COMPSTAT 2010, Paris

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

Francesco Audrino and Kameliya Filipova

University of St. Gallen

August 26, 2010

Overview

- Introduction
  [Motivation, Related literature, Open issues]
- Modeling framework
  [Conditional dynamics]
- Estimation procedure
  [Best subset selection, Threshold estimation]
- Empirical results
  [Optimal structure, Stylized facts, Forecasting]
- Conclusion

Introduction

Motivation

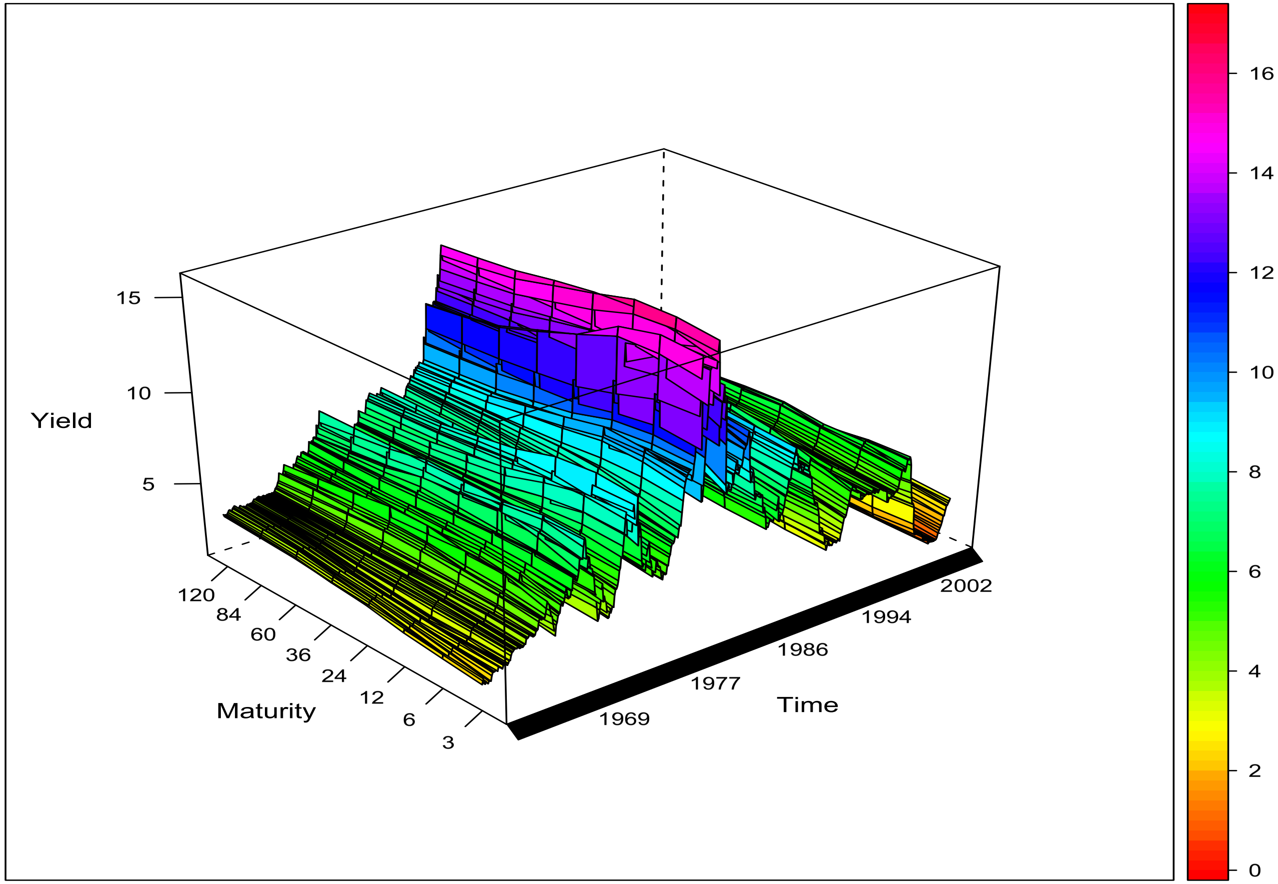
Modelling the time-varying dynamics of interest rates is crucial for many diverse task, such as

- ☐ pricing assets and their financial derivatives;
- ☐ managing financial risk;
- ☐ conducting monetary policy;
- ☐ forecasting.

To describe the term structure behavior, a wide variety of methods has been proposed.

Why yet another term structure model?

A first look at the data



U.S. monthly Treasury Bill data taken from CRSP database for the period January 1961 - June 2005.

Two views of the term structure

Finance view:

- latent factor models - aim at perfectly fitting the term structure at any point in time in order to ensure that no arbitrage opportunities exist
- but ... factors are latent. They don't model how yields respond to macro variables.

Affine Term Structure Models (ATSMs) [Duffie and Kan (1996); Dai and Singleton (2000)] and their extensions - "essentially" ATSMs [Duffee (2002)], "extended" ATSMs, Quadratic Term Structure Models (QTSMs), Dynamic Term Structure Models (DTSMs), . . .

Macroeconomic view:

- short rate is set by the central bank, which adjusts the rate to achieve its economic stabilization goals
- but ... they fit the historical interest data poorly.

Dynamic Stochastic General Equilibrium (DSGE) models, Taylor (1993) rule and its extensions Clarida, Gali and Gertler (2000).

Macro-finance models

Main Idea: Since macroeconomic variables are correlated with yields, incorporating these economic factors into a pure finance model usually improves the predictive performance and provides macroeconomic linkage.

Common setup: reduced-form no-arbitrage models with continuous-time or discrete time (mostly Gaussian) diffusions, following the tradition of DTSMs (ATSMs, QTSMs)

Related Literature: Ang and Piazzesi (2003); Dewachter, Lyrio, and Maes (2006); Dewachter and Lyrio (2006); Hoerdahl, Tristani, and Vestin (2006); Rudebusch and Wu (2007); Ang, Boivin, and Dong (2007); Ang, Bekaert, and Wei (2007); Kim and Wright (2005); Buraschi and Jiltsov (2007); DAmico, Kim, and Wei (2008) among many others.

However, the question how yields are associated with macro variables remains open.

Open questions

While just a small number of factors are sufficient to model the cross sectional variation of yields, a couple of questions still remain open.

- What is the number of factors needed to build a good model for the time series dynamics?
- How yields are associated with macro variables?
- Is there any predictability of the macro variables on top of the latent factors? If yes, then how many and which macroeconomic factors should be included in the model?
- Do these variables always have the same impact on the yields with different maturities?

Contributions

We propose regime-switching multifactor model model for the term structure dynamics over time which

- for every maturity we are able to identify or infer, in a purely data-driven way, the most important macroeconomic and latent variables driving both the local dynamics and the regime shifts;
- is able to replicate the most important stylized facts;
- while it remains highly competitive in terms of in- and out-of-sample forecasting performance.

As such, the modeling framework offers a clear interpretation and regime specification.

Regime-switching models

Regime-switching models describe better the nonlinearities in the yields' drift and the volatility found in the historical interest rate data.

Related literature: Ang and Bekaert (2002), Bansal and Zhou (2002), Dai, Singleton, and Yang (2007), Bansal, Tauchen, and Zhou (2004), Audrino and De Giorgi (2007) , Audrino (2006), Rudebusch and Wu (2007); Ang, Bekaert, and Wei (2007); ...

However instead of using the common Markovian regime–switching framework, the regimes could be constructed as multiple tree-structured thresholds partitioning the predictor space into relevant disjoint regions. [Tong and Lim (1980); Audrino and Bühlmann (2001); Audrino (2006); Audrino and Trojani (2006)]

Why threshold models?

- The probability to be at any given time in a specific regime is related to some relevant macroeconomic and/or term structure variables. [monetary policy conduction]
- The regimes are determined endogenously. [forecasting]
- We are able to disentangle macroeconomic from monetary policy changes. [clear regime interpretation]
- Provide better out-of-sample fit than the Markovian regime switching. [see, for example, Audrino (2006), Audrino and Medeiros (2010)]

Modeling framework

Our approach in brief

To infer the yield curve behavior, we use a model with four distinct features:

- ☐ to capture the cross sectional dynamics of the yield curve we employ latent term structure factors;
- ☐ we allow heteroskedasticity in the error term;
- ☐ motivated by the interpretability and the improved forecasting performance of the macro-factor literature in comparison to the pure finance approach, we incorporate macroeconomic variables;
- ☐ our model accommodates regime-switching behavior, but still allows interpretation and clear endogenous regime specification.

Model specification

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. We assume the following model for the term structure dynamics

$$\Delta y(t, n_\tau) = \mu(\Phi_{t-1,n_\tau}; \psi_{n_\tau}) + \varepsilon_{t,n_\tau}, \quad \tau = 1, \dots, T,$$

where

- $\mu(\Phi_{t-1,n_\tau}; \psi_{n_\tau})$ is a parametric function representing the conditional mean;
- $\varepsilon_{t,n_\tau} = \sqrt{h(\Phi_{t-1,n_\tau}; \psi_{n_\tau})} z_t$ is the error term of the yields' returns with maturity n_τ. $(z_t)_{t \in \mathbb{Z}}$ is a sequence of iid random variables with zero mean and unit variance, and $h(\Phi_{t-1,n_\tau}; \psi_{n_\tau})$ is the time-varying conditional variance.

Conditional dynamics

The conditional dynamics of the yields is given by

$$\mu_{t,n_\tau} = \sum_{j=1}^{K_{n_\tau}} (\alpha^j_{0,n_\tau} + \alpha^j_{1,n_\tau}\Delta y(t-1,n_\tau) + \beta^{\mathbf{j}\prime}_{\mathbf{n}_\tau}\mathbf{x_{t-1}} + \gamma^{\mathbf{j}\prime}_{\mathbf{n}_\tau}\mathbf{x^{ex}_{t-1}}) I_{[\Phi_{t-1,n_\tau}\in\mathcal{R}^j_{n_\tau}]},$$

$$h_{t,n_\tau} = \sum_{j=1}^{K_{n_\tau}} (\omega^j_{n_\tau} + a^j_{n_\tau}\epsilon^2_{t-1,n_\tau} + b^j_{n_\tau}h_{t-1,n_\tau}) I_{[\Phi_{t-1,n_\tau}\in\mathcal{R}^j_{n_\tau}]},$$

where

- $\psi_{n_\tau} = (\alpha^j_{0,n_\tau}, \alpha^j_{1,n_\tau}, \beta^{\mathbf{j}\prime}_{\mathbf{n}_\tau}, \gamma^{\mathbf{j}\prime}_{\mathbf{n}_\tau}, \omega^j_{n_\tau}, a^j_{n_\tau}, b^j_{n_\tau}, j = 1, \ldots, K_{n_\tau})$ is a parameter vector which parameterizes the local dynamics in the different regimes;
- x_{t-1} and x^{ex}_{t-1} are the relevant endogenous and exogenous variables at time $t-1$, respectively;
- K_{n_τ} is the number of regimes for maturity n_τ (estimated from the data).

Model estimation

Model estimation - Step 1

In order obtain an estimate for the unknown (true) parameters ψ we employ a two step procedure.

Step 1: Best subset selection - the main idea is to retain only a subset of the most informative variables, and to eliminate the noise variables from the model.

Advantages of the dimensionality reduction technique

- interpretability - we would like to identify a smaller subset that contains the most relevant information;
- prediction accuracy - including all possible prediction variables often leads to poor forecasts, due to the increased variance of the estimates in a model that is too complex;
- avoid data-mining problems;
- inline with the term structure literature.

Model estimation - Step 2

Step 2: Regime specification - the regimes are built as multiple tree-structured thresholds partitioning the predictor space G into relevant disjoint regions [Audrino and Bühlmann (2001); Audrino (2006); Audrino and Trojani (2006)].

Similar to CART the **estimation procedure** involves the following steps:

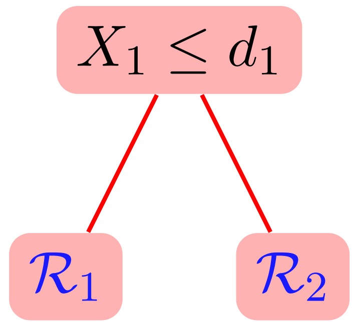
(i) Growing a large tree (a tree with a large number of nodes). The threshold selection is based on optimizing the conditional negative log-likelihood.

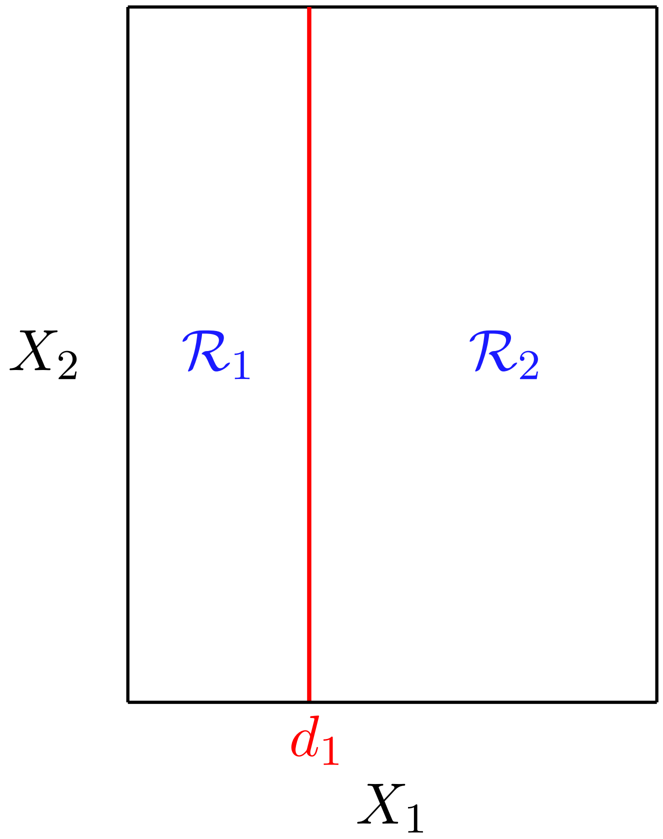
(ii) Combining some of the branches of this large tree to generate a series of sub-trees of different sizes (varying numbers of nodes).

(iii) Selecting an optimal tree via the application of measures of accuracy of the tree (BIC, AIC, C_p,...).

Illustration

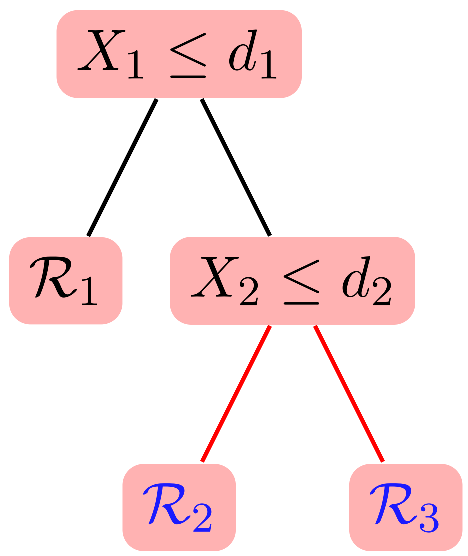
`Number of regimes: 2`

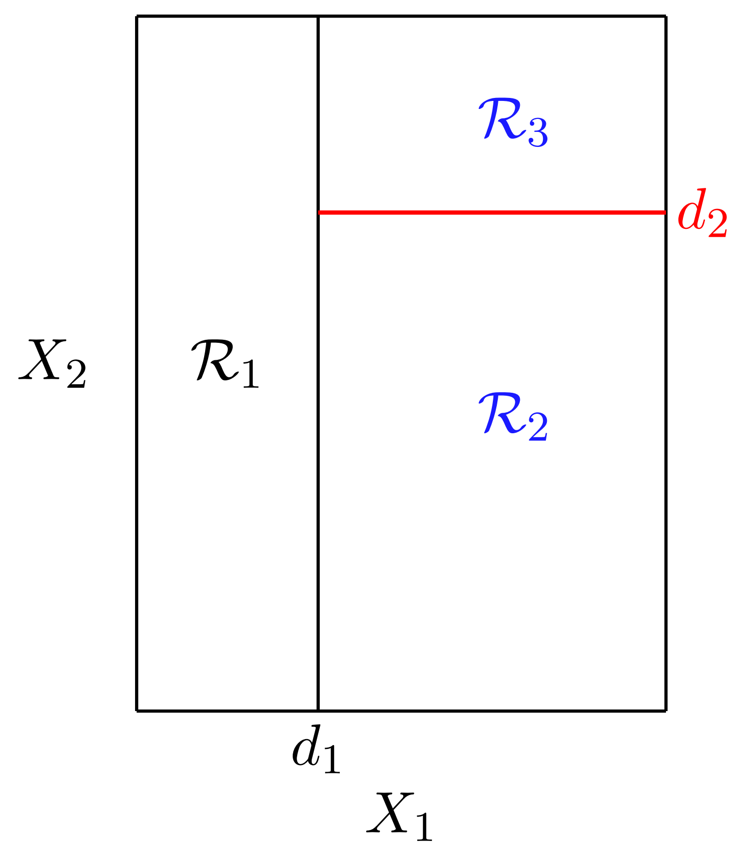




Illustration

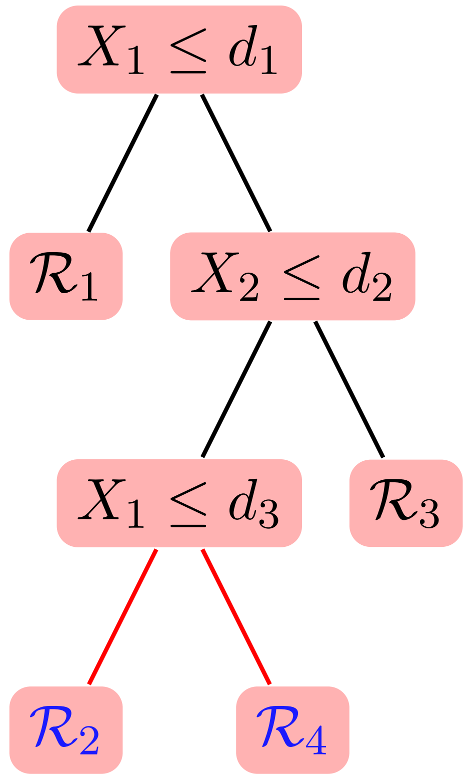
`Number of regimes: 3`

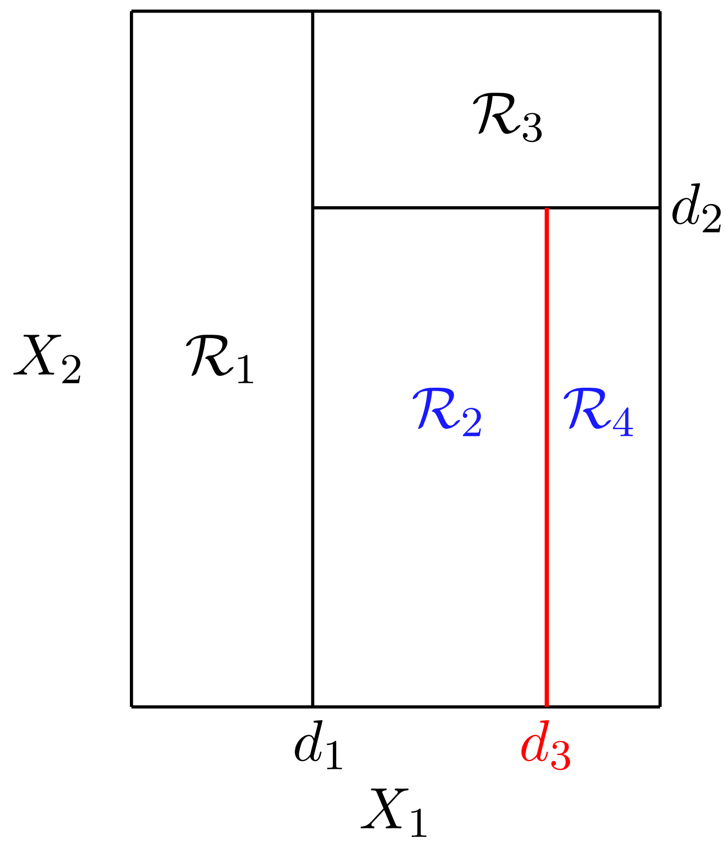




Illustration

Number of regimes: 4





Improving the forecasting ability: Bagging

Bagging is a machine learning technique aimed at reducing the variance and thus improving the forecasting performance of estimators such as trees. It involves the following steps:

- □ generate a large number of time series bootstrap resamples from the data;
- □ for each bootstrap sample apply the two–step procedure described above;
- □ average the forecasts of the conditional mean.

Initially bagging has been developed for cross sectional data [Breiman (1996)] and later extended to the time series framework. [see, for example, Inoue and Kilian (2004); Hillebrand and Medeiros (2007); Audrino and Medeiros (2008)]

Empirical Results

Data

Term structure data: U.S. Treasury bills (January 1952 - June 2005) 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.

Macroeconomic data: (January 1960 - December 2008) available from the Datastream International

- inflation: consumer price index (CPI), production price index (PPI);
- real activity: HELP, unemployment (UE), industrial production (IP).

In addition we construct the empirical proxies for:

- term structure level [10Y yield] and slope [10Y-3M yield];
- variance and conditional volatility of the macroeconomic data.

What is driving the yield curve predictability?

Optimal local mean dynamics

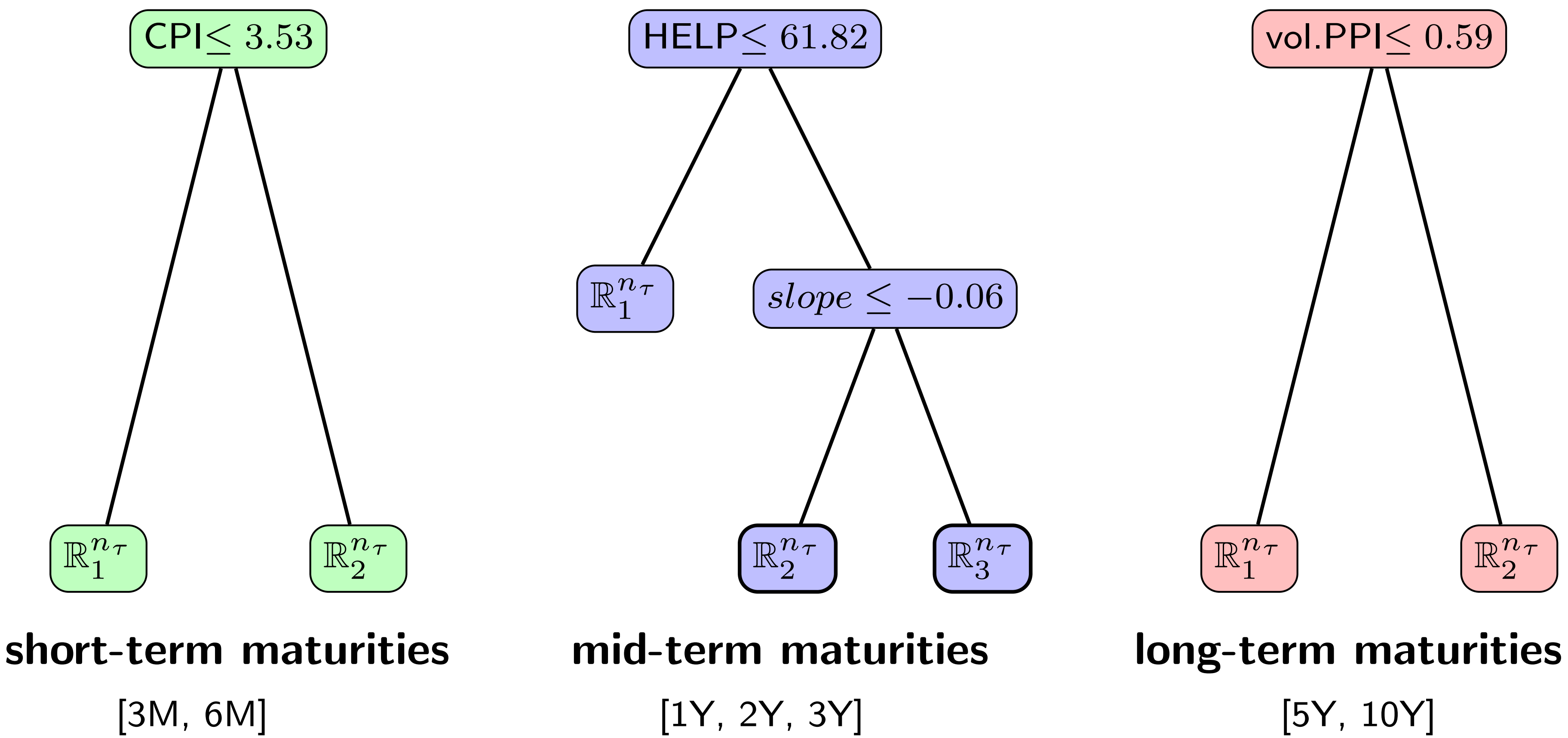
Maturity	Δy_{n_τ}	slope	level	PPI	HELP	HELP.sq	vol.PPI	vol.CPI
3M	$\star$	$\star$	$\star$	$\star$		$\star$	$\star$	
6M	$\star$	$\star$	$\star$	$\star$		$\star$	$\star$	
1Y	$\star$							
2Y	$\star$							
3Y	$\star$							
5Y			$\star$		$\star$		$\star$	$\star$
7Y			$\star$		$\star$		$\star$	$\star$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

Table 1: Yields local dynamics found for every maturity selected from a large number of potential term structure and macroeconomic predictors via best subset selection technique.

Clear pattern - the results can be summarized into 3 groups - short, mid-term and long-term maturities.

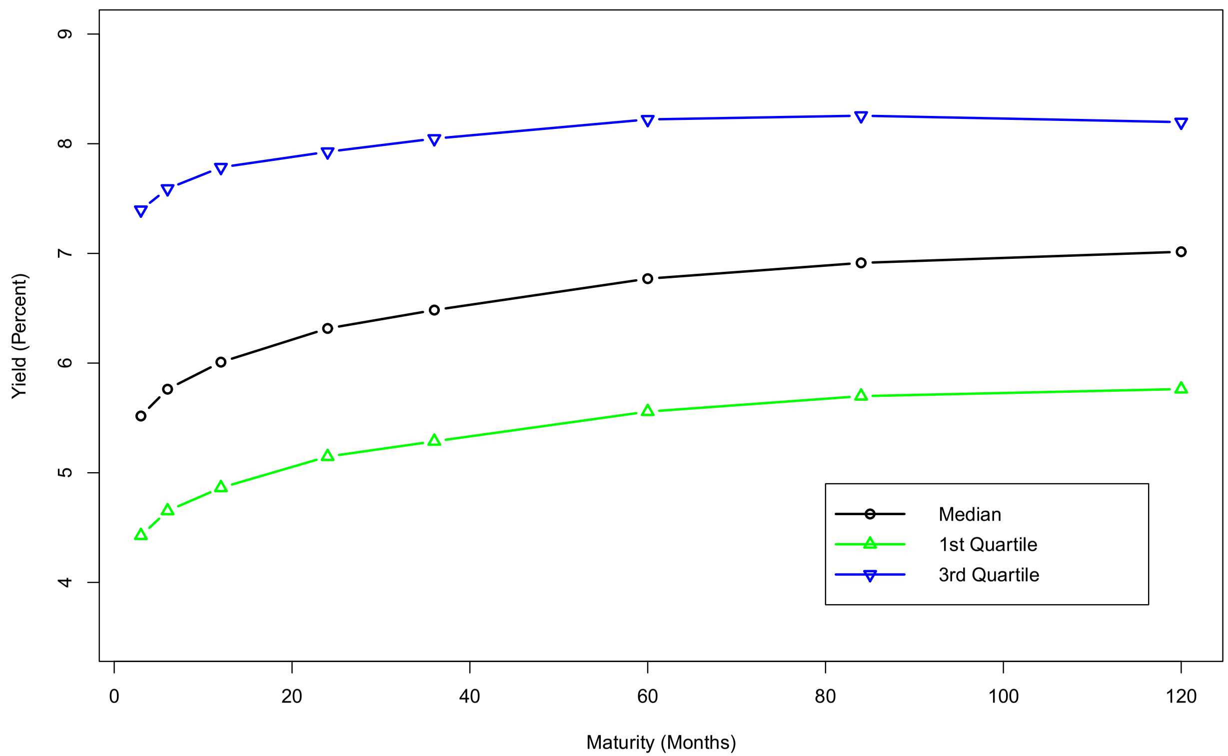
What is driving the yield curve predictability?

Optimal threshold structure



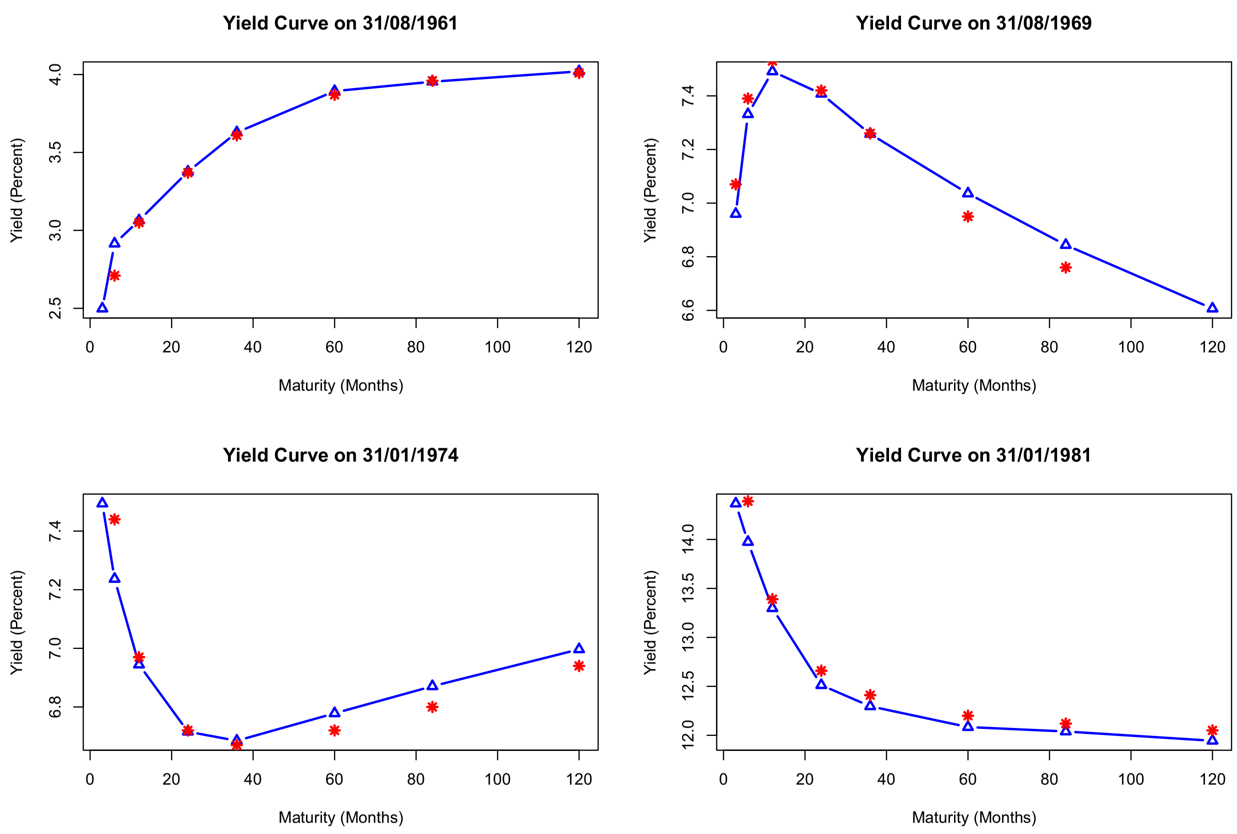
Similar to the local dynamics, we find a the same clear pattern.

Stylized facts: average yield curve



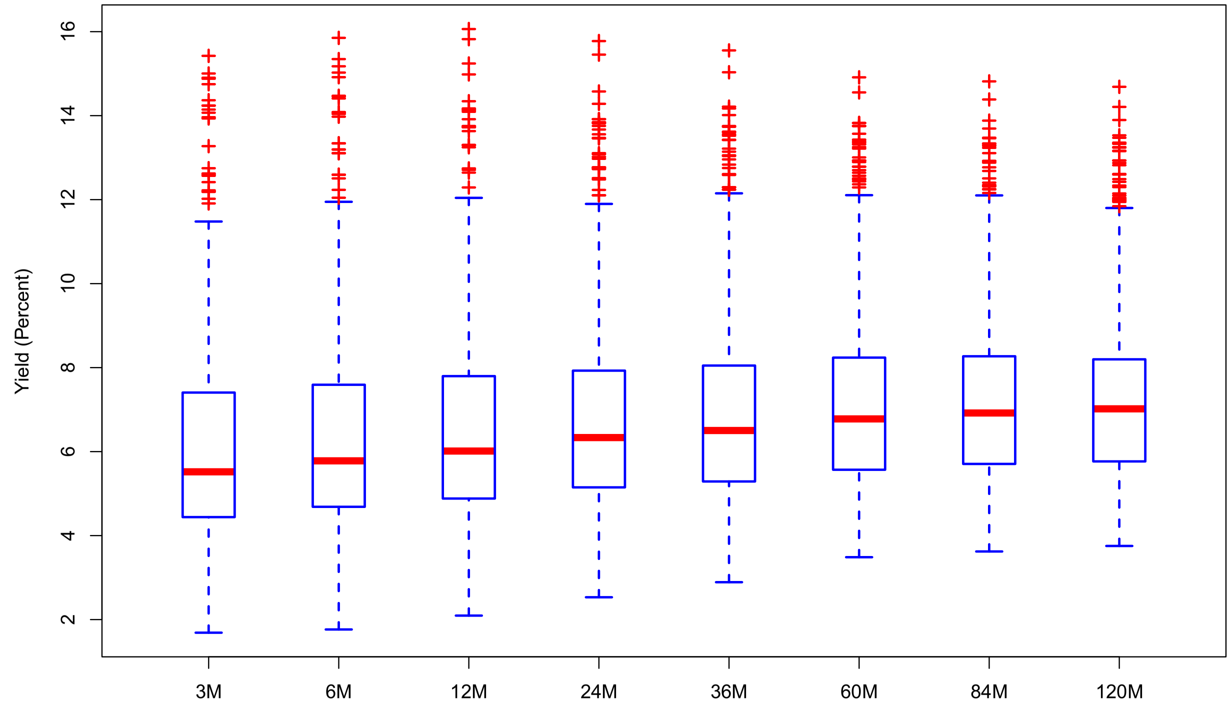
The average yield curve is upward sloping and concave.

Stylized facts: shapes of the yield curve



The yield curve assumes a variety of shapes through time - upward sloping, downward sloping, humped and inverted humped.

Stylized facts: volatility and persistency



The short end of the yield curve is more volatile than the long end and long rates are more persistent than short rates.

Out-of-sample performance

Out-of-sample MSE for the Bagged Models

	Macro Tree	Best Subset	NS AR(1)	Audrino Tree	Gray's RS
3M	0.0068 (0.595)	0.0820 (0)	0.5781 (0)	0.0128 (0.126)	0.1440 (0)
6M	0.0099 (0.526)	0.0368 (0.012)	0.4329 (0)	0.0196 (0.023)	0.0798 (0.004)
1Y	0.0284 (0.537)	0.0653 (0)	0.2420 (0)	0.0357 (0.512)	0.3754 (0)
2Y	0.0824 (0.642)	0.0905 (0.284)	0.0845 (0.591)	0.0887 (0.450)	0.3112 (0)
3Y	0.1149 (0.667)	0.1449 (0.484)	0.1550 (0.368)	0.1142 (0.652)	0.2941 (0)
5Y	0.1242 (0.657)	0.1434 (0.074)	0.1679 (0.027)	0.1230 (0.695)	0.2607 (0)
7Y	0.1155 (0.684)	0.1155 (0.684)	0.4108 (0)	0.1116 (0.672)	0.2707 (0.105)
10Y	0.0918 (0.629)	0.0995 (0.234)	0.1230 (0.047)	0.0951 (0.460)	0.2093 (0)

Table 2: p-values for SPA test [Hansen (2005)] are presented in parenthesis. Out-of-sample period January 2002 - July 2005.

Conclusion

Conclusion

We present a methodology to build and estimate a discrete–time regime–switching model of interest rates that

- incorporates latent and macroeconomic factors;
- 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

In terms of out-of-sample forecasting the bagged versions of our model are significantly better than almost all of the alternative approaches.