(0.001)	(1.01e-04)	·
g	-0.010	0.247	0.962	0	0.041
	(0.003)	(0.064)	(0.010)	·	(0.001)

Short rate (r_t):

	γ_0	γ_{pi}	γ_g	ρ
Regime 1 (CPI $\leq$ 0.0335)	-0.004	-0.068	-0.001	0.970
	(3.3e-4)	(0.012)	(0.001)	(0.008)
Regime 2 (CPI > 0.0335) & (UNEMPL $\leq$ 0.089)	0.007	0.105	-0.010	0.970
	(0.001)	(0.009)	(0.001)	(0.008)
Regime 3 (CPI > 0.0335) & (UNEMPL > 0.089)	0.007	0.153	-0.009	0.970
	(0.001)	(0.009)	(0.001)	(0.008)

Risk Premia Parameters (λ_t):

	λ_0	π	g	r	π^2	πg	g^2
π	-0.015	-0.019	-0.005				
	(3.0e-5)	(0.002)	(0.006)				
g	-0.005	0.012	-0.040				
	(2.1e-4)	(0.021)	(0.015)				
r	0.030	-0.008	0.007	-0.011	0.016	0.034	0.001
	(5.0e-5)	(0.007)	(0.002)	(1.9e-4)	(0.076)	(0.570)	(0.017)

Table 3.6: This table presents parameter estimates for the best–fitting restricted 3–regime model, specified in the text. Asymptotic standard errors, presented in parenthesis, are calculated using the outer product method. The estimates are obtained on monthly data over the sample period 1964:01 – 2011:11.

Parameter Estimates: Unrestricted Model

Macroeconomic Fundamentals (π_t, g_t):

	μ	π	g	π	g
π	2.8e-5	0.995	-0.005	0.004	0
	(0.001)	(0.006)	(0.001)	(1.01e-04)	·
g	-0.010	0.247	0.962	0	0.041
	(0.003)	(0.064)	(0.010)	·	(0.001)

Short rate (r_t):

	γ_0	γ_π	γ_g	ρ
Regime 1 (CPI ≤ 0.033)	0.001	-0.028	-0.002	0.895
	(4.2e-5)	(0.015)	(0.001)	(0.012)
Regime 2 (CPI > 0.033)	0.001	0.095	-0.010	0.995
	(4.0e-5)	(0.015)	(0.001)	(0.010)

Risk Premia Parameters (λ_t):

	λ_0	π	g	r	π^2	πg	πr
π	-0.015	-0.018	-0.004				
	(1.1e-4)	(3.9e-3)	(1.9e-4)				
g	-0.005	0.011	-0.006				
	(2.6e-3)	(0.003)	(0.037)				
r	0.025	-0.008	0.006	-0.008	-0.009	0.0132	0.029
	(3.1e-3)	(0.001)	(2.7e-3)	(1.4e-3)	(0.007)	(0.002)	(0.064)

Table 3.7: This table presents parameter estimates for the best–fitting unrestricted 2–regime model, specified in the text. Asymptotic standard errors, presented in parenthesis, are calculated using the outer product method. The estimates are obtained on monthly data over the sample period 1964:01 – 2011:11.

proceed discussing the overall impact of the individual macroeconomic factors on the term structure.

The optimal unrestricted and restricted specifications yield similar market price of risk results. The price of inflation risk is negative on average, implying higher compensation for investors. This feature becomes even more pronounced in periods of high inflation. Inflation is also riskier when unemployment and short rate are higher. The price of unemployment growth rate is positive on average, suggesting a decrease in the future short rate dynamics.

Term Structure Factors

To assess the impact of the individual risk factor on each yield of maturity n, we use Equation (3.17) and compute the factor loadings A, B and C. Figure 3.4 superimposes the resulting term structure.

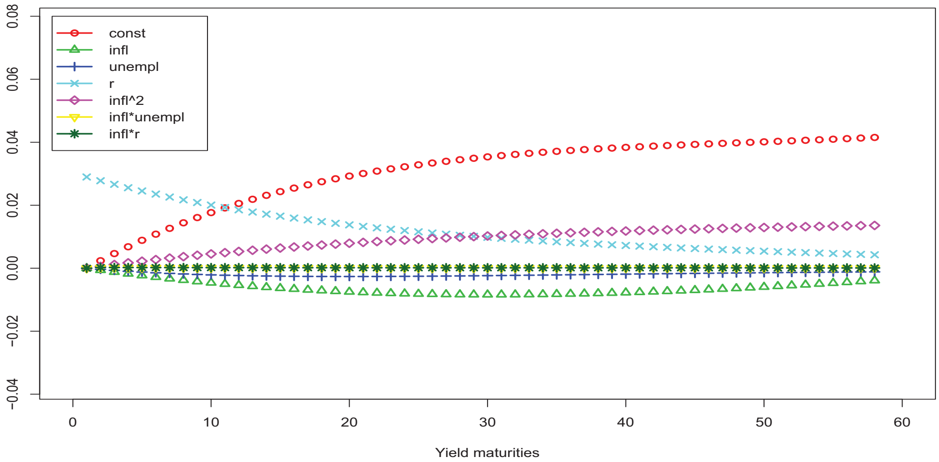
Estimates from both the restricted and unrestricted models suggest a similar positive slow decaying response to short rate. The intercept, inflation and the squared inflation factor loadings are somewhat different across both specifications. While in the restricted short rate representation higher inflation implies higher yields with the maximum loading around a maturity of 20 months, in the 2–regime unrestricted model the impact of inflation increases with time to maturity. The unemployment growth rate weight for is negative with a peak at 24 months for the unrestricted representation. Despite the differences discussed above, in practise the sum of the three factors $A(n)$, $B_\pi(n)$ and $B_g(n)$ result almost identical term structure dynamics. One explanation for this finding is that the long–term inflation and unemployment expectations implied by the classical VAR representation, used to model the underlying macroeconomic state dynamics, are almost constant (see Kim (2008)).

We now move to a discussion about the factor loadings of the higher order components $C(n)$. Those terms are typically not present in the classical affine term structure models. It is important to stress that in our framework those terms are not predetermined, but arise naturally as a result (by-product) of the regime–switching short rate dynamics. In this way, we can identify (select) the most important higher order terms, avoiding full model specification, which typically leads to overparametrization. Moreover, our framework allows us to establish a tight link between yields and directly interpretable economic quantities without resorting to purely latent factors.

Several empirical findings are worth emphasizing. First, we find that inflation volatility is positively related to bonds, increasing with time to maturity of yields. Second, the empirical correlation loading between inflation and unemployment is almost zero for all maturities. The presence of this term can be reconciled with the stock–bond correlation models in the spirit of Campbell, Sunderam, and Viceira (2010), David and Veronesi (2009), where depending on the joint dynamics of consumption (unemployment) and inflation nominal bonds can act as "inflation bets or deflation hedges". The fact that this term has zero

Term Structure Factor Loadings

Panel A: Unrestricted Model



Panel B: Restricted Model

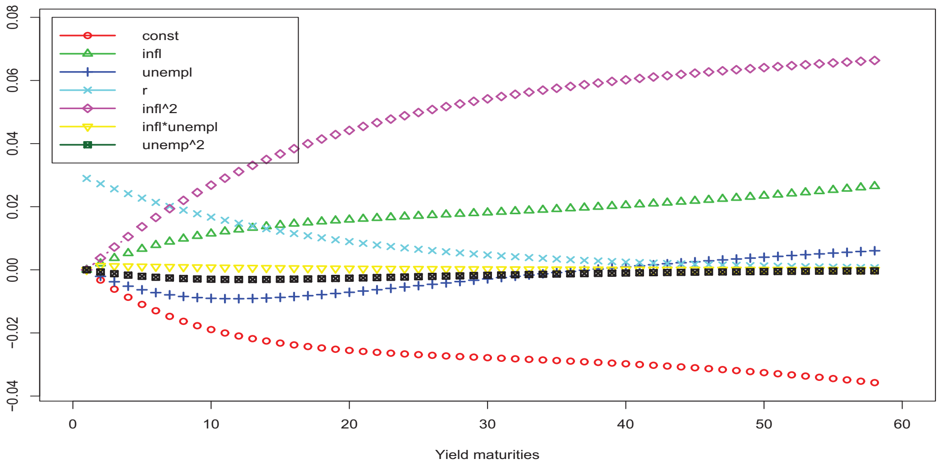


Figure 3.4: The two graphs report factor loadings from the optimal unrestricted (Panel A) and restricted (Panel B) three–regime Taylor rule specification. The weights are scaled to correspond to one standard deviation movements in the factors.

effect in pricing bonds, but has a big impact on yields' time variation makes it unspanned. Indeed, a big class of recent literature discusses the presence of unspanned factors in the term structure. Duffee (2011), for example, extract those hidden factor using a Kalman filter technique, Joslin, Priebsch, and Singleton (2009) employ two macro factors, inflation and real activity, and model them as orthogonal to yields. Ludvigson and Ng (2009b) extract them from a large panel of macroeconomic indexes. As mentioned above, the key difference between those models and ours is that in our framework this factor is selected in a purely endogenous way and is consequence of the regime–switching structure. The volatility of unemployment growth rate, $-C_{\pi r}(n)$, in the restricted model and the weight $-C_{\pi r}(n)$ in the unrestricted specification has almost no effect on the cross section of yields, contributing additional sources of unspanning.

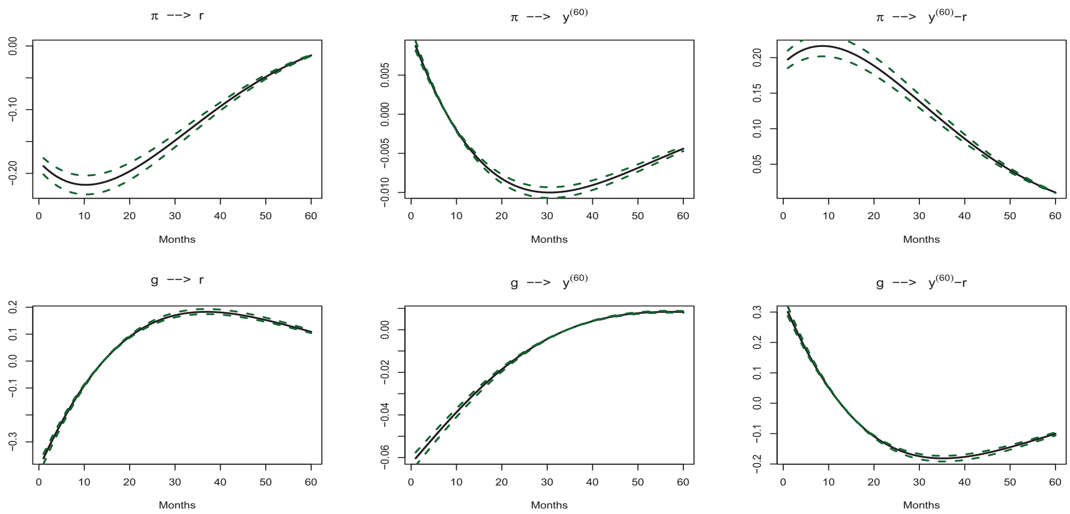
Impulse Response

While inflation and unemployment matter for the term structure dynamics, recent empirical studies (see, e.g., Beber and Brandt (2010)) show that bond returns' response to macroeconomic news differ in expansion and recession. Indeed, the previous subsection provides several insights into the nonlinear response of macroeconomic news to the term structure. By computing impulse responses, we now discuss how bond yields change with economic surprises. While the traditional affine term structure models generate symmetric response of the yield curve to good and bad news, our modeling framework allows asymmetric effects on yields.

It is interesting to see how the yield curve respond to macroeconomic risk under each regime. As mentioned in the previous section, not all of the market price of risk factors can be uniquely determined. Whenever necessary, we split the factor proportionally among the regimes. Therefore, we interpret the resulting dynamics under each individual regime with some caution. Figure 3.5, Panel A, plots impulse responses of the unrestricted model's short rate r_t, 5–year nominal bond $y^{(60)}$, to unemployment growth rate and inflation conditional on one regime being in place.

Looking at the unrestricted model specification, in times when inflation is low (Figure 3.5, Panel A), a positive unemployment shock increases both short and long term yields and decreases the term spread. When inflation is medium or high, a shock in unemployment suggests a slight initial decrease in yields. The key differences between the two regimes, however, lie in the reaction of the short rate to inflation and the repond of the long yield to unemployment growth rate. This is intriguing but consistent with the finding of Campbell, Sunderam, and Viceira (2010). The fear of deflation (hyperinflation) causes an increase (decrease) in the short rate's level over time. This fact results also a change in term spread's slope.

Panel A: Unrestricted Model, Regime 1 ($CPI \leq 0.0335$)



Panel B: Unrestricted Model, Regime 2 ($CPI > 0.0335$)

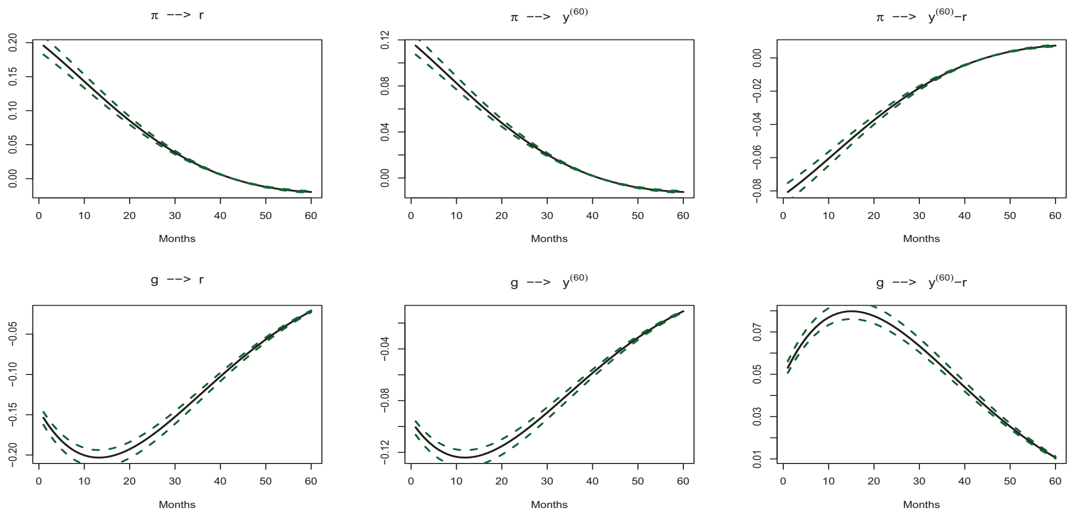
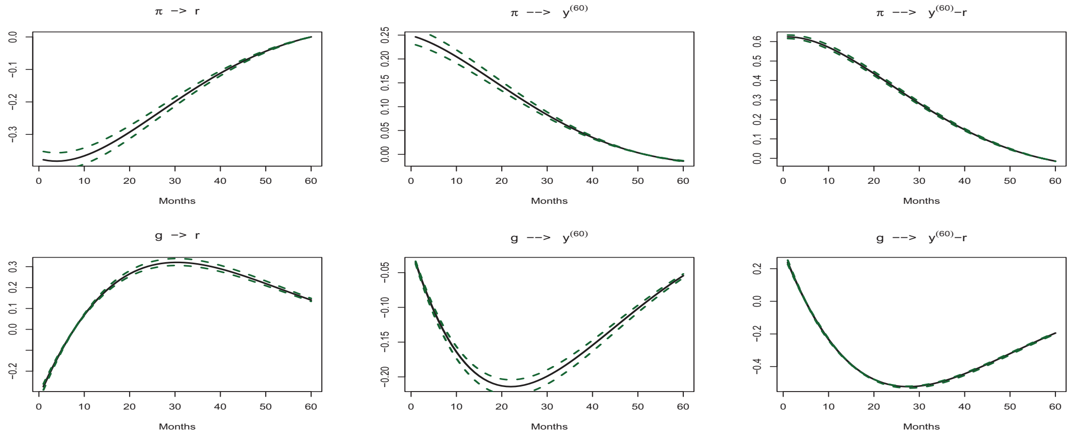
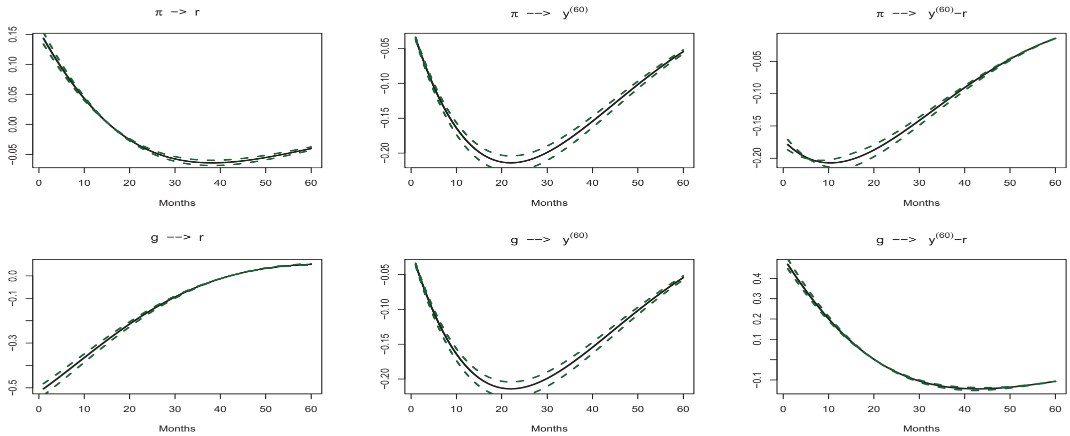


Figure 3.5: The figure shows one standard deviation impulse responses of the short rate, r_t (column 1), 5-year nominal bond, $y^{(60)}$, (column 2) and term spread $y^{(60)} - r_t$ (column 3) to inflation, π_t and unemployment growth, g_t, respectively. The individual Panels A and B display impulse responses conditional on the first and the second regime, respectively. The first row of each Panel shows the respond to inflation, whereas the second row present the impulse response to unemployment growth.

Panel A: Restricted Model, Regime 1 ($CPI \leq 0.033$)



Panel B: Restricted Model, Regime 2 ($CPI > 0.033$ and $UNEMPL \leq 0.089$)



Panel C: Restricted Model, Regime 3 ($CPI > 0.033$ and $UNEMPL > 0.089$)

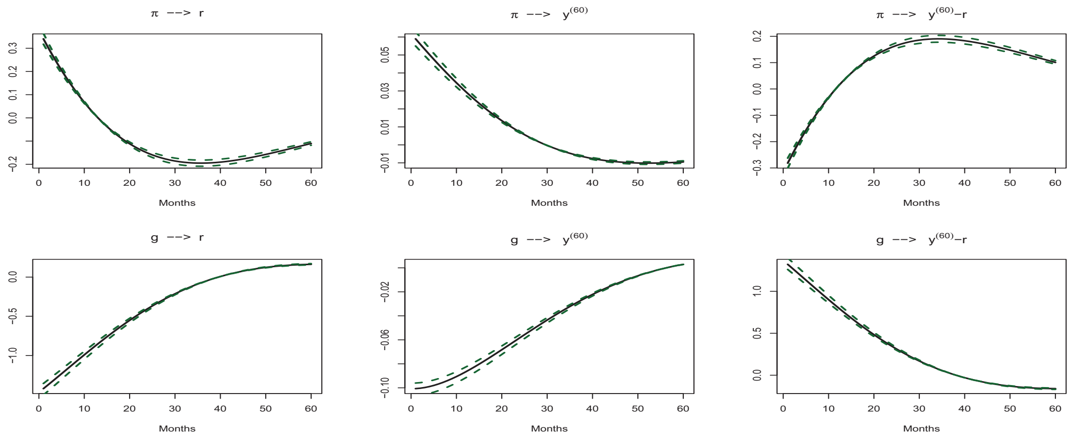


Figure 3.6: The figure shows one standard deviation impulse responses of the short rate, r_t (co lumn 1), 5-year nominal bond, $y^{(60)}$, (column 2) and term spread $y^{(60)} - r_t$ (column 3) to inflation, π_t and unemployment growth, g_t, respectively. The individual Panels A, B and C display impulse responses conditional on the first and second and the third regime, respectively. The first row of each Panel shows the respond to inflation, whereas the second row present the impulse response to unemployment growth.

Impulse responses of the restricted model conditional on the three regimes are presented in Figure 3.6. Similar to the unrestricted model, depending on the level of inflation, we document substantial differences in monetary policy's reaction to inflation shocks. More precisely, shocks to inflation causes and an increase (decrease) in short rate in times when inflation is low (high).

The reaction of the short rate to unemployment growth rate is quite similar across the three regimes. However, this is not valid for long term yield. While in the first regime and the third regime a shock in unemployment growth rate has a pronounced increasing effect on the long term yield, it has a different impact on the slope. This is quite intuitive: High unemployment growth typically occurs in periods right after recession. Since expansions are characterized with upward sloping yield curve, the unemployment shocks' effect is consistent with the business cycle patterns.

Bond Pricing

In this subsection we evaluate our optimal unrestricted and restricted regime switching models' ability for pricing and forecasting yields. Table 3.8 and Table 3.9 compare the in– and out–of–sample performance of our regime–switching models to their one regime counterpart. For completeness, we also add random walk forecasting results.

It is well known that the traditional dynamic term structure models fail to outperform out-of-sample the random walk predictions, especially over short horizons (see, e.g., Duffee (2002) and Ang and Piazzesi (2003)). Thus, by including random walk forecasts, we make an indirect comparison of the predictive ability between our models and the classical affine term structure models.

We look at four different forecasting horizons: 1, 3 , 6 and 12 months. Overall, the performance of the two regime–switching models is quite similar. The in–sample results (see Table 3.8) show that both regime-switching models lead to substantial RMSE reduction, for maturities below one year. The superior performance of the regime-switching model over different forecasting horizons retains in our out-of-sample study (Table 3.9), yet only for the short maturities.

At first sign the performance of our regime–switching models could be a bit surprising in the context of the results found in the classical macro-finance affine term structure model literature (see e.g. Dewachter and Lyrio (2006), Ang and Piazzesi (2003)), where predictability improves with time to maturity and forecasting horizon. However, by relying only on most recent macroeconomic information, we are able to capture in a more accurate way the short-run monetary policy related Phillips curve movements. As discussed by Fama (2006), to match adequately the dynamics of the long term maturities, a slow moving, highly persistent latent component is needed. This component bares the intuition of Bansal and Yaron's (2004) long–run risk economy and is associated with the central banks credibility

In–Sample Forecasting Performance: Yield Curve

Yield Maturity	Forecasting Horizon	Random Walk	2 regimes unrestricted Taylor Rule	1 Regime Taylor Rule	3 regimes restricted Taylor Rule
3M	1	0.0041	0.0032	0.0054	0.0044
6M	1	0.0055	0.0056	0.0071	0.0054
1Y	1	0.0056	0.0113	0.0114	0.0108
2Y	1	0.0049	0.0181	0.0182	0.0172
5Y	1	0.0040	0.0240	0.0270	0.0205
3M	3	0.0109	0.0091	0.0096	0.0091
6M	3	0.0098	0.0088	0.0087	0.0090
1Y	3	0.0097	0.0120	0.0118	0.0120
2Y	3	0.0087	0.0178	0.0181	0.0171
5Y	3	0.0069	0.0239	0.0270	0.0200
3M	6	0.0135	0.0132	0.0143	0.0133
6M	6	0.0132	0.0117	0.0126	0.0122
1Y	6	0.0127	0.0128	0.0130	0.0133
2Y	6	0.0115	0.0174	0.0172	0.0168
5Y	6	0.0093	0.0236	0.0268	0.0192
3M	12	0.0196	0.0186	0.0194	0.0196
6M	12	0.0183	0.0165	0.0175	0.0182
1Y	12	0.0174	0.0155	0.0162	0.0176
2Y	12	0.0158	0.0175	0.0182	0.0181
5Y	12	0.0130	0.0232	0.0267	0.0181

Table 3.8: This table compares the in–sample forecasting performance of four different models as described in the text. The root mean squared error (RMSE) for a given yield maturity $n = 3, 6, 12, 24, 60$ months and given forecasting horizon $h = 1, 3, 6, 12$ months is computed as $RMSE(n,h) = \sqrt{\frac{1}{T-h}\left(\sum_{t=1}^{T-h}\mathbb{E}_t(y(t+h,n)) - \tilde{y}(t+h,n)\right)^2}$, where $\mathbb{E}_t(y(t+h,n))$ is the expected model implied yield and $\tilde{y}(t+h,n)$ denotes its realized value. Sample period 1964:01 – 2011:11, with a total of $T = 575$ observations.

Out–of–Sample Forecasting Performance: Yield Curve

Yield Maturity	Forecasting Horizon	Random Walk	2 regimes unrestricted Taylor Rule	1 Regime Taylor Rule	3 regimes restricted Taylor Rule
3M	1	0.0028	0.0020	0.0028	0.0028
6M	1	0.0026	0.0041	0.074	0.0047
1Y	1	0.0026	0.0106	0.0183	0.0128
2Y	1	0.0028	0.0203	0.0304	0.0277
5Y	1	0.0029	0.0328	0.0359	0.0326
3M	3	0.0061	0.0044	0.0057	0.0045
6M	3	0.0060	0.0065	0.0093	0.0072
1Y	3	0.0060	0.0122	0.0195	0.0143
2Y	3	0.0060	0.0213	0.0311	0.0285
5Y	3	0.0051	0.0331	0.0361	0.0407
3M	6	0.0105	0.0091	0.0108	0.0091
6M	6	0.0102	0.0104	0.0128	0.0110
1Y	6	0.0098	0.0147	0.0215	0.0168
2Y	6	0.0093	0.0227	0.0321	0.0300
5Y	6	0.0074	0.0304	0.0364	0.0397
3M	12	0.0182	0.0169	0.0182	0.0174
6M	12	0.0174	0.0178	0.0211	0.0183
1Y	12	0.0161	0.0210	0.0264	0.0231
2Y	12	0.0141	0.0277	0.0355	0.0352
5Y	12	0.0094	0.0353	0.0376	0.0430

Table 3.9: This table compares the out–of–sample forecasting performance of four different models as described in the text. The root mean squared error (RMSE) for a given yield maturity $n = 3, 6, 12, 24, 60$ months and given forecasting horizon $h = 1, 3, 6, 12$ months is computed as $RMSE(n,h) = \sqrt{\frac{1}{T-h}\left(\sum_{t=1}^{T-h}\mathbb{E}_t(y(t+h,n) - \tilde{y}(t+h,n)\right)^2}$, where $\mathbb{E}_t(y(t+h,n))$ is the expected model implied yield and $\tilde{y}(t+h,n)$ denotes its realized value. The model is estimated over the in–sample period 1964:01 – 2004:12, and evaluated over the out–of–sample period 2005:01 – 2011:11. $T = 71$ is the total number of out–of–sample observations.

and policymakers perceptions of the long–term inflation target.

We resign including a latent factor for several reasons. First, while the inclusion of latent component improves the empirical fit of the mid and long end of the yield curve, it comes at a cost of direct economic interpretability. The improved fit, conceals also a deeper problem. Due to the highly correlated and very persistent yield structure, when estimating the model with a latent component(s), the model quickly allocates the latent factor to reducing the pricing errors on yields, making observable macroeconomic variables only marginally significant. This fact, however, is a statistical feature, rather than an economically interpretable phenomenon.

We do not see the increasing RMSE for maturities above one year implied by our model as a drawback. In fact, our empirical study complements the standard dynamic term structure framework by uncovering the macroeconomic nature of the short–run yield fluctuation, without resorting to any latent factors.

Bond Risk Premia and the Yield Curve

In this section, we study two further implications of our regime–switching framework, namely the flexibility of our model to reproduce different yield curve patterns and bond excess return predictability.

Figure 3.7 shows the ability of our models to reproduce different yield curve shapes: upward sloping, downward sloping, humped and inverted humped. While the optimal restricted and unrestricted regime–switching model fits differ at some time periods, overall they produce similar results.

An important result in the term structure literature to evaluate the goodness of a model is related to bond returns. Using a methodology, introduced in the term structure literature by Stambaugh (1988), Cochrane and Piazzesi (2005) recently show that a single linear combination of forward rates is the best predictor of bond excess returns at a wide range of maturities. The aim here is to test if our models reproduce the observed (tent shape) pattern. Since our modeling framework is mainly devoted to the short end of the yield curve, we implement Cochrane and Piazzesi's (2005) analysis in somewhat different setting. Instead of working with a one-year short rate and a one-year holding period, we take three month bond as a short rate. To avoid data overlap, we consider a three month holding period.

Despite those changes, the predictability of bond excess returns is small in our models. Regressing 3–month bond excess returns on 3-, 6- and 9-month forward rates, we are able to reproduce the observed pattern. However, the estimates from the model–implied regressions are larger in comparison to those estimated from the data. Figure 3.8 and Table 3.10 show the results.

The linear combination of data-implied forward rates explains about 25% of 3-month

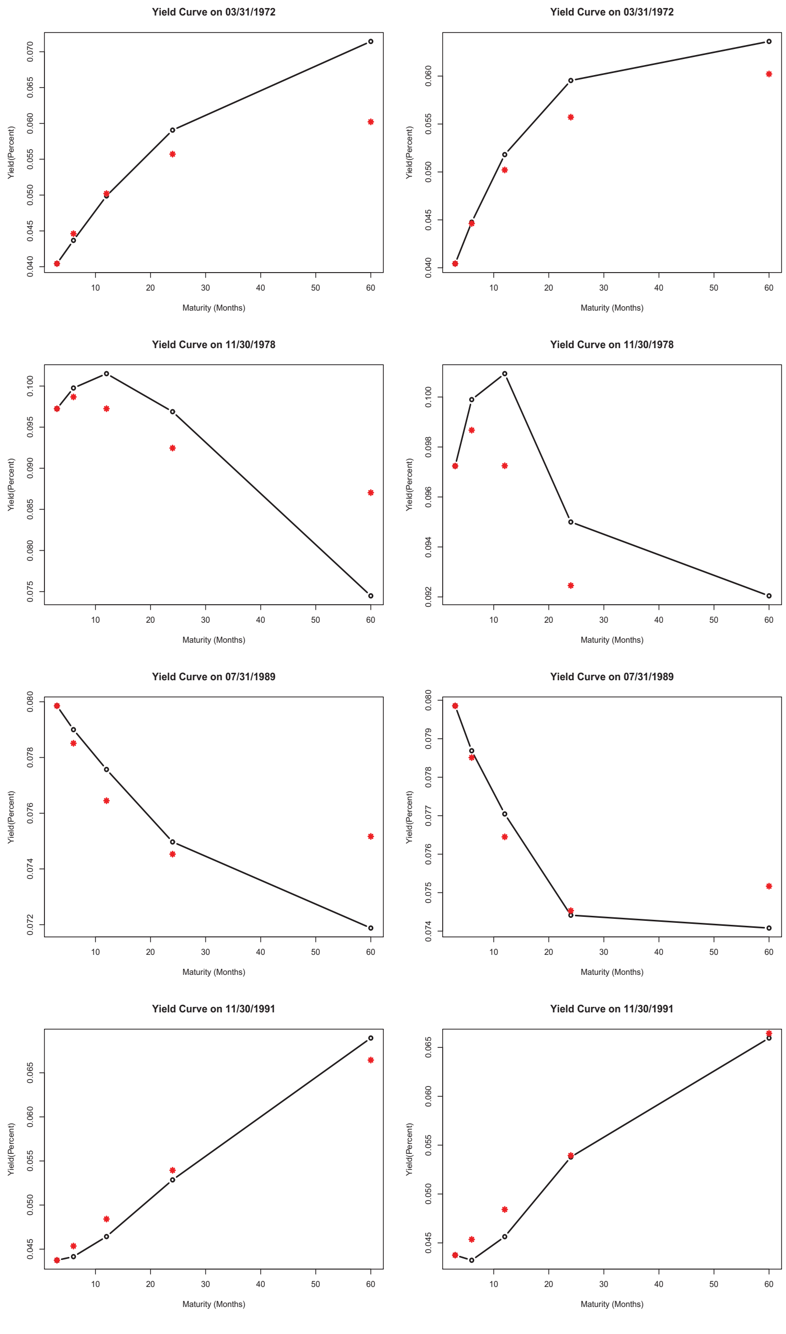


Figure 3.7: The figure shows fitted (model-based) yield curves (dotted lines) for selected dates , together with actual yields (stars). The left panel presents the results from the optimal restricted 3-regime restricted model. The right panel shows the resulting curve from the optimal 2- regime unrestricted model.

Regression Coefficients of 3-month Excess Returns on Forward Rates

Model	Data Implied	2 regimes unrestricted	Linear model	3 regimes restricted
Intercept	-0.002	-0.001***	0.001	-0.004
	(-1.462)	(-4.264)	(0.536)	(-1.456)
$f1$	-0.643***	-0.245	0.974	0.409*
	(-5.673)	(-1.910)	(0.105)	(0.224)
$f2$	0.547**	0.621	1.651	0.792
	(2.763)	(1.999)	(0.084)	(0.208)
$f3$	0.147	-0.394**	0.68	0.390
	(1.223)	(-2.020)	(0.067)	(0.197)
R^2	0.259	0.145	0.09	0.140

Table 3.10: The table shows parameter estimates from regressions of 3-month excess returns on 3-,6- and 9-month forward rates, denoted by f1,f2,f3, respectively. The results in columns are from (i) the real data–implied model (column 2); (ii) the 2-regime unrestricted regime-switching model (column 3); (iii) the global linear model (column 4); and (iv) the 3-regime restricted regime-switching model (column 5). t -statistics (in parenthesis) are based on heteroscedastic-consistent standard errors. Asterisks * ,** ,*** denote significance at the 10%, 5%, 1% level, respectively. The estimates are obtained on monthly data over the sample period 1964:01 – 2011:11.

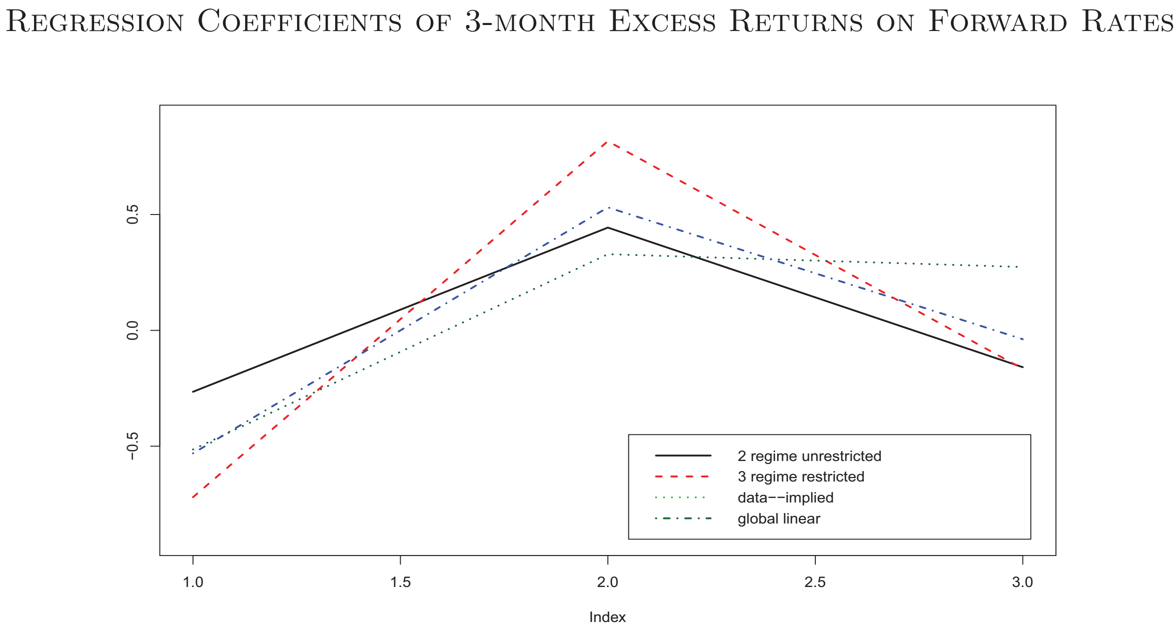


Figure 3.8: The figure shows parameter estimates from four different model implied bond excess returns on their corresponding model–implied forward rates. The models taken into consideration are (i) real data–implied model; (ii) 2-regime unrestricted regime-switching model; (iii) the global linear model; and (iv) 3-regime restricted regime-switching model. The estimates are obtained on monthly data over the sample period 1964:01 – 2011:11.

bond excess return, whereas the R^2 of our regime-switching models, account for up to 14.5% of the bond excess returns.

3.6 Generalizations and Possible Applications

This work can be extended in many directions. The first and probably the most intuitive extension is to include a long–term macroeconomic component as an additional yield curve risk factor. As already mentioned in the chapter, the absence of long–term factor(s) makes our analysis less attractive for modeling the long end of the yield curve. While incorporating a latent component, as typically done in term structure literature is straightforward, a more involved task is to link it directly to macroeconomic variables. In the spirit of Dewachter, Lyrio, and Maes (2006), Dewachter and Lyrio (2006) or Kozicki and Tinsley (2005) this component can be extracted by decomposing inflation (and other macro factors) into a permanent and transitory component. Survey data can come at favor uncovering the long–run macro risks in the economy. In this way we will be able to use macro information related to different economic frequencies.

Our results show the importance of nonlinearities that capture changing roles of macro variables in different economic conditions. However, the fact that when modeling the transition probabilities we rely only on the last observation, relates our regimes to business cycles. If the final goal is to uncover smoother long–term changes in interest rates, not only the most recent, but also lagged macroeconomic information should be added.

While our study focuses on the first moments, an important caveat is the absence of stochastic volatility. Without loss of generality stochastic volatility in both yield and macro factors could be modeled as a separate state variable(s) as typically done in the classical affine term structure framework (see, e.g., Dai and Singleton (2002), Duffee (2002)).

It is important to stress that all of the above–mentioned extensions require absolutely no alteration in our modeling framework or estimation procedure. Those generalizations are simply suggestions aimed at improving the empirical fit and/or providing better understanding of the various economic risks driving the yield curve dynamics.

The regime–switching modeling framework can be successfully applied in at least one more direction, closely related to our study. A benchmark result in the literature (see, e.g., Cochrane and Piazzesi (2005); Stambaugh (1988)) shows that a single linear combination of forward rates is the best predictor of bond returns. Unfortunately, the economic interpretation of that factor remains an open question. While a linear combination of macroeconomic factors fail to provide any reasonable explanation in that direction, the flexible way in which we model the regime shifts might be useful to uncover a possible nonlinear relationship between macro factors and bond risk premia.

.1 Restricted Model Derivation

The price of a single-period zero coupon bond at time t, denoted by $P(t,1)$ satisfies:

$$\begin{aligned} P(t,1) &= \mathbb{E}_t^{\mathbb{P}}(M_{t+1}) \\ &= \mathbb{E}_t^{\mathbb{P}}(\exp(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t\varepsilon_{t+1}^m) \\ &= \exp(-r_t - \frac{1}{2}\lambda_t'\lambda_t + \frac{1}{2}\lambda_t'\lambda_t) = \exp(-r_t) \end{aligned}$$

We assume that under the historical distribution $\mathbb{P}$ the state variables follow the processes:

$$\begin{aligned} \pi_{t+1} &= \mu_\pi^{\mathbb{P}} + \phi_{\pi\pi}^{\mathbb{P}}\pi_t + \phi_{\pi g}^{\mathbb{P}}g_t + \varepsilon_{t+1}^\pi, \\ g_{t+1} &= \mu_g^{\mathbb{P}} + \phi_{g\pi}^{\mathbb{P}}\pi_t + \phi_{gg}^{\mathbb{P}}g_t + \varepsilon_{t+1}^g, \\ r_{t+1} &= (\gamma_{01}^{\mathbb{P}} + \gamma_{\pi 1}^{\mathbb{P}}\pi_t + \gamma_{g1}^{\mathbb{P}}g_t + \rho_r^{\mathbb{P}}r_t)(1 - a_\pi - b_\pi\pi_t) + (\gamma_{02}^{\mathbb{P}} + \gamma_{\pi 2}^{\mathbb{P}}\pi_t + \gamma_{g2}^{\mathbb{P}}g_t + \rho_r^{\mathbb{P}}r_t)(a_\pi + b_\pi\pi_t)(1 - a_g - b_g g_t) + \\ &\quad + (\gamma_{03}^{\mathbb{P}} + \gamma_{\pi 3}^{\mathbb{P}}\pi_t + \gamma_{g3}^{\mathbb{P}}g_t + \rho_r^{\mathbb{P}}r_t)(a_\pi + b_\pi\pi_t)(a_g + b_g g_t) + \varepsilon_{t+1}^r. \end{aligned}$$

The market price of risk is of the form:

$$\begin{aligned} \lambda_{\pi,t+1} &= \sigma_\pi^{-1}(\lambda_{0\pi} + \lambda_{\pi\pi}\pi_t + \lambda_{\pi g}g_t) \\ \lambda_{g,t+1} &= \sigma_g^{-1}(\lambda_{0g} + \lambda_{g\pi}\pi_t + \lambda_{gg}g_t) \\ \lambda_{r,t+1} &= \sigma_r^{-1}((\lambda_{0r1} + \lambda_{r\pi 1}\pi_t + \lambda_{rg1}g_t + \lambda_{rr}r_t)(1 - a_\pi - b_\pi\pi_t) + (\lambda_{0r2} + \lambda_{r\pi 2}\pi_t + \lambda_{rg2}g_t + \lambda_{rr}r_t)(a_\pi + b_\pi\pi_t)(1 - a_g - b_g g_t) \\ &\quad + (\lambda_{0r3} + \lambda_{r\pi 3}\pi_t + \lambda_{rg3}g_t + \lambda_{rr}r_t)(a_\pi + b_\pi\pi_t)(a_g + b_g g_t)) \end{aligned}$$

Under the risk neutral distribution $\mathbb{Q}$ the dynamics of π_t, g_t and r_t is given by:

$$
\begin{aligned}
\pi_{t+1} &= \mu_\pi^{\mathbb{Q}} + \phi_{\pi\pi}^{\mathbb{Q}}\pi_t + \phi_{\pi g}^{\mathbb{Q}} g_t + \varepsilon_{t+1}^{\pi} \\
g_{t+1} &= \mu_g^{\mathbb{Q}} + \phi_{g\pi}^{\mathbb{Q}}\pi_t + \phi_{gg}^{\mathbb{Q}} g_t + \varepsilon_{t+1}^{g} \\
r_{t+1} &= (\gamma_{01}^{\mathbb{Q}} + \gamma_{\pi 1}^{\mathbb{Q}}\pi_t + \gamma_{g1}^{\mathbb{Q}} g_t + \rho_r^{\mathbb{Q}} r_t)(1 - a_\pi - b_\pi \pi_t) + (\gamma_{02}^{\mathbb{Q}} + \gamma_{\pi 2}^{\mathbb{Q}}\pi_t + \gamma_{g2}^{\mathbb{Q}} g_t + \rho_r^{\mathbb{Q}} r_t)(a_\pi + b_\pi \pi_t)(1 - a_g - b_g g_t) \\
&\quad + (\gamma_{03}^{\mathbb{Q}} + \gamma_{\pi 3}^{\mathbb{Q}}\pi_t + \gamma_{g3}^{\mathbb{Q}} g_t + \rho_r^{\mathbb{Q}} r_t)(a_\pi + b_\pi \pi_t)(a_g + b_g g_t) + \varepsilon_{t+1}^{r}.
\end{aligned} \tag{27}
$$

The link between the historical and risk neutral distributions is given by

$$
\begin{aligned}
\mu^{\mathbb{Q}} &= \mu^{\mathbb{P}} - \lambda_0 \\
\Phi^{\mathbb{Q}} &= \Phi^{\mathbb{P}} - \lambda_1,
\end{aligned}
$$

where

$$
\begin{aligned}
\mu^{\mathbb{Q}} &= (\mu_\pi^{\mathbb{Q}}, \mu_g^{\mathbb{Q}}, \gamma_{01}^{\mathbb{Q}}(1 - a_\pi) + \gamma_{02}^{\mathbb{Q}} a_\pi (1 - a_g) + \gamma_{03}^{\mathbb{Q}} a_\pi a_g)' \\
\mu^{\mathbb{P}} &= (\mu_\pi^{\mathbb{P}}, \mu_g^{\mathbb{P}}, \gamma_{01}^{\mathbb{P}}(1 - a_\pi) + \gamma_{02}^{\mathbb{P}} a_\pi (1 - a_g) + \gamma_{03}^{\mathbb{P}} a_\pi a_g)' \\
\lambda_0 &= (\lambda_{0\pi}, \lambda_{0g}, \lambda_{0r1}(1 - a_\pi) + \lambda_{0r2} a_\pi (1 - a_g) + \lambda_{0r3} a_\pi a_g)'
\end{aligned}
$$

$$
\Phi^{\mathbb{Q}} = \begin{pmatrix} \phi_{\pi\pi}^{\mathbb{Q}} & \phi_{\pi g}^{\mathbb{Q}} & 0 \\ \phi_{g\pi}^{\mathbb{Q}} & \phi_{gg}^{\mathbb{Q}} & 0 \\ \phi_{r\pi}^{\mathbb{Q}} & \phi_{rg}^{\mathbb{Q}} & \rho_r^{\mathbb{Q}} \end{pmatrix}
$$

$$
\Phi^{\mathbb{P}} = \begin{pmatrix} \phi_{\pi\pi}^{\mathbb{P}} & \phi_{\pi g}^{\mathbb{P}} & 0 \\ \phi_{g\pi}^{\mathbb{P}} & \phi_{gg}^{\mathbb{P}} & 0 \\ \phi_{r\pi}^{\mathbb{P}} & \phi_{rg}^{\mathbb{P}} & \rho_r^{\mathbb{P}} \end{pmatrix}
$$

$$\lambda_1 = \begin{pmatrix} \lambda_{\pi\pi} & \lambda_{\pi g} & 0 \\ \lambda_{g\pi} & \lambda_{gg} & 0 \\ \lambda_{r\pi} & \lambda_{rg} & \lambda_{rr} \end{pmatrix}$$

where

$$\begin{aligned}
\phi^{\mathbb{Q}}_{r\pi} &= \gamma^{\mathbb{Q}}_{01}(-b_\pi) + \gamma^{\mathbb{Q}}_{02} b_\pi(1-a_g) + \gamma^{\mathbb{Q}}_{03} b_\pi a_g + \gamma^{\mathbb{Q}}_{\pi 1}(1-a_\pi) + \gamma^{\mathbb{Q}}_{\pi 2} a_\pi(1-a_g) + \gamma^{\mathbb{Q}}_{\pi 3} a_\pi a_g \\
\phi^{\mathbb{Q}}_{rg} &= \gamma^{\mathbb{Q}}_{02} a_\pi(-b_g) + \gamma^{\mathbb{Q}}_{03} a_\pi b_g + \gamma^{\mathbb{Q}}_{g1}(1-a_\pi) + \gamma^{\mathbb{Q}}_{g2} a_\pi(1-a_g) + \gamma^{\mathbb{Q}}_{g3} a_\pi a_g \\
\phi^{\mathbb{P}}_{r\pi} &= \gamma^{\mathbb{P}}_{01}(-b_\pi) + \gamma^{\mathbb{P}}_{02} b_\pi(1-a_g) + \gamma^{\mathbb{P}}_{03} b_\pi a_g + \gamma^{\mathbb{P}}_{\pi 1}(1-a_\pi) + \gamma^{\mathbb{P}}_{\pi 2} a_\pi(1-a_g) + \gamma^{\mathbb{P}}_{\pi 3} a_\pi a_g \\
\phi^{\mathbb{P}}_{rg} &= \gamma^{\mathbb{P}}_{02} a_\pi(-b_g) + \gamma^{\mathbb{P}}_{03} a_\pi b_g + \gamma^{\mathbb{P}}_{g1}(1-a_\pi) + \gamma^{\mathbb{P}}_{g2} a_\pi(1-a_g) + \gamma^{\mathbb{P}}_{g3} a_\pi a_g \\
\lambda_{r\pi} &= \lambda_{0r1}(-b_\pi) + \lambda_{0r2} b_\pi(1-a_g) + \lambda_{0r3} b_\pi a_g + \lambda_{r\pi 1}(1-a_\pi) + \lambda_{r\pi 2} a_\pi(1-a_g) + \lambda_{r\pi 3} a_\pi a_g \\
\lambda_{rg} &= \lambda_{0r2} a_\pi(-b_g) + \lambda_{0r3} a_\pi b_g + \lambda_{r\pi 1}(1-a_\pi) + \lambda_{r\pi 2} a_\pi(1-a_g) + \lambda_{r\pi 3} a_\pi a_g
\end{aligned}$$

To ease notation, below we drop the superscript $\mathbb{Q}$. Further, the equation for the short rate can be conveniently written as:

$$r_{t+1} = \beta_0 + \beta_\pi \pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{gg} g_t^2 + \beta_{\pi g}\pi_t g_t + \varepsilon^r_{t+1},$$

where

$$\begin{aligned}
\beta_0 &= \gamma_{01}(1-a_\pi) + \gamma_{02} a_\pi(1-a_g) + \gamma_{03} a_\pi a_g \\
\beta_\pi &= \gamma_{01}(-b_\pi) + \gamma_{02} b_\pi(1-a_g) + \gamma_{03} b_\pi a_g + \gamma_{\pi 1}(1-a_\pi) + \gamma_{\pi 2} a_\pi(1-a_g) + \gamma_{\pi 3} a_\pi a_g \\
\beta_g &= \gamma_{02} a_\pi(-b_g) + \gamma_{03} a_\pi b_g + \gamma_{g1}(1-a_\pi) + \gamma_{g2} a_\pi(1-a_g) + \gamma_{g3} a_\pi a_g \\
\beta_r &= \rho_r \\
\beta_{\pi\pi} &= \gamma_{\pi 1}(-b_\pi) + \gamma_{\pi 2} b_\pi(1-a_g) + \gamma_{\pi 3} b_\pi a_g \\
\beta_{gg} &= \gamma_{g2} a_\pi(-b_g) + \gamma_{g3} a_\pi b_g \\
\beta_{\pi g} &= \gamma_{g1} b_\pi + \gamma_{02} b_\pi(-b_g) + \gamma_{03} b_\pi b_g + \gamma_{\pi 2} a_\pi(-b_g) + \gamma_{\pi 3} a_\pi b_g + \gamma_{g2} b_\pi(1-a_g) + \gamma_{g3} b_\pi a_g
\end{aligned}$$

All results below are derived under the risk neutral measure. Further, we conjecture that the price of zero-coupon bond is exponentially quadratic of the form:

$$P(t,n) = \exp\Big(A(n) + B_\pi(n)\pi_t + B_g(n)g_t + B_r(n)r_t + C_{\pi\pi}(n)\pi_t^2 + C_{gg}(n)g_t^2 + C_{\pi g}(n)\pi_t g_t\Big)$$

and

$$P(t+1,n-1) = \exp\Big(A(n-1) + B_\pi(n-1)\pi_{t+1} + B_g(n-1)g_{t+1} + B_r(n-1)r_{t+1} + C_{\pi\pi}(n-1)\pi_{t+1}^2 + C_{gg}(n-1)g_{t+1}^2 + C_{\pi g}(n-1)\pi_{t+1}g_{t+1}\Big)$$

From the bond pricing formula it immediately follows that

$$\begin{aligned}
P(t,n) &= \mathbb{E}_t(M_{t,t+1}P(t+1,n-1)) \\
&= \mathbb{E}_t\Big(\exp(-r_t + p(t+1,n-1))\Big) \\
&= \mathbb{E}_t\Big(\exp(-r_t + A(n-1) + B_\pi(n-1)\pi_{t+1} + B_g(n-1)g_{t+1} + B_r(n-1)r_{t+1} + C_{\pi\pi}(n-1)\pi_{t+1}^2 + C_{gg}(n-1)g_{t+1}^2 + C_{\pi g}(n-1)\pi_{t+1}g_{t+1})\Big) \\
&= \mathbb{E}_t\Bigg(\exp\big(-r_t + A(n-1) + B_\pi(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t + \varepsilon_{t+1}^\pi) + B_g(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t + \varepsilon_{t+1}^g) \\
&\quad + B_r(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{gg}g_t^2 + \beta_{\pi g}\pi_t g_t + \varepsilon_{t+1}^r) + C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t + \varepsilon_{t+1}^\pi)^2 \\
&\quad + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t + \varepsilon_{t+1}^\pi)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t + \varepsilon_{t+1}^g) + C_{gg}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t + \varepsilon_{t+1}^g)^2)\Bigg) \\
&= \Big(\exp(-r_t - \frac{1}{2}\lambda_t'\lambda_t + A(n-1) + B_\pi(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + B_g(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\
&\quad + B_r(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{gg}g_t^2 + \beta_{\pi g}\pi_t g_t) + C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)^2 \\
&\quad + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) + C_{gg}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t)^2)\Big) \\
&\quad \times \mathbb{E}_t(\big(\exp(h'\varepsilon_{t+1} + \varepsilon_{t+1}'\Gamma\eta_{t+1})\big),
\end{aligned}$$

where $\varepsilon = (\varepsilon^{\pi}_{t+1}, \varepsilon^{g}_{t+1}, \varepsilon^{r}_{t+1})' \sim N(0, \Sigma\Sigma')$;

$$\Sigma_\eta = \begin{pmatrix} \sigma_\pi & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & \sigma_r \end{pmatrix}$$

$$h = \begin{pmatrix} B_\pi(n-1) + 2C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + C_{\pi g}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\ B_g(n-1) + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + 2C_{gg}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\ B_r(n-1) \end{pmatrix}$$

and

$$\Gamma = \begin{pmatrix} C_{\pi\pi}(n-1) & \frac{1}{2}C_{\pi g}(n-1) & 0 \\ \frac{1}{2}C_{\pi g}(n-1) & C_{gg}(n-1) & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

To compute the conditional expectation of a quadratic Gaussian variable above, we use the following result:

$$\begin{aligned} \mathbb{E}\left(\exp(h\eta_{t+1} + \eta'_{t+1}\Gamma\eta_{t+1})\right) &= \frac{\det(\Sigma\Sigma')^{-1/2}}{\det((\Sigma\Sigma')^{-1} - 2\Gamma)^{1/2}} \exp\left(\frac{1}{2}h((\Sigma\Sigma')^{-1} - 2\Gamma)^{-1}h'\right) \\ &= \exp\left(-\frac{1}{2}\ln\det(I - 2\Sigma\Sigma'\Gamma) + \frac{1}{2}h((\Sigma\Sigma')^{-1} - 2\Gamma)^{-1}h'\right), \end{aligned}$$

For a detailed treatment in quadratic forms in random variables, their distributions, moments and various properties we refer to Mathai, Provost, and Hayakawa (1995).

To ease notation, let us denote by d_{ij} the ij-th element of the matrix $D = ((\Sigma\Sigma')^{-1} - 2\Gamma)^{-1}$.

$$D = \left(\begin{pmatrix} \frac{1}{\sigma_\pi^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_g^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_r^2} \end{pmatrix} - \begin{pmatrix} C_{\pi\pi}(n-1) & \frac{1}{2}C_{\pi g}(n-1) & 0 \\ \frac{1}{2}C_{\pi g}(n-1) & C_{gg}(n-1) & 0 \\ 0 & 0 & 0 \end{pmatrix}\right)^{-1} = \begin{pmatrix} \frac{1}{\sigma_\pi^2} - C_{\pi\pi}(n-1) & -\frac{1}{2}C_{\pi g}(n-1) & 0 \\ -\frac{1}{2}C_{\pi g}(n-1) & \frac{1}{\sigma_g^2} - C_{gg}(n-1) & 0 \\ 0 & 0 & \frac{1}{\sigma_r^2} \end{pmatrix}^{-1}$$

$$
\begin{aligned}
P(t,n) = \exp\Bigg(& - r_t + A(n-1) + B_\pi(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g} g_t) + B_g(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg} g_t) \\
& + B_r(n-1)(\beta_0 + \beta_\pi \pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{gg} g_t^2 + \beta_{\pi g}\pi_t g_t) + C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g} g_t)^2 \\
& + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g} g_t)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg} g_t) + C_{gg}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg} g_t)^2 - \frac{1}{2}\ln\det(I - 2\Sigma\Sigma'\Gamma) \\
& + \frac{1}{2} d_{11}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g} g_t) + C_{\pi g}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg} g_t))^2 \\
& + \frac{1}{2} d_{22}(B_g(n-1) + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g} g_t) + 2C_{gg}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg} g_t))^2 + \frac{1}{2} d_{33}(B_r^2(n-1)) \\
& + d_{12}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g} g_t) + C_{\pi g}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg} g_t))(B_g(n-1) + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g} g_t) \\
& \quad + 2C_{gg}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg} g_t))\Bigg)
\end{aligned}
$$

Matching coefficients on the constant, π_t, g_t, r_t, π_t^2, g_t^2 and $\pi_t g_t$ we obtain the following recursive equation for $A(n)$ and $B_\pi(n)$, $B_g(n)$, $B_r(n)$, $C_{\pi\pi}(n)$, $C_{gg}(n)$ and $C_{\pi g}(n)$

$$
A(n) = \begin{bmatrix}
A(n-1) + B_\pi(n-1)\mu_\pi + B_g(n-1)\mu_g + B_r(n-1)\beta_0 + C_{\pi\pi}(n-1)\mu_\pi^2 + C_{\pi g}(n-1)\mu_\pi\mu_g + C_{gg}(n-1)\mu_g^2 - \frac{1}{2}\ln\det(I - 2\Sigma\Sigma'\Gamma) \\
+\frac{1}{2} d_{11}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g)^2 + \frac{1}{2} d_{22}(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g)^2 + \frac{1}{2} d_{33} B_r^2(n-1) \\
+d_{12}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g)(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g)
\end{bmatrix}
$$

$$
B_\pi(n) = \begin{bmatrix}
B_\pi(n-1)\phi_{\pi\pi} + B_g(n-1)\phi_{g\pi} + B_r(n-1)\beta_\pi + 2C_{\pi\pi}(n-1)\mu_\pi\phi_{\pi\pi} + C_{\pi g}(n-1)(\mu_g\phi_{\pi\pi} + \mu_\pi\phi_{g\pi}) + 2C_{gg}(n-1)\mu_g\phi_{g\pi} \\
+\frac{1}{2} d_{11}(4C_{\pi\pi}(n-1)\phi_{\pi\pi}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g) + 2C_{\pi g}\phi_{g\pi}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g)) \\
+\frac{1}{2} d_{22}(2C_{\pi g}(n-1)\phi_{\pi\pi}(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g) + 2C_{gg}(n-1)\phi_{g\pi}(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g)) \\
+d_{12}((2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi})(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g) \\
+(C_{\pi g}(n-1)\phi_{\pi\pi} + 2C_{gg}(n-1)\phi_{g\pi})(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g))
\end{bmatrix}
$$

$$B_g(n) = \left[\begin{array}{c} B_\pi(n-1)\phi_{\pi g} + B_g(n-1)\phi_{gg} + B_r(n-1)\beta_g + 2C_{\pi\pi}(n-1)\mu_\pi\phi_{\pi g} + C_{\pi g}(n-1)(\mu_g\phi_{\pi g} + \mu_\pi\phi_{gg}) + 2C_{gg}(n-1)\mu_g\phi_{gg} \\ +\frac{1}{2}d_{11}(4C_{\pi\pi}(n-1)\phi_{\pi g}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g) + 2C_{\pi g}(n-1)\phi_{gg}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g)) \\ +\frac{1}{2}d_{22}(2C_{\pi g}(n-1)\phi_{\pi g}(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g) + 4C_{gg}(n-1)\phi_{gg}(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g))) \\ +d_{12}((2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg})(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi + 2C_{gg}(n-1)\mu_g) \\ +(C_{\pi g}(n-1)\phi_{\pi g} + 2C_{gg}(n-1)\phi_{gg})(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g)) \end{array} \right]$$

$$B_r(n) = B_r(n-1)\beta_r - 1$$

$$C_{\pi\pi}(n) = \left[\begin{array}{c} B_r(n-1)\beta_{\pi\pi} + C_{\pi\pi}(n-1)\phi_{\pi\pi}^2 + C_{\pi g}(n-1)\phi_{\pi\pi}\phi_{g\pi} + C_{gg}(n-1)\phi_{g\pi}^2 \\ +\frac{1}{2}d_{11}((2C_{\pi\pi}(n-1)\phi_{\pi\pi})^2 + (C_{\pi g}(n-1)\phi_{g\pi})^2 + 4C_{\pi\pi}(n-1)C_{\pi g}(n-1)\phi_{\pi\pi}\phi_{g\pi}) \\ +\frac{1}{2}d_{22}((C_{\pi g}(n-1)\phi_{\pi\pi})^2 + (2C_{gg}(n-1)\phi_{g\pi})^2 + 4C_{\pi g}(n-1)C_{gg}(n-1)\phi_{\pi\pi}\phi_{g\pi}) \\ +d_{12}(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi})(C_{\pi g}(n-1)\phi_{\pi\pi} + 2C_{gg}(n-1)\phi_{g\pi}) \end{array} \right]$$

$$C_{gg}(n) = \left[\begin{array}{c} B_r(n-1)\beta_{gg} + C_{\pi\pi}(n-1)\phi_{\pi g}^2 + C_{\pi g}(n-1)\phi_{\pi g}\phi_{gg} + C_{gg}(n-1)\phi_{gg}^2 \\ +\frac{1}{2}d_{11}((2C_{\pi\pi}(n-1)\phi_{\pi g})^2 + (C_{\pi g}(n-1)\phi_{gg})^2 + 4C_{\pi\pi}(n-1)C_{\pi g}(n-1)\phi_{\pi g}\phi_{gg}) \\ +\frac{1}{2}d_{22}((C_{\pi g}(n-1)\phi_{\pi g})^2 + (2C_{gg}(n-1)\phi_{gg})^2 + 4C_{\pi g}(n-1)C_{gg}(n-1)\phi_{\pi g}\phi_{gg}) \\ +d_{12}(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg})(C_{\pi g}(n-1)\phi_{\pi g} + 2C_{gg}(n-1)\phi_{gg}) \end{array} \right]$$

$$C_{\pi g}(n) = \left[\begin{array}{c} B_r(n-1)\beta_{\pi g} + 2C_{\pi\pi}(n-1)\phi_{\pi\pi}\phi_{\pi g} + C_{\pi g}(n-1)(\phi_{\pi\pi}\phi_{gg} + \phi_{\pi g}\phi_{g\pi}) + 2C_{gg}(n-1)\phi_{g\pi}\phi_{gg} \\ +\frac{1}{2}d_{11}2(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi})(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg}) \\ +\frac{1}{2}d_{22}2(C_{\pi g}(n-1)\phi_{\pi\pi} + 2C_{gg}(n-1)\phi_{g\pi})(C_{\pi g}(n-1)\phi_{\pi g} + 2C_{gg}(n-1)\phi_{gg})) \\ +d_{12}((2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi})(C_{\pi g}(n-1)\phi_{\pi g} + 2C_{gg}(n-1)\phi_{gg}) \\ +(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg})(C_{\pi g}(n-1)\phi_{\pi\pi} + 2C_{gg}(n-1)\phi_{g\pi})) \end{array} \right]$$

.2 Unrestricted Model Derivation

We assume that under the historical distribution $\mathbb{P}$ the state variables follow the processes:

$$
\begin{aligned}
\pi_{t+1} &= \mu_\pi^{\mathbb{P}} + \phi_{\pi\pi}^{\mathbb{P}}\pi_t + \phi_{\pi g}^{\mathbb{P}} g_t + \varepsilon_{t+1}^{\pi} \\
g_{t+1} &= \mu_g^{\mathbb{P}} + \phi_{g\pi}^{\mathbb{P}}\pi_t + \phi_{gg}^{\mathbb{P}} g_t + \varepsilon_{t+1}^{g} \\
r_{t+1} &= (\gamma_{01}^{\mathbb{P}} + \gamma_{\pi 1}^{\mathbb{P}}\pi_t + \gamma_{g1}^{\mathbb{P}} g_t + \rho_{r1}^{\mathbb{P}} r_t)(1 - a_\pi - b_\pi\pi_t) + (\gamma_{02}^{\mathbb{P}} + \gamma_{\pi 2}^{\mathbb{P}}\pi_t + \gamma_{g2}^{\mathbb{P}} g_t + \rho_{r2}^{\mathbb{P}} r_t)(a_\pi + b_\pi\pi_t) + \varepsilon_{t+1}^{r}.
\end{aligned}
$$

The market price of risk is of the form:

$$
\begin{aligned}
\lambda_{\pi,t+1} &= \sigma_\pi^{-1}(\lambda_{0\pi} + \lambda_{\pi\pi}\pi_t + \lambda_{\pi g} g_t) \\
\lambda_{g,t+1} &= \sigma_g^{-1}(\lambda_{0g} + \lambda_{g\pi}\pi_t + \lambda_{gg} g_t) \\
\lambda_{r,t+1} &= \sigma_r^{-1}(\lambda_{0r1} + \lambda_{r\pi 1}\pi_t + \lambda_{rg1} g_t + \lambda_{rr1} r_t)(1 - a_\pi - b_\pi\pi_t) + (\lambda_{0r2} + \lambda_{r\pi 2}\pi_t + \lambda_{rg2} g_t + \lambda_{rr2} r_t)(a_\pi + b_\pi\pi_t)
\end{aligned}
$$

Under the risk neutral distribution $\mathbb{Q}$ the dynamics of π_t, g_t and r_t is given by:

$$
\begin{aligned}
\pi_{t+1} &= \mu_\pi^{\mathbb{Q}} + \phi_{\pi\pi}^{\mathbb{Q}}\pi_t + \phi_{\pi g}^{\mathbb{Q}} g_t + \varepsilon_{t+1}^{\pi} \\
g_{t+1} &= \mu_g^{\mathbb{Q}} + \phi_{g\pi}^{\mathbb{Q}}\pi_t + \phi_{gg}^{\mathbb{Q}} g_t + \varepsilon_{t+1}^{g} \\
r_{t+1} &= (\gamma_{01}^{\mathbb{Q}} + \gamma_{\pi 1}^{\mathbb{Q}}\pi_t + \gamma_{g1}^{\mathbb{Q}} g_t + \rho_{r1}^{\mathbb{Q}} r_t)(1 - a_\pi - b_\pi\pi_t) + (\gamma_{02}^{\mathbb{Q}} + \gamma_{\pi 2}^{\mathbb{Q}}\pi_t + \gamma_{g2}^{\mathbb{Q}} g_t + \rho_{r2}^{\mathbb{Q}} r_t)(a_\pi + b_\pi\pi_t) + \varepsilon_{t+1}^{r}.
\end{aligned}
$$

The link between the historical and risk neutral distributions is given by

$$
\begin{aligned}
\mu^{\mathbb{Q}} &= \mu^{\mathbb{P}} - \lambda_0 \\
\Phi^{\mathbb{Q}} &= \Phi^{\mathbb{P}} - \lambda_1,
\end{aligned}
$$

where

$$\mu^{\mathbb{Q}} = (\mu_{\pi}^{\mathbb{Q}}, \mu_g^{\mathbb{Q}}, \gamma_{01}^{\mathbb{Q}}(1-a_\pi) + \gamma_{02}^{\mathbb{Q}} a_\pi)'$$
$$\mu^{\mathbb{P}} = (\mu_{\pi}^{\mathbb{P}}, \mu_g^{\mathbb{P}}, \gamma_{01}^{\mathbb{P}}(1-a_\pi) + \gamma_{02}^{\mathbb{P}} a_\pi)'$$
$$\lambda_0 = (\lambda_{0\pi}, \lambda_{0g}, \lambda_{0r1}(1-a_\pi) + \lambda_{0r2} a_\pi)'$$

$$\Phi^{\mathbb{Q}} = \begin{pmatrix} \phi_{\pi\pi}^{\mathbb{Q}} & \phi_{\pi g}^{\mathbb{Q}} & 0 \\ \phi_{g\pi}^{\mathbb{Q}} & \phi_{gg}^{\mathbb{Q}} & 0 \\ \gamma_{01}^{\mathbb{Q}}(-b_\pi) + \gamma_{02}^{\mathbb{Q}} b_\pi + \gamma_{\pi 1}^{\mathbb{Q}}(1-a_\pi) + \gamma_{\pi 2}^{\mathbb{Q}} a_\pi & \gamma_{g1}^{\mathbb{Q}}(1-a_\pi) + \gamma_{g2}^{\mathbb{Q}} a_\pi & \rho_{r1}^{\mathbb{Q}}(1-a_\pi) + \rho_{r2}^{\mathbb{Q}} a_\pi \end{pmatrix}$$

$$\Phi^{\mathbb{P}} = \begin{pmatrix} \phi_{\pi\pi}^{\mathbb{P}} & \phi_{\pi g}^{\mathbb{P}} & 0 \\ \phi_{g\pi}^{\mathbb{P}} & \phi_{gg}^{\mathbb{P}} & 0 \\ \gamma_{01}^{\mathbb{P}}(-b_\pi) + \gamma_{02}^{\mathbb{P}} b_\pi + \gamma_{\pi 1}^{\mathbb{P}}(1-a_\pi) + \gamma_{\pi 2}^{\mathbb{P}} a_\pi & \gamma_{g1}^{\mathbb{P}}(1-a_\pi) + \gamma_{g2}^{\mathbb{P}} a_\pi & \rho_{r1}^{\mathbb{P}}(1-a_\pi) + \rho_{r2}^{\mathbb{P}} a_\pi \end{pmatrix}$$

$$\lambda_1 = \begin{pmatrix} \lambda_{\pi\pi} & \lambda_{\pi g} & 0 \\ \lambda_{g\pi} & \lambda_{gg} & 0 \\ \lambda_{0r1}(-b_\pi) + \lambda_{0r2} b_\pi + \lambda_{r\pi 1}(1-a_\pi) + \lambda_{r\pi 2} a_\pi & \lambda_{rg1}(1-a_\pi) + \lambda_{rg2} a_\pi & \lambda_{rr1}(1-a_\pi) + \lambda_{rr2} a_\pi \end{pmatrix}$$

Further, the equation for the short rate can be conveniently written as:

$$r_{t+1} = \beta_0 + \beta_\pi \pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi} \pi_t^2 + \beta_{\pi g} \pi_t g_t + \beta_{\pi r} \pi_t r_t + \varepsilon_{t+1}^r,$$

where

$$
\begin{aligned}
\beta_0 &= \gamma_{01}(1-a_\pi) + \gamma_{02}a_\pi \\
\beta_\pi &= \gamma_{01}(-b_\pi) + \gamma_{02}b_\pi + \gamma_{\pi 1}(1-a_\pi) + \gamma_{\pi 2}a_\pi \\
\beta_g &= \gamma_{g1}(1-a_\pi) + \gamma_{g2}a_\pi \\
\beta_r &= \rho_{r1}(1-a_\pi) + \rho_{r2}a_\pi \\
\beta_{\pi\pi} &= \gamma_{\pi 1}(-b_\pi) + \gamma_{\pi 2}b_\pi \\
\beta_{\pi g} &= \gamma_{g1}(-b_\pi) + \gamma_{g2}b_\pi \\
\beta_{\pi r} &= \rho_{r1}(-b_\pi) + \rho_{r2}b_\pi
\end{aligned}
$$

All results below are derived under the risk neutral measure. To ease notation, we drop the superscript $^{\mathbb{Q}}$. Further, we conjecture that the price of zero-coupon bond is exponentially quadratic of the form:

$$P(t,n) = \exp\Big(A(n) + B_\pi(n)\pi_t + B_g(n)g_t + B_r(n)r_t + C_{\pi\pi}(n)\pi_t^2 + C_{\pi g}(n)\pi_t g_t + C_{\pi r}(n)\pi_t r_t\Big)$$

and

$$P(t+1,n-1) = \exp\Big(A(n-1) + B_\pi(n-1)\pi_{t+1} + B_g(n-1)g_{t+1} + B_r(n-1)r_{t+1} + C_{\pi\pi}(n-1)\pi_{t+1}^2 + C_{\pi g}(n-1)\pi_{t+1}g_{t+1} + C_{\pi r}(n-1)\pi_{t+1}r_{t+1}\Big)$$

From the bond pricing formula it immediately follows that

$$
\begin{aligned}
P(t,n) &= \mathbb{E}_t(M_{t,t+1}P(t+1,n-1)) \\
&= \mathbb{E}_t\Big(\exp(-r_t + p(t+1,n-1))\Big) \\
&= \mathbb{E}_t\Big(\exp(-r_t + A(n-1) + B_\pi(n-1)\pi_{t+1} + B_g(n-1)g_{t+1} + B_r(n-1)r_{t+1} + C_{\pi\pi}(n-1)\pi_{t+1}^2 + C_{\pi g}(n-1)\pi_{t+1}g_{t+1} + C_{\pi r}(n-1)\pi_{t+1}r_{t+1})\Big) \\
&= \mathbb{E}_t\Bigg(\exp\big(-r_t + A(n-1) + B_\pi(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t + \varepsilon_{t+1}^\pi) + B_g(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t + \varepsilon_{t+1}^g) \\
&\quad + B_r(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{gg}g_t^2 + \beta_{\pi g}\pi_t g_t + \varepsilon_{t+1}^r) + C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t + \varepsilon_{t+1}^\pi)^2 \\
&\quad + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t + \varepsilon_{t+1}^\pi)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t + \varepsilon_{t+1}^g) \\
&\quad + C_{\pi r}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t + \varepsilon_{t+1}^\pi)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t + \varepsilon_{t+1}^r))\Bigg) \\
&= \Big(\exp(-r_t - \frac{1}{2}\lambda_t'\lambda_t + A(n-1) + B_\pi(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + B_g(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\
&\quad + B_r(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t) + C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)^2 \\
&\quad + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) + C_{\pi r}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}r_t)\Big) \\
&\quad \times \mathbb{E}_t(\big(\exp(h'\varepsilon_{t+1} + \varepsilon_{t+1}'\Gamma\varepsilon_{t+1})\big),
\end{aligned}
$$

where $\varepsilon = (\varepsilon_{t+1}^\pi, \varepsilon_{t+1}^g, \varepsilon_{t+1}^r)' \sim N(0, \Sigma\Sigma')$;

$$
\Sigma_\eta = \begin{pmatrix} \sigma_\pi & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & \sigma_r \end{pmatrix}
$$

$$
h = \begin{pmatrix} B_\pi(n-1) + 2C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + C_{\pi g}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\ +C_{\pi r}(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t) \\ B_g(n-1) + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) \\ B_r(n-1) + C_{\pi r}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) \end{pmatrix}
$$

and

$$\Gamma = \begin{pmatrix} C_{\pi\pi}(n-1) & \frac{1}{2}C_{\pi g}(n-1) & \frac{1}{2}C_{\pi r}(n-1) \\ \frac{1}{2}C_{\pi g}(n-1) & 0 & 0 \\ \frac{1}{2}C_{\pi r}(n-1) & 0 & 0 \end{pmatrix}$$

To ease notation, let us denote by d_{ij} the ij-th element of the matrix $D = ((\Sigma\Sigma')^{-1} - 2\Gamma)^{-1}$.

$$D = \left(\begin{pmatrix} \frac{1}{\sigma_\pi^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_g^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_r^2} \end{pmatrix} - \begin{pmatrix} C_{\pi\pi}(n-1) & \frac{1}{2}C_{\pi g}(n-1) & \frac{1}{2}C_{\pi r}(n-1) \\ \frac{1}{2}C_{\pi g}(n-1) & 0 & 0 \\ \frac{1}{2}C_{\pi r}(n-1) & 0 & 0 \end{pmatrix} \right)^{-1} = \begin{pmatrix} \frac{1}{\sigma_\pi^2} - C_{\pi\pi}(n-1) & -\frac{1}{2}C_{\pi g}(n-1) & -\frac{1}{2}C_{\pi r}(n-1) \\ -\frac{1}{2}C_{\pi g}(n-1) & \frac{1}{\sigma_g^2} & 0 \\ -\frac{1}{2}C_{\pi r}(n-1) & 0 & \frac{1}{\sigma_r^2} \end{pmatrix}^{-1}$$

$$\begin{aligned}
P(t,n) = \exp\Big(& -r_t + A(n-1) + B_\pi(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + B_g(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\
& + B_r(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t) + C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)^2 \\
& + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) + C_{\pi r}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t) \\
& - \frac{1}{2}\ln\det(I - 2\Sigma\Sigma'\Gamma) + \frac{1}{2}d_{11}\big(B_\pi(n-1) + 2C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + C_{\pi g}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\
& \quad + C_{\pi r}(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t)\big)^2 \\
& + \frac{1}{2}d_{22}(B_g(n-1) + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t))^2 + \frac{1}{2}d_{33}(B_r(n-1) + C_{\pi r}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t))^2 \\
& + d_{12}\big(\big(B_\pi(n-1) + 2C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + C_{\pi g}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\
& \quad + C_{\pi r}(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t)\big)(B_g(n-1) + C_{\pi g}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t)) \\
& + d_{13}\big(\big(B_\pi(n-1) + 2C_{\pi\pi}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t) + C_{\pi g}(n-1)(\mu_g + \phi_{g\pi}\pi_t + \phi_{gg}g_t) \\
& \quad + C_{\pi r}(n-1)(\beta_0 + \beta_\pi\pi_t + \beta_g g_t + \beta_r r_t + \beta_{\pi\pi}\pi_t^2 + \beta_{\pi g}\pi_t g_t + \beta_{\pi r}\pi_t r_t)\big)(B_r(n-1) + C_{\pi r}(n-1)(\mu_\pi + \phi_{\pi\pi}\pi_t + \phi_{\pi g}g_t))\Big)
\end{aligned}$$

Matching coefficients on the constant, π_t, g_t, r_t, π_t^2, $\pi_t g_t$ and $\pi_t r_t$ we obtain the following recursive equation for $A(n)$ and $B_\pi(n)$, $B_g(n)$, $B_r(n)$, $C_{\pi\pi}(n)$, $C_{\pi g}(n)$ and $C_{\pi r}(n)$

$$A(n) = \left[\begin{array}{c} A(n-1) + B_\pi(n-1)\mu_\pi + B_g(n-1)\mu_g + B_r(n-1)\beta_0 + C_{\pi\pi}(n-1)\mu_\pi^2 + C_{\pi g}(n-1)\mu_\pi\mu_g + C_{\pi r}(n-1)\mu_g\beta_0 - \frac{1}{2}\ln\det(I - 2\Sigma\Sigma'\Gamma) \\ +\frac{1}{2}d_{11}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0)^2 + \frac{1}{2}d_{22}(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi)^2 + \frac{1}{2}d_{33}(B_r(n-1) + C_{\pi r}(n-1)\mu_\pi)^2 \\ +d_{12}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0)(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi) \\ +d_{13}(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0)(B_r(n-1) + C_{\pi r}(n-1)\mu_\pi) \end{array} \right]$$

$$B_\pi(n) = \left[\begin{array}{c} B_\pi(n-1)\phi_{\pi\pi} + B_g(n-1)\phi_{g\pi} + B_r(n-1)\beta_\pi + 2C_{\pi\pi}(n-1)\mu_\pi\phi_{\pi\pi} + C_{\pi g}(n-1)(\mu_g\phi_{\pi\pi} + \mu_\pi\phi_{g\pi}) + C_{\pi r}(n-1)(\mu_\pi\beta_\pi + \beta_0\phi_{\pi\pi}) \\ +\frac{1}{2}d_{11}2(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C_{\pi r}(n-1)\beta_\pi)(B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0) \\ +\frac{1}{2}d_{22}2(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi)C_{\pi g}(n-1)\phi_{\pi\pi} + \frac{1}{2}d_{33}2(B_r(n-1) + C_{\pi r}(n-1)\mu_\pi)C_{\pi r}(n-1)\phi_{\pi\pi} \\ +d_{12}((B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C\pi r(n-1)\beta_0)C_{\pi g}(n-1)\phi_{\pi\pi} \\ +(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C\pi r(n-1)\beta_\pi)(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi)) \\ +d_{13}((B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C\pi r(n-1)\beta_0)C_{\pi r}(n-1)\phi_{\pi\pi} \\ +(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C\pi r(n-1)\beta_\pi)(B_r(n-1) + C_{\pi r}(n-1)\mu_\pi)) \end{array} \right]$$

$$B_g(n) = \left[\begin{array}{c} B_\pi(n-1)\phi_{\pi g} + B_g(n-1)\phi_{gg} + B_r(n-1)\beta_g + 2C_{\pi\pi}\mu_\pi\phi_{\pi g} + C_{\pi g}(n-1)(\mu_g\phi_{\pi g} + \mu_\pi\phi_{gg}) + C_{\pi r}(n-1)(\mu_\pi\beta_g + \beta_0\phi_{\pi g}) \\ +\frac{1}{2}d_{11}2(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg} + C_{\pi r}(n-1)\beta_g)(B_\pi(n-1) + 2C_{\pi\pi}\mu_\pi + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0) \\ +\frac{1}{2}d_{22}2(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi)C_{\pi g}(n-1)\phi_{\pi g} + \frac{1}{2}d_{33}2(B_r(n-1) + C_{\pi r}(n-1)\mu_\pi)C_{\pi r}(n-1)\phi_{\pi g} \\ +d_{12}((B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C\pi r(n-1)\beta_0)C_{\pi g}(n-1)\phi_{\pi g} \\ +(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg} + C\pi r\beta_g)(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi)) \\ +d_{13}((B_\pi(n-1) + 2C_{\pi\pi}(n-1)\mu_\pi + C_{\pi g}(n-1)\mu_g + C\pi r(n-1)\beta_0)C_{\pi r}(n-1)\phi_{\pi g} \\ +(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg} + C\pi r(n-1)\beta_\pi)(B_r(n-1) + C_{\pi r}(n-1)\mu_\pi)) \end{array} \right]$$

$$B_r(n) = \left[\begin{array}{c} B_r(n-1)\beta_r + C_{\pi r}(n-1)\mu_\pi\beta_r + \frac{1}{2}d_{11}2C_{\pi r}(n-1)\beta_r(B_\pi(n-1) + 2C_{\pi\pi}\mu_\pi + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0) \\ +d_{12}C_{\pi r}(n-1)\beta_r(B_g(n-1) + C_{\pi g}(n-1)\mu_\pi) + d_{13}C_{\pi r}(n-1)\beta_r(B_r(n-1) + C_{\pi r}(n-1)\mu_\pi) - 1 \end{array} \right]$$

$$C_{\pi\pi}(n) = \left[\begin{array}{c} B_r(n-1)\beta_{\pi\pi} + (C_{\pi\pi}(n-1)\phi_{\pi\pi})^2 + C_{\pi g}(n-1)\phi_{\pi\pi}\phi_{g\pi} + C_{\pi r}(n-1)(\phi_{\pi\pi}\beta_{\pi} + \mu_{\pi}\beta_{\pi\pi}) \\ +\frac{1}{2}d_{11}((2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C_{\pi r}(n-1)\beta_{\pi})^2 + 2(B_{\pi}(n-1) + 2C_{\pi\pi}(n-1)\mu_{\pi} + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0)\beta_{\pi\pi}) \\ +\frac{1}{2}d_{22}(C_{\pi g}(n-1)\phi_{\pi\pi})^2 + \frac{1}{2}d_{33}(C_{\pi r}(n-1)\phi_{\pi\pi})^2 \\ +d_{12}(C_{\pi g}(n-1)\phi_{\pi\pi}(2C_{\pi\pi}\phi_{\pi\pi} + C_{\pi g}\phi_{g\pi} + C_{\pi r}(n-1)\beta_{\pi}) \\ +\beta_{\pi\pi}(B_g(n-1) + C_{\pi g}(n-1)\mu_{\pi})) \\ +d_{13}C_{\pi r}(n-1)\phi_{\pi\pi}(2C_{\pi\pi}\phi_{\pi\pi} + C_{\pi g}\phi_{g\pi} + C_{\pi r}(n-1)\beta_{\pi}) \\ +\beta_{\pi\pi}(B_r(n-1) + C_{\pi r}(n-1)\mu_{\pi})) \end{array} \right]$$

$$C_{\pi g}(n) = \left[\begin{array}{c} B_r(n-1)\beta_{\pi g} + 2C_{\pi\pi}(n-1)\phi_{\pi g}\phi_{\pi\pi} + C_{\pi g}(n-1)(\phi_{\pi\pi}\phi_{gg} + \phi_{\pi g}\phi_{g\pi}) + C_{\pi r}(n-1)(\phi_{\pi\pi}\beta_g + \phi_{\pi g}\beta_{\pi} + \mu_{\pi}\beta_{\pi g}) \\ +\frac{1}{2}d_{11}(2(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C_{\pi r}(n-1)\beta_{\pi}) \\ (2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg} + C_{\pi r}(n-1)\beta_g) \\ +2(B_{\pi}(n-1) + 2C_{\pi\pi}(n-1)\mu_{\pi} + C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0)C_{\pi r}(n-1)\beta_{\pi g}) \\ +\frac{1}{2}d_{22}2(C_{\pi g}^2(n-1)\phi_{\pi\pi}\phi_{\pi g}) + \frac{1}{2}d_{33}2(C_{\pi r}^2(n-1)\phi_{\pi\pi}\phi_{\pi g}) \\ +d_{12}(C_{\pi g}(n-1)\phi_{\pi g}(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C_{\pi r}(n-1)\beta_{\pi}) + C_{\pi g}(n-1)\phi_{\pi\pi}(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg} + C_{\pi r}(n-1)\beta_g) \\ +\beta_{\pi g}(B_g(n-1) + C_{\pi g}(n-1)\mu_{\pi})) \\ +d_{13}(C_{\pi r}(n-1)\phi_{\pi g}(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C_{\pi r}(n-1)\beta_{\pi}) + C_{\pi r}(n-1)\phi_{\pi\pi}(2C_{\pi\pi}(n-1)\phi_{\pi g} + C_{\pi g}(n-1)\phi_{gg} + C_{\pi r}(n-1)\beta_g) \\ +\beta_{\pi g}(B_r(n-1) + C_{\pi r}(n-1)\mu_{\pi})) \end{array} \right]$$

$$C_{\pi r}(n) = \left[\begin{array}{c} B_r(n-1)\beta_{\pi r} + C_{\pi r}(n-1)(\phi_{\pi\pi}\beta_r + \mu_{\pi}\beta_{\pi r}) \\ +\frac{1}{2}d_{11}(2(B_{\pi}(n-1) + 2C_{\pi\pi}(n-1)\mu_{\pi} \\ +C_{\pi g}(n-1)\mu_g + C_{\pi r}(n-1)\beta_0)C_{\pi r}(n-1)\beta_{\pi r} + (2(2C_{\pi\pi}(n-1)\phi_{\pi\pi} + C_{\pi g}(n-1)\phi_{g\pi} + C_{\pi r}(n-1)\beta_{\pi})C_{\pi r}(n-1)\beta_r)) \\ +d_{12}(C_{\pi g}(n-1)\phi_{\pi\pi}C_{\pi r}(n-1)\beta_r + +\beta_{\pi r}(B_g(n-1) + C_{\pi g}(n-1)\mu_{\pi})) \\ +d_{13}(C_{\pi r}(n-1)\phi_{\pi\pi}C_{\pi r}(n-1)\beta_r + +\beta_{\pi r}(B_r(n-1) + C_{\pi g}(n-1)\mu_{\pi})) \end{array} \right]$$

Note that in the unrestricted model representation higher order components are still present. However, as discussed in the main body their impact is statistically and economically negligible.

.3 Unscented Kalman Filter

This part of the Appendix describes in details the unscented Kalman filter procedure.

As mentioned in the chapter, to estimate the model parameters on the observed yields, inflation and unemployment, we cast the model into a state-space form and infer the three unobserved state variables. The transition equation is given by:

$$x_{t+1} = F(x_t, u) + \varepsilon_{t+1}, \tag{28}$$

where the state variable $x_t = (\pi_t, g_t, r_t)'$ represents the unobserved inflation, unemployment and short rate dynamics. The explicit form of the F function is given in the text equations 3.20

The measurement equation dynamics is of the form:

$$m_t = H(x_t, u) + \vartheta_{t+1}, \tag{29}$$

where $m_t = (\tilde{\pi}_t, \tilde{g}_t, \tilde{y}(t,3), \tilde{y}(t,6), \tilde{y}(t,12), \tilde{y}(t,24), \tilde{y}(t,60))'$ is the collection of observed inflation, observed unemployment, and five observed yields with time to maturity 3, 6, 12, 24, and 60 months, respectively. The covariance matrix of the transition equation is diagonal:

$\varepsilon_{t+1} = (\varepsilon^{\pi}_{t+1}, \varepsilon^{g}_{t+1}, \varepsilon^{r}_{t+1})' \sim N(0, \Sigma\Sigma')$,

$$\Sigma = \begin{pmatrix} \sigma_\pi & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & \sigma_r \end{pmatrix}$$

Similar to the transition equation, the measurement shocks are normally distributed and are uncorrelated with each other. The variance–covariance matrix of the measurement equation is of the form:

$$Cov(\vartheta) = \begin{pmatrix} \widetilde{\sigma}^2_\pi & 0 & \mathbf{0}_{5\times 5} \\ 0 & \widetilde{\sigma}^2_g & \mathbf{0}_{5\times 5} \\ 0 & 0 & \widetilde{\sigma}^2_y \mathbf{I}_{5\times 5}. \end{pmatrix}$$

Furthermore, we assume that the shocks of the measurement equation are uncorrelated with the transition equation's innovations.

The presence of nonlinearities in the measurement and transition equations motivate our choice for using UKF.

The basis of the UKF is the use of unscented transformation. In essence, the unscented transformation (UT) is a method for calculating the statistics of a random variable, which undergoes a nonlinear transformation (see Julier and Uhlmann (1997)). Below, we explain unscented transformation the procedure.

Unscented Transformation

Let x has a mean $\overline{x}$ and covariance $\overline{P}$. We build a matrix of $2L+1$ sigma vectors χ as follows:

$$
\begin{aligned}
\chi_0 &= \overline{x} \\
\chi_i &= \overline{x} + \left(\sqrt{(L+\kappa)P_x}\right)_i, \qquad i = 1, \ldots, L \\
\chi_i &= \overline{x} - \left(\sqrt{(L+\kappa)P_x}\right)_{i-L}, \quad i = L+1, \ldots, 2L,
\end{aligned}
$$

where L is the dimension of the state, κ is a composite scaling parameter governing the dispersion of the sigma points around the mean. $\left(\sqrt{(L+\kappa)P_x}\right)_i$ is the i-th row or column of the matrix square root of $(L+\kappa)P_x$. Sigma points are propagated through the measurement function $H(\cdot)$ to generate $\mathcal{M}$. The mean and the covariance of m_t are approximated by:

$$
\begin{aligned}
\overline{m} &\approx \sum_{i=0}^{2L} W_i^{\mu} \mathcal{M}_i \\
P_m &\approx \sum_{i=0}^{2L} W_i^{c} (\mathcal{M}_i - \overline{m})(\mathcal{M}_i - \overline{m})', \qquad i = 1, \ldots, L
\end{aligned}
$$

where W_μ and W^c are the weights for the mean and the covariance matrix.

$$
\begin{aligned}
W_0^{\mu} &= \frac{\kappa}{L+\kappa} \\
W_0^{c} &= \frac{\kappa}{L+\kappa} + 1 - \alpha^2 + \beta, \\
W_i^{\mu} = W_i^{c} &= \frac{\kappa}{2(L+\kappa)} \qquad i = 1, \ldots, 2L
\end{aligned}
$$

α and β are tune in parameters, determining the higher order moments of the distribution.

The UKF Algorithm

The specification of the UKF algorithm formalized below is especially designed for zero-mean additive noise. For general treatment, we refer to Wan and van der Merwe (2001).

(i) Initialize with

$$\hat{x}_0 = \mathbb{E}[x_0]$$
$$P_0 = \mathbb{E}[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)']$$

for $k \in 1, \ldots, \infty$:

(ii) Calculate sigma points:

$$\chi_{k-1} = [\hat{x}_{k-1} \quad \hat{x}_{k-1} + \sqrt{(L+\kappa)P_{k-1}} \quad \hat{x}_{k-1} - \sqrt{(L+\kappa)P_{k-1}}] \tag{30}$$

(iii) Time update

$$\chi^*_{k|k-1} = F(\chi_{k-1})$$
$$\hat{x}^-_k = \sum_{i=0}^{2L} \chi^*_{k|k-1}$$
$$P^-_k = \sum_{i=0}^{2L} (\chi^*_{k|k-1} - \hat{x}^-_k)(\chi^*_{k|k-1} - \hat{x}^-_k)' + Cov(\varepsilon)$$

(a) Redraw sigma points

$$\chi_{k-1} = [\hat{x}_{k-1} \quad \hat{x}_{k-1} + \sqrt{(L+\kappa)P_{k-1}} \quad \hat{x}_{k-1} - \sqrt{(L+\kappa)P_{k-1}}]$$

$$\mathcal{M}_{k|k-1} = H(\chi_{k-1})$$
$$\hat{m}^-_k = \sum_{i=0}^{2L} \mathcal{M}_{k|k-1}$$

(iv) Update measurement equation:

$$P^-_{m_k m_k} = \sum_{i=0}^{2L} (\mathcal{M}_{ik|k-1} - \hat{m}^-_k)(\mathcal{M}_{ik|k-1} - \hat{m}^-_k)' + Cov(\vartheta)$$
$$P_{x_k m_k} = \sum_{i=0}^{2L} ({}_{ik|k-1} - \hat{m}^-_k)(\mathcal{M}_{ik|k-1} - \hat{m}^-_k)'$$
$$\mathcal{K}_k = P_{x_k m_k} P^{-1}_{m_k m_k}$$
$$\hat{x}_k = \hat{x}^-_k + \mathcal{K}_k(m_k - \hat{m}^-_k)$$
$$P_k = P_{k-1} - \mathcal{K}_k P^-_{m_k m_k} \mathcal{K}'_k$$

Bibliography

Ahn, D., R. Dittmar, and R. Gallant (2002): "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, 15, 243–288.

Aït-Sahalia, Y. (1996a): "Non-Parametric Pricing of Interest Rate Derivative Securities," *Econometrica*, 64(3), 527–560.

——— (1996b): "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9(2), 385–426.

Ang, A., and G. Bekaert (2002): "Regime Switches in Interest Rates," *Journal of Business and Economic Statistics*, 20(2), 163–182.

Ang, A., G. Bekaert, and M. Wei (2008): "The Term Structure of Real Rates and Expected Inflation," *Journal of Finance*, 64(2), 797–849.

Ang, A., J. Boivin, S. Dong, and R. Loo-Kung (2010): "Monetary policy shift and the term structure," *Review of Economic Studies*, 2 (47), 429–457.

Ang, A., and M. Piazzesi (2003): "No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 50(4), 745–787.

Ang, A., M. Piazzesi, and M. Wei (2006): "What Does the Yield Curve Tell Us About GDP Growth," *Journal of Econometrics*, (131), 359–403.

Audrino, F. (2006): "Tree-Structured Multiple Regime in Interest Rates," *Journal of Business and Economic Statistics*, 24(3), 338–353.

Audrino, F., and P. Bühlmann (2001): "Tree-Structured GARCH Models," *Journal of the Royal Statistical Society (Series B)*, 63(4), 727–744.

Audrino, F., and E. De Giorgi (2007): "Beta Regimes for the Yield Curve," *Journal of Financial Econometrics*, 5(3), 456–490.

Audrino, F., and M. Medeiros (2011): "Modeling and forecasting short-term interest rates: The benefits of smooth regimes, macroeconomic variables and bagging," *Journal of Applied Econometrics*, 6, 999–1022.

Audrino, F., and F. Trojani (2006): "Estimating and Predicting Multivariate Volatility Thresholds in Global Stock Markets," *Journal of Applied Econometrics*, 21(3), 345–369.

Baillie, R., T. Bollerslev, and H. Mikkelsen (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics*, 74, 3–30.

Bansal, R., G. Tauchen, and H. Zhou (2004): "Regime Shifts in Term Structure, Expectation Hypothesis Puzzle, and the Real Business Cycle," *Journal of Business and Economic Statistics*, 22(4), 396–409.

Bansal, R., and H. Zhou (2002): "Term Structure of Interest Rates with Regime Shifts," *Journal of Finance*, 57(5), 1997–2043.

Beber, A., and M. C. Brandt (2010): "When it cannot get better or worse: The asymmetric impact of good and bad news on bond returns in expansions and recessions," *Review of Finance*, 14(3), 119–155.

Bibikov, R., and M. Chernov (2008): "Monetary policy regimes and the term structure of interest rates.," Discussion paper, Columbia University.

——— (2010): "No Arbitrage Macroeconomic Determinants of the Yield Curve," *Journal of Econometrics*, 159(1), 166–182.

Breiman, L. (1996): "Bagging Predictors," *Machine Learning*, 24, 124–140.

Bühlmann, P., and B. Yu (2002): "Analyzing Bagging," *Annals of Statistics*, 30, 927–961.

Campbell, J., A. Sunderam, and L. Viceira (2010): "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds," working paper.

Carr, P., and L. Wu (2007): "Stochastic skew in currency options.," *Journal of Financial Economics*, 86, 213–247.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, 115(1), 147–180.

Cochrane, J., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95, 138–60.

Cogley, T., and T. Sargent (2005): "Drifts and volatilities: monetary policies and outcomes in the post wwii us," *Review of Economic Studies*, 8(2), 262–302.

Cox, J., W. Ingersoll, and S. Ross (1985): "A Theory of the Term Structure of Interest Rate," *Econometrica*, 53(2), 385–407.

da Rosa, J., A. Veiga, and M. Medeiros (2008): "Tree-structured smooth transition regression models," *Computational Statistics and Data Analysis*, 52, 2469–2488.

Dai, Q., and K. Singleton (2000): "Discrete-time Term Structure Models with Generalized Market Prices of Risk," *Journal of Finance*, 55(5), 1943–1978.

——— (2002): "Expectations Puzzles, Time-varying Risk Premia, and Affine Model of the Term Structure," *Journal of Financial Economics*, 63(2), 415–441.

Dai, Q., K. Singleton, and W. Yang (2007): "Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields," *Review of Financial Studies*, 20(5), 1669–1706.

David, A., and P. Veronesi (2009): "What Ties return Volatilities to Price Valuations and Fundamentals," working paper.

De Pooter, M., F. Ravazzolo, and D. Van Dijk (2007): "Predicting the Term Structure of Interest Rates: Incorporating Parameter Uncertainty, Model Uncertainty and Macroeconomic Information," Discussion Paper 028/4, Tinbergen Institute.

Dewachter, H., and M. Lyrio (2006): "Macro Factors and Term Structure of Interest Rates," *Journal of Money, Credit and Banking*, 38(1), 119–140.

Dewachter, H., M. Lyrio, and K. Maes (2006): "A Joint Model for the Term Structure of Interest Rates and the Macroeconomy," *Journal of Applied Econometrics*, 21(4), 439–462.

Diebold, F., and C. Li (2006): "Forecasting the term structure of government bond yields," *Journal of Econometrics*, 130(2), 337–364.

Diebold, F., and R. Mariano (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistic*, 13, 253–263.

Diebold, F., and M. Nerlove (1989): "The dynamics of exchange rate volatility: a multivariate latent factor ARCH model," *Journal of Applied Econometrics*, 4(1), 1–21.

Diebold, F., G. Rudebusch, and S. Aruoba (2006): "The macroeconomy and the yield curve: a dynamic latent factor approach," *Journal of Econometrics*, 131(1-2), 309–338.

DUFFEE, G. (2002): “Term Premia and Interest Rate Forecasts in Affine Models,” *Journal of Finance*, 57(1), 405–443.

——— (2011): “Information in (and Not in) the Term Structure,” *Review of Financial Studies*, forthcoming.

EVANS, M., AND K. LEWIS (1995): “Do Expected Shifts in Inflation Affects Estimates of the Long-Run Fisher Relation,” *Journal of Finance*, 50(1), 225–253.

FAMA, E. (2006): “The Behavior of Interest Rates,” *Review of Financial Studies*, 19, 359–379.

FIORENTINI, G., E. SENTANA, AND N. SHEPHARD (2004): “Likelihood-based estimation of latent generalized ARCH structures,” *Econometrica*, pp. 1481–1517.

FORNI, M., M. HALLIN, M. LIPPI, AND L. REICHLIN (2004): “The generalized dynamic factor model consistency and rates,” *Journal of Econometrics*, 119(2), 231–255.

FORNI, M., AND L. REICHLIN (1998): “Lets get Real: A factor Analytic Approach to Disaggregated Business Cycle Dynamics,” *Review of Economic Studies*, 65, 453–473.

GAGLIARDINI, P., AND C. GOURIEROUX (2009): “Efficiency in Large Dynamic Panel Models with Common Factor,” Working Paper Series.

GARCIA, R., AND P. PERRON (1996): “An Analysis of the Real Interest Rate Under Regime Shifts,” *Review of Economics and Statistics*, 78(1), 111–125.

GOURIEROUX, C., AND R. SUFANA (2003): “Wishart Quadratic Term Structure Models,” Discussion paper, University of Toronto.

GRAY, S. (1996): “Modeling the Conditional Distribution of Interest Rates in a Regime-Switching Process,” *Journal of Financial Economics*, 42(1), 27–62.

GÜRKAYNAK, R., B. SACK, AND J. WRIGHT (2006): “The U.S. Treasury Yield Curve: 1961 to the Present,” Working paper.

HAMILTON, J. (1988): “Rational Expectations Econometric Analysis of Changes in the Regimes: An Investigation of the Term Structure of Interest Rates,” *Journal of Economic Dynamics and Control*, 12, 385–423.

HANSEN, P. R. (2005): “A test for superior predictive ability,” *Journal of Business and Economic Statistics*, 23, 365–380.

HARVEY, A., E. RUIZ, AND E. SENTANA (1992): “Unobserved component time series models with ARCH disturbances,” *Journal of Econometrics*, 52(1-2), 129–157.

HASSELTOFT, H. (2008): "Stock, Bonds, and Long-Run Consumption Risk," working paper.

HASTIE, T., R. TIBSHIRANI, AND J. FRIEDMAN (2001): *The elements of statistical learning: data mining, inference and prediction*, Springer Series in Statistics. Springer, Canada.

HOERDAHL, P., O. TRISTANI, AND D. VESTIN (2006): "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics," *Journal of Econometrics*, 131, 405–444.

INOUE, A., AND L. KILIAN (2004): "Bagging Time Series Models," CEPR Discussion Papers 4333, C.E.P.R. Discussion Papers.

JARDET, C., A. MONFORT, AND F. PEGORARO (2011): "No-Arbitrage Near-Cointegrated VAR(p) Term Structure Models, Term Premia and GDP Growth," Discussion paper, Banque de France, CNAM, CREST.

JOSLIN, S., M. PRIEBSCH, AND K. SINGLETON (2009): "Risk-Premium Accounting in Macro-Dynamic Term Structure Models," Working paper.

JULIER, S., AND J. UHLMANN (1997): "A New Extension of the Kalman Filter to Non- linear Systems," *Proceedings SPIE Signal Processing, Sensor Fusion, and Target Recognition*, VI(3068), 182–193.

KALMAN, R. (1960): "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, 82, 35–45.

KIM, D. (2008): "Challenges in Macro-Finance Modeling," *Federal Reserve Bank of St. Louis Review*, 91, 519–44.

KIM, D., AND A. ORPHANIDES (2005): "Term Structure Estimation with Survey Data on Interest Rate Forecasts," CEPR Discussion Papers 5341, C.E.P.R. Discussion Papers.

KOZICKI, S., AND P. TINSLEY (2005): "Permanent and Transitory Policy Shocks in an Empirical Macro Model with Asymmetric Information," *Journal of Economic Dynamics and Control*, 29, 1985–2015.

KÜNSCH, H. (1989): "The jackknife and the bootstrap for general stationary observations," *Annals of Statistics*, 17, 1217–1241.

LEIPPOLD, M., AND L. WU (2003): "Design and Estimation of Quadratic Term Structure Models," *European Finance Review*, 7, 47–73.

LITTERMAN, R., AND J. SCHEINKMAN (1991): "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, 1, 51–61.

LUDVIGSON, S., AND S. NG (2009a): "A Factor Analysis of Bond Risk Premia," *Handbook of Applied Econometrics*, p. forthcoming.

——— (2009b): "Marco Factors in Bond Risk Premia," *Review of Financial Studies*, 22(12), 5027–5067.

MATHAI, A., S. PROVOST, AND T. HAYAKAWA (1995): *Bilinear Forms and Zonal Polynomials, Lecture Notes in Statistics.* Springer-Verlag, New York.

MOENCH, E. (2008): "Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach," *Journal of Econometrics*, 146(1), 26–43.

RUDEBUSCH, G., AND T. WU (2007): "Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models," *Journal of Money, Credit and Banking*, 39, 395–422.

——— (2008): "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy," *Economic Journal*, 118(530), 906–926.

SENTANA, E. (1991): "Time-varying volatility and returns on ordinary shares: an empirical investigation," Unpublished Doctoral Dissertation, University of London.

——— (2004): "Factor representing portfolios in large asset markets," *Journal of Econometrics*, 119(2), 257–289.

SENTANA, E., G. CALZOLARI, AND G. FIORENTINI (2008): "Indirect estimation of large conditionally heteroskedastic factor models, with an application to the Dow 30 stocks," *Journal of Econometrics*, 146(1), 10–25.

SIMS, C., AND T. ZHA (2006): "Were there regime switches in U.S. monetary policy," *American Economic Review*, 96, 54–81.

SOLA, M., AND J. DRIFFILL (1994): "Testing the Term Structure of Interest Rates Using a Vector Autoregression with Regime Switching," *Journal of Economic Dynamics and Control*, 18, 601–628.

STAMBAUGH, R. (1988): "The information in forward rates," *Journal of Financial Economics*, 21, 41–70.

STANTON, R. (1997): "A Non-Parametric Model for the Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance*, 52(5), 1973–2002.

STOCK, J., AND M. WATSON (2005): "Implications of Dynamic Factor Models for VAR Analysis," Working Paper, Princeton University.

Taylor, J. (1993): "Discretion versus policy rules in practice. Carnegie-Rochester Conference Series on Public Policy," *Journal of Econometrics*, 39, 195–214.

Tong, H. (1978): *On a Threshold Model*, Pattern Recognition and Signal Processing. Sijthoff and Noordhoff, Amsterdam.

Tong, H., and K. Lim (1980): "Threshold aAutoregressive, Limit Cycles and Cyclical Data," *Journal of the Royal Statistical Society, Series B*, 42, 245–292.

van Dijk, D., T. Tersvirta, and P. Franses (2002): "Smooth transition autoregressive models - - a survey of recent developments," *Econometrics Review*, 21, 1—47.

Wan, E., and R. van der Merwe (2001): "Kalman Filtering and Neural Network," *John Wiley & Sons, Inc.*

Wright, J. (2009): "Term Premiums and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset," Working Paper Series.

Curriculum Vitae

Name	Kameliya Filipova
Date of Birth	January 25th 1981
Nationality	Bulgarian
Contact	Spisergasse 43 9000 St. Gallen Switzerland phone: +41 (0)78 901 11 20 email: kameliya.filipova@gmail.com

Education

University of St. Gallen **April 2007 - March 2012**
Ph.D. in Economics and Finance
Topic: Multiple Threshold Regimes and Macroeconomic Predictors for Analyzing and Pricing Interest Rate-Dependent Instruments
Dissertation Committee: F. Audrino (University of St. Gallen) (advisor)
G.Tauchen (Duke University), M. Lechner (University of St. Gallen)
P. Söderlind (University of St. Gallen)

Duke University, NC, USA **January 2011 - July 2011**
Visiting Scholar at the Economics Department

University of Sofia, Bulgaria **October 1999 - June 2006**
Department of Mathematics and Computer Science
M.S. in Applied Mathematics
Master Thesis: Dependence Modeling in Risk Management

Academic Experience

University of St. Gallen **February 2007 - March 2012**

Research Assistant

Institute of Mathematics and Statistics

University of St. Gallen

Teaching Assistant

Tutoring Bachelor level course *Econometrics* (2008, 2009, 2010)

Tutoring Bachelor level course *Statistics* (2008, 2009)

Tutoring Bachelor level course *Mathematics* (2010)

Professional Experience

UBS, Zürich **since April 2012**

Quantitative Analyst

Bulgarian Telecommunication Company **July 2005 - January 2007**

Credit Risk and Revenue Assurance Analyst

Languages

Bulgarian (mother tongue), German (fluent), English (fluent), Russian (good knowledge)

Grants

Swiss National Science Foundation visiting scholar fellowship, Duke University, North Carolina, USA, January 2011 – July 2011

Sofia University, scholarship for perspective student, 2004 - 2006