Universität St.Gallen

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

Francesco Audrino and Kameliya Filipova

May 2009 Discussion Paper no. 2009-10

Department of Economics University of St. Gallen

Editor:	Martina Flockerzi University of St. Gallen Department of Economics Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25 Fax +41 71 224 31 35 Email vwaabtass@unisg.ch
Publisher:	Department of Economics University of St. Gallen Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25 Fax +41 71 224 31 35
Electronic Publication:	http://www.vwa.unisg.ch

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

Francesco Audrino and Kameliya Filipova

Author's address: Prof. Dr. Francesco Audrino
Institute of Mathematics and Statistics
Bodanstrasse 6
CH-9000 St. Gallen
Phone +41 71 224 2431
Fax +41 71 224 2894
Email francesco.audrino@unisg.ch
Website www. istituti.usilu.net/audrinof/web/

Author's address: Kameliya Filipova
Institute of Mathematics and Statistics
Bodanstrasse 6
CH-9000 St. Gallen
Phone +41 71 224 2477
Fax +41 71 224 2894
Email kameliya.filipova@unisg.ch

Abstract

We propose an empirical approach to determine the various economic sources driving the US yield curve. 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 fitting in-sample and predicting out-of-sample the yield curve in comparison to several alternative approaches.

Keywords

Yield curve modeling and forecasting; Macroeconomic variables; Tree-structured models; Threshold regimes; GARCH; Bagging.

JEL Classification

C22; C51; C53; E43; E44

1 Introduction

Over the past three decades financial economists, macroeconomists, and market practitioners have all attempted to build good models for the yield curve. Depending on the different researchers' modeling strategies and goals in constructing the yield curve model (for example bond pricing, policy analysis, interest rate forecasting), the resulting models vary enormously in their form and fit. While macroeconomists focus mainly on building equilibrium models describing the relationship between the yields and various macroeconomic indices/variables (mainly measures of inflation and real economic activity), the traditional finance term structure literature decomposes the yield curve into a small set of latent variables and ignores the macroeconomic nature.

The connection between the macroeconomic and financial views of the term structure has been a very fertile area for recent research. The macroeconomic linkage and the improved forecasting performance of macro variables on top of latent factors have given rise to a new modeling framework, the so-called macro-finance models. Early works in this field include for example Rudebusch (1995) and Balduzzi, Bertola, and Foresi (1997) who introduce latent term structure models including the central bank's target rate as a factor. Studies such as Estrella and Mishkin (1997) and Evans and Marshall (1998) use VARs with yields of various maturities together with macro variables. Ang and Piazzesi (2003) propose models that combine two macroeconomic variables (real activity and inflation) as state variables together with three unobserved factors. They find that the macro factors explain up to 85 percent of the short and middle parts of the yield curve and a significantly smaller portion (around 40 percent) of the long-end of the yield curve. Using output and unemployment as macro factors, Ludvigson and Ng (2007) were able to explain more than 25 percent of the yield curve variation. Other important contributions in that area include for example Dewachter, Lyrio, and Maes (2006), Dewachter and Lyrio (2006), Hoerdahl, Tristani, and Vestin (2006), and Rudebusch and Wu (2008).

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.[1] Factor analysis of the unconditional variance-covariance matrix of yields commonly suggests the number of latent factors needed to explain the cross-sectional dynamics. In addition, standard macroeconomic intuition is typically used to determine the macro factors entering the yield curve 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 over time.

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 on real rates, 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 (1996), 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

[1] 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.

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 paper 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 paper 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 our paper we do not restrict the local dynamics to follow Cox, Ingersoll, and Ross (1985) process, but allow for a more flexible data-driven structure selected by a given decision rule.

Our contribution to the term structure literature is twofold. First, our approach enables an interpretable and statistically accurate identification of the most important predictors and the regime structure driving the yield curve dynamics over time for each maturity. Second, it 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,[2] 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

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

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.

The remainder of this paper is organized as follows: Section 2.1 and Section 2.2 present the modeling framework we use for fitting and forecasting the term structure. Section 2.3 describes the techniques we employ for model estimation. The role of bagging is discussed in Section 2.4. In Section 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 4 concludes.

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.[3] 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

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

endogenous regime specification.

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

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

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. More formally, ε_{t,n_τ} 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.

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[4] 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:

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

where $\psi_{\mathbf{n}_\tau} = (\alpha^j_{0,n_\tau}, \alpha^j_{1,n_\tau}, (\beta^{\mathbf{j}}_{\mathbf{n}_\tau})', (\gamma^{\mathbf{j}}_{\mathbf{n}_\tau})', \omega^j_{n_\tau}, a^j_{n_\tau}, b^j_{n_\tau}, j = 1,\dots,K_{n_\tau})$ is a $((m+4)\times K_{n_\tau})$-dimensional vector of the unknown (true) parameters $\tau = 1,\dots,T$. $I(\cdot)$ is the indicator function and $\mathcal{R}^j_{n_\tau}$ represents a region of the partition $\mathcal{P}_{n_\tau} = \{\mathcal{R}^1_{n_\tau},\dots,\mathcal{R}^{K_{n_\tau}}_{n_\tau}\}$ of the state space G_{n_τ} of $\Phi_{t,n_\tau} = \{(\Delta y(t,n_\tau), \mathbf{x'_t}, \mathbf{x^{ex\prime}_t})' \in \mathbb{R}^1\times\mathbb{R}^{m_1}\times\mathbb{R}^{m_2}\}$ such that

$$\mathcal{P}_{n_\tau} = \{\mathcal{R}^1_{n_\tau},\dots,\mathcal{R}^{K_{n_\tau}}_{n_\tau}\},\ G_{n_\tau} = \cup_{j=1}^{K_{n_\tau}}\mathcal{R}^j_{n_\tau},\ \mathcal{R}^i_{n_\tau} \underset{(i\neq j)}{\cap} \mathcal{R}^j_{n_\tau} = \varnothing \quad \tau = 1,\dots,T.$$

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

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.[5] Consequently, this approach leads to severe

[4]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.

[5]See for example Kim and Orphanides (2005) for discussion of this.

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.

2.3.1 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 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 paper 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.[6]

There are at least four reasons why we favor employing a dimensionality reduction

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

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.

2.3.2 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 paper we will mention only the main steps of the binary tree construction and estimation. However, an exact description, illustrative examples, and further applications of the algorithm can be found for example in Audrino and Bühlmann (2001), Audrino (2006), and Audrino and Trojani (2006).

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 optimizing the conditional negative 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.

2.4 Improving the forecasting ability: Bagging

One of the major problems with the two-step procedure presented in the previous section is the high variance of the forecasts. The reason for this instability lies in the hierarchical nature of the tree process: the effect of an error on the top of the split is propagated down to all the splits below it. One way to overcome this problem is to average forecasts from a large number of models selected by the given decision rule. This is actually the main idea of bagging (short for bootstrap aggregating), proposed by Breiman (1996). Bagging is a variance reduction technique aimed at improving the predictive performance of unstable estimators such as 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 (2008)).

The dramatic reduction of the prediction error for a wide range of models with a similar (unstable) 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.

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

Construct B bootstrap samples denoted by

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

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

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

3 Empirical Results

We start this section with a brief description of the data we use for the empirical part of the paper. 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.

3.1 Data

The *term structure data* consist of one-month 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 for a total of 534 monthly observations. 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 provides a fairly detailed description of the data.

[Table 1 about here.]

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 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.

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.

3.2 What is driving the Yield Curve Predictability?

3.2.1 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 2.

[Table 2 about here.]

Judging from the results presented in Table 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 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. According 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,

$$r_t = \beta(\mathbb{E}_t[\pi_{t+1}] - \pi_t^\star) + \gamma \mathbb{E}_t(z_{t+1}) + \rho r_{t-1}. \tag{2}$$

For the linear dynamics of our resulting model, the combination of the yield curve's level and the level and conditional volatility of inflation (vol.PPI) might be thought of as a proxy for the difference between the expected and the target inflation. However, the exact behavior of the two measures is rather difficult to disentangle. The reason is that both expected inflation as well as the Federal Reserve inflation target are in general unobservable. In addition, the linear combination of the expected inflation (intuitively measured by the inflation level, the conditional volatility of inflation, and the level of the yield curve), the square of the leading real activity index (HELP), and the slope of the yield curve may be considered as an empirical proxy for the 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 (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.

3.2.2 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 3.

[Table 3 about here.]

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,[7] 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.

[7] 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.

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.

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 4.

[Table 4 about here.]

[Figure 1 about here.]

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 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.[8] 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 confirms this finding. The

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

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 5.

[Table 5 about here.]

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 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 6.

[Table 6 about here.]

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.

3.2.3 Stylized Facts

An adequate term structure model should not only give insight into the economic forces driving the dynamics of the yields with different maturities, but it should also be in line with the most important stylized facts. In this section we test our model's ability to replicate the following stylized facts: *(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 2 and Figure 3 provide a graphical representation of the above-mentioned facts.

[Figure 2 about here.]

The upper panel of Figure 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 2, we can easily draw one more conclusion - the distribution of yields around their median is asymmetric with a longer right tail.

[Figure 3 about here.]

Next, Figure 3 presents four fitted yield curves for some selected dates. Apparently, 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 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 2 Panel A as well as the threshold structure given in Table 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.

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 level; *(iii)* two dynamic specifications of Nelson-Siegel 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. We perform out-of-sample forecasts over the period January 2002 - June 2005 for a total of 42 observations.

In this paper, 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) and the mean absolute error (MAE). The measures are given by:

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

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 unstable 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.[9] The results are presented in Table 7.

[Table 7 about here.]

To assess the statistical differences in the out-of-sample performances of the different models and their bagged versions, we perform a series of tests for superior predictive ability introduced by Hansen (2005). The results are summarized in Table 8.

[Table 8 about here.]

Comparing the one-month-ahead out-of-sample results of the different models (see Table 7), without considering bagging, we find that our model has overall good performance at all eight maturities both in terms of MSE and MAE. Matters improve dramatically, once we apply bagging. The SPA p-values presented in Table 8 reveal that the forecasts

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

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.[10] 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.

4 Conclusion

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

[10] 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.

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.

References

Aït-Sahalia, Y. (1996): "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., 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 (2008): “Smooth Regimes, Macroeconomic Variables, and Bagging for the Short-Term Interest Rate Process,” University of St. Gallen Department of Economics working paper series 2008-16.

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

Balduzzi, P., G. Bertola, and S. Foresi (1997): “A Model of Target Changes and the Term Structure of Interest Rates,” *Journal of Monetary Economics*, 39(2), 223–249.

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, 396–409.

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

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

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

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

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

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

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.

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., 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, 405–443.

Estrella, A., and F. Mishkin (1997): "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank," *European Economic Review*, 41, 13751401.

Evans, C., and D. Marshall (1998): "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory," *Carnegie-Rochester Conference Series on Public Policy*, 49(1), 53–111.

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.

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.

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

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.

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.

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.

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

Litterman, R., and J. Scheinkman (1991): "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, 1, 51–61.

Ludvigson, S., and S. Ng (2007): "Macro Factors in Bond Risk Premia," NBER Working Paper.

Rudebusch, G. (1995): "Federal Reserve interest rate targeting, rational expectations, and the term structure," *Journal of Monetary Economics*, 35(2), 245–274.

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.

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.

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.

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

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	16.06739	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: 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, the unemployment rate, and the US gross domestic product, 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.

Panel A: Best Subset Selection

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

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	$HELP \leq 61.82$ $HELP > 61.82$ and $slope \leq -0.0662$ $HELP > 61.82$ and $slope > -0.0662$	3
2Y	$HELP \leq 61.82$ $HELP > 61.82$ and $slope \leq -0.0662$ $HELP > 61.82$ and $slope > -0.0662$	3
3Y	$HELP \leq 61.82$ $HELP > 61.82$ and $slope \leq -0.0662$ $HELP > 61.82$ and $slope > -0.0662$	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 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.

Short-term Maturities' Parameter Estimates

Optimal Regime Structure	Variable	3 Months Coefficient	3 Months t-statistic	6 Months Coefficient	6 Months t-statistic
$CPI_{t-1} \leq 3.5316$	const	0.1729	29.5631	0.2101	31.1557
	Δy	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	0.0385	27.0689	0.0462	27.0864
	ϵ^2	0.1347	4.2985	0.1068	3.5909
	σ^2	0.0000	0.0046	0.0006	0.8040
$CPI_{t-1} > 3.5316$	const	0.2131	9.5540	0.2541	46.8190
	Δy	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	0.1800	22.7663	0.1598	10.3497
	ϵ^2	0.8275	51.1376	0.5685	44.9099
	σ^2	0.0077	0.2294	0.2093	11.5955
LB_5^2		7.2069	(0.2057)	6.1919	(0.2880)
LB_{10}^2		15.033	(0.1309)	14.0398	(0.1712)
LB_{15}^2		16.6322	(0.3413)	15.9012	(0.3886)

Table 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_i^2 denotes the Ljung-Box statistic for serial correlation of the squared residuals out to i lags. p-values are in parentheses.

Mid-term Maturities' Parameter Estimates

Regime Structure	Variable	1 Year Coefficient	1 Year t-statistic	2 Years Coefficient	2 Years t-statistic	3 Years Coefficient	3 Years t-statistic
	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
$HELP \leq 61.82$	ω	0.0071	1.4674	0.0061	2.2538	0.0032	2.2138
	ϵ^2	0.5131	1.1935	0.3814	1.5724	0.2993	1.7977
	σ^2	0.1387	0.8706	0.2040	1.0613	0.3263	2.1271
	const	-0.0568	-0.8287	-0.0468	-0.6951	-0.0159	-0.2608
$HELP > 61.82$	Δ y	-0.2591	-1.2252	-0.2654	-1.7572	-0.2788	-1.9412
$slope \leq -0.0662$	ω	0.0317	1.0229	0.0000	0.0000	0.0000	0.0000
	ϵ^2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	σ^2	1.0709	20.5100	1.1434	26.8699	1.1303	32.5502
	const	0.0206	1.2551	0.0131	0.6677	0.0131	0.6852
$HELP > 61.82$	Δ y	0.0635	1.3306	0.0613	1.2087	0.0405	0.8011
$slope > -0.0662$	ω	0.0040	0.9031	0.0164	1.9815	0.0183	2.2763
	ϵ^2	0.0338	2.6079	0.0721	2.9940	0.0607	2.6512
	σ^2	0.9161	9.7915	0.7935	11.7520	0.7852	10.9538
LB_5^2		5.3243	(0.3776)	4.5076	(0.4789)	4.1067	(0.5342)
LB_{10}^2		5.8277	(0.8295)	10.7420	(0.3780)	6.4213	(0.7787)
LB_{15}^2		15.8527	(0.3919)	14.1564	(0.5137)	9.3152	(0.8605)

Table 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_i^2 denotes the Ljung-Box statistic for serial correlation of the squared residuals out to i lags. p-values are in parentheses.

Long-term Maturities' Parameter Estimates

		5 Years		10 Years	
Optimal Regime Structure	Variable	Coefficient	t-statistic	Coefficient	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	0.0005	0.2769	0.0003	1.9806
	ϵ^2	0.6736	2.4245	0.0001	0.0011
	σ^2	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	0.0094	1.1868	0.0056	1.6716
	ϵ^2	0.1036	2.9457	0.0930	2.3062
	σ^2	0.8334	11.0224	0.8543	13.8319
LB_5^2		4.6216	(0.4638)	4.8898	(0.4295)
LB_{10}^2		9.7054	(0.4667)	10.4846	(0.3991)
LB_{15}^2		10.8031	(0.7664)	13.5046	(0.5634)

		7 Years	
Optimal Regime Structure	Variable	Coefficient	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	0.0002	0.0602
	ϵ^2	0.1856	2.5730
	σ^2	0.8496	13.5133
LB_5^2		21.4143	(0.0007)
LB_{10}^2		39.9147	(0.0000)
LB_{15}^2		57.6575	(0.0000)

Table 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_i^2 denotes the Ljung-Box statistic for serial correlation of the squared residuals out to i lags. p-values are in parentheses.

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 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.

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)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.0068	0.0820	0.5781	0.6626	0.0147	0.1315	0.0128	0.1440
6M	0.0099	0.0368	0.4329	0.5083	0.0171	0.1135	0.0196	0.0798
1Y	0.0284	0.0653	0.2420	0.3014	0.0346	0.1486	0.0357	0.3754
2Y	0.0824	0.0905	0.0845	0.1253	0.0950	0.2045	0.0887	0.3112
3Y	0.1149	0.1449	0.1550	0.1439	0.1236	0.2295	0.1142	0.2941
5Y	0.1242	0.1434	0.1679	0.1323	0.1345	0.1905	0.1230	0.2607
7Y	0.1155	0.1155	0.4108	0.6082	0.1219	0.1510	0.1116	0.2707
10Y	0.0918	0.0995	0.1230	0.1731	0.1035	0.1204	0.0951	0.2093

	Out-of-sample MAE							
Maturity	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.0949	3.2111	0.4937	0.5457	0.0843	0.0948	0.1237	0.1487
6M	0.1218	3.3961	0.3804	0.4308	0.0976	0.1267	0.2691	0.1995
1Y	0.1479	0.1431	0.2450	0.2805	0.1474	0.1693	0.1669	0.6552
2Y	0.2411	0.2322	0.2894	0.2266	0.2410	0.2529	0.2389	0.5194
3Y	0.2655	0.2654	0.2794	0.2689	0.2782	0.2884	0.2650	0.5111
5Y	0.2772	0.4071	0.2625	0.3858	0.2943	0.2928	0.2790	0.2675
7Y	0.2679	0.2679	0.4582	0.5795	0.2764	0.2757	0.2855	0.3157
10Y	0.2395	0.2525	0.2460	0.2708	0.2561	0.2515	0.2436	0.2638

	Out-of-sample MAE for the bagged models							
Maturity	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.0644	0.2595	0.7361	0.7946	0.0843	0.3249	0.0821	0.3552
6M	0.0813	0.1643	0.6214	0.6808	0.0976	0.2945	0.1242	0.1654
1Y	0.1326	0.2142	0.4387	0.5000	0.1474	0.3378	0.1562	0.5774
2Y	0.2236	0.2301	0.2612	0.2826	0.2410	0.3787	0.2360	0.4926
3Y	0.2600	0.2577	0.3094	0.3047	0.2782	0.4047	0.2618	0.4654
5Y	0.2698	0.3018	0.3222	0.2815	0.2943	0.3624	0.2714	0.4450
7Y	0.2564	0.2564	0.4686	0.7115	0.2764	0.3251	0.2561	0.3257
10Y	0.2282	0.2379	0.2876	0.3584	0.2561	0.2840	0.2378	0.3513

Table 7: Results of out-of-sample 1-month-ahead forecasting using eight models and their bagged versions, as described in detail in the text. The results are based on the out-of-sample period January 2002 - June 2006, for a total of 42 observations.

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)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.5952	0.0000	0.0000	0.0000	0.2474	0.0000	0.1265	0.0000
6M	0.5260	0.0123	0.0000	0.0000	0.4006	0.0000	0.0231	0.0039
1Y	0.5371	0.0000	0.0000	0.0000	0.4637	0.0000	0.5118	0.0000
2Y	0.6423	0.2842	0.5915	0.0973	0.3876	0.0000	0.4496	0.0000
3Y	0.6677	0.4843	0.3679	0.3726	0.4399	0.0063	0.6521	0.0000
5Y	0.6578	0.0740	0.0278	0.6332	0.5135	0.0000	0.6947	0.0000
7Y	0.6845	0.6845	0.0000	0.0000	0.5557	0.0521	0.6721	0.1050
10Y	0.6290	0.2346	0.0473	0.0000	0.3647	0.0226	0.4608	0.0000

Maturity	Out-of-Sample SPA test for the MAE							
	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.0595	0.0000	0.0000	0.0000	0.4035	0.0507	0.0000	0.0000
6M	0.0000	0.0000	0.0000	0.0000	0.4551	0.0459	0.0000	0.0000
1Y	0.5061	0.5553	0.1090	0.0000	0.4091	0.1827	0.0516	0.0000
2Y	0.2909	0.4371	0.3017	0.6873	0.4292	0.5103	0.2179	0.0000
3Y	0.4048	0.3371	0.5405	0.5583	0.4352	0.3805	0.5171	0.0000
5Y	0.4965	0.0000	0.0000	0.1128	0.4150	0.4945	0.7182	0.5721
7Y	0.6146	0.6146	0.0000	0.0000	0.5185	0.5391	0.0605	0.0760
10Y	0.3893	0.2474	0.4487	0.0101	0.3538	0.1757	0.0944	0.2051

Maturity	Out-of-Sample SPA test for the MAE of the bagged models							
	Macro Tree	Best Subset	NS AR(1)	NS VAR(1)	RW	VAR(1)	Audrino Tree	Gray's RS
3M	0.5433	0.0000	0.0000	0.0000	0.4035	0.0000	0.2360	0.0000
6M	0.3680	0.0000	0.0000	0.0000	0.4551	0.0000	0.0000	0.0000
1Y	0.5688	0.0000	0.0000	0.0000	0.4091	0.0000	0.1832	0.0000
2Y	0.6851	0.3987	0.6193	0.1475	0.4292	0.0000	0.3010	0.0000
3Y	0.6458	0.5351	0.3744	0.3710	0.4352	0.0000	0.5464	0.0000
5Y	0.6580	0.0000	0.0461	0.5154	0.4150	0.0067	0.5669	0.0000
7Y	0.6967	0.6967	0.0000	0.0000	0.5185	0.0000	0.5161	0.1026
10Y	0.7041	0.1806	0.0099	0.0000	0.3538	0.0102	0.1039	0.0000

Table 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 2006, for a total of 42 observations.

Mid-term Maturity Regimes and NBER Business Cycles

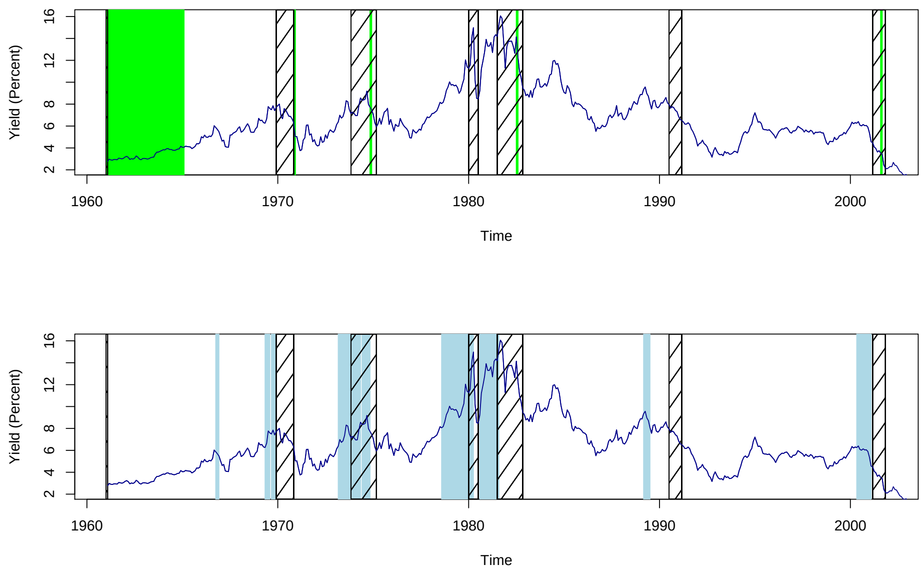
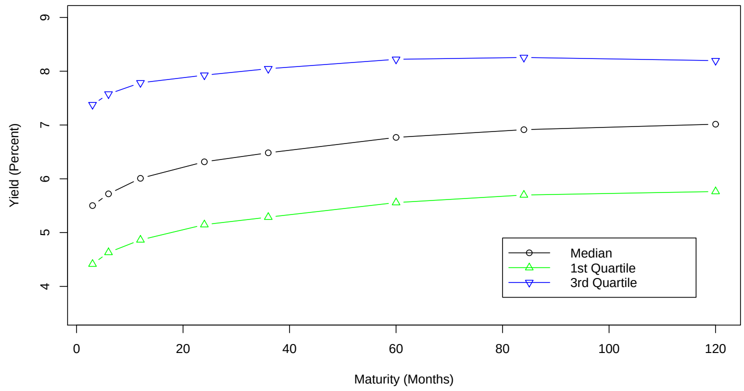


Figure 1: The top and the bottom panels plot the one-year yield time series for the period January 1961-December 2001. The gray bars in the top panel overlay periods with low real activity HELP $\leq$ 61.82 as found in Regime 1. The gray bars in the bottom panel overlay periods with medium and high real 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.

Panel A: Median Yield Curve



Panel B: Boxplots Yield Curve

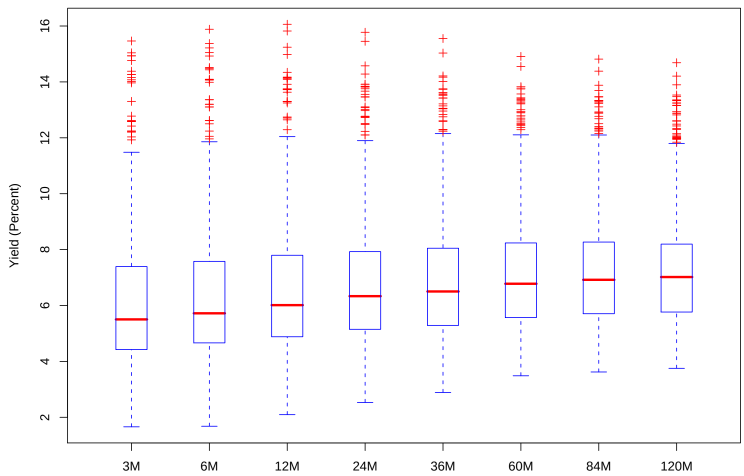


Figure 2: Panel A shows the median fitted (data-based) yield curve with interquartile range (25th and 75th percentiles). Panel B presents Boxplots for the fitted (model-based) yields for every maturity. The data span the time period January 1961-December 2001, for a total of 492 observations.

Selected Fitted Yield Curves

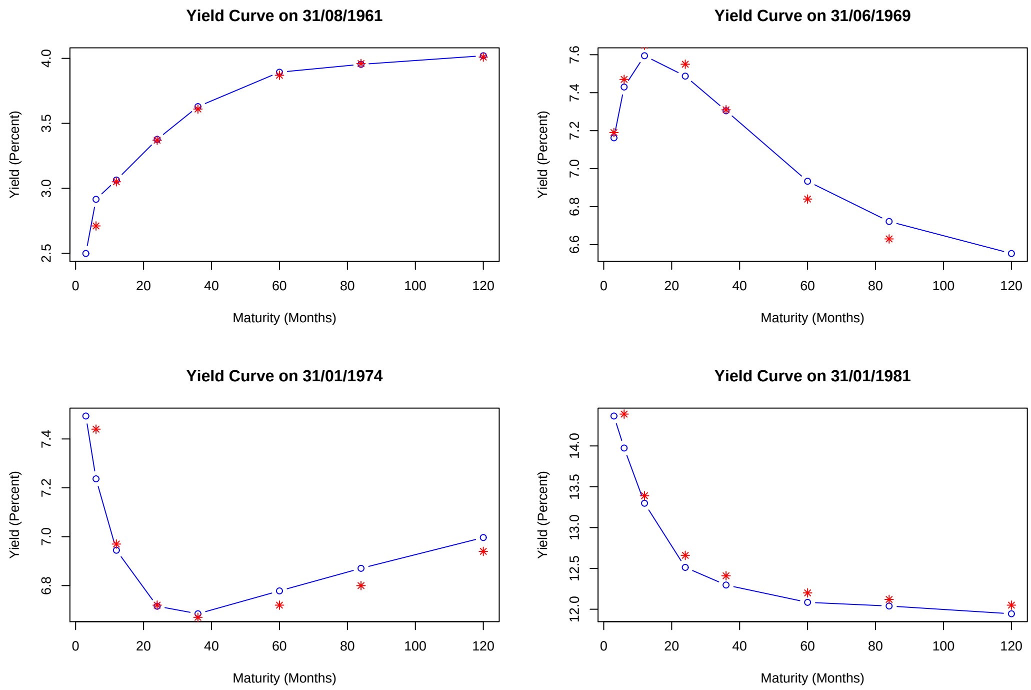


Figure 3: Fitted (model-based) yield curves for selected dates (dotted lines), together with actual yields (stars). See text for details.