Essays on Macro-Financial Linkages

Essays on Macro-Financial Linkages

Rafael B. De Rezende

Dissertation for the Degree of Doctor
of Philosophy, Ph.D., in Finance
Stockholm School of Economics, 2014

Essays on Macro-Financial Linkages

ISBN 978-91-7258-931-5 (printed)
ISBN 978-91-7258-932-2 (pdf)

Front cover illustration:
© Luciano Queiroz (Shutterstock), "Old Vintage House at Serra da Canastra National Park - Minas Gerais – Brazil"

Back cover illustration:
© Arctistic, 2012

Printed by:
Ineko, Göteborg, 2014

Keywords:
Expectations hypothesis; term structure of interest rates; ex ante macroeconomic risks; quantile-based risk measures, out-of-sample forecasting; bond risk premia; excess bond returns; quantile regression; term spread; GDP growth; recessions; yield curve; in-sample fitting; Nelson-Siegel; quantile autoregression.

Foreword

This volume is the result of a research project carried out at the Department of Finance at the Stockholm School of Economics (SSE).

This volume is submitted as a doctor's thesis at SSE. In keeping with the policies of SSE, the author has been entirely free to conduct and present his research in the manner of his choosing as an expression of his own ideas.

SSE is grateful for the financial support provided by the Swedish Bank Research Foundation (BFI) which has made it possible to fulfill the project.

Göran Lindqvist
Director of Research
Stockholm School of Economics

Magnus Dahlquist
Professor and Head of the
Department of Finance
Stockholm School of Economics

Acknowledgements

This thesis marks the end of my PhD studies at the Stockholm School of Economics, a fantastic journey full of excitement and challenges. This journey could not be completed without the support of friends, family and people around me. My thanks go to all of you.

Special gratitude goes to Magnus Dahlquist. As my supervisor, Magnus has helped me immensely throughout the years. I benefited a lot from his knowledge about financial economics and from our discussions about new research ideas and results. His mentoring was essential throughout my PhD. I will always remember Magnus asking me to simplify my papers!

I also would like to give special thanks to Michael Halling and Lars E.O. Svensson. Michael was very enthusiastic about my research, very generous with his time and was always open to give advices or for a quick chat. Lars had always something to teach me. I truly admire his wisdom, kindness and ability to explain difficult concepts using simple words. Talk to Lars for one hour and learn the equivalent to one day.

I am also indebted to the staff and faculty at the Finance Department and SHoF, who gave me all the support I needed for producing my research in a fantastic academic environment. I would like to thank, in particular, Anneli Sandbladh, Roméo Tédongap, Peter Englund, Cristian Huse, Mike Burkart, Jungsuk Han, Irina Zviadadze, Clas Bergström and Per Strömberg.

Furthermore, I have benefited greatly from my fellow Ph.D students at SSE. These include my friends Qing Xia, who was my office mate during the whole PhD program, and Ricardo Aliouchkin, who was always available for a late lunch and a good chat; and the mates Adam Farago, Nikita Koptyug, Mariana Khapko, Andrejs Delmans, Jieying Li and Patrick Augustin.

Last and most importantly, I would like to thank my family. My parents Viviana and Eduardo, and my sister Juliana, have been a source of unconditional love throughout my whole life. Avani, Egon and Maria Julia, who are now part of the family, for all the pleasant moments in Stockholm. And finally, my wife Bartira, who has been my source of support and encouragement. Her love, joy and light kept my soul alive throughout this project. I dedicate this thesis to her.

Stockholm, June 2014

Rafael B. De Rezende

Contents

List of Tables

List of Figures

Introduction

The present thesis is a collection of research papers on the analysis of the term structure of interest rates with a focus at the intersection of macroeconomics and finance. The topic is important not only to academics, but also to investors and policymakers. For instance, investors may want to predict the future dynamics of the term structure of interest rates to determine their optimal portfolio allocation across different assets. Policymakers are often interested in extracting market expectations of policy rates and other macroeconomic variables from long-term interest rates, and in taking actions to influence them, as the whole term structure is important for the investment and borrowing decisions of households and businesses. An emphasis is put on the analysis of the failure of the expectations hypothesis of the term structure of interest rates, which asserts that long-term rates are simple averages of short-rate expectations, and that risk premium, if not equal to zero, is constant over time.

The first paper *"Risks in macroeconomic fundamentals and bond return predictability"* contributes to the macro-finance literature by studying the macroeconomic forces behind risk premia in the US government bond market. Emphasis is put on the measurement of risks underlying the macroeconomy and their relationship with the term structure of interest rates. More specifically, the paper provides evidence that measures of macroeconomic risks related to expectations, uncertainty and downside (upside) macroeconomic risks are able to explain variation in bond risk premia across maturities. Factors, referred to as macro risk factors, extracted from these measures are found to be powerful predictors of bond excess returns. In addition, it is shown that they provide new information about bond risk premia when compared to forward rates and current macroeconomic conditions, which are often used to explain time-variation in expected bond excess returns.

The paper also contributes to a vast literature on the use of factors beyond the yield curve to describe the behavior of the term structure of interest rates using affine models. I document that macro risk factors provide information about variation in bond risk premia that is not spanned by the yield curve and discuss its practical implications for the identification and estimation of the term premium component in long-term yields using these models. Accordingly, the estimation of an affine term structure model with unspanned macro risk factors reveals that they carry significant prices of risk and generate time-varying and countercyclical term premium, as suggested by economic theory.

It happens, however, that in-sample predictive power does not necessarily imply out-of-sample predictability. Recent research evaluating the predictability of excess returns in government bond and equity markets has documented that many variables that are shown to work quite well in-sample have not had the same success in an out-of-sample framework. The second paper *"Out-of-sample bond excess returns predictability"* builds on this argument to assess the statistical and economic significance of out-of-sample forecasts of bond excess returns based on forward rates, macroeconomic variables and risks in macroeconomic outcomes. Results suggest that macroeconomic variables, risks in macroeconomic outcomes as well as the combination of these different sources of information outperform a constant model of no-predictability that is consistent with the expectations hypothesis of the term structure of interest rates. These results are confirmed when using macroeconomic data available in real-time, indicating that the predictability of bond returns is not driven by data revisions, as suggested by recent research.

The research papers above relating bond risk premium (or term premium) to subsequent developments in macroeconomic variables are part of a much larger literature on the predictive power of the slope of the yield curve and its relationship with future GDP growth, recessions and inflation. This is the object of study of the third paper, *"Re-examining the predictive power of the yield curve with quantile regression"*, in co-authorship with Mauro S. Ferreira. More specifically, in this paper we use quantile regression methods to re-examine the predictive ability of the yield curve slope (or term spread) with respect to future GDP growth throughout the entire conditional distribution of the latter variable. The simple, but novel approach, allows us to reach a number of interesting conclusions. The term spread is found to be a better predictor of negative to intermediate conditional GDP growth in the US, confirming its usefulness for predicting recessions and future economic activity. In addition, changes in the predictive relationship towards longer horizons and structural breaks at upper percentiles and dated around 1984 suggest the Fed started to respond tougher and in great advance to inflationary pressures resulted from excessive growth after the mid-1980's. Motivated by these findings we use quantile regression to forecast GDP growth and recessions probabilities in an out-of-sample scheme. Results suggest that quantile models deliver more accurate forecasts than competitors in both exercises. We also compare the performance of the spread-based models to professional forecasters. Superiority is found for mid and longer horizons, but higher accuracy in shorter horizons was also observed in the period prior the 2008/2009 recession. We conclude that the predictive power of the yield curve remains, despite the changes in the GDP growth - term spread relationship.

The fourth paper, *"Modeling and forecasting the yield curve by an extended Nelson-Siegel class of models: a quantile regression approach"*, also in co-authorship with Mauro S. Ferreira, deals with yield curve prediction. More specifically, the paper compares the in sample fitting and the out of sample forecasting performances of four different Nelson-Siegel type of models: Nelson-Siegel, Bliss, Svensson, and a five factor model we propose in order to enhance fitting flexibility. The introduction of the fifth factor resulted in superior adjustment to the Brazilian data. For the forecasting exercise the paper contrasts the performances of the term structure models in association to four different econometric methods: quantile autoregression evaluated at the median, VAR, AR, and a random walk. Results suggest that the quantile procedure delivered the best results for longer forecast horizons, which may be explained by robustness of the quantile regression method, especially because the paper deals with very volatile financial data characterized by the presence of extreme values that tend to bias mean estimators.

Chapter 1

Risks in macroeconomic fundamentals and bond return predictability[1]

Rafael B. De Rezende

ABSTRACT. I extract factors from quantile-based risk measures estimated for US macroeconomic variables and document that risks in macroeconomic fundamentals contain valuable information about bond risk premia. Macro risk factors predict bond excess returns with power above and beyond the Cochrane-Piazzesi and Ludvigson-Ng factors, with results being verified statistically as well as economically. Importantly, macro risk factors generate countercyclical bond risk premia and capture unspanned predictability in bond excess returns. These results provide further support for the idea that predictability of bond excess returns cannot be completely summarized by the yield curve. Risks in macroeconomic fundamentals should also be taken into account.

Keywords: ex ante macroeconomic risks; bond risk premia; quantile-based risk measures.
JEL Classifications: G12, G17, G11, E43, E44

1.1 Introduction

Empirical research in financial economics has revealed significant predictable variation in expected excess returns of US government bonds, a violation of the expectations hypothesis. Understanding this variation and its relationship with the economy has been an important question in economics and finance, and an active

[1] I would like to thank Magnus Dahlquist, Lars E.O. Svensson and Michael Halling for comments and suggestions that significantly improved this paper. I am also grateful to Ádám Faragó, Roméo Tédongap, Erik Hjalmarsson, Andrejs Delmans, Nikita Koptyug, Ricardo Aliouchkin and seminar participants at the Swedish House of Finance, the XXI Finance Forum, the National PhD Workshop in Finance 2013 and the 9th BMRC-QASS Conference on Macro and Financial Economics for comments and suggestions. I kindly thank the Swedish Bank Research Foundation (BFI) for financial support.

area of ongoing research. Fama (1984), Fama and Bliss (1987), Stambaugh (1988) and Cochrane and Piazzesi (2005) find that yield spreads and forward rates predict excess bond returns with R^2s ranging from 10% to 40%. Ludvigson and Ng (2009) and Cooper and Priestley (2009) document that macroeconomic variables carry information about bond risk premia not contained in financial variables.

In this paper I study the links between the macroeconomy and bond risk premia from a forward looking perspective. I start by asking the following questions. How do distributions of *future* macroeconomic outcomes evolve over time? Do they reveal the existence of any potential risks in the macroeconomy such as risks of extreme macroeconomic outcomes, downside (upside) risks and macroeconomic uncertainty? If so, do these potential macroeconomic risks carry any information about bond risk premia?

In order to answer these questions, I provide three simple measures that allow me to empirically study how conditional distributions of future macroeconomic variables evolve over time. First, as a measure of central tendency, the *median* (Med) is a natural candidate for describing the center of predicted conditional distributions, in particular when data are highly asymmetric.[2] The second measure, the *interquantile range* (IQR), captures how spread out conditional distributions are, while the third measure, the *interquantile skewness* (IQS), as the corresponding quantile-based metric of conditional skewness, can be used to gauge the degree of conditional distributions' asymmetry.

The three measures share the property of having appealing economic interpretations. First, as one way of measuring the typical values variables may assume in the future, the median arises as a natural metric of macroeconomic *point expectations*. The interquantile range allows capturing the current level of *uncertainty* about future macroeconomic outcomes, while the interquantile skewness provides a natural characterization of the *downside (upside) risks* regarding the future state of the economy. I concentrate my analysis on the top and bottom 5% conditional quantiles, meaning that both IQR and IQS also allow capturing information on macroeconomic *tail risks*, providing an even richer description of the risks involving future macroeconomic variables. From here on I refer to these measures − Med, IQR and IQS − as measures of *ex ante macroeconomic risks*, where "ex ante" stands for the fact that all measures can be obtained at date t.

Using simple quantile regression methods (Koenker and Basset, 1978), I esti-

[2] As is well known the median may be preferable to the mean if the distribution is long-tailed. The median lacks the sensitivity to extreme values of the mean and may represent the position (or location) of an asymmetric distribution better than the mean. For similar reasons in the regression context one may be interested in median regressions.

mate risk measures for six variables closely related to business cycles in the US: GDP price index, real GDP, unemployment, industrial production, housing starts and corporate profits; and document several interesting features of their predicted conditional distributions. First, I find that the interquantile range for all variables, except inflation, shows pronounced countercyclical behavior, indicating increasing levels of uncertainty regarding the future state of the US economy during recessions. Using different approaches this result is also documented by Jurado, Ludvigson and Ng (2013) and Bansal and Shaliastovich (2013). Interestingly, for the last great recession, the level of uncertainty for housing starts rises right from 2004, the year when subprime mortgage lending increased dramatically in the US. I also document that ex ante lower tail risks for GDP and industrial production move cyclically, while the opposite holds for the other variables. When it comes to asymmetries, I find that predicted conditional distributions for inflation and unemployment (industrial production and housing starts) growth are mostly positively (negatively) skewed, indicating the presence of consistent upside (downside) risks for these variables.

The estimated risk measures also reveal themselves as important determinants of bond risk premia variation. I extract a few factors, referred to as macro risk factors, that summarize almost all the information about bond risk premia contained in the estimated risk measures and verify that they predict future bond excess returns across different maturities with R^2s ranging from 20% to 30%. Importantly, the macro risk factors have economic interpretations in terms of the ex ante macroeconomic risks I estimate. Point expectations regarding economic activity variables, uncertainty about GDP growth and downside (upside) risks with respect to housing starts, unemployment and inflation are shown to be important determinants of bond risk premia variation in the US.

I also form a single macro risk factor following the idea of Cochrane and Piazzesi (2005, CP hereafter). The new single factor predicts excess bond returns with power beyond CP and Ludvigson and Ng (2009, LN hereafter) factors. For instance, the single factor explains variation in excess bond returns with R^2s of up to 31%. Combining it with the CP and LN factors increases models' predictive power to levels around 45%, indicating that risks in macroeconomic fundamentals capture information about bond risk premia that is not embedded in forward rates and current macroeconomic variables. Importantly, the new factor shows a pronounced countercyclical behavior, consistent with theoretical models asserting that investors must be compensated for risks associated with recessions (Campbell and Cochrane, 1999; Wachter, 2006; Bansal and Yaron, 2004; Rudebusch and Swanson, 2009). Much of this evidence can be explained by the countercyclical behavior of the ex ante macro risks I estimate.

Consistent with recent research, I also find that macro risk factors capture predictability in bond excess returns that is largely unspanned by the yield curve (Duffee, 2011; Joslin, Priebsch and Singleton, forthcoming). This result has important implications for the identification of the term premium component of yields using affine term structure models, as models of this class commonly disregard the information about expected excess returns contained in factors beyond the yield curve (Ang and Piazzesi, 2003; Ang, Dong, and Piazzesi, 2007; Rudebusch and Wu, 2008). Accordingly, the estimation of an affine model with unspanned macro risk factors reveal that they carry significant prices of risk and generate time-varying and countercyclical term premia.

My findings have important implications for both finance and macroeconomics. In finance, a clearer understanding of the determinants of bond risk premia helps explain why investors demand higher compensation for bearing the risk of holding long-duration bonds, in particular during bad times. Furthermore, the notion that the information in the whole distribution of future macroeconomic outcomes matters for explaining bond risk premia variation challenges the prevailing use of conditional point expectations and uncertainty alone underlying many asset pricing models. For macroeconomics, the better identification of the term premium component of yields helps to clarify the relationship between short and long interest rates and facilitates the understanding of the transmission mechanisms of monetary policy, as the whole yield curve is important for the investment and borrowing decisions of households and businesses.

Related literature

The present study adds to the literature examining the failure of the expectations hypothesis of the term structure of interest rates and its determinants (Fama, 1984; Fama and Bliss, 1987; Stambaugh, 1988; Cochrane and Piazzesi, 2005; Ludvigson and Ng, 2009; Cooper and Priestley, 2009; Huang and Shi, 2012; Cieslak and Povala, 2012; among others). My main contribution is to show that risks in macroeconomic fundamentals are able to explain variation in bond risk premia. Furthermore, I show that the information they provide is, to a large extent, unrelated to that contained in financial and current macroeconomic variables.

This article also relates to the works by Bansal and Shaliastovich (2013), Buraschi and Whelan (2012), Wright (2011), Dick, Schmeling and Schrimpf (2013) and Chun (2011) which document that bond risk premia are influenced by expectations and uncertainties about the future state of the economy. [3] A drawback of most of these studies, however, is that they have relied exclusively on survey-based proxies such as the consensus forecast and the dispersion of the cross-sectional distribution of analysts' forecasts to measure these variables, a strategy that has a number of weaknesses. First, as respondents typically sampled are professional forecasters, surveys do not necessarily represent expectations of financial market participants. Second, some analysts may provide potentially strategic forecasts or omit relevant forecasting information (Ottaviani and Sorensen, 2006), while surveys also commonly suffer from the small number of cross sectional observations at certain dates. Differently, I estimate macroeconomic expectations and uncertainties by relying on quantile regression methods (Koenker and Basset, 1978), accommodating a rich information set composed by analysts' forecasts, expectations of consumers and by financial variables that are known to be good predictors of several macroeconomic indicators. This makes feasible the use of information that is more likely to span the unobservable information sets of bond market investors when forming their macroeconomic predictions. Moreover, I analyze the role played by asymmetries and tail risks.

In a recent paper Colacito, Ghysels and Meng (2013) rely on the first three moments of the cross-sectional distribution of analysts' forecasts to show that conditional point expectations and skewness are able to explain the equity premium. They rationalize their results using a long-run risk model that introduces asymmetry in the distribution of expected consumption growth rates and show that their model can account for a number of observed features of the distribution of equity returns. The present study is closely related to their work as I also look at asymmetries in distributions of macroeconomic forecasts. However, in addition to GDP growth, I analyze the role played by a number of other macro indicators and study risk premia in the government bonds market. Furthermore, I do not rely exclusively on survey data to estimate my objects of interest.

Another related strand of research aims at measuring macroeconomic risks, which is commonly done using volatility. My measures go beyond volatility as I also look at skewness and tail risks. In a similar spirit, Kitsul and Wright (2012) rely on CPI based options to construct probability densities for inflation and use them to

[3] Macroeconomic uncertainty and disagreement are terms that have been used interchangeably in this literature. For instance, Buraschi and Whelan (2012) study both theoretically and empirically the links between macroeconomic disagreement, or differences in beliefs, and bond markets. Their empirical measure of macroeconomic disagreement - the mean absolute deviation of professional forecasts - however, can be also interpreted as a measure of macroeconomic uncertainty as it simply measures the dispersion of the cross-sectional distribution of forecasts as in many other papers (Lahiri and Liu, 2006; Giordani and Söderlind, 2003; Wright, 2011).

measure deflation and high inflation risks. De Rezende and Ferreira (2013) rely on quantile regression and the term spread to forecast probabilities of future recessions. Gaglianone and Lima (2012) use quantile regression to construct density forecasts for macro variables and use them to estimate risks of high unemployment rates. Christensen, Lopez and Rudebusch (2011) rely on Treasury Inflation Protected Securities to measure deflation probabilities.

Other papers measure macroeconomic risks from the distribution of forecasts provided by surveys. Garcia and Werner (2010) extract measures of inflation risks such as asymmetry and uncertainty from the cross-sectional distribution of professional forecasts. In a similar spirit, Giordani and Söderlind (2003) look at uncertainty only. Andrade, Ghysels and Idier (2012) propose new measures of inflation tail risk, uncertainty and asymmetry similar to the ones used in this paper. The authors rely on inflation probability distributions obtained from each forecaster to estimate their measures of inflation risk. Differently, I estimate my risk measures using quantile regression methods (Koenker and Basset, 1978) and discuss how this approach allows extending the notion of risks in macroeconomics to any variable of interest.

The rest of the paper is organized as follows. The next section introduces the measures of macroeconomic risks used in the paper and discusses how they are estimated; the third section presents the econometric framework proposed for predicting excess bond returns; in the fourth section I discuss the main results of the paper; and the last section concludes.

1.2 Measures of ex ante macroeconomic risks

1.2.1 Median, interquantile range and interquantile skewness

I provide three simple measures, each with three distinguishing features, namely, robustness to outliers, the ability to capture time variation in future conditional distributions and finally that they can be obtained for any h-period ahead macro variable, $z_{t,t+h}$, using information known at date t. I start by defining my first object of interest, the median. Let $z_{t,t+h}$ denote the annual log rate of change of macroeconomic variable Z during the period t to $t+h$, and $F_{z_{t,t+h}}(x)$ be its cumulative distribution function (CDF) conditional on date t information Ω_t,

$$F_{z_{t,t+h}}(x) = Pr\left(z_{t,t+h} \leq x|\Omega_t\right) \quad (1.1)$$

Let also $q_{z_{t,t+h}}(\tau) = F^{-1}_{z_{t,t+h}}(\tau)$ be its conditional quantile associated with probability $\tau \in (0,1)$, assuming that $F_{z_{t,t+h}}(x)$ is strictly increasing. I then define

$$Med^h_t = q_{z_{t,t+h}}(0.5) \tag{1.2}$$

as the median of $F_{z_{t,t+h}}$, measured at time t. The median is one of a number of ways of summarizing typical values that can be assumed by $z_{t,t+h}$. Unlike the mean or the mode, however, the median presents the appealing property of robustness, being an attractive candidate for forecasting $z_{t,t+h}$, especially in the presence of outliers and conditional asymmetries in the data.

The second measure is the interquantile range of the conditional distribution of $z_{t,t+h}$. As the simplest robust measure of data dispersion, the interquantile range provides a natural way of gauging how spread out the conditional distribution of $z_{t,t+h}$ is. More precisely, given $q_{z_{t,t+h}}(\tau)$, the interquantile range of the conditional distribution of $z_{t,t+h}$ associated to the level τ, $\tau < 0.5$, is defined as

$$IQR^h_t(\tau) = q_{z_{t,t+h}}(1-\tau) - q_{z_{t,t+h}}(\tau) \tag{1.3}$$

The third measure is based on Hinkley's (1975) generalization of Bowley's (1920) robust coefficient of asymmetry (skewness). It is defined as the interquantile skewness of the conditional distribution of $z_{t,t+h}$ associated to level τ, with $\tau < 0.5$, or more precisely,

$$IQS^h_t(\tau) = \frac{(q_{z_{t,t+h}}(1-\tau) - q_{z_{t,t+h}}(0.5)) - (q_{z_{t,t+h}}(0.5) - q_{z_{t,t+h}}(\tau))}{q_{z_{t,t+h}}(1-\tau) - q_{z_{t,t+h}}(\tau)} \tag{1.4}$$

The normalization in the denominator ensures that the measure assumes values between -1 and 1. If the right quantile is further from the median than the left quantile, then IQS is positive indicating that there is a higher probability that $z_{t,t+h}$ will be above the median than below, while the opposite yields a negative coefficient. Also an advantage of this measure is that because it does not cube any values, it is more robust to outliers than the conventional third-moment formula (Kim and White, 2004). Other papers that have used the interquantile skewness in empirical macro and finance include Conrad, Dittmar and Ghysels (2013), White, Kim, and Manganelli (2008), Ghysels, Plazzi, and Valkanov (2010) and Andrade, Ghysels and Idier (2012).

1.2.2 Economic interpretation and estimation

The risk measures defined above share one additional property $-$ that of having appealing economic interpretations in terms of measures of macroeconomic point expectations, uncertainty and downside (upside) macroeconomic risks. To motivate, I assume that there exists a continuum of individuals in the economy who are interested in making predictions about $z_{t,t+h}$. Following the literature on forecast optimality (Granger, 1969; Granger and Newbold, 1986; Christoffersen and Diebold, 1997; Patton and Timmermann, 2007; Gaglianone and Lima, 2012, Gneiting, 2011), I consider that each individual chooses an optimal forecast $\widehat{z}^i_{t,t+h}$ by minimizing an expected individual loss function L^i. In this case, disagreement among individuals about $z_{t,t+h}$ will result from the heterogeneity of loss functions and, as discussed below, optimal estimates for Med, IQR and IQS can be obtained through the estimation of simple quantile regressions for certain individuals' corresponding percentiles.[4]

More precisely, I assume that each individual i makes his optimal prediction $\widehat{z}^i_{t,t+h}$ by minimizing an expected asymmetric loss function L^i defined as below,

Assumption: *the loss function is of the form*

$$L^i\left(\tau^i, z, \widehat{z}\right) = \begin{cases} \left(\tau^i - 1\right)\left[\mathrm{f}(z_{t,t+h}) - \mathrm{f}(\widehat{z}_{t,t+h})\right], \; if \; z_{t,t+h} < \widehat{z}_{t,t+h} \\ \tau^i\left[\mathrm{f}(z_{t,t+h}) - \mathrm{f}(\widehat{z}_{t,t+h})\right], \; otherwise \end{cases} \tag{1.5}$$

where $\tau^i \in (0,1)$ *and* f *is a strictly increasing real function.*

The parameter τ^i describes the degree of asymmetry in the individual i's loss function. Values less than one half indicate that overpredicting induces greater loss to the individual than underpredicting by the same magnitude. In the symmetric case τ^i equals one half and the individual i's cost due to over- and underpredictions are the same.

It can be shown that if individual i's loss function is of the form given in (1.5), then his optimal forecast $\widehat{z}^i_{t,t+h}$ is simply the respective conditional quantile of $z_{t,t+h}$ (Gneiting, 2011; Komunjer, 2005; Komunjer and Vuong, 2010), or more specifically

$$\widehat{z}^i_{t,t+h} = q_{z_{t,t+h}}(\tau^i)$$

[4] This approach was firstly proposed by Gaglianone and Lima (2012) for constructing density forecasts for macro variables using a location-scale model and very general individual loss function. In my approach, the shape of the distribution of $z_{t,t+h}$ is also allowed to vary over time. This is important to estimate time-varying interquantile skewness.

Optimal individuals' predictions are then based on the same information set and differences in opinion are explained by differences in loss functions or, more specifically, by differences in the parameter τ^i.[5] Importantly, the real function f which enters individuals' losses need not be known. This means that quantiles are optimal forecasts for a large number of loss functions obtained by letting f vary in the set of monotone increasing functions. This result is very powerful as strong assumptions on the form of the loss function used by individuals are not needed.[6]

Assuming that the quantiles of $z_{t,t+h}$ can be approximated by a linear function as $q_{z_{t,t+h}}(\tau) = \beta(\tau)' x_t$, where x_t is a $k \times 1$ vector of covariates and $\beta(\tau)$ is a $k \times 1$ vector of coefficients to be estimated, it follows that parameters $\beta(\tau)$ will depend solely on L^i and optimal estimates for Med, IQR and IQS can then be simply obtained through the estimation of quantile regressions as

$$q_{z_{t,t+h}}(\tau) = \beta(\tau)' x_t \tag{1.6}$$

for certain individuals'-corresponding percentiles $\tau \in (0,1)$, where $\beta(\tau)$ can be estimated as in Koenker and Basset (1978) (see Appendix 1.A for details).

This means that Med, IQR and IQS can be viewed as specific characteristics of the cross-sectional distribution of individuals' optimal forecasts and can then be interpreted as measures of ex ante macroeconomic risks. Med can serve as a measure of the consensus forecast (median) that is commonly used as a proxy for macroeconomic point expectations. IQR can be viewed as a measure of uncertainty or individuals' disagreement about $z_{t,t+h}$, while IQS can serve as a measure of downside (upside) macroeconomic risks as negative values for IQS, for instance, indicate that more individuals believe that $z_{t,t+h}$ will be below its median value than above. Finally, it is also crucial to point out that when evaluated at percentiles close to zero, IQR and IQS share the attractive property of also capturing information on both the upper and lower tails of the conditional distribution of $z_{t,t+h}$, that is, they can also be used to model tail risks regarding this variable, allowing for a rich characterization of the risks involving the future state of the economy. Macroeconomic tail risks enables the estimation of risks of extreme macroeconomic outcomes such as, for example, risks of a large drop in economic activity, high

[5]Patton and Timmermann (2010) point out that differences in opinion are not primarily driven by differences in information.

[6]Empirical evidence on asymmetric loss has been found in forecasts for different economic variables made by different economic agents. See Elliot, Kmounjer and Timmermann (2005) for evidences on governments' budget deficits forecasts made by the IMF and the OECD; Elliot, Kmounjer and Timmermann (2008) on GDP growth forecasts generated by professional forecasters; Capistrán (2008) on inflation forecasts made by the Federal Reserve; Clatworthy, Peel and Pope (2012) for evidence on firms' earnings forecasts generated by analysts.

inflationary pressures or even a boom in the housing market, which can have important implications for explaining risk premia in equity and bond markets (Gourio, 2013; Gabaix, 2012; Tsai and Wachter, 2013).

Although linear, model (1.6) allows for great flexibility. The advantage relies on the estimation of one regression for each conditional quantile of the response variable, meaning that covariates x_t are allowed to affect the shape of the distribution of $z_{t,t+h}$, which may be Gaussian, but can also assume non-standard forms. Figure 1.1 illustrates this with several quantile lines estimated for inflation and growth in GDP, unemployment, industrial production, housing starts and corporate profits. Notice that due to the flexibility of the quantile regression approach, predicted conditional distributions are allowed to assume quite interesting shapes, while they are also able to capture several interesting features of the data as, for instance, the increasing levels of uncertainties, tail risks as well as downside (upside) risks in these variables, in particular around recessions dates. Notice that while the median is able to match realized values on many dates, it misses important periods of macroeconomic stress. The indicators of macroeconomic tail risks, on the other hand, seem to capture extreme movements in macroeconomic variables with higher accuracy. This result is evident for several variables during the 2008/2009 recession.

Variables entering the vector x_t were chosen in such a way to maximize the benefits of a large information set while minimizing the curse of dimensionality problem that may limit any forecasting model (Stock and Watson, 2005). In this paper, I follow Gaglianone and Lima (2012) who propose the use of analysts' consensus forecasts to construct density forecasts for macroeconomic variables using quantile regressions, but I augment their specification with information from additional predictors as in Aiolfi, Capistrán and Timmermann (2011). More specifically, I consider a specification that combines the equal-weighted survey forecast, or consensus forecast, with three other covariates that are known to contain information about $z_{t,t+h}$,

$$x_t^{'} = \left(1,\, z_t^{SPF,h},\, Mich\,Expect_t,\, 5 - year\,term\,spread_t,\, Baa\,corp\,spread_t\right) \quad (1.7)$$

where $z_t^{SPF,h}$ is the h period ahead consensus (mean) forecast for variable z obtained from the Survey of Professional Forecasters (SPF hereafter) reported at time t, $Mich\ Expect_t$ is the University of Michigan consumer expectations index (MCEI hereafter), $5 - year\ term\ spread_t$ is the 5-year TBond rate spread over the 3-month TBill rate (5yTS hereafter) and $Baa\ corp\ spread_t$ is the Moody's Baa corporate rate spread over the 3-month TBill rate (BaaCS hereafter).

Recent works studying the links between bond risk premia and macroeconomic expectations and uncertainties have relied exclusively on survey based proxies such as the consensus forecast and the dispersion of the cross-sectional distribution of analysts' forecasts for measuring these variables (Buraschi and Whelan, 2012; Wright, 2011; Dick, Schmeling and Schrimpf, 2013; Chun, 2011).[7] A drawback of this strategy, however, is that survey' respondents typically sampled are professional forecasters meaning that they may not accurately represent expectations of financial market participants. Moreover, some analysts may provide potentially strategic forecasts or omit relevant forecasting information (Ottaviani and Sorensen, 2006), while surveys commonly suffer from a small number of cross sectional observations at certain dates.

The main advantage of the approach I propose is the possibility of estimating these variables using information that is more likely to span the unobservable information sets of bond market participants. While $z_t^{SPF,h}$ is a good source of information about analysts' expectations (Capistrán and Timmermann, 2009), MCEI, which has been shown to be a good predictor of future macro variables (Ang, Bekaert and Wei, 2007), is able to capture consumers' expectations about the short and long-term levels of the US economy. Moreover, 5yTS and BaaCS are well known predictors of future inflation and economic activity (Estrella and Hardouvellis, 1991; Mishkin, 1990; Stock and Watson, 2003; Friedman and Kuttner, 1998), as they may contain information on agents' perceptions about the likelihood of business bankruptcy and default (Friedman and Kuttner, 1998) and about the future reactions of the Fed towards inflation and economic activity. Another advantage is that optimal predictions can be made available for a continuum of individuals since an infinite number of quantile regressions can be estimated. In this case, values assumed by coefficients $\beta(\tau)$ at each percentile $\tau \in (0,1)$ will simply define the weights each individual will put on each covariate when predicting $z_{t,t+h}$.

The macro variables were selected according to their availability in the SPF data set from 1968:Q4, the date when the survey was initiated. This means that the risk measures I propose are estimated for GDP price index, real GDP, unemployment rate, industrial production, housing starts and corporate profits after tax (see appendix 1.D for more details about the data). Despite being few in number, these six variables contain a large amount of information about business cycles and inflationary pressures, which have been shown to be closely related to movements in bond risk premia (Ludvigson and Ng, 2009). The sample ranges from 1968:Q4 to 2011:Q4. Since I will be predicting bond excess returns accumulated over the

[7]See footnote 2.

following year starting from t, when estimating risk measures, h is set equal to 4 (four quarters).

Med is obviously estimated for $\tau = 0.5$. For estimating IQR and IQS I set $\tau = 0.05$. In principle, other values of τ could be considered, but typically the case of $\tau = 0.05$ allows capturing the tails of conditional distributions of $z_{t,t+4}$, meaning that $F_{z_{t,t+4}}$ can be richly characterized through the estimation of Med, IQR and IQS only. This procedure yields a 18×1 column vector m_t of macro risks observed at time t (ex ante) for time $t+4$, where three measures are estimated for each of the six macro variables.

1.2.3 Ex ante macroeconomic risks in the US: some interesting new facts

Figures 1.2, 1.3 and 1.4 show the eighteen estimated measures of macro risks observed at time t together with $q_{z_{t,t+4}}(.05)$ and $q_{z_{t,t+4}}(.95)$ lines. NBER-dated recessions are shown as shaded bars. Notice that their time series reveal several interesting features. First, the range of ex ante conditional distributions of growth in GDP, unemployment, industrial production and housing starts, show pronounced countercyclical behavior, indicating the presence of increasing levels of uncertainties regarding these variables during bad times. Using a different approach this result is also documented by Jurado, Ludvigson and Ng (2013) and Bansal and Shaliastovich (2013). This pattern is also observed for tail risks. Although lower and upper tails show similar dynamics for most variables, risks of extreme declines in GDP, industrial production and housing starts and extreme rises in unemployment and inflation show more pronounced behavior and increase substantially during recessions periods.

It is also worth commenting on the behavior of uncertainty regarding housing starts during the recession of 2008/2009. While we see sharp increases for this variable during all previous NBER recessions, when it comes to the last recession the level of uncertainty shows consistent increases right from 2004, the year when the subprime mortgage lending rose dramatically in the US. Another result is that inflation uncertainty increases with the level of expected inflation as documented by Golob (1994), Garcia and Perron (1996) and Capistrán and Timmermann (2009), while it seems to decrease quickly during periods of economic slowdowns, when the level of expected inflation follows the same trend.

When it comes to asymmetries, notice that predicted conditional distributions for inflation and unemployment (industrial production and housing starts) growth are mostly positively (negatively) skewed, indicating the presence of consistent ex ante upside (downside) risks for these variables. This last feature is also verified in Table 1.1 - Panel A, which shows descriptive statistics for macro risks. Mean values indicate that consistent upside risks for inflation and unemployment, and downside risks for GDP, industrial production, housing starts and corporate profits are present. In addition, ex ante lower tail risks for GDP, industrial production and housing starts is more volatile (with higher standard deviation) than upper tail risks. The opposite seems to be the case for inflation, unemployment and corporate profits.

In order to have a better understanding of how ex ante risks for each of the six macro variables relate to business cycles, Figure 1.5 shows correlations between estimated risk measures and GDP growth, both measured at time t. Blue circles indicate statistically significant correlation coefficients. Observe that most ex ante risks show strong relations to GDP growth. Tail risks as well as median predictions regarding GDP and industrial production are positively related to GDP growth. The opposite seems to be the case for unemployment, housing starts and corporate profits. Uncertainty for all variables, except inflation, show strong and negative correlations to movements in GDP, strengthening my previous findings that macroeconomic uncertainty is countercyclical. Also, observe that GDP growth relates positively to ex ante downside risks for inflation, GDP and industrial production, revealing that the current level of the economy may have an effect on skewness risks for these variables. This is also true for housing starts and corporate profits, although correlations show negative signs, that is, when the economy is slowing down, ex ante upside risks for these variables tend to rise.

1.3 Predicting excess bond returns

I focus on one-year log returns on an n-year zero-coupon Treasury bond in excess of the annualized yield on a 1-year zero coupon bond. These are constructed from the Fama-Bliss discount bond yields data set for maturities up to five-years, and from the Treasury zero-coupon bond yields data set of Gürkaynak, Sack, and Wright (2007) (GSW) for maturities from six to ten years. The sample ranges from 1968:Q4 to 2011:Q4.[8] As both the SPF and the Michigan Survey reports are released by the middle of the quarter, I use yields for the end of the second month of each quarter.[9] For $t = 1,...,T$, excess returns are denoted as $rx^n_{t,t+4} = r^n_{t,t+4} - y^1_t = -(n-1)\, y^{n-1}_{t+4} + ny^n_t - y^1_t$, where $r^n_{t,t+4}$ is the one-year log holding-period return on

[8]For the period 1968Q4 - 1971Q3 yields for maturities from eight to ten years were obtained by extrapolating the Gürkaynak, Sack and Wright (2007) data set using Svensson's (1997) parametrization and the estimated parameters provided by the authors.

[9]The Michigan Survey is conducted at a monthly frequency beginning from January 1978.

an n-year bond purchased at time t and sold one year after at time $t+1$ (or $t+4$ quarters) and y_t^n is the log yield on the n-year bond.

Table 1.1 - Panel B shows descriptive statistics for the 1-year yield and the 2-year to 10-year excess bond returns. Notice that the average term structure of excess returns is positively sloped and standard deviations increase with maturities, suggesting that investors require higher premia for investing in longer (riskier) bonds. In addition, returns are negatively skewed and exhibit positive excess kurtosis. The Robust Jarque-Bera test of normality, however, does not reject the null hypothesis of normality for excess returns, which also show high persistence as indicated by the first order autocorrelation coefficients.

For predicting excess bond returns I then propose the following regression model,

$$rx_{t,t+4}^n = \alpha_0 + \alpha' m_t + \vartheta' g_t + \varepsilon_{t,t+4} \tag{1.8}$$

where α and ϑ are 18×1 vectors of coefficients, m_t is a 18×1 vector of estimated macro risks measured at time t (ex ante), three for each of the six macro variables, and g_t can include any other potential predictor such as the single forward factor of Cochrane and Piazzesi (2005) or the single macro factor of Ludvigson and Ng (2009). The risk measures I include in m_t are Med, IQR and IQS. Since IQR and IQS were both estimated using $\tau = 0.05$ they implicitly embed information about tail risks, meaning that tail risks do not necessarily need to be included in m_t.

Although regression (1.8) allows the use of all the information available on macroeconomic risks to explain variation in bond risk premia, it quickly becomes impractical since there are at least 2^{18} possible combinations of predictors to consider. Furthermore, it is highly likely that the high dimension assumed by (1.8) will deteriorate its out-of-sample forecasts (Stock and Watson, 2002a, 2002b, 2005), obfuscating any sign of out-of-sample predictability. Nevertheless, as a remedy to these problems, substantial dimensionality reduction can be achieved by extracting a few factors that summarize almost all the information about $rx_{t,t+4}^n$ contained in the panel of estimated risk measures. In this paper, I follow Stock and Watson (2002a, 2002b) and Ludvigson and Ng (2007, 2009, 2010) and use a factor model estimated by Principal Component Analysis (see Appendix 1.B for details). The initial number of factors to be estimated is set by Bai and Ng (2002) information criteria, while factors that are effectively important for predicting $rx_{t,t+4}^n$ can be optimally selected using Schwarz (1978) Bayesian information criteria (SBIC)[10].

[10]This is the procedure adopted by Ludvigson and Ng (2007, 2009, 2010). Also, Stock and Watson (2002b) point out that minimizing the SBIC yields the preferred set of factors. I also tested the Hannan and

This leads to the following regression

$$rx^n_{t,t+4} = \alpha_0 + \alpha' MRF_t + \vartheta' g_t + \varepsilon_{t,t+4} \tag{1.9}$$

where MRF_t is a vector of estimated macro risk factors and α_0 and α are parameters to be estimated by OLS[11]. The advantage of this approach is that we can summarize almost all important information about $rx^n_{t,t+4}$ contained in m_t in a few variables, MRF_t.

1.4 Empirical results

Do risks in macro fundamentals explain variation in bond risk premia?

Bai and Ng (2002) information criteria indicate that the panel of estimated macro risks is well described by eight principal components (or factors) from which three were formally chosen (using SBIC) among all the 2^8 possible specifications for $rx^n_{t,t+4} = \alpha_0 + \alpha' MRF_t + \varepsilon_{t,t+4}$. The selected factors were the first, the fourth and the sixth first principal components, forming the vector $MRF_t = (MRF_{1t}, MRF_{4t}, MRF_{6t})'$. In principle, other combinations of factors could also be used, but I focus my analysis on MRF_t since this is the combination that delivers the highest explanatory power (optimal SBIC) for $rx^n_{t,t+4}$, while I also find that this particular combination has economic meaning, as I discuss below. Following Cochrane and Piazzesi (2005) I also test whether a single linear combination of these factors has predictive power for excess returns across maturities. I define this object as the "single macro risk factor", $SMRF$, which can be constructed from a simple linear regression of average excess returns (across maturities ranging from 2-year to 10-year) on MRF_t

$$\begin{gathered} \overline{rx}_{t,t+4} = \theta_0 + \theta_1 MRF_{1t} + \theta_2 MRF_{4t} + \theta_3 MRF_{6t} + \varepsilon_{t,t+4} \\ SMRF_t = \widehat{\theta}' MRF_t \end{gathered} \tag{1.10}$$

Table 1.2 shows results with both MRF and $SMRF$ as predictors. Newey-West t-stats computed with 6 lags are shown in parentheses. The small-sample performance of statistics was also verified and 95% bootstrap confidence intervals for coefficient estimates, Wald statistics and adjusted-R^2s are provided in square brackets. Results reveal that factors have high predictive power for $rx^n_{t,t+4}$ for all maturities with R^2s ranging from 0.20 for the 2-year bond to 0.30 for the 10-year bond. Factor MRF_4 presents the highest statistical significance followed by MRF_1. MRF_6 is

Quinn (1979) (HQIC) criteria, which delivered the same set of optimal factors as SBIC.

[11] I disregard the use of hats in MRF_t to ease notation.

not significant, although it seems important for predicting $rx^n_{t,t+4}$ according to SBIC.[12] The single factor also shows high predictive power with R^2s slightly higher than *MRF* regressions. Results remain robust when we analyze the small-sample significance of estimated coefficients. Notice that MRF_1 is no longer significant for the 2-year excess return. The Wald statistic, however, remains highly significant, indicating that all factors are jointly significant, even in small samples.

Since factors are orthogonal by construction, we can characterize their relative importance in the vector MRF_t by simply investigating the absolute value of the coefficients on each factor in regression (1.10). After running (1.10) I find the following values for coefficients estimates: $\widehat{\theta}_1 = 2.128$, $\widehat{\theta}_2 = -2.264$ and $\widehat{\theta}_3 = 1.052$; revealing that the first and the forth factors are the most important predictors.

It is well known that factors do not correspond exactly to a precise economic concept. Nonetheless, it is useful to show that *MRF* capture relevant information about macro risks. I do so here by briefly characterizing macro risk factors as they relate to each of my estimated risk measures. This analysis is based on marginal R^2s obtained by regressing each of the 18 variables in m_t onto the three factors, one at a time.

Figure 1.6 displays computed R^2s as bar plots, with Panel A showing R^2s grouped by macro variables and Panel B showing R^2s grouped by risk measures. Results reveal that the first factor loads on all variables, but R^2s are higher for risks on unemployment, industrial production and GDP, that is, variables related to economic activity. The forth factor is highly related to risks on housing starts, more specifically, to downside (upside) risks, although it also manifests a strong relationship with GDP-IQR and Unemp-IQS. The sixth factor is clearly significantly related to risks associated to inflation, with Inf-IQS explaining a large portion of its variation. Notice also from Panel B that while the first factor seems to be mostly related to expectations, the forth and sixth factors are strongly related to downside (upside) risks.

Figure 1.7 plots the time series of MRF_1, MRF_4 and MRF_6 against the respective macro risk that is most related to each factor. In order to verify that the first factor is indeed a real activity risk factor I picked Unemp-Med, while MRF_4 and MRF_6 are plotted against Hous-IQS and Inf-IQS, respectively. Shaded bars indicate NBER recessions. Figure 1.7 indicates that MRF_1 is highly related to Unemp-Med, with the two series presenting a correlation of -98%. The correlation with GDP-Med is 96% and with Unemp-IQR is -90%, which indicates that MRF_1 has strong relation to risks in economic activity. MRF_4 is clearly negatively correlated with Hous-IQS

[12]The Hannan and Quinn (1979) (HQIC) criteria delivered the same set of optimal factors as SBIC.

with a coefficient of -49%. The correlations with GDP-IQR and Unemp-IQS are both 47%. Factor MRF_6 on the other hand shows strong comovement with Inf-IQS, and the correlation between the two series is 55%. These results lead us to classify MRF_4 (MRF_6) as a housing (inflation) skewness factor, although it is also possible to interpret MRF_4 as a GDP uncertainty or an unemployment skewness factor.

Beyond the median

I have provided evidence that risks in macroeconomic fundamentals derived from Med, IQR and IQS are able to explain movements in expected excess bond returns. Recent empirical evidence has shown that macroeconomic expectations obtained from survey based consensus forecasts (mean or median) are able to explain bond risk premia (Chun, 2011; Piazzesi, Salomao and Schneider, 2013; Dick, Schmeling and Schrimpf, 2013; Buraschi and Whelan, 2012). Thus, a natural question that arises is whether IQR and IQS provide information about risk premia that is not contained in simple mean or median forecasts. If so, there is strong evidence that information beyond the median is indeed important for explaining movements in bond premia.

Rather than focusing on survey consensus forecasts, I extract median forecasts by estimating regressions as (1.6) for the six macro variables, as previously done.[13] When evaluated at $\tau = 0.5$, equation (1.6) provides a measure that is similar to the median of individuals' forecasts provided by surveys. For purposes of comparison with the macro risk factors previously estimated I then estimate median factors, MeF, and a single median factor, $SMeF$, by applying PCA to the $T \times 6$ panel of estimated medians. Bai and Ng (2002) information criteria indicates that this panel is well described by three principal components from which all the three were formally chosen (using SBIC) as previously done. The single median factor was then obtained as

$$\begin{gathered} \overline{rx}_{t,t+4} = \kappa_0 + \kappa_1 MeF_{1t} + \kappa_2 MeF_{2t} + \kappa_3 MeF_{3t} + \varepsilon_{t,t+4} \\ SMeF_t = \widehat{\kappa}' MeF_t \end{gathered} \tag{1.11}$$

Table 1.3 shows the results of this exercise. As has been recently documented, conditional median forecasts represented here by SMeF show high predictive power for $rx^n_{t,t+4}$ for all maturities with R^2s ranging from 0.12 to 0.25 and highly significant estimates. SMeF loads more heavily on excess returns at longer maturities and its predictive power increases with n. However, notice that all the significance

[13] I use the conditional median instead of the conditional mean $E\left(z_{t,t+4}|\Omega_t\right) = \beta' x_t$ because of its robustness property against the conditional asymmetries existent in the data.

of SMeF switches to SMRF when the single macro risk factor is included as additional predictor. This result is somewhat expected given that SMRF embeds the information in SMeF about $rx^n_{t,t+4}$. Notice, however, that R^2s also increase substantially, indicating that IQR and IQS indeed provide additional information about bond risk premia variation.

Comparison with classical bond return predictors

Cochrane and Piazzesi (2005, 2008) show that a single factor, which they make observable through a linear combination of forward rates, captures substantial variation in expected excess returns on bonds with different maturities. Similarly, Ludvigson and Ng (2009) find that a single factor formed from a linear combination of individual macro factors has forecasting power for future excess returns, beyond the predictive power contained in forward rates. In this subsection I then compare the predictive abilities of SMRF, CP and LN factors.

As in Cochrane and Piazzesi (2008), CP was formed from the linear combination of the 1-year yield and forward rates from two to ten years,

$$\begin{gathered} \overline{rx}_{t,t+4} = \delta_0 + \delta_1 y^1_t + ... + \delta_{10} fw^{10}_t + \varepsilon_{t,t+4} \\ CP_t = \widehat{\delta}' fw_t \end{gathered} \tag{1.12}$$

where fw^n_t is the n-year forward rate defined as $fw^n_t = -(n-1)\, y^{n-1}_t + n y^n_t$.

LN was obtained as a linear combination of macro factors extracted from a large macroeconomic data set (131 variables). When forming LN I used the data set provided by Ludvigson and Ng (2010) but I set October 1968 as the starting date to enable direct comparisons with the other predictors studied in the paper.[14] Quarterly frequency was obtained by selecting observations for the second month of each quarter. LN was then constructed by running average bond returns on the best subset of macro factors estimated by Principal Component Analysis,

$$\begin{gathered} \overline{rx}_{t,t+4} = \varphi_0 + \varphi_1 F_{1t} + \varphi_2 F_{2t} + \varphi_3 F_{6t} + \varepsilon_{t,t+4} \\ LN_t = \widehat{\varphi}' F_t \end{gathered} \tag{1.13}$$

where $\widehat{\varphi}$ is a line vector of estimated parameters and F_t is a column vector of estimated macro factors, where I also disregard the use of hats to ease notation.[15]

Results are shown in Table 1.4. As documented by Cochrane and Piazzesi (2005, 2008) I find that CP captures a large portion of variation in expected excess returns

[14] The data set was downloaded from Sydney C. Ludvigson's web page: http://www.econ.nyu.edu/user/ludvigsons/.

[15] Following Ludvigson and Ng (2009) I also included F^3_{1t} in the set of macro factors.

with R^2s ranging from 0.21 to 0.32. When CP regressions are augmented with SMRF, notice that both variables reveal strong statistically significant predictive power, with R^2s increasing substantially and reaching 0.40 for the 10-year return. These results reveal that the factor I propose contains additional information about bond risk premia, despite the forward looking nature of forward rates.

The LN factor also has high explanatory power with R^2s ranging from 0.17 to 0.21 and highly significant estimates. Notice, however, that when SMRF is included as additional predictor, LN estimates decrease considerably, together with its statistical significance, while R^2s values jump substantially. As an example, R^2s increase from 0.17 to 0.38 for the 10-year return when including SMRF. The increases are quite large, especially for longer maturities, indicating the SMRF and LN capture information about bond risk premia that is somewhat independent.

I also test regressions that include all three single factors jointly. As documented by Ludvigson and Ng (2009), including LN to CP regressions increase R^2s to levels close to 0.4. Notice, however, that R^2s are even higher when augmenting regressions with SMRF, with highly significant coefficients from the 2-year maturity according to asymptotic t-stats, and from the 5-year maturity in small samples. Notice also that LN loses significance from the 3-year maturity in small-samples.

In general, results suggest that, to a large extent, SMRF captures information about expected excess bond returns that is not contained in CP and LN factors. This indicates that macroeconomic expectations, uncertainties, macroeconomic downside (upside) risks and tail risks are important determinants of bond risk premia in the US and are also able to capture information about bond risk premia that is somewhat unrelated to that contained in forward rates and current macroeconomic variables.

Are bond risk premia countercyclical?

From a theoretical point of view, Campbell and Cochrane (1999) and Wachter (2006) provide an explanation for the link between time-varying bond risk premia and the business cycle. Simply speaking, the rationale behind their argument is that investors have a slow-moving external habit, so when the economy falls into a recession, the risk of running below the minimum level of consumption increases and investors become more risk-averse, which leads risk premia to go up during bad times.

In light of this, we can gain some economic intuition of how bond premia implied by risks in macro fundamentals behave by examining its fluctuations over business cycles. More specifically, I show that movements in the single macro risk factor, a measure of the average bond risk premia across maturities, are closely connected to

NBER-dated business-cycle phases. Figure 1.8 - Panel A plots the 4-quarter moving average of SMRF. In general, we see declines in bond premium during expansions and sharp increases during recessions. Observe also that the increases in the risk premium during the 1990-1991 and 2001 recessions are somewhat more modest than those during recessions in the 80's and the recession of the late 2000's. This makes sense, since these two recessions were milder relative to the others. Overall, Figure 1.8 - Panel A demonstrates that macroeconomic risks produce a bond risk premium that closely tracks NBER business-cycle phases.

In order to complement the evidence shown in Figure 1.8 - Panel A, Panel B shows lead/lag relations between the bond premium and growth rates of three macroeconomic variables - real GDP, industrial production and unemployment rates - that are closely related to business cycles. I keep my bond premium indicator fixed at date t and then lead and lag the economic indicators. Notice that correlations turn negative and positive as macro variables are leaded/lagged. While a drop in economic activity leads an increase in bond premium, a rise in bond premium tends to lead an improvement in future economic activity. These correlations are statistically significant and demonstrate that bond premia implied by risks in macroeconomic fundamentals are closely related to movements in the real economy.

Are macro risk factors *unspanned*?

Several recent papers have considered the possibility that some factors in the term structure of interest rates are unspanned in the sense that while they are irrelevant for explaining the cross-sectional variation of current yields, they are important for forecasting future interest rates and explain variation in bond risk premia (Duffee, 2011; Joslin, Priebsch and Singleton, forthcoming; Ludvigson and Ng, 2009; Kim, 2008). As shown, macro risk factors are able to predict bond returns, but are they unspanned factors? In this section, I provide a possible answer to this question.

It is customary in the term-structure literature to summarize the information in yields using the three first principal components (PC hereafter) as they explain virtually all of the variation in the yield curve (Litterman and Scheinkman, 1991). Thus, the first evidence of the unspanning property of MRF can be provided by regressing PC and/or MRF onto yields and verifying their explanatory power. If macro risk factors are able to explain variation in current yields with levels comparable to PC, they may not be unspanned factors. Table 1.5 provides results for this exercise. While PC is able to explain about 0.99 of the variation in current yields, MRF regressions show moderate to low R^2s. Also adding MRF to PC regressions keeps R^2s unaltered, indicating that the new factors do not add any information about current yields.

Another possibility is to verify whether the new risk factors contain information about bond risk premia that is in some degree independent of that contained in the yield curve. Table 1.5 shows R^2s for regressions of PC and/or MRF onto excess returns. While PC shows predictive power with R^2s ranging from 0.07 to 0.19, regressions with PC and MRF deliver R^2s ranging from 0.30 to 0.36. In other words, to some extent, macro risk factors and the yield curve contain different information about bond risk premia.

Joslin, Priebsch and Singleton (forthcoming) also suggest examining the following spanning condition,

$$MRF_{jt} = \omega_0 + \omega' PC_t, \; j = 1,4,6 \tag{1.14}$$

by projecting each risk factor onto PC. Projections of MRF_1, MRF_4 and MRF_6 onto PC give R^2s of 0.68, 0.05 and 0.28. Augmenting the dimension of the principal components to five only raises R^2s to 0.72, 0.05 and 0.34, indicating that a large portion of variation in MRF arises from variables distinct from PC. This is especially true for MRF_4.

To sum up, there is strong evidence indicating that macro risk factors contain information about bond risk premia that is unspanned by the yield curve, suggesting that predictability of bond excess returns cannot be completely summarized by the cross-section of yields or forward rates. This result has important implications for the estimation of the term premium component of yields using affine term structure models, as many models of this class commonly disregard the information about expected excess returns contained in factors other than the yield curve (Ang and Piazzesi, 2003; Ang, Dong, and Piazzesi, 2007; Rudebusch and Wu, 2008). In fact, information in current macro variables (Wright, 2011; Joslin, Priebsch and Singleton, forthcoming; Ludvigson and Ng, 2009) and in conditional distributions of future macroeconomic outcomes, as I have shown, also need to be taken into account.

In Section 1.5 I provide results on the estimation of an affine term structure model along the lines of Joslin, Priebsch and Singleton (forthcoming) with MRF and PC as state variables, and discuss further results on the unspanning features of MRF. In general terms, estimated parameters governing expected excess returns show that shocks to MRF_4 (MRF_6) have a negative (positive) and significant impact on risk premia through the level risk. Additionally, shocks to all macro risk factors cause off-setting movements in the term premium and expected short-rate components of current long-yields, leaving them statistically unaffected.[16] These results provide further evidence that macro risk factors have a component that is unspanned by the yield curve, meaning that they are indeed able to affect term premium estimates obtained from affine term structure models.

Robustness Tests

A natural question that may arise is whether specification (1.7) is really capturing the true quantiles of $z_{t,t+4}$. In order to verify this I apply the backtest of Gaglianone et al (2011) (GLLS hereafter) who propose a framework to evaluate the performance of a VaR model through quantile regression methods. While the most common backtests are based on simple hit indicators that signal whether a particular threshold was exceeded, the GLLS backtest allows identifying to what extent a VaR model indicates increases in risk exposure, which is a key issue to any risk model. The authors also show through Monte-Carlo simulations that the GLLS backtest shows increased finite sample power in comparison to the most common backtests existent in the literature. The GLLS test is implemented through the estimation of the following quantile regression

$$q_{z_{t,t+4}}(\tau) = \phi_0(\tau) + \phi_1(\tau)\widehat{q_{z_{t,t+4}}}(\tau)$$

where the null hypothesis of correct specification of the quantile model at level τ is given by $H_0 : (\phi_0(\tau), \phi_1(\tau)) = (0,1)$. H_0 can be tested through the VQR test statistic proposed by GLLS, which follows a chi-square distribution with 2 degrees of freedom. If the model is correctly specified H_0 is not rejected implying that $q_{z_{t,t+4}}(\tau) = \widehat{q_{z_{t,t+4}}}(\tau)$.

Table 1.6 shows p-values of the GLLS test for several percentiles. Since Med, IQR and IQS were obtained from quantile functions estimated for $\tau =$ 0.05, 0.5, 0.95, these are the most important test results, but the implementation of the test for other typical values reveals that my quantile specification is well identified at other percentiles as well. Results in Table 1.6 show that specification (1.7) produces conditional quantiles forecasts, $\widehat{q_{z_{t,t+4}}}(\tau)$, that are statistically indistinguishable from the true conditional quantiles of $z_{t,t+4}$, which suggests that my risk measures are being precisely estimated and also that (1.7) is able to accurately capture the conditional distributions of $z_{t,t+4}$.

The most natural question, however, is whether the high predictability I have found is coming exclusively from variables in the vector x_t. If this is the case,

[16] Duffee (2011) points out that factors whose impacts on term premium and short-rate expectations cancel each other out may be considered unspanned.

there is no reason to add the complexity of estimating measures of ex ante macro risks and then using them to forecast excess bond returns. A suitable test for this issue is provided by simply checking the predictive power of SMRF when controlling directly for the information in x_t. In order to guard against the possibility of overfitting, the information in predictors x_t can be summarized by estimating predictor factors, Fx, and a single predictors factor, SFx, by applying PCA to the $T \times 9$ panel of predictors formed by the six consensus forecasts, $z_t^{SPF,4}$, and MCEI, 5yTS and BaaCS. Bai and Ng (2002) indicates that this panel is well described by seven factors from which three (first, third and fifth principal components) were formally chosen using SBIC. These three factors form the vector Fx. Factor SFx is a linear combination of Fx and provides a variable that can directly be used to control for the information in quantile predictors x_t. Results of this exercise are provided in the Appendix 1.E. In general, they show that although SFx contains high predictive power for $rx_{t,t+4}^n$, adding SMRF to regressions increases R^2s substantially to levels almost identical to the ones shown by Table 1.2, with statistical significance shifting to SMRF. These results suggest that the high predictability found is largerly due to the extra information obtained from the estimation of Med, IQR and IQS.

Another convenient robustness test I make available is the assessment of the predictive power of macro risk factors (MRF and SMRF) when risk measures are estimated using alternative approaches. To guard against the possibility of inadequacy of quantile regressions estimated at $\tau = 0.05$, 0.95 Appendix 1.F shows results using two alternative estimation procedures. First, I assess the predictive power of ex ante macroeconomic risks when risk measures are estimated for $\tau = 0.10$. While the use of $\tau = 0.10$ somewhat misses some information about tail risks, it places less weight on extreme data points, guarding against possible instabilities of quantile regressions estimated at tails. In the second procedure, risk measures are estimated for $\tau = 0.05$ using the Wang, Li and He (2012) approach, which integrates quantile regression with Extreme Value Theory and is suitable for quantile curves at tails. Their procedure is explained in details in the Appendix 1.F. Results show that the statistical significance of macro risk factors (MRF and SMRF) and the magnitudes of their predictive power remain very high with levels comparable to those shown in tables 1.2, 1.3 and 1.4, which corroborate my previous findings.

1.5 An affine term structural model with macro risk factors

Motivated by my findings that bond risk premia is driven by risks in macro fundamentals, in this section I examine time variation in term premia. From a monetary policy perspective, understanding time variation in term premia is important as term premia obfuscates the relationship between short-term interest rates controlled by central banks and long-term interest rates, while it also makes it difficult to measure expectations of future short-term rates using the yield curve. I do so by estimating a Gaussian affine term structure model along the lines of Joslin, Priebsch and Singleton (forthcoming) (JPS hereafter), where state variables are composed by the first three principal components of yields and the three macro risk factors, MRF_1, MRF_4 and MRF_6. I treat the macro risk factors as unspanned in the state vector (Duffee, 2011; Joslin, Priebsch and Singleton, forthcoming), as I have found that they explain variation in risk premia, but are irrelevant for explaining the cross-sectional variation in current yields.

Model specification

Following the macro-finance literature since Ang and Piazzesi (2003), I assume that the $p \times 1$ vector of state variables X_t follows a VAR(1) process under the objective probability measure $\mathbb{P}$,

$$X_{t+1} = \mu + \Phi X_t + \Sigma \varepsilon_{t+1} \tag{1.15}$$

where $\varepsilon_t \sim iid\, N(0, I_p)$, the state vector consists of the first three principal components of yields and MRF_1, MRF_4 and MRF_6, and Σ is an $p \times p$ lower triangular matrix. The pricing kernel is assumed to be conditionally lognormal

$$M_{t+1} = exp\left(-r_t - \frac{1}{2}\lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1}\right) \tag{1.16}$$

where $r_t = \delta_0 + \delta_1' X_t$ is the three-month interest rate and the $p \times 1$ vector of risk prices is affine in state variables, $\lambda_t = \lambda_0 + \lambda_1 X_t$. Under the risk-neutral measure $\mathbb{Q}$ the state vector follows the dynamics,

$$X_{t+1} = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_t + \Sigma \varepsilon_{t+1} \tag{1.17}$$

where $\mu^{\mathbb{Q}} = \mu - \Sigma\lambda_0$ and $\Phi^{\mathbb{Q}} = \Phi - \Sigma\lambda_1$.

It then follows that under no-arbitrage bond prices are exponential affine functions of the state variables, $P_t^n = exp\left(A_n + B_n^{'} X_t\right)$, where A_n is a scalar and B_n is an $p \times 1$ vector that satisfy the recursions

$$\begin{aligned} A_{n+1} &= -\delta_0 + A_n + B_n^{'} \mu^{\mathbb{Q}} + \tfrac{1}{2} B_n^{'} \Sigma \Sigma^{'} B_n \\ B_{n+1} &= \Phi^{\mathbb{Q}'} B_n - \delta_1 \end{aligned} \tag{1.18}$$

which start from $A_1 = -\delta_0$ and $B_1 = -\delta_1$.

I treat the macro risk factors as unspanned factors. To illustrate, partition X_t as $(X_{1t}^{'}, X_{2t}^{'})^{'}$, where X_{1t} and X_{2t} are $p_1 \times 1$ and $p_2 \times 1$ vectors consisting of the first three principal components of yields and MRF_1, MRF_4 and MRF_6, respectively. Set also the last p_2 elements of δ_1 and the upper-right $p_1 \times p_2$ block of $\Phi^{\mathbb{Q}}$ to be equal to zero. Then the last p_2 elements of B_n will be equal to zero and bond prices reduce to

$$P_t^n = exp\left(A_n + B_{1n}^{'} X_{1t}\right) \tag{1.19}$$

where B_{1n} consists of the first p_1 elements of B_n. The result of this is that factors in X_{2t} are important for forecasting future yields, but only factors in X_{1t} are important for pricing bonds at time t. Model implied yields are then computed as $y_t^n = -n^{-1} log\, P_t^n = -n^{-1}(A_n + B_{1n}^{'} X_{1t})$.

Estimation

The estimation approach follows JPS and Joslin, Singleton and Zhu (2011) with parameters being estimated by MLE. Due to the separation result of the likelihood function derived in Joslin, Singleton and Zhu (2011) parameters in μ and Φ are estimated separately from those governing the risk neutral pricing of bonds, which can be done by a simple OLS. For estimating the remaining identified parameters, i.e. Σ, $\mu_1^{\mathbb{Q}}$, $\Phi_{11}^{\mathbb{Q}}$, δ_0 and $\delta_{1,1}$, where $\mu_1^{\mathbb{Q}}$ and $\delta_{1,1}$ are the $p_1 \times 1$ vectors of $\mu^{\mathbb{Q}}$ and δ_1, and $\Phi_{11}^{\mathbb{Q}}$ is the upper-left $p_1 \times p_1$ block of $\Phi^{\mathbb{Q}}$, it is assumed that observed yields are equal to the model-implied yields plus i.i.d. Gaussian measurement errors. As in JPS, the model is first reparameterized in terms of a $p_1 \times 1$ latent state vector $\mathscr{S}$ that follows a VAR with zero intercept and diagonal slope coefficient matrix equal to $I_1 + \Lambda^{\mathbb{Q}}$, and an equation for the short-rate that assumes the form $r_t = r_\infty^{\mathbb{Q}} + 1\mathscr{S}_t$, where 1 is a line vector of ones. The likelihood is then maximized with respect to these parameters and original parameters' estimates can be retrieved from $r_\infty^{\mathbb{Q}}$ and $\Lambda^{\mathbb{Q}}$ as in JPS.[17] In the estimation I use the three-month interest rate and yields from one to ten years. In order to correct for the possibility of existence of small-sample bias in VAR parameters (Duffee and Stanton, 2004; Kim and Wright, 2005; Kim

[17] Appendix B in Joslin, Priebsch and Singleton (2010) specifies how to do this.

and Orphanides, 2005) I use the bootstrap approach following Bauer, Rudebusch and Wu (2012).[18]

The model shows a good fit. Fitting errors measured in terms of MSE are small and equal to 0.0039, indicating that the first three principal components together are able to account for almot all the cross-sectional variation in yields and that no other factor is required for this purpose.

It is also worth computing a Wald statistic testing the hypothesis that macro risk factors do not enter matrix Φ in (1.15). The hypothesis is highly rejected with Wald statistic equal to 2065.9 and p-value virtually equal to zero, indicating that MRF_1, MRF_4 and MRF_6 do help to predict future interest rates, and motivating their inclusion in the state vector under the $\mathbb{P}$ measure.

Table 1.7 shows the full set of parameter estimates for the affine model, including the ones governing the expected excess returns, λ_0 and λ_1. Bootstraped standard errors are in parentheses.[19] The prices of level (PC1) and slope (PC2) risks have a significant negative constant component, implying that investors on average require positive expected excess returns for holding the level and slope portfolios. In addition to level and slope risks being nonzero unconditionally, I find that the level risk varies significantly as a function of MRF_4 and MRF_6. The loading on these factors have positive and negative coefficients, respectively, implying that shocks to MRF_4 (MRF_6) have a negative (positive) impact on risk premia. Another finding is that the slope carries a significant price of risk. The level factor, the slope factor itself, as well as the curvature (PC3) factor all significantly affect the price of slope risk over time. Coefficients on the level and curvature factors are positive, indicating that expected excess returns on the slope portfolio is decreasing in the level and curvature of yields. Contrary, the coefficient on the slope factor shows a negative coefficient.

Decomposing long term yields

Long term yields can be represented as the sum of future nominal short-rate expectations plus a term premium defined as the average of risk premia of declining

[18]I also test the indirect inference approach proposed by Bauer, Rudebusch and Wu (2012), which is supposed to remove higher-order bias. I found that the two methods deliver almost the same results in the particular case of this paper. I use the bootstrap approach due to its simplicity and faster computation.

[19]The method used to bootstrap standard errors is as follows. First I resample the OLS residuals in the state equation and randomly choose a starting value among the T observations to construct a bootstrap sample for state variables using the original state equation parameters. Then, using the maximum likelihood estimates of the parameters, I simulate a path of the term structure for the whole sample and estimate the model based on these simulated data. These steps are repeated 1000 times delivering empirical probability distributions for all parameters from which bootstrap standard deviations can be easily computed.

maturities.[20] After estimating the parameters for the affine model, the n-year term premium can be computed as the difference between the n-year implied yield under $\mathbb{Q}$ and the average of expected short-rate up to year n under the $\mathbb{P}$ measure.

This decomposition is ilustrated by Figure 1.9 for the 10-year yield. Two aspects are noteworthy. First, long-term term premium implied by risks in macro fundamentals have a marked countercyclical behaviour showing declines during expansions and increases during recessions. Observe that increases have been more pronounced since the early 90's, even though the 10-year yield has shown a decreasing pattern since then. Second, these movements have been closely followed by decreases in short-rate expectations, indicating that the two components seem to move in opposite directions, in particular, during bad times. As an example of this, notice that while the term premium raised sharply during the recession of the late 2000's, short-rate expectations declined abruptly to levels close to zero. The correlation between the two series is high and negative: -51%.

Consistent with previous findings concerning SMRF (the return risk premia), term premium is highly countercyclical. Figure 1.9 - Panel B shows lead/lag relations between the term premium and growth rates for real GDP, industrial production and unemployment rates. Contemporaneous correlations with real GDP, industrial production and unemployment growth rates are -32%, -27% and 32%, respectively, and are highly statistically significant. In addition, cross-correlations turn negative and positive as macro variables are leaded/lagged, indicating that bond premia implied by risks in macroeconomic fundamentals are closely related to movements in the real economy, as suggested by theory (Campbell and Cochrane, 1999; Wachter, 2006; Bansal and Yaron, 2004; Rudebusch and Swanson, 2009).

Impulse response analysis

Duffee (2011) points out that factors whose impacts on term premium and short-rate expectations cancel each other may be considered unspanned, as shocks have no impact on current bond yields. As Figure 1.9 suggests, term premium and short-rate expectations implied by expected macro risks move in opposite directions. It is then worth verifying how shocks to macro risk factors affect these two terms separately and, consequently, yields.

Figure 1.10 shows impulse response functions (IRFs) for the term premium, short-rate expectations and the 10-year yield to one-standard deviation shocks in MRF_1, MRF_4 and MRF_6. Notice that shocks in all macro factors cause off-setting

[20] y_t^n can be decomposed as $y_t^n = \frac{1}{n}E_t(r_t + r_{t+1} + ... + r_{t+n-1}) + \frac{1}{n}[E_t(rx_{t+1}^n) + E_t(rx_{t+2}^{n-1}) + ... + E_t(rx_{t+n-1}^2)]$.

movements in the term premium and expected short-term interest rates, leaving current yields statistically unaffected. While shocks in MRF_1 and MRF_4 (MRF_6) drive term premium down (up), they bring the expectations component up (down). Following Duffee (2011), these results provide even stronger evidence that macro risk factors are indeed unspanned by the yield curve.

It is also worth interpreting the impulse response functions shown by Figure 1.10. Notice that a positive shock in MRF_1 drives term premium down by about 10 basis points after which it gradually reverts back. This is consistent with the notion that higher expected economic activity leads to lower risk premium. A similar but stronger effect is observed for MRF_4 with term premium decreasing by about 33 basis points. A shock in MRF_6, on the other hand, causes an increase in term premium of about 20 basis points, which is consistent with the idea that higher upside inflation risks raise risk premium for long-term bonds as investors will demand a higher premium to compensate for inflation risk. The magnitude of the impacts over the expectations component are of lower magnitude and mostly statistically insignificant.

1.6 Conclusions

I provide evidence that risks in macroeconomic fundamentals revealed from predicted distributions of *future* macroeconomic outcomes contain valuable information about bond risk premia. I extract factors, referred to as macro risk factors, from quantile-based risk measures estimated for variables closely related to business cycles and find that they predict excess bond returns across maturities with R^2s ranging from 20% to 30%.

From macro risk factors I construct a measure of variation in bond risk premia, referred to as “single macro risk factor”. The new single factor is highly countercyclical and predicts future excess bond returns with power above and beyond that of the Cochrane-Piazzesi and Ludvigson-Ng factors. These results provide evidence that risks in macroeconomic fundamentals explain a large portion of variation in bond risk premia, with information that is, to a large extent, unrelated to that contained in forward rates and current macro variables.

In addition, I document that macro risk factors capture unspanned predictability in bond excess returns, and discuss its practical implications for the identification and estimation of the term premium component in long-term yields using affine term structure models. Accordingly, the estimation of an affine model with unspanned macro risk factors reveals that they carry significant prices of risk and generate time-varying and countercyclical term premium. An impulse response analysis allows reaching similar conclusions. Shocks in all macro risk factors are found to cause off-setting movements in term premium and expected short-term interest rates, leaving current yields statistically unaffected.

To sum up, this study provides further support for the idea that predictability of bond excess returns cannot be completely summarized by the yield curve. Risks in macroeconomic fundamentals should also be taken into account.

Table 1.1: Descriptive statistics

Notes: Panel A shows summary statistics for the estimated ex ante macroeconomic risks. Panel B shows summary statistics for the 1-year yields and 2-year to 10-year excess bond returns. The statistics reported are the mean, standard deviation, skewness, excess kurtosis, the p-value of a Robust Jarque-Bera (RJB) test for normality and the 1st and 4th sample autocorrelations. Critical values for the RJB test were obtained empirically through 4000 Monte-Carlo simulations. Mean values are reported in percentage point basis.

Panel A

	$infl$		gdp		$unemp$	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
q(.05)	2.224	1.616	−0.894	2.032	−0.129	0.085
Med	3.523	1.941	2.737	1.298	−0.003	0.097
q(.95)	6.090	2.213	5.383	1.361	0.236	0.204
IQR	3.866	0.954	6.278	1.476	0.366	0.125
IQS	0.302	0.215	−0.151	0.110	0.262	0.203

	ip		hs		$cprof$	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
q(.05)	−5.774	5.725	−32.515	20.446	−13.654	5.800
Med	2.751	2.202	1.176	12.588	9.063	8.378
q(.95)	6.942	2.323	22.600	19.649	30.116	8.699
IQR	12.717	3.711	55.115	24.351	43.770	3.889
IQS	−0.286	0.228	−0.248	0.129	−0.040	0.283

Panel B

	Mean	Std. Dev.	Skewness	Exc. Kurtosis	pv-RJB	ρ_1	ρ_4
y^1	5.840	2.918	0.378	0.434	0.014	0.936	0.794
rx^2	0.595	1.714	−0.244	0.331	0.214	0.754	0.202
rx^3	1.038	3.121	−0.277	0.324	0.233	0.749	0.151
rx^4	1.424	4.324	−0.260	0.385	0.284	0.756	0.138
rx^5	1.548	5.232	−0.183	0.171	0.575	0.742	0.099
rx^6	1.964	6.287	−0.157	0.300	0.624	0.752	0.077
rx^7	2.031	7.144	−0.135	0.486	0.459	0.747	0.056
rx^8	2.168	8.028	−0.115	0.615	0.264	0.746	0.038
rx^9	2.273	8.900	−0.093	0.738	0.149	0.746	0.023
rx^{10}	2.354	9.765	−0.068	0.847	0.076	0.745	0.010

Table 1.2: Predictive power of macro risk factors

Notes: This table shows the predictive power of MRF and SMRF. t-stats computed using Newey-West standard errors with six lags are reported in parentheses and $\overline{R}^2$ refers to the adjusted-R^2. Wald statistics were also computed using Newey-West variance-covariance matrices with six lags. 95% confidence intervals for estimated coefficients and $\overline{R}^2$s, and p-values for Wald statistics are reported in square brackets. These were obtained through a residual-based block bootstrap with 4999 replications and overlapping blocks of size equal to six. Confidence intervals for coefficients were obtained using an asymptotic refinement based on the t-stat *(percentile-t method)* with bootstrapped t-stats computed using Newey-West standard errors with six lags.

	MRF_1	MRF_4	MRF_6	$SMRF$	$\overline{R}^2$	Wald
	0.413	−0.667	0.198		0.210	0.000
	(2.287)	(−4.020)	(0.867)		–	–
	[−0.01; 0.84]	[−1.05; −0.28]	[−0.36; 0.75]		[0.04; 0.39]	[0.010]
rx^2				0.241	0.210	
				(4.664)	–	
				[0.12; 0.36]	[0.01; 0.35]	
	0.811	−1.209	0.488		0.230	0.000
	(2.463)	(−4.676)	(1.226)		–	–
	[0.06; 1.55]	[−1.83; −0.61]	[−0.43; 1.39]		[0.04; 0.39]	[0.005]
rx^3				0.461	0.233	
				(5.063)	–	
				[0.25; 0.67]	[0.02; 0.36]	
	1.664	−1.986	1.006		0.271	0.000
	(3.019)	(−5.012)	(1.580)		–	–
	[0.41; 3.01]	[−2.93; −1.04]	[−0.51; 2.51]		[0.06; 0.41]	[0.000]
rx^5				0.843	0.278	
				(6.078)	–	
				[0.52; 1.16]	[0.04; 0.40]	
	2.554	−2.681	1.303		0.291	0.000
	(3.490)	(−4.918)	(1.571)		–	–
	[0.89; 4.20]	[−3.94; −1.39]	[−0.58; 3.17]		[0.07; 0.43]	[0.000]
rx^7				1.194	0.299	
				(6.470)	–	
				[0.78; 1.61]	[0.05; 0.43]	
	3.884	−3.534	1.620		0.306	0.000
	(3.891)	(−4.810)	(1.526)		–	–
	[1.52; 6.22]	[−5.19; −1.88]	[−0.81; 4.02]		[0.08; 0.45]	[0.000]
rx^{10}				1.667	0.313	
				(6.423)	–	
				[1.08; 2.22]	[0.06; 0.45]	

Table 1.3: Predictive power of SMRF and SMeF

Notes: This table shows the predictive power of SMRF and SMeF. t-stats computed using Newey-West standard errors with six lags are reported in parentheses and $\overline{R}^2$ refers to the adjusted-R^2. 95% confidence intervals for estimated coefficients and $\overline{R}^2$s obtained through a residual-based block bootstrap as detailed by Table 1.2 (Notes) and Appendix 1.C are reported in square brackets.

	SMRF	*SMeF*	$\overline{R}^2$
rx^2		0.212	0.125
		(3.230)	–
		[0.06; 0.36]	[0.00; 0.27]
	0.245	−0.006	0.205
	(3.386)	(−0.080)	–
	[0.08; 0.41]	[−0.18; 0.18]	[0.02; 0.36]
rx^3		0.426	0.152
		(3.591)	–
		[0.16; 0.71]	[0.00; 0.29]
	0.437	0.036	0.229
	(3.557)	(0.253)	–
	[0.16; 0.73]	[−0.29; 0.37]	[0.03; 0.38]
rx^5		0.819	0.200
		(4.508)	–
		[0.40; 1.24]	[0.02; 0.33]
	0.732	0.165	0.277
	(4.047)	(0.783)	–
	[0.33; 1.16]	[−0.31; 0.63]	[0.05; 0.41]
rx^7		1.203	0.231
		(5.134)	–
		[0.68; 1.74]	[0.04; 0.35]
	0.965	0.342	0.303
	(4.031)	(1.203)	–
	[0.45; 1.50]	[−0.30; 0.98]	[0.07; 0.43]
rx^{10}		1.723	0.254
		(5.301)	–
		[1.01; 2.46]	[0.05; 0.38]
	1.275	0.585	0.320
	(4.058)	(1.565)	–
	[0.59; 1.98]	[−0.24; 1.40]	[0.08; 0.45]

Table 1.4: Predictive power of SMRF, CP and LN factors

Notes: This table shows the predictive power of the CP, LN and SMRF factors. t-stats computed using Newey-West standard errors with six truncation lags are reported in parentheses and $\overline{R}^2$ refers to the adjusted-R^2. 95% confidence intervals for estimated coefficients and $\overline{R}^2$ s obtained through a residual-based block bootstrap as detailed in Table 1.2 (Notes) and Appendix 1.C are reported in square brackets. Regressions in which LN is included in the set of predictors are estimated using the sample 1968Q4 - 2007Q4.

	SMRF	*CP*	$\overline{R}^2$	*SMRF*	*LN*	$\overline{R}^2$	*SMRF*	*CP*	*LN*	$\overline{R}^2$
rx^2		0.242	0.214		0.304	0.213		0.195	0.190	0.354
		(5.204)	–		(4.957)	–		(3.897)	(2.704)	–
		[0.15;0.34]	[0.08;0.37]		[0.16;0.44]	[0.05;0.36]		[0.08;0.31]	[0.03;0.35]	[0.19;0.50]
	0.152	0.153	0.271	0.192	0.187	0.333	0.111	0.139	0.155	0.383
	(2.924)	(3.395)	–	(4.245)	(3.098)	–	(2.146)	(2.559)	(2.276)	–
	[0.03;0.27]	[0.06;0.25]	[0.13;0.63]	[0.08;0.30]	[0.05;0.33]	[0.07;0.45]	[−0.02;0.23]	[0.01;0.27]	[0.00;0.032]	[0.13;0.67]
rx^3		0.468	0.242		0.545	0.208		0.384	0.321	0.375
		(5.338)	–		(5.104)	–		(4.353)	(2.770)	–
		[0.28;0.65]	[0.10;0.39]		[0.31;0.77]	[0.05;0.35]		[0.18;0.58]	[0.06;0.59]	[0.22;0.52]
	0.288	0.300	0.303	0.371	0.318	0.346	0.210	0.277	0.255	0.406
	(2.984)	(3.407)	–	(4.409)	(2.870)	–	(2.111)	(2.901)	(2.211)	–
	[0.07;0.49]	[0.12;0.49]	[0.14;0.64]	[0.17;0.57]	[0.05;0.58]	[0.08;0.46]	[−0.03;0.45]	[0.04;0.50]	[−0.02;0.53]	[0.15;0.67]
rx^5		0.814	0.260		0.863	0.189		0.692	0.459	0.384
		(5.159)	–		(5.595)	–		(4.299)	(2.989)	–
		[0.48;1.13]	[0.12;0.42]		[0.53;1.21]	[0.04;0.33]		[0.31;1.07]	[0.10;0.80]	[0.23;0.52]
	0.562	0.486	0.344	0.693	0.438	0.363	0.415	0.479	0.329	0.428
	(3.970)	(3.355)	–	(4.827)	(2.862)	–	(2.723)	(3.051)	(2.138)	–
	[0.24;0.88]	[0.18;0.79]	[0.15;0.64]	[0.36;1.03]	[0.09;0.81]	[0.10;0.48]	[0.05;0.78]	[0.09;0.86]	[−0.05;0.72]	[0.16;0.67]
rx^7		1.185	0.296		1.182	0.188		1.004	0.596	0.406
		(5.669)	–		(5.579)	–		(4.506)	(3.126)	–
		[0.74;1.62]	[0.15;0.45]		[0.73;1.63]	[0.03;0.33]		[0.48;1.52]	[0.17;1.02]	[0.25;0.55]
	0.768	0.737	0.380	0.986	0.577	0.375	0.575	0.709	0.416	0.450
	(3.905)	(3.662)	–	(4.753)	(2.606)	–	(2.635)	(3.227)	(2.138)	–
	[0.34;1.21]	[0.32;1.16]	[0.17;0.65]	[0.51;1.46]	[0.09;1.06]	[0.11;0.49]	[0.06;1.11]	[0.21;1.26]	[−0.04;0.89]	[0.17;0.68]
rx^{10}		1.686	0.321		1.559	0.172		1.445	0.715	0.411
		(6.018)	–		(5.012)	–		(4.768)	(2.918)	–
		[1.08;2.25]	[0.17;0.48]		[0.90;2.20]	[0.03;0.30]		[0.77;2.12]	[0.16;1.27]	[0.25;0.57]
	1.045	1.076	0.403	1.428	0.683	0.379	0.839	1.015	0.453	0.461
	(3.784)	(4.052)	–	(4.773)	(2.091)	–	(2.712)	(3.532)	(1.759)	–
	[0.46;1.64]	[0.53;1.62]	[0.18;0.66]	[0.75;2.12]	[−0.02;1.41]	[0.12;0.50]	[0.11;1.58]	[0.32;1.70]	[−0.14;1.04]	[0.18;0.68]

Table 1.5: Evidence of MRFs as unspanned factors

Notes: This table shows the predictive/explanatory power of $PC_t = (PC_{1t}, PC_{2t}, PC_{3t})'$ and $MRF_t = (MRF_{1t}, MRF_{4t}, MRF_{6t})'$ for $rx^n_{t,t+4}$ and y^n_t. Only $\overline{R}^2$s are provided.

	PC	MRF	$PC+MRF$		PC	MRF	$PC+MRF$
rx^2	0.076	0.210	0.303	y^2	0.999	0.325	0.999
rx^3	0.076	0.233	0.303	y^3	0.999	0.305	0.999
rx^5	0.117	0.278	0.324	y^5	0.999	0.289	0.999
rx^7	0.155	0.299	0.347	y^7	0.999	0.287	0.999
rx^{10}	0.191	0.313	0.364	y^{10}	0.999	0.283	0.999

Table 1.6: GLLS test of quantile model performance

Notes: This table shows p-values for the GLLS test of quantile model performance with specification $q_{z_{t,t+h}}(\tau) = \beta(\tau)' x_t$, $x_t' = (1, z_t^{SPF,h}, Mich\,Expect_t, 5 - year\,term\,spread_t, Baa\,corp\,spread_t)$. The test is implemented for percentiles $\tau = 0.05, 0.2, 0.35, 0.5, 0.65, 0.8, 0.95$.

	Macro Variables (z)					
tau (τ)	$infl$	gdp	$unemp$	ip	hs	$cprof$
0.05	1.0	1.0	1.0	1.0	0.62	1.0
0.20	1.0	1.0	1.0	1.0	1.0	1.0
0.35	1.0	1.0	1.0	1.0	0.54	1.0
0.50	1.0	1.0	1.0	1.0	1.0	1.0
0.65	1.0	1.0	1.0	1.0	1.0	1.0
0.80	1.0	1.0	1.0	1.0	1.0	1.0
0.95	1.0	1.0	1.0	1.0	1.0	1.0

Table 1.7: Affine model - parameter estimates

Notes: This table shows the full parameters estimates for the affine model with macro risk factors MRF1, MRF4 and MRF6. Bootstrap standard errors are shown in parentheses.

μ'	0.497	−0.245	0.231	−0.535	0.169	0.297	$\mu^{\mathbb{Q}'}$	2.398	0.757	0.079			
	(0.988)	(0.310)	(0.127)	(0.197)	(0.274)	(0.265)		(0.880)	(0.222)	(0.223)			
Φ	0.957	0.055	0.261	−0.205	0.339	−0.538	$\Phi^{\mathbb{Q}}$	0.991	0.242	−0.348			
	(0.035)	(0.242)	(0.623)	(0.352)	(0.214)	(0.217)		(0.010)	(0.016)	(0.024)			
	0.015	0.752	0.843	0.194	−0.059	0.081		−0.008	0.888	0.514			
	(0.010)	(0.074)	(0.192)	(0.104)	(0.065)	(0.066)		(0.005)	(0.025)	(0.016)			
	−0.001	0.042	0.592	−0.069	0.006	−0.001		0.004	0.009	0.805			
	(0.004)	(0.031)	(0.080)	(0.044)	(0.027)	(0.028)		(0.002)	(0.004)	(0.023)			
	0.002	0.113	0.157	0.808	−0.034	0.051							
	(0.007)	(0.048)	(0.125)	(0.070)	(0.041)	(0.045)							
	0.005	0.021	−0.396	0.075	0.833	0.077							
	(0.009)	(0.062)	(0.166)	(0.090)	(0.056)	(0.057)							
	−0.018	−0.114	0.516	0.179	−0.090	0.727							
	(0.009)	(0.065)	(0.166)	(0.094)	(0.055)	(0.061)							
Σ	4.204	−1.012	0.026	−0.288	0.184	0.175	δ_0	0.016					
	(2.146)	(0.833)	(0.090)	(0.246)	(0.295)	(0.264)		(0.000)					
	−1.012	0.974	−0.237	0.414	−0.389	−0.075	δ_1'	0.083	−0.155	0.144			
	(0.833)	(0.315)	(0.035)	(0.099)	(0.105)	(0.096)		(0.007)	(0.008)	(0.008)			
	0.026	−0.237	0.115	−0.136	0.140	0.007	λ_0'	−0.927	−1.708	−1.461			
	(0.090)	(0.035)	(0.020)	(0.036)	(0.039)	(0.038)		(0.440)	(0.714)	(1.059)			
	−0.288	0.414	−0.136	0.313	−0.193	−0.079	λ_1	−0.017	−0.091	0.297	−0.100	0.165	−0.262
	(0.246)	(0.099)	(0.036)	(0.660)	(0.347)	(0.210)		(0.011)	(0.066)	(0.216)	(0.080)	(0.075)	(0.078)
	0.184	−0.389	0.140	−0.193	0.436	0.022		0.017	−0.212	0.557	0.017	0.026	−0.057
	(0.295)	(0.105)	(0.039)	(0.347)	(0.987)	(0.352)		(0.007)	(0.074)	(0.225)	(0.076)	(0.071)	(0.070)
	0.175	−0.075	0.007	−0.079	0.022	0.342		0.000	−0.116	−0.324	−0.106	0.056	−0.065
	(0.263)	(0.096)	(0.038)	(0.210)	(0.352)	(0.532)		(0.013)	(0.092)	(0.325)	(0.116)	(0.114)	(0.111)

Figure 1.1: Predicted conditional distributions

Notes: charts show predicted conditional distributions for inflation and growth in the real GDP, unemployment, industrial production, housing starts and corporate profits using quantile models estimated for $\tau = 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95$. The prediction horizon is set equal to four ($h = 4$) and the predictors used are $x_t' = \left(1,\, z_t^{SPF,h},\, Mich\,Expect_t,\, 5-year\,term\,spread_t,\, Baa\,corp\,spread_t\right)$. Red lines indicate the predicted median and blue lines indicate the realized values.

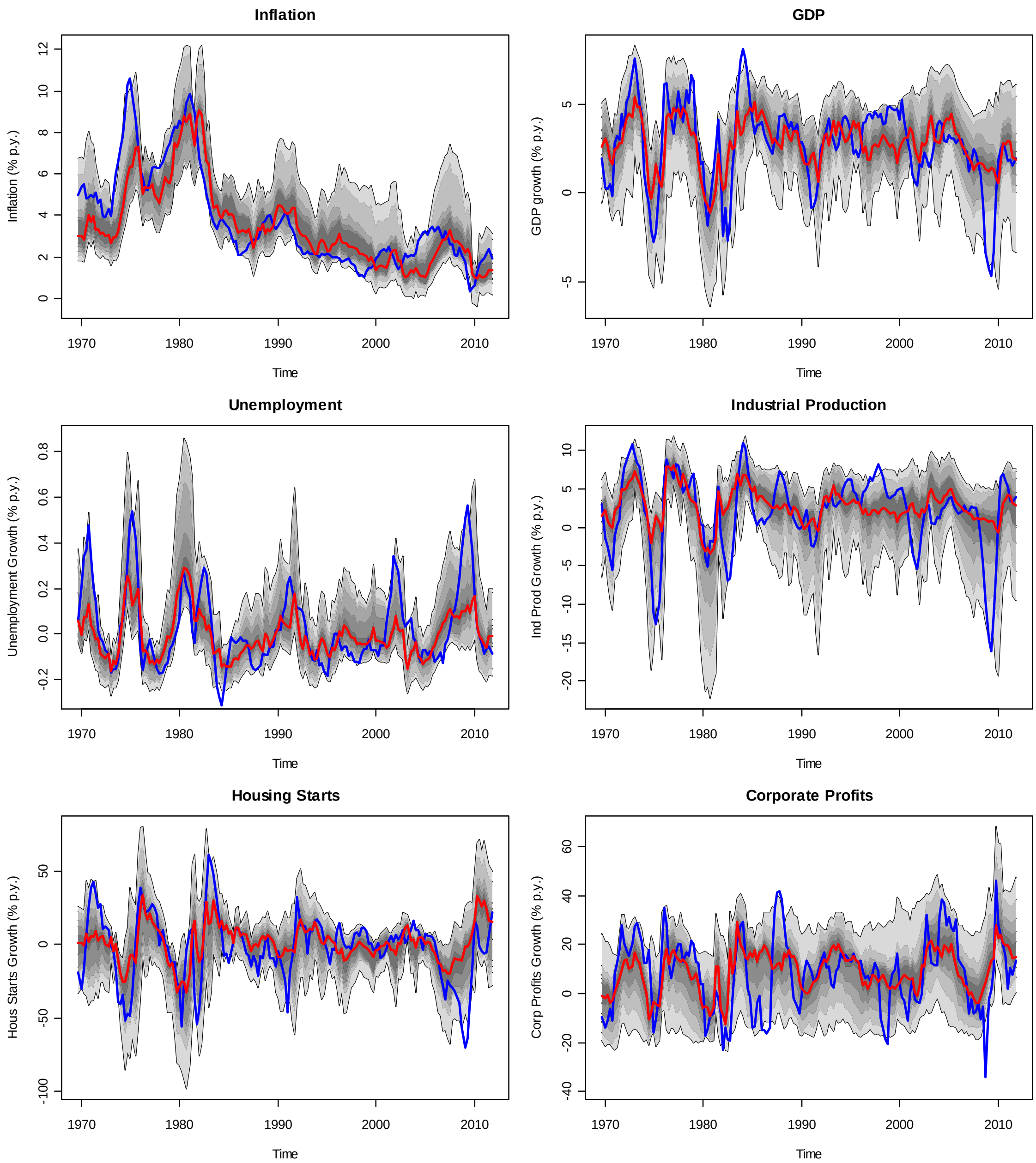
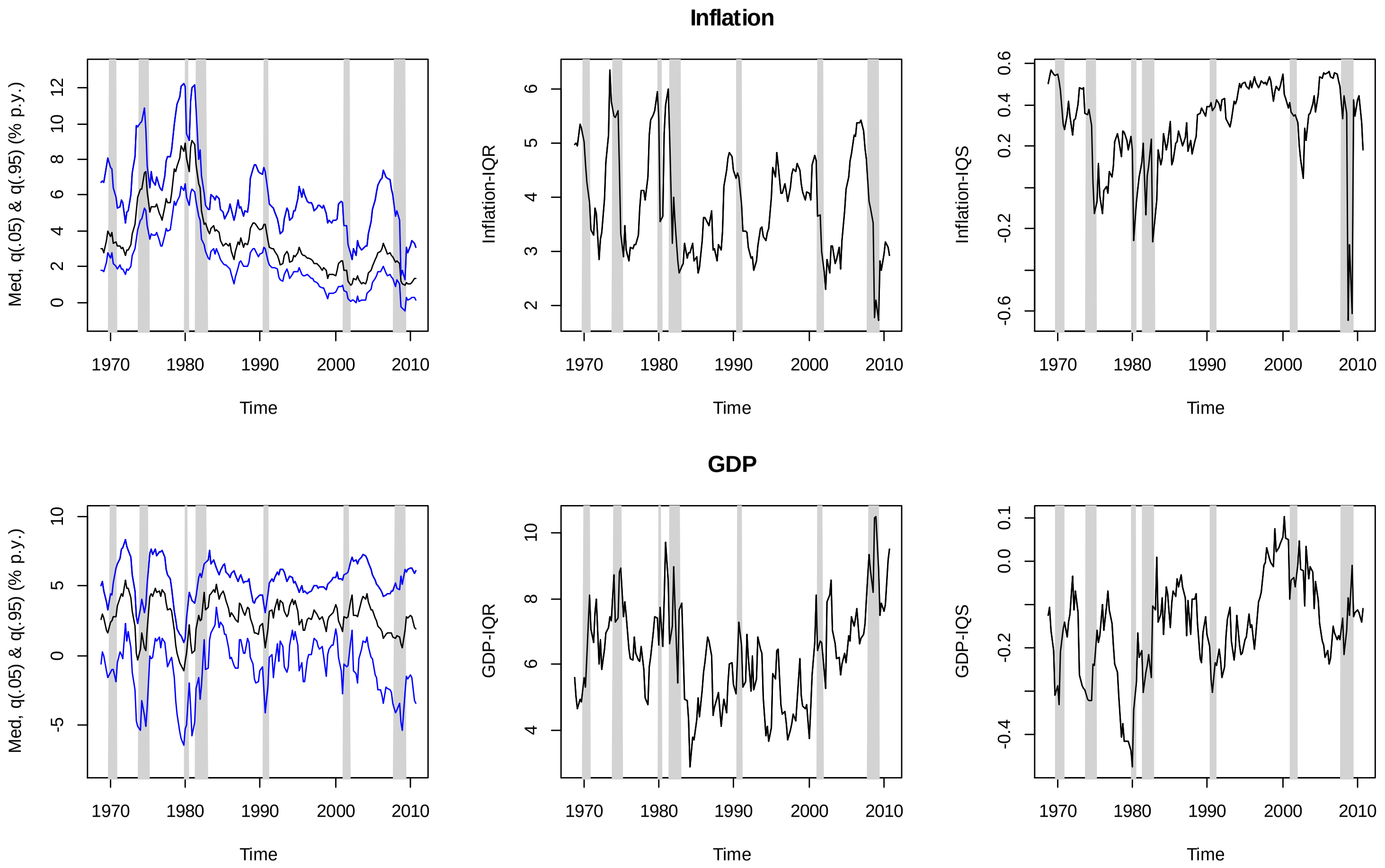


Figure 1.2: Ex ante macroeconomic risks 1

Notes: charts show estimated ex ante macroeconomic risks. The left column shows q(.05), q(.95) along with Med. The middle column shows IQR(.05) and the right column shows IQS(.05).



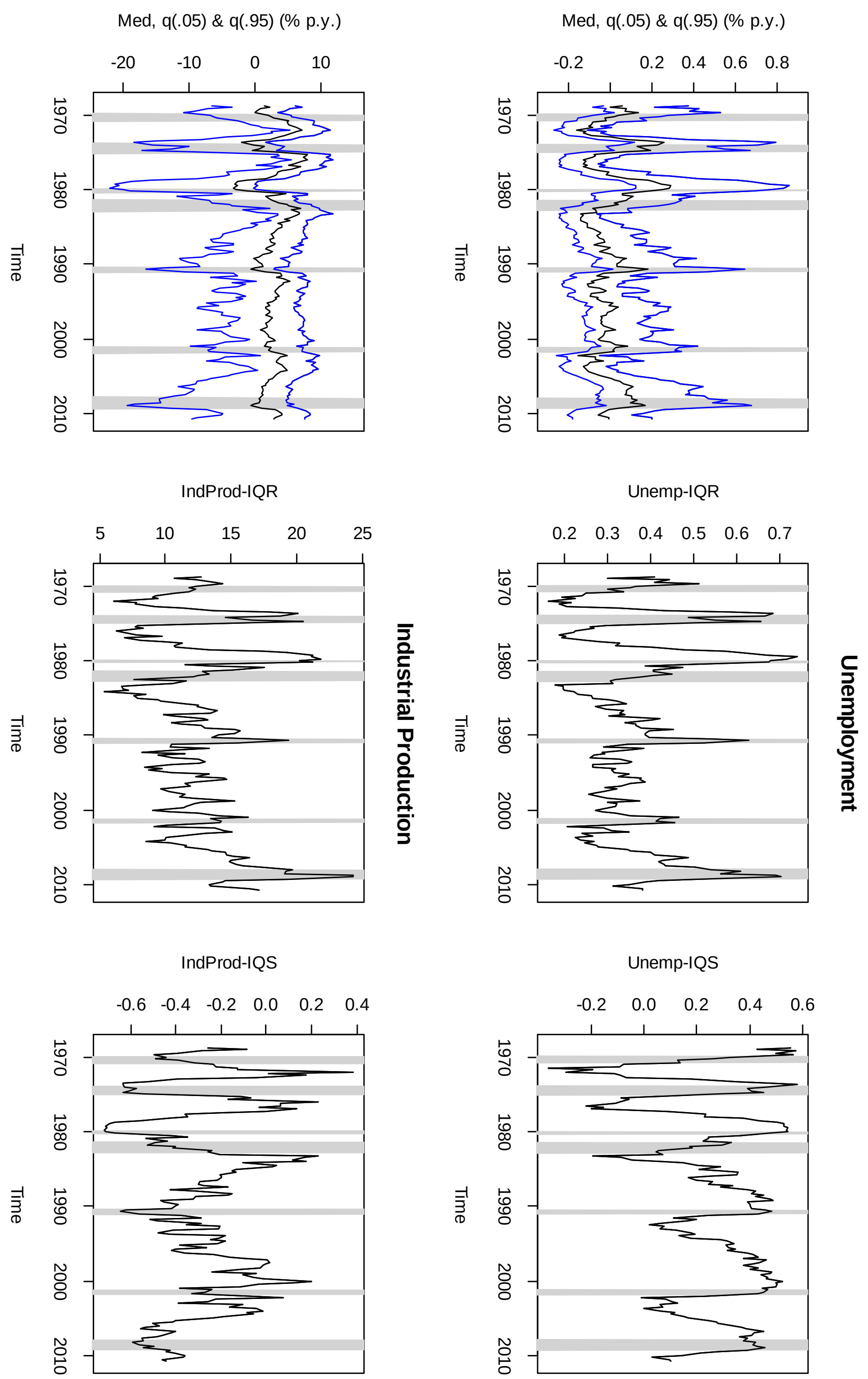
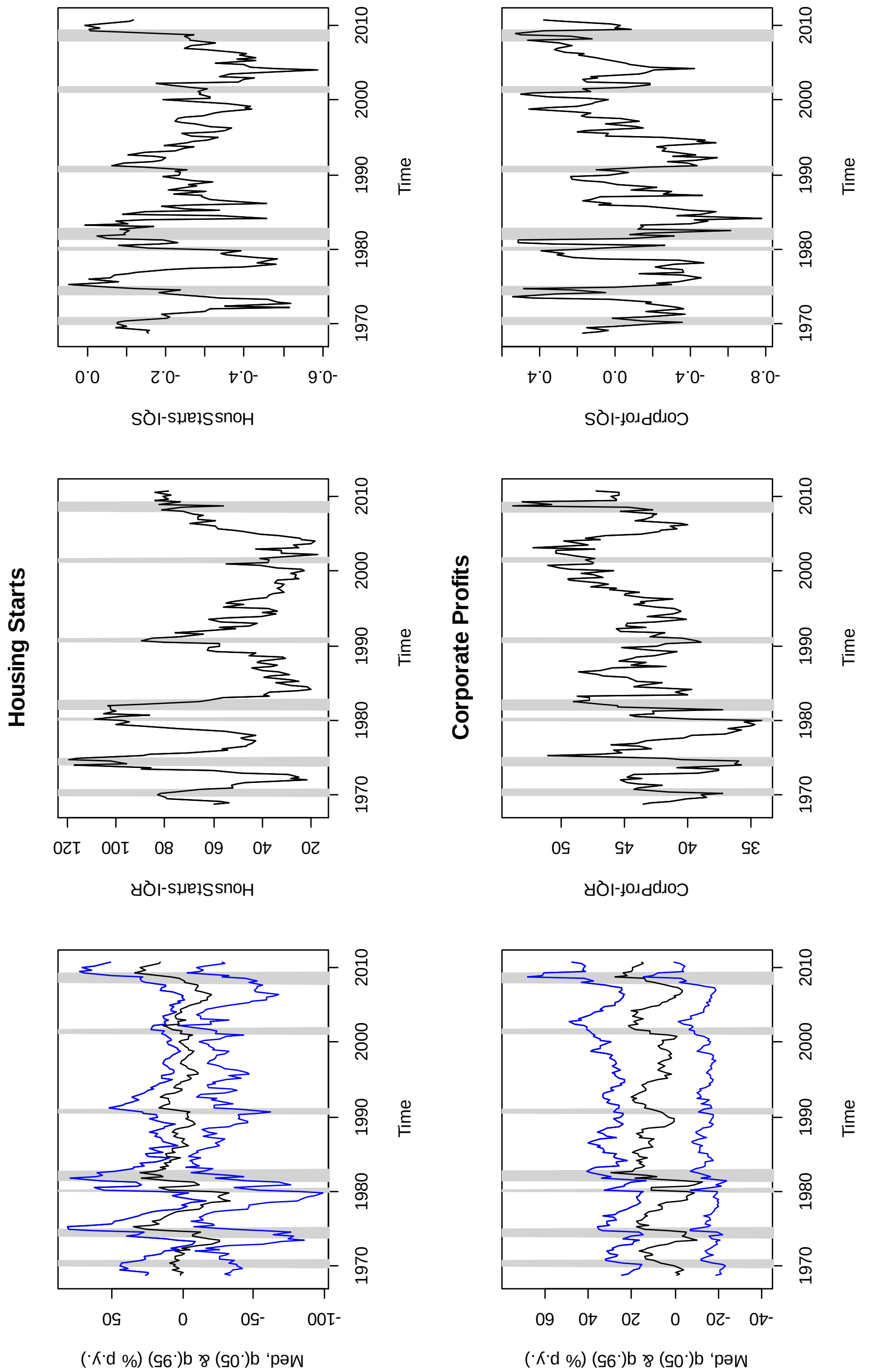


Figure 1.3: Ex ante macroeconomic risks 2

Notes: As in Figure 1.2.

Figure 1.4: Ex ante macroeconomic risks 3



Notes: As in Figure 1.2.

Figure 1.5: Correlations of ex ante macro risks and GDP growth

Notes: graphs show Pearson's correlations between the estimated ex ante macroeconomic risks and GDP growth. Circles indicate statistical significant correlations at the 5% level.

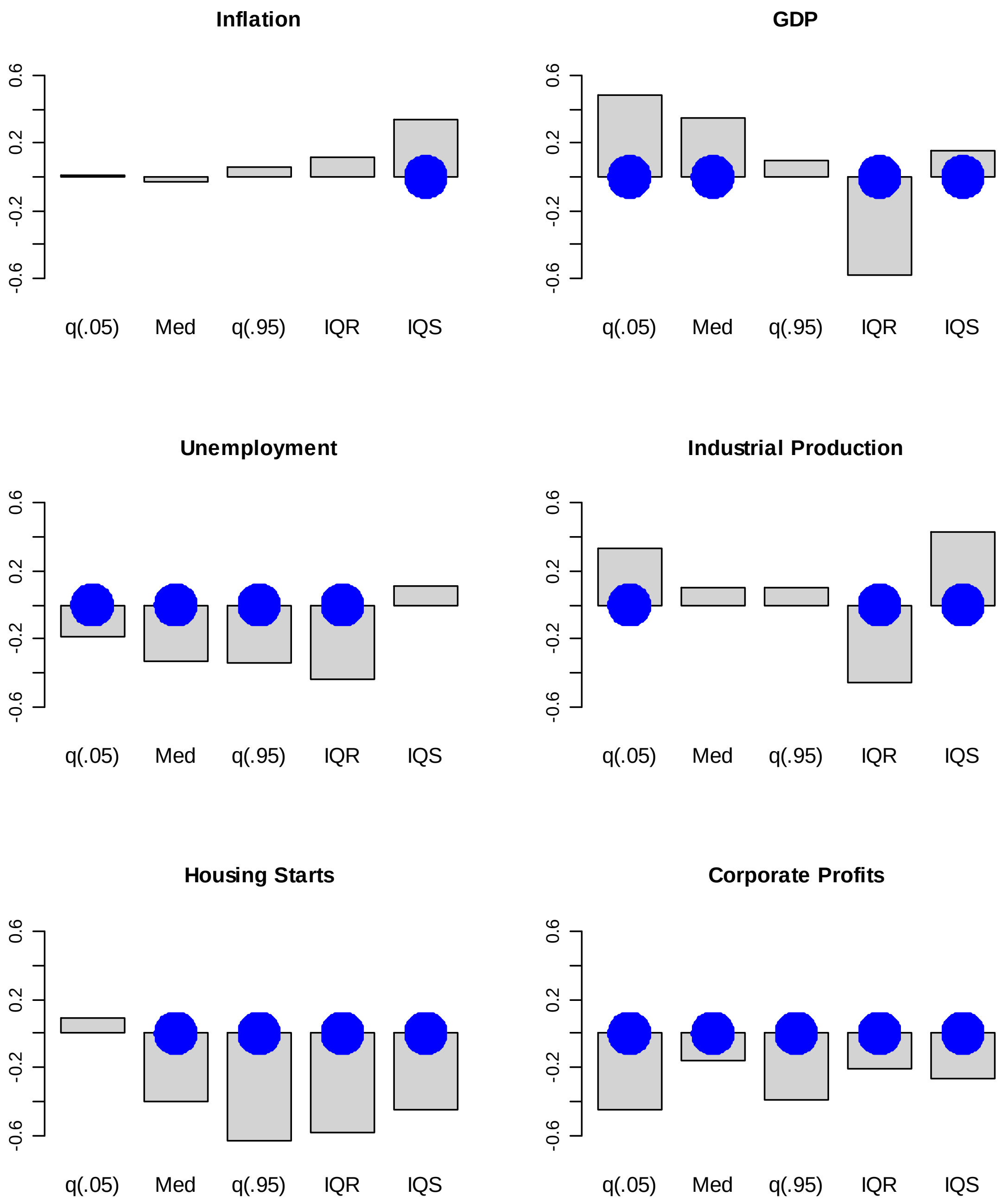


Figure 1.6: Economic Interpretation of macro risk factors - Marginal R^2s

Notes: charts show R^2s from regressions of each estimated ex ante macroeconomic risk onto MRF1, MRF4 and MRF6. Panel A: R^2s are grouped by macroeconomic variables. Panel B: R^2s are grouped by risk measures.

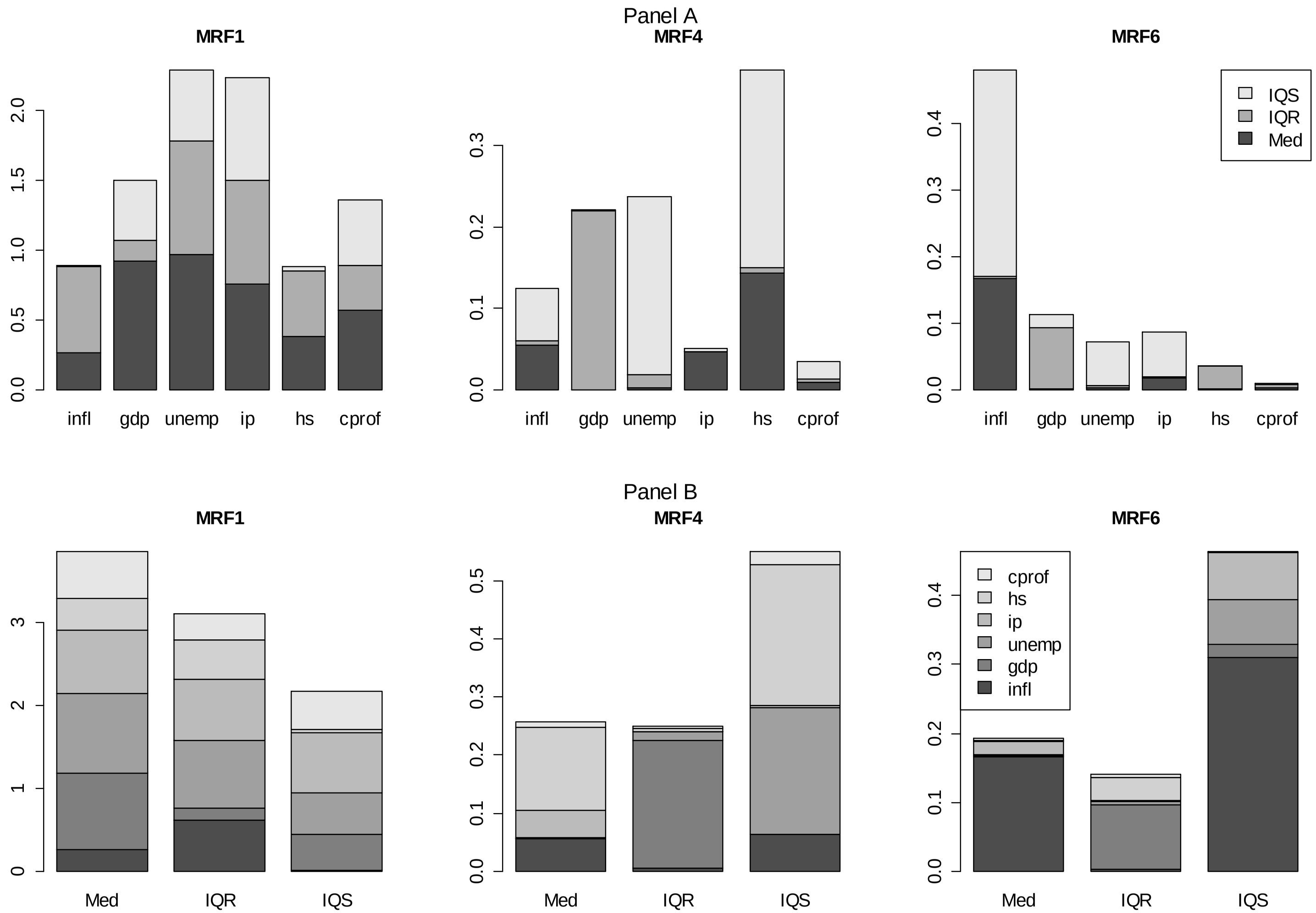


Figure 1.7: Comovements of macro risk factors and key ex ante macro risks

Notes: Standardized units are reported. Shadings areas denote NBER dated recessions. MRF1, MRF4 and MRF6 denote the first, fourth and sixth macro risk factors. Unemp-Med denotes the unemployment median, Housing-IQS and Inflation-IQS denote the housing starts and inflation interquantile skewness, respectively.

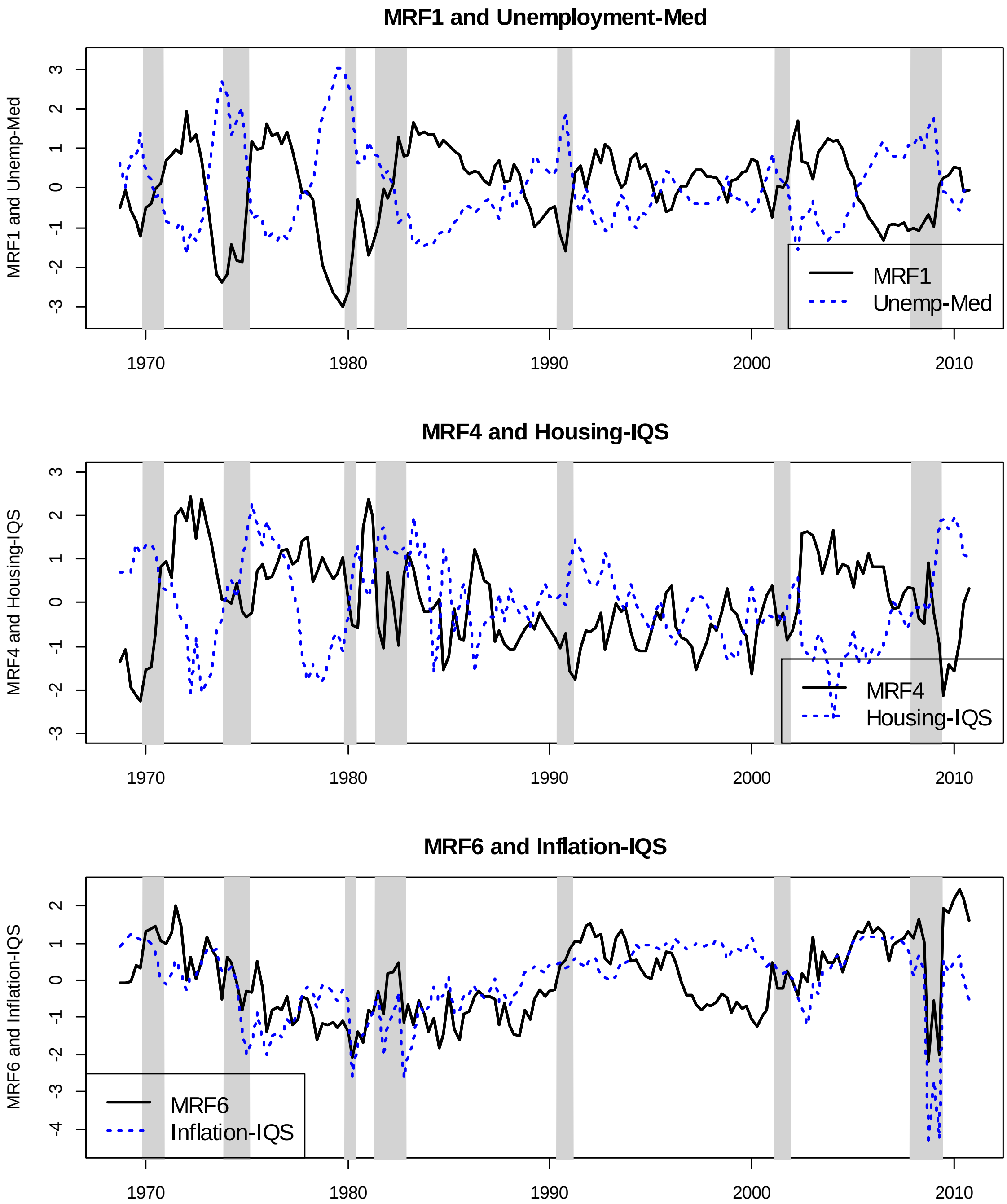


Figure 1.8: Bond risk premium estimate

Notes: Panel A shows the estimated SMRF, a measure of the average risk premia across maturities. Shadings areas denote NBER dated recessions. The risk premium is smoothed using exponential-weighted moving average. In Panel B graphs show lead/lag correlations between the non-smoothed risk premium and growth rates of key economic activity indicators. The risk premium is at date t while growth rates are at time t + l, where l refers to lead (if negative) and lags (if positive). Leads and lags are shown in an annual frequency.

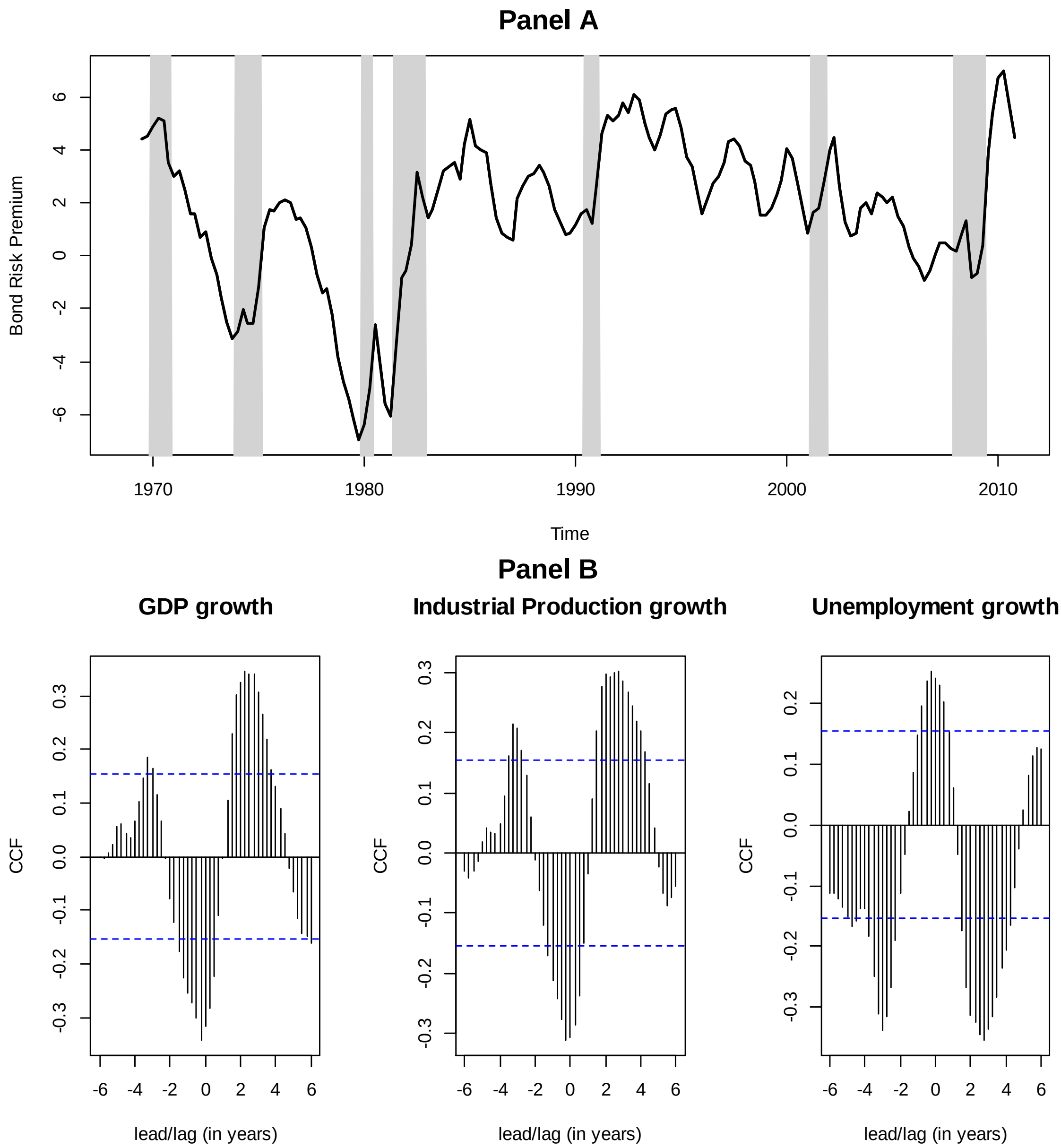


Figure 1.9: 10-year yield decomposition

Notes: Panel A shows NBER dated recessions and time series for the 10-year yield (solid black), 10-year short-rate expectations (dotted blue) and 10-year term premium (solid blue) estimated from the affine term structure model with macro risk factors. The term premium is smoothed using exponential-weighted moving average. In Panel B graphs show lead/lag correlations between the non-smoothed 10-year term premium and growth rates of key economic activity indicators. The term premium is at date t while growth rates are at time t + l, where l refers to lead (if negative) and lags (if positive). Leads and lags are shown in annual frequency.

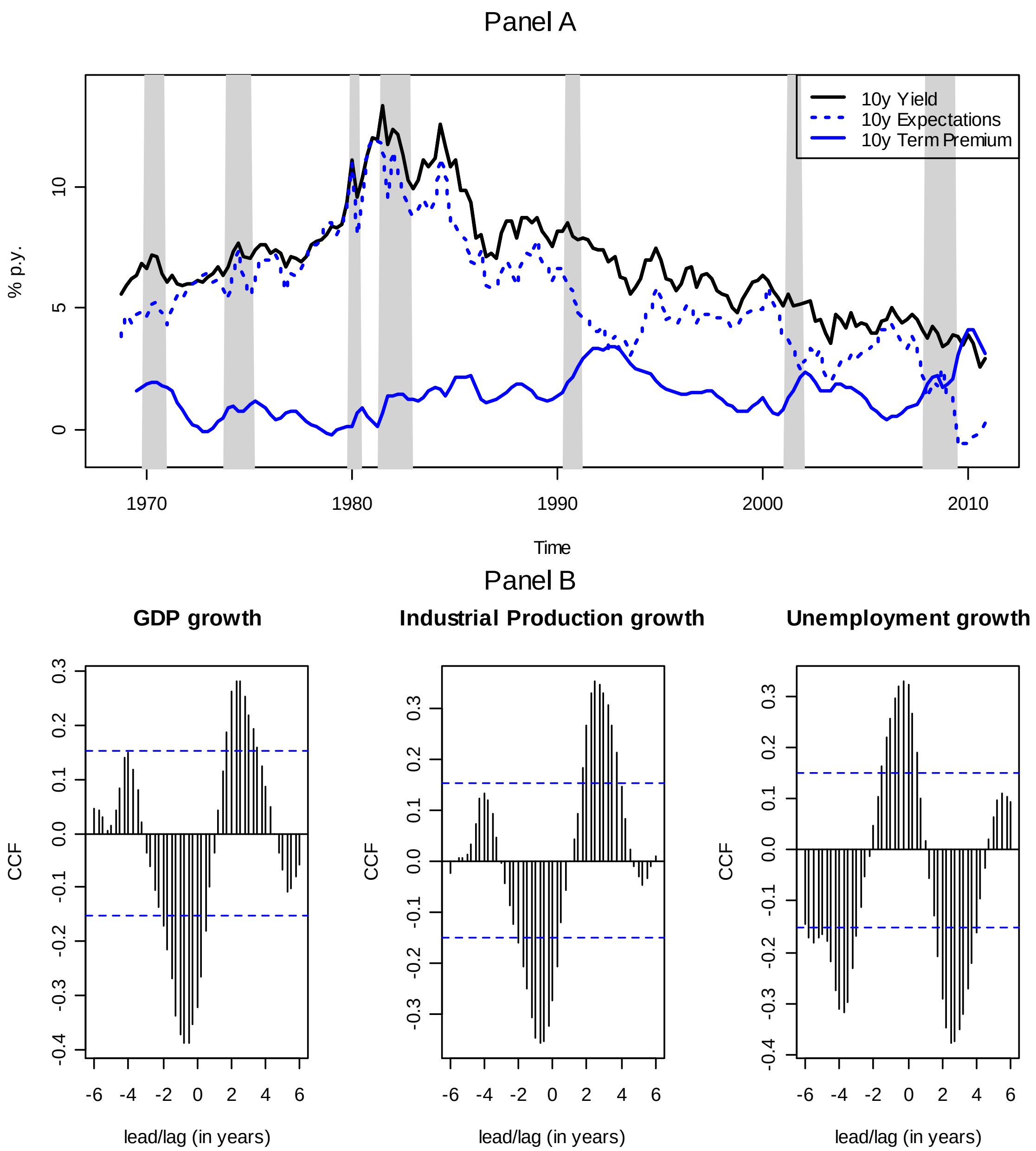
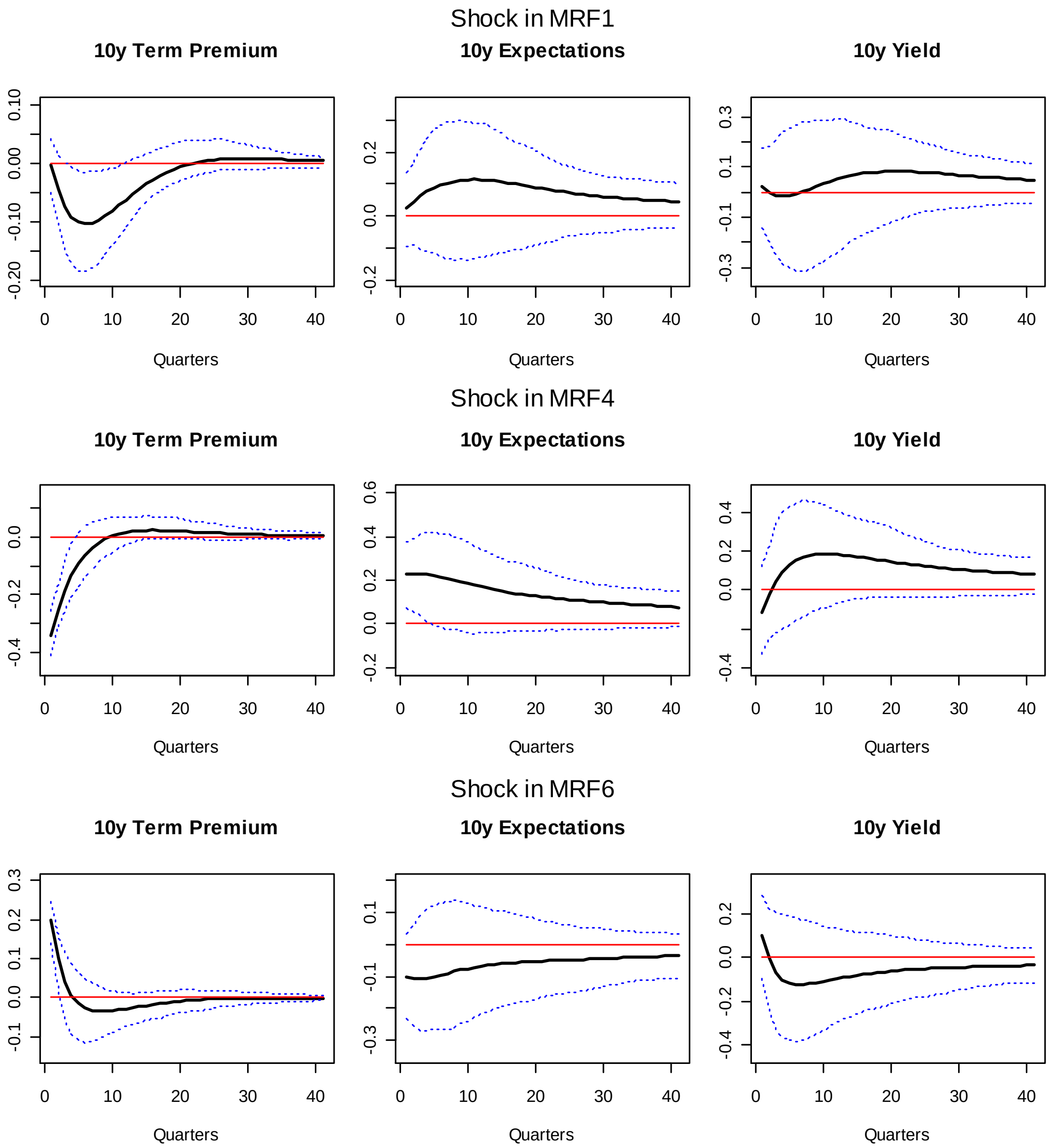


Figure 1.10: Impulse response functions

Notes: This figure shows impulse response functions for the 10-year term premium, 10-year average short-rate expectations and 10-year yield to one standard-deviation shocks in MRF1, MRF4 and MRF6. 95% percent confidence intervals are shown as dotted lines.



Appendices

1.A. Quantile regression estimation

Given $z_{t,t+4} = \beta' x_t + e_t$, the τth quantile regression estimator $\widehat{\beta}(\tau)$ minimizes the following asymmetric loss function

$$V_T = \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(z_{t,t+4} - \beta' x_t\right) = \frac{1}{T}\left[\tau \sum_{z_{t,t+4}\geq\beta' x_t}\left|z_{t,t+4} - \beta' x_t\right| + (1-\tau)\sum_{z_{t,t+4}<\beta' x_t}\left|z_{t,t+4} - \beta' x_t\right|\right]$$

where $\rho_\tau(e) = (\tau - 1_{\{e<0\}})e$ is the check function. $\widehat{\beta}(\tau)$ does not have a closed form, so the minimization problem is solved using the Barrodale-Roberts simplex algorithm for L_1 (Least Absolute Deviation) regressions described in Koenker and d'Orey (1987, 1994). In order to guarantee the monotonicity of $F_{z_{t,t+4}}$, a set of quantile regressions as (1.6) is first estimated for $\tau = 0.01, 0.02, ..., 0.99$ and then the "rearrangement" procedure of Chernozhukov, Fernandez-Val and Galichon (2010) is applied across quantiles. The "rearrangement" procedure is performed as follows. Starting with a model $q_{z_{t,t+4}}(\tau)$ for the conditional quantiles of $z_{t,t+4}$ given x_t, estimate the conditional quantile regression $\widehat{q_{z_{t,t+4}}}(\tau)$. Then use the estimated curve to construct a new random variable $z^*_{t,t+4} \equiv q^*_{z_{t,t+4}}(U)$, where $U \sim iid\, U(0,1)$ is a uniform random variable on $(0,1)$, and estimate its quantile function $q^*_{z_{t,t+4}}(\tau)$ as

$$q^*_{z_{t,t+4}}(\tau) = \widehat{F}^{-1}_{z_{t,t+4}}(\tau) = inf\left\{d : \widehat{F}_{z_{t,t+4}}(d) \geq \tau\right\} \; with\, \widehat{F}_{z_{t,t+4}}(d) \equiv \int_0^1 1\left\{q^*_{z_{t,t+4}}(\tau) \leq d\right\} d\tau$$

which is naturally monotone. Besides guaranteeing the monotonicity of $F_{z_{t,t+4}}$ across quantiles, this procedure also delivers more precisely estimated quantile curves. Chernozhukov, Fernandez-Val and Galichon (2010) show that a "rearranged" curve is closer to the true quantile curve in finite samples than the original one with estimation errors being reduced by up to 14%.

1.B. Macro risk factors estimation

It is assumed that each of the 18 estimated macro risks contained in $\widehat{m}_t$ has a factor structure, i.e.

$$\widehat{m}_{lt} = \lambda_l' MRf_t + e_{lt}$$

where MRf_t is an $s \times 1$ dimensional vector of common macro risk factors, λ_l is a $s \times 1$ vector of factor loadings and e_{lt} denotes an idiosyncratic component. In matrix notation,

$$\widehat{M} = MRf\Lambda + e$$

where $\widehat{M}$ is a $T \times 18$ matrix, MRf is an $T \times s$ matrix of latent macro risk factors, Λ is an $s \times 18$ matrix of factor loadings and e is a $T \times 18$ matrix of idiosyncratic components.

As MRf_t is not observed, it needs to be replaced by estimates $\widehat{MRf}_t$, which is obtained via standard PCA. I start by allowing for s factors in the estimation. Then under the restriction that $\Lambda^{'}\Lambda/18 = I_s$, the factor loadings matrix $\widehat{\Lambda} = \left(\widehat{\lambda}_1, ..., \widehat{\lambda}_{18}\right)$ is estimated by $\sqrt{18}$ times the eigenvectors corresponding to the s largest eigenvalues of the matrix $\widehat{M}^{'}\widehat{M}$. The corresponding factor estimates are then given by $\widehat{MRf}_t = \widehat{M}\Lambda^{'}/18$. As it is usually recommended in factor analysis, all variables in $\widehat{M}$ are standardized prior to estimation. The dimension s of $\widehat{MRf}_t$ is set using Bai and Ng (2002) information criteria, while $\widehat{MRF}_t$ is optimally selected using SBIC after running $rx^n_{t,t+4}$ on all the possible 2^s combinations of factors in $\widehat{MRf}_t$.

1.C. Small Sample Inference

The small-sample performance of test statistics in forecasting regressions with overlapping data is especially important when the right-hand-side variables are highly serially correlated (Bekaert, Hodrick and Marshall, 1997). Even though factors in the vector $\widehat{MRF}_t$ are not as highly persistent as forward rates, for example (see Table A.1.1), a bootstrap analysis is performed. I use a residual-based block bootstrap to assess the small sample properties of test statistics. Bootstrap samples of $rx^n_{t,t+4}$ are obtained by creating bootstrap samples for factors MRF1, MRF4, MRF6, SMRF, SMeF, LN and CP in the first place. Let $\widehat{m}_{lt} = \widehat{\lambda}^{\prime}_{li}\widehat{MRf}_t + \widehat{e}_{lt}$, where $\widehat{\lambda}_l$ and $\widehat{MRf}_t$ are the principal components estimates of λ_l and MRf_t, and $\widehat{e}_{lt}$ is the estimated idiosyncratic error. For each $l = 1, ..., 18$, I estimate an AR(1) model $\widehat{e}_{lt} = \psi_0 + \psi_1 \widehat{e}_{lt-1} + u_{lt}$, sample u^*_{it} from u_{lt} by letting $u^*_{l1} = u_{l1}$ and use the estimated autoregression to obtain $\widehat{e}^*_{lt}$. With $\widehat{e}^*_{lt}$ in hands it is then straightforward to build $\widehat{m}^*_{lt}$ from $\widehat{m}^*_{lt} = \widehat{\lambda}^{\prime}_l\widehat{MRf}_t + \widehat{e}^*_{lt}$, yielding the $T \times 18$ panel $\widehat{M}^*$. Applying PCA to $\widehat{M}^*$ yields $\widehat{MRf}^*_t$ and $\widehat{MRF}^*_t = \left(\widehat{MRF}^*_{1t}, \widehat{MRF}^*_{4t}, \widehat{MRF}^*_{6t}\right)^{'}$, which is then used to obtain $\widehat{SMRF}^*_t$. Applying the same procedure on the panel of macroeconomic variables provided by Ludvigson and Ng (2010) and on the panel of expectations medians, yields LN^*_t and $\widehat{SMeF}^*_t$. CP^*_t is obtained by first approximating it by an AR(1) process, and then sampling the residuals of the autoregression. Bootstrap samples of $rx^{n}_{t,t+4}$ can now be generated from $rx^{n*}_{t,t+4} = \widehat{\delta}_0 + \widehat{\delta}' G^*_t + \varepsilon^*_{t,t+4}$, where G^*_t is a set of bootstrapped regressors, $\varepsilon^*_{t,t+4}$ is sampled from $\varepsilon_{t,t+4} = rx^{n}_{t,t+4} - \widehat{\delta}_0 - \widehat{\delta}' G^*_t$ using overlapping blocks of size equal to six and $\widehat{\delta}$ are the least squares estimates reported in Table 1.2, 1.3 and 1.4. After running the regression of $rx^{n*}_{t,t+4}$ on G^*_t, the bootstrap coefficients $\widehat{\delta}^*_0$ and $\widehat{\delta}^*$ are obtained. This procedure is repeated 4999 times producing empirical distributions for estimated parameters, t-statistics, Wald statistics and $\overline{R}^2$s. In order to be valid, the bootstrap t and Wald statistics were computed as $t^*_j = \frac{\widehat{\delta}^*_j - \widehat{\delta}_j}{s\left(\widehat{\rho}^*_j\right)}$ and $Wald^* = \left(\widehat{\delta}^* - \widehat{\delta}\right)' V\left(\widehat{\delta}^*\right)^{-1}\left(\widehat{\delta}^* - \widehat{\delta}\right)$, where $s\left(\widehat{\delta}^*_j\right)$ and $V\left(\widehat{\delta}^*\right)$ were obtained using a Newey-West HAC estimator with truncation lag equal to six. Asymptotically refined confidence intervals for $\widehat{\delta}_j$ are obtained by computing a percentile-t 95% confidence interval such as $\left[\widehat{\delta}_j - \left|t^*_{0.975}\right| \times s\left(\widehat{\delta}_j\right); \widehat{\delta}_j + \left|t^*_{0.025}\right| \times s\left(\widehat{\delta}_j\right)\right]$.

1.D. Data

Table A.1.2 describes the data used in this study. It provides the name of each variable with its respective code, period, a short description and the data source. In order to match the data obtained from the Survey of Professional Forecasters (SPF) some of the macro variables were constructed by merging different series. For example, gdp was built from merging Real GNP - for the period 1968Q4-1991Q4 - with Real GDP - for the period 1992Q4-2011Q4. The real-time macro data provided by the Philadelphia Fed is merged already, except $cprof$.

1.E. Controlling for the information in x_t

This appendix presents results on the predictive power of SMRF when controling directly for the information in x_t. In order to guard against the possibility of overfitting in out-of-sample forecasting the information in predictors x_t is summarized by estimating predictors factors, Fx, and a single predictors factor, SFx, by applying PCA to the $T \times 9$ panel of original predictors formed by the six consensus forecasts $z^{SPF,h}_t$, MCEI, 5yTS and BaaCS. Bai and Ng (2002) indicates that this panel is well described by seven factors from which three (first, third and fifth principal component) were formally chosen using SBIC. These three factors form the vector Fx. Factor SFx is a linear combination of Fx. It not only provides a variable that can directly be used to control for the information in the original predictors, but also guards against the possibility of overfitting in out-of-sample forecasting. Results of this exercise are provided by Tables A.1.3. Results show that although SFx contains high predictive power, adding SMRF to regressions increases R^2s substantially to levels almost identical to the ones shown by Table 1.2, with statistical significance of SFx also shifting to SMRF. These results indicate that the high predictability found is, to a large extent, due to the extra information obtained from the estimation of Med, IQR and IQS, which capture the different features of the conditional distribution of the macro variables.

1.F. Alternative estimation procedures for IQR and IQS

The baseline quantile regression estimation used in the paper relies on the modified Barrodale and Roberts algorithm for L_1 (or Least Absolute Deviation) regressions described in Koenker and d'Orey (1987, 1994) in conjunction with the monotone rearrangement procedure of Chernozhukov, Fernandez-Val and Galichon (2010). It is well known in the literature of quantile estimation that regression quantiles can show instabilities at tails due to paucity and sparsity of data (He, 1997; Wang, Li and He, 2012). For this reason I present results for two alternative estimation procedures:

(i) $q_{z_{t,t+4}}(\tau)$ is estimated for $\tau = 0.10$ using the baseline approach. Using $\tau = 0.10$ places less weight on extreme data points and assures more robust estimated quantile regressions;

(ii) $q_{z_{t,t+4}}(\tau)$ is estimated for $\tau = 0.05$ using the approach proposed by Wang, Li and He (2012) (WLH hereafter) which integrates quantile regression and Extreme Value Theory and is suitable for quantile curves at tails. This procedure is explained below.

Wang, Li and He (2012) procedure explained

The estimation is performed without assuming common slopes for $q_{z_{t,t+4}}(\tau)$. I focus here on the procedure for the estimation of conditional high quantiles since a low quantile of $z_{t,t+4}$ can be viewed as a high quantile of $-z_{t,t+4}$. First I define a sequence $\tau_j = j/(T+1)$, $j = T-b,...,v$, where $v = T-[T^a]$ with $[T^a]$ denoting the integer part of T^a, $T = 169$ is the sample size and $a > 0$ and $b > a$ are constants fixed as $a = 0.1$, as suggested by Wang, Li and He (2012), and $b = 25$. For each $j = T-b,...,v$ I estimate $\beta(\tau_j)$ in (1.6) following the baseline approach. Then, for each $t = 1,...,T$ I define $\widehat{q_j} = \hat{\beta}(\tau_j)' x_t$ and estimate the parameter γ as

$$\widehat{\gamma} = \frac{1}{b-[T^a]} \sum_{j=[T^a]}^{k} log \frac{\widehat{q_{T-j}}}{\widehat{q_{T-b}}} \tag{1.20}$$

The robust quantile function $q_{z_{t,t+4}}(\tau_s)^{\star}$ is then obtained as $\widehat{q_{z_{t,t+4}}}(\tau_s)^{\star} = \left(\frac{1-\tau_{T-b}}{1-\tau_s}\right)^{\widehat{\gamma}} \widehat{q}_{T-b}$, where τ_s is the percentile of interest. In our case, $\tau_s = 0.05$.

Results using alternative estimation procedures

I show results predictive results for regressions with the following specifications: (i) MRF, (ii) SMRF, (iii) SMRF and CP, (iv) SMRF and LN and (v) SMRF, CP and LN. I only provide results with asymptotic inference. Tables A.1.4 and A.1.5 show results for the baseline approach and percentile equal to 0.10. The optimal vector of macro risk factors selected by SBIC now includes the fifth factor instead of the sixth one. Statistical significance of MRF and SMRF and R^2s remain very high for all bond maturities and results are comparable to the ones shown in tables 1.3 and 1.4. Tables A.1.6 and A.1.7 show results for the WLH approach. The SBIC now selects four factors as predictors: MRF1, MRF4, MRF5 and MRF8. Notice that with the exception of MRF8, all factors are highly statistically significant, and the same is true for Wald statistics. The predicitive powers of MRF and SMRF remain very high with R^2s ranging from 0.17 to 0.30. When comparing the predictive power of SMRF with LN and CP, results are only a bit weaker than the ones shown in Table 1.4, but still high significance and high R^2s are obtained, especially from the 5-year maturity. These results corroborate my previous findings.

Table A.1.1: Sample autocorrelations

	$\widehat{MRF}_1$	$\widehat{MRF}_4$	$\widehat{MRF}_6$	y^1	fw^5	fw^{10}	F_1	F_2	F_6
ρ_1	0.88	0.79	0.78	0.94	0.94	0.95	0.70	0.49	0.18
ρ_3	0.58	0.46	0.55	0.86	0.88	0.89	0.25	0.33	0.22
ρ_5	0.26	0.28	0.46	0.75	0.81	0.82	0.03	0.09	0.03

Table A.1.2: Data description

Variables	Name and/or code as in data source	Period	Description	Source
π^{SPF}	Price Index for GNP/GDP (PGDP)	1968Q4 to 2011Q4	Index level, SA GNP deflator prior to 1992, GDP deflator 1992-1995, GDP Chain-type price index 1996-present	SPF
gdp^{SPF}	Real GNP/GDP (RGDP)	1968Q4 to 2011Q4	Billions of real dollars, SA Real GNP prior to 1992, Real GDP 1992-present	SPF
$unemp^{SPF}$	Civilian Unemployment Rate (UNEMP)	1968Q4 to 2011Q4	Percentage points, SA	SPF
ip^{SPF}	Industrial Production Index (INDPROD)	1968Q4 to 2011Q4	Index level, SA	SPF
hs^{SPF}	Housing Starts (HOUSING)	1968Q4 to 2011Q4	Millions of units, SA	SPF
$cprof^{SPF}$	Corporate Profits After Tax (CPROF)	1968Q4 to 2011Q4	Billions of dollars, SA, Excludes IVA and CCAdj prior to 2006, Includes IVA and CCAdj 2006-present	SPF
$Mich\,Expect$	Reuters/University of Michigan, Consumer Expectations Index	1968Q4 to 2011Q4	Volume index, NSA	Univ. of Michigan/ Thomson Reuters
$5-year\,term\,spread$	5 - year term spread (GS5 - TB3MS)	1968Q4 to 2011Q4	5-Year Treasury Constant Maturity Rate minus 3-Month Treasury Bill: Secondary Market Rate	FRED
$Baa\,corp\,spread$	Baa corporate bond spread (BAA - TB3MS)	1968Q4 to 2011Q4	Moody's Seasoned Baa Corporate Bond Yield minus 3-Month Treasury Bill: Secondary Market Rate	FRED
π	GNP: Implicit Price Deflator (GNPDEF) GDP: Implicit Price Deflator (GDPDEF) GDP: Chain-type Price Index (GDPCTPI)	1968Q4 to 1991Q4 1992Q1 to 1995Q4 1996Q1 to 2011Q4	Index level, SA	FRED
gdp	Real Gross National Product (GNPC96) Real Gross Domestic Product (GDPC96)	1968Q4 to 1991Q4 1992Q1 to 2011Q4	Level, SA	FRED
$unemp$	Civilian Unemployment Rate (UNRATE)	1968Q4 to 2011Q4	SA	FRED
ip	Industrial Production Index (INDPRO)	1968Q4 to 2011Q4	Index, SA	FRED
hs	Housing Starts, Total, New Privately Owned Housing Units Started (HOUST)	1968Q4 to 2011Q4	Thousands of Units, SA	FRED
$cprof$	Corporate profits after tax (A055RC0A144NBEA) Corporate Profits After Tax with Inventory Valuation Adjustment (IVA) and Capital Consump. Adjustm. (CCAdj) (CPATAX)	1968Q4 to 2006Q4 2006Q1 to 2011Q4	Billions of Dollars, SA	FRED
Real-Time macro data	ROUTPUT, P, RUC, IPT, HSTARTS, NCPROFAT, NCPROFATW	1968Q4 to 2011Q4	Various	Philadelphia Fed
$y^1,\ldots,y^5$	Fama-Bliss Discount Bond Yields	1968Q4 to 2011Q4	100 bp	CRSP
$y^6,\ldots,y^{10}$	Gürkaynak, Sack and Wright Bond Yields	1968Q4 to 2011Q4	100 bp	Federal Reserve Board

Table A.1.3: Predictive power of SMRF and SFx

Notes: This table shows the predictive power of SMRF and SFx. t-stats computed using Newey-West standard errors with six lags are reported in parentheses and $\overline{R}^2$ refers to the adjusted-R^2.

	SMRF	*SFx*	$\overline{R}^2$
rx^2		0.225	0.126
		(3.157)	–
	0.224	0.027	0.206
	(2.911)	(0.291)	–
rx^3		0.436	0.144
		(3.243)	–
	0.423	0.063	0.230
	(3.045)	(0.361)	–
rx^5		0.816	0.180
		(3.567)	–
	0.748	0.154	0.277
	(3.478)	(0.541)	–
rx^7		1.196	0.208
		(3.914)	–
	1.003	0.310	0.302
	(3.517)	(0.792)	–
rx^{10}		1.722	0.232
		(4.078)	–
	1.328	0.548	0.319
	(3.607)	(1.080)	–

Table A.1.4: Predictive Power of Macro Risk Factors - baseline with $\tau = 0.10$

Notes: This table shows the predictive power of MRF and SMRF when $q_{z_{t,t+4}}(\tau)$ is estimated using $\tau = 0.10$. t-stats computed using Newey-West standard errors with six lags are reported in parentheses. Wald statistics were also computed using Newey-West variance-covariance matrices with six lags.

	MRF_1	MRF_4	MRF_5	$SMRF$	$\overline{R}^2$	Wald
rx^2	0.397	−0.494	0.447		0.192	0.000
	(2.002)	(−2.341)	(2.034)		–	–
				0.228	0.182	
				(4.080)	–	
rx^3	0.803	−0.864	0.903		0.214	0.000
	(2.302)	(−2.366)	(2.357)		–	–
				0.445	0.210	
				(4.444)	–	
rx^5	1.682	−1.223	1.729		0.255	0.000
	(2.921)	(−2.250)	(2.772)		–	–
				0.832	0.262	
				(5.306)	–	
rx^7	2.605	−1.416	2.487		0.282	0.000
	(3.371)	(−1.929)	(2.998)		–	–
				1.195	0.291	
				(5.803)	–	
rx^{10}	3.954	−1.633	3.502		0.310	0.000
	(3.756)	(−1.686)	(3.188)		–	–
				1.700	0.315	
				(5.879)	–	

Table A.1.5: Predictive Power of SMRF, CP and LN factors - baseline with $\tau = 0.10$

Notes: This table shows the predictive power of CP, LN and SMRF factors when $q_{z_{t,t+4}}(\tau)$ is estimated using $\tau = 0.10$. t-stats computed using Newey-West standard errors with six truncation lags are reported in parentheses. Regressions in which LN is included in the set of predictors are estimated using the sample 1968Q4 - 2007Q4.

	SMRF	*CP*	$\overline{R}^2$	*SMRF*	*LN*	$\overline{R}^2$	*SMRF*	*CP*	*LN*	$\overline{R}^2$
rx^2		0.242	0.215		0.304	0.213		0.195	0.190	0.354
		(5.205)	–		(4.957)	–		(3.897)	(2.704)	–
	0.121	0.164	0.245	0.180	0.202	0.321	0.092	0.145	0.167	0.374
	(2.058)	(3.286)	–	(3.681)	(3.374)	–	(1.613)	(2.557)	(2.364)	–
rx^3		0.468	0.242		0.545	0.208		0.384	0.321	0.375
		(5.338)	–		(5.1045)	–		(4.353)	(2.771)	–
	0.240	0.314	0.279	0.362	0.338	0.342	0.192	0.280	0.272	0.400
	(2.158)	(3.227)	–	(4.155)	(3.175)	–	(1.827)	(2.726)	(2.279)	–
rx^5		0.814	0.261		0.863	0.189		0.692	0.459	0.384
		(5.159)	–		(5.595)	–		(4.299)	(2.989)	–
	0.516	0.486	0.321	0.684	0.472	0.361	0.395	0.476	0.359	0.423
	(2.974)	(2.927)	–	(4.957)	(3.461)	–	(2.689)	(2.822)	(2.317)	–
rx^7		1.185	0.296		1.182	0.188		1.004	0.596	0.406
		(5.669)	–		(5.579)	–		(4.507)	(3.126)	–
	0.724	0.723	0.360	0.991	0.615	0.380	0.572	0.692	0.451	0.449
	(2.940)	(3.069)	–	(5.242)	(3.268)	–	(2.911)	(2.975)	(2.334)	–
rx^{10}		1.686	0.321		1.559	0.172		1.445	0.715	0.411
		(6.018)	–		(5.012)	–		(4.768)	(2.918)	–
	1.029	1.029	0.390	1.435	0.739	0.385	0.835	0.990	0.504	0.460
	(2.955)	(3.275)	–	(5.203)	(2.682)	–	(2.897)	(3.205)	(2.007)	–

Table A.1.6: Predictive Power of Macro Risk Factors - WLH

Notes: This table shows the predictive power of MRF and $SMRF$ when $q_{z_{t,t+4}}(\tau)$ is estimated using $\tau = 0.05$ and the WLH approach. t-stats computed using Newey-West standard errors with six lags are reported in parentheses. Wald statistics were also computed using Newey-West variance-covariance matrices with six lags.

	MRF_1	MRF_4	MRF_5	MRF_8	$SMRF$	$\overline{R}^2$	Wald
rx^2	0.432	−0.555	0.263	0.058		0.175	0.000
	(2.304)	(−3.550)	(1.977)	(0.312)		–	–
					0.225	0.166	
					(3.802)	–	
rx^3	0.860	−0.958	0.534	0.242		0.187	0.000
	(2.457)	(−3.800)	(2.309)	(0.704)		–	–
					0.436	0.190	
					(4.075)	–	
rx^5	1.758	−1.451	0.987	0.768		0.230	0.000
	(3.010)	(−4.072)	(2.556)	(1.346)		–	–
					0.823	0.242	
					(4.718)	–	
rx^7	2.700	−1.809	1.537	1.117		0.262	0.000
	(3.545)	(−3.928)	(2.937)	(1.529)		–	–
					1.195	0.275	
					(5.245)	–	
rx^{10}	4.076	−2.286	2.196	1.610		0.292	0.000
	(3.983)	(−3.728)	(3.155)	(1.739)		–	–
					1.711	0.302	
					(5.405)	–	

Table A.1.7: Predictive Power of SMRF, CP and LN factors - WLH

Notes: This table shows the predictive power of CP, LN and SMRF factors when $q_{z_{t,t+4}}(\tau)$ is estimated using $\tau = 0.05$ and the WLH approach. t-stats computed using Newey-West standard errors with six truncation lags are reported in parentheses. Regressions in which LN is included in the set of predictors are estimated using the sample 1968Q4 - 2007Q4.

	SMRF	*CP*	$\overline{R}^2$	*SMRF*	*LN*	$\overline{R}^2$	*SMRF*	*CP*	*LN*	$\overline{R}^2$
rx^2		0.242	0.215		0.304	0.213		0.195	0.190	0.354
		(5.205)	–		(4.957)	–		(3.897)	(2.704)	–
	0.111	0.175	0.240	0.179	0.212	0.309	0.0805	0.154	0.173	0.368
	(1.972)	(3.852)	–	(4.100)	(3.500)	–	(1.581)	(2.855)	(2.495)	–
rx^3		0.468	0.242		0.545	0.208		0.384	0.321	0.375
		(5.338)	–		(5.104)	–		(4.353)	(2.771)	–
	0.217	0.337	0.272	0.350	0.365	0.320	0.155	0.304	0.288	0.390
	(2.006)	(3.787)	–	(4.222)	(3.349)	–	(1.602)	(3.318)	(2.471)	–
rx^5		0.814	0.261		0.863	0.189		0.692	0.459	0.384
		(5.159)	–		(5.595)	–		(4.299)	(2.989)	–
	0.483	0.524	0.313	0.666	0.519	0.336	0.333	0.519	0.388	0.409
	(2.717)	(3.466)	–	(4.785)	(3.499)	–	(2.410)	(3.334)	(2.511)	–
rx^7		1.185	0.296		1.182	0.188		1.004	0.596	0.406
		(5.669)	–		(5.579)	–		(4.507)	(3.126)	–
	0.698	0.765	0.354	0.959	0.686	0.350	0.473	0.758	0.494	0.433
	(2.774)	(3.533)	–	(4.835)	(3.347)	–	(2.516)	(3.541)	(2.580)	–
rx^{10}		1.686	0.321		1.559	0.172		1.445	0.715	0.411
		(6.018)	–		(5.012)	–		(4.768)	(2.918)	–
	1.012	1.077	0.386	1.405	0.833	0.355	0.717	1.074	0.562	0.443
	(2.839)	(3.695)	–	(4.742)	(2.849)	–	(2.562)	(3.795)	(2.292)	–

Chapter 2

Out-of-sample bond excess returns predictability[1]

Rafael B. De Rezende

ABSTRACT. This article investigates the out-of-sample predictability of bond excess returns. I assess the statistical and economic significance of forecasts generated by empirical models based on forward interest rates, macroeconomic variables and risks in macroeconomic outcomes. Results suggest that macroeconomic variables, risks in macroeconomic outcomes as well as the combination of these different sources of information outperform a constant model of no-predictability. These results are confirmed when using macroeconomic data available in real-time, suggesting that the predictability of bond returns is not driven by data revisions.

Keywords: out-of-sample; bond risk premia; statistical predictability; economic predictability.
JEL Classifications: G11, G12, E43, E44

2.1 Introduction

Empirical research in financial economics has revealed significant predictable variation in expected excess returns of US government bonds, a violation of the expectations hypothesis. Understanding this variation and its relationship with the economy has been an important question in economics and finance, and an active area of ongoing research. Fama (1984), Fama and Bliss (1987), Stambaugh (1988) and Cochrane and Piazzesi (2005, 2008) find that yield spreads and forward rates

[1] I would like to thank Magnus Dahlquist, Lars E.O. Svensson and Michael Halling for comments and suggestions that significantly improved this paper. I am also grateful to Ádám Faragó, Roméo Tédongap, Erik Hjalmarsson, Andrejs Delmans, Nikita Koptyug and Ricardo Aliouchkin for comments and suggestions. I kindly thank the Swedish Bank Research Foundation (BFI) for financial support.

predict excess bond returns with R^2s ranging from 10% to 40%. Ludvigson and Ng (2009) and Cooper and Priestley (2009) document that macroeconomic variables carry information about bond risk premia not contained in financial variables. De Rezende (2014) finds that measures of risks in macroeconomic outcomes provide information about bond risk premia variation that is not embedded in forward rates and macroeconomic variables.

In-sample predictive power, however, does not necessarily imply out-of-sample predictability (Inoue and Kilian, 2004, 2006). Thornton and Valente (2012) show that, although working quite well in-sample, the forward rate factor of Cochrane and Piazzesi (2005) generates poor out-of-sample bond excess returns forecasts. Similarly, in the context of equity premium prediction, Welch and Goyal (2008) show that a long list of predictors commonly found in the literature are unable to deliver out-of-sample forecasts that are consistently superior to a simple random walk model.

This paper contributes to the existing literature by examining the failure of the expectation hypothesis of the term structure of interest rates in an out-of-sample setting. More specifically, I look at the predictive abilities of the forward rate factor of Cochrane and Piazzesi (2005, CP hereafter), the macro factor of Ludvigson and Ng (2009, LN hereafter) and the macro risk factor of De Rezende (2014, SMRF hereafter) and compare them to a constant model of no-predictability in a genuine out-of-sample exercise.

Since statistical predictability does not imply economic predictability (Leitch and Tanner, 1991; Della Corte, Sarno and Thornton, 2008; Della Corte, Sarno, and Tsiakas, 2009), I also assess the economic value of the three predictors. The analysis is based on a classical portfolio choice problem, in which I consider a risk-averse investor who exploits the predictive power of factors to invest in a portfolio of bonds. I quantify the portfolio management fee that the investor would be willing to pay to have access to the additional information available in a predictive regression model relative to a constant model. The other measures of portfolio performance I consider are the Sharpe ratio and the risk-adjusted measure proposed by Goetzmann et al. (2007), which has is shown to be robust to a number of manipulations.

Results reveal that the factors LN and SMRF are largely more accurate than the constant model of no-predictability. While LN performs quite well in shorter maturities, SMRF performs best in intermediate and long maturities. For instance, I find that prediction errors are reduced by up to 29% when relying on SMRF. The LN factor, in turn, delivers predictions that are up to 12% more accurate than the constant model. More interestingly, the combination of the two factors generates improvements ranging from 15% to 35%. The CP factor, on the other hand, is shown to perform poorly and does not beat the constant model for any bond maturity. The combination of the three factors, however, increases predictive power substantially with prediction errors being reduced by levels ranging from 20% to 37%. These results provide further evidence that the three factors capture somewhat independent information about bond risk premia variation, as documented by De Rezende (2014) in an in-sample exercise.

When evaluated economically, results indicate that SMRF delivers utility gains that are around 3% superior than the constant model. Slightly lower numbers are found for the LN factor, but combining the two variables increases utility gains to levels up to 3.7%. Similar results are found for the other measures of portfolio performance. Consistent with the statistical analysis, the CP factor shows a poor performance, which is in line with the findings by Thornton and Valente (2012). Improvements concerning the CP factor are only obtained when augmenting its regressions with factors SMRF, LN or both.

These results are confirmed when using macroeconomic data available in real-time. Although the use of revised macroeconomic data overstate the information set available to investors at the time predictions are made, I document that the predictive power of a SMRF constructed in real-time is not strongly affected, suggesting that the predictability of bond returns is not necessarily driven by data revisions. Ghysels, Horan and Moench (2012) conclude differently when relying on the Ludvigson and Ng (2009) macro factors.

This paper is part of large literature investigating the failure of the expectations hypothesis of the term structure of interest rates (Fama, 1984; Fama and Bliss, 1987; Stambaugh, 1988; Cochrane and Piazzesi, 2005; Ludvigson and Ng, 2009; Cooper and Priestley, 2009; Huang and Shi, 2012; Cieslak and Povala, 2012; among others). Although Thornton and Valente (2012) find no evidence of time-variation in out-of-sample bond risk premia based on forward rates, this paper identifies two important sources of out-of-sample time variation in bond risk premia: current macroeconomic conditions and risks in macroeconomic outcomes.

This paper also relates to a literature examining time-variation in risk premium in equity markets. Welsh and Goyal (2008) investigates the predictive power of a large number of financial and macroeconomic based variables and find no systematic evidence of out-of-sample excess equity return predictability. Campbell and Thompson (2008), Rapash, Strauss and Zhou (2010), Henkel, Martin and Nardari (2011) and Dangl and Halling (2012) investigate the same set of predictors applying different econometric models and find evidence of predictability in an out-of-sample setting. Differently, this paper investigates variation in risk premia in government bond markets and find evidence of statistical as well as economic out-of-sample predictability.

The rest of the paper is organized as follows. The next section describes the framework for predicting excess bond returns; section three describes the framework used to assess the bond return predictions; the forth section presents the main results of the paper; and the last section concludes.

2.2 Predicting bond excess returns

In line with the existing literature, I focus on one-year log returns on an n-year zero-coupon Treasury bond in excess of the annualized yield on a 1-year zero coupon bond. More specifically, for $t = 1, ..., T$, one-year excess returns are denoted as $rx^n_{t,t+4} = r^n_{t,t+4} - y^1_t = -(n-1)\,y^{n-1}_{t+4} + ny^n_t - y^1_t$, where $r^n_{t,t+4}$ is the one-year log holding-period return on an n-year bond purchased at time t and sold one year after at time $t+1$ year and y^n_t is the log yield on the n-year bond.

Several studies have uncovered that expected bond excess returns vary over time and that they are a quantitatively important source of fluctuations in the bond market. In this study, I selected three models that have been successful in explaining variation in bond excess returns. More specifically, I look at the predictive power of the single forward factor of Cochrane and Piazzesi (2005), the macro factor of Ludvigson and Ng (2009) and the macro risk factor of De Rezende (2014).

As in Cochrane and Piazzesi (2008), CP was formed from a linear regression of average excess returns (across maturities ranging from 2-year to 10-year) on the 1-year yield and forward rates from two to ten years,

$$\begin{gathered} \overline{rx}_{t,t+4} = \delta_0 + \delta_1 y^1_t + ... + \delta_{10} fw^{10}_t + \varepsilon_{t,t+4} \\ CP_t = \widehat{\delta}' fw_t \end{gathered} \tag{2.1}$$

where fw^n_t is the n-year forward rate defined as $fw^n_t = -(n-1)\,y^{n-1}_t + ny^n_t$.

LN was obtained from a linear combination of macro factors extracted from a large macroeconomic data set (131 variables). When forming LN I used the data set provided by Ludvigson and Ng (2010) but I set October 1968 as the starting date to enable direct comparisons with the other predictors studied in the paper.[2] Quarterly frequency was obtained by selecting observations for the second month of each quarter. LN was then constructed by running the average bond returns on the

[2]The data set was downloaded from Sydney C. Ludvigson's web page: http://www.econ.nyu.edu/user/ludvigsons/.

best subset of macro factors selected using Schwarz (1978) Bayesian information criteria (SBIC),

$$\begin{aligned} \overline{rx}_{t,t+4} &= \varphi' F_t + \varepsilon_{t,t+4} \\ LN_t &= \widehat{\varphi}' F_t \end{aligned} \tag{2.2}$$

where $\widehat{\varphi}$ is a line vector of estimated parameters and F_t is a column vector of macro factors estimated by Principal Component Analysis, where I also disregard the use of hats to ease notation.[3]

The single macro risk factor of De Rezende (2014), SMRF, was obtained from a linear combination of macro risk factors, as specified in De Rezende (2014). These are obtained by estimating a factor model on a set of measures of macroeconomic risks estimated from quantile regressions and that capture expectations, uncertainty and downside (upside) risks for six key US macro variables. SMRF was constructed by running the average bond returns on the best subset of macro risk factors selected optimally using SBIC,

$$\begin{aligned} \overline{rx}_{t,t+4} &= \theta_0 + \theta_1 MRF_{1t} + \theta_2 MRF_{4t} + \theta_3 MRF_{6t} + \varepsilon_{t,t+4} \\ SMRF_t &= \widehat{\theta}' MRF_t \end{aligned} \tag{2.3}$$

where $\widehat{\theta}$ is a line vector of estimated parameters and MRF_t is a column vector of macro risk factors also estimated by Principal Component Analysis.

Bond excess returns can then be predicted by estimating the following regression by OLS,

$$rx^n_{t,t+4} = \alpha_0 + \alpha' Z_t + \varepsilon_{t,t+4} \tag{2.4}$$

where Z_t can assume different forms: $SMRF_t$, CP_t, LN_t, $SMRF_t + CP_t$, $SMRF_t + LN_t$, $CP_t + LN_t$, $SMRF_t + CP_t + LN_t$.

When $\alpha = 0$ bond excess returns are not predictable and are constant over time. This model is consistent with the expectations hypothesis of the term structure of interest rates and is a natural candidate for testing its validity in an out-of-sample setup.

[3]Following Ludvigson and Ng (2009) I also included F^3_{1t} in the set of macro factors.

2.3 Assessing bond excess returns predictions

2.3.1 The statistical approach

The predictors evaluated are SMRF, CP, LN and their respective combinations. I conduct several model comparisons. First, I assess the incremental predictive power of each predictor and their respective combinations relative to a constant model of no-predictability. In the second round of comparisons I test whether adding a factor to a specific regression increases its predictive power. Ludvigson and Ng (2009) shows that including LN in a regression with the CP factor increases out-sample predictability substantially. In this paper, I test whether adding SMRF to CP, LN and CP+LN regressions increases their predictive power. In this case, I compare the out-of-sample forecasting performance of an "unrestricted" specification including SMRF and the other predictors to the performance of a "restricted" model (the null) which includes only CP, LN or both.

It is important to clarify how this exercise is implemented. The first estimation window uses data from 1968Q4 to 1989Q1, meaning that forecasts of excess bond returns were generated for the period 1990Q1 - 2011Q4. I implement a recursive forecasting scheme in which all parameters are re-estimated as a new observation becomes available. In each recursion, factors used in the construction of SMRF and LN are also re-estimated and picked optimally (using SBIC), taking into consideration the possibility that different factors may be chosen in different samples. When the LN factor is included in the set of predictors, the out-of-sample portion of data ends at 2007Q4.

For both evaluations I use the out-of-sample R^2 statistic, R^2_{oos}, suggested by Campbell and Thompson (2008). The R^2_{oos} statistic measures the reduction in Mean Squared Prediction Error (MSPE) obtained with a predictor based ("unrestricted") model relative to the constant ("restricted") model. Thus, when $R^2_{oos} > 0$, the predictor based ("unrestricted") model outperforms the constant ("restricted") model according to the MSPE metric.[4]

A more rigorous comparison, however, can be assessed by relying on the MSPE-adjusted test statistic proposed by Clark and West (2007) (CW hereafter) and the MSE-F statistic of equal forecast performance of McCracken (2007) (MC hereafter),

[4] The R^2_{oos} statistic is given by $R^{2,j}_{oos} = 1 - \frac{\sum_{t=R}^{T}\left(rx^n_{t,t+4} - \widehat{rx}^{n,j}_{t,t+4}\right)^2}{\sum_{t=R}^{T}\left(rx^n_{t,t+4} - \widehat{rx}^{n,b}_{t,t+4}\right)^2}$, where $\widehat{rx}^{n,j}_{t,t+4}$ is a forecast generated from model $j = SMRF, CP, LN, SMRF+CP, SMRF+LN, CP+LN, SMRF+CP+LN$ and $\widehat{rx}^{n,b}_{t,t+4}$ is the forecast generated from the benchmark, with $b = constant, CP, LN, CP+LN$. R is the length of the initial sample window used for estimating parameters and T the total sample size.

which are both suitable for cases when one is comparing forecasts generated from nested models. In this case, a rejection of the null hypothesis implies that additional regressors contain out-of-sample predictive power regarding $rx^n_{t,t+4}$.

The advantage of relying on the MSPE-adjusted test of CW, however, is that it corrects for finite sample bias in MSPE comparison between nested models. The correction accounts for the fact that when considering two nested models, the smaller model has an unfair efficiency advantage relative to the larger one because it imposes zero parameters that are zero in population, while the alternative introduces noise into the forecasting process that will, in finite samples, inflate the MSPE. Without correcting the test statistic the researcher may, therefore, erroneously conclude that the smaller model is better, resulting in size distortions where the larger model is rejected too often. The MSPE-adjusted statistic makes a correction that addresses this finite sample bias, and the correction is why it is possible for the larger model to outperform the benchmark even when the computed MSPE differences are positive.

2.3.2 The economic value approach

Statistical predictability does not mechanically imply economic predictability as it does not explicitly account for the risk borne by an investor over the out-of-sample period (Leitch and Tanner, 1991; Della Corte, Sarno and Thornton, 2008; Della Corte, Sarno, and Tsiakas, 2009). It is then useful to also assess the economic value of predictors relative to the constant model of no-predictability using an asset allocation framework. Here, I closely follow the work by Thornton and Valente (2012) who evaluate the economic value of forward rates.

The analysis is based on a classical portfolio choice problem, in which I consider an investor who exploits the predictability of excess returns to optimally invest in a portfolio comprising $J+1$ bonds: a risk-free 1-year bond and J risky n-year bonds. The investor constructs a dynamically rebalanced portfolio by choosing weights to maximize the trade-off between mean and variance in the portfolio return. More specifically, at each date t the investor solves the following problem,

$$\max_{\mathbf{w_t}} \mathbf{w_t}' \mu_{t,t+4} - \frac{\rho}{2} \mathbf{w_t}' \sum\nolimits_{t,t+4} \mathbf{w_t} \tag{2.5}$$

with solution equal to $\mathbf{w_t} = \frac{1}{\rho} \sum_{t,t+4}^{-1} \mu_{t,t+4}$, where $\mathbf{w_t} = \left(w_t^2, ..., w_t^{10}\right)'$ is the $J \times 1$ vector of weights on the risky bonds, $\mu_{t,t+4}$ and $\sum_{t,t+4}$ are the conditional expectation and the conditional variance-covariance matrix of the $J \times 1$ vector of excess bond returns $\mathbf{rx}_{t,t+4}$ and ρ is a parameter governing the degree of investor's risk aversion.

I limit the weights for each of the n-year risky-bonds by $-1 \leq w_t^n \leq 1$ to avoid extreme investments (Welsh and Goyal, 2008; Dangl and Halling, 2012), but allow for the full proceeds of short sales (Vayanos and Weill, 2008; Thornton and Valente, 2012). The weight on the 1-year bond is equal to $1-\mathbf{w}_{\mathbf{t}}^{'}\iota$, where ι is a $J\times 1$ vector of ones. Conditional expected bond excess returns, $\mu_{t,t+4}$, are generated using the constant model of no-predictability and various other predictor based models. Volatility forecasts are obtained by assuming that the conditional covariance matrix of the residuals of each model, $\sum_{t,t+4} = E\left(\varepsilon_{t,t+4}\varepsilon_{t,t+4}^{'}\right)$ with $\varepsilon_{t,t+4} = \left(\varepsilon_{t,t+4}^{2},...,\varepsilon_{t,t+4}^{10}\right)$, is constant up to time t, $\widehat{\sum}_{t,t+4} = \widehat{\sum}$. Although simple, this approach works quite well in practice (Thornton and Valente, 2012).

The economic value of predictors is assessed by using power utility in wealth as in Campbell and Viceira (2002).[5] The average utility of the investor is then given by

$$\overline{U}(\cdot) = \frac{1}{T-R}\sum_{t=R}^{T}\left[r_{t,t+4}^{p} - \frac{(\varsigma-1)}{2}\mathbf{w}_{\mathbf{t}}^{'}\sum_{t,t+4}\mathbf{w}_{\mathbf{t}}\right] \tag{2.6}$$

where $r_{t,t+4}^{p} = y_t^1 + \mathbf{w}_{\mathbf{t}}^{'}\mu_{t,t+4}$ is the log return on the bond portfolio and ς denotes investor's degree of relative risk aversion (RRA), which plays the same role as ρ in (15) and is set such that $\rho = \varsigma - 1$. R is the length of the initial window used for estimating parameters and T the total sample size.

As in Campbell and Thompson (2008) and Rapash, Strauss and Zhou (2010), the economic value of models is obtained by evaluating the average utility gain, UG, of investing in a portfolio constructed using model j relative to a portfolio built using the constant model, that is,

$$UG^{j} = \sum_{t=R}^{T}\left[r_{j,t,t+4}^{p} - \frac{(\varsigma-1)}{2}\mathbf{w}_{\mathbf{j,t}}^{'}\sum_{j,t,t+4}\mathbf{w}_{\mathbf{j,t}}\right] - \sum_{t=R}^{T}\left[r_{c,t,t+4}^{p} - \frac{(\varsigma-1)}{2}\mathbf{w}_{\mathbf{c,t}}^{'}\sum_{c,t,t+4}\mathbf{w}_{\mathbf{c,t}}\right] \tag{2.7}$$

where $j = SMRF, CP, LN, SMRF+CP, SMRF+LN, CP+LN, SMRF+CP+LN$ and c refers to the constant model. The utility gain can be interpreted as the portfolio management fee that an investor would be willing to pay to have access to the additional information available in a predictive regression model relative to the information in the historical average of the bond premium alone.

Another frequently used measure of performance in mean-variance analysis is the Sharpe ratio. In this paper, I follow Sharpe (1994) and compute a modified version of the original ratio known as Information ratio (IR), which is defined as the ratio of portfolio returns above the benchmark (the constant) to its volatility, that is,

[5] Results using quadratic utility are not qualitatively different.

$$IR^j = \frac{1}{T-R}\sum_{t=R}^{T} \frac{\left(r^p_{j,t,t+4} - r^p_{c,t,t+4}\right)}{\sqrt{var\left(r^p_{j,t,t+4} - r^p_{c,t,t+4}\right)}} \tag{2.8}$$

However, while Sharpe ratios are commonly used, they exhibit some drawbacks. Abnormalities like excess kurtosis, outliers or skewness on the distribution of returns can be problematic for the statistic, as standard deviation computation does not have the same effectiveness when these problems exist. Also they can be manipulated in various ways (Goetzmann et al., 2007). As an alternative, I follow Thornton and Valente (2012) and Goetzmann et al. (2007) and also compute a measure of risk-adjusted performance of predictors' based portfolios relative to the constant strategy,

$$GISW^j = \frac{1}{(2-\varsigma)}\left\{log\left[\frac{1}{T-R}\sum_{t=R}^{T}\left(\frac{r^p_{j,t,t+4}}{1+y^1_{t+4}}\right)^{2-\varsigma}\right]\right\} - log\left[\frac{1}{T-R}\sum_{t=R}^{T}\left(\frac{r^p_{c,t,t+4}}{1+y^1_{t+4}}\right)^{2-\varsigma}\right] \tag{2.9}$$

2.4 Empirical results

2.4.1 Data and preliminary results

Yields are obtained from the Fama-Bliss data set for maturities up to five-years and from the Gürkaynak, Sack, and Wright (2007) (GSW) data set for maturities from six to ten years. The sample spans the period 1968:Q4 - 2011:Q4.[6] As data from the Survey of Professional Forecasters and the Michigan Survey that are needed to construct the SMRF factor of De Rezende (2014) are released by the middle of the quarter, I use yields for the end of the second month of each quarter.[7]

Table 2.1 - Panel B shows descriptive statistics for the 1-year yield and the 2-year to 10-year excess bond returns. Notice that the average term structure of excess returns is positively sloped and standard deviations increase with maturities, suggesting that investors require higher premia for investing in longer (riskier) bonds. In addition, returns are negatively skewed and exhibit positive excess kurtosis. The Robust Jarque-Bera test of normality, however, does not reject the null hypothesis

[6]For the period 1968Q4 - 1971Q3 yields for maturities from eight to ten years were obtained by extrapolating the Gürkaynak, Sack and Wright (2007) data set using Svensson's (1997) parametrization and the estimated parameters provided by the authors.

[7]The Michigan Survey is conducted at a monthly frequency beginning from January 1978.

of normality for excess returns, which also show high persistence as indicated by the first order autocorrelation coefficients.

The preliminary exploration of the data is completed by performing in-sample predictions over the sample 1968–2011. For this exercise, the SBIC selected the first, fourth and sixth macro risk factors, and the first, second and sixth macro factors, forming the vectors $MRF_t = (MRF_{1t}, MRF_{4t}, MRF_{6t})'$ and $MRF_t = (F_{1t}, F_{2t}, F_{6t})'$, respectively. Results are shown in Table 2.2 and Table 2.3. As documented by Cochrane and Piazzesi (2005, 2008), CP captures a large portion of variation in expected excess returns with R^2s ranging from 0.21 to 0.32. The LN and the SMRF factors also show high explanatory power with R^2s ranging from 0.17 to 0.21 and 0.21 to 0.31, respectively, and highly significant estimates.

When CP regressions are augmented with SMRF, notice that both variables reveal strong statistically significant predictive power, with R^2s increasing substantially and reaching 0.40 for the 10-year return. As documented by Ludvigson and Ng (2009), including LN to CP regressions increase R^2s to levels ranging from 0.35 to 0.41. Regressions with SMRF and LN together also reveal a very good fit with R^2s increasing substantially when SMRF or LN are included as additional predictors. For example, R^2s increase from 0.17 to 0.38 for the 10-year return when adding SMRF to the LN regression, whereas we observe an increase from 0.21 to 0.33 for the 2-year return when adding LN.

I also test regressions that include all three single factors jointly. In this case, R^2s are even higher and range from 0.38 to 0.46. Notice from Table 2.3 that R^2s increase whenever any additional factor is added and statistical significance is, in general, maintained. These results suggest that the three factors capture somewhat independent information about bond risk premia variation.

2.4.2 Statistical predictability

We now verify whether the in-sample predictability verified above holds in an out-of-sample setting. As observed from Table 2.4 - Panel A, beating the constant model is not an easy task. All predictors, except LN, fail to forecast rx^n for shorter maturities. Results change as we move our attention to longer maturities though. While CP continues to perform poorly, the LN and SMRF factors show impressive results, with higher accuracy being achieved by SMRF. In this case, R^2_{oos}s turn positive from rx^5 (rx^3) in the sample period of 1990-2011 (1990-2007) and reach 0.292 (0.278) for the 10-year bond return. Notice also that LN and SMRF outperform the constant model with high statistical significance according to both CW and MC tests.

Combining different predictors also improves regressions' predictive power substantially. Models SMRF+CP+LN and SMRF+LN, in particular, are the most successful ones and generate R^2_{oos}s ranging from 0.20 to 0.36 and from 0.22 to 0.26, respectively, indicating that the expectations hypothesis is also rejected in an out-of-sample setting.

Table 2.4 - Panel B shows results of tests verifying whether the higher accuracy obtained from including SMRF in predictive regressions is statistically significant. Notice that augmenting regressions with SMRF improves predictability remarkably with highly statistically significant results according to both CW and MC tests. Notice that improvements are quite large for longer maturities with differences in R^2_{oos}s reaching 0.296 and 0.319 for CP and LN regressions, respectively. This is also true for regressions that include all the three predictors together. In this case, notice that R^2_{oos}s are positive from the 5-year maturity and are highly statistically significant.

In order to check the stability of these results over time Figure 2.1 - Panel A shows R^2_{oos}s computed recursively against the constant model of no-predictability. Notice that, the levels of predictive power for most models show high stability and statistical significance over the full period of evaluation. Some exceptions are found though. For instance, the models which include the CP factor among predictors show decreasing levels of predictive power. The LN model performed not so well up to the early 2000's, from when R^2_{oos}s start increasing substantially. SMRF regressions as well as models that include SMRF as an additional predictor, on the other hand, show quite high R^2_{oos}s over the full period of evaluation. In general, the most successful model is SMRF+CP+LN.

These results reveal that the in-sample predictability can be somewhat replicated in an out-of-sample framework. The LN and SMRF factors show high degrees of predictive power and predictability is especially strong when combining different predictors, which are shown to capture somewhat different information about bond risk premia variation. An exception is the CP factor alone. Although CP shows high predictive power in-sample, this is not verified out-of-sample. This result is in line with the findings of Thornton and Valente (2012) who show, from an asset allocation perspective, that forward-rate models do not successfully beat a constant model of no-predictability. My results, on the other hand, reveal that this result holds for forward-rate models only. Factors based on macroeconomic variables as well as on indicators of macroeconomic risks are able to generate out-of-sample predictability. This result holds when models are evaluated economically, as discussed in the next subsection.

2.4.3 Economic predictability

Results for the asset allocation assessment are shown in Table 2.5. As observed, predictors that provide consistent economic value in terms of utility gains relative to the constant model are SMRF and LN only, or models in which one of these variables is present. SMRF alone performs quite well and is superior to LN when $\varsigma = 3$, with utility gains of 2.6 compared 2.46. Utility gains of an investor who had relied on the CP model is negative though, a result that is consistent with the findings of Thornton and Valente (2012).[8] Notice also that utility gains always increase when augmenting regressions with SMRF or LN. Although the model SMRF+CP+LN performs quite well, the model that performs best is SMRF+LN with utility gains achieving 3.71 and 3.23, depending on the degree of RRA.

When we analyze the risk-adjusted measure of portfolio performance, GISW, results remain in line with those obtained with the utility gain approach, except for the fact that the portfolio formed with the SMRF strategy performs better than the one formed using LN, independently of the degree of RRA. Positive GISWs are not obtained when augmenting CP regressions with SMRF, even though we see some big improvements. Augmenting it with LN, on the other hand, delivers GISW. As observed in the utility gain approach, the model that performs best is SMRF+LN. Results with Information ratios are similar to the ones obtained with GISW, except for the fact that the LN regression is superior than all the others

The stability of these results over time is verified in Figure 2.1 - Panel B, which shows utility gains computed recursively. For the sample 1990Q1:2011Q4 results show that the economic predictability of models that include SMRF as predictor are quite stable. For instance, the SMRF portfolio delivers high and more stable utility gains when compared to other predictors. Utility gains for portfolios that include LN are less stable and are downward trended. Notice, however, that the model SMRF+LN is the one that generates the highest levels of recursive utility gains to the investor, as also suggested by Table 2.5.

To sum up, we observe that SMRF and LN regressions are able to generate quite high utility gains and risk-adjusted measures of portfolio performance. Furthermore, augmenting regressions with both SMRF and LN always improves their respective portfolios' performances. In addition, I find that the portfolio based on the SMRF+LN specification is the one that performs best. These results confirm the evidence of out-of-sample bond return predictability described in the subsection 2.4.2.

[8] Notice that, differently from Thornton and Valente (2012), the values for Util. Gains and GISW shown in Table 2.5 are in percentage points.

2.4.4 Real-time macro data

It is well known that macroeconomic data are subject to publication delays and revisions, meaning that the information set available to market participants at the time forecasts are made is not the necessarily the same implied by the use of final revised macroeconomic data. This raises the question of whether the predictive information contained in ex ante macroeconomic risks is due to the use of final revised data.[9] In order to examine this issue, in this subsection, I re-assess the predictive power of SMRF constructed on the basis of a truly real-time exercise, where I consider only data available at the time forecasts are generated.

For this analysis, I collected *advance* real-time macro data vintages from the Federal Reserve Bank of Philadelphia database. While they are subject to greater measurement error, the use of *advance* vintages makes more sense, as the other variables used in this study are available by the end of the second month of each quarter, which is exactly the month the Bureau of Economic Analysis (BEA) makes its *advance* estimates available to the public. Due to the unavailability of a large macroeconomic data set in real-time, I only provide results for the SMRF and CP factors. In order to take into account the misalignments between macroeconomic and financial data that may occur due to publication lags in the former, I use the "jumping-off point" strategy of Faust and Wright (2012) and treat the SPF current-quarter forecast (nowcast) as the last observation available for macro variables. Besides permitting the alignment of data and, consequently, an easier forecast comparison among models with different data structures, Faust and Wright (forthcoming) show that using the "jumping-off point" strategy can improve out-of-sample inflation forecasts generated from a large number of econometric models. For this analysis, I take their findings as given and apply their approach to the other macro variables as well.

Results are shown in Table 2.6. Panels A and B provide results for the statistical exercise and results for the economic evaluation are provided by Panel C. Although some of the forecasting power is lost when considering real-time data, SMRF still shows high predictive power, in particular, for longer maturities. Regardless of the evaluation period, notice that positive and highly statistically significant R^2_{oos}s are obtained for the 5- to 10-year excess returns when compared to the constant model. In addition, observe that although R^2_{oos}s obtained from SMRF regressions are nega-

[9] In a recent paper, Ghysels, Horan and Moench (2012) re-assessed the predictive power of the macro factors of Ludvigson and Ng (2009) using a real-time large macroeconomic data set. Although the time period and variables entering their data set are not the same as in Ludvigson and Ng (2009), the authors document that the additional predictive information of factors extracted from revised macroeconomic data largely disappears in a truly real-time out-of-sample forecasting exercise.

tive for the 2- and 3-year maturities, the MSPE-adjusted test of CW still indicates statistically significant results. When we analyze the improvements obtained from augmenting CP regressions with SMRF observe that, as in the previous analysis with final revised data, R^2_{oos}s are always positive and highly significant.

When we move our attention to the economic evaluation, results are very good and, in general, slightly superior to the ones obtained with final revised data. SMRF regressions generate utility gains ranging from 3% to 3.56% when compared to the constant model. Information ratios are bit lower than before, but still high and ranging from 0.41 to 0.46. When we analyze the Goetzmann et al. (2007) measure, GISW, results are also quite high and in line with those obtained with the utility approach. These results indicate that an investor who had relied on the SMRF regression to invest in a portfolio of US government bonds during the period from 1990Q1 to 2011Q4 (2007Q4) would had obtained high utility gains and risk-adjusted returns when compared to the historical average. In addition, notice that augmenting CP regressions with SMRF improves their economic performances substantially with utility gains and information ratios turning even positive. Recursive R^2_{oos}s and utility gains were also computed and are provided by Figure 2.2. The stability of results over time are comparable to those shown in Figure 2.1.

2.5 Conclusions

I investigate the out-of-sample predictability of bond excess returns using different sources of risk premia variation. I assess the statistical and economic significance of the forecasting ability of empirical models based on forward interest rates, macroeconomic variables and risks in macroeconomic outcomes. Although no predictive power is found for the forward rate factor, results suggest that factors based on macroeconomic variables and risks in macroeconomic outcomes are able to outperform a constant model of no-predictability. The macro risk factor is up to 29% more accurate than the constant model, while the macro factor delivers prediction errors that are up to 12% smaller. More interestingly, the combination of the two (three) factors generates improvements ranging from 15% to 35% (20% to 37%), providing further evidence that the three factors capture somewhat independent information about bond risk premia variation (De Rezende, 2014). These results are confirmed economically using a classical portfolio choice problem in which I analyze the performance of bonds portfolios chosen optimally by an investor who exploits the predictive power of factors.

I also investigate whether the predictability found is maintained when factors are constructed using macroeconomic data available in real-time, which reflect the information available to market participants when predictions are made. Results indicate that most of the predictive power of risks in macroeconomic outcomes is preserved, suggesting that the predictability of Treasury returns is not driven by data revisions.

Table 2.1: Descriptive statistics

Notes: This table shows summary statistics for the 1-year yields and 2-year to 10-year excess bond returns. The statistics reported are the mean, standard deviation, skewness, excess kurtosis, the p-value of a Robust Jarque-Bera (RJB) test for normality and the 1st and 4th sample autocorrelations. Critical values for the RJB test were obtained empirically through 4000 Monte-Carlo simulations. Mean values are in percentage point basis.

	Mean	Std. Dev.	Skewness	Exc. Kurtosis	pv-RJB	ρ_1	ρ_4
y^1	5.840	2.918	0.378	0.434	0.014	0.936	0.794
rx^2	0.595	1.714	−0.244	0.331	0.214	0.754	0.202
rx^3	1.038	3.121	−0.277	0.324	0.233	0.749	0.151
rx^4	1.424	4.324	−0.260	0.385	0.284	0.756	0.138
rx^5	1.548	5.232	−0.183	0.171	0.575	0.742	0.099
rx^6	1.964	6.287	−0.157	0.300	0.624	0.752	0.077
rx^7	2.031	7.144	−0.135	0.486	0.459	0.747	0.056
rx^8	2.168	8.028	−0.115	0.615	0.264	0.746	0.038
rx^9	2.273	8.900	−0.093	0.738	0.149	0.746	0.023
rx^{10}	2.354	9.765	−0.068	0.847	0.076	0.745	0.010

Table 2.2: In-sample predictability - SMRF, CP and LN factors 1

Notes: This table shows the predictive power of the CP, LN and SMRF factors. t-stats computed using Newey-West standard errors with six truncation lags are reported in parentheses and $\overline{R}^2$ refers to the adjusted-R^2. 95% confidence intervals for estimated coefficients and $\overline{R}^2$ s are reported in square brackets. These were obtained through a residual-based block bootstrap with 4999 replications and overlapping blocks of size equal to six. Confidence intervals for coefficients were obtained using an asymptotic refinement based on the t-stat *(percentile-t method)* with bootstrapped t-stats computed using Newey-West standard errors with six lags. Regressions in which LN is included in the set of predictors are estimated using the sample 1968Q4 - 2007Q4.

	SMRF	$\overline{R}^2$	*CP*	$\overline{R}^2$	*LN*	$\overline{R}^2$
	0.241	0.210	0.242	0.214	0.304	0.213
rx^2	(4.664)	–	(5.204)	–	(4.957)	–
	[0.12; 0.36]	[0.01; 0.35]	[0.15; 0.34]	[0.08; 0.37]	[0.16; 0.44]	[0.05; 0.36]
	0.461	0.233	0.468	0.242	0.545	0.208
rx^3	(5.063)	–	(5.338)	–	(5.104)	–
	[0.25; 0.67]	[0.02; 0.36]	[0.28; 0.65]	[0.10; 0.39]	[0.31; 0.77]	[0.05; 0.35]
	0.843	0.278	0.814	0.260	0.863	0.189
rx^5	(6.078)	–	(5.159)	–	(5.595)	–
	[0.52; 1.16]	[0.04; 0.40]	[0.48; 1.13]	[0.12; 0.42]	[0.53; 1.21]	[0.04; 0.33]
	1.194	0.299	1.185	0.296	1.182	0.188
rx^7	(6.470)	–	(5.669)	–	(5.579)	–
	[0.78; 1.61]	[0.05; 0.43]	[0.74; 1.62]	[0.15; 0.45]	[0.73; 1.63]	[0.03; 0.33]
	1.667	0.313	1.686	0.321	1.559	0.172
rx^{10}	(6.423)	–	(6.018)	–	(5.012)	–
	[1.08; 2.22]	[0.06; 0.45]	[1.08; 2.25]	[0.17; 0.48]	[0.90; 2.20]	[0.03; 0.30]

Table 2.3: In-sample predictability - SMRF, CP and LN factors 2

Notes: As in Table 2.2.

	SMRF	*CP*	$\overline{R}^2$	*SMRF*	*LN*	$\overline{R}^2$	*CP*	*LN*	$\overline{R}^2$	*SMRF*	*CP*	*LN*	$\overline{R}^2$
	0.152	0.153	0.271	0.192	0.187	0.333	0.195	0.190	0.354	0.111	0.139	0.155	0.383
rx^2	(2.924)	(3.395)	–	(4.245)	(3.098)	–	(3.897)	(2.704)	–	(2.146)	(2.559)	(2.276)	–
	[0.03;0.27]	[0.06;0.25]	[0.13;0.63]	[0.08;0.30]	[0.05;0.33]	[0.07;0.45]	[0.08;0.31]	[0.03;0.35]	[0.19;0.50]	[−0.02;0.23]	[0.01;0.27]	[0.00;0.032]	[0.13;0.67]
	0.288	0.300	0.303	0.371	0.318	0.346	0.384	0.321	0.375	0.210	0.277	0.255	0.406
rx^3	(2.984)	(3.407)	–	(4.409)	(2.870)	–	(4.353)	(2.770)	–	(2.111)	(2.901)	(2.211)	–
	[0.07;0.49]	[0.12;0.49]	[0.14;0.64]	[0.17;0.57]	[0.05;0.58]	[0.08;0.46]	[0.18;0.58]	[0.06;0.59]	[0.22;0.52]	[−0.03;0.45]	[0.04;0.50]	[−0.02;0.53]	[0.15;0.67]
	0.562	0.486	0.344	0.693	0.438	0.363	0.692	0.459	0.384	0.415	0.479	0.329	0.428
rx^5	(3.970)	(3.355)	–	(4.827)	(2.862)	–	(4.299)	(2.989)	–	(2.723)	(3.051)	(2.138)	–
	[0.24;0.88]	[0.18;0.79]	[0.15;0.64]	[0.36;1.03]	[0.09;0.81]	[0.10;0.48]	[0.31;1.07]	[0.10;0.80]	[0.23;0.52]	[0.05;0.78]	[0.09;0.86]	[−0.05;0.72]	[0.16;0.67]
	0.768	0.737	0.380	0.986	0.577	0.375	1.004	0.596	0.406	0.575	0.709	0.416	0.450
rx^7	(3.905)	(3.662)	–	(4.753)	(2.606)	–	(4.506)	(3.126)	–	(2.635)	(3.227)	(2.138)	–
	[0.34;1.21]	[0.32;1.16]	[0.17;0.65]	[0.51;1.46]	[0.09;1.06]	[0.11;0.49]	[0.48;1.52]	[0.17;1.02]	[0.25;0.55]	[0.06;1.11]	[0.21;1.26]	[−0.04;0.89]	[0.17;0.68]
	1.045	1.076	0.403	1.428	0.683	0.379	1.445	0.715	0.411	0.839	1.015	0.453	0.461
rx^{10}	(3.784)	(4.052)	–	(4.773)	(2.091)	–	(4.768)	(2.918)	–	(2.712)	(3.532)	(1.759)	–
	[0.46;1.64]	[0.53;1.62]	[0.18;0.66]	[0.75;2.12]	[−0.02;1.41]	[0.12;0.50]	[0.77;2.12]	[0.16;1.27]	[0.25;0.57]	[0.11;1.58]	[0.32;1.70]	[−0.14;1.04]	[0.18;0.68]

Table 2.4: Out-of-Sample predictability - statistical significance

Notes: Panel A reports R^2_{oos} statistics for predictor based models against a constant. $R^2_{oos} > 0$ indicates outperformance of predictor based models. Panel B reports R^2_{oos} of models with SMRF ("unrestricted") against models without SMRF ("restricted").$R^2_{oos} > 0$ indicates outperformance of models augmented with SMRF ("unrestricted" models). In both panels, ($\star$), ($\star\star$), ($\star\star\star$) indicate statistical significance according to the MSPE-adjusted test of Clark and West (2007) at 10%, 5% and 1%, respectively. ($\dagger$), ($\dagger\dagger$), ($\dagger\dagger\dagger$) indicate statistical significance according to the MSE-F test of McCracken (2007) at 10%, 5% and 1%, respectively. The R^2_{oos} statistic is defined as

$$R^{2,j}_{oos} = 1 - \frac{\Sigma^T_{t=R}\left(rx^n_{t,t+4} - \widehat{rx}^{n,j}_{t,t+4}\right)^2}{\Sigma^T_{t=R}\left(rx^n_{t,t+4} - \widehat{rx}^{n,b}_{t,t+4}\right)^2},$$

where $\widehat{rx}^{n,j}_{t,t+4}$ is a forecast generated from model $j = SMRF, CP, LN, SMRF+CP, SMRF+LN, CP+LN, SMRF+CP+LN$ and $\widehat{rx}^{n,b}_{t,t+4}$ is the forecast generated from the benchmark, with $b = constant, CP, LN, CP+LN$.

Panel A - against constant										
	1990Q1-2011Q4			1990Q1-2007Q4						
	SMRF	*CP*	*SMRF+CP*	*SMRF*	*CP*	*SMRF+CP*	*LN*	*SMRF+LN*	*CP+LN*	*SMRF+CP+LN*
rx^2	$-0.099^{\star}$	-0.373	-0.224	$-0.039^{\star}$	-0.220	-0.107	$0.118^{\star\star\dagger\dagger\dagger}$	$0.152^{\star\star\dagger\dagger\dagger}$	$0.220^{\star\star\star\dagger\dagger\dagger}$	$0.202^{\star\star\star\dagger\dagger\dagger}$
rx^3	$-0.011^{\star\star}$	-0.354	-0.160	$0.014^{\star\star}$	-0.176	$-0.044^{\star\star}$	$0.105^{\star\star\dagger\dagger\dagger}$	$0.173^{\star\star\star\dagger\dagger\dagger}$	$0.221^{\star\star\star\dagger\dagger\dagger}$	$0.217^{\star\star\star\dagger\dagger\dagger}$
rx^5	$0.134^{\star\star\star\dagger\dagger\dagger}$	-0.315	$-0.039^{\star\star}$	$0.133^{\star\star\star\dagger\dagger\dagger}$	$-0.100^{\star}$	$0.081^{\star\star\star\dagger\dagger\dagger}$	$0.089^{\star\star\dagger\dagger\dagger}$	$0.246^{\star\star\star\dagger\dagger\dagger}$	$0.257^{\star\star\star\dagger\dagger\dagger}$	$0.295^{\star\star\star\dagger\dagger\dagger}$
rx^7	$0.219^{\star\star\star\dagger\dagger\dagger}$	-0.343	$-0.004^{\star\star}$	$0.205^{\star\star\star\dagger\dagger\dagger}$	$-0.123^{\star}$	$0.106^{\star\star\star\dagger\dagger\dagger}$	$0.089^{\star\star\dagger\dagger\dagger}$	$0.309^{\star\star\star\dagger\dagger\dagger}$	$0.262^{\star\star\star\dagger\dagger\dagger}$	$0.338^{\star\star\star\dagger\dagger\dagger}$
rx^{10}	$0.292^{\star\star\star\dagger\dagger\dagger}$	-0.353	$0.048^{\star\star\star\dagger\dagger\dagger}$	$0.278^{\star\star\star\dagger\dagger\dagger}$	$-0.134^{\star\star}$	$0.150^{\star\star\star\dagger\dagger\dagger}$	$0.046^{\star\star\dagger\dagger\dagger}$	$0.351^{\star\star\star\dagger\dagger\dagger}$	$0.247^{\star\star\star\dagger\dagger\dagger}$	$0.367^{\star\star\star\dagger\dagger\dagger}$

Panel B - against "restricted" model				
	1990Q1-2011Q4	1990Q1-2007Q4		
	SMRF+CP vs CP	*SMRF+CP vs CP*	*SMRF+LN vs LN*	*SMRF+CP+LN vs CP+LN*
rx^2	$0.108^{\star\star\star\dagger\dagger\dagger}$	$0.092^{\star\star\dagger\dagger\dagger}$	$0.038^{\star\star\dagger\dagger}$	-0.024
rx^3	$0.143^{\star\star\star\dagger\dagger\dagger}$	$0.112^{\star\star\dagger\dagger\dagger}$	$0.076^{\star\star\star\dagger\dagger\dagger}$	$-0.005^{\star}$
rx^5	$0.210^{\star\star\star\dagger\dagger\dagger}$	$0.165^{\star\star\star\dagger\dagger\dagger}$	$0.173^{\star\star\star\dagger\dagger\dagger}$	$0.052^{\star\dagger\dagger\dagger}$
rx^7	$0.252^{\star\star\star\dagger\dagger\dagger}$	$0.204^{\star\star\star\dagger\dagger\dagger}$	$0.242^{\star\star\star\dagger\dagger\dagger}$	$0.102^{\star\star\dagger\dagger\dagger}$
rx^{10}	$0.296^{\star\star\star\dagger\dagger\dagger}$	$0.250^{\star\star\star\dagger\dagger\dagger}$	$0.319^{\star\star\star\dagger\dagger\dagger}$	$0.160^{\star\star\dagger\dagger\dagger}$

Table 2.5: Out-of-Sample predictability - economic significance

Notes: This table reports several statistics used for evaluating the economic significance of excess bond return predictors. Util. Gain (GISW) refers to the utility gain (risk-adjusted performance) of a portfolio constructed using SMRF, CP, LN, SMRF+CP, SMRF+LN, CP+LN, SMRF+CP+LN relative to a portfolio built using the constant model that is associated with the validity of the expectations hypothesis. IR refers to the information ratio.

	1990Q1-2011Q4			1990Q1-2007Q4						
	SMRF	*CP*	*SMRF + CP*	*SMRF*	*CP*	*SMRF + CP*	*LN*	*SMRF + LN*	*CP + LN*	*SMRF + CP + LN*
	$\varsigma = 4$			$\varsigma = 4$						
Util. Gain	3.310	−1.923	1.016	2.883	−1.929	0.872	2.950	3.706	2.867	3.201
IR	0.481	−0.889	−0.030	0.419	−0.887	−0.064	0.521	0.492	0.302	0.429
GISW	2.124	−3.139	−0.395	1.691	−3.107	−0.554	1.091	1.793	0.777	1.123
	$\varsigma = 3$			$\varsigma = 3$						
Util. Gain	3.040	−2.338	0.626	2.604	−2.353	0.477	2.463	3.226	2.267	2.610
IR	0.479	−0.887	−0.031	0.417	−0.887	−0.065	0.516	0.490	0.300	0.427
GISW	2.209	−3.073	−0.305	1.776	−3.057	−0.466	1.107	1.863	0.816	1.158

Table 2.6: Out-of-Sample predictability - real time macro data

Notes: SMRF is constructed using macroeconomic data available in real-time. For details regarding Panel A and Panel B see Notes in Table 2.4. For details regarding Panel C see Notes in Table 2.5. Results for the LN factor are not reported due to the unavailability of a large macroeconomic data set available in real-time that is needed to construct the LN factor.

	Panel A - against constant					
	1990Q1-2011Q4			1990Q1-2007Q4		
	SMRF	*CP*	*SMRF+CP*	*SMRF*	*CP*	*SMRF+CP*
rx^2	−0.231	−0.373	−0.301	−0.179⋆	−0.220	−0.175
rx^3	−0.131⋆⋆	−0.354	−0.237	−0.110⋆⋆	−0.176	−0.108⋆⋆
rx^5	0.035⋆⋆⋆†††	−0.315	−0.108⋆⋆	0.040⋆⋆†††	−0.100⋆	0.042⋆⋆⋆†††
rx^7	0.128⋆⋆⋆†††	−0.343	−0.071⋆⋆	0.126⋆⋆⋆†††	−0.123⋆	0.081⋆⋆⋆†††
rx^{10}	0.198⋆⋆⋆†††	−0.353	−0.016⋆⋆⋆	0.199⋆⋆⋆†††	−0.134⋆⋆	0.135⋆⋆⋆†††

	Panel B - against "restricted" model	
	1990Q1-2011Q4	1990Q1-2007Q4
	SMRF+CP vs CP	*SMRF+CP vs CP*
rx^2	0.052⋆⋆†††	0.036⋆††
rx^3	0.086⋆⋆⋆†††	0.058⋆⋆†††
rx^5	0.157⋆⋆⋆†††	0.130⋆⋆⋆†††
rx^7	0.202⋆⋆⋆†††	0.182⋆⋆⋆†††
rx^{10}	0.249⋆⋆⋆†††	0.237⋆⋆⋆†††

	Panel C - economic predictability					
	1990Q1-2011Q4			1990Q1-2007Q4		
	SMRF	*CP*	*SMRF+CP*	*SMRF*	*CP*	*SMRF+CP*
	$\varsigma = 4$			$\varsigma = 4$		
Util. Gain	3.563	−1.923	1.179	3.307	−1.929	1.324
IR	0.460	−0.889	0.026	0.415	−0.887	0.048
GISW	2.203	−3.139	−0.252	1.894	−3.107	−0.089
	$\varsigma = 3$			$\varsigma = 3$		
Util. Gain	3.278	−2.338	0.856	3.006	−2.353	1.004
IR	0.458	−0.887	0.025	0.412	−0.887	0.047
GISW	2.296	−3.073	−0.105	1.988	−3.057	0.050

Figure 2.1: Recursive estimates of R^2_{oos} and utility gains

Notes: Panel A: charts show recursive R^2_{oos} computed for the period 1995Q1-2011Q4. Asterisks indicate statistical significance at 5% according to the MSPE-adjusted statistic of Clark and West (2007). Panel B: charts show recursive utility gains accrued by an investor investing in a portfolio of US government bonds. R^2_{oos} and utility gains are computed against a constant model of no-predictability.

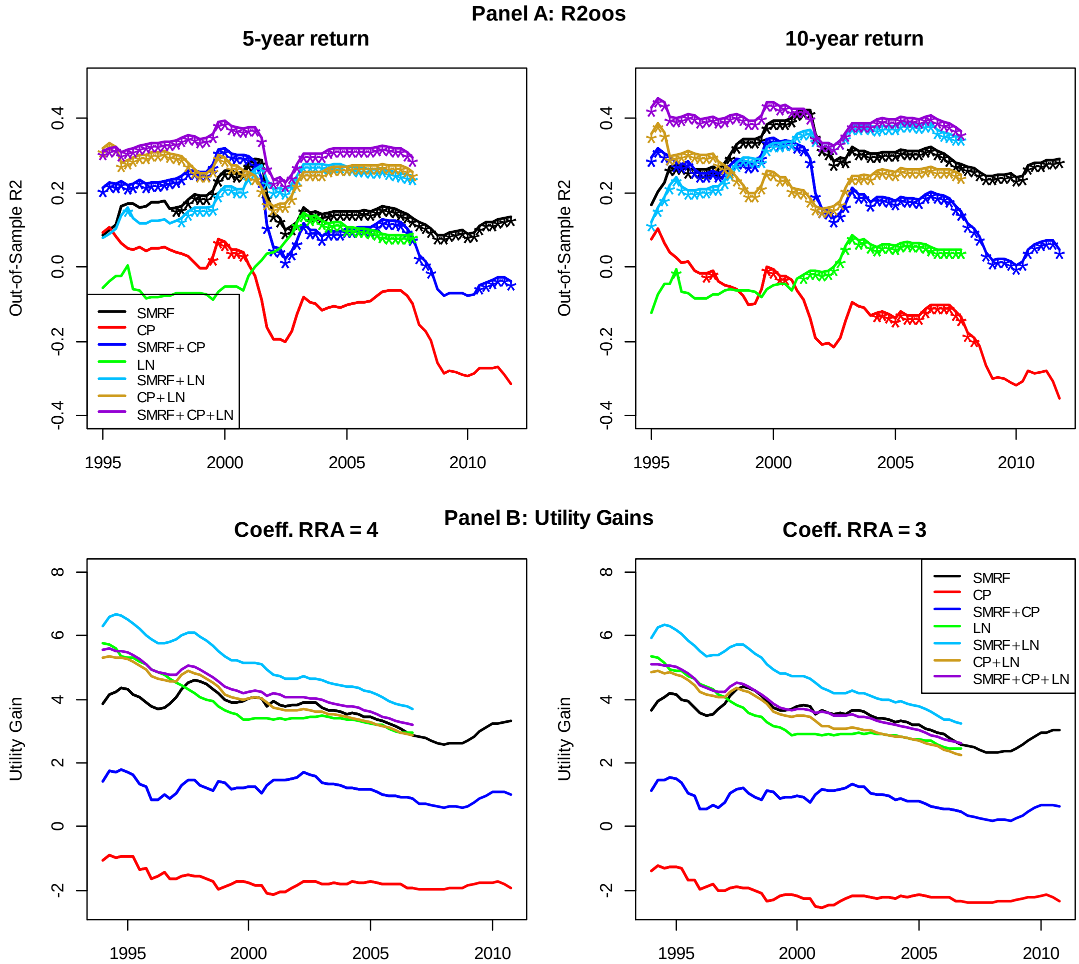
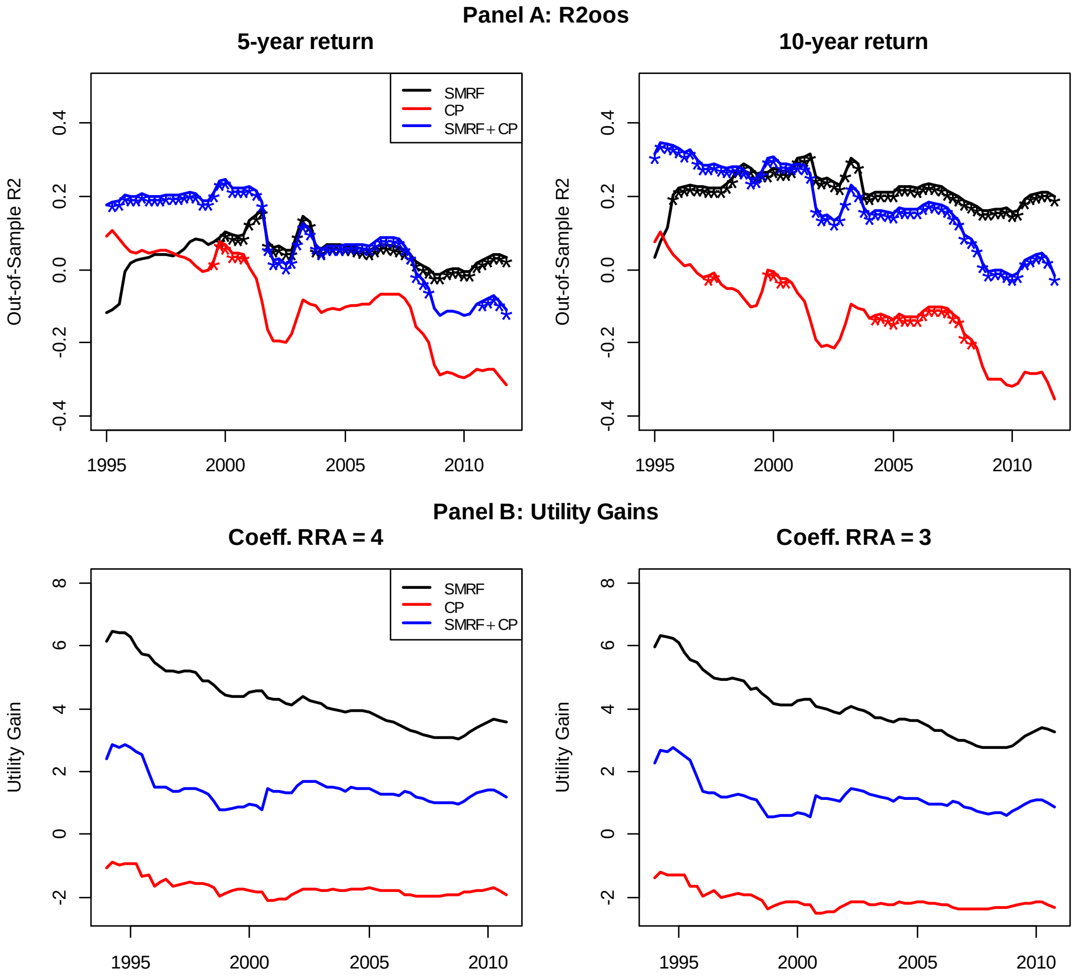


Figure 2.2: Recursive estimates of R^2_{oos} and utility gains in real-time

Notes: Panel A: charts show recursive R^2_{oos} computed for the period 1995Q1-2011Q4 in a fully real-time exercise. Asterisks indicate statistical significance at 5% according to the MSPE-adjusted statistic of Clark and West (2007). Panel B: charts show recursive utility gains accrued by an investor investing in a portfolio of US government bonds. R^2_{oos} and utility gains are computed against a constant model of no-predictability.



Chapter 3

Re-examining the predictive power of the yield curve with quantile regression[1]

Rafael B. De Rezende   Mauro S. Ferreira

ABSTRACT. We use quantile regression to re-examine the predictive power of the term spread with respect to GDP growth and future recessions. The term spread is a better predictor of lower and intermediate conditional GDP growth, confirming its usefulness for predicting economic activity. Changes in the predictive relationship towards longer horizons and structural breaks at upper percentiles suggest the Fed started to respond tougher and in greater advance to inflationary pressures resulted from excessive growth after the mid-1980's. Motivated by these findings we use quantile regression to forecast GDP growth and recessions probabilities in an out-of-sample scheme. Quantile models deliver more accurate forecasts than competitors in both exercises. Superiority against professional forecasters is found for mid and longer horizons, but higher accuracy in shorter horizons is also observed in the period prior the 2008/2009 recession. The predictive power of the yield curve remains.

Keywords: Quantile regression; term spread; forecasting; GDP growth; recessions.
JEL Classifications: E32; E37; E43; E44

3.1 Introduction

The relationship between economic activity and the term structure of interest rates has been widely studied in the last decades. A large number of studies has concen-

[1]We would like to thank Magnus Dahlquist, Jonathan Wright and Enrique Sentana for comments that significantly improved this paper. We also thank seminar participants at the Stockholm School of Economics, the European Meeting of the Econometric Society 2012, the North American Summer Meeting of the Econometric Society 2012 and the Brazilian Time Series and Econometrics School 2013 for further comments and suggestions. Rafael B. De Rezende kindly thank the Swedish Bank Research Foundation (BFI) for financial support.

trated on the predictive power of the difference between long and short yields (the term spread or yield curve slope) regarding future GDP growth (Laurent, 1988); Harvey, 1988, 1989; Estrella and Hardouvelis, 1991), among several others). Others have verified whether the term spread informs about the probabilities of future recessions (Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998; Bernard and Gerlach, 1998; Wright, 2006), among others).

Despite some evidence that movements in bond premia may be linked to economic activity (Ludvigson and Ng, 2009; Joslin, Priebsch and Singleton, forthcoming; Rudebusch and Swanson, 2012), the rationale behind the predictive ability of the term spread rests mainly on the forward looking behaviour of market participants that anticipate future reactions of the central bank (Ang, Piazzesi and Wei, 2006; Rudebusch, Sack and Swanson, 2007). A likely future recession implies the central bank will aggressively reduce interest rates to counteract GDP contractions and disinflationary pressures. By anticipating such scenario, current long rates become smaller than short ones, resulting in a negative spread in the present. Indeed, despite some evidence that parameter instability has weakened the predictive power of the yield curve regarding future output growth since the mid-1980's (Estrella, Rodrigues, and Schich, 2003; Stock and Watson, 2003; Giacomini and Rossi, 2006), the term spread has predicted every recession after the mid-1960s, with only one "false alarm" (an yield curve inversion that was not followed by a recession). This performance justifies its use as one of the most important leading indicators.

In this paper we re-examine the predictive ability of the term spread with respect to GDP growth by studying the entire conditional distribution of the latter variable using quantile regression methods. One of our main concerns is that the term spread - output growth relationship may be subject to asymmetries. Although some papers have explored asymmetries between the term spread and output growth through threshold models (Galbraith and Tkacz, 2000; Duarte, Venetis and Paya, 2005), quantile regression provides information on a different type of asymmetry, as discussed by Koenker and Hallock (2001), Koenker and Xiao (2004, 2006) and Ferreira (2011). Under this method it is possible, for instance, to verify how each conditional percentile of the dependent variable is explained by a set of regressors. In this paper we verify a yet not documented (to our knowledge) asymmetry between our variables of interest: the term spread is able to better predict lower conditional GDP growth, which discourage the use of the term spread as a reasonable variable to predict probability of booms, while confirms its ability to assess the probability of recessions and moderate GDP growth.

We also verify the stability of predictive regressions across percentiles. We apply the structural break tests proposed by Qu (2008) and Oka and Qu (2011), and identify changes happening mostly in the second quarter of 1984, which is consistent with previous research focusing on conditional mean regressions. More interestingly, the breaks were mostly found at higher conditional percentiles, associated to larger GDP growth. This may explain why models focused on providing recessions forecasts, which are commonly associated to negative future GDP growth, have performed decently despite the break usually found in conditional mean models.

Since we verify the existence of (conditional) asymmetries in the GDP growth - term spread relationship, it seems natural to rely on nonstandard probability distributions to computing the probability of future recessions. Contrasting with most of the works in this literature that relies on standard probit regressions (assuming conditional normality of the GDP growth), we estimate these probabilities from quantile regressions. The advantage relies on the flexibility of the approach as we estimate one quantile equation for each percentile of output growth, meaning that probabilities of future recessions can be computed from empirical distributions, which may follow a Gaussian pattern, but can also assume nonusual shapes. Our analyses also suggest that the existence of conditional asymmetries may lead OLS estimators to deliver biased GDP growth point forecasts, suggesting the use of more robust methods such as the median estimators. Given the evidence of instability of regression quantiles in mid and high percentiles we conduct this exercise using two samples - 1955-2011 and 1985-2011 - while we forecast recessions using the full sample only.

In our work, we were also cautious about specification. In particular, we incorporated nonlinearities that have been previously found by studies using threshold regression models. For instance, Galbraith and Tkacz (2000) and Duarte, Venetis and Paya (2005) verified that the positive output growth - term spread relationship weakens if the spread exceeds a certain value. Following Koenker (2005) in the context of quantile regression, we capture this nonlinearity using a quadratic specification, which allows reaching similar conclusion.

Our study thus builds on two main exercises: the study of the conditional distribution of GDP growth accross percentiles and the out-of-sample forecasting of GDP growth and recessions probabilities using quantile regressions. When forecasting GDP growth results show that (i) conditional median models are highly more accurate than similar OLS regressions for almost all horizons and than an autoregressive benchmark for mid and long horizons; (ii) including a quadratic spread variable increase forecasting power; (iii) for the post-1985 data, higher accuracy is obtained for longer horizons. When forecasting recessions results reveal that (i) quantile models are more accurate than probits for a large number of specifications and horizons; (ii) the quadratic spread variable does not necessarily increase accuracy, but a lagged GDP growth variable does. In both exercises we also verify that quantile models are more accurate than professional forecasters for mid and longer horizons, but higher accuracy in shorter horizons was also obtained in the period prior the late 2000's recession. These results indicate the continuing predictive power of the yield curve.

This study is part of a large literature on the predictive power of the slope of the yield curve (Harvey, 1988, 1989; Estrella and Hardouvelis, 1991; Estrella and Mishkin, 1998; Bernard and Gerlach, 1998; Stock and Watson, 2003; Wright, 2006, among several others). Our paper, however, goes beyond by proposing forecasting models based on quantile regressions. We also innovate by providing a more complete analysis of the relationship between the term spread and GDP growth as we look at the entire conditional distribution of the latter variable.

We also add to a debate aiming at verifying a possible change in the Fed's reaction function after the Volcker's presidency. Among several studies pursuing this task, it is worthwhile mentioning the works by Clarida, Galí and Gertler (2000), Sims and Zha (2002), Kim and Nelson (2006), and Boivin, Kiley and Mishkin (2010), who verified that output became less responsive to changes in the fed funds after the mid-1980's. Kim and Nelson (2006) suggest that this change may be due to the Fed being more responsive to inflation during this period. The structural break tests we apply indicate shifts occurring mostly in the second quarter of 1984 and at higher percentiles of the conditional distribution of GDP growth. In addition, we identify changes in the predictive relationship towards longer horizons. These results are consistent with the view that the Fed started to respond tougher and in greater advance to inflationary pressures resulted from excessive growth after the mid-1980's

Another related strand of research aims at measuring macroeconomic risks. Our approach allows computing probabilities that GDP growth will fall in certain intervals, from which we measure recession risks. In a similar spirit, Kitsul and Wright (2012) rely on CPI based options to construct probability densities for inflation and use them to measure deflation and high inflation risks. Gaglianone and Lima (2012) use quantile regression to construct density forecasts for macro variables and use them to estimate risks of high unemployment rates. Christensen, Lopez and Rudebusch (2011) rely on Treasury Inflation Protected Securities to measure deflation probabilities.

The remaining of the paper is organized as follows. In the next section we revisit the term spread - GDP growth relationship using quantile regressions; the third section presents the out-of-sample framework; the fourth section describes the forecasting results; and the last section concludes.

3.2 The term spread - GDP growth relationship

In this section we revisit the term spread - GDP growth relationship. Our main variable of interest is the quarterly (annualized) real GDP growth between quarters t and $t-1$, which we denote by y_t and is computed according to

$$y_t = 400 \times [log(GDP_t) - log(GDP_{t-1})] \tag{3.1}$$

Following the literature, real GDP growth can be predicted by estimating the following equation by OLS

$$y_t = \alpha_0 + \alpha' X_{t-h} + u_t \tag{3.2}$$

where u is a random shock, α is a $k \times 1$ vector of coefficients and X_{t-h} is a $k \times 1$ vector of covariates, which include the term spread, but may also include other variables such as lagged GDP growth that is commonly used to deal with autocorrelation in y_t.[2]

In our specifications we also allow for a quadratic term spread variable. Its inclusion is motivated by previous findings, such as those of Galbraith and Tkacz (2000) and Venetis, Paya and Peel (2003), revealing that the intensity of the positive relation between GDP growth and the term spread diminishes if the former crosses a certain threshold. The quadratic spread variable is used to approximate such nonlinearity, as suggested by Koenker (2005) in a general context of quantile regression. When modeling y_t, we thus allow for four different specifications: (i) $X_{t-h} = (Spread_{t-h})$, implying in $\alpha = \alpha_1$, (ii) $X_{t-h} = (Spread_{t-h}, Spread^2_{t-h})$, so that $\alpha = (\alpha_1, \alpha_2)$, (iii) $X_{t-h} = (Spread_{t-h}, y_{t-h})$ with $\alpha = (\alpha_1, \alpha_3)$, and (iv) a specification with all covariates included, $X_{t-h} = (Spread_{t-h}, Spread^2_{t-h}, y_{t-h})$.[3] To be consistent with threshold models, one should expect $\alpha_1 > 0$ and $\alpha_2 < 0$, resulting in a concave relation between the spread and future GDP growth.

While OLS regression models are often used to approximate the conditional

[2] Stock and Watson (2003), Hamilton and Kim (2002) and Ang, Piazzesi and Wei (2006) are among some authors that also use lagged GDP growth in their forecasting equations.

[3] We also tested specifications with the Federal Funds rate or the 3-month Treasury Bill rate as regressors. The estimated coefficients on both variables showed very small values and no significance was achieved for any forecast horizon. Moreover, the regressions showed poor performance in the out-of-sample forecasting exercise.

mean of y_t, quantile regression models are suitable to approximate its conditional quantiles. Following the linear specification in (3.2) and denoting $Q_u(\tau)$ as the τ-th quantile of u, the τ-th quantile of y_t, conditioned on X_{t-h}, can be obtained by

$$Q_{y_t}(\tau|X_{t-h}) = \alpha_0(\tau) + \alpha(\tau)' X_{t-h} \tag{3.3}$$

where, $\alpha_0(\tau) = \alpha_0 + Q_u(\tau)$ and $\alpha(\tau)$ can be estimated according to Koenker and Basset (1978). In particular, the conditional median of y_t is approximated after setting $\tau = 0.5$.

3.2.1 Mean and median regressions

We collect yield data from the FRED database. As in the extant literature the term spread series is computed as the difference between the 10-year Treasury Bond and the 3-month Treasury Bill interest rates. The quarterly frequency is obtained by averaging the monthly spread over each quarter. For the in-sample analyses performed in this section we use real GDP data collected from the Fed of Philadelphia. More specifically, we use the vintage that became available on the third quarter of 2011 and we decided to work with two samples normally used in this literature: a full sample ranging from 1955Q1 to 2011Q2, and a subsample ranging from 1985Q1 to 2011Q2. This subsample coincides with the Great Moderation period and is justified by several works reporting a breakdown in the predictive ability of the term spread after the mid-1980s (see Estrella, Rodrigues, and Schich, 2003; Stock and Watson, 2003; Giacomini and Rossi, 2006).

In Table 3.1 we report predictive results for OLS and QR(0.5) regressions. In order to save space, we only show results for specifications (iii) and (iv), which incorporate y_{t-h} as covariate. As commonly found in the literature the spread keeps a strong and positive relation with future GDP growth. Estimates are generally highly significant and indicate that the spread is able to predict GDP growth for a number of horizons over the two samples considered. Interestingly, the significance of the quadratic spread variable indicates the existence of a nonlinear relationship between our two variables of interest. Similar to Galbraith and Tkacz (2000) and Venetis, Paya and Peel (2003), we verify a concave relationship with $\alpha_1 > 0$ and $\alpha_2 < 0$, indicating that its intensity diminishes after a certain level of the spread is crossed. Notice also that the quadratic spread regressions show superior fitting than their counterparts as indicated by the smaller Akaike Information Criterion (AIC) statistics.

Results also indicate the ability of the spread to signal about future economic activity with greater advance after the mid-1980's. While significant spread estimates are observed from $h = 1$ to $h = 6$ in the full sample, significance is verified from $h = 4$ in the subsample. This result is robust regardless the estimation method and specification. In addition, coefficients linking the spread to future GDP growth are larger for horizons up to four quarters ahead in the full sample, while the largests coefficients are found for longer horizons when using post-1985 data. These results indicate the continuing predictive power of the yield curve.[4]

3.2.2 A more complete analysis: the conditional distribution of GDP growth

Analyzing the conditional distribution of GDP growth across percentiles provides a more complete picture of its relation with the term spread. In each box of Figure 3.1 we plot quantile regression estimates using the two samples considered previously. The first row brings estimates of the spread variable in the standard linear model and the second and third rows present estimates in a quadratic model. In both cases we allowed for the presence of lagged growth among the regressors. Figure 3.2 shows the goodness of fit of models by presenting the adjusted $R1(\tau)$ statistics of Koenker and Machado (1999) computed across percentiles.[5]

Results in Figure 3.1 reinforce previous findings. The term spread keeps a strong relation with future output growth and this is shown over almost the whole conditional distribution of the latter variable. In addition, for horizons up to $h = 4$, we verify that the spread is significantly related to larger ranges of the conditional distribution of GDP growth when using the full sample. Contrary, for the period 1985-2011, statistical significance is largely observed across percentiles from $h = 4$. These patterns occur in the linear and in the quadratic models, but are stronger once we deal with nonlinearity. Results in Figure 3.2 reinforce these conclusions. In addition, they indicate the ability of the spread to predict a large portion of the conditional distribution of economic activity with greater advance when using the post-1985 data.

Equally interesting is the fact that figures 3.1 and 3.2 also reveal the existence of a strong asymmetric relationship between our variables of interest. While the spread is able to predict negative, low and intermediate GDP growth, it informs little

[4]Haubrich and Dombrosky (1996) and Aguiar-Conraria, Martins and Soares (2012), despite of recognizing a change in the intensity of the relationship, also found information content in the yield spread about future economic activity during the Great Moderation.

[5]The $\overline{R1}(\tau)$ statistic of Koenker and Machado (1999) is similar in nature to R^2, as it informs (for each percentile) the goodness of a fit compared to a regression on the intercept $\alpha_0(\tau)$. The closer $\overline{R1}(\tau)$ is to 1, the better the adjustment.

about high conditional output growth. This result is evident for all specifications and explains why the spread has been a more successful predictor of probability of recessions than of output growth.[6]

In order to verify the stability of regressions across percentiles we also applied the structural break tests proposed by Qu (2008) and Oka and Qu (2011). Table 3.2 shows results for the statistic SQ, which tests the null hypothesis of no break in a specific quantile equation. Table 3.3 shows results for the statistic DQ, which tests the null of no break in multiple quantiles. SQ(1) and DQ(1) are used to test for the absence of only one break, while SQ(2|1) and DQ(2|1) test for the absence of a second break given the first one was found. In both tests we allowed for a maximum of two breaks.[7] When rejecting H_0, the tests also estimate the break date.

Table 3.2 reveals that most of the detected breaks happen at higher percentile equations. The smallest percentile in which breaks were detected was $\tau = 0.35$ when forecasting 4 quarters ahead with the linear model. For the linear model with $h = 8$ and the quadratic with $h = 4$, breaks were found only in regressions for $\tau \geq 0.65$. In the case of the quadratic model and $h = 8$, breaks were verified for $\tau \geq 0.5$. In line with the existing literature, most of the estimated break dates occur in the mid-1980's. Out of the 19 breaks found, 9 were in 1984Q2 and 1 in 1984Q1. The test that searches for breaks in multiple equations detected breaks in 1984Q2.

These results suggest the spread started to signal about future economic activity in a different manner around 1984. Besides indicating a change in the predictive ability of the spread towards longer horizons, this shift happened for higher conditional GDP growth, which is consistent with a market view that the Fed would respond tougher and in greater advance to inflationary pressures resulted from excessive growth. Additionally, the fact that breaks were not observed at lower percentile equations suggest that the conduct of monetary policy was not expected to be modified in the presence of future adverse cycles.

Accordingly, most of the literature focused on changes in the conduct of monetary policy after the Volcker's presidency describes the Fed being tougher to inflation. Boivin, Hiley, and Mishkin (2010) discuss that a more intense and faster reaction of output gap to movements in short-rates was the pattern during 1962-1979, but a less intense, slower and more persistent reaction was observed from 1984 to 2008. Kim and Nelson (2006) suggest that this change may be due to the Fed being more responsive to inflation during this period. Similar conclusions were also reached by

[6]See Wheelock and Wohar (2009) for a survey on the predictive power of the term spread regarding future GDP growth and recessions.

[7]Following the previous analyses, these tests were conducted for specifications with the presence of a lagged GDP growth among covariates.

Clarida, Galí and Gertler (2000), Galí, Lopez-Salido, and Valles (2003), Boivin and Giannoni (2006) and Galí and Gambetti (2009).

Our results also validate the use of the entire sample to estimate probability of future recessions, given the absence of breaks in low percentile equations. This explains why binary models focused on estimating probabilities of future recessions conditioned on current spread have been more stable over time, as found by Estrella, Rodrigues, and Schich (2003) and emphasized by Rudebusch and Williams (2009). It seems, however, that similar exercise to estimate probability of booms, or even intermediate growth, would be more problematic since breaks were identified at central and higher percentiles.

3.3 Out-of-sample forecasting

Since we verify the existence of (conditional) asymmetries in the GDP growth - term spread relationship, it seems natural to rely on nonstandard probability distributions to computing the probability of future recessions. Contrasting with most of the works in this literature that relies on standard probit regressions (assuming conditional normality of the GDP growth), we estimate these probabilities from quantile regressions. The advantage relies on the flexibility of the approach as we estimate one quantile equation for each percentile of output growth, meaning that probabilities of future recessions can be computed from empirical distribution, which may follow a Gaussian pattern, but can also assume nonusual shapes.

Our analyses also suggest that the existence of conditional asymmetries may lead OLS estimates to deliver biased GDP growth point forecasts, suggesting the use of more robust methods such as the median estimators. Given the evidence of instability of regression quantiles in mid and high percentiles we forecast recessions using the full sample only, but conduct output growth predictions using the post-1985 data as well.

We conduct forecasting analyses using real-time and final revised real GDP data collected from the Fed of Philadelphia. For the real-time data we decided to work with the *advance* data, which is the Bureau of Economic Analysis' (BEA) first estimate for the previous quarter. While advance data are subject to greater measurement error, their use makes more sense as we are also willing to address comparisons to forecasts reported by professional forecasters in the Survey of Professional Forecasters (SPF henceforward), who know the advance data when submitting their projections.[8]

[8]This happens because the Fed of Philadelphia sends its survey questionnaires right after the advance

The forecasts presented in this work will be contrasted to the final revised data consisting of the vintage that became available on the third quarter of 2011. At this quarter the BEA released an extensive revision that incorporated important methodological changes in the way it estimates national accounts, so we consider this vintage a good proxy for our final revised data.

3.3.1 Real GDP growth

Real GDP growth point forecasts can be readily computed from prediction models (3.3) and (3.4). More specifically, given the OLS estimates $\hat{\alpha}_0$ and $\hat{\alpha}$, mean forecasts of y_t in $t+h$, conditioned on X_t, can be computed as

$$E\left(y_{t+h}|X_t\right) = \hat{\alpha}_0 + \hat{\alpha}' X_t \tag{3.4}$$

Similarly, median forecasts of y_t in $t+h$ can be obtained using

$$\hat{Q}_{y_{t+h}}\left(0.5|X_t\right) = \hat{\alpha}_0\left(0.5\right) + \hat{\alpha}\left(0.5\right)' X_t \tag{3.5}$$

where $\hat{\alpha}_0(\tau)$ and $\hat{\alpha}(\tau)$ are estimated according to Koenker and Basset (1978). In this paper, OLS and QR(0.5) forecasts are confronted against each other, against mean and median forecasts reported in the SPF, and also against a direct autoregressive model, $\hat{y}_{t+h} = \hat{\alpha}_0 + \hat{\alpha}_1 y_t$.

3.3.2 Recessions

When it comes to forecasting recessions, we first have to define what a recession is in our study. We follow Rudebusch and Williams (2009) and use the following rule linking real GDP changes to recessions: the economy is in recession at quarter t, implying $R_t = 1$, if a negative quarterly real GDP growth ($y_t < 0$) is observed; $R_t = 0$, otherwise. The reason for not using the NBER recession series is that NBER recessions are observed ex-post, meaning that it is not suitable for our real-time exercise. Nevertheless, our rule produces 22 recessionary quarters that match the 34 NBER recession quarters in the period 1955Q1-2011Q2, with only 11 false alarms.[9]

Given each forecast $\hat{Q}_{y_{t+h}}(\tau|X_t)$, $\forall\tau \in [\underline{\tau}, \overline{\tau}]$, the probability of a recession in $t+h$ computed using quantile regressions can then be obtained as follows

estimates become public. Since these questionnaires must be returned within the following two or three weeks, which is before the release of the next revision, professional forecasters' information set, when submitting their projections, include the advance data.

[9] As stated by the NBER (2003): *"The NBER considers real GDP to be the single measure that comes closest to capturing what it means by 'aggregate economic activity'. The [NBER] therefore places considerable weight on real GDP and other output measures"*.

$$\hat{P}_t^{QR}(R_{t+h}=1)=\hat{P}_t^{QR}(y_{t+h}\leq 0)=sup\left\{\tau\in[\underline{\tau},\overline{\tau}]:\hat{Q}_{y_{t+h}}(\tau|X_t)\leq 0\right\}$$

When $\hat{Q}_{y_{t+h}}(\tau|X_t)>0, \forall\tau\in[\underline{\tau},\overline{\tau}]$, the model indicates $\hat{P}_t^{QR}(R_{t+h}=1)=\hat{P}_t^{QR}(y_{t+h}\leq 0)=0$.[10] In order to estimate these probabilities, we set $\tau\in[\underline{\tau}=0.0025,\overline{\tau}=0.9975]$ and allow for our four sets of covariates. In addition, we avoid crossings of $\hat{Q}_{y_{t+h}}(\tau|X_t)$ across percentiles through the "rearrangement" procedure of Chernozhukov, Fernandez-Val and Galichon (2010).

As it is the standard procedure in this literature, we also forecast recession probabilities from Probit models as below

$$\hat{P}_t^{PR}(R_{t+h}=1)=\Phi\left(\hat{\theta}_0+\hat{\theta}'X_t\right) \tag{3.6}$$

where Φ is the standard normal cumulative distribution function.

QR forecasts are compared to Probit forecasts with same specifications, SPF recession forecasts and to a simple probit model, $\hat{P}_t^{PR}(R_{t+h}=1)=\Phi\left(\hat{\theta}_0+\hat{\theta}_1 y_t\right)$.

3.3.3 Assessing the accuracy of forecasts

Real GDP growth point forecasts are compared using the Root Mean Squared Forecast Error (RMSFE) measure. When comparing recessions probabilities we follow Diebold and Rudebusch (1989) and employ two different measures: the Root Quadratic Probability Score (RQPS) and the Log Probability Score (LPS). Letting T be the number of out-of-sample forecasts, and i a particular model, these two last score measures are respectively computed as

$$RQPS^i=\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left[\hat{P}_t^i(R_{t+h}=1)-R_{t+h}\right]^2}$$

$$LPS^i=-\frac{1}{T}\sum_{t=1}^{T}\left[(1-R_{t+h})\,ln\left(1-\hat{P}_t^i(R_{t+h}=1)\right)+R_{t+h}ln\left(\hat{P}_t^i(R_{t+h}=1)\right)\right]$$

A more rigorous comparison between each model, however, can be assessed by relying on the Harvey, Leybourne and Newbold (1997) test (HLN henceforward), which allows verifying if the difference between the average forecast errors of two competing models is statistically significant. HLN test is based on a modification

[10]Here we define any event in which $y_t\leqslant 0$ as a recession. The use of the operator $\leq$ instead of $<$ does not really make any difference in our results.

of the Diebold and Mariano (DM, 1995) statistic that corrects for its tendency to be over-sized in finite samples, in particular. The adjustment consists of estimating the variance of the mean loss differentials using the rectangular lag window with a lag truncation parameter equal to $h-1$ and then multiplying the DM statistic by $\sqrt{\left(T+1-2h+T^{-1}h\left(h-1\right)\right)/T}$. Through a Monte Carlo experiment, the authors show that the modified statistic performs considerably better than the original one, providing important size corrections.

Although the HLN test was originally designed for non-nested models, Clark and McCracken (2012), when comparing the use of alternative HAC estimators in nested models, find that comparing the HLN test statistic to standard normal critical values delivers a test that has size fairly close to the nominal level in both population and finite-samples. In this study we thus follow Harvey, Leybourne and Newbold (1997) approach.[11] As suggested by these authors, we compare the HLN statistic to critical values from a t_{T-1} distribution. Our assessments indicated that the t_{T-1} distribution yields results that are a bit more conservative also when comparing nested models.

In order to pursue such analysis, we rely on $D_{t+h}^{i,j}$, which is the loss differentials between models i and j of forecasting errors made at $t+h$. We compute the sample loss differentials, according to RMSFE, RQPS and LPS, using the following equations:

$$\hat{D}_{t+h}^{i,j}\left(MSFE\right)=\left(\hat{y}_{t+h}^{i}-y_{t+h}\right)^{2}-\left(\hat{y}_{t+h}^{j}-y_{t+h}\right)^{2}$$

$$\hat{D}_{t+h}^{i,j}\left(QPS\right)=\left(\hat{P}_{t}^{i}\left(R_{t+h}=1\right)-R_{t+h}\right)^{2}-\left(\hat{P}_{t}^{j}\left(R_{t+h}=1\right)-R_{t+h}\right)^{2}$$

$$\begin{aligned}\hat{D}_{t+h}^{i,j}\left(LPS\right)=-\left(1-R_{t+h}\right)\left[ln\left(1-\hat{P}_{t}^{i}\left(R_{t+h}=1\right)\right)-ln\left(1-\hat{P}_{t}^{j}\left(R_{t+h}=1\right)\right)\right]-\\-R_{t+h}\left[ln\left(\hat{P}_{t}^{i}\left(R_{t+h}=1\right)\right)-ln\left(\hat{P}_{t}^{j}\left(R_{t+h}=1\right)\right)\right]\end{aligned}$$

Models i and j perform equally if the mean loss differential is zero; $H_0:E\left(D_t^{i,j}\right)=0,\ \forall t$. As we are interested in verifying whether model i delivers more accurate out of sample forecasts than model j, we state the alternative as $H_1:E\left(D_t^{i,j}\right)<0,\ \forall t$.[12]

The tests just described are not suitable for real-time data though. Clark and

[11] This test is also used by Faust and Wright (forthcoming) when forecasting inflation.

[12] The HLN test is not available for a RMSFE (RQPS) loss function, so we report the results from the MSFE (QPS) version of the test.

McCracken (2009) recently proposed an alternative test of equal predictive accuracy for real-time data, the construction of which requires further assumptions on the nature of the data revisions and evidence that these assumptions are met in the real-time data. This test is not designed for recursive forecasts and could not be applied in our context even if our real-time data satisfied the assumptions of Clark and McCracken (2009). As an alternative we apply the accuracy tests on ex-post (our *final*) revised data, but we also report tests' results for the real-time forecasts.

3.4 Forecasting results

Tables 3.4, 3.5, and 3.6 summarize our findings, where we report the statistical significance of the several tests we conduct. Asterisks indicate superiority of median forecasts against those carried by the OLS estimator. The symbol ∘ indicates statistical significance against a common benchmark that follows an AR structure. Superiority of quadratic specification against similar linear model is indicated by †, and • shows if the model analyzed delivered superior forecasts than those of SPF.

3.4.1 GDP Growth

Full sample

Table 3.4 shows results using the full sample. We report the ratio of the RMSFE of each model relative to the benchmark. The forecasting exercise is carried according to the recursive method, meaning that each forecast model is re-estimated after incorporating the newest data. For real-time data, we used the most recent vintage available in each recursion. The first estimation window ranges from 1955Q1 to 1985Q4 when using final revised data, and up to 1986Q1 when relying on real-time data.

Table 3.4 reveals several interesting results. First, the spread shows higher accuracy than the benchmark autoregressive model. This is true for horizons from $h = 4$, but results are stronger for $h = 5$, when statistical significance is achieved for median regressions only. More interestingly, however, is the fact that median regressions are highly superior than OLS regressions for almost all the horizons and regression specifications. Notice that statistiscal significance is achieved for a large number of horizons - from $h = 1$ to $h = 5$, more specifically - which shows the advantage of relying on median models when performing GDP growth point forecasts.

Another appealing result we found is that adding the quadratic spread variable to both mean and median regressions improves forecasting performance of models substantially. Notice that for $h = 1,...,6$ quadratic regressions are always superior to its counterparts with statistical superiority being achieved for $h = 1,...,5$ and various models. These results are in line with those reported in Table 3.1 showing statistically significant coefficients on $spread^2$ at short and mid horizons. Also, in accordance to Table 3.1, models which include the lagged GDP growth among covariates perform better than their counterparts for $h = 1,2$. For $h \geq 3$, its inclusion deteriorates prediction power.

Results described above show that the conditional median brings important accuracy gains to traditional GDP growth - term spread models, which suggests that the existence of the conditional asymmetries diagnosed in Section 3.2 may be problematic to traditional OLS predictive regressions. Additionally, taking into account nonlinearities in the form of quadratic regressions reduces forecast errors.

Post-1985 sample

Motivated by our findings in Section 3.2 we also check the forecasting performances of models using the sub-sample 1985Q1-2011Q2. This period, referred to as "Great Moderation", has been characterized by a increased stability of various macroeconomic variables and some authors have attributed the observed decline in the predictive power of the spread to a change in the conduct of monetary policy after the Volcker's presidency (Stock and Watson, 2003; Estrella, Rodrigues, and Schich, 2003; D'Agostino, Giannone and Surico, 2006).[13]

We show post-1985 results in Table 3.5. The first window used to estimate parameters is 1985Q1-1991Q4. We start verifying smaller RMSFE ratios against the benchmark when $h = 4$, but significance was only achieved for $h = 8$, for both estimation methods (QR and OLS). Notice also that the higher accuracy of the conditional median compared to the conditional mean is maintained for almost all horizons and models' specifications, even though statistical significance is only achieved for QR 2 at $h = 8$. In general, the most accurate model is QR 1.

Again, if we relied on the traditional linear OLS, the conclusion would be in accordance to several studies documenting a breakdown in the forecasting performance of the term spread after the mid-1980's (Stock and Watson, 2003; Estrella,

[13]Notice that we consider the post-2007 period part of the "great moderation". As Clark (2009) points out, the higher volatility of several macro variables during the recent crisis was mostly driven by temporary large shocks to oil prices and financial markets, which did not mark the end of the great moderation period. In addition, despite having the Fed switching its main instrument of monetary policy from short-term policy rate to large-scale asset purchases since the crisis, the use of such unconventional monetary policy has not changed its degree of transparency and credibility, which are associated to the macroeconomic stability observed since the mid 1980's.

Rodrigues and Schich, 2003; D'Agostino, Giannone and Surico, 2006 among others). Differently, our results indicate that the forecasting power of the yield curve remains, despite a shift towards longer horizons. This is in accordance with the results in Table 3.1 which also indicate the ability of the spread to predict output growth with greater advance in the 1985-2011 period.

Comparison with professional forecasters

We now compare forecasts of econometric models with those of SPF. For this survey, panelists are asked to provide forecasts for the current quarter and to one to four quarters ahead. Note that at the time answers are reported panelists know the advance BEA report.

Results are shown in the last two columns of tables 3.4 and 3.5. From Table 3.4 we observe that professional forecasters provide very accurate GDP growth forecasts in short horizons, delivering smaller forecasts erros than all the spread based models. Notice, however, that forecasters accuracy diminish monotonically with the forecast horizon. For instance, from $h = 4$ we already observe superiority of median models over both mean and median SPF forecasts. Results are stronger for $h = 5$ with several mean and median specifications showing lower forecast errors. Notice that statistical significance was only achieved by QR 1 and QR 2 models though. When using post-1985 data, median models were able to beat the SPF for $h = 5$, but results are not statistically significant.

We also verified how these forecasts performed over time in a recursive forecasting scheme. Results for the full sample are shown in Figure 3.3, with asterisks indicating statistical significance against professional forecasts at the 10% level. The first observation that comes up is how results contrast to those reported in Table 3.4. Although Table 3.4 reveals that professional forecasters were substantially superior than econometric models in short horizons, Figure 3.3 draws a different picture when we look for the period before the last great recession of 2008-2009. For instance, for $h = 2$ spread based models and forecasters performed comparatively well. When $h = 3$, however, several models were able to outperform SPF forecasts, with statistical significance being achieved by QR 2. In the case of $h = 5$, econometric models resulted in smaller forecast errors since 1990, with statistical significance observed in most quarters.

This robustness check reveals how our conclusions, at least for $h = 2,3,4$, would have changed had we not included the last recession in our sample. In particular, during at least a decade one would perform a better forecast if using either models 2 and 4, and even more so if relying on QR median forecasts.

3.4.2 Probability of recession

Full sample

Results are shown in Table 3.6, where we report RQPS and LPS statistics computed for each model relative to the probability inferred by a simple probit model. As in the GDP growth exercise, the first window used to estimate parameters is 1955Q1-1985Q4. Then, at each recursion models are re-estimated using the new information available.

Table 3.6 reveals several interesting results. First, spread based models show higher accuracy than the benchmark probit from $h = 2$ when using real-time data and from $h = 3$ when using final revised data, with statistical significance being achieved from $h = 3$. More interestingly, however, is the fact that quantile models show higher accuracy than probits for almost all horizons and regression specifications, which shows the advantage of relying on the flexibility of quantile models for forecasting recessions even though statistical superiority is not observed.

Another result we observe is that, although models including the quadratic spread do not deliver superior forecasts, specifications which include lagged GDP growth perform considerably better than its counterparts when $h = 1, 2, 3$. This is especially true for quantile models, as QR 4 shows, in general, higher accuracy than other models for these particular horizons.

A natural question that may arise is the ability of our models to predict the great recession of 2008/2009. Figure 3.4 provides a good visualization of this exercise by showing dispersions for GDP growth forecasts in real time for the period 2003Q1-2011Q2. Given the very similar performances of models, we show results generated when $X_t = (Spread_t, Spread_t^2)$, for $h = 2, 4, 6, 8$. Notice that GDP growth downside risks considerably before and during the recession period, indicating a higher probability of negative output growth. For $h = 3$ the model misses the recession a little by anticipating it, while for $h = 4, 6, 8$ it certainly hits the NBER recession with high accuracy, indicating that our very simple spread-based QR model was able to warn about risks of a recession occurring in 2008/2009 with at least two years in advance.

Another interesting feature observed in Figure 3.4 is that while forecasts of high GDP growth barely modify, the same is not true for lower conditional output growth. This result is basically capturing a conditional asymmetric behavior in the term spread - GDP growth relationship. This asymmetry also reveals a type of heterokedasticity that is not normally identified by standard time series models.

Comparison with professional forecasters

We now report comparisons against the SPF. Results are also shown in Table 3.6 and are stronger than when forecasting GDP growth. Relying on the term spread results in statistically more accurate forecasts than those reported by professional forecasters for $h = 4,5$. Opposite results are observed for $h = 1,2,3$, with the exception of model QR 1 for $h = 3$, when evaluated by LPS.

In order to verify the stability of these results Figure 3.5 reports RQPS and LPS statistics computed recursively for quantile models, the standard probit AR and the SPF. Similarly to what was found for GDP growth, Figure 3.5 reveals that for the period before the last recession QR models were more accurate than professional forecasters. For instance, for $h = 2$ QR models show higher accuracy over the whole 2000's when evaluated using RQPS and since the early 1990's when evaluated using LPS, although statistical significance was not observed. For $h = 3$, however, the higher accuracy of QR models was statistically significant since the early 1990's and the same is observed for $h = 5$.

In line with Rudebusch and Williams (2009), which conduct the same analysis using a simple probit model, these results indicate the high accuracy of spread models when forecasting recessions. As shown, this is also true when forecasting GDP growth and reinforce our findings regarding the continuing predictive power of the yield spread.

3.5 Conclusions

We use quantile regression to re-examine the term spread ability to predict recessions and GDP growth by studying the whole conditional distribution of the latter variable. Our analyses reveal a yet not documented asymmetry between our variables of interest: while the spread is able to predict negative and lower GDP growth, it informs little about high output growth. We also verify changes in the predictive relationship towards longer horizons after the mid-1980's and identify structural breaks happening mostly at intermediate and higher quantiles. This is consistent with a market view that the Fed started to respond tougher and in greater advance to inflationary pressures resulted from excessive growth around 1984, which is in line with a large literature verifying a change in the reaction function of the Fed towards inflation after the Volcker's presidency.

Since we verify the existence of (conditional) asymmetries in the GDP growth - term spread relationship we then use quantile regressions to forecast GDP growth and probabilities of future recessions. Contrasting with most of the works that relies on standard probit regressions (assuming conditional normality of the GDP growth), quantile regressions offer greater flexibility as recessions probabilities can be computed from empirical distributions. In addition, the existence of conditional asymmetries may lead OLS estimates to deliver biased GDP growth point forecasts, suggesting the use of median estimators.

Results reveal that quantile models deliver more accurate forecasts than competitors in both exercises. For the post-1985 data, results indicate higher accuracy in longer horizons, which is consistent with a more forward looking Fed. We also find superiority against professional forecasters for mid and longer horizons, but higher accuracy in shorter horizons is also observed in the period prior the 2008/2009 recession. All together, our findings suggest that the predictive power of the yield curve remains.

Table 3.1: OLS and QR(0.5) regressions results

Notes: Standard errors (in parentheses) for OLS coefficients are computed by the heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator of Andrews and Monahan (1992). Standard errors for QR are computed by paired bootstrap using 2000 replications. $\star$ denotes statistical significance at (***)1%, (**)5% and (*)10%, respectively.

Sample		OLS				QR(0.5)			
		$spread_{t-h}$	$spread^2_{t-h}$	y_{t-h}	AIC	$spread_{t-h}$	$spread^2_{t-h}$	y_{t-h}	AIC
1955 − 2011	$h=1$	0.476** (0.196)		0.335*** (0.075)	1196.9	0.199 (0.156)		0.295*** (0.103)	1183.3
		1.204** (0.333)	−0.311*** (0.104)	0.325*** (0.080)	1191.3	1.758** (0.722)	−0.484** (0.210)	0.231*** (0.089)	1175.6
	$h=2$	0.719*** (0.217)		0.194** (0.083)	1202.7	0.537** (0.212)		0.178** (0.071)	1183.5
		1.825*** (0.279)	−0.476*** (0.092)	0.181** (0.076)	1187.6	1.984*** (0.415)	−0.528*** (0.117)	0.161*** (0.055)	1162.1
	$h=4$	0.688*** (0.255)		−0.007 (0.079)	1205.8	0.382** (0.198)		−0.010 (0.071)	1186.0
		1.350** (0.621)	−0.286 (0.201)	−0.015 (0.078)	1202.2	2.101*** (0.601)	−0.514*** (0.166)	−0.020 (0.174)	1171.8
	$h=6$	0.419** (0.199)		−0.034 (0.074)	1203.0	0.255 (0.161)		−0.019 (0.063)	1179.5
		0.646** (0.304)	−0.100 (0.106)	−0.037 (0.075)	1204.4	0.832* (0.440)	−0.223 (0.149)	−0.041 (0.060)	1179.6
	$h=8$	0.171 (0.276)		−0.076 (0.081)	1194.9	0.060 (0.209)		−0.028 (0.079)	1169.7
		0.446 (0.651)	−0.124 (0.202)	−0.080 (0.080)	1195.9	0.730 (0.664)	−0.236 (0.196)	−0.072 (0.073)	1169.1
1985 − 2011	$h=1$	0.061 (0.135)		0.482*** (0.152)	479.8	0.020 (0.177)		0.307*** (0.116)	469.6
		0.366 (0.580)	−0.091 (0.161)	0.484*** (0.142)	481.6	−0.414 (0.946)	0.128 (0.251)	0.303** (0.120)	471.4
	$h=2$	0.173 (0.158)		0.417*** (0.097)	486.9	0.056 (0.206)		0.300** (0.116)	469.7
		0.481 (0.623)	−0.092 (0.184)	0.420*** (0.092)	488.7	1.011 (0.819)	−0.255 (0.213)	0.300*** (0.106)	470.3
	$h=4$	0.441* (0.262)		0.146 (0.114)	500.6	0.293 (0.234)		0.038 (0.090)	480.0
		1.807*** (0.629)	−0.409** (0.191)	0.157 (0.113)	498.3	1.570** (0.076)	−0.398** (0.199)	0.057 (0.104)	474.3
	$h=6$	0.713** (0.354)		0.051 (0.067)	495.6	0.453** (0.210)		0.023 (0.066)	476.6
		2.285** (1.077)	−0.479* (0.269)	0.055 (0.064)	491.7	2.180* (1.139)	−0.499* (0.298)	−0.039 (0.059)	473.9
	$h=8$	0.725* (0.370)		0.043 (0.077)	495.7	0.423* (0.220)		0.071 (0.075)	477.2
		3.535** (1.328)	−0.867** (0.347)	0.042 (0.068)	477.4	1.999*** (0.738)	−0.511** (0.198)	0.023 (0.062)	460.2

Table 3.2: Breaks in regression quantiles - SQ test

Notes: This table shows estimated statistics for the break test of Qu (2008). SQ(1) refers to the test statistic for the presence of 1 break in the specified quantile regression. SQ(2|1) tests for the existence of a second break given the first one was found. ⋆⋆ denotes significance at the 5% level. The sample period is 1955Q1-2011Q2.

		Quantiles	0.1	0.2	0.35	0.5	0.65	0.8	0.9
Linear	$h=4$	SQ(1)	1.112	1.408	1.771**	1.610**	1.962**	1.995**	2.083**
		SQ(2|1)	–	–	1.788**	1.653**	1.793**	1.058	0.855
		Break Date			62:4,08:2	84:2,08.3	84:2,08.3	84:2	84:2
	$h=8$	SQ(1)	1.077	1.157	1.345	1.463	1.816**	1.688**	2.015**
		SQ(2|1)	–	–	–	–	1.520**	1.479	0.986
		Break Date	–	–	–	–	66:1,00:2	84:2	84:2
Quadratic	$h=4$	SQ(1)	1.079	1.160	1.173	1.262	1.821**	1.995**	2.197**
		SQ(2|1)	–	–	–	–	1.143	1.096	1.060
		Break Date	–	–	–	–	84:2	84:2	84:1
	$h=8$	SQ(1)	1.158	1.349	1.376	1.706**	1.723**	1.542	1.878**
		SQ(2|1)	–	–	–	1.200	1.994**	–	1.035
		Break Date	–	–	–	07:2	66:1,00:2	–	84:2

Table 3.3: Breaks in regression quantiles - DQ test

Notes: This table shows estimated statistics for the break test of Oka and Qu (2011), allowing for a maximum of two breaks. DQ(1) refers to the test statistic of 1 break in multiple regression quantiles. DQ(2|1) tests for the existence of a second break given the first was detected. ⋆⋆ denotes significance at the 5% level. The sample period is 1955Q1-2011Q2.

Specification	Linear Model		Quadratic Model	
	$h=4$	$h=8$	$h=4$	$h=8$
DQ(1)	1.035**	0.918**	0.952**	0.920
DQ(2|1)	0.940	0.905	0.914	–
Break Date	84:2	84:2	84:2	–

Table 3.4: GDP growth forecasting results

Notes: This table shows out-of-sample GDP growth forecasting results in the form of RMSFE ratios relative to an AR model. OLS (QR) 1, OLS (QR) 2, OLS (QR) 3 and OLS (QR) 4 denote specifications (i) $X_t = Spread_t$, (ii) $X_t = (Spread_t, Spread_t^2)$, (iii) $X_t = (Spread_t, y_t)$ and (iv) $X_t = (Spread_t, Spread_t^2, y_t)$, respectively. Blue/underscored values indicate outperformance of QR models against OLS with same specification. Large numbers indicate the most accurate forecasts model. $\star$ denotes significance [(***)1%, (**)5%, (*) 10%] of QR against OLS. $\circ$ denotes significance [($^{\circ\circ\circ}$)1%, ($^{\circ\circ}$)5%, ($^{\circ}$)10%] against AR. $\dagger$ denotes significance [(†††)1%, (††)5%, (†)10%] of quadratic specification against linear specification. $\bullet$ denotes significance [($^{\bullet\bullet\bullet}$)1%, ($^{\bullet\bullet}$)5%, ($^{\bullet}$)10%] against SPF.

		OLS 1	OLS 2	OLS 3	OLS 4	QR 1	QR 2	QR 3	QR 4	SPFMEAN	SPFMED
$h=1$	REALTIME	1.469	1.316†††	1.208	1.092†††	1.358*	1.281*††	1.134**	1.128	0.654	0.653
	FINREV	1.663	1.500†††	1.281	1.156†††	1.466***	1.408***†	1.251	1.243	–	–
$h=2$	REALTIME	1.414	1.204†††	1.311	1.127††	1.222***	1.150*†	1.137***	1.076	0.812	0.825
	FINREV	1.606	1.367†††	1.445	1.231††	1.357***	1.266***†	1.250***	1.133***†	–	–
$h=3$	REALTIME	1.219	1.090††	1.232	1.105††	1.079***	1.025**	1.125**	1.088	0.886	0.931
	FINREV	1.287	1.150††	1.301	1.168††	1.118***	1.037***	1.160***	1.132	–	–
$h=4$	REALTIME	1.085	1.001†	1.117	1.030†	0.986**	0.932**†	1.017**	0.955$^{\star\star\star\dagger\circ}$	0.955	0.971
	FINREV	1.158	1.073†	1.207	1.121†	1.016***	0.935***†	1.125**	1.030**	–	–
$h=5$	REALTIME	0.985	0.950	1.033	0.988	0.930$^{\star\star\bullet\bullet}$	0.921$^{\star\circ\bullet}$	0.993	0.976	0.997	0.991
	FINREV	1.022	0.978	1.103	1.050†	0.944***	0.907$^{\star\star\star\dagger\circ}$	1.022**	0.976***†	–	–
$h=6$	REALTIME	0.949	0.944	0.959	0.953°	0.962	0.977	0.972	0.974	–	–
	FINREV	0.971	0.967	0.992	0.986	0.966	0.966	0.981	0.973	–	–
$h=8$	REALTIME	0.990	1.011	0.995	1.007	0.996	1.028	1.028	1.043	–	–
	FINREV	0.977	0.984	1.003	1.004	0.969	0.950*	1.022	1.013†	–	–

Table 3.5: GDP growth forecasting results - post-1985 data

		OLS 1	OLS 2	OLS 3	OLS 4	QR 1	QR 2	QR 3	QR 4	SPFMEAN	SPFMED
$h=1$	REALTIME	1.180	1.182	1.021	1.014	1.149	1.154	1.030	1.051	0.612	0.607
	FINREV	1.328	1.363	1.037	1.056	1.347	1.363	1.160	1.191	–	–
$h=2$	REALTIME	1.139	1.156	1.052	1.055	1.085	1.103	1.026	1.039	0.794	0.799
	FINREV	1.252	1.296	1.064	1.085	1.247	1.282	1.040	1.075	–	–
$h=3$	REALTIME	1.068	1.088	1.073	1.089	1.027	1.056	1.057	1.071	0.865	0.915
	FINREV	1.015	1.026	1.101	1.103	1.028	1.008	1.082	1.001	–	–
$h=4$	REALTIME	0.996	1.010	1.033	1.028	0.932	0.974	1.016	1.011	0.899	0.902
	FINREV	1.020	1.004	1.029	1.008	1.015	0.989	1.020	0.988	–	–
$h=5$	REALTIME	1.000	1.146	0.997	1.210	0.969	1.164	0.977	1.324	0.976	0.983
	FINREV	0.963	1.033	0.999	1.108	0.953	1.090	0.977	1.117	–	–
$h=6$	REALTIME	0.963	1.070	1.013	1.103	0.928	1.097	0.968	1.106	–	–
	FINREV	0.955	1.027	1.036	1.042	0.950	1.084	1.012	1.114	–	–
$h=8$	REALTIME	0.913	$0.817^{\dagger\dagger\circ}$	0.964	$0.891^{\dagger\circ}$	0.893°	$0.774^{\star\star\dagger\dagger\circ\circ}$	0.960	$0.895^{\dagger\dagger\dagger\circ}$	–	–
	FINREV	0.921	$0.799^{\dagger\circ}$	0.983	$0.869^{\dagger}$	0.951	0.818	1.012	$0.879^{\dagger\dagger}$	–	–

Notes: As in Table 3.4.

Table 3.6: Recessions forecasting results

Notes: This table shows out-of-sample recessions forecasting results in the form of RQPS/LPS ratios relative to a simple probit model. PRB (QR) 1, PRB (QR) 2, PRB (QR) 3 and PRB (QR) 4 denote specifications (i) $X_t = Spread_t$, (ii) $X_t = (Spread_t, Spread_t^2)$, (iii) $X_t = (Spread_t, y_t)$ and (iv) $X_t = (Spread_t, Spread_t^2, y_t)$, respectively. Blue/underscored values indicate outperformance of QR models against Probit with same specification. Large numbers indicate the most accurate forecast model. $\star$ denotes significance [$(^{\star\star\star})1\%$, $(^{\star\star})5\%$, $(^{\star})10\%$] of QR against Probit with same specification. $\circ$ denotes significance [$(^{\circ\circ\circ})1\%$, $(^{\circ\circ})5\%$, $(^{\circ})10\%$] against simple probit. † denotes significance [$(^{\dagger\dagger\dagger})1\%$, $(^{\dagger\dagger})5\%$, $(^{\dagger})10\%$] of quadratic specification against linear specification. $\bullet$ denotes significance[$(^{\bullet\bullet\bullet})1\%$, $(^{\bullet\bullet})5\%$, $(^{\bullet})10\%$] against SPF.

			PRB 1	PRB 2	PRB 3	PRB 4	QR 1	QR 2	QR 3	QR 4	SPF
$h=1$	RealTime	RQPS	1.421	1.458	1.344	1.382	1.442	1.439	1.314	1.345	0.758
		LPS	1.543	1.589	1.406	1.446	1.523	1.540	1.353	1.390	0.863
	FinRev	RQPS	1.154	1.196	1.055	1.102	1.165	1.172	1.037	1.029	–
		LPS	1.119	1.157	0.966	0.995	1.105	1.121	0.962	0.956	–
$h=2$	RealTime	RQPS	0.886	0.887	0.875	0.875	0.894	0.882	0.886	0.873	0.668
		LPS	0.847	0.839	0.799	0.791	0.804	0.795	0.783	0.777	0.644
	FinRev	RQPS	1.068	1.070	1.013	1.011	1.060	1.054	0.962	0.982	–
		LPS	1.071	1.069	0.945	0.943	1.008	0.998	0.890	0.906	–
$h=3$	RealTime	RQPS	0.775°	0.775°	0.774°	0.774°	0.776°	0.769°	0.783	0.765°	0.732
		LPS	0.746	0.748	0.742	0.744	0.713°	0.724°	0.716°	0.716°	0.715
	FinRev	RQPS	0.882°	0.889°	0.886$^{\circ\circ}$	0.893°	0.884°	0.879°	0.886$^{\circ\circ}$	0.878$^{\circ\circ}$	–
		LPS	0.860	0.867	0.857	0.863	0.836$^{\circ\circ}$	0.834$^{\circ\circ}$	0.835$^{\circ\circ}$	0.832$^{\circ\circ}$	–
$h=4$	RealTime	RQPS	0.678$^{\circ\circ\bullet\bullet}$	0.681$^{\circ\circ\bullet\bullet}$	0.684$^{\circ\circ}$	0.686$^{\circ\circ}$	0.701°	0.685$^{\circ\circ\bullet}$	0.698°	0.684$^{\circ\circ\bullet}$	0.771
		LPS	0.596$^{\circ\circ\bullet}$	0.600$^{\circ\circ}$	0.603$^{\circ\circ}$	0.606$^{\circ\circ}$	0.596$^{\circ\circ\bullet\bullet}$	0.593$^{\circ\circ\bullet}$	0.593$^{\circ\circ\bullet\bullet}$	0.591$^{\circ\circ\bullet}$	0.692
	FinRev	RQPS	0.832$^{\circ\circ\circ}$	0.836$^{\circ\circ\circ}$	0.855$^{\circ\circ\circ}$	0.859$^{\circ\circ\circ}$	0.831$^{\circ\circ\circ}$	0.835$^{\circ\circ\circ}$	0.848$^{\circ\circ\circ}$	0.845$^{\circ\circ\circ}$	–
		LPS	0.787$^{\circ\circ\circ}$	0.789$^{\circ\circ\circ}$	0.813$^{\circ\circ}$	0.815$^{\circ\circ}$	0.770$^{\circ\circ\circ}$	0.766$^{\circ\circ\circ}$	0.781$^{\circ\circ\circ}$	0.783$^{\circ\circ\circ}$	–
$h=5$	RealTime	RQPS	0.670$^{\circ\circ\bullet\bullet\bullet}$	0.679$^{\circ\circ\bullet\bullet\bullet}$	0.671$^{\circ\circ\bullet\bullet\bullet}$	0.681$^{\circ\circ\bullet\bullet\bullet}$	0.690$^{\circ\circ\bullet\bullet\bullet}$	0.679$^{\circ\circ\bullet\bullet\bullet}$	0.689$^{\circ\circ\bullet\bullet\bullet}$	0.684$^{\circ\circ\bullet\bullet\bullet}$	0.801
		LPS	0.533$^{\circ\circ\bullet\bullet\bullet}$	0.548$^{\circ\circ\bullet\bullet\bullet}$	0.535$^{\circ\circ\bullet\bullet\bullet}$	0.551$^{\circ\circ\bullet\bullet\bullet}$	0.545$^{\circ\circ\bullet\bullet\bullet}$	0.542$^{\circ\circ\bullet\bullet\bullet}$	0.546$^{\circ\circ\bullet\bullet\bullet}$	0.552$^{\circ\circ\bullet\bullet\bullet}$	0.664
	FinRev	RQPS	0.811$^{\circ\circ\circ}$	0.822$^{\circ\circ\circ}$	0.818$^{\circ\circ\circ}$	0.828$^{\circ\circ\circ}$	0.820$^{\circ\circ\circ}$	0.834$^{\circ\circ\circ}$	0.815$^{\circ\circ\circ}$	0.827$^{\circ\circ\circ}$	–
		LPS	0.749$^{\circ\circ\circ}$	0.764$^{\circ\circ\circ}$	0.757$^{\circ\circ\circ}$	0.774$^{\circ\circ\circ}$	0.755$^{\circ\circ\circ}$	0.763$^{\circ\circ\circ}$	0.752$^{\circ\circ\circ}$	0.760$^{\circ\circ\circ}$	–
$h=6$	RealTime	RQPS	0.694$^{\circ\circ}$	0.688$^{\circ\circ}$	0.698$^{\circ\circ}$	0.692$^{\circ\circ}$	0.695$^{\circ\circ}$	0.698$^{\circ\circ}$	0.693$^{\circ\circ}$	0.696$^{\circ\circ}$	–
		LPS	0.621$^{\circ\circ}$	0.608$^{\circ\circ}$	0.623$^{\circ\circ}$	0.611$^{\circ\circ}$	0.610$^{\circ\circ}$	0.610$^{\circ\circ}$	0.611$^{\circ\circ}$	0.609$^{\circ\circ}$	–
	FinRev	RQPS	0.832$^{\circ\circ\circ}$	0.828$^{\circ\circ\circ}$	0.834$^{\circ\circ\circ}$	0.830$^{\circ\circ\circ}$	0.823$^{\circ\circ\circ}$	0.830$^{\circ\circ\circ}$	0.835$^{\circ\circ\circ}$	0.838$^{\circ\circ\circ}$	–
		LPS	0.788$^{\circ\circ\circ}$	0.774$^{\circ\circ\circ}$	0.789$^{\circ\circ\circ}$	0.776$^{\circ\circ\circ}$	0.780$^{\circ\circ\circ}$	0.773$^{\circ\circ\circ}$	0.793$^{\circ\circ\circ}$	0.782$^{\circ\circ\circ}$	–
$h=8$	RealTime	RQPS	0.762$^{\circ\circ}$	0.759$^{\circ\circ}$	0.740$^{\circ\circ}$	0.737$^{\circ\circ}$	0.761$^{\circ\circ}$	0.784$^{\circ\circ}$	0.733$^{\circ\circ}$	0.744$^{\circ\circ}$	–
		LPS	0.695$^{\circ\circ}$	0.686$^{\circ\circ}$	0.668$^{\circ\circ}$	0.659$^{\circ\circ}$	0.686$^{\circ\circ}$	0.706°	0.655$^{\circ\circ}$	0.662$^{\circ\circ}$	–
	FinRev	RQPS	0.973	0.965††	0.902$^{\circ\circ}$	0.893$^{\circ\circ}$	0.905$^{\circ\circ}$	0.913$^{\circ\circ}$	0.904$^{\circ\circ}$	0.900$^{\circ\circ}$	–
		LPS	0.960	0.932†††	0.870$^{\circ\circ\circ}$	0.851$^{\dagger\circ\circ\circ}$	0.872$^{\circ\circ\circ}$	0.871$^{\circ\circ\circ}$	0.864$^{\circ\circ}$	0.857$^{\dagger\circ\circ\circ}$	–

Figure 3.1: Quantile regressions estimates

Notes: Row 1 shows estimated coefficients for *spread* when $X_{t-h} = (Spread_{t-h}, y_{t-h})$. Rows 2 and 3 show estimated coefficients for *spread* and $spread^2$ when $X_{t-h} = (Spread_{t-h}, Spread^2_{t-h}, y_{t-h})$. Standard errors were obtained by paired bootstrap with 2000 replications. Asterisks indicate statistical significance at 10%.

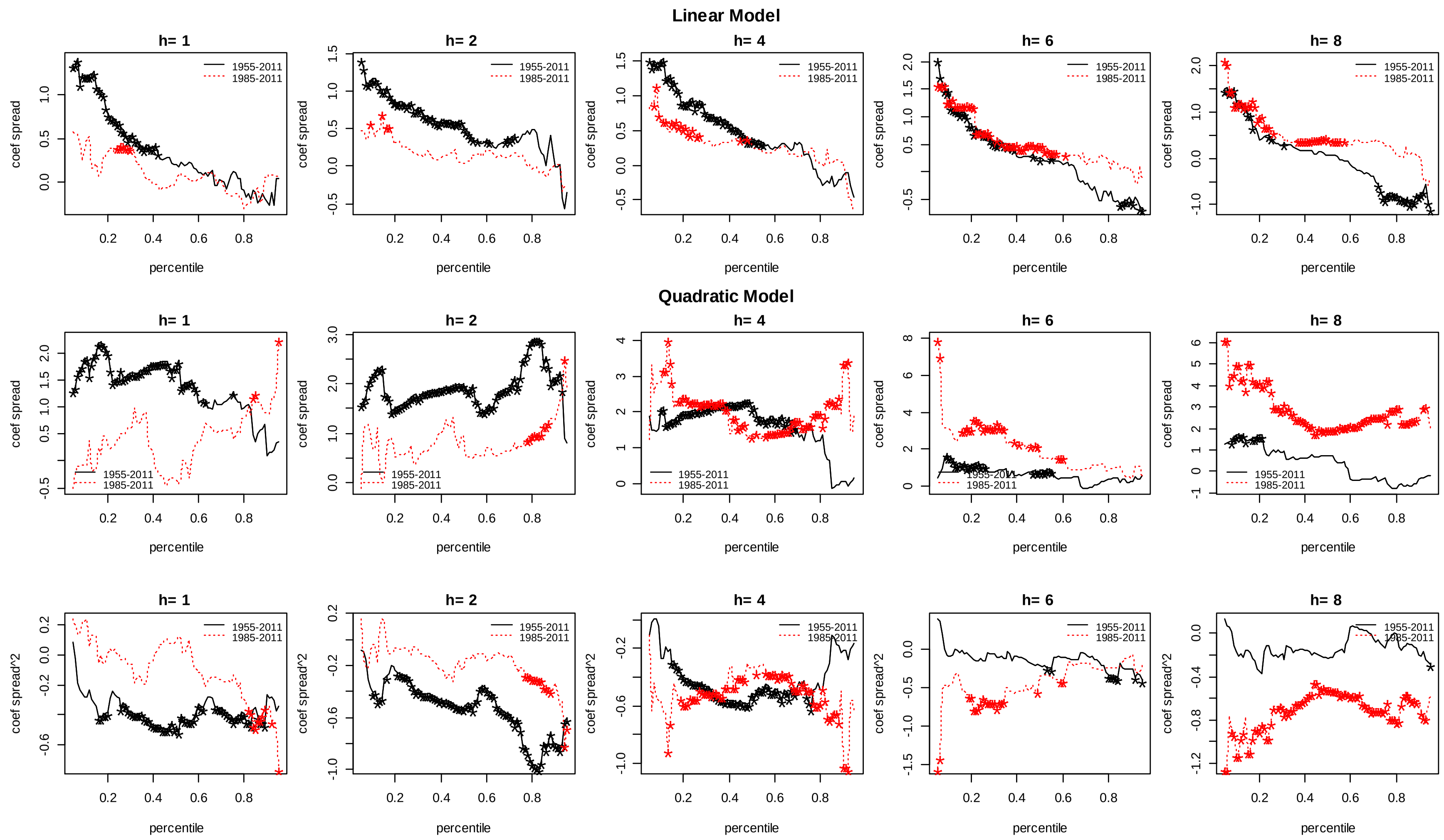


Figure 3.2: Adjusted $R1\left(\tau\right)$ statistics

Notes: This figure shows the $\overline{R1}\left(\tau\right)$ statistics of Koenker and Machado (1999) computed for the following specifications: (i) $X_{t-h}=Spread_{t-h}$, (ii) $X_{t-h}=(Spread_{t-h},Spread_{t-h}^{2})$, (iii) $X_{t-h}=(Spread_{t-h},y_{t-h})$ and (iv) $X_{t-h}=(Spread_{t-h},Spread_{t-h}^{2},y_{t-h})$.

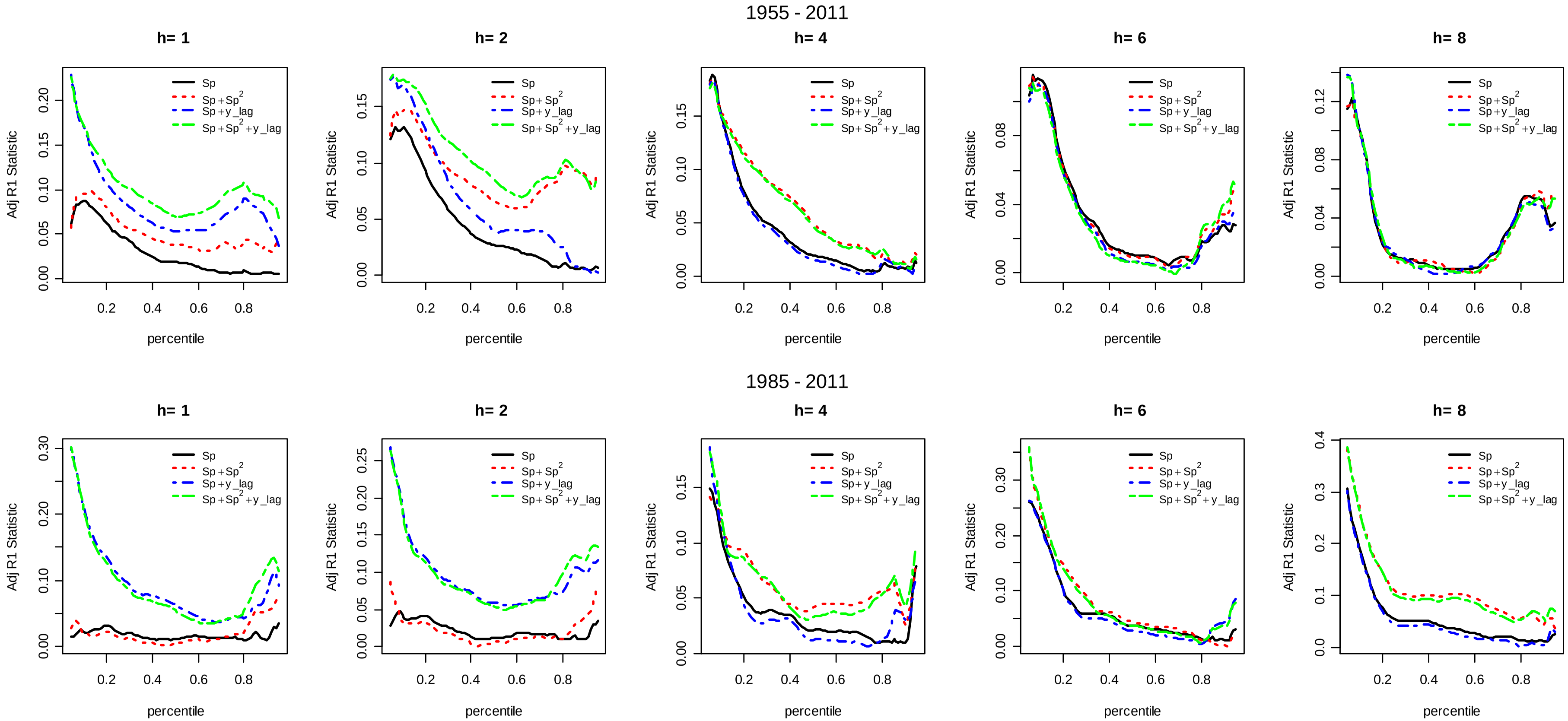


Figure 3.3: Recursive GDP growth forecasting in real-time

Notes: Panels (a), (b), (c) and (d) show recursive GDP growth forecast errors for 1, 2, 3 and 5 quarters ahead, respectively. Mean (Median) indicates SPF mean (median) and OLS (QR(0.5)) recursive forecast errors. The sample period is 1955Q1-2011Q2. Shaded areas show NBER recessions and asterisks denote statistical significance against SPF at the 10% level.

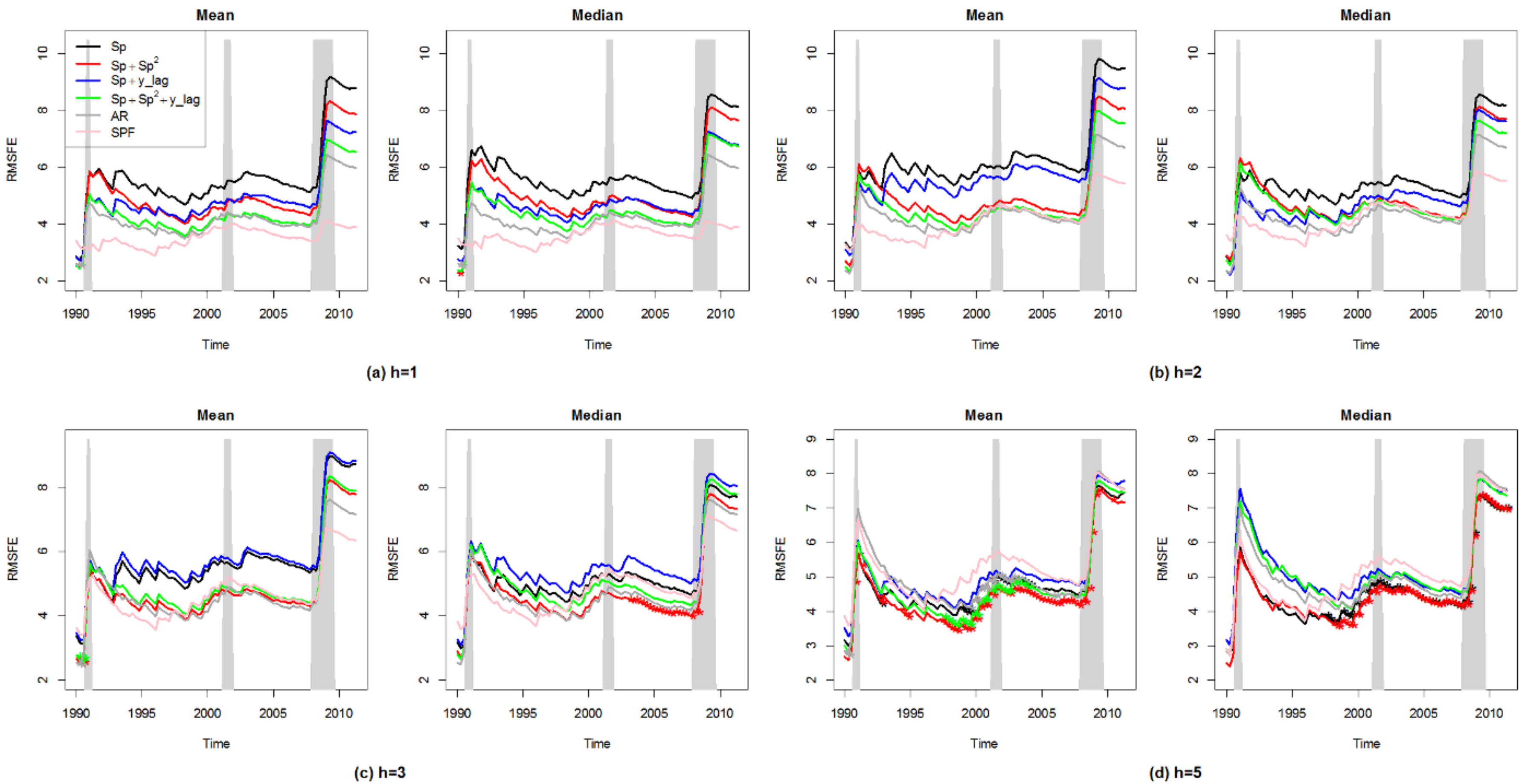


Figure 3.4: Forecasting the great recession in real-time

Notes: This figure shows forecasted GDP growth dispersions in real-time. We show results for $\tau = 0.02, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.98$, $h = 2, 4, 6, 8$ and specification $X_t = (Spread_t, Spread_t^2)$. The red line refers to median forecasts. The sample period is 1955Q1-2011Q2.

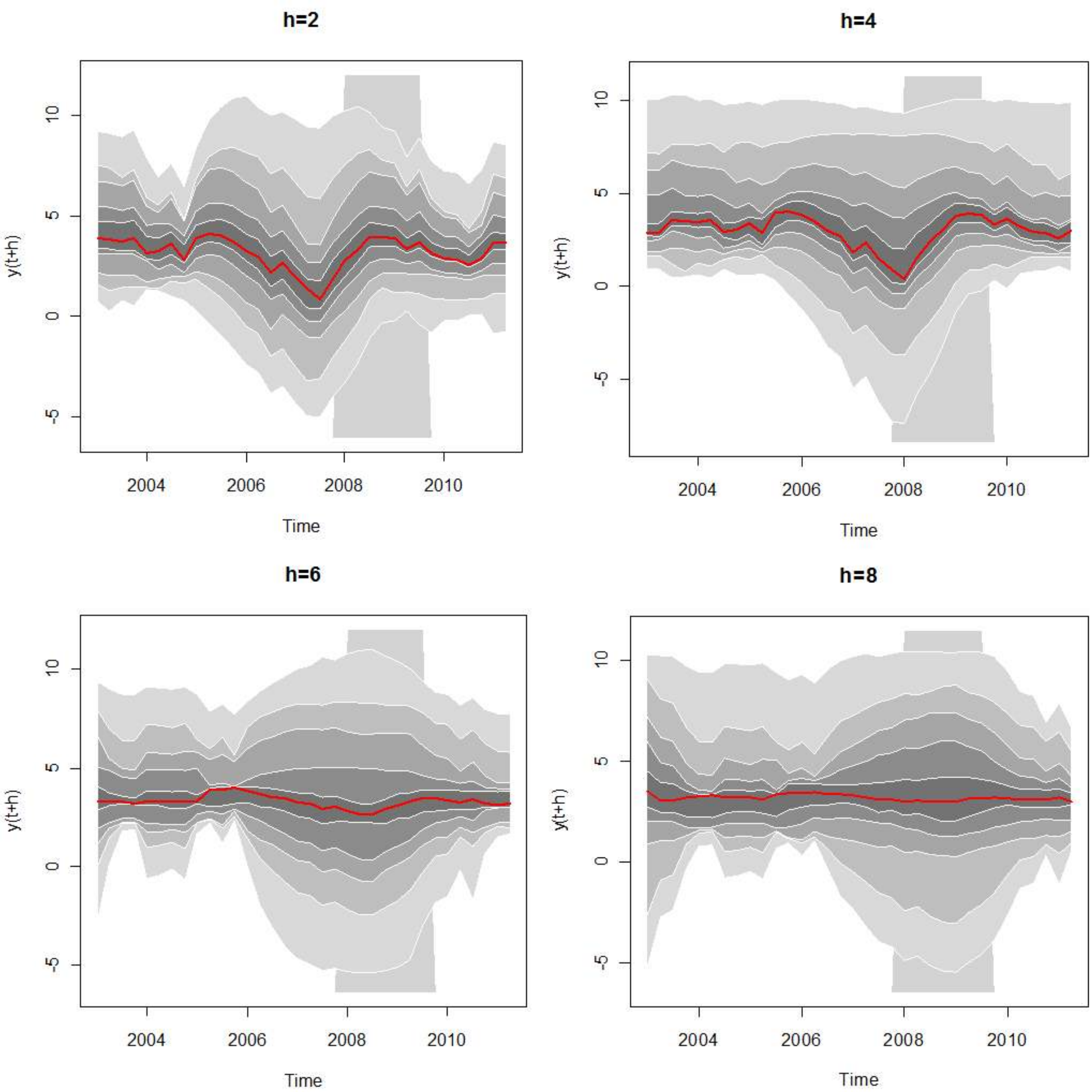
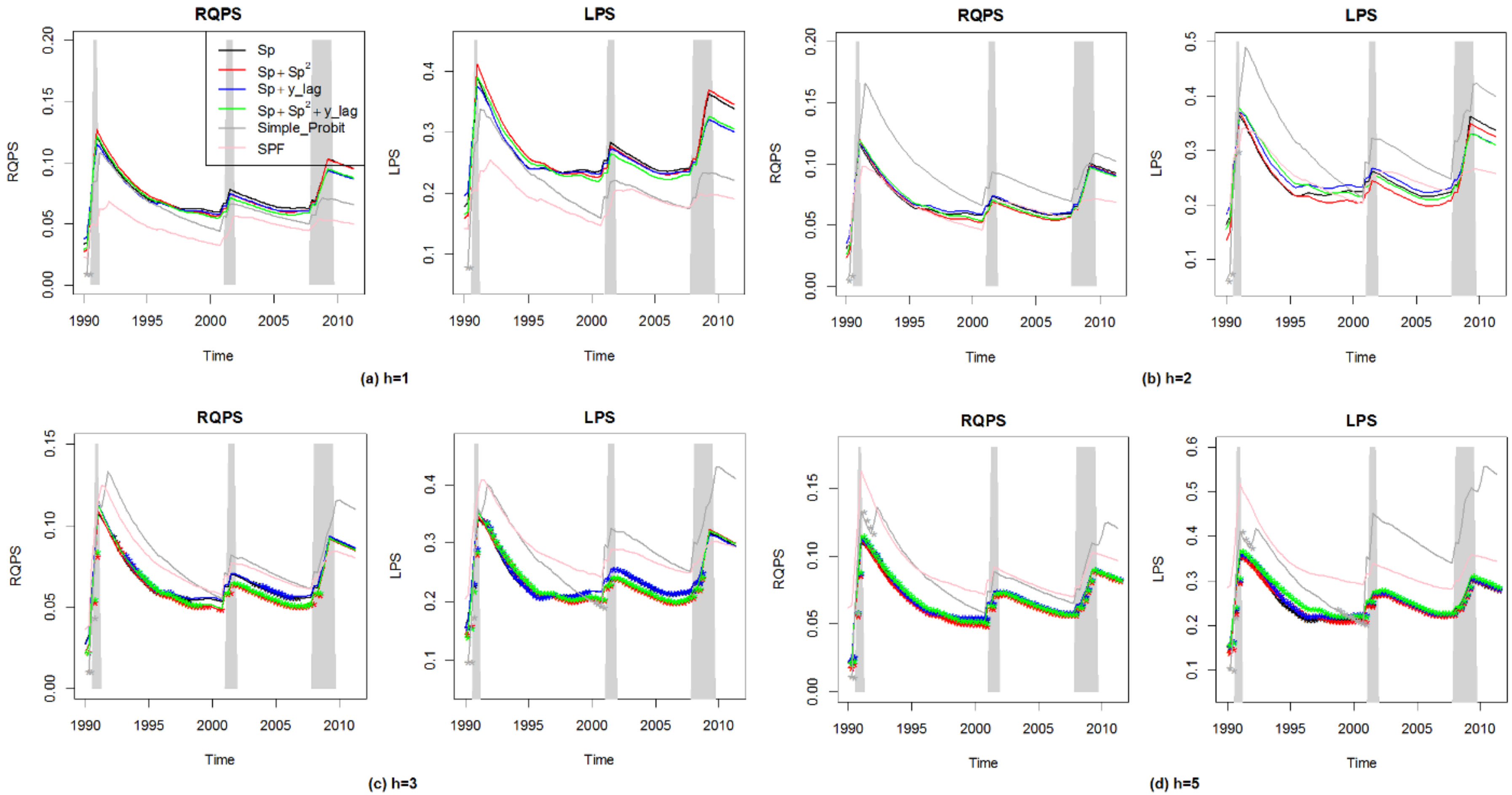


Figure 3.5: Recursive recessions forecasting in real-time

Notes: Panels (a), (b), (c) and (d) show recursive forecast errors of QR recessions probabilities for 1, 2, 3 and 5 quarters ahead, respectively. The sample period is 1955Q1-2011Q2. Shaded areas show NBER recessions ans asterisks denote statistical significance against SPF at the 10% level.



Chapter 4

Modeling and forecasting the yield curve by an extended Nelson-Siegel class of models: a quantile autoregression approach[1]

Rafael B. De Rezende Mauro S. Ferreira

ABSTRACT. This paper compares the in sample fitting and the out of sample forecasting performances of four different Nelson-Siegel type of models: Nelson-Siegel, Bliss, Svensson, and a five factor model we propose in order to enhance fitting flexibility. The introduction of the fifth factor resulted in superior adjustment to the data. For the forecasting exercise the paper contrasts the performances of the term structure models in association to the following econometric methods: quantile autoregression evaluated at the median, VAR, AR, and a random walk. The quantile procedure delivered the best results for longer forecast horizons.

Keywords: yield curve, in-sample fitting, out-of-sample forecasts, Nelson-Siegel, quantile autoregression

JEL Classifications: C53; E43; E47

---

[1]This is an early version of the paper "Modeling and Forecasting the Yield Curve by an Extended Nelson-Siegel Class of Models: A Quantile Autoregression Approach", Journal of Forecasting, v.32, n.2, pp.111-123, 2013. We would like to thank seminar participants at the Universidade Federal de Minas Gerais, Forecasting in Rio, Latin American Meeting of the Econometric Society 2008, Brazilian Meeting of Econometrics 2008 and the Brazilian Meeting of Finance 2008 for comments and suggestions.

4.1 Introduction

The yield curve plays a central role in macroeconomics and finance. For macroeconomists, it carries information on a variety of variables, such as expected inflation and future GDP (Estrella and Hardouvelis, 1991 and Estrella and Mishkin, 1996, 1998). In finance, the yield curve allows marking to market, pricing derivatives, hedging, among other uses. Thus, it is not surprising the substantial research effort in estimating and forecasting the yield curve.

The three most popular estimation approaches are: the affine equilibrium models (Vasicek, 1977, Cox et al., 1985, Duffie and Kan, 1996), no-arbitrage models (Hull and White, 1990, Heath et al., 1992), and statistical and parametric models (McCulloch, 1971, 1975; Fisher et al., 1995; Vasicek and Fong, 1982; Nelson and Siegel, 1987; Svensson, 1994; Bliss, 1997). The first two have not been very successful in terms of forecasting. According to Duffee (2002), equilibrium models only pay attention to instantaneous short rates, resulting in poor yield curve predictions. The same is true for arbitrage-free models, as they are specialized in fitting yield curves at a particular point in time, leaving aside their dynamics.

The Nelson-Siegel class of parametric models, on the other hand, has delivered good estimation and forecasting properties, which has increased its popularity among users. The Bank for International Settlements (BIS, 2005) reports that nine of the thirteen main central banks of the world rely on the Nelson and Siegel (1987) model (NS, hereafter) and/or on the Svensson (1994) model (SV, hereafter) for yield curve estimation, while Gürkaynak et al. (2007) suggest the use of the SV by the Federal Reserve Board. For forecasting purposes, Diebold and Li (2006) show that a dynamic version of the NS model predicts the US yield curve more accurately than other competing models, especially at longer horizons. Möench (2008) partially confirms these results, while Bolder (2006) shows the superiority of the NS when compared to affine equilibrium and other parametric models.

Motivated by the practical use of the Nelson-Siegel models, this paper compares the in-sample fitting and the out-of-sample predictive power of some of these models: NS, Bliss (1987) (BL, henceforth), SV and a five-factor model (FF, hereafter) we propose. The decision to include a fifth factor as an extension to the SV model is an attempt to enhance its flexibility in order to improve in-sample fitting and its predictive power.

The paper also innovates when compares the forecasting performance of a first order quantile autoregressive model - QAR, estimated at the median, to other standard procedures: AR(1), VAR(1), and a random walk (RW) applied directly on the term structure yields. The use of a robust estimation method, such as QAR, reduces estimation bias provoked by outliers which normally affects out of sample forecasting performance. This robustness seems particularly attractive when dealing with financial variables that normally suffer from wide oscillations, even more in emerging markets.

We conduct the analysis for the Brazilian zero-coupon data, which characteristics are described in the next section. Although most of the term structure literature focuses on data for the richest countries, the international investors' increasing interest in the high yields of emerging economies, which has been accentuated due to their recent conquest of macroeconomic and institutional stability, naturally requires a better comprehension of such financial markets. This paper takes a step in this direction when analyzing the Brazilian case, one of the most popular destinies for international investors seeking high returns in the emerging world.

In the third section we present three standard term structure models before introducing the FF model. The fourth section describes the in sample and the out of sample estimation procedures while the fifth section discusses the results. The sixth and last section brings our concluding remarks.

We anticipate the main results of our study: (i) the FF shows superior in-sample fitting; (ii) for the out of sample exercise, both the SV and BL are more accurate; (iii) among the econometric forecasting procedures, QAR is superior to AR and to VAR; (iv) the random walk beats QAR for 1 day ahead prediction, but (v) QAR shows superior performance for 1 and 3 months horizons.

4.2 The data

We use daily data on the implicit yield curves extracted from swap operations between the Brazilian Interbank Deposit rate and the Predetermined Interest, which is a fixed coupon rate. The Interbank Deposit rate is a weighted average of daily rates on the interbank lending operations. This swap, referenced by ID-PRE, is probably the most important and liquid instrument of the Brazilian fixed-income market. The contract has the same characteristics of a risk-free zero-coupon since BM&FBOVESPA provides full assurance.

Our sample ranges from March 16 of 2000 to October 15 of 2007, constituting 1883 business days. Because very long maturity contracts have only recently become more liquid, we decided to work with the following 15 vertices: 1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 18, 24, 36, 48 and 60 months (in running days). Figure 4.1 provides a three-dimensional picture of the data. Table 4.1 reports descriptive statistics for the yields and for the empirical factors: level, slope and curvature. Data reveals the average yield curve is positively sloped and volatility increases with maturities. For all maturities, yields are highly persistent over time. They also seem to depart from a normal distribution, as suggested by the positive skewness, either high or low excess kurtosis, and also by the low p-values obtained from a Jarque-Bera test of normality. This last pattern is also observed for all empirical factors. The level, however, presents higher persistency and volatility than the slope and the curvature.

In order to familiarize the readers with the data, we contrast, in Figure 4.2, the time series of the empirical factors (left scale) with the Brazilian risk premium (right scale), measured in basis point by the Emerging Markets Bond Index for Brazil (EMBI Brazil), which is computed by the JP Morgan. It is straightforward to visualize the close comovements between these series, reflecting how the Brazilian interest rate policy was strongly affected by the country risk perception. This pattern, observed in several emerging markets, has been studied by Uribe and Yue (2006). Following an adverse shock (normally a contagion, a default or even an increase in the political uncertainty), investors take their money out of the country, provoking a huge currency depreciation that passes through to the local inflation rates. Central banks following an inflation targeting regime, which is the Brazilian situation, react by increasing the policy interest rate, which ends up oscillating according to the country risk perception.

This is the explanation for why the empirical factors and the EMBI Brazil varied so abruptly from the beginning of 2001 until the middle of 2003. During this period Brazil suffered contagion from the Argentine's sovereign default in December of 2001 and from the increase in the international risk aversion following the 9/11 attack. The country also had to deal with a local political uncertainty during the presidential election in October 2002, since the leading candidate, Lula, and his Labor Party had announced before they would deviate from orthodox economic policies. The analysis of Uribe and Yue (2006) also explains the coincident decline in the risk premium and the empirical factors. As Lula positively surprised by honoring debt contracts, increasing budget surplus and maintaining the inflation targeting framework, the risk premium declined, attracting foreign investors back to Brazil and appreciating the local currency, which helped lowering the inflation. As a result, the level of term structure of interest rates declined with the perspective that the Central Bank of Brazil would reduce further the interest rate.

4.3 Nelson-Siegel class models

Basic definitions

The term structure of interest rates can be described in terms of the spot (or zero-coupon) rate, the discount rate and the forward rate. A forward rate $f(\overline{\tau}, \tau^*)$ is the interest rate of a forward contract on an investment that will be initiated $\overline{\tau}$ periods in the future and will mature τ^* periods beyond the start date of the contract. The instantaneous forward rate $f(\overline{\tau})$ is defined as $lim_{\tau^* \to 0} f(\overline{\tau}, \tau^*) = f(\overline{\tau})$, from which one obtains the forward curve, $f(\tau)$. The spot rate $y(\overline{\tau})$, implicit in a zero-coupon bond with maturity τ, is defined as $y(\overline{\tau}) = \frac{1}{\tau} \int_0^{\overline{\tau}} f(x).dx$, from which one gets the spot curve (or zero-coupon curve), $y(\tau)$. The discount curve, formed by the present value of a zero-coupon bond paying the nominal value of \$1.00 after τ periods, is obtained as the following: $d(\tau) = e^{-y(\tau)\tau}$. It is then straightforward to move from one curve to another using the relations $d(\tau) = exp\left[-\int_0^{\tau} f(x).dx\right]$ and $f(\tau) = -\frac{d'(\tau)}{d(\tau)}$.

Three factor Nelson-Siegel (1987) model

Nelson and Siegel (1987) fitted a zero-coupon curve at a particular point in time using the following model:

$$y(\tau) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right) + \varepsilon^{\tau} \tag{4.1}$$

Later, Diebold and Li (2006) introduced dynamics to this model by regressing, period by period, yields $y(\tau)$ on the exponential components of (4.1), which resulted in time series for its coefficients β_1, β_2, and β_3. Following Litterman and Scheinkman (1991), Diebold and Li (2006) interpreted them as latent factors of level (β_{1t}), slope (β_{2t}) and curvature (β_{3t}), while the exponential terms (inside parentheses of equation (4.1)) were interpreted as factor loadings whose shapes depend on the decaying parameter (λ_t). For matters of simplicity, Diebold and Li (2006) fixed λ_t at $\lambda = 0.0609$ for every t, which allowed (4.1) to be estimated by OLS. In Figure 4.3(a) we show the shapes these exponential components can assume.

Despite its importance, the Nelson-Siegel model has gone through several modifications to enhance flexibility in order to capture a wider variety of curve shapes. These improvements were mainly done by incorporating additional factors and decaying parameters.

Bliss (1997) three factor model

Bliss (1997) tried to improve the term structure fitting by incorporating two different decaying parameters: $\lambda_1$and λ_2. The spot curve at each time t is then given by

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau} - e^{-\lambda_{2t}\tau}\right) + \varepsilon^{\tau} \tag{4.2}$$

The factor loadings of equation (4.2) are shown in Figure 4.3(b).

Svensson (1994) four factor model

Svensson (1994) included another exponential term in the model that is similar to the third one, but with a different decaying parameter. The SV four factor model in its dynamic form can be written as:

$$\begin{aligned} y_t(\tau) \;&=\; \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau} - e^{-\lambda_{1t}\tau}\right) + \\ &\quad \beta_{4t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau} - e^{-\lambda_{2t}\tau}\right) + \varepsilon^{\tau} \end{aligned} \tag{4.3}$$

The fourth component can be interpreted as a second curvature. The factor loadings of equation (4.3) are shown in Figure 4.3(c).

Five factor model

We propose a five factor model (FF) is a natural extension of the SV and includes a term that resembles the second loading, but with a different decaying parameter.

$$\begin{aligned} y_t(\tau) \;&=\; \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau}\right) + \\ &\quad \beta_{4t}\left(\frac{1-e^{-\lambda_{1t}\tau}}{\lambda_{1t}\tau} - e^{-\lambda_{1t}\tau}\right) + \beta_{5t}\left(\frac{1-e^{-\lambda_{2t}\tau}}{\lambda_{2t}\tau} - e^{-\lambda_{2t}\tau}\right) + \varepsilon^{\tau} \end{aligned} \tag{4.4}$$

The third factor loading, shown in Figure 4.3(d), can be interpreted as a second slope. We expect the FF model to enhance the fitting of more complex and twisted curves.

4.4 Estimation method

The in-sample fitting

The models presented in the previous section can be nested in the following representation:

$$Y_t = X_t \beta_t + \varepsilon_t \tag{4.5}$$

where $Y_t = [y_t(\tau_1), y_t(\tau_2), ..., y_t(\tau_N)]'$ is a N x 1 column vector representing the term structure of N interest rates at time t; X_t is the N x F matrix of factor loadings with F being the number of factors in each model; β_t is the F x 1 vector of latent factors; and ε_t is the N x 1 vector of errors with typical element ε_{nt}, for $n = 1, ..., N$, satisfying $\varepsilon_{nt} \sim iid N(0, \sigma_n^2)$. Fitting the term structure according to (4.6) requires, for each period t, estimating the vector and the decaying parameters λ_t, λ_{1t}, and λ_{2t}. Following Nelson and Siegel (1987) and Diebold and Li (2006), we keep these last parameters constant for all t, i.e. $\lambda_t = \lambda$, $\lambda_{1t} = \lambda_1$, and $\lambda_{2t} = \lambda_2$. Diebold and Li (2006) decided to fix λ at the 30 months maturity, but we adopted a less arbitrary strategy. For choosing the decaying parameter for the NS model, we initially constrained the range of λ between 0.03 and 0.42, as these values correspond to a maximum of the curvature loadings at the 5-year and 0.05-year (0.6 months) maturities, respectively. Given these boundaries, we then constructed the set $\Omega = \{0.029 + 0.001l\}_{l=1}^{391}$. Given $\lambda \in \Omega$ and the correspondent matrix of factor loadings X_t, the vector β_t is estimated by OLS at each period t. We chose the optimal decaying factor that minimizes the average of the Root Mean Squared Error (RMSE). More specifically, $\widehat{\lambda}$ is the solution of the following problem:

$$\widehat{\lambda} = \underset{\lambda \in \Omega}{arg\,min} \left\{ \frac{1}{N} \sum_{n=1}^{N} \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left(y_t(\tau_n) - \widehat{y}_t\left(\tau_n, \lambda, \widehat{\beta}_t\right) \right)^2} \right\} \tag{4.6}$$

where T is the number of yield curves in the sample. In the case of the BL, SV and FF we solved a similar problem, but now we had to find the values $\left(\widehat{\lambda}_1, \widehat{\lambda}_2\right)$ from the set $\Lambda = \{(\lambda_1, \lambda_2) \,|\, \lambda_1 \in \Omega, \lambda_2 \in \Omega\}$ i.e., the cartesian product $\Omega \times \Omega$. Thus, $\left(\widehat{\lambda}_1, \widehat{\lambda}_2\right)$ solves the following problem:

$$\left(\widehat{\lambda}_1, \widehat{\lambda}_2\right) = \underset{(\lambda_1, \lambda_2) \in \Lambda}{arg\,min} \left\{ \frac{1}{N} \sum_{n=1}^{N} \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left(y_t(\tau_n) - \widehat{y}_t\left(\tau_n, \lambda_1, \lambda_2, \widehat{\beta}_t\right) \right)^2} \right\} \tag{4.7}$$

Table 4.2 shows the optimal values of the decaying parameters for each model.

The out-of-sample forecasting

We use the first T^* yield curves to set $\widehat{\lambda}$ and $\left(\widehat{\lambda}_1, \widehat{\lambda}_2\right)$ based on the procedure described in the previous subsection. As the algorithm selects the optimal values for the decaying parameters, it simultaneously defines a time series of estimated factors $\widehat{\beta}_{f,t}$ for $t = 1,...,T^*$ and $f = 1,...,F$. Given these time series we estimate the equation

$$\widehat{\beta}_{f,t} = \mu_f + \phi_f \widehat{\beta}_{f,t-h} + v_{f,t} \tag{4.8}$$

where h is the forecast horizon, μ_f and ϕ_f are parameters to be estimated and $v_{f,t}$ is the residual whose restrictions we will comment on later.

Let $m = 0,...,(T - T^* - h)$. Given $\widehat{\lambda}$ and $\left(\widehat{\lambda}_1, \widehat{\lambda}_2\right)$, equations (4.5) and (4.8) can then be used to forecast each factor in the following iteratively manner: i) use the data available up to $T^* + m$ and estimate the parameters $\widehat{\mu}_f^{(T^*+m)}$ and $\widehat{\phi}_f^{(T^*+m)}$ according to (4.8), where the indexed superscript indicates the last observation used in the regression; ii) obtain $\widetilde{\beta}_{f,T^*+h+m|T^*+m} = \widehat{\mu}_f^{(T^*+m)} + \widehat{\phi}_f^{(T^*+m)} \widehat{\beta}_{f,T^*+m}$, where $\widetilde{\beta}_{f,T^*+h+m|T^*+m}$ is the h steps ahead forecasted value of $\widehat{\beta}_{f,T^*+h+m}$, given the information available at $T^* + m$; iii) forecast the vector of yields $\widehat{Y}_{T^*+h+m|T^*+m}$ using $\widehat{Y}_{T^*+h+m|T^*+m} = \widehat{X}\left(\widehat{\lambda}\right) \widetilde{\beta}_{T^*+h+m|T^*+m}$ or $\widehat{Y}_{T^*+h+m|T^*+m} = \widehat{X}\left(\widehat{\lambda}_1, \widehat{\lambda}_2\right) \widetilde{\beta}_{T^*+h+m|T^*+m}$.

We are particularly interested in contrasting the forecasting power of the following autoregressive models: AR(1), VAR(1), and QAR(1). When equation (4.8) is estimated according to an AR(1), $v_{f,t} \sim iid N\left(0, \sigma^2_{v_f}\right)$, for each f. In the case of a QAR(1), we only assume independence of $v_{f,t}$ over t. In both cases, we assume independence of the errors across regressions, i.e., $v_{f_1,t} \perp v_{f_2,t}$, for any $f_1 \neq f_2$. This last assumption is dropped in the case of the VAR(1) estimates, in which case (4.8) needs to be rewritten so that, instead of estimating μ_f, ϕ_f and $\sigma^2_{v_f}$, we respectively estimate the matrices M, Φ, and Ω. The VAR(1) is represented by

$$\widehat{\beta}_t = M + \Phi \widehat{\beta}_{t-h} + v_t \tag{4.9}$$

where $\widehat{\beta}_t$ is an F x 1 vector of all factors; M is an F x 1 vector representing the intercept of each equation; Φ is an F x F matrix; v_t is an F x 1 vector of serially uncorrelated residuals; and Ω is the F x F variance-covariance matrix where it is allowed $cov\left(v_f, v_g\right) \neq 0$ for any $v_f \neq v_g$. We compare the forecasting performance of each combination of term structure model / econometric method against each other and also against the predictions of a random walk applied directly on the yields: $\widehat{Y}_{\tau_n,T^*+h+m|T^*+m} = \widehat{Y}_{\tau_n,T^*+m}$.

4.5 Empirical results

The forecasting exercise is carried for horizons $h = 1$, $h = 21$ (1 month), and $h = 63$ (3 months). We have a sample of T=1883 yield curves, from which we set $T^* = 1400$. Table 4.2 shows the values for $\widehat{\lambda}$ and $\left(\widehat{\lambda}_1, \widehat{\lambda}_2\right)$ that are used in the in-sample and in the out-of-sample exercises.

In-sample fitting

Table 4.3 brings the RMSE for all four term structure models. The first thing to observe is that the FF has the smallest RMSE in all maturities analyzed. In some cases it is only a bit smaller than the results of the SV model, which also always delivers smaller RMSE than those reached by NS and BL. Incorporating the fifth factor resulted in large improvement for the shortest (1 month) and longest (60 months) maturities, exactly those considered the most difficult to fit. For these two vertices, the SV delivered an RMSE that was, respectively, 75% and 170% higher than the FF model. In the same direction, the SV resulted in much smaller RMSEs at these extreme vertices than the BL and NS. It is strongly suggested that more flexibility improves the curve adjustment, especially at extreme maturities.

In order to compare the performance of each model over time, we computed the RMSE for the entire curve of yields at each period t. The FF delivered smaller RMSE than the SV in 84.65% of the sample. The superiority against BL and NS occurred 99.78% and 99.31% of the sample, respectively. Figure 4.4(a) presents the time series of these RMSE, from which we observe FF normally reaching the smallest value.

We also computed a time series of Bayesian Information Criterion (BIC) to verify whether the greater flexibility that resulted in smaller RMSE compensates for losing degrees of freedom due to addition of extra parameters. Again, the result favors the FF in 65.90% of time when compared to the SV model. The BIC of the FF was also smaller than those of BL and NS in 89.33% and 91.82% of our sample, respectively. Figure 4.4(b) presents the BIC time series, from where we can inspect the superiority of the FF model.

To gain better intuition for why more flexible models have delivered better results, we verify, in Figure 4.5, their fitting against different yield curve shapes. Figures 4.5(a), 4.5(b) and 4.5(c) show that all models perform similarly well when facing less twisted curves. However, for curves presenting at least one inflection point, as indicated by Figures 4.5(d), 4.5(e), and 4.5(f), the most flexible models (SV and FF) do a better adjustment. In these last three figures we also observe that FF performs better than SV when exposed to more twisted curves.

Out-of-sample forecasting

Table 4.4 reports forecasting RMSEs for each term structure model and forecasting method (equation 4.8, for AR and QAR; equation 4.9 for VAR). Table 4.5 reports similar statistics for the random walk applied directly on yield levels. These tables need to be read together. Shaded boxes indicate the best forecasting model at each maturity; bold values, only present in Table 4.4, indicate outperformance over the RW forecasts.

As a pattern we observe the random walk reaching the lowest RMSEs for the 1-day horizon and all maturities, except the 24- and 60-month. For these two vertices, the best models were SV-AR and SV-QAR, respectively. The BL-QAR delivered the lowest RMSEs for the 1-month horizon and all maturities, expect the 1- and 60-month, in which the best models were NS-QAR and SV-QAR, respectively. Finally, the SV-QAR reached the lowest RMSEs for the 3-month horizon and all maturities.

The results show that the various combinations of term structure model/econometric method rarely delivered lower RMSEs than those of the random walk. For the 1-day horizon, the RW was outperformed in three occasions only: the 24-month maturity when using SV-AR and SV-QAR and the 60-month maturity when using SV-QAR. For the 1- and 3-month horizons, the random walk delivered smaller RMSEs than almost all forecasts based on the AR and VAR models. The only exception happened for the 1-month maturity, in which case all the four term structure models had smaller RMSEs when relying on VAR.

Results are quite different when we use QAR to predict 1 and 3 months ahead. For these horizons the random walk showed larger RMSEs than those computed by at least one of the term structure models. NS-QAR and BL-QAR had smaller RMSEs than RW for all maturities, expect the 60-month when forecasting 1 and 3 months ahead, respectively. The use of SV-QAR resulted in smaller RMSEs than that of the RW when forecasting 3-months ahead. SV-QAR also delivered smaller RMSEs when forecasting the 24 and the 60-month maturities 1 day and 1 month ahead. This advantage of the QAR model certainly has to do with the robustness of the median that is not affected by the presence of outliers and extreme values that are commonly observed in financial data.

When we simply compare the models against themselves, we find that the BL-

QAR generates, on average, 1 month ahead forecasts that are 40% more accurate than the best non-QAR model, the NS-AR. Similar comparison shows SV-QAR generating 3 month ahead predictions that are 32% more accurate than the FF-AR, which was the best non-QAR method.

It also interesting to note that the gain in terms of in-sample fitting obtained with the inclusion of the second slope term in the FF model was not translated into better forecasting performance. In general, the FF resulted in higher RMSEs, which may have been caused by overparameterization (Diebold and Li, 2006).

In Table 4.6 we present forecasts RMSE ratios for the RW, SV-QAR and BL-QAR. We only compare BL-QAR and SV-QAR against RW and against each other because these were the most accurate models for the 1-month and 3-month horizons, respectively. The SV-QAR also had the smallest RMSEs on average for the 1-day horizon.

Table 4.6 also shows the results for the Diebold and Mariano (1995) test, which compares the forecasting accuracy based on the null hypothesis that both models generate equal forecast mean squared errors. Asterisks indicate the DM test rejects the null at 1% (***), 5% (**), and 10% (*) significance levels. It is also important to note when comparing models SV-QAR and BL-QAR with the RW that values smaller than one indicate that these models outperform the RW by showing smaller RMSEs. When contrasting SV-QAR against BL-QAR, values smaller than one indicates that SV-QAR outperforms BL-QAR.

1-day ahead forecasting

The RW was superior than SV-QAR when forecasting the 1, 6 and 12-month maturities, if we consider a 1% confidence level, but the DM test does not allow concluding superiority of the RW forecasts for 3, 24 and 60-month maturities. Comparing to BL-QAR, the random walk is not superior only when predicting the 3-month maturity yield. When confronting SV-QAR against BL-QAR results are mixed but the DM test indicates superiority of SV-QAR when forecasting the 1, 6, 24 and 60-month maturities. BL-QAR seems superior only when predicting the 12-month yield and the DM test does not reject the null in the case of the 3-month maturity.

1-month ahead forecasting

The superiority of BL-QAR over RW is confirmed as the DM test rejects the null in favor of a smaller RMSE for the BL-QAR for all maturities, but 60-month. Opposite result was found when confronting SV-QAR against the RW. The RW delivered more accurate forecasts for all maturities, except for the 24 and 60-month maturities. When comparing SV-QAR and BL-QAR the test indicates the last model is more accurate for all maturities, except the 24 and 60-month. For the 24-month maturity the test did not reject the null, but for 60-month rejection happened at 5%, indicating superiority of SV-QAR.

3-month ahead forecasting

When comparing SV-QAR with RW, the test rejected the null for all maturities, but 60-month, indicating superior forecasting performance of SV-QAR. For the longest maturity the test did not reject the null. The DM test also did not reject the null at any maturity when confronting BL-QAR with RW. In the case of SV-QAR against BL-QAR the test favored the SV-QAR for 1, 3, 12, and 60-month maturity. For the 6 and 24-month maturities there was no rejection of the null.

4.6 Conclusions

We compare, for the Brazilian yield curve data, the in-sample adjustment and the out-of-sample forecasting performances of four different Nelson-Siegel type of models, following Diebold and Li (2006). In order to improve the ability to fit more twisted curves very commonly observed in term structures of emerging markets we propose a new five factor (FF) model, which is a natural extension of Svensson (1994). We also innovate by using quantile autoregression (QAR), evaluated at the median, to forecast the yields.

The FF model delivered the best results in terms of in-sample fitting. Visual inspection of a few very twisted curves suggests this superiority arriving from better adjustment at the most extreme yields. The FF performed poorly in out of sample forecasting, which may be due to overparameterization (Diebold and Li, 2006).

For 1 day ahead forecasting, no combination of term structure model/forecasting method beat the random walk applied directly on yield levels. The Diebold and Mariano test, however, indicated no superiority of the random walk forecasts against the Svensson‘s model (SV) combined with a QAR, for the 3, 24 and 60 months maturity.

For one month ahead forecasting, we found the Bliss model (BL)-QAR delivering superior results than the random walk for all maturities, but 60 months, for which we found no difference between the two methods. For three months ahead forecasting the best model was SV-QAR, which beat the random walk for all maturities, but the 60-month.

While the best term structure model to beat the random walk depends on the prediction horizon, the forecasts based on QAR evaluated at the median were consistently superior. The robustness of quantile regression method against outliers probably justifies our findings, especially because we have dealt with very volatile financial data, characterized by the presence of extreme values that tend to bias mean estimators.

The analysis of this paper can be extended in at least three directions. First, no-arbitrage restrictions (Christensen et al., 2007) could be applied to more flexible models such as FF, SV and BL. Second, macroeconomic variables could be used to improve the in sample adjustment and the out of sample forecasting performances of models, as suggested by Diebold et al. (2006). Finally, quantile regression, or any other robust estimation procedure, could be included among our standard forecasting methods menu.

Table 4.2: Optimal parameters λ

Notes: This table shows the optimal parameters λ obtained for the in-sample and out-of-sample exercises.

Models	In-sample fitting	Out-of-sample forecasting
NS	0.097	0.095
BL	0.048; 0.114	0.048; 0.112
SV	0.084; 0.222	0.084; 0.229
FF	0.042; 0.320	0.046; 0.259

Table 4.1: Interest Rate Descriptive Statistics

Notes: This table shows summary statistics for the Brazilian Swap ID x PRE yields. The sample ranges from March 16, 2000 to October 15, 2007. We report the mean, standard deviation, skewness, excess kurtosis, minimum, maximum, the p-value of the Jarque-Bera test for normality and the 1st, 21th and 63th sample autocorrelations.

Maturities	Mean	SD	Skew	Kurt	Min.	Max.	JB p	ρ_1	ρ_{21}	ρ_{63}
1	17.771	3.525	0.547	40.556	11.050	26.950	0.000	0.997	0.967	0.781
2	17.873	3.642	0.486	19.785	11.040	27.390	0.000	0.997	0.963	0.782
3	17.983	3.764	0.437	1.637	11.010	27.770	0.000	0.996	0.960	0.784
4	18.105	3.900	0.399	−14.370	11.020	28.170	0.000	0.996	0.959	0.784
5	18.186	4.018	0.399	−20.813	11.030	28.490	0.000	0.996	0.960	0.784
6	18.259	4.136	0.426	−21.699	10.990	28.990	0.000	0.996	0.959	0.783
7	18.330	4.248	0.454	−21.522	10.910	29.450	0.000	0.996	0.959	0.782
8	18.405	4.366	0.494	−18.014	10.850	30.130	0.000	0.997	0.958	0.779
9	18.473	4.481	0.538	−12.578	10.810	30.650	0.000	0.997	0.958	0.778
12	18.670	4.818	0.679	8.476	10.710	32.690	0.000	0.997	0.957	0.776
16	19.098	5.497	0.893	40.142	10.490	36.900	0.000	0.997	0.957	0.773
24	19.442	6.043	1.025	61.567	10.350	39.280	0.000	0.997	0.958	0.775
36	20.094	6.998	1.144	73.944	10.090	43.380	0.000	0.998	0.958	0.777
48	20.582	7.672	1.192	80.481	9.970	45.980	0.000	0.998	0.958	0.782
60 (Level)	20.944	8.130	1.199	80.924	9.840	47.430	0.000	0.998	0.958	0.787
Slope	3.174	6.171	1.429	163.078	−4.670	27.690	0.000	0.997	0.897	0.708
Curvature	−1.374	2.468	−0.781	15.961	−10.020	4.550	0.000	0.983	0.795	0.588

Table 4.3: RMSE of in-sample fitting

Model	Average	Maturities (months)								
		1	2	4	6	8	12	24	36	48	60
NS	0.1735	0.2755	0.1252	0.1464	0.1554	0.1201	0.1540	0.2903	0.2182	0.1034	0.2401
BL	0.1657	0.2513	0.0996	0.1430	0.1410	0.1134	0.1646	0.2485	0.2087	0.1058	0.2321
SV	0.1119	0.1236	0.0679	0.0977	0.0611	0.0705	0.1462	0.1284	0.1610	0.0988	0.1370
FF	0.0859	0.0706	0.0673	0.0680	0.0555	0.0641	0.1206	0.1264	0.0666	0.0829	0.0507

Notes: This table shows the in-sample fitting errors of models NS, BL, SV and FF.

Table 4.4: RMSEs of out-of-sample forecasts – NS class models

Notes: This table shows the RMSEs of the out-of-sample forecasts generated from combinations of term structure models (NS, BL, SV and FF) and forecasting methods (AR(1), VAR(1) and QAR(1)). Bold values indicate which models beat the RW and shaded-boxes indicate the best forecasting model for each maturity.

Maturities (months)	AR(1)			VAR(1)			QAR(1)		
	1 day	1 month	3 months	1 day	1 month	3 months	1 day	1 month	3 months
					NS				
1	0.0976	0.4418	1.2760	0.1011	**0.1929**	**0.8368**	0.1015	**0.1914**	**0.8511**
3	0.0378	0.5216	1.3795	0.0411	0.4383	1.3625	0.0373	**0.2653**	**0.8623**
6	0.0955	0.6267	1.5268	0.1076	0.7214	1.9709	0.0940	**0.3616**	**0.9140**
12	0.0886	0.6696	1.6747	0.1001	0.9603	2.6254	0.0871	**0.4076**	**0.9905**
24	0.1570	0.7732	1.9533	0.1465	1.1207	3.0799	0.1590	**0.5457**	**1.2081**
60	0.1836	1.0848	2.5413	0.1901	1.2858	3.2215	0.1801	0.7297	1.5508
Average	0.1100	0.6863	1.7253	0.1144	0.7866	2.1828	0.1098	**0.4169**	**1.0628**
					BL				
1	0.0833	0.4906	1.3359	0.0859	**0.1957**	**0.7971**	0.0907	**0.1938**	**0.9175**
3	0.0398	0.5825	1.4968	0.0422	0.4404	1.3572	0.0376	**0.2521**	**0.8962**
6	0.0938	0.7069	1.7080	0.0998	0.7226	1.9853	0.0820	**0.3428**	**0.9188**
12	0.0889	0.7750	1.9182	0.0984	0.9592	2.6252	0.0872	**0.3971**	**0.9921**
24	0.1475	0.8804	2.1965	0.1439	1.1140	3.0292	0.1558	**0.5415**	**1.2039**
60	0.1780	1.1306	2.6899	0.1827	1.2943	3.2551	0.1700	0.7203	1.5272
Average	0.1052	0.7610	1.8909	0.1088	0.7877	2.1748	0.1039	**0.4079**	**1.0759**
					SV				
1	0.0634	0.6135	1.4585	0.0548	**0.2678**	**0.7627**	0.0589	0.4838	**0.6728**
3	0.0382	0.7009	1.6472	0.0386	0.4883	1.3676	0.0389	0.4897	**0.7116**
6	0.0604	0.8066	1.8198	0.0647	0.7691	2.0036	0.0560	0.5375	**0.7872**
12	0.0932	0.8218	1.8751	0.1087	1.0326	2.6143	0.0911	0.5208	**0.8615**
24	**0.1216**	0.8359	1.9631	0.1298	1.2261	3.0229	0.1223	0.5694	**1.0827**
60	0.1469	1.0090	2.3820	0.1627	1.4020	3.2344	**0.1451**	**0.6710**	**1.3443**
Average	0.0873	0.7979	1.8576	0.0932	0.8643	2.1676	0.0854	0.5454	**0.9100**
					FF				
1	0.1212	1.4974	1.3616	0.0456	**0.2512**	**0.7485**	0.0746	1.0858	1.6758
3	0.1084	1.2797	1.1305	0.0373	0.4887	1.3642	0.0842	1.0179	1.5496
6	0.1339	1.0427	**1.0080**	0.0630	0.8052	2.0445	0.0768	0.9188	1.4211
12	0.1501	0.8940	1.1139	0.1032	1.1138	2.7100	0.1175	0.9003	1.4218
24	0.1706	0.8361	1.3941	0.1280	1.3178	3.1171	0.1441	0.9300	1.5017
60	0.2059	0.8473	2.0040	0.1644	1.5334	3.4143	0.1605	0.8769	1.4778
Average	0.1484	1.0662	1.3354	0.0902	0.9184	2.2331	0.1096	0.9550	1.5080

Table 4.5: RMSE of out-of-sample forecasts – RW model

Notes: This table shows the RMSEs of the out-of-sample forecasts generated from a random-walk model. Shaded-boxes indicate the best forecasting model for each maturity.

Maturities (months)	RW		
	1 day	1 month	3 months
1	0.0386	0.3851	1.1251
3	0.0369	0.3832	1.1120
6	0.0522	0.4011	1.0812
12	0.0845	0.4650	1.1139
24	0.1234	0.5993	1.2872
60	0.1464	0.6821	1.3910
Average	0.0803	0.4860	1.1851

Table 4.6: Ratios of forecast RMSEs, Diebold-Mariano tests

Notes: This table shows the forecast RMSE ratios and the Diebold–Mariano forecast accuracy tests for the best forecasting models. The null hypothesis is that the two forecasts being compared have the same mean squared errors. Asterisks denote significance relative to the asymptotic null distribution (10% (*), 5% (**) and 1% (***)).

Maturities	SV-QAR(1)/RW			BL-QAR(1)RW			SV-QAR(1)/BL-QAR(1)		
	1 day	1 month	3 months	1 day	1 month	3 months	1 day	1 month	3 months
1	1.5257***	1.2564***	0.5979**	2.3491***	0.5032***	0.8155	0.6495***	2.4968***	0.7332**
3	1.0539	1.2779***	0.6399**	1.0167	0.6579***	0.8060	1.0366	1.9424***	0.794**
6	1.0714***	1.3398***	0.7281*	1.5698***	0.8545**	0.8498	0.6825***	1.5681***	0.8568
12	1.0781***	1.1201*	0.7734*	1.0319**	0.854**	0.8906	1.0448***	1.3116***	0.8684*
24	0.9907	0.95*	0.8411*	1.2625***	0.9035**	0.9353	0.7847***	1.0515	0.8994
60	0.9914	0.9838	0.9665	1.1614***	1.0560	1.0979	0.8536***	0.9316**	0.8803*

Figure 4.1: ID x PRE Swap Yield Curves

Notes: This figure shows the ID x PRE swap yield curves. The sample ranges from March 16, 2000 to October 15, 2007 and the maturities are (in months): 1, 2, 3, 4, 5, 6, 7, 8, 9, 12, 16, 24, 36, 48, and 60.

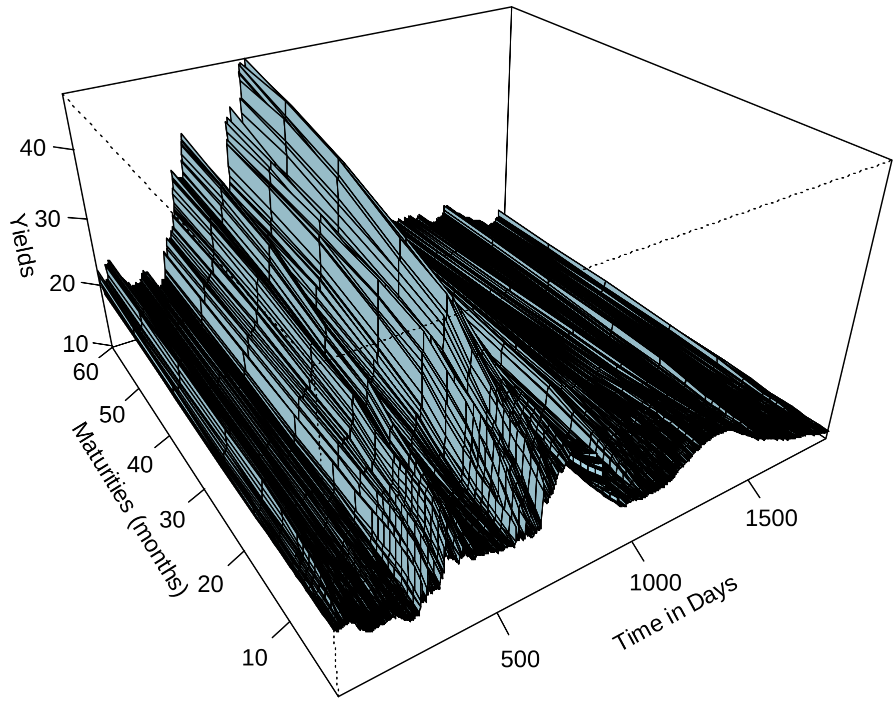


Figure 4.2: Empirical Factors and EMBI – Brazil

Notes: Empirical factors are left scaled and the EMBI Brazil is right scaled. The level is defined as the 60-month yield, the slope as the 60-month less 1-month yield and the curvature as twice the 1-year yield less the sum of the 1-month and 60-month yields.

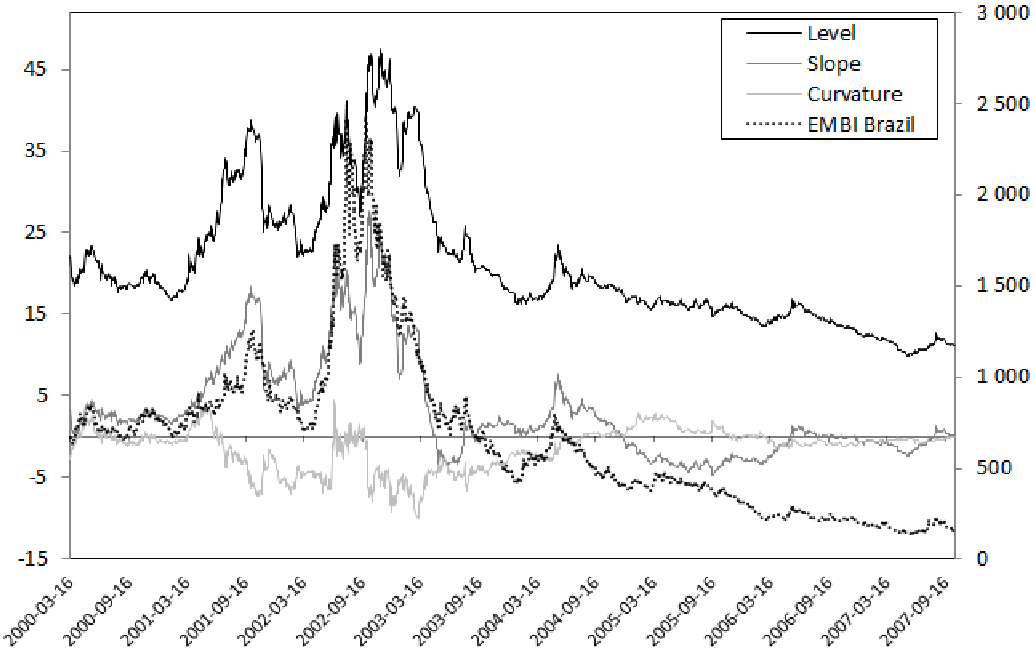


Figure 4.3: Loadings of the NS Class Models

Notes: This figure shows factor loadings for the NS, Bl, SV and FF models computed considering $\lambda = 0.097$, $\lambda_1 = 0.048$ and $\lambda_2 = 0.114$, $\lambda_1 = 0.084$ and $\lambda_2 = 0.222$, $\lambda_1 = 0.042$ and $\lambda_2 = 0.320$, respectively.

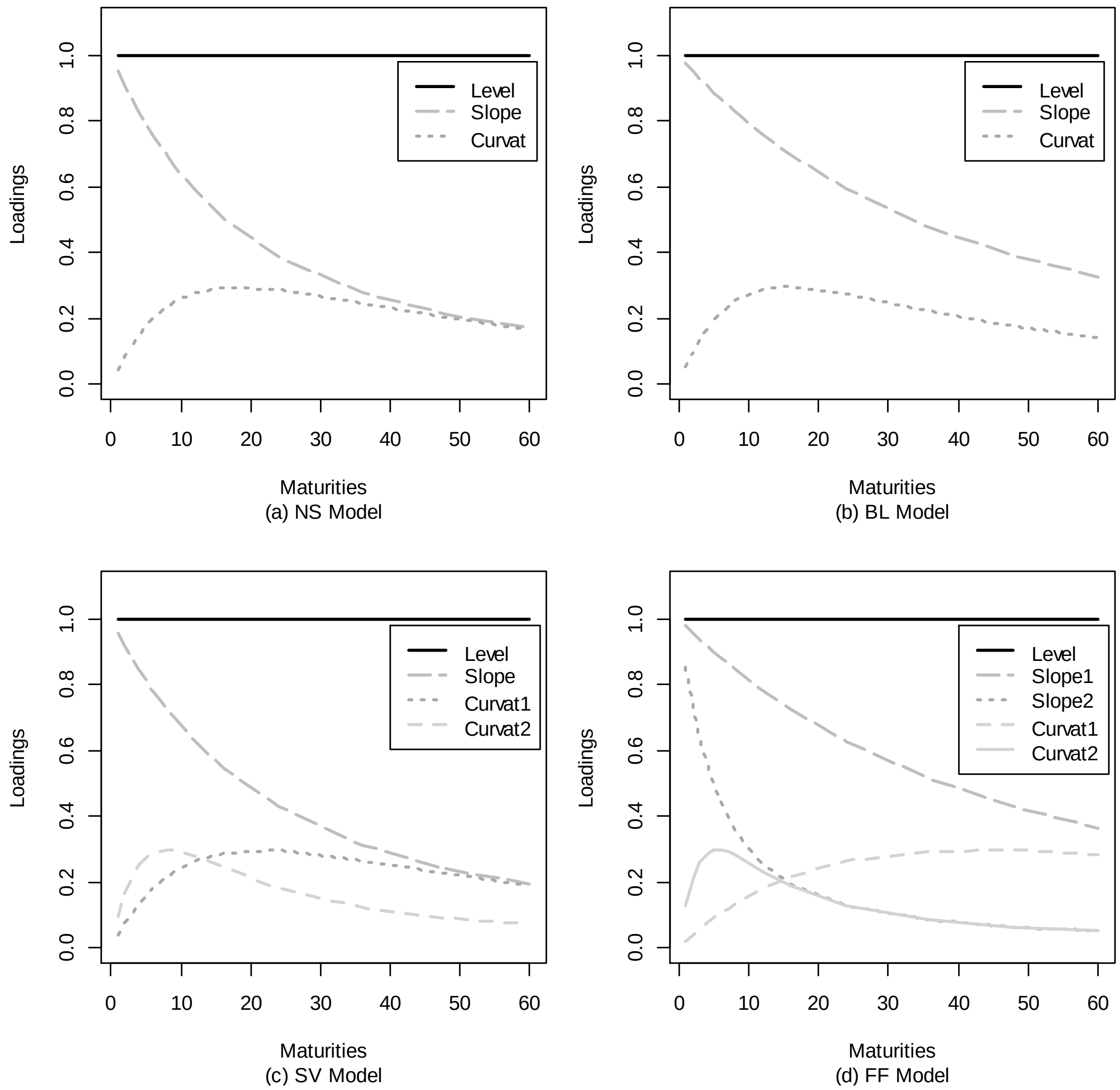
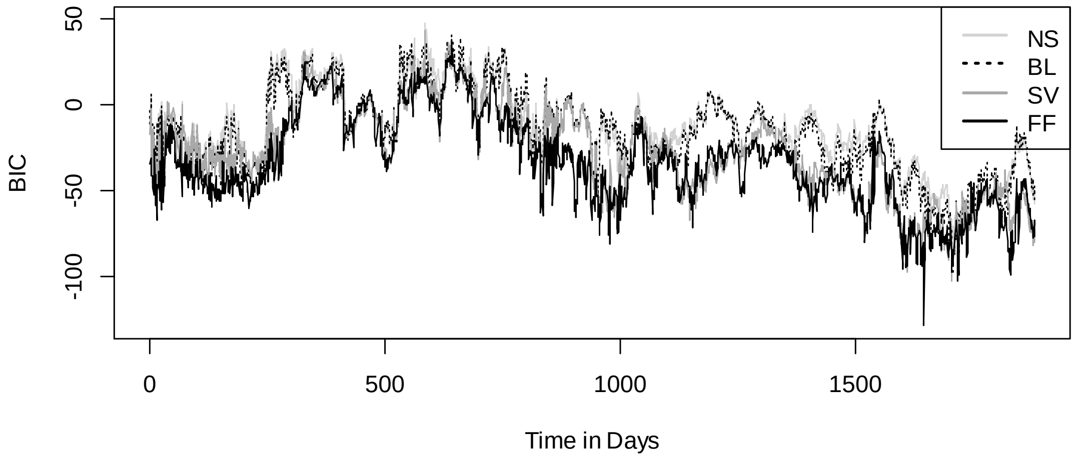
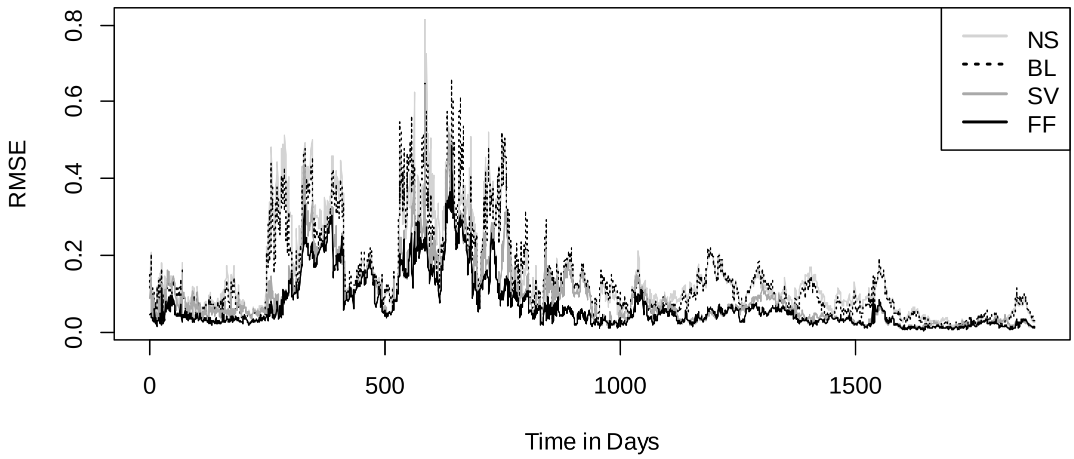


Figure 4.4: Fitting RMSE and BIC – NS Class Models

Notes: This figure shows fitting RMSE and BIC statistics estimated for each day of the sample that ranges from March 16, 2000 to October 15, 2007.



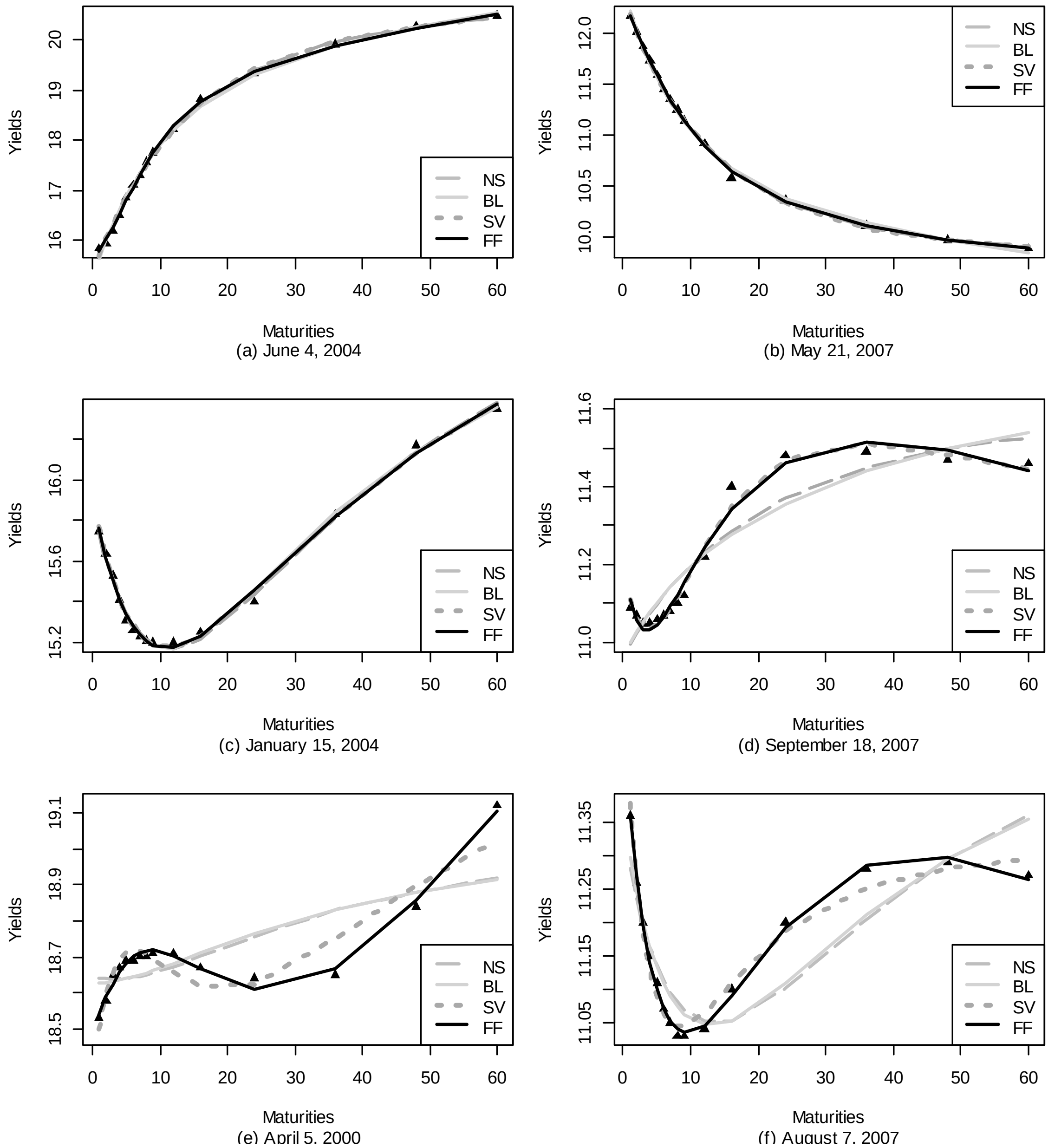
(a) Bayesian Criterion - BIC



(b) RMSE

Figure 4.5: Fitted Yield Curves in specific days

Notes: This figure shows observed and fitted yield curves for the following days: (a) June 4, 2004; (b) May 21, 2007; (c) January 15, 2004; (d) September 18, 2007; (e) April 4, 2000; and (f) August 7, 2007.



(a) June 4, 2004

(b) May 21, 2007

(c) January 15, 2004

(d) September 18, 2007

(e) April 5, 2000

(f) August 7, 2007

Bibliography

[1] Aguiar-Conraria, L., Martins, M. M., and Soares, M. J. (2012), "The yield curve and the macro-economy across time and frequencies", Journal of Economics Dynamics and Control, 36(12), p. 1950-1970.

[2] Aiolfi, M., Capistrán, C. and Timmermann, A. (2011). "Forecast Combination". In M.P. Clements and D.F. Hendry (Eds.) The Oxford Handbook of Economic Forecasting, Oxford University Press.

[3] Andrade P., Ghysels E. and Idier J. (2012). "Tails of inflation forecasts and tales of monetary policy", mimeo, Banque de France.

[4] Andrade P., Fourel, V., Ghysels E. and Idier J. (forthcoming). "The financial content of inflation risks in the euro area". International Journal of Forecasting.

[5] Ang, A. and Piazzesi, M. (2003). "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables". Journal of Monetary Economics, v. 50, p. 745-787.

[6] Ang, A., Bekaert G. and Wei, M. (2007). "Do macro variables, asset markets, or surveys forecast inflation better?". Journal of Monetary Economics, v. 54, p. 1163-1212.

[7] Ang, A., Dong, S. and Piazzesi, M. (2007). "No-arbitrage Taylor Rules". NBER Working Paper No. 13448.

[8] Bai, J. and Ng, S. (2002). "Determining the Number of Factors in Approximate Factor Models". Econometrica, v. 70, p. 191–221.

[9] Bank for International Settlements (2005). "Zero-Coupon Yield Curves: Technical Documentation". BIS papers No. 25. http://www.bis.org/publ/bppdf/bispap25.pdf

[10] Bansal, R., and Yaron, A. (2004). “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles”. Journal of Finance, v. 59, p. 1481–1509.

[11] Bansal, R. and Shaliastovich, I. (2013). “A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets”. Review of Financial Studies, v. 26, p. 1-33.

[12] Bauer, M. D., Rudebusch, G. D. and Wu, C. (2012). “Correcting Estimation Bias in Dynamic Term Structure Models”. Journal of Business & Economic Statistics, v. 30, p. 454-467.

[13] Beechey, M. and Wright, J. (2009). “The High-Frequency Impact of News on Long-Term Yields and Forward Rates: Is it Real?”. Journal of Monetary Economics, v. 56, p. 535-544.

[14] Bernanke, B. (2004). “The Great Moderation”. Speech at the Meeting of the Eastern Economic Association, Washington DC, February 20.

[15] Bernard, H. and Gerlach, S. (1998). “Does the Term Structure Predict Recessions? The International Evidence.” International Journal of Finance and Economics, v. 3(3), pp. 195-215.

[16] Bekaert, G., Hodrick R. J. and Marshall D. (1997). “On Biases in Tests of the Expectations Hypothesis of the Term Structure of Interest Rates”. Journal of Financial Economics, v. 44, p. 309-348.

[17] Bank for International Settlements. 2005. Zero-Coupon Yield Curves: Technical Documentation. BIS papers No. 25. http://www.bis.org/publ/bppdf/bispap25.pdf

[18] Boivin, J., & Giannoni, M. (2006), “DSGE models in a data-rich environment”, National Bureau of Economic Research, WP 12772.

[19] Boivin, J., Kiley, M. T., and Mishkin, F. S. (2010), “How has the monetary transmission mechanism evolved over time?”, National Bureau of Economic Research, WP 15879.

[20] Bolder D. J. (2006). “Modelling Term-Structure Dynamics for Risk Management: A Practitioner’s Perspective”. Bank of Canada Working Paper No. 2006–48.

[21] Bowley, A. L. (1920). “Elements of Statistics”. New York: Scribner.

[22] Buraschi, A. and Whelan P. (2012). “Term Structure Models and Differences in Beliefs”. Available at: “http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1887644”.

[23] Campbell, J. Y. and J. H. Cochrane. (1999). “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior”. Journal of Political Economy, v. 107, p. 205–251.

[24] Campbell, J. Y. and Thompson, S. B. (2008). “Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?”. Review of Financial Studies, v. 21, p. 1509–1531.

[25] Campbell, J. Y. and Viceira, L. M. (2002). “Strategic Asset Allocation”. Oxford University Press.

[26] Capistrán, C. (2008). “Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious?”. Journal of Monetary Economics, v. 55, p. 1415–1427.

[27] Capistrán, C. and Timmermann, A. (2009). “Forecast Combination with Entry and Exit of Experts”. Journal of Business and Economic Statistics, v. 27, p. 428-440.

[28] Cheznozhukov, V., Fernandez-Val, I. and Galichon, A. (2010). “Quantile and Probability Curves Without Crossing”. Econometrica, v. 78, p. 1093-1125.

[29] Christensen J., Diebold F. X., Rudebusch G. (2007). “The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models”. NBER Working Papers No. 13611.

[30] Christensen, R., Lopez, J. and Rudebusch, G. (2011). “Extracting deflation probability forecasts from treasury yields”, FRBSF Working Paper 2011-10.

[31] Christoffersen, P. F. and Diebold, F. (1997). “Optimal prediction under asymmetric loss”. Econometric Theory, v. 13, p. 808-817.

[32] Chun, A. L. (2011). “Expectations, Bond Yields and Monetary Policy”. Review of Financial Studies, v. 24(1), p. 208-247.

[33] Clarida, R., Galí, J. and Gertler, M. (2000). "Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory", The Quarterly Journal of Economics, v. 115(1), p. 147-180.

[34] Clark, T. (2009). “Is the Great Moderation Over? An Empirical Analysis”. Economic Review, Federal Reserve Bank of Kansas City, p. 5-42

[35] Clark, T. E. and McCracken, M. W. (2001). “Tests of equal forecast accuracy and encompassing for nested models" Journal of Econometrics, v. 105, p. 85-110.

[36] Clark, T. and McCracken (2012). “Advances in Forecast Evaluation”. Working Paper Series 2011-025B, Federal Reserve Bank of St. Louis

[37] Clark, T. E. and West, K. D. (2007). “Approximately Normal Tests for Equal Predictive Accuracy in Nested Models”. Journal of Econometrics, v.138, p. 291-311.

[38] Clatworthy, M. A., Peel, D. and Pope, P. F. (2012). “Are Analysts’ Loss Functions Asymmetric?”. Journal of Forecasting, v. 31, p. 736-756.

[39] Cochrane, J. H. and Piazzesi, M. (2005). “Bond risk premia". American Economic Review, v. 95, p. 138-160.

[40] Cochrane, J. H. and Piazzesi, M. (2008). “Decomposing the Yield Curve". Working Paper, University of Chicago.

[41] Colacito, R., Ghysels, E. and Meng, J. (2013). ”Skewness in Expected Macro Fundamentals and the Predictability of Equity Returns: Evidence and Theory”. Available at: “http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2185659”.

[42] Conrad, J., Dittmar, R. and Ghysels, E. (2013). “Ex Ante Skewness and Expected Stock Returns”. Journal of Finance, v. 68, p. 85-124.

[43] Cooper, I. and R. Priestley (2009). “Time-varying risk premiums and the output gap". Review of Financial Studies, v. 22, p. 2801.

[44] Cox J, Ingersoll J.E. and Ross S.A. (1985). “A Theory of the Term Structure of Interest Rates”. Econometrica, v. 53, p. 385–407.

[45] D’Agostino, A., Gianonne, D. and Surico, P. (2006). “(Un)Predictability and Macroeconomic Stability”. Working Paper Series 605, European Central Bank.

[46] Dangl, T. and Halling, M. (2007). “Predictive Regressions with Time-Varying Coefficients”. Journal of Financial Economics, v. 106, p. 157-181.

[47] De Rezende, R. and Ferreira, M. (2013). “Re-examining the Forecasting Power of the Yield Curve with Quantile Regression”. mimeo.

[48] Della Corte, P., Sarno, L. and Thornton, D. L. (2008). “The Expectations Hypothesis of the Term Structure of Very Short-term Rates: Statistical Tests and Economic Value”. Journal of Financial Economics, v. 89, p. 158–174.

[49] Dick, C., Schmeling, M. and Schrimpf, A. (2013). “Macro Expectations, Aggregate Uncertainty, and Expected Term Premia”. European Economic Review v. 58, p. 58-80.

[50] Diebold F. X. and Li C. (2006). “Forecasting the Term Structure of Government Bond Yields”. Journal of Econometrics, v. 130, p. 337–364.

[51] Diebold, F. X. and Mariano, R. S. (1995). “Comparing Predictive Accuracy”. Journal of Business and Economics Statistics, v. 13(3), p. 253-265.

[52] Diebold, F. X., and Rudebusch, G. D. (1989). “Scoring the Leading Indicators”. Journal of Business, v. 62(3), p. 369-391.

[53] Diebold F. X., Rudebusch G. D. and Aruoba B. (2006). “The Macroeconomy and the Yield Curve: a Dynamic Latent Factor Approach”. Journal of Econometrics, v. 131, p. 309–338.

[54] Duarte, A., Venetis, I. A. and Paya, I. (2005) “Predicting Real Growth and the Probability of Recession in the Euro Area Using the Yield Spread.” International Journal of Forecasting, v. 21(2), p. 262-277.

[55] Duffee, G. and Stanton, H. (2004). “Estimation of Dynamic Term Structure Models”. Working Paper, Haas School of Business.

[56] Duffee, G. (2011). “Information in (and not in) the term structure”. Review of Financial Studies, v. 24(9), p. 2895-2934.

[57] Duffie D. and Kan, R. (1996). “A Yield Factor Model of Interest Rates”. Mathematical Finance, v. 6, p. 379–406.

[58] Elliot, G., Timmermann, A. and Komunjer, I. (2005). "Estimating Loss Function Parameters". Review of Economic Studies, v. 72, p. 1107-1125.

[59] Elliot, G., Timmermann, A. and Komunjer, I. (2008). “Biases in Macroeconomic Forecasts: Irrationality or Asymmetric Loss?”. Journal of the European Economic Association, v. 6(1), p. 122-157.

[60] Estrella, A. and Hardouvelis, G. A. (1991). “The Term Structure as a Predictor of Real Economic Activity”. Journal of Finance, v. 46(2), p. 555-76.

[61] Estrella, A. and Mishkin, F. S. (1998). “Predicting U.S. Recessions: Financial Variables as Leading Indicators.” Review of Economics and Statistics, v. 80(1), p. 45-61.

[62] Estrella, A., Rodrigues, A. P. and Schich, S. (2003). “How Stable Is the Predictive Power of the Yield Curve? Evidence from Germany and the United States.” Review of Economics and Statistics, v. 85(3), p. 629-644.

[63] Fama, E. F. (1984). “The information in the term structure”. Journal of Financial Economics, v. 6(4), p. 509-528.

[64] Fama, E. F. and Bliss, R.R. (1987). “The information in long-maturity forward rates". American Economic Review, v. 77, p. 680-692.

[65] Faust J. and Wright, J. (forthcoming). “Forecasting Inflation”. Prepared for the Handbook of Forecasting.

[66] Ferreira, M. S. (2011) “Capturing asymmetries in real exchange rates with quantile autoregression”. Applied Economics, v. 43(3), p. 327-340.

[67] Fleming, J., Kirby, C. and Ostdiek, B. (2001). “The Economic Value of Volatility Timing”. Journal of Finance, v. 56, p. 329-352.

[68] Fisher M., Nychka D. and Zervos D. (1995). “Fitting the Term Structure of Interest Rates with Smoothing Splines”. Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series No. 1995–1.

[69] Friedman, B. and Kuttner, K. (1998). “Indicator Properties of the Paper-Bill Spread: Lessons from Recent Experience”. Review of Economics and Statistics, v. 80, p. 34-44.

[70] Gabaix, X. (2012). “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance”. Quarterly Journal of Economics, v. 127(2), p. 645-700.

[71] Gaglianone, W., Lima L. R., Linton, O., and Smith D. R. (2011). “Evaluating Value-at-Risk Models via Quantile Regressions”. Journal of Business and Economic Stastistics, v. 29, p. 150-160.

[72] Gaglianone, W. and Lima, L.R. (2012). "Constructing Density Forecasts from Quantile Regressions". Journal of Money, Credit and Banking, v. 44(8), p. 1589-1607.

[73] Galbraith, J. W. and Tkacz, G. (2000). "Testing for Asymmetry in the Link Between the Yield Spread and Output in the G-7 Countries." Journal of International Money and Finance, v. 19(5), p. 657-72.

[74] Galí, J., and Gambetti, L. (2009), "On the sources of the great moderation", American Economic Journal: Macroeconomics, v. 1(1), p. 26-57.

[75] Galí, J., Lopez-Salido, J. D., and Valles, J. (2003). "Technology shocks and monetary policy: assessing the Fed's performance", Journal of Monetary Economics, v. 50(4), p. 723-743.

[76] Garcia, R. and Perron, P. (1996). "An Analysis of the Real Interest Rate under Regime Shifts". The Review of Economics and Statistics, v. 78(1), p. 111-125.

[77] Garcia, J. A. and Werner, T. (2010). "Inflation Risks and Inflation Risk Premia". ECB Working Paper 1162.

[78] Ghysels, E., A. Plazzi, and R. Valkanov. (2010). "Conditional skewness of stock market returns in developed and emerging markets and its economic fundamentals". Swiss Finance Institute Research Paper No. 11-06. Available at SSRN: http://ssrn.com/abstract= 1761446.

[79] Ghysels, E., Casidhe H. and Moench, E. (2012). "Forecasting through the Rear-View Mirror: Data Revisions and Bond Return Predictability". Staff Reports 581, Federal Reserve Bank of New York.

[80] Giacomini, R. and Rossi, B. (2006). "How Stable Is the Forecasting Performance of the Yield Curve for Output Growth?" Oxford Bulletin of Economics and Statistics, v. 68(1), p. 783-795.

[81] Giordani, P. and Söderlind, P. (2003). "Inflation forecast uncertainty". European Economic Review, v. 47, p. 1037-1059.

[82] Gimeno R. and Nave J. (2006). "Genetic Algorithm Estimation of Interest Rate Term Structure". Documentos de Trabajo del Banco de España No. 0634.

[83] Gneiting, T. (2011). “Quantiles as optimal point forecasts". International Journal of Forecasting, v. 27, p. 197-207.

[84] Goetzmann, W., Ingersoll, J., Spiegel, M. and Welch, I. (2007). “Portfolio Performance Manipulation and Manipulation-proof Performance Measures”. Review of Financial Studies, v. 20, p. 1503-1546.

[85] Golob, J. (1994). “Does inflation uncertainty increase with inflation?”. Economic Review, issue Q III, p. 27-38.

[86] Gourio, F. (2013). “Credit Risk and Disaster Risk”. American Economic Journal: Macroeconomics, v. 5, p. 1-34.

[87] Granger, C. (1969). “Prediction with a generalized cost of error function”. Operational Research Quarterly, v. 20, p. 199-207.

[88] Granger, C. and Newbold, P. (1986). “Forecasting time series”. 2nd Edition. Academic Press.

[89] Gürkaynak, R., Sack, B. and Swanson, E. (2005). “The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models”. American Economic Review, v. 95(1), p. 425-436.

[90] Gürkaynak, R., Sack, B. and Wright, J. (2007). “The U.S. Treasury Yield Curve: 1961 to the Present”. Journal of Monetary Economics, v. 54, p. 2291-2304.

[91] Hamilton, J. and Kim, D. H. (2002). "A Reexamination of the Predictability of Economic Activity Using the Yield Spread". Journal of Money, Credit and Banking, v. 34(2), p. 340-360.

[92] Hannan, E. and Quinn, B. (1979). “The Determination of the Order of an Autoregression”. Journal of Royal Statistical Society B, v. 41, p. 190-195.

[93] Harvey, C. R. (1988). “The Real Term Structure and Consumption Growth,” Journal of Financial Economics, v. 22(2), p. 305-33.

[94] Harvey, C. R. (1989). “Forecasts of Economic Growth From the Bond and Stock Markets.” Financial Analysts Journal, v. 45(5), p. 38-45.

[95] Harvey, D. I., Leybourne, S. J., and Newbold, P. (1997). “Testing the equality of prediction mean squared errors”. International Journal of Forecasting v. 13(2), p. 281–291.

[96] Haubrich, J.G., Dombrosky, A.M. (1996), “Predicting Real Growth Using the Yield Curve", Economic Review, v. 32, p. 26-34.

[97] Heath D., Jarrow R. and Morton A. (1992). “Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation”. Econometrica, v. 60, p. 77–105.

[98] Hinkley, D. V. (1975). “On power transformations to symmetry”. Biometrika, v. 62, p. 101-111.

[99] Hotta L. K. (1993). “The effect of additive outliers on the estimates of aggregated and disaggregated ARIMA models”. International Journal of Forecasting, v. 9, p. 85–93.

[100] Huang, J. Y. and Shi, Z. (2012). “Determinants of Bond Risk Premia”. Working paper, Penn State University.

[101] Hull J. and White A. (1990). “Pricing interest-rate-derivative securities”. Review of Financial Studies, v. 3, p. 573–592.

[102] Inoue, A. and Kilian, L. (2004). “In-sample or Out-of-sample Tests of Predictability: Which One Should We Use?”. Econometric Reviews, v. 23, p. 371-402.

[103] Inoue, A. and Kilian, L. (2006). “On the Selection of Forecasting Models”. Journal of Econometrics, v. 130, p. 273-306.

[104] Joslin, S., Priebsch, M. and Singleton, K. (2010). “Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks”. Working Paper, Graduate School of Business, Stanford University.

[105] Joslin, S., Priebsch, M. and Singleton, K. (forthcoming). “Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks”. Journal of Finance .

[106] Jurado, K., Ludvigson, S. and Ng, S. (2013). “Measuring Uncertainty”. Working Paper, Dept. of Economics, New York University.

[107] Kim, D. H. and Orphanides, A. (2005). “Term Structure Estimation with Survey Data on Interest Rate Forecasts”. Computing in Economics and Finance 474, Society for Computational Economics

[108] Kim, C-L., and Nelson, C.R. (2006), "Estimation of a Forward-Looking Monetary Policy Rule: A Time-Varying Parameter Model using Ex-Post Data", Journal of Monetary Economics, v. 53(8), p. 1949-1966.

[109] Kim, D. H. and Wright, J. (2005). "An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates". Finance and Economics Discussion Series 2005-33, Board of Governors of the Federal Reserve System.

[110] Kim, D. H. (2008). "Challenges in Macro-Finance Modeling." Federal Reserve Board Finance and Economics Discussion Series, 2008-6.

[111] Kim, T. H. and White, H. (2004). "On more robust estimation of skewness and kurtosis". Finance Research Letters, v. 1(1), p. 56-73.

[112] Kitsul, Y. and Wright, J. H. (forthcoming). "The economics of options-implied inflation probability density functions". Journal of Financial Economics.

[113] Koenker, R. (2005). "Quantile Regression". Econometric Society Monographs No. 38, Cambridge University Press.

[114] Koenker, R. and Bassett, G. (1978). "Regression Quantiles". Econometrica, v. 46(1), p. 33–50.

[115] Koenker, R. and d'Orey, V. (1987). "Computing Regression Quantiles". Journal of the Royal Statistical Society C, v. 36, p. 383-393.

[116] Koenker, R. and d'Orey, V. (1994) "Remark on Algorithm AS 229: Computing dual regression quantiles and regression rank scores". Applied Statistics, v. 43, p. 410-414.

[117] Koenker, R. and Hallock, K. F. (2001). "Quantile Regression". Journal of Economic Perspectives, v. 15(4), p. 143–156.

[118] Koenker, R. and Machado, J.A.F. (1999). "Goodness of fit and related inference processes for quantile regression". Journal of the American Statistical Association, v. 94, p. 1296–1310.

[119] Koenker, R. and Xiao, Z. (2004) "Unit root quantile autoregression inference". Journal of the American Statistical Association, v. 99, p. 775-787.

[120] Koenker, R. and Xiao, Z. (2006) "Quantile Autoregression". Journal of the American Statistical Association, v. 101, p. 980-990.

[121] Komunjer, I. (2005). “Quasi-Maximum Likelihood Estimation for Conditional Quantiles". Journal of Econometrics, v. 128, p. 137-164.

[122] Komunjer, I. and Vuong, Q. (2010). “Efficient Estimation in Dynamic Conditional Quantile Models". Journal of Econometrics, v. 157, p. 272-285.

[123] Lahiri K, Liu F. (2006). “Modeling multi-period inflation uncertainty using a panel of density forecasts”. Journal of Applied Econometrics, v. 21, p. 1199-1220.

[124] Laurent, R. D. (1988). “An Interest Rate-Based Indicator of Monetary Policy.” Federal Reserve Bank of Chicago Economic Perspectives, v. 12(1), p. 3-14.

[125] Laurini M. and Hotta L. K. (2007). “Bayesian Extensions to the Diebold and Li Term Structure Model”. IBMEC Working Paper No. 40.

[126] Ledolter J. (1989). “The effect of additive outliers on the forecasts from ARIMA models”. International Journal of Forecasting, v. 5, p. 231–240.

[127] Litterman R. and Scheinkman J. (1991). “Common Factors Affecting Bond Returns”. Journal of Fixed Income, v. 1, p. 54–61.

[128] Ludvigson, S.C. and Ng, S. (2007). “The Empirical Risk-Return Tradeoff: A Factor Analysis Approach”. The Journal of Financial Economics, v. 83, p. 171-222.

[129] Ludvigson, S.C. and Ng, S. (2009). “Macro factors in bond risk premia". Review of Financial Studies, v. 22, p. 5027-5067.

[130] Ludvigson, S.C. and Ng, S. (2010). “A Factor Analysis of Bond Risk Premia”. Handbook of Empirical Economics and Finance, e.d. by Aman Uhla and David E. A. Giles, p. 313-372.

[131] Litterman, R. and Scheinkman, J. (1991). “Common factors affecting bond returns”. Journal of Fixed Income, v. 1, p. 54-61.

[132] McCracken, M. W. (2007). “Asymptotics for Out-of-Sample Tests of Granger Causality”. Journal of Econometrics, v. 140, p. 719-752.

[133] McCulloch J. H. (1971). “Measuring the Term Structure of Interest Rates”. Journal of Business, v. 44, p. 19–31.

[134] McCulloch J. H. (1975). “The tax adjusted yield curve”. Journal of Finance, v. 30, p. 811–830.

[135] Mishkin, F. S. (1990). “What does the term structure tell us about future inflation?”. Journal of Monetary Economics, v. 25, p. 77-95.

[136] Möench E. (2008). “Forecasting the Yield Curve in a Data-Rich Environment: a No-Arbitrage Factor Augmented VAR Approach”. Journal of Econometrics, v. 146, p. 26–43.

[137] National Bureau of Economic Research (2003). “The NBER’s Business-Cycle Dating Procedure”, unpublished manuscript.

[138] Nelson C. R. and Siegel A. F. (1987). “Parsimonious Modeling of Yield Curves”. Journal of Business, v. 60, p. 473–489.

[139] Newey, W. K. and West, K. D. (1987). “A simple, positive semi-denite, heteroskedasticity and autocorrelation consistent covariance matrix”. Econometrica, v . 55, p. 703-708.

[140] Oka, T. and Qu, Z. (2011). “Estimating Structural Changes in Regression Quantiles”. Journal of Econometrics, v. 162, p. 248-267.

[141] Ottaviani, M. and Sorensen, P. N. (2006). “The strategy of professional forecasting”. Journal of Financial Economics, v. 81, p. 441-466.

[142] Patton, A.J. and Timmermann, A. (2007). “Testing Forecast Optimality under Unknown Loss”. Journal of the American Statistical Association, v. 102, p. 1172-1184.

[143] Patton, A.J. and Timmermann, A. (2010). “Why do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion”. Journal of Monetary Economics, v. 57, p. 803-820

[144] Piazzesi, M., Salomao, J and Schneider, M. (2013). “Trend and Cycle in Bond Premia”. Working Paper, Depart. of Economics, Stanford University.

[145] Qu, Z. (2008). “Testing for Structural Change in Regression Quantiles”. Journal of Econometrics, v. 148, p. 170-184.

[146] Rudebusch, G., Sack, B. and Swanson, E. (2007). "Macroeconomic implications of changes in the term premium," Review, Federal Reserve Bank of St. Louis, p. 241-270.

[147] Rudebusch, G. and Swanson, E. (2012). “The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks”. American Economic Journal: Macroeconomics, v. 4, p. 105-143.

[148] Rudebusch, G. and Williams, J. (2009). “Forecasting Recessions: The Puzzle of the Enduring Power of the Yield Curve”. Journal of Business and Economic Statistics, v. 27(4), p. 492-503.

[149] Rudebusch, G. and Wu, T. (2008). ”A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy”. Economic Journal, v. 118, p. 906-926.

[150] Sharpe, W. F. (1994). “The Sharpe Ratio”. The Journal of Portfolio Management, v. 21, p. 49–58.

[151] Sims, C. A. and Zha, T. (2006). "Were There Regime Switches in U.S. Monetary Policy?" American Economic Review, v. 96(1), p. 54-81.

[152] Stambaugh, R. (1988). “The Information in Forward Rates: Implications for Models of the Term Structure”. Journal of Financial Economics, v. 21(1), p. 41-70.

[153] Stock, J. and Watson, M. (2002a). “Forecasting using principal components from a large number of predictors" Journal of the American Statistical Association, v. 97, p. 1167-1179.

[154] Stock, J. and Watson, M. (2002b). “Macroeconomic Forecasting Using Diffusion Indexes”. Journal of Business and Economic Statistics, v. 20, p. 147-162.

[155] Stock, J. H. and Watson, M. W. (2003). “Forecasting Output and Inflation: The Role of Asset Prices.” Journal of Economic Literature, v. 41(3), p. 788-829.

[156] Stock, J. and Watson, M. (2005). “Forecasting with many predictors”. Handbook of Economic Forecasting, v. 1, p. 515-550.

[157] Svensson, L. E. O. (1994). “Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994”. NBER Working Paper No. 4871.

[158] Schwarz, G. (1978). “Estimating the Dimension of a Model”. Annals of Statistics, v. 6, p. 461-464.

[159] Thornton, D. L. and Valente, G. (2012). “Out-of-Sample Predictions of Bond Excess Returns and Forward Rates: an Asset Allocation Perspective”. Review of Financial Studies, v. 25(10), p. 3141-3168.

[160] Tsai, J. and Wachter, J. (2013). “Rare booms and disasters in a multi- sector endowment economy”. Working paper, University of Pennsylvania.

[161] Uribe M. and Yue, V. (2006). Country Spreads and emerging countries: Who drives who?" Journal of International Economics, v. 69, p. 6-36.

[162] Vasicek, O. A. (1977). “An Equilibrium Characterization of the Term Structure”. Journal of Financial Economics, v. 5, p. 177–188.

[163] Vasicek, O. A, and Fong, H. G. (1982). “Term structure modeling using exponential splines”. Journal of Finance, v. 37, p. 339–348.

[164] Vayanos, D. and Weill, P. O. (2008). “A Search-based Theory of the On-the-run Phenomenon”. Journal of Finance, v. 63, p. 1361-1398.

[165] Venetis, I. A., Paya, I. and Peel, D. A. (2003). “Re-Examination of the Predictability of Economic Activity Using the Yield Spread: A Nonlinear Approach.” International Review of Economics and Finance, v. 12(2), p. 187-207.

[166] Wachter, J. (2006). “A Consumption Based Model of the Term Structure of Interest Rates”. Journal of Financial Economics, v. 79, p. 365–99.

[167] Wang, H., Li, D. and He, X. (2012). “Estimation of high conditional quantiles for heavy-tailed distributions”. Journal of the American Statistical Association, v. 107, p. 1453-1464.

[168] Welsh, I. and Goyal, A. (2008). “A comprehensive look at the empirical performance of equity premium prediction". Review of Financial Studies, v. 21, p. 1455-1508.

[169] Wheelock, D. and Wohar, M. (2009). “Can the Term Spread Predict Output Growth and Recessions? A Survey of the Literature”. Federal Reserve Bank of St. Louis Review, v. 91(5), p. 419-440.

[170] Wright, J. H. (2006). “The Yield Curve and Predicting Recessions.” Finance and Economics Discussion Series No. 2006-07, Federal Reserve Board of Governors, www.federalreserve.gov/pubs/feds/2006/200607/200607pap.pdf.

[171] Wright, J. H. (2011). “Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset”. American Economic Review, v. 101, p. 1514-1534.