Risks in macroeconomic fundamentals and excess bond returns predictability*

Rafael B. De Rezende†

This Version: October 20, 2015‡

Abstract

I extract factors from quantile-based macroeconomic risk measures and document that macroeconomic risks such as expectations, uncertainty, downside and upside risks and tail risks contain valuable information about bond risk premia. Macro risk factors account for up to 31% of the variation in excess bond returns, generate countercyclical bond risk premia and are largely unrelated to the Cochrane-Piazzesi and Ludvigson-Ng factors. The high predictive power is confirmed statistically and economically in an out-of-sample setting and hold when factors are constructed using macroeconomic data available in real-time. All together, these findings suggest that risks in macroeconomic fundamentals are an important source of fluctuations in the US government bond market.

Keywords: expectations hypothesis; term structure of interest rates; ex ante macroeconomic risks; bond risk premia; macro risk factors.

JEL Classifications: G12, G11, E43, E44

*I am grateful to Magnus Dahlquist, Lars E.O. Svensson, Michael Halling, Refet Gürkaynak, Roméo Tédongap, Erik Hjalmarsson, Ferre De Graeve, Ádám Faragó as well as seminar participants at the Stockholm School of Economics, Bank of England, Sveriges Riksbank, Getúlio Vargas Foundation São Paulo, the World Congress of the Econometric Society 2015 and the XXI Finance Forum for comments and suggestions. I kindly thank the Swedish Bank Research Foundation (BFI) for financial support.

†Sveriges Riksbank (Central Bank of Sweden), Brunkebergstorg 11, 103 37 Stockholm, Sweden. Email: rafael.rezende@riksbank.se. Phone number: +46 767256582.

‡First Draft: April 30, 2013

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. 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 that are not embedded in financial variables. These findings imply that risk premia are time-varying and account for a significant portion of fluctuations in the US government bond market.

This paper addresses two questions. First, can movements in bond risk premia be empirically explained by macroeconomic risks such as risks of extreme macroeconomic outcomes, macroeconomic expectations, downside and upside risks, and macroeconomic uncertainty? Second, if so, do such risks contain any information about risk premia that is not already embedded in current financial and macroeconomic data?

The first question is central to test empirically the assertions of theoretical asset-pricing models that take macroeconomic risks into account. Such models suggest that investors care about the temporal distribution of risk and imply that time-variation in risk premia are driven by time-varying volatility and skewness in expected inflation and expected real growth (Bansal and Yaron, 2004; Bansal and Shaliastovich, 2013; Colacito et al. 2015). For example, using a long-run model framework Colacito et al. (2015) theorize that investors like high expected utility levels and positive asymmetry about future consumption growth rates, but also dislike uncertainty and negative asymmetry. The second question is important for understanding whether such risks provide additional information about variation in bond risk premia compared to financial and macroeconomic indicators. Recent papers have found factors that do not lie in the span of the term structure of interest rates, but that are still important for explaining bond risk premia (Kim, 2008; Ludvigson and Ng, 2009; Duffee, 2011; Joslin, Priebsch and Singleton, 2014). As macroeconomic variables, macroeconomic risks may be unspanned factors. Therefore, uncovering new empirical information about variation in bond risk premia is of great interest.

Neverthless, despite the growing body of theoretical work in this area, there is still little empirical evidence of a direct link between risks underlying macroeconomic variables and risk premia in government bond markets. Currently, the literature has uncovered information about bond risk premia variation contained in measures of macroeconomic expectations and macroeconomic uncertainty, but the amount of information is still not significantly strong, and the information content is not necessarily different from that provided by financial and current macroeconomic indicators.

There are several possible reasons why it may be difficult to find a strong link between macroeconomic risks and bond risk premia. First, the empirical literature has primarily focused on only a few risk measures such as macroeconomic expectations and uncertainty (Chun, 2011; Wright, 2011; Buraschi and Whelan, 2012; Bansal and Shaliastovich, 2013; Dick, Schmeling and Schrimpf, 2013).[1] Theory, however, suggests that skewness risks as well as tail risks account for a significant amount of fluctuations in asset market risk premia (Bollerslev and Todorov, 2011; Bollerslev, Todorov and Xu, 2014; Colacito et al., 2015), indicating the importance of taking this information into consideration. Second, existing studies have measured macroeconomic risks for only one or two macro variables. However, it is common knowledge that financial market participants typically consider a number of macroeconomic indicators when making investment decisions, meaning that considering a small number of variables may be insufficient. Third, macroeconomic risks are latent variables and are difficult to measure. Most existing studies have proposed measures based on the cross-sectional distribution of analysts' forecasts, but surveys respondents are typically professional forecasters and the information contained in their expectations may not fully represent the information that is relevant to financial market participants. In addition, some respondents may provide strategic forecasts or omit relevant forecast information (Ottaviani and Sorensen, 2006), and surveys also commonly suffer from a small number of cross-sectional observations at certain dates.

This paper considers ways to circumvent these difficulties. First, it considers a complete set of macroeconomic risk measures. More specifically, three quantile based measures are used to model the first three moments of the conditional distributions of future macroeconomic outcomes. These measures assume appealing economic interpretations in terms of macroeconomic expectations, uncertainty and downside (upside) macroeconomic risks. Since the analysis is concentrated on the top and bottom 5% conditional quantiles, the measures also capture information on macroeconomic tail risks, providing a much richer description of the risks involving the future state of the economy. Second, it uses quantile regression methods to estimate macroeconomic risks. This reduces the reliance on surveys as macroeconomic risks can be estimated using information that is more likely to span the unobservable information sets of bond market participants. Moreover, as no parametric form is imposed on the conditional distribution of the error term the approach shows high flexibility, capturing various features of the data and allowing one to produce accurate density forecast, from which macroeconomic risks can be obtained (Galvão, 2011; Gaglianone and Lima, 2012; De Rezende and Ferreira, 2014). Lastly, the risk measures are estimated for several macro variables and are effectively summarized in a small number of factors using the methodology of dynamic factor analysis. All together, these allow for a much richer information base of risks in macroeconomic

[1] 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).

fundamentals than what has been possible in prior empirical studies.

Results indicate that excess bond returns can be indeed predicted by risks in macroeconomic fundamentals. The estimated factors, referred to as macro risk factors, predict future excess bond returns across maturities with R^2s ranging from 20% to 31%. Importantly, macro risk factors can also be interpreted economically. Point expectations of real economic activity, uncertainty about real GDP growth, and downside and upside risks in housing starts and the unemployment rate are shown to be important determinants of bond risk premia in the US. Moreover, macro risk factors capture predictability in excess bond returns that is largely unspanned by the yield curve (Duffee, 2011; Joslin, Priebsch and Singleton, 2014), as macro risk factors are found to contain predictive information beyond the yield curve, while a large part of their variation remains unexplained by current yields.

Following Cochrane and Piazzesi (2005, CP hereafter), I also form a single macro risk factor. The new single factor explains variation in excess bond returns with R^2s of up to 31% and shows higher predictive power than an expectations factor. This shows the importance of accounting for other types of macroeconomic risks. In addition, results suggest that the new factor is superior to the CP and Ludvigson and Ng (2009, LN hereafter) factors. Combining them together results in levels of predictability around 45%, indicating that risks in macroeconomic fundamentals capture information about bond risk premia that is not embedded in forward rates or current macroeconomic indicators. Importantly, the new factor shows a pronounced countercyclical behavior, consistent with theoretical models asserting that investors must be compensated for macroeconomic risks associated with recessions (Campbell and Cochrane, 1999; Bansal and Yaron, 2004; Wachter, 2006; Rudebusch and Swanson, 2009). Much of this evidence can be explained by the countercyclical behavior of the macro risk measures I estimate.

These findings are also verified in an out-of-sample exercise. Results reveal that the single macro risk factor generates out-of-sample predictions that are more accurate than those produced by a constant model of no-predictability, with prediction errors being reduced by up to 29%. These results are superior than those achieved by the CP factor for all maturities and by the LN factor for intermediate to longer maturities. Also, adding the new factor to CP and LN regressions substantially increases predictive power. Prediction errors are reduced by 11% to 32%, providing even stronger evidence that risks in macroeconomic fundamentals contain additional information about variation in bond risk premia when compared to current financial and macroeconomic indicators. These results are confirmed economically in a classical portfolio choice problem. A portfolio of bonds constructed from the single macro risk factor delivers utility gains and positive risk-adjusted measures of portfolio performance when compared to a constant model. The only predictor that provides comparable results is the LN factor. Results also hold when factors are constructed using macroeconomic data available in real-time, indicating that the predictability of excess bond returns is not necessarily driven by data revisions, as suggested by Ghysels, Horan and Möench (2014).

The findings presented in this paper have important implications for both finance and macroeconomics. By tying time-variation in bond risk premia to risks in macroeconomic fundamentals, this paper provides an empirical ground for structural asset-pricing models that rationalize asset market risk premia. The findings also demonstrate the importance of accounting for information about risks in macroeconomic fundamentals to obtain a better identification of the term premium component of yields. This helps to clarify the relationship between short and long interest rates, facilitating 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.

The rest of the paper is organized as follows. The next section briefly reviews the related literature not discussed above. The third section introduces the measures of macroeconomic risks used in the paper and discusses their estimation. The fourth section presents the econometric framework proposed for predicting excess bond returns. The fifth section discusses the main results of the paper. The last section concludes.

2 Related literature

This work is related to research that looks at connections between bond yields and macroeconomic risks. Chun (2011) incorporates analysts' forecasts as factors in an affine term structure model for the US and finds that survey expectations about inflation, output growth and future policy rates are able to explain movements in bond yields. Moreover, expectations about GDP growth are found to account for a large amount of variation in risk premia. Wright (2011) finds that declining term premia have been the major source of the downtrends in government bond yields and forward rates observed globally in the last decades. Finally, he attibutes this trend to declining inflation uncertainty. This work considers a set of macroeconomic risk measures that goes beyond expectations and uncertainty, and estimates them for a number of macroeconomic indicators using quantile regressions. This allows for a much richer information base of risks in macroeconomic fundamentals than what has been possible in prior empirical studies, permiting us to reach a number of novel results.

Researchers and policy makers have recognized the importance of going beyond traditional point forecasts and have recently looked at density forecasts from which estimates of macroeconomic risks can be obtained. A recent strand of research aims at measuring such risks. For instance, 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. 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 these to estimate risks of high unemployment rates. Christensen, Lopez and Rudebusch (2011) rely on Treasury Inflation Protected Securities (TIPS) to measure deflation probabilities. In this paper, I estimate a pool of macroeconomic risk measures that goes from simple median forecasts to measures of uncertainty, skewness and tail risks.

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 skewness that are 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 risk measures using quantile regression methods (Koenker and Basset, 1978) and discuss how this approach allows extending the notion of macroeconomic risks to any variable of interest.

3 Measures of macroeconomic risks

3.1 Median, interquantile range and interquantile skewness

I start by providing three simple risk measures that share the distinguishing feature of being able to capture time variation in conditional distributions of any h-period ahead macro variable, $z_{t,t+h}$. My first object of interest is the median. Let $z_{t,t+h}$ denote the annual log rate of change in 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,

$$F_{z_{t,t+h}}(x) = Pr\left(z_{t,t+h} \leq x|\Omega_t\right) \tag{1}$$

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,

$$Med^h_t = q_{z_{t,t+h}}(0.5) \tag{2}$$

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.[2]

The second measure is the interquantile range of the conditional distribution of $z_{t,t+h}$. As

[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.

the simplest robust measure of data dispersion, the interquantile range provides a natural way of gauging how spread out is the conditional distribution of $z_{t,t+h}$. 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

$$IQR_t^h(\tau) = q_{z_{t,t+h}}(1-\tau) - q_{z_{t,t+h}}(\tau) \tag{3}$$

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,

$$IQS_t^h(\tau) = \frac{\left(q_{z_{t,t+h}}(1-\tau) - q_{z_{t,t+h}}(0.5)\right) - \left(q_{z_{t,t+h}}(0.5) - q_{z_{t,t+h}}(\tau)\right)}{q_{z_{t,t+h}}(1-\tau) - q_{z_{t,t+h}}(\tau)} \tag{4}$$

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. An additional 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 White, Kim, and Manganelli (2008), Ghysels, Plazzi, and Valkanov (2010), Andrade, Ghysels and Idier (2012) and Conrad, Dittmar and Ghysels (2013).

3.2 Estimation and economic interpretation

The risk measures defined above can easily be estimated using linear regression techniques. One simple and tractable approach is Koenker and Basset (1978)'s quantile regression method, which is suitable for approximating conditional quantiles of the response variable through estimated quantile functions. I consider here that $q_{z_{t,t+h}}(\tau)$ can be approximated by a model of the form,

$$q_{z_{t,t+h}}(\tau) = \beta(\tau)' x_t \tag{5}$$

where x_t is a $k \times 1$ vector of covariates and $\beta(\tau)$ is a $k \times 1$ vector of parameters to be estimated according to Koenker and Basset (1978) (see Appendix A for details).

Variables entering the vector x_t were chosen in 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 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 $z_{t,t+h}$

$$x_t^{'} = \left(1,\, z_t^{SPF,h},\, Mich\,Expect_t,\, 5 - year\,term\,spread_t,\, Baa\,corp\,spread_t\right) \tag{6}$$

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 risks have relied exclusively on surveys to obtain estimates of macroeconomic risks (Chun, 2011; Wright, 2011; Buraschi and Whelan, 2012; Dick, Schmeling and Schrimpf, 2013).[3] A limitation of this strategy, however, is that the typically sampled surveys' respondents are professional forecasters meaning that the information contained in their expectations may not necessarily fully represent the information that is relevant to financial market participants. Moreover, some analysts may potentially provide strategic forecasts or omit relevant forecasting information (Ottaviani and Sorensen, 2006), while surveys also commonly suffer from a small number of cross sectional observations at certain dates which may bias risk measures estimates.

The main advantage of using quantile regression is the possibility of estimating these variables using information that is more likely to span the unobservable information set 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 about market participants' perceptions of the likelihood of business bankruptcy and default (Friedman and Kuttner, 1998), as well as about future Federal Reserve's reactions to inflation and economic activity.

Another advantage of model (5) is its great flexibility. The appeal 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 conditional distributions of $z_{t,t+h}$, which may be Gaussian, but can also assume non-standard forms. Figure 1 illustrates this with several quantile lines estimated for

[3] These studies, however, focus only on measures of macroeconomic expectations and uncertainty. These measures are proxied by the average (or median) of forecasts, also known as consensus forecast, and by the dispersion of the cross-sectional distribution of forecasts at each date.

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 interesting shapes and to capture several interesting features of the data as, for instance, the increasing levels in dispersion, skewness and tail movements around recessions. Notice also that while the median is able to match realized values at many dates, it misses important periods of macroeconomic stress. The tails of $z_{t,t+h}$, on the other hand, seem to capture extreme movements in macro variables with higher accuracy. This result is evident during the 2008/2009 recession.

This means that Med, IQR and IQS can then be interpreted as measures of macroeconomic risks. Med, as a way of characterizing typical values assumed by $z_{t,t+h}$, can serve as a measure of macroeconomic point expectations. IQR can be viewed as a measure of uncertainty about $z_{t,t+h}$ at time t, while IQS can be interpreted as a measure of downside (upside) macroeconomic risks, as negative values for IQS, for instance, indicate that there is a higher probability 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 also share the attractive property of 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 capture information about macroeconomic tail risks, allowing for a rich characterization of risks involving the future state of the economy. Risks of extreme macroeconomic outcomes such as large drops in economic activity, high inflationary pressures or even a boom in the housing market, may have important implications for risk premia in equity and bond markets (Bollerslev and Todorov, 2011; Gabaix, 2012; Tsai and Wachter, 2013; Bollerslev, Todorov and Xu, 2014).

Macro variables were selected according to their availability in the SPF data set since when the survey was initiated. This means that the risk measures are estimated for inflation measured by the GDP price index and growth in real GDP, unemployment rate, industrial production, housing starts and corporate profits after tax (see appendix D for more details about the data). The estimation of the risk measures for a larger set of macroeconomic indicators eliminates the reliance on a small number of imperfectly measured proxies for macroeconomic risks and allows me to exploit a much richer information base of risks in macroeconomic fundamentals than what has been possible in prior empirical studies in this literature.

The sample ranges from 1968:Q4 to 2011:Q4. Since I will be predicting excess bond returns accumulated over the following year starting from t, h is then 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.

3.3 Ex ante macroeconomic risks in the US: stylized facts

Figures 2 and 3 show the eighteen estimated measures of macro risks observed at time t together with $q_{z_{t,t+4}}(.05)$ and $q_{z_{t,t+4}}(.95)$. NBER-dated recessions are shown as shaded bars. Notice that the risk measures' estimated time series reveal several interesting features. First, the interquantile range of the conditional distributions of growth in real GDP, unemployment rate, industrial production and housing starts, show pronounced countercyclical behavior, indicating the presence of increasing levels of uncertainties regarding future developments in these variables during bad times. Using a different approach, this result is also documented by Jurado, Ludvigson and Ng (2015) 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 real GDP, industrial production and housing starts along with extreme rises in unemployment rate and inflation show more pronounced behavior and increase substantially during recessions.

It is also worth commenting on the behavior of uncertainty for housing starts during the recession of 2008/2009. While we see sharp increases in this variable during all previous NBER-dated recessions, when it comes to the the recession of 2008/2009, 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 - 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 real 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 4 shows the correlations between estimated risk measures and real GDP growth, both measured at time t. Blue circles indicate statistically significant correlation coefficients. Observe that most ex ante risks show strong relationships to real GDP growth. Tail risks as well as median predictions regarding real GDP and industrial production are positively related to real 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 real GDP, strengthening my previous findings that macroeconomic uncertainty is countercyclical. Also, observe that real GDP growth relates positively to ex ante downside risks for inflation, real 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.

4 Predicting excess bond returns

I focus on 1-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 of 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.[4] 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.[5] For $t = 1, \ldots, T$, 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 an n-year bond purchased at time t and sold one year after at time $t+1$ (or $t+4$ quarters) and y^n_t is the log yield on the n-year bond.

Table 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,

$$rx^n_{t,t+4} = \alpha_0 + \alpha' m_t + \vartheta' g_t + \varepsilon_{t,t+4} \tag{7}$$

where α and ϑ are 18×1 vectors of coefficients, m_t is a 18×1 vector of estimated macro risks

[4] 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.

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

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 (7) allows for the use of all the information about macroeconomic risks available, 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 (7) 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 $rx^n_{t,t+4}$ 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 $rx^n_{t,t+4}$ can be optimally selected using Schwarz (1978) Bayesian information criteria (SBIC).[6] This leads to the following regression,

$$rx^n_{t,t+4} = \alpha_0 + \alpha' MRF_t + \vartheta' g_t + \varepsilon_{t,t+4} \tag{8}$$

where MRF_t is a vector of estimated macro risk factors and α_0 and α are parameters to be estimated by OLS[7]. 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.

5 Empirical results

5.1 In-sample evidence

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

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

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

$MRF_t = (MRF_{1t}, MRF_{4t}, MRF_{6t})^{'}$. 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

$$\begin{gathered} \overline{rx}_{t,t+4} = \theta_0 + \theta_1 MRF_{1t} + \theta_2 MRF_{4t} + \theta_3 MRF_{6t} + \varepsilon_{t,t+4} \\ SMRF_t = \widehat{\theta}' MRF_t \end{gathered} \tag{9}$$

Table 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 shows the highest statistical significance followed by MRF_1. MRF_6 is not significant, although it seems important for predicting $rx^n_{t,t+4}$, according to SBIC.[8] 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 (9). After running (9) 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 fourth 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 5 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 real GDP, that is, variables related to economic activity. The fourth factor is highly related to the

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

downside and upside risks to housing starts, although it also manifests a strong relationship with GDP-IQR and Unemp-IQS. The sixth factor is clearly significantly related to risks associated with inflation, with Inf-IQS explaining a large portion of its variation. From Panel B, notice also that while the first factor seems to be mostly related to expectations, the fourth and sixth factors are strongly related to downside and upside risks.

Figure 6 shows 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, MRF_1 is plotted against Unemp-Med, while MRF_4 and MRF_6 are plotted against Hous-IQS and Inf-IQS, respectively. Shaded bars indicate NBER-dated recessions. Figure 6 shows 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 is strongly related to risks in economic activity. MRF_4 is clearly negatively correlated with Hous-IQS 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. The correlation between the two series is 55%. These results lead us to classify MRF_4 (MRF_6) as a housing (inflation) skewness factor, though MRF_4 may be also interpreted as a GDP uncertainty or unemployment skewness factor.

Beyond the median

I have provided evidence that risks in macroeconomic fundamentals derived from Med, IQR and IQS estimated for various variables 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 in explaining movements in bond premia.

Rather than focusing on survey consensus forecasts, I extract median forecasts by estimating median regressions as (5) for the six macro variables.[9] Equation (5) 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)'s information criteria indicates that this panel is well described by three principal components, which were finally all chosen using the SBIC criteria, as previously done. More specifically, the single median factor was obtained as,

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

$$
\begin{gathered}
\overline{rx}_{t,t+4} = \kappa_0 + \kappa_1 MeF_{1t} + \kappa_2 MeF_{2t} + \kappa_3 MeF_{3t} + \varepsilon_{t,t+4} \\
SMeF_t = \widehat{\kappa}' MeF_t
\end{gathered} \tag{10}
$$

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

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,

$$
\begin{gathered}
\overline{rx}_{t,t+4} = \delta_0 + \delta_1 y^1_t + \ldots + \delta_{10} fw^{10}_t + \varepsilon_{t,t+4} \\
CP_t = \widehat{\delta}' fw_t
\end{gathered} \tag{11}
$$

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 as a linear combination of macro factors extracted from a large macroeconomic data set (131 variables). When forming LN, I use the data set provided by Ludvigson and Ng (2010) but set October 1968 as the starting date in order to enable direct comparisons with the other predictors studied in the paper.[10] The data are set at quarterly frequency 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,

$$
\begin{gathered}
\overline{rx}_{t,t+4} = \varphi_0 + \varphi_1 F_{1t} + \varphi_2 F_{2t} + \varphi_3 F_{6t} + \varepsilon_{t,t+4} \\
LN_t = \widehat{\varphi}' F_t
\end{gathered} \tag{12}
$$

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

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.[11]

Results are shown in Table 4. As documented by Cochrane and Piazzesi (2005, 2008), I find that CP captures a large portion of variation in expected excess returns 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 an 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 the three single factors jointly. As documented by Ludvigson and Ng (2009), including LN to CP regressions increases 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 according to bootstrap standard errors. In addition, notice that LN estimates lose significance from the 3-year maturity, according to bootstrap standard errors.

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 and upside risks as well as 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 the information 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 to higher risk premia during bad times.

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

In light of this, we can gain some economic intuition about bond premia implied by risks in macroeconomic fundamentals by examining how they behave over business cycles. More specifically, I show that movements in the single macro risk factor, a measure of average bond risk premia across maturities, is closely connected to NBER-dated business-cycle phases. Figure 7 - Panel A shows the 4-quarter moving average of SMRF. In general, we see declines in bond premium during expansions and sharp increases during recessions. Notice also that the increases in risk premium observed during the 1990-1991 and 2001 recessions are somewhat more modest than those observed during the recessions of the 1980's and the late 2000's. This makes sense because these two recessions were milder relative to the others. Overall, Figure 7 - Panel A suggests that macroeconomic risks produce bond risk premia that closely track NBER-dated business-cycle phases.

Panel B complements the evidence shown in Figure 7 - Panel A and shows lead/lag relations between bond premium and growth rates for three macroeconomic variables closely related to business cycles: real GDP, industrial production and unemployment rates. The bond premium indicator is kept fixed at date t and the economic indicators are then led and lagged. Notice that correlations turn negative/positive as macro variables are led/lagged. While a drop in economic activity leads to 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 the 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 economy are unspanned by the term structure of interest rates 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 (Kim, 2008; Ludvigson and Ng, 2009; Duffee, 2011; Joslin, Priebsch and Singleton, 2014). As demonstrated above, macro risk factors are able to predict excess bond returns. In this section, I then explore whether macro risk factos are alo unspanned factors.

It is customary in the term-structure literature to summarize the information in yields using its three first principal components (PC hereafter) as they explain virtually all 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 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 the bond risk premia that is in some degree independent of that contained in the yield curve. Table 5 shows R^2s for the regressions of PC and/or MRF onto excess returns. While PC shows predictive power with R^2s that range 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 (2014) also suggest examining the following spanning condition

$$MRF_{jt} = \omega_0 + \omega' PC_t, \; j = 1,4,6 \tag{13}$$

which projects 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, respectively. 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 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 identified by the cross-section of yields or forward rates alone. 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 beyond 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, 2014; Ludvigson and Ng, 2009) and in conditional distributions of future macroeconomic outcomes should also be taken into account.

Online Appendix C provides the results for the estimation of an affine term structure model along the lines of Joslin, Priebsch and Singleton (2014) and further discusses 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.[12] 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. Online Appendix C also provides estimates of the 10-year yield term premium obtained from the affine model with macro risk factors. Consistent with previous findings on the SMRF (return risk premia), term premia implied by risks in macro fundamentals are highly countercyclical, which is consistent with existing theory.

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

Robustness Tests

A natural question that may arise, however, is whether specification (6) is really capturing the true quantiles of $z_{t,t+4}$. In order to verify this, I apply the backtest proposed by Gaglianone et al (2011) (GLLS hereafter) which evaluates 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 for the identification of the extent to which 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 higher power in finite samples compared to the most common existing backtests. 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 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. 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 6 show that specification (7) produces conditional quantiles forecasts, $\widehat{q_{z_{t,t+4}}}(\tau)$, that are statistically indistinguishable from the true conditional quantiles of $z_{t,t+4}$. This suggests that the risk measures Med, IQR and IQS are being precisely estimated and also that (7) is able to accurately capture the conditional distributions of $z_{t,t+4}$.

The most natural question, however, is whether the high degree of predictability in excess bond returns is coming exclusively from variables in the vector x_t. If this is the case, 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 verifying the predictive power of SMRF when controlling 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 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) criteria 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, while SFx is a linear combination of Fx. SFx not only provides a variable that can be used to control for the information in quantile predictors x_t, but also guards against the possibility of overfitting in out-of-sample forecasting. The in-sample and out-sample results of this exercise are provided in the Online Appendix A. In general, in-sample results show, that although SFx contains high predictive power for $rx^n_{t,t+4}$, adding SMRF to regressions increases R^2s substantially to levels almost identical to the ones shown by Table 2, with statistical significance shifting to SMRF. Out-of-sample results are even more informative. SFx regressions generate very inaccurate forecasts and results improve dramatically when SMRF is included as an additional predictor. These results indicate that the high degree of predictability found is largerly due to the extra information obtained from the estimation of Med, IQR and IQS.

The assessment of the predictive power of macro risk factors (MRF and SMRF) when risk measures are estimated using alternative approaches is another convenient robustness test . To guard against the possibility of inadequacy of quantile regressions estimated at $\tau = 0.05, 0.95$ the Online Appendix B presents 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 the 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 Online Appendix B. 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 2, 3 and 4, which corroborate my previous findings.

5.2 Out-of-Sample evidence

Statistical Predictability

With the exception of Ludvigson and Ng (2009, 2010) and Thornton and Valente (2012), most papers in this literature only provide predictive results in an in-sample framework. As models in-sample predictive performance tends to be poorly related to their ability to generate meaningful out-of-sample predictions (Inoue and Kilian, 2004, 2006), in this subsection, I examine the predictability of excess bond returns in an out-of-sample setting. For this exercise, all parameters are estimated recursively using only information available at the time forecasts are generated. In each recursion, factors used in the construction of SMRF and LN are also reestimated and optimally selected (using SBIC), taking into consideration the possibility that different factors may be chosen in different samples.

SMRF, CP, LN and their respective combinations are the predictors evaluated. 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. 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 the data. In the second round of comparisons, I test whether adding SMRF to CP, LN and CP+LN regressions indeed 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. Forecasts were generated for the period 1990Q1 - 2011Q4. When the LN factor is included in the set of predictors, the out-of-sample portion of the 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.[13]

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), 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.

Results are shown in Table 7 - Panel A. As observed, beating the constant model is not an easy task. All predictors, except LN, fail when forecasting rx^n for shorter maturities. However, results change as we move our attention to longer maturities. While CP continues to perform poorly,

---

[13]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 $j = SMRF, CP, LN, SMRF+CP, SMRF+LN, SMRF+CP+LN$ 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.

SMRF shows impressive results with R^2_{oos}s turning positive from rx^5 (rx^3) in the sample period of 1990-2011 (1990-2007) and reaching 0.292 (0.278) for the 10-year bond return. Notice that SMRF outperforms the constant model with high statistical significance according to both CW and MC tests when predicting the 3, 5, 7 and 10-year returns. In addition, combining SMRF with other predictors always improves regressions' predictive power. Models SMRF+CP+LN and SMRF+LN, for instance, show impressive results. The two specifications 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.

In the second round of comparisons, I test whether improvements obtained from adding SMRF to CP and LN regressions are statistically significant. Results of this exercise are shown in Table 7 - Panel B. Notice that augmenting regressions with SMRF remarkably improves predictability 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, adding SMRF generates positive R^2_{oos}s from the 5-year maturity and results are highly statistically significant.

In order to check the stability of results over time, Figure 8 - 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 against the constant model over the full period of evaluation. Some exceptions are found. For instance, models which include the CP factor among predictors show decreasing levels of predictive power. The LN model does not performe so well until the early 2000's, when R^2_{oos}s start increasing substantially. On the other hand, SMRF regressions as well as models that include SMRF as an additional predictor show R^2_{oos}s that are quite high over the full period of evaluation. In general, the most successful model is SMRF+CP+LN.

These results provide even stronger evidence that risks in macroeconomic fundamentals are able to explain variation in bond risk premia. The LN and SMRF factors show high degrees of predictive power and predictability is especially strong when combining the three factors together, confirming my previous finding that the three predictors capture somewhat independent information about bond risk premia variation.

Economic Predictability

Statistical predictability does not mechanically imply economic predictability because 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). Thus, in this subsection, I 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,

$$\max_{\mathbf{w_t}} \mathbf{w_t}^{'} \mu_{t,t+4} - \frac{\rho}{2} \mathbf{w_t}^{'} \sum_{t,t+4} \mathbf{w_t} \tag{14}$$

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 $1 - \mathbf{w_t}^{'} \iota$, 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).[14] The average utility of the investor is then given by

$$\overline{U}(\cdot) = \frac{1}{T-R} \sum_{t=R}^{T} \left[r_{t,t+4}^{p} - \frac{(\varsigma - 1)}{2} \mathbf{w_t}^{'} \sum_{t,t+4} \mathbf{w_t} \right] \tag{15}$$

where $r_{t,t+4}^{p} = y_t^1 + \mathbf{w_t}^{'} \mu_{t,t+4}$ 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 (14) 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 the model 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,

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

$$UG^j = \sum_{t=R}^{T}\left[r^p_{j,t,t+4} - \frac{(\varsigma-1)}{2}\mathbf{w}_{\mathbf{j},\mathbf{t}}^{'}\sum\nolimits_{j,t,t+4}\mathbf{w}_{\mathbf{j},\mathbf{t}}\right] - \sum_{t=R}^{T}\left[r^p_{c,t,t+4} - \frac{(\varsigma-1)}{2}\mathbf{w}_{\mathbf{c},\mathbf{t}}^{'}\sum\nolimits_{c,t,t+4}\mathbf{w}_{\mathbf{c},\mathbf{t}}\right] \tag{16}$$

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.

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

$$IR^j = \frac{1}{T-R}\sum_{t=R}^{T}\frac{\left(r^p_{j,t,t+4} - r^p_{c,t,t+4}\right)}{\sqrt{var\left(r^p_{j,t,t+4} - r^p_{c,t,t+4}\right)}} \tag{17}$$

However, while Sharpe ratios are commonly used, they have drawbacks. Abnormalities like excess kurtosis, outliers or skewness on the distribution of returns can be problematic for the statistic, as standard deviation computation is not as effetive when these problems are present. Also Sharpe ratios can be manipulated (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,

$$GISW^j = \frac{1}{(2-\varsigma)}\left\{log\left[\frac{1}{T-R}\sum_{t=R}^{T}\left(\frac{r^p_{j,t,t+4}}{1+y^1_{t+4}}\right)^{2-\varsigma}\right]\right\} - log\left[\frac{1}{T-R}\sum_{t=R}^{T}\left(\frac{r^p_{c,t,t+4}}{1+y^1_{t+4}}\right)^{2-\varsigma}\right] \tag{18}$$

Results of this exercise are shown in Table 8. Out-of-sample predictions were generated as before. As observed, only SMRF and LN or models in which one of these variables is present provide consistent economic value in terms of utility gains relative to the constant model. 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 are negative though, a result that is consistent with the findings of Thornton and Valente (2012).[15] Notice also that utility gains always increase when augmenting regressions with SMRF or LN. Although SMRF+CP+LN model performs quite well, SMRF+LN performs best with utility gains of 3.71 and 3.23, depending on the degree of RRA.

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

When we analyze the risk-adjusted measure of portfolio performance, GISW, results are consistent 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. On the other hand, augmenting the remaning models with SMRF always improves their respective portfolios' performances. Results with IRs are similar to the ones obtained with GISW.

The stability of these results over time is verified in Figure 8 - Panel B, which shows utility gains computed recursively. For the sample 1990Q1:2011Q4, results show that the economic predictability of models that include SMRF 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 SMRF+LN model is the one that generates the highest levels of recursive utility gains to the investor, as also suggested by Table 5.

To sum up, we observe that SMRF regressions are able to generate quite high utility gains and risk-adjusted measures of portfolio performance. Augmenting CP and LN regressions with SMRF is also found to always improve their respective portfolios' performances. In addition, the portfolio based on the SMRF+LN specification is the one that performs best. These results confirm the statistical evidence of out-of-sample bond return predictability.

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 necessarily the same that is 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.[16] 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, *advance* real-time macro data vintages are collected 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

[16] In a recent paper, Ghysels, Horan and Moench (2014) 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.

(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 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 (2012) 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 9. 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, particularly 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 negative 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. IRs are a 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 in the Online Appendix D. In general, results are quite similar to those shown in Figure 8. All together, these findings suggest that the predictability of bond returns is not necessarily driven by data revisions, as suggested by Ghysels, Horan and Moench (2014) when relying on the Ludvigson and Ng (2009) macro factors.

6 Conclusions

This paper contributes to the literature on the determinants of bond risk premia by showing that risks in macroeconomic fundamentals have important predictive power for US excess bond returns. To do so, I estimate a complete set of macroeconomic risk measures that are able to capture macroeconomic expectations, downside and upside macroeconomic risks, macroeconomic uncertainty and tail risks, using quantile regression methods. The approach reveals a number of advantages. First, it allows one to exploit a much richer information base on macroeconomic risks than previous studies, which focused on macroeconomic expectations and uncertainty only. Second, it eliminates the reliance on a small number of imperfectly measured proxies of macroeconomic risks based on survey forecasts. Third, risk measures are estimated for several macro variables and are effectively summarized in a small number of factors using the methodology of dynamic factor analysis. All together, these allow for a much richer information base on risks in macroeconomic fundamentals than what has been possible in prior empirical studies, permiting us to reach a number of novel results.

Two aspects of my findings are emphasized. First, in contrast to the existing empirical literature, I find strong predictable variation in excess bond returns that is associated with macroeconomic risks, beyond that of expectations and uncertainty. Second, specifications using forward rates and current macroeconomic variables omit important information about bond risk premia variation. The macro risk factors I estimate are shown to have substantial predictive power independent of that in the Cochrane-Piazzesi and the Ludvigson-Ng factors. When information contained in macroeconomic risks, current macroeconomic variables and the forward rates is combined together, I find large violations of the expectations hypothesis. In addition, macro risk factors capture predictability in excess bond returns that is largely unspanned by the yield curve .

These results are also verified statistically and economically in an out-of-sample setting. The new macro risk factors produce out-of-sample predictions with R^2s of up to 29% compared to the -10% and 12% delivered by the Cochrane-Piazzesi and Ludvigson-Ng factors, respectively. In addition, adding the new factor to CP and LN regressions reduces prediction errors by up to 32%, confirming my in-sample findings that risks in macroeconomic fundamentals capture information about bond risk premia that is largely unrelated to that contained in forward rates and current macroeconomic indicators. These high levels of predictability hold when factors are constructed using macroeconomic data available in real-time, indicating that the predictability of excess bond returns is not driven by data revisions, as suggested by Ghysels, Horan and Möench (2014).

To sum up, this study provides empirical support for asset-pricing models that rationalize asset market risk premia using macroeconomic risks. All together, my findings suggest that risks in macroeconomic fundamentals are an important source of fluctuations in the US government bond market.

Appendices

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 in (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 the "rearranged" curve more closely matches the true quantile curve in finite samples than the not rearranged one and reduces estimation errors by up to 14%.

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,

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

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 are 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 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 $\widehat{MRf}_t$.

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 (see Table A.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 first creating bootstrap samples for factors MRF1, MRF4, MRF6, SMRF, SMeF, LN and CP. 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 $rx^{n*}_{t,t+4} = \widehat{\delta}_0 + \widehat{\delta}^{\prime}G^*_t + \varepsilon^*_{t,t+4}$, where G^*_t is a set of bootstrapped regressors, $\varepsilon^*_{t,t+4}$ is sampled from $\varepsilon_{t,t+4} = rx^n_{t,t+4} - \widehat{\delta}_0 - \widehat{\delta}^{\prime}G^*_t$ using overlapping blocks of size equal to six and $\widehat{\delta}$ are the least squares estimates reported in Table 2, 3 and 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 a 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 $\left[\widehat{\delta}_j - \left|t_{0.975}^*\right| \times s\left(\widehat{\delta}_j\right); \widehat{\delta}_j + \left|t_{0.025}^*\right| \times s\left(\widehat{\delta}_j\right)\right]$.

Table A.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

D. Data

Table A.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 by merging Real GNP in the period 1968Q4-1991Q4 with Real GDP in the period 1992Q4-2011Q4. The real-time macro data provided by the Federal Reserve of Philadelphia are already merged, except $cprof$.

Table A.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
$MichExpect$	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
$y^6,\ldots,y^{10}$	Gurkaynak, Sack and Wright Bond Yields	1968Q4 to 2011Q4	100 bp	Federal Reserve Board

References

[1] 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.

[2] Andrade P., Ghysels E. and Idier J. (2014). "Inflation riks measures and their information content", mimeo, Banque de France.

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

[4] 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.

[5] 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.

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

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] Bollerslev, T. and Todorov, V. (2011). "Tails, Fears and Risk Premia". Journal of Finance, v. 66, p. 2165-2211.

[13] Bollerslev, T, Todorov, V. and Xu, L. (2015). "Tail Risk Premia and Return Predictability". Journal of Financial Economics, v. 118, p. 113-134.

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

[15] 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”.

[16] 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.

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

[18] 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.

[19] 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.

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

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

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

[23] 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.

[24] 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.

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

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

[27] Colacito, R., Ghysels, E., Meng, J. and Siwasarit, W. (2015). ”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”.

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

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

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

[31] 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.

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

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

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

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

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

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

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

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

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

[41] 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.

[42] 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.

[43] 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.

[44] Galvão, A. (2011). “Quantile regression for dynamic panel data with fixed effects”. Journal of Econometrics, v. 164, p. 142-157.

[45] 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.

[46] 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.

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

[48] 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.

[49] 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.

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

[51] 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.

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

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

[54] Gurkaynak, 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.

[55] Gurkaynak, 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.

[56] 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.

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

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

[59] 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.

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

[61] Jondeau, E. and Rockinger, M. (2008). “The Economic Value of Distributional Timing”. Mimeo, Swiss Finance Institute.

[62] 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.

[63] Joslin, S., Priebsch, M. and Singleton, K. (2014). “Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks”. Journal of Finance, v. 69, p. 1197-1233.

[64] Jurado, K., Ludvigson, S. and Ng, S. (2015). “Measuring Uncertainty”. American Economic Review, v. 105(3), p. 1177-1216.

[65] 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.

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

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

[68] Kitsul, Y. and Wright, J. H. (2012). “The economics of options-implied inflation probability density functions”. mimeo.

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

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

[71] 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.

[72] 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

[73] Leitch, G., and Tanner, J. E. (1991). “Economic Forecast Evaluation: Profits versus the Conventional Measures”. American Economic Review, v. 81, p. 580-590.

[74] 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.

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

[76] 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.

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

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

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

[80] 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.

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

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

[83] Rapach, D. E., Strauss, J. K. and Zhou, G. (2010). “Out-of-sample equity premium prediction: Combination forecasts and links to the real economy”. Review of Financial Studies, v. 23, p. 821-862.

[84] 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.

[85] 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.

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

[87] 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.

[88] 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.

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

[90] 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.

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

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

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

[94] 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 (forthcoming).

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

[96] 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.

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

[98] 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.

[99] 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.

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

Table 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 mean, standard deviation, skewness, excess kurtosis, p-value of a Robust Jarque-Bera (RJB) test for normality and the 1st and 4th sample autocorrelations are reported. Critical values for the RJB test were empirically obtained 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 2: In-sample predictability - 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 3: In-sample predictability - 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 in Table 2 (Notes) and Appendix 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 4: In-sample predictability - 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 2 (Notes) and Appendix C are reported in square brackets. Regressions in which LN is included in the set of predictors are estimated over the sample 1968Q4 - 2007Q4.

	SMRF	*CP*	$\overline{R}^2$	*SMRF*	*LN*	$\overline{R}^2$	*SMRF*	*CP*	*LN*	$\overline{R}^2$
		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]
rx^2	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]
		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]
rx^3	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]
		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]
rx^5	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]
		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]
rx^7	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]
		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]
rx^{10}	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 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 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 7: 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. (†), (††), († † †) 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

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

where $\widehat{rx}^{n,j}_{t,t+4}$ is a forecast generated from model $j = SMRF, CP, LN, SMRF+CP, SMRF+LN, SMRF+CP+LN$ 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 8: 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 9: Out-of-Sample predictability - real time macro data

Notes: SMRF is constructed using real-time macroeconomic data. For details regarding Panel A and Panel B, see Notes in Table 7. For details regarding Panel C, see Notes in Table 8. Results for the LN factor are not available due to the inexistence of a large macroeconomic data set in real-time.

	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	$0.035^{***\dagger\dagger\dagger}$	−0.315	-0.108^{**}	$0.040^{**\dagger\dagger\dagger}$	-0.100^{*}	$0.042^{***\dagger\dagger\dagger}$
rx^7	$0.128^{***\dagger\dagger\dagger}$	−0.343	-0.071^{**}	$0.126^{***\dagger\dagger\dagger}$	-0.123^{*}	$0.081^{***\dagger\dagger\dagger}$
rx^{10}	$0.198^{***\dagger\dagger\dagger}$	−0.353	-0.016^{***}	$0.199^{***\dagger\dagger\dagger}$	-0.134^{**}	$0.135^{***\dagger\dagger\dagger}$

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

	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 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 give the predicted median and blue lines indicate the realized values.

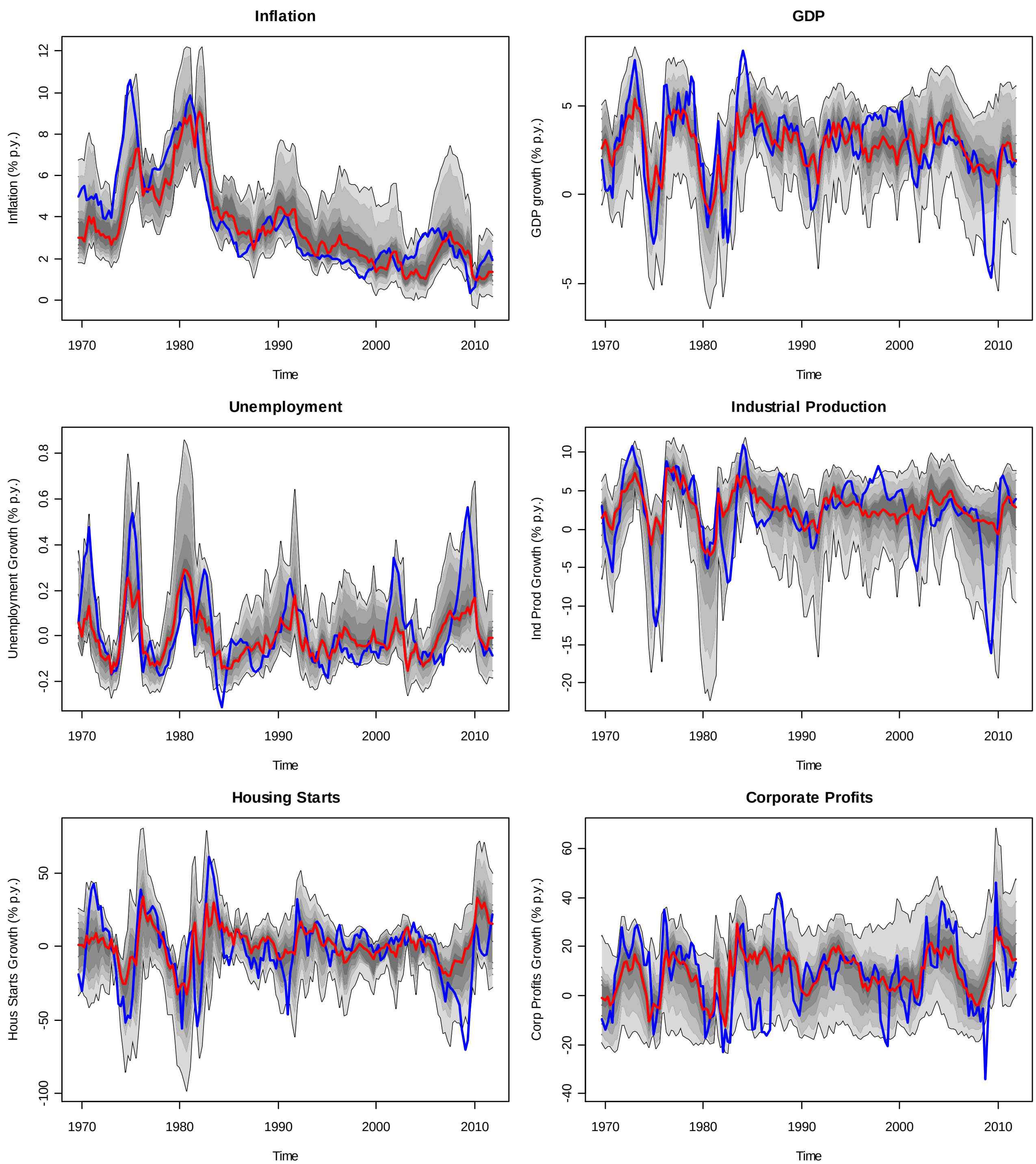

Figure 2: Ex ante macroeconomic risks 1

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).

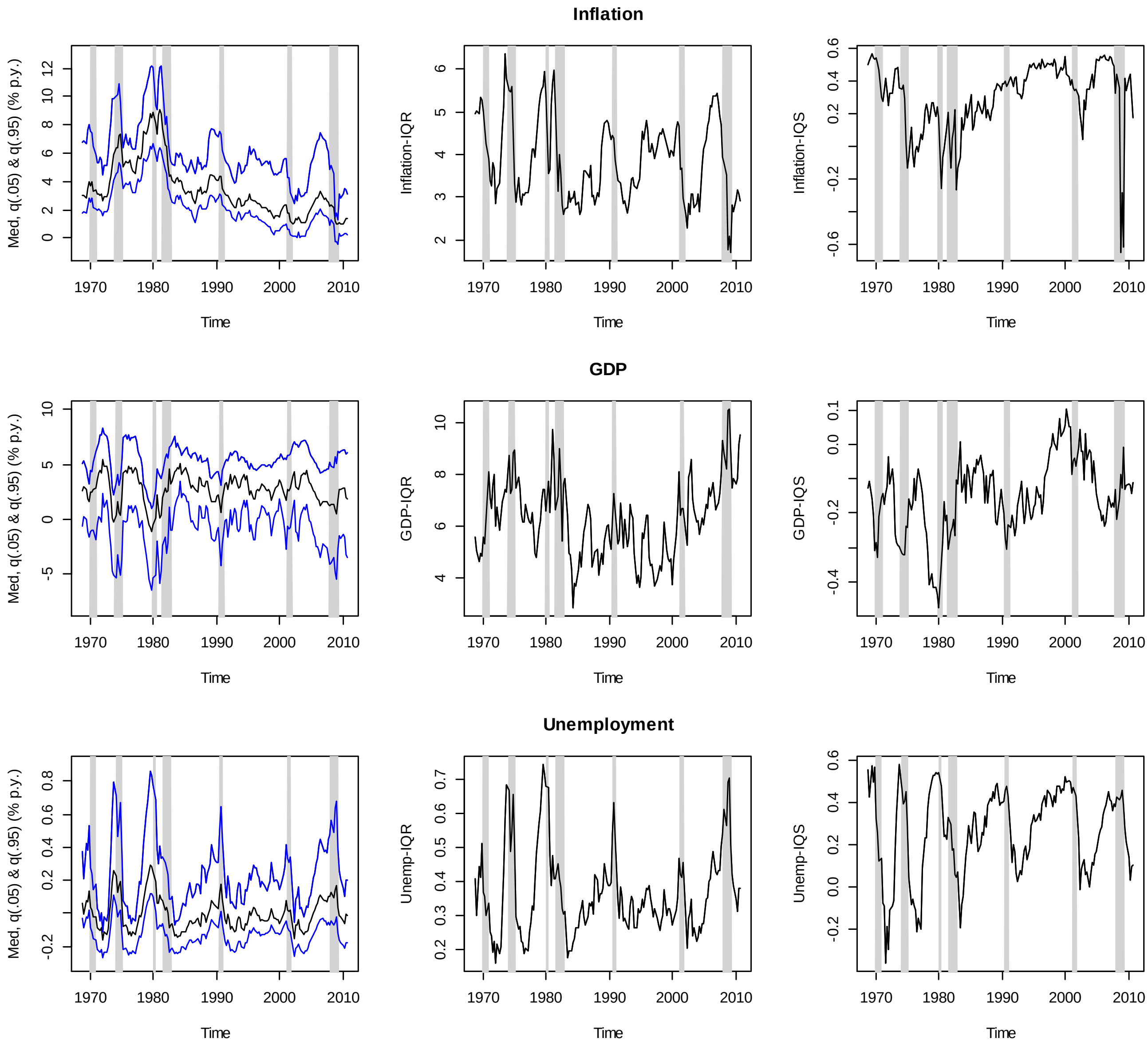

Figure 3: Ex ante macroeconomic risks 2

Notes: As in Figure 2.

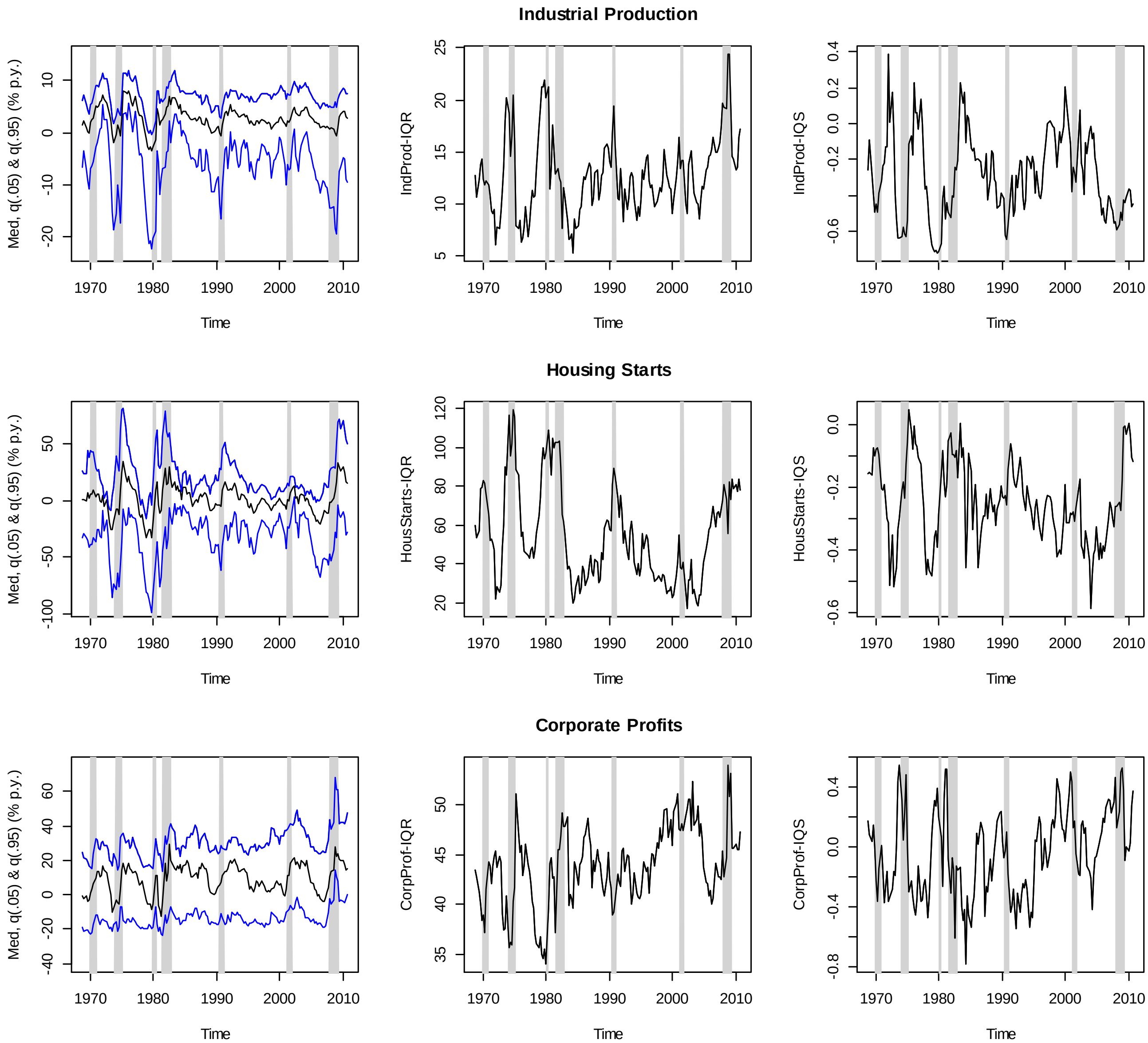

Figure 4: Correlations of ex ante macro risks and GDP growth

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

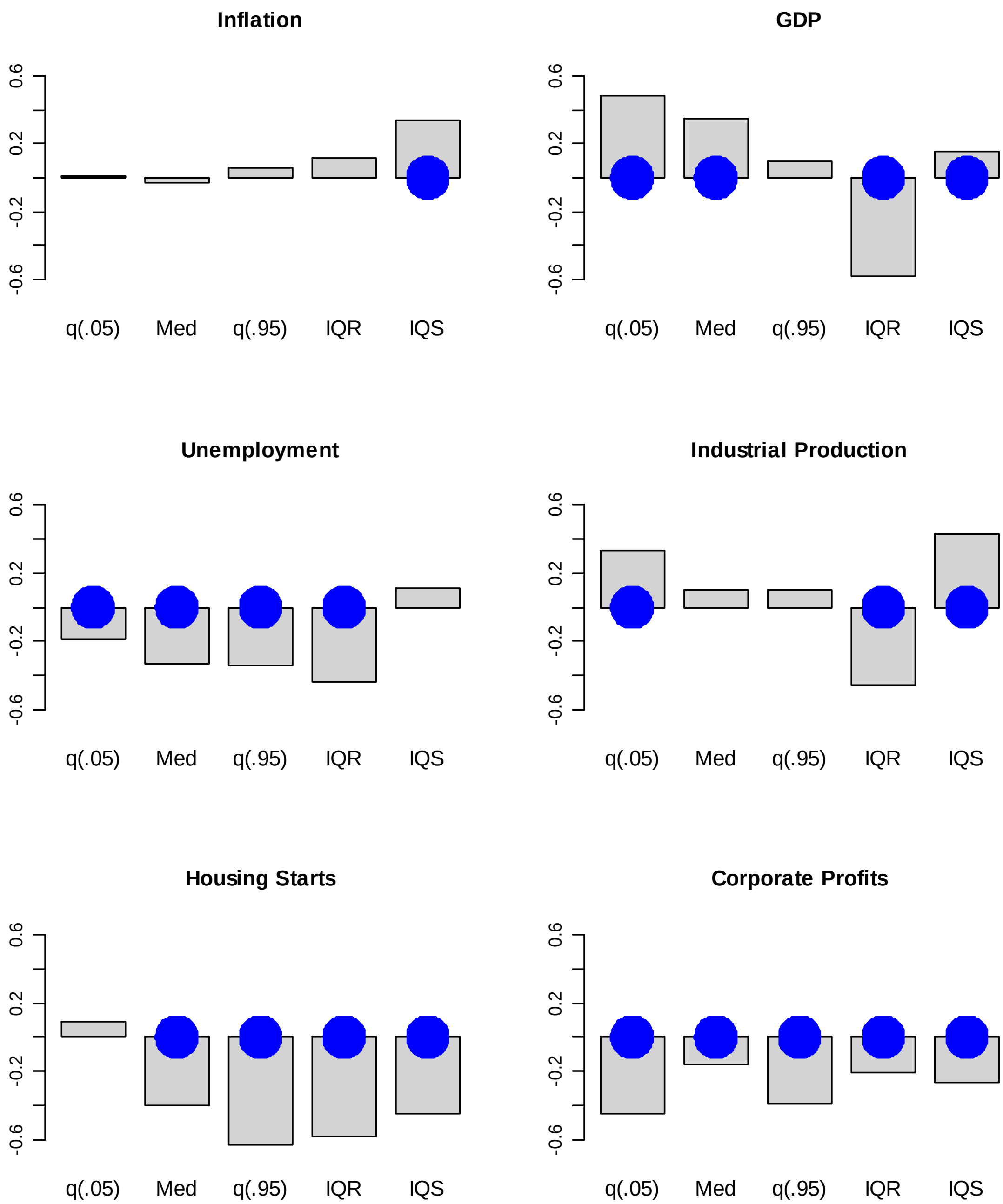

Figure 5: Economic Interpretation of macro risk factors - Marginal R^2s

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.

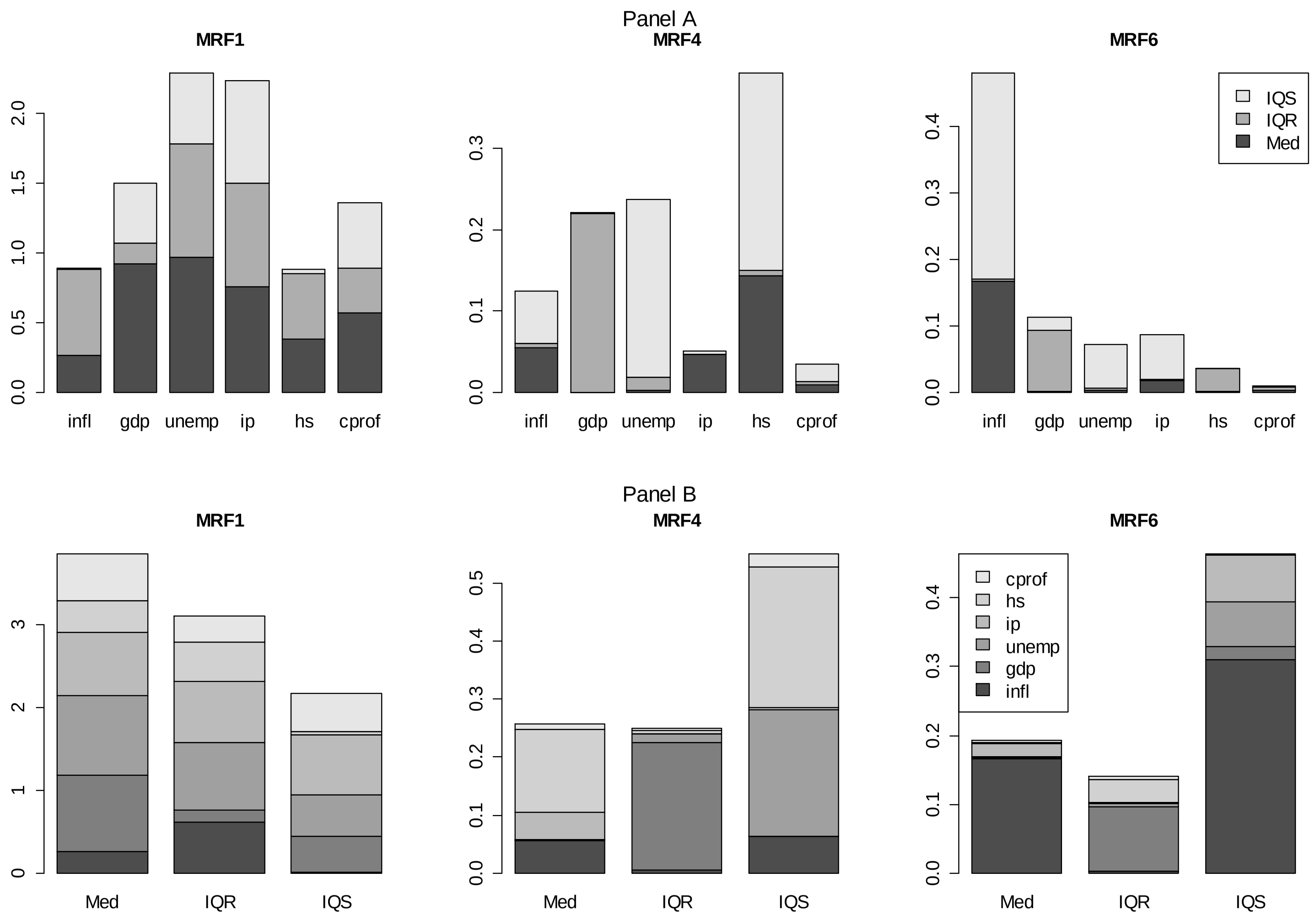

Figure 6: Comovements of macro risk factors and key ex ante macro risks

Notes: Standardized units are reported. Shaded 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.

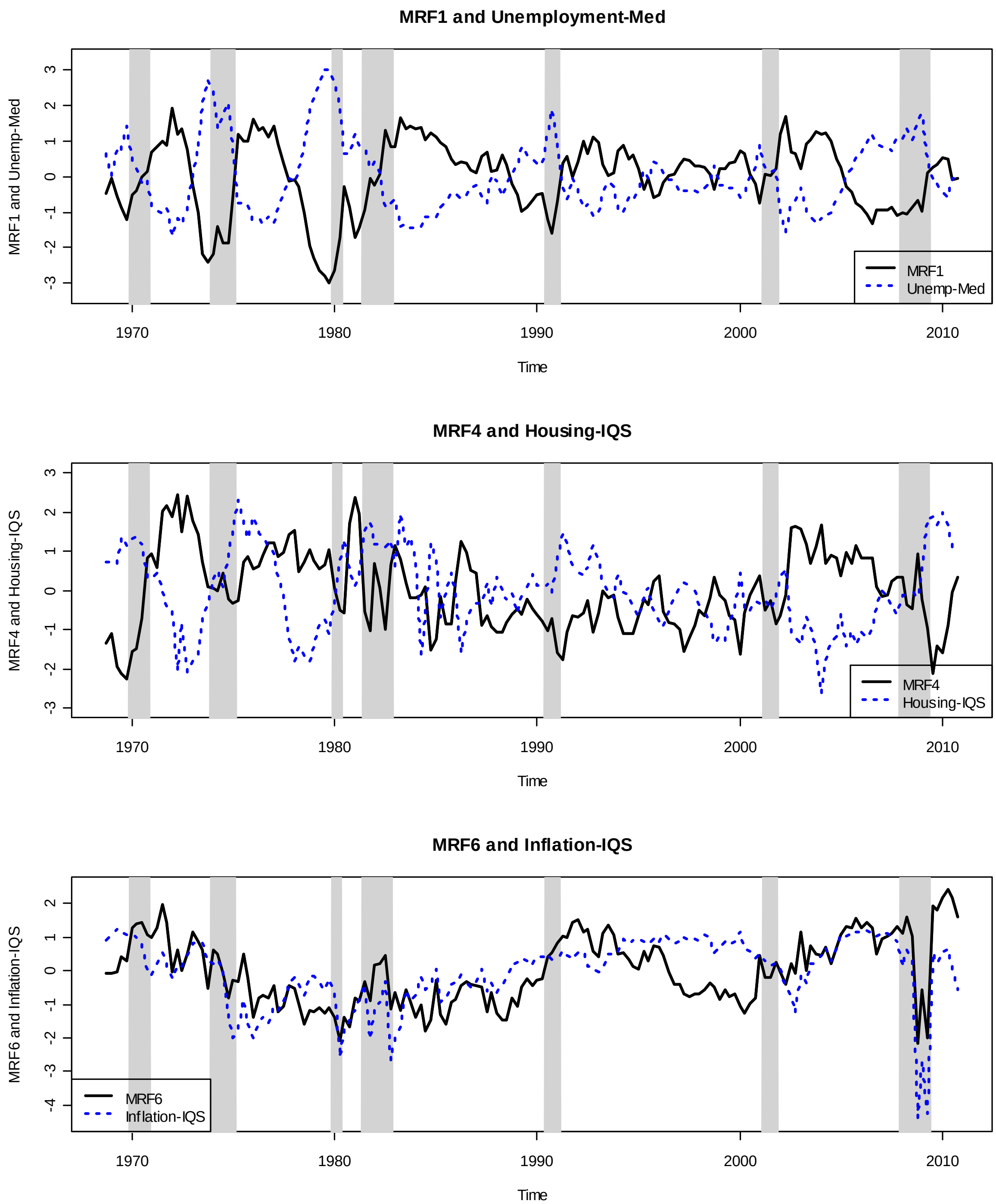

Figure 7: Bond risk premium estimate

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 given in annual frequency.

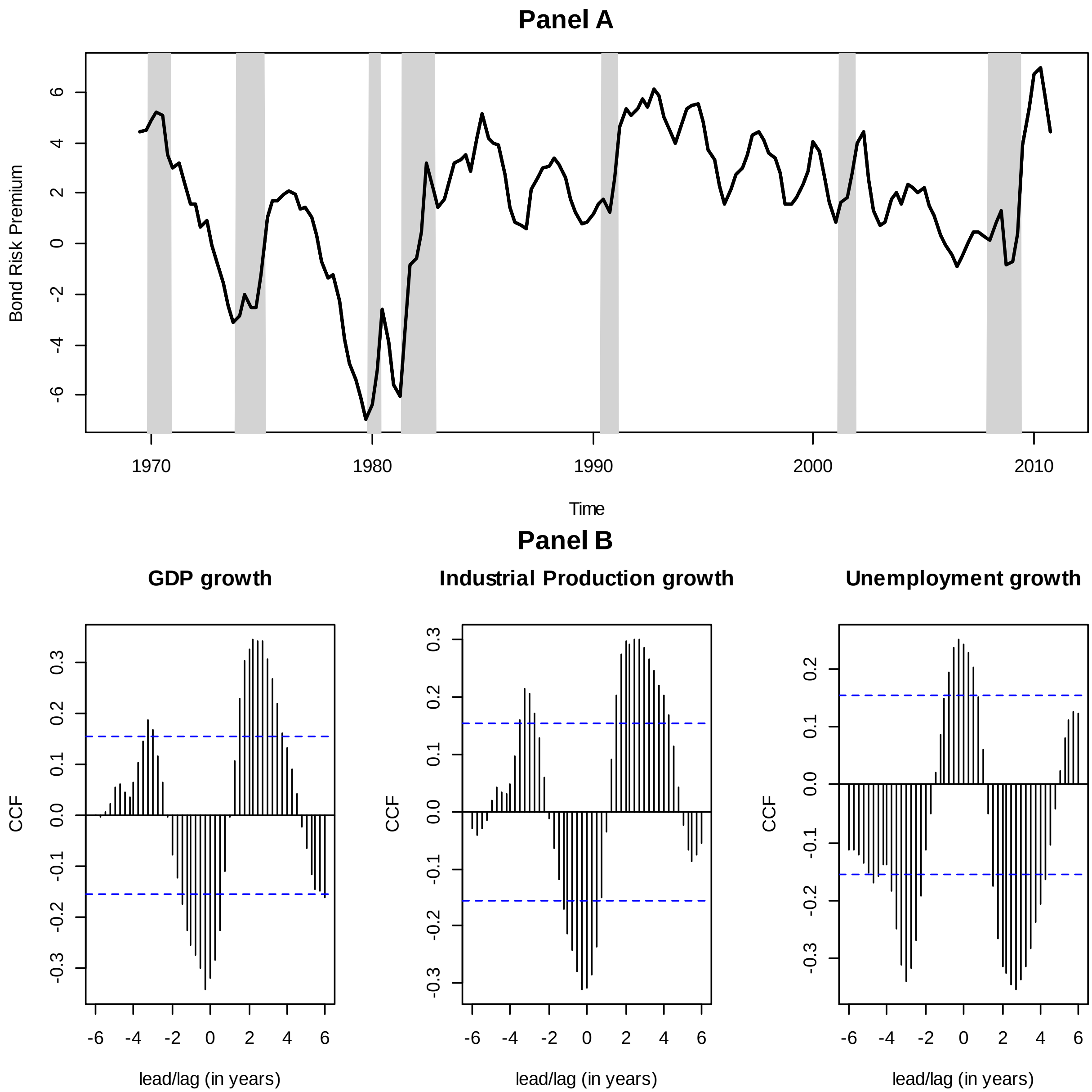

Figure 8: Recursive estimates of R^2_{oos} and utility gains

Notes: Charts in Panel A give 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). Charts in Panel B give 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.