The Cross-Section of Subjective Bond Risk Premia

Andrea Buraschi, Ilaria Piatti, and Paul Whelan

andrea.buraschi@imperial.ac.uk
ilaria.piatti@sbs.ox.ac.uk
pawh.fi@cbs.dk

ABSTRACT

This paper studies the properties of bond risk premia in the cross-section of subjective expectations. To this end we exploit an extensive dataset of yield curve forecasts from financial institutions that allows the identification of heterogeneous $\mathbb{P}$-dynamics. We present a number of novel findings. First, consensus beliefs are a misleading statistic due to a rich dynamics in the cross-section. Second, contrary to evidence presented for stock markets, but consistent with rational expectations, the relation between expectations and realisations is positive, and this result holds for the entire cross-section. Third, we show that optimistic beliefs are more spanned by bond prices and, at the same time, they are the most accurate. Moreover, we show that, out-of-sample, optimistic beliefs outperform popular forecasting models and thus represent a valid measure of bond risk premia that can be used to avoid issues related to in-sample fitting of ex-post returns, when evaluating models. As an application of this result, we study the link between survey forecasts and proxies for state-variables arising in structural models and uncover a number of statistically significant relationships in favour of rational expectations models.

JEL classification: E3, E4, G10, G12, G14, G17
Keywords: Rational Expectations, Cross-Section of Beliefs, Bond Risk Premia, Spanning, Expectation Formation.
First version: January, 2015.
This version: May, 2016.

Andrea Buraschi is Chair of Finance at Imperial College Business School; Ilaria Piatti is at Saïd Business School, University of Oxford; and Paul Whelan is at Copenhagen Business School.

I. Introduction

A large asset pricing literature finds compelling evidence of predictability in several asset markets. A stream of the literature interprets this result as evidence of a time-varying risk premium that can be understood in the context of rational general equilibrium models. A second stream of the literature, on the other hand, argues that several characteristics of this predictability are more likely due to the existence of behavioral biases affecting the dynamics of subjective beliefs, informational frictions, or both. In this paper, we use a detailed data set of investors' forecasts about future interest rates to obtain a direct measure of expected (subjective) bond risk premia. This allows us to study the link between a direct measure of expected bond excess returns and alternative model-implied risk premia proposed by the literature. We use the results of this exercise to address a number of questions about the properties of expected bond risk premia and the formation of expectations.

What are the time-series and cross-sectional properties of expected returns as perceived by investors in real-time? Such questions are important for an emerging literature in financial economics that employs data on actual expectations to investigate testable implications of economic models. Most studies find that survey data contain useful information to predict future economic activity. Some studies argue that agents are better in forecasting some economic variables, as for example economic growth and inflation, than sophisticated econometric models (see e.g. Ang, Bekaert, and Wei (2007) and Aioli and Timmermann (2011)). On the other hand, Greenwood and Schleifer (2014) show that consensus expectations about future stock market returns are negatively correlated with actual realisations, and Koijen, Schmeling, and Vrugt (2015) reach a similar conclusion in the context of global equities, currencies and global fixed income returns across countries. Both these studies then argue that this result is difficult to reconcile with rational expectation models. In contrast to Greenwood and Schleifer (2014) and Koijen, Schmeling, and Vrugt (2015), we focus on bond markets and investigate the efficiency of agents expectations taking advantage of a unique feature of our dataset that identifies agent specific forecasts across both maturity and time-series dimension. This allows us to distinguish between alternative models of formation of expectations and provide evidence about the cross-sectional properties of expected returns that previous studies have been unable to explore.

Other studies have looked at the dynamics of private sector expectations about interest rates and at the dynamics of the corresponding forecast errors, see e.g. Cieslak and Povala (2012) for fed fund rate forecasts and Piazzesi, Salomao, and Schneider (2015) for bond risk premia. While these studies focus only on the consensus forecast, i.e. the median of the cross-sectional distribution of subjective expectations, we show the importance of studying the dynamics and drivers of the full cross-section of agents' beliefs and we argue that the use of consensus expectations to proxy for the expectations of the marginal investor might be misleading.

We begin by constructing measures of expected bond risk premia (EBR) from professional market participants' expectations regarding future yields. Specifically, we use Treasury coupon bond yield forecasts at the agent specific level to obtain a set of constant maturity 1-year zero-coupon bond yield expectations. Individual agent EBRs are then obtained by subtracting the date t observable risk free rate from expected price changes. With these measures at hand we document a number of novel findings.

First, we document a large unconditional heterogeneity in the cross-section of EBR point forecasts. For the median (Q2) forecaster EBRs are 0.26% for 2-year bonds and 1.08% for 10-year bonds. For first quartile (Q1) forecasters EBRs range from -0.06% to -1.56%, and from 0.53% to 3.55% for third quartile (Q3) forecasters, between 2 and 10-year maturities, respectively.

Second, we find clear evidence of persistence in agents forecasts. For example, a forecaster in the first quartile of the cross-sectional distribution of 2-year EBR has a probability of almost 75% to stay in the first quartile the following month, and this probability is about 73% for the 10-year EBR, which is about three times what it should be under the null hypothesis of no persistence. Moreover, some agents are consistently good in their forecasts about excess bond returns, however, on average over the last 25 years agents have been surprised by larger than expected excess bond returns. We find that agents who are the most optimists about economic growth or inflation are also most likely to be in the lowest quartile of the cross-sectional distribution of EBR forecast, and vice versa. This relation is consistent with the idea that good states of the economy are generally characterised by increasing yields, at least at short maturity, decreasing bond prices and thus lower expected excess returns. We also find, however, that this relation with macroeconomic forecasts holds only for the most extreme quantiles of the distribution of subjective bond risk premia.

Third, we can formally test, and strongly reject, the hypothesis that bond risk premia are constant, and we extend this result to the whole cross-section of subjective risk premia showing that expected bond excess returns are time-varying across all deciles of the cross-sectional distribution of forecasters, but agents in the right tail of the distribution believe that expected returns are four times more persistent, and hence more predictable, than agents in the left tail. In general, we find that the beliefs of optimistic agents, identified by higher deciles of the cross-sectional EBR distribution, are on average very well spanned by current bond prices, while the opposite is true for pessimistic investors. For example, for the 2-year bond, regressions of different quantiles of the EBR distribution on the level, slope and curvature of the yield curve produce an R-squared of almost 80% for the most optimist agent (the 90^{th} percentile) and only 25% for the pessimist (the 10^{th} percentile). This result is consistent with two conjectures. The first is the existence of market frictions such as short selling constraints (as in Hong, Sraer, and Yu (2013)). If pessimists cannot sell short, bond prices would just reflect the beliefs of optimists. A second alternative conjecture is based on the hypothesis of market selection in competitive markets. If optimists had been consistently more accurate than pessimists, they would have been accumulating more economic weight in the pricing kernel. To disentangle empirically these two alternative hypothesis we study the dynamic link between subjective expectations and ex-post realisations, carefully distinguishing across deciles in the distribution of subjective expectations.

Fourth, simple predictive regressions of realised excess returns on subjective risk premia show that forecasters tend to under-predict bond excess returns and that the predictive power is relatively low. However, the relation between expectations and realisations is always positive, contrary to what Greenwood and Schleifer (2014) document in the context of the stock market, and to what Koijen, Schmeling, and Vrugt (2015) find in the context of global equities, currencies and global fixed income returns across countries. We also show that the root mean square errors (RMSEs) of the forecasts are monotonically decreasing from the 10^{th} decile to the 80^{th} decile of the EBR distribution, suggesting that optimistic agents outperform pessimistic agents and are the most accurate forecasters in this sample period. This empirical evidence, jointly with the earlier result that the optimist's beliefs are better spanned by time-t bond yields, is very interesting. Indeed it helps to distinguish between the two conjectures illustrated above. In models with short-selling constraints, agents who are active in buying the assets are those who are willing to pay excessively and earn a negative risk premium ex-post. On the other hand, in rational models with competitive markets the marginal agent is the one with the most accurate expectations. His larger accuracy makes him accumulating a bigger relative wealth share. Our results support this second class of models.

Fifth, we find that the out-of-sample performance of the survey-implied bond risk premia are highly competitive in forecasting future realised excess returns relative to some popular reduced form models. Indeed, considering the right tail of the distribution of survey forecasts we find subjective bond risk premia significantly outperform projections implied by either Cochrane and Piazzesi (2005) or Ludvigson and Ng (2009) forecasting factors, or even a combination of them both, for all bond maturities. This findings suggests that surveys can indeed be used to build reliable measures of bond risk premia in real time and thus avoid issues related to in-sample versus out-of-sample model fitting. As an application of this result, we test the relationship between proxies for state-variables arising in rational expectation models and subjective expectations of bond excess returns. To summarise, we find a significant role for rational risk premium proxies that have been proposed by the literature. Moreover, in most cases the empirical sign of the factor loading is consistent with predictions from theory. Finally, taken together in a multivariate regressions these proxies are explaining in excess of 30% of the variation in subjective expected returns. This result stands in contrast to the findings of Greenwood and Schleifer (2014) in the context of equity markets and suggests that rational expectation models cannot be dismissed so easily.

We provide evidence that the survey forecast accuracy is particularly good for the optimists. However, there might be nothing special about being an optimist, aside from having been sufficiently lucky to be more accurate in the rather special 1988-2015 sample period. Their luck might reverse in the next 30 years. Thus, we revisit our results by distinguishing explicitly between periods in which agents have been surprised negatively and positively. We ask the following question: are bond yields spanning the beliefs of the optimists all the time (as suggested by models with short-selling constraints) or is the spanning result reverse when the pessimists are more accurate in their forecasts (as suggested by rational expectation models with heterogeneous agents)? To answer this question we study the relative spanning of optimists and pessimists' beliefs against their ex-ante relative accuracy and we find a very significantly positive relation: a regression of relative spanning, defined as the difference in the R-squared of regressions of EBR on the three principal components of the term structure for the 90^{th} and 10^{th} percentile, on ex-ante difference in RMSE yields an adjusted R-squared of 61.3% and a strongly significant positive slope coefficient.

Finally, we show that during periods of increasing interest rates (which correspond to good states for the U.S. economy in this sample period) the distribution of forecast errors is symmetric around zero. On the other hand, following periods of decreasing short term rates, all agents, including the most optimistic, are surprised by larger excess bond returns. This is consistent with the findings in Cieslak and Povala (2012) who analyze the survey forecast expectations of the fed fund rate and show that the largest errors are negative and occur during and after NBER recessions. We also propose and test some alternative theoretical explanations for the observed bias and state-dependence in forecast errors. A large literature in behavioral finance frequently argues that forecasters form irrational beliefs. Often this argument is tested in the context of extrapolative learning models. The substantial persistence in beliefs reported in the first part of the paper and the predictability of the forecasts errors is - prima facie - consistent with this conjecture. A second stream of the literature has studied rational agents who face informational rigidities. Finally, the observed dynamics and cross section of forecast errors is potentially consistent with models in which forecasters have identical and complete information but asymmetric loss functions with heterogeneity in the degree of loss aversion, or with forecasters engaging in forecast smoothing for reputational considerations.

The paper proceeds as follows. Section II presents our data and provides a description of subjective bond risk premia. In Section III we study the extent to which expected bond returns are linked to the current term structure. Section IV discusses the forecasting power of expected excess bond returns for future realised excess returns and the cross-sectional variations in the forecast accuracy. Section V analyses the dynamics of the forecast errors in order to document the efficiency of the forecasters and the presence of potential biases, and we also propose and test potential theoretical explanations for the state dependence in the forecast errors. Section VI concludes.

II. Descriptive Analysis

This section briefly introduces the data and provides a description of subjective bond excess returns. All data are monthly, from January 1988 to July 2015.

A. Survey data

We construct measures of expected bond risk premia (EBR) directly from professional market participants' expectations regarding future yields. The BlueChip Financial Forecasts (BCFF) is a monthly survey providing extensive panel data on the expectations of professional economists working at leading financial institutions about all maturities of the yield curve and economic fundamentals, such as GDP and inflation.[1] The contributors are asked to provide point forecasts at horizons that range from the end of the current quarter to 5 quarters ahead (6 from January 1997).

BCFF represents the most extensive dataset currently available to investigate the role of expectations formation in asset pricing. It is unique with respect to alternative commonly studied surveys along at least four dimensions. First, the dataset is available at a monthly frequency, while other surveys, such as the Survey of Professional Forecasters' (SPF) is available only at quarterly frequency. This increases the power of asset pricing tests. Second, the number of participants in the survey is large and stable over time. In our sample it is 40 on average, with a standard deviation of about 4.2. Moreover, even considering those forecasters contributing to all maturities and all horizons it never falls below 30. On the other hand, in the SPF the distribution of respondents displays significant variability: the mean number of respondents is around 40, the standard deviation is 13 and in some years the number of contributors is as low as 9. While in the early 70s the number of SPF forecasters was around 60, it decreased in two major steps in the mid 1970s and mid 1980s to as low as 14 forecasters in 1990.[2] Third, Bluechip has always been administered by the same agency, while other surveys, such as SPF, have been administered by different agencies over the years. Moreover, SPF changed some of the questions in the survey, and some of

[1] In our analysis we use agent specific forecasts for the Federal Funds rate, Treasury bills with maturities 3-months/6-months/1-year, Treasury notes with maturities 1,2,5,10-years, and the 30-year Treasury bond.

[2] If one restricts the attention to forecasters who participated to at least 8 surveys, this limits the number of data points considerably.

these changes crucially affected the forecasting horizon.[3] Fourth, the survey is conducted in a short window of time, between the 25th and 27th of the month and mailed to subscribers within the first 5 days of the subsequent month. This allows the empirical analysis to be unaffected by biases induced by staleness or overlapping observations between returns and responses.

We use forecasts at the agent specific level to obtain a set of constant maturity zero-coupon bond yield forecasts, for horizon from 1 to 6 quarters ahead, and focus on the one year ahead projections.[4] Over the whole sample there are 160 forecasters for which we can compute the whole expected term structure of zero-coupon yields and on average they contribute to the cross-section for about 82 months.

B. The cross-sectional distribution of subjective bond excess returns

As common in the literature, we use p_t^n to denote the logarithm of the time-t price of a risk-free zero-coupon bond that pays one unit of the numeraire n-years in the future. Spot yields and forward rates are then defined as $y_t^n = -\frac{p_t^n}{n}$ and $f_t^n = p_t^n - p_t^{n-1}$. We refer to gross and excess returns on a n-period bond by using the notation $r_{t+1}^n = p_{t+1}^{n-1} - p_t^n$ and $rx_{t+1}^n = r_{t+1}^n - y_t^1$, respectively.

Let us denote the expected excess bond returns (EBR) implied by survey forecaster i with an horizon of one year as $erx_{i,t}^n = E_t^i\left[rx_{t+1}^n\right]$, so that

$$erx_{i,t}^n = E_t^i\left[p_{t+1}^{n-1}\right] - p_t^n - y_t^1. \tag{1}$$

Individual expected bond excess returns $erx_{i,t}^n$ can therefore be obtained from individual

[3] For a detailed discussion on the issues related to SPF, see D'Amico and Orphanides (2008) and Giordani and Soderlind (2003).

[4] Since forecast data consist of yields to maturity of coupon-bearing bonds, we construct curves of *expected* zero coupon discount rates via a bootstrap approach. First, we obtain a set of equally spaced (semiannual frequency) yields to maturity by interpolating available yields with the Akima (1970) algorithm. Next, we bootstrap Federal Funds, Treasury Bills, and coupon notes, and the 30 year bond to obtain a set of (simple, semiannual) zero coupon yields. Finally, we convert yields to their continuously compounded counterparts. The final output is a monthly panel data of expected (quarterly horizons out to 1.25-years) zero coupon (continuously compounded) discount rates at evenly spaced maturities between 1 and 10-years (we disregard maturities greater than 10-years).

yield forecasts by observing that:

$$erx_{i,t}^{n} = -(n-1) \times \underbrace{E_t^i\left[y_{t+1}^{n-1}\right]}_{\substack{\text{Survey Yield}\\ \text{Forecasts}}} + ny_t^n - y_t^1. \tag{2}$$

For realized bond data we use zero-coupon bond yields provided by Gürkaynak, Sack, and Wright (2006) which are available from the Federal Reserve website.

First, we document that there exists large unconditional heterogeneity in the cross-section of EBR point forecasts. Table I provides summary statistics for the median (the consensus) and the first and third quartile of the (1-year) EBR distribution for the 2, 5 and 10-year bonds. Throughout the rest of the paper, we refer to *optimists* as those agents whose expected excess returns are above the median and *pessimists* as those whose expected excess returns are below the median.[5] Excluding the mean, the unconditional properties of the three quartiles are quite similar. In all cases the volatility and kurtosis are increasing in bond maturity. However, the *spread* between the Q1 and Q3 unconditional expected excess bond returns is large and sharply increasing with maturity. While consensus (Q2) and optimistic (Q3) investors believe in a positive risk premium, pessimistic (Q1) investors believe in a negative risk premium.

The conditional properties of the cross-sectional distribution of EBR display rich dynamics in the time-series. The top panel of Figure 1 displays the min, Q1, median, Q3 and max of the cross-sectional distribution of EBR for 5-year maturity bonds. This demonstrates there exists significant time-varying heterogeneity around the consensus (Q2) forecast. The bottom panel of Figure 1 makes this point clear for all maturities by plotting the cross-sectional standard deviation of EBR for 2, 5, and 10-year bonds standardized by their full-sample mean EBR. There are a few interesting take-aways from this plot: first, longer maturity bonds display a clear downward trend in dispersion over time, while no trend exists for shorter maturity bonds. Secondly, dispersion tends to rise at the onset of recessionary periods and drop again as the economy recovers.[6] Third, there was a large drop in dispersion during the Quantitative Easing period, which reversed during 2014 and continues to rise

[5]Note that optimists expect higher bond returns, thus lower yields, while pessimists are expecting higher bond yields.

[6]The countercyclicality of the dispersion in beliefs is consistent with the empirical evidence in Patton and Timmermann (2010) and Buraschi, Trojani, and Vedolin (2014), among others.

until the end of our sample.

B.1. Persistence in forecasters' optimism and pessimism

A second interesting question is related to the extent to which individual forecasters are regularly in one particular quartile of the cross-sectional distribution of subjective expected bond returns. Figure 2 plots the time series average of four individual forecasters' positions in the cross-sectional distribution of subjective expected bond returns, for maturities between 2 and 10 years. This plot is suggestive of some persistence in the individual forecasts.

In order to address this question more systematically, we rank all forecasters according to whether in a given month t their expected bond return is in the first, second, third or fourth quartile of the cross-sectional distribution. We repeat this exercise for all months in the sample and compute transition probabilities. In other words, we compute the probability that forecasters in a given quartile at time t stay in that particular quartile in $t+1$ or move to a different quartile of the distribution. We do that separately for two different bond maturities (2 and 10 years) in Table II. If views are not persistent, all the entries in Table II should be approximately equal to 25%, while we expect the diagonal elements to be significantly higher than 25% in presence of persistent EBRs, and this is exactly what we find, in particular for the most extreme quantiles, Q1 and Q4. For example, a forecaster in the first quartile of the cross-sectional distribution of 2-year EBR has a probability of 75% to stay in the first quartile the following month, and this probability is 73% for the 10-year EBR, which is about three times what it should be under the null hypothesis of no persistence. In all cases, the probability of remaining in the same quartile is significantly higher than 25% at a level of 5%. The results suggest that forecasters are persistently optimistic or pessimistic relative to the consensus excess return. This is consistent with the results in Patton and Timmermann (2010) with regards to macroeconomic forecasts. The persistence in forecasters' optimism and pessimism about expected excess bond returns allows us to focus on different quantiles of the cross-sectional distribution instead of individual agents when we study the behaviour of optimistic versus pessimistic agents in the rest of the paper.

What is the extent to which expected bond returns are linked to expectations about future economic fundamentals? We first use the above methodology to compute transition probabilities for GDP and CPI forecasts in our sample (see Table III). Two results emerge:

first, macro forecasts are also extremely persistent, and the transition probabilities are of the same magnitude as for the EBR forecasts.[7] Second, since we know the name of the forecaster, we can study the link between each individual yield forecasts and their macroeconomic forecasts. We find that agents who are marginally more optimistic or pessimistic about macroeconomic variables are not consistently in one particular quartile of the cross-sectional EBR distribution, as shown in Table IV. There seem though to be an interesting pattern at the corners of these tables: for example, an analyst in the first quartile of the EBR distribution will be also in the fourth quartile of the GDP (or CPI) distribution with a probability between 37% and 41%, depending on bond maturity, which is significantly higher than 25%. Macro optimists are thus most likely in the lowest quartile of the cross sectional distribution of EBR forecast, and vice versa. This relation is consistent with the idea that good states of the economy are generally characterised by increasing yields, at least at short maturity, decreasing bond prices and thus lower expected excess returns, but the probabilities are not impressive, suggesting that the drivers of beliefs about the yield curve and the macroeconomy (GDP and inflation) are largely different.[8]

C. Time-varying expected returns

An extensive literature in fixed income studies the properties of bond risk premia and argues that these are time varying. Empirical proxies of conditional bond risk premia usually either require the specification of a model or they use ex-post data on bond returns. The limit of arguments based on the central limit theorem is of course the lack of sufficiently long data samples. For this reason, some studies have argued that the results are not statistically convincing.

Our data allows us to study bond risk premia using directly the dynamics of expectations that are obtained in a model independent way. We can formally test the null hypothesis that bond risk premia are constant and reject the null at the 1% confidence level. The results are very strong and support the hypothesis that expected excess bond returns are indeed time

[7]The evidence of persistence in excess bond returns and macroeconomic forecasts is even stronger than what Patton and Timmermann (2010) document for macroeconomic forecasts using data from the Consensus Economics Inc, at a quarterly frequency.

[8] Interestingly, unreported results also show that optimism or pessimism about GDP growth is not related to optimism or pessimism about inflation: joint probabilities are close to 25% for all elements of the joint transition matrix.

varying, i.e. excess bond returns are predictable. Table V reports the results of regressions of the second quartile of the cross-sectional distribution of subjective excess returns (the consensus) of 2, 5 and 10-year zero-coupon bonds on a constant and their own lag at the 1-year horizon. The slope coefficients are significantly different from 0 and 1 at all levels. It is interesting to look at the results of the same regression for the different quantiles of the cross-sectional distribution of expected bond returns. Figure 3 plots the cross-sectional distribution of EBR 1-year autoregression coefficients and associated R^2. The results are striking: moving from the 10^{th} percentile to the 90^{th} percentile, for all bond maturities, the autocorrelation coefficient is monotonically increasing. Considering the R^2 in these regressions, forecasters in the right tail of the distribution believe EBR are about 4 times more predictable than forecasters in the left tail of the distribution. In other words, EBR pessimists believe expected returns are less persistent (hence less predictable in an R^2 sense) than optimists.

III. Subjective Expectations and the Yield Curve

An important stream of the fixed income literature discusses the spanning properties of the term structure of interest rates. This literature addresses the important question of whether a sufficiently rich cross-section of current prices reveals enough information which is relevant for the dynamics of bond returns. In traditional dynamic general equilibrium models, for instance, in absence of frictions bond prices span the priced risk factors. Cochrane and Piazzesi (2005) provide supporting evidence for this conjecture by showing that a tent-shaped combination of forward rates helps to explain bond excess returns. In the context of heterogeneous economies, equilibrium bond prices could be affected by the beliefs of the optimists, the pessimist, the wealth-weighted average of all beliefs, or none of the above. In the heterogeneous beliefs models studied by Hong, Sraer, and Yu (2013), for instance, agents are subject to short selling constraints. This implies that equilibrium prices span the beliefs of the agents who are the most optimist in terms of the assets returns. In the general equilibrium models with disagreement and no frictions (as in Basak (2005), Buraschi and Jiltsov (2006), Jouini and Napp (2006), Xiong and Yan (2010), Chen, Joslin, and Tran (2012), Buraschi and Whelan (2012), among others), on the other hand, bond prices span the wealth-weighted average of optimists and pessimists beliefs. This weight depends on the wealth share of the agents in the economy and therefore is a function of the accuracy of their forecasts made in the past. Indeed, on the basis of the survival argument of Friedman, the spanned beliefs should neither be those of the optimist nor of the pessimist but those of agents whose beliefs are the most rational (i.e. closest to the actual physical probability).

In this section, we use information on agents beliefs from both the time series and the cross section to address which beliefs are spanned by current bond prices. To proceed parsimoniously, we decompose the yield curve up to 10 years maturity in a small number of (orthogonal) principle components. These factors are often labelled in the literature as level, slope, and curvature, based on how shocks to these factors affect the shape of the yield curve (see, for example, Litterman and Scheinkman (1991), Dai and Singleton (2003), or Joslin, Singleton, and Zhu (2011)).[9]

Then, we run regressions of EBRs for different deciles of the distribution onto these factors:

$$erx^n_{i,t} = \beta^n_{i,0} + \beta^n_{i,1} Level_t + \beta^n_{i,2} Slope_t + \beta^n_{i,3} Curv_t + \epsilon^n_{i,t}. \tag{3}$$

As a comparison, we also run regressions using the realized 1-year excess bond returns, $hprx^n_{t+1}$. Table VI and Figure 4 summarize the estimated regression coefficients and adjusted R-squared for $n = 2, 5$ and 10 years.

Table VI, Panel A, reports the results for consensus beliefs. We find that for the 2-year bond all coefficients are statistically significant at the 1% level and explain 57.47% of the variation in EBR. The loadings on level and slope factors are positive while the loading on the curvature factor is negative. Considering 5-year and 10-year bonds the loadings on level and slope remain statistically significant and positive, while the loading on curvature becomes insignificant. Level and slope factors jointly explain 34% and 22% of variation on 5 and 10-year bonds, respectively. In unreported univariate regressions, we find that close to half of this explanatory power is due to the level, while half is due to the slope of the term structure.

Figure 4 summarizes the regression when we distinguish across different deciles of the

[9]As usual, we find a level factor explains the vast majority of variation ($\sim$ 85%), a slope factor which we rotate such that a positive shock raises the long end of lower the short end of the term structure, and a curvature factor for which shocks raise mid maturities relative to short and long maturities.

EBR distribution. Agents who are pessimistic about future bond returns are in the lower deciles of the distribution. Figure 4 plots the factor loadings and adjusted R-squares. Consistent with Table V, optimistic investors expectations are very well spanned by the cross-section of the yield curve, while the opposite is true for pessimistic investors. Also, we note an interesting pattern for the conditional impact of the curvature factor: for 10-year bonds it is positive and monotonically increasing across percentiles, while for the 2 and 5-year bonds it is negative and monotonically decreasing. Moreover, the curvature coefficient is significantly different from zero across deciles for the 2-year bond but not for the 5-year bond, and only for the most optimistic deciles for what concerns the 10-year bond. This result is interesting since many studies attribute the curvature factor to the impact of monetary policy (see, for example, Piazzesi (2005)).

Panel B of Table VI reports return predictability regressions using ex-post realized returns as a proxy for ex-ante bond risk premia. Consistent with the large literature on bond return predictability, we find that the slope of the yield curve reveals important information about bond risk premia (see Campbell and Shiller (1991) and Fama and Bliss (1987)). For 10-year zero-coupon bonds, we obtain an R-squared of 18% with a t-statistic on the slope factor significant at the 1% level. The level, however, contains no information regarding future realized return.

If one compares the results in the two panels of Table VI, the difference is striking. On the basis of Panel B, one might be tempted to conclude that the amount of spanning is somewhat limited. On the other hand, when one considers direct measures of subjective expected returns, there is strong evidence that the variation in subjective bond risk premia is largely spanned by date t yield factors. Moreover, while for realized returns the explanatory power is only due to the slope, for subjective returns explanatory power is coming from both the level and slope of the term structure.

Taken together, two conclusions emerge from these findings. First, one should be careful when trying to infer the effects of changes in the current yield curve on expectations. This is due to the large heterogeneity in expectations and their heterogeneous impact on equilibrium prices. Second, the beliefs of optimists appears better spanned by contemporaneous prices than the beliefs of pessimists. This result is intriguing and consistent with two conjectures. The first is the existence of market frictions such as short selling constraints (as in Hong, Sraer, and Yu (2013)). If pessimists cannot sell short, bond prices would just reflect the beliefs of optimists. A second alternative conjecture is based on the hypothesis of market selection in competitive markets. As Alchian (1950) argued, "*Realized profits, not maximum profits, are the mark of success and viability. It does not matter through what process of reasoning or motivation such success was achieved. The fact of its accomplishment is sufficient. This is the criterion by which the economic system selects survivors: those who realize positive profits are the survivors; those who suffer losses disappear.*" If optimists had been consistently more accurate than pessimists, they would have been accumulating more economic weight in the pricing kernel. To distinguish empirically between these two alternative hypothesis, in the next section we study the link between subjective time-t expections and actual time-$(t+1)$ realizations, after carefully distinguishing across deciles in the distribution of subjective expectations.

IV. Predictive Performance

A. Forecasting regressions

The natural starting point for our analysis of the survey's forecasting performance is a simple predictive regression of realized excess returns on the consensus risk premium from survey forecasts:

$$rx_{t+1}^{n} = \alpha_c^n + \beta_c^n erx_{c,t}^n + \epsilon_{c,t+1}^n. \tag{4}$$

Table VII reports the results of this regression, for bond maturities of 2, 5 and 10 years. If survey expectations measure true expected excess returns, they should predict future realized excess returns with an intercept α_c^n of zero and a slope coefficient β_c^n of one. We find that the slope coefficients, β_c^n, are positive, but they are all smaller than one and the difference is statistically significant for the 10-year bond. Moreover, the intercept, α_c^n, is always positive and statistically significant. This implies that the consensus forecast tends to under-predict bond excess returns. The predictive power, measured in terms of adjusted R-squared, is relatively low, between 2% and 5% depending on bond maturity. However, the relation between expectations and realizations is always positive, contrary to what Greenwood and Schleifer (2014) document in the context of the stock market, and to what Koijen, Schmeling, and Vrugt (2015) find in the context of global equities, currencies and global fixed income returns across countries.

We have shown in the previous sections that consensus forecasts are not always representative of the cross-sectional distribution of forecasters. Therefore, we also run predictive regressions for each different decile $i = 0.10, \ldots, 0.90$:

$$rx_{t+1}^{n} = \alpha_i^n + \beta_i^n erx_{i,t}^{n} + \epsilon_{i,t+1}^{n}. \tag{5}$$

Figure 5 shows regression coefficients and R^2 of regressions (5) for each decile. The intercepts, α_i^n, are all positive and significant; the slope coefficients are always positive for all deciles of the distribution. At the same time, we find that they are lower than one for all quantiles and bond maturities, ranging between about 0.3 and 0.6, and the R-squares vary between 1% and 6%. The R-squares show interesting patterns in the cross section: for the short-term bond, the slope coefficient and the R-squared are increasing when moving from pessimist to optimistic agents while the opposite is true for long-term bonds. The intercept instead is monotonically decreasing for all bond maturities, and it is only marginally significant for the most optimistic forecasters (above the 80^{th} percentile).[10] These findings also suggest that expectations of future excess bond returns are indeed correlated with future realization of bond risk premia.

B. *Forecast accuracy*

To further investigate the different accuracy across the distribution, Panel A of Table VIII reports the RMSEs for deciles $i = 0.10, \ldots, 0.90$ and for bond maturities $n = 2, 5, 10$:

$$RMSE_i^n(Surv) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(rx_{t+1}^{n} - erx_{i,t}^{n}\right)^2}. \tag{6}$$

This panel shows that over the full sample RMSEs are monotonically decreasing from the 10^{th} decile to the 80^{th} decile. This means optimistic agents outperform pessimistic agents and are the most accurate forecasters in this sample period.

[10] In the interest of space, we do not show the cross-sectional distribution of t-statistic. The slope coefficients are not significant at the 5% level for the 2-year bond, while they are significant for the 5-year bond across deciles and for the 10-year bond only up to $i = 0.30$.

At this point it is interesting to study how accurate survey forecasts are in comparison with statistical models usually employed in the fixed income literature to predict excess returns. Therefore, we also study the forecast accuracy of two reduced form predictability factors studied extensively in the literature:

- The Cochrane and Piazzesi (2005) return forecasting factor, which is a tent-shaped linear combination of forward rates that has been shown to contain information about future bond returns, and subsumes information contained in the level, slope and curvature of the term structure. We denote this factor CP.

- The real macro factor uncovered by Ludvigson and Ng (2009) in a panel of macro economic aggregates that links cyclical fluctuations in bond market premiums to economic activity. We update the Ludvigson and Ng (2009) dataset through 2015 (where possible) but throw away any information on prices so that our panel only contains information on stationary growth rates. The predicting factor is then the first principle component of this panel which we denote LN.

Panel B of Table VIII reports the in-sample RMSEs of the risk premia implied by these two statistical models, again for bond maturities of 2, 5 and 10 years. Comparing the in-sample RMSE of models with the RMSE of surveys it is evident that the models outperform even the best forecasters in-sample. However, the in-sample estimates of the model-implied risk premia are obtained by fitting a regression of the ex-post observed realized excess returns on the factors, and thus they make use of information which is not available to the forecasters in real time. Therefore, a fairer comparison between models and surveys should look at the out-of-sample performance.

Goyal and Welch (2008) report the different performance of several well-known models in-sample and out-of-sample in predicting stock returns. Their results show that running model specification tests using short data sets is challenging and in-sample test statistics can be misleading. The out-of-sample performance of statistical models in bond markets is less studied, but we find that even in bond markets in-sample and out-of-sample goodness-of-fit statistics can be largely different.[11] Panels C and D of Table VIII show the out-of-sample

[11] We also find that the out-of-sample RMSE of the models is quite sensitive to the sample period considered and to the choice of the starting date for the out-of-sample period.

RMSE of statistical models relative to the consensus and to the optimistic (the 90^{th} decile) forecaster. We choose January 1998 as a starting date for our out-of-sample forecasts and compute model-implied expectations recursively with an expanding window. Interestingly, the optimist performs better than all models at all maturities. For example for the CP factor the relative RMSE, i.e. $RMSE^n(Model)/RMSE^n_{0.9}(Surv)$, are about 1.13 for the 2 and 5-year bond and about 1.09 for the 10-year bond. Notice that survey forecasts are out-of-sample by construction: agents form their expectations of time $t+1$ returns only using information available at time t.

Overall, we find that out-of-sample several sets of agents in the decile distribution can outperform statistical models in forecasting bond excess returns and this is particularly true for the optimists. This last finding provides additional evidence that surveys can be used to build reliable measures of bond risk premia, especially when looking at the appropriate quantile of the cross-sectional distribution of agents' beliefs.

This empirical evidence, jointly with the earlier result that time-t bond yields span the beliefs of the optimist, is very interesting. Indeed it helps to distinguish between two alternative classes of models. In models with short-selling constraints, agents who are active in buying the assets are those who are willing to pay excessively and earn a negative risk premium ex-post. On the other hand, in rational models with competitive markets the marginal agent is the one with the most accurate expectations. His larger accuracy makes him accumulating a bigger relative wealth share. The joint evidence from Table VIII, Panel A, and Figure 4 supports this second class of models.

C. *The dynamics of subjective forecasts*

There might be nothing special about being an optimist, aside from having been sufficiently lucky to be more accurate in the rather special 1988-2015 sample period. Their luck may reverse in the next 30 years. Thus, we revisit our results by distinguishing explicitly between periods in which agents have been surprised negatively and positively. We ask the following question: Are bond yields spanning the beliefs of the optimists all the time, as suggested by models with short-selling constraints, or is the spanning result reverse when the pessimists are more accurate in their forecasts, as in rational expectation models with heterogeneous agents?

First, we study the dynamics of the subjective return forecasts by computing the RMSE in Equation (6) for the consensus forecaster on rolling windows of 120 months, for bond maturities $n = 2$ and 10 years. The upper panel of Figure 6 displays the RMSE realized at the last date within each rolling window, standardised in order to make the dynamics for the two maturities comparable. Indeed, we find that the predictive ability of the survey forecasts display interesting time variation. In the early part of the sample, before the burst of the dot-com bubble, the RMSEs for all maturities tend to increase, but the error on the 10-year bond started from an higher level, since it was already above its mean in 1998. Post dot-com bubble, and before the financial crisis, the predictive errors of subjective expectations of bond returns decrease and they are quite correlated across bond maturities. Between 2007 and 2011 instead the RMSEs on the 10-year bond behaves very differently from that on the 2-year bond: the RMSE on the short maturity bonds increases during the financial crisis but then falls again, while the 10-year RMSE decreases during the recession and then rises sharply. It is interesting to note that all RMSEs are declining post 2012, but the RMSE on 10-year bonds remains above average. This period coincides with the second part of quantitative easing policies by the Federal Reserve (the Federal Reserve Maturity Extension Program - known as Operation Twist - and QE3) which was accompanied by significant effort to provide forward guidance to the market.[12]

It is interesting to look at the relation between the time variation in the forecast accuracy and the spanning of beliefs by the cross-section of bond prices, discussed in Section III. The bottom panel of Figure 6 displays the R-squared in the spanning regression (3) for the consensus forecaster on rolling windows of 120 months, for bond maturities $n = 2$ and 10 years. As shown in the previous section, the spanning R-squared is much higher for the short than for the long-term bonds, and the difference is particularly noticeable in the first half of our sample. Starting from 2007, the degree of spanning decreases drastically for all bond maturities, from about 70% (60% for the 10-year bond) to around 30%, and then starts to

[12]On the 21'st September 2011, the FOMC announced the implementation of 'Operation Twist' whereby the Fed intended to purchase 400 billion of bonds with maturities of 6 to 30 years and to sell an equal amount of bonds with maturities of 3 years or less. The objective of this program was to provide liquidity while putting downward pressure on yields without expanding the Fed balance sheet. On the 13^{th} of September 2012, the Fed announces a third round of quantitative easing with an open-ended commitment to purchase 40 billion agency mortgage backed securities per month 'until the labor market improves substantially'. QE3 was intended to operate alongside Operation Twist, allowing the Fed to neutralize longer term securities purchases by selling shorter term Treasuries.

pick up again slowly from 2012.

While Figure 6 shows the time variation in the forecast accuracy and spanning for the consensus forecast across maturities, it is interesting to look at the cross-section of subjective beliefs and study the relation between the time variation in the forecast accuracy of optimists and pessimists and the spanning or their beliefs by the cross-section of bond prices. Rational models with disagreement and no frictions would predict that the agents who have been more accurate in the past accumulate more weight in the pricing kernel, so that their beliefs are ex-post more spanned by bond prices. In order to test this hypothesis, Figure 7 displays the time series of forecast accuracy (upper panel) for the pessimistic (D10), consensus (Q2) and optimistic (D90) agents, and the subsequent spanning of their forecasts (bottom panel), for the 2-year bond, for which the overall level of spanning is the highest, based on the results in Figure 6. It is clear from Figure 7 that periods in which agents are more accurate in their forecasts are followed by a larger degree of spanning of their beliefs by current bond prices.[13] Moreover, the ranking of the degree of spanning of the agents' beliefs is consistent with their past accuracy's ranking. In particular, the RMSEs of the optimist and of the consensus are very similar in the first part of the sample, until just before the recent financial crisis, and so are their R-squares in the spanning regressions. In the last part of the sample the optimist consistently makes lower forecast errors and so he accumulates a larger weight in the pricing kernel, which is then reflected in a larger degree of spanning compared to the consensus. Similar dynamics hold for the 5 and 10-year bonds (not reported), except that between 2005 and 2007 the order in the beliefs' spanning is reversed: the R-squared of the pessimist is higher than the R-squared of the consensus and of the optimist, despite their past forecast errors are lower. Panel (b) of Figure 6 suggests that the information included in the term structure might not be enough to eplain agents' forecasts on the long-term bonds, and we address this issue in the next subsection. The inclusion of additional variables might also restore the link between accuracy and spanning for the longer maturity bonds.

A more detailed picture of this link between accuracy and spanning can be obtained by focusing on the difference between optimist and pessimist. Figure 8 shows the relative spanning of optimist (D90) and pessimist (D10) beliefs against their ex-ante relative accuracy,

[13] Note that times in the upper and bottom panels have been aligned for simplicity of comparison, and correspond to the times in which the ex-ante forecasts are realized.

for the 2-year bond excess returns. The relative spanning is defined as the difference in the R-squared of regressions of EBR on the three principal component of the term structure for the 90^{th} and 10^{th} percentile, now using rolling windows of 60 months. The relative accuracy is defined as the difference in RMSE of the 10^{th} and 90^{th} percentile, over the same rolling windows, but with a lag of 1 year. Indeed, we find a very significantly positive relation: a regression of relative spanning on ex-ante difference in RMSE yields an adjusted R-squared of about 61.3% and a positive slope coefficient with a t-statistic, Newey-West corrected, of more than 10.[14] For the 5 and 10-year bonds the link is less strong, but it becomes significant if we increase the lag between current beliefs and past errors. This might be consistent with a situation in which errors on long term bond excess return forecasts are not realized in the following year but instead compounded with future 1-year forecast errors on bonds with decreasing maturity for the duration of the original bond. Therefore, RMSEs on long term bonds might not immediately affect the weight of the agents in the pricing kernel.

D. Rational expectation models vs subjective risk premia

The empirical evaluation of rational expectation models is traditionally conducted by proxying expected risk premia with sample average of future returns: $E(rx_t)$ is proxied by $\frac{1}{T}\sum_{t=1}^{T} rx_t$ and $E(rx_{t+\tau}|F_t)$ with sample projections of future realizations $rx_{t+\tau}$ onto observables in F_t. This is potentially problematic since one could argue that sample projections based on future realizations can potentially be quite different than true investors expectations. Thus, direct measures of subjective expectations provide useful information to compare alternative specification of bond risk premia. Under the assumption that erx_t measures expectations of bond excess returns accurately, alternative models should be ranked based on their ability to explain the dynamics of erx_t, as opposed to rx_{t+1}.

The previous subsection showed that, out-of-sample, survey-implied bond risk premia are highly competitive in forecasting future realised excess returns relative to some popular reduced form models. These findings provide evidence that surveys can indeed be used to build reliable measures of bond risk premia, and suggests that they can be used to evaluate

[14] Since the left-hand side variable in this regression is estimated from other regressions, on overlapping samples, the properties of the t-statistic might not be standard, even if we correct for autocorrelation in the errors using Newey-West standard errors. However, the scatter plot shows a clear link between left- and right-hand side variables in this regression, even if the value of the t-statistic and R-squared of the regression must be taken with caution.

models in real time, avoiding the issues related to in-sample and out-of-sample model fitting. In fact, if surveys provide a good measure of bond risk premia, the (absolute) correlation between this measure and each factor would provide us with a simple and fully nonparametric way to separate good models from the bad, as long as the sign of the correlation is consistent with the economic intuition behind the factor.

From table VIII we see that optimistic forecasters (those in the right tail of the distribution) are especially accurate in their predictions. In the following, we evaluate a set of models by taking the risk premium implied by our optimistic agent's beliefs - the 90^{th} percentile of the cross-sectional subjective excess return distribution - and by running the following regressions:

$$erx^n_{0.9,t} = a^n_{0.9} + b^n_{0.9} F_t + \epsilon^n_{0.9,t}, \tag{7}$$

where the factors, F_t, we consider are grouped into three categories: (a) proxies for state-variables that arise in structural models, (b) volatility related factors, and (c) alternative predictor variables that have been studied by the literature.

Structural Factors

- Buraschi and Whelan (2016) test the predictions of a rational expectations model with multiple risk tolerant agents. These authors show that disagreement about real growth rates is a significant determinant of expected excess bond returns and, in addition, study the role of inflation disagreement (see Ehling, Gallmeyer, Heyerdahl-Larsen, and Illeditsch (2013) and Hong, Sraer, and Yu (2013) for a theoretical discussion). We denote their proxies for real and inflation disagreement as $DiB(g)$ and $DiB(\pi)$, respectively.

- From Fontaine and Garcia (2012) we take the funding liquidity factor (Liq), which is a time-varying risk factor in economies in which financial intermediaries face priced shocks to funding conditions.

- Vayanos and Vila (2009) and Greenwood and Vayanos (2014) study term structure models in which risk-averse arbitrageurs absorb shocks to the demand and supply for Treasury bonds. These shocks alter the price of duration risk and thus affect both bond yields and expected returns. Building on these authors work Malkhozov, Mueller, Vedolin, and Venter (2016) study a model with endogenous supply shocks which predicts that the outstanding quantity of mortgage-backed securities duration is positively linked to future excess bond returns. We denote their duration factor $Dura$.

- In economies with external habit preferences, such as Campbell and Cochrane (1999), time variation in risk compensation arises because of an endogenously time-varying price of risk. Shocks to the current endowment affect the wedge between consumption and habit, i.e. the consumption surplus, which induces a time-varying expected returns. To obtain a proxy for consumption surplus ($Surp$) we follow Wachter (2006) and use a weighted average of 10 years of monthly consumption growth rates:

$$Surplus = \sum_{j=1}^{120} \phi^j \Delta c_{t-j},$$

  where the weight is set to $\phi = 0.97^{1/3}$ to match the quarterly autocorrelation of the P/D ratio in the data.[15]

- In long-run risk economies with recursive preferences (see e.g. Bansal and Yaron (2004)), time variation in risk compensation arises from economic uncertainty (second moments) of the conditional growth rate of fundamentals. We obtain a proxy for economic uncertainty following Bansal and Shaliastovich (2013). First, we use our survey data on consensus expectation of GDP growth and inflation and fit a bivariate $VAR(1)$. Then, we regress the sum of the squared residuals between t and $t+12$ months on time-t yields. Finally, we take the square root of the fitted values as an estimate of conditional volatility of expected real growth ($LRR(g)$) and expected inflation ($LRR(\pi)$).

Volatility Factors

An important stream of the literature studies the empirical performance of stochastic volatility models in terms of their ability to forecast expected future yield changes. For example, Dai and Singleton (2000) provide a detailed study of the completely affine class of

[15]For consumption data we obtain seasonally adjusted, real per-capita consumption of nondurables and services from the Bureau of Economic Analysis.

term structure models in which elements of the state vector that affect bond volatility also affect expected returns. In an equilibrium context Le and Singleton (2013) emphasise that priced volatility risks also carry over to structural models where the state vector follows an affine diffusion.[16] Motivated by this literature we consider three proxies for volatility risk:

- The intra-month sum of squared returns on a constant maturity 30-day Treasury bill as a proxy for short rate volatility, denoted by $\sigma_y(3m)$.
- The treasury variance risk premium (average across maturity) as studied by Mueller, Vedolin, Sabtchevsky, and Whelan (2016) in the context of a continuous time long-run risk economy, which we denote $TVRP$.
- The realised treasury jump risk measure first studied by Wright and Zhou (2009), updated and extended across maturities, which we denote $Jump$.

Finally, we also study the Cochrane and Piazzesi (2005) and Ludvigson and Ng (2009) forecasting factors as outlined above. Tables IX and X report results from regressions of 2-year and 10-year (D90) subjective excess bond returns on factors.

Considering regression specification (i) we see that disagreement about real growth loads positively on bond risk premia with a t-statistic significant at the 1% level. Inflation disagreement does not enter significantly for 2-year bond risk premia but is significant at the 1% level for 10-year bond risk premia and enters with a negative sign. Taken together disagreement on real growth and inflation explain 8% and 13% of the variation in 2 and 10-year subjective bond risk premia, respectively. Moreover, the factor loadings are consistent with the rational disagreement models of Buraschi and Whelan (2016) for real growth and Hong, Sraer, and Yu (2013) for inflation. Regressions on the liquidity factor of Fontaine and Garcia (2012) show a statistically significant negative relationship, consistent with the interpretation that negative shocks to this factor are bad news for funding conditions, and thus raise expected returns. This result is particularly strong for the 10-year bond risk premium with a factor loading significant at the 1% level and an adjusted R-squared of 12%. Regression (iii) shows a positive statistical relationship with duration risk with $\overline{R}^2$ of 14% and 2%, on 2 and 10-year

[16] Such models include the class of long-run risk models (Bansal and Yaron (2004), Bollerslev, Tauchen, and Zhou (2009), or Bansal and Shaliastovich (2013)), habit models (Wachter (2006) or Buraschi and Jiltsov (2007)) or models with heterogeneous agents (Buraschi and Whelan (2012) or Piatti (2014)).

bonds. Again, the sign of the factor loading is consistent with the theoretical predictions of Malkhozov, Mueller, Vedolin, and Venter (2016) in which positive shocks to the supply of MBS affect the price of interest rate risk. Regression (*iv*) reports estimates on our proxy for consumption surplus. The results suggest no statistical relationship for either maturity, however, the sign of the regression is consistent with the theoretical prediction that negative shocks to surplus raise expected returns on risky assets. The next rows show our proxies for economic uncertainty are explaining a large proportion of the variance of survey expected returns, with $\overline{R}^2$ of 22% and 31%, on 2 and 10-year bonds. Both real and inflation uncertainty are highly significant and enter with positive signs. In terms of theoretical predictions the sign on inflation uncertainty is consistent with the rational non-neutrality model of Bansal and Shaliastovich (2013), while the positive sign on real growth is not supported by (current) theory (see, for example, Bansal and Yaron (2004)). Specification (*vi*) runs a multivariate horse race between the above structural factors and shows that for 2-year bond all factors are statistically significant except $Surp$, while for the 10-year bond only disagreement and long-run risk factors are significant. Taken jointly, the structural factors are explaining a large proportion of the variance of expected returns, with an adjusted R-squared of 49% and 39% on 2 and 10-year bonds, respectively.

Specification (*vii*) reports estimates from multivariate regressions on the volatility factors discussed above. While the explanatory power of these factors is not large, the signs on $TVRP$ and $Jump$ are consistent with the theoretical predictions and empirical tests (on ex-post data) reported by Mueller, Vedolin, Sabtchevsky, and Whelan (2016) and Wright and Zhou (2009). Finally, (*viii*) shows a highly significant positive relationship between survey expectations and the Cochrane and Piazzesi (2005) return forecasting factor, in particular for the short-term bond. In terms of explanatory power this finding is particularly strong with approximately 1/2 and 1/4 of the variance on 2 and 10-year subjective excess returns explained by this combination of forward rates, and this is consistent with the large degree of spanning of survey expectations by term structure factors described in Section III. The macro factor instead is only significant for the long-term bond and with the positive sign dictated by economic theory.

In the context of the equity market, Greenwood and Schleifer (2014) find that several rational expectation models are negatively correlated with survey expectations of stock market returns. They interpret their result as clear evidence of a rejection of rational expectations models: "*We can reject this hypothesis with considerable confidence. This evidence is inconsistent with the view that expectations of stock market returns reflect the beliefs or requirements of a representative investor in a rational expectations model.*" On the other hand, we find significant positive correlation between proxies of expected excess returns obtained from some of the rational expectation models and expectations of bond excess returns erx_t. Moreover, while some specifications do worse than subjective expectations some models do remarkably well. This suggests that, at least in the context of bond markets, rational expectation models cannot be dismissed so quickly.

V. Forecast Errors

A. The dynamics of forecast errors

Figure 9 shows the time series of forecast errors, i.e. $fe^n_{i,t+1} = rx^n_{t+1} - erx^n_{i,t}$, for different quantiles of the cross-sectional distribution of expected bond returns, for the 2 and 10-year bond maturities. They are significantly positive in the early 90s, in 1995, in the early 2000s and in the run-up to the 2008/2009 financial crisis. On the other hand, they were consistently negative in 1988, 1993/1994 and 1999. Unconditionally, however, the distribution of $fe^n_{i,t+1}$ is largely skewed towards positive errors. Figure 10 shows the average forecast error for the different deciles of the cross-sectional distribution of expected bond returns, for the 2, 5 and 10-year bond maturities: all agents, except the most optimistic, make positive errors on average.

Is the distribution of forecast errors state-dependent? We split the sample in two parts to capture persistent periods of increasing and decreasing interest rates, respectively. We compute the exponential moving average of the monthly change in the one year yield over the previous 12 months. Considering the whole sample, there are 198 months in which the exponential moving average of the 1-year yield change is decreasing and 112 in which it is increasing. Figure 11 shows that during periods of increasing interest rates (which correspond to good states for the U.S. economy in this sample period) the distribution of forecast errors is symmetric around zero. An almost equal mass of agents commit positive and negative forecast errors. On the other hand, following periods of decreasing short term rates, all agents, including the most optimistic, are surprised by larger excess bond returns.[17] Note that we define optimists in terms of expected returns, which means that they are expecting lower bond yields.

Moreover, the consensus agent commits much larger forecast errors in absolute terms in bad times. This is consistent with the findings in Cieslak and Povala (2012) who analyze the survey forecast expectations of the fed fund rate and show that "*most pronounced errors are negative and typically occur during and after NBER recessions as forecasters largely fail in predicting the extent of subsequent monetary easing*". When the 1-year yield is increasing, $fe_{i,t+1}$ for the consensus forecaster are -0.50%, -0.68% and -0.37% for the 2, 5 and 10-year bond, respectively. In contrast, the mean of the forecast errors of the consensus forecaster in phases of decreasing interest rates are 1.17%, 4.09% and 5.38% for the 2, 5 and 10-year bond, respectively. These errors are extremely large considering that the corresponding average realized excess return in the same periods are -0.27%, -0.44% and 0.91%.

Is the distribution of forecast errors predictable? We run a simple regression of forecast errors $fe^n_{i,t+1}$ on observable variables at time t :

$$fe^n_{i,t+1} = \alpha^n_i + \beta^n_i X_t + \epsilon^n_{i,t+1}. \tag{8}$$

Regressions of this type have been used to test the *rational expectation hypothesis* (see e.g. Jitmaneeroj and Wood (2013)) since forecast errors should not be predictable using time t information if forecasters are rational. In the interest of parsimony and simplicity, we include in X_t the level, slope and curvature factors of the current term structure extracted from the three principal components of the yield curve at time t. Table XI reports results of regression (8) for all deciles of cross-sectional distribution of expected bond returns. We find that the slope of the term structure predicts $fe^n_{i,t+1}$ with a positive factor loading for $n = 10$ years, and the coefficient is significant for all deciles except the most optimistic one. The higher the slope, the larger the positive surprise in terms of bond excess returns. This is consistent with the previous finding. In bad times short term interest rates fall faster that long term bond yields, thus producing steeper yield curves. These are periods in which agents incorrectly forecast that long term yields will not drop much. This dynamics in the

[17]Results are robust to the choice of time periods for the moving average.

forecast errors is consistent with several explanations. First, agents underestimate the extent to which the Federal Reserve is willing to aggressively drop short term rates. Second, agents overestimate the willingness of the Federal Reserve to start normalizing short interest rates for a prolonged period of time. Third, agents have incorrectly assumed that the initial real economic shock was just temporary.

B. Economic interpretations

What is the source of the observed bias and state-dependence in forecast errors? In what follows, we discuss and compare some alternative explanations. A large literature in behavioral finance frequently argues that forecasters form irrational beliefs. Often this argument is tested in the context of extrapolative learning models. The substantial persistence in beliefs reported in the first part of the paper and the predictability of the forecasts errors is - prima facie - consistent with this conjecture. A second stream of the literature has studied rational agents who face informational rigidities. Finally, the observed dynamics and cross-section of forecast errors is potentially consistent with models in which forecasters have identical and complete information but asymmetric loss functions with heterogeneity in the degree of loss aversion, or with forecasters engaging in forecast smoothing for reputational considerations. Notice that all these alternative theoretical models are also consistent with the substantial persistence in beliefs.

B.1. Extrapolative behaviour

Greenwood and Schleifer (2014) show evidence of a high positive correlation between investor expectations about stock market returns and past realized returns. Malmendier and Nagel (2011) show that experience affects risk aversion and individuals who experienced large drops in asset prices behave as if more risk averse, thus requiring a larger risk premium. Both studies suggest the existence of extrapolative components in the formation of individual expectations. However, the first argues that the relationship between past returns and future expected returns is positive, while the second argues that the relation between past returns and expected future excess returns is negative.

We study these hypotheses in the context of our data by testing if past bond excess returns drive current subjective risk premia. We use our direct measure of expectation of bond returns and we run a regression of $erx_{i,t}^n$ on lagged realized bond excess returns:

$$erx_{i,t}^n = a_i^n + b_i^n rx_t^n + \epsilon_{i,t}^n. \tag{9}$$

As in the previous section, we run this regression for the consensus forecast, as well as for the different deciles of the cross-sectional distribution of forecasts. The slope coefficients b_i^n and the R-squared of the regression for all the deciles of the cross-sectional distribution of forecasts are represented in Figure 12. The b_i^n coefficient is insignificant across all percentiles for the 2 and 5-year bond, but it is negative and significant (except for the most optimistic decile) for the 10-year bond. The R-squares are very small (between 0 and 2%) but for the long-term bond they increase for the pessimists up to about 10%.

We extend regression (9) by controlling for additional variables included in the time-t information set of the agents:

$$erx_{i,t}^n = a_i^n + b_i^n rx_t^n + c_i^n X_t + \epsilon_{i,t}^n. \tag{10}$$

For brevity, we consider only the information about the current term structure, which has proven relatively successful in spanning expected bond excess returns. Just including the slope (the second principal component) of the term structure as a control, lagged realized excess returns are still significantly negatively correlated with bond risk premia for the 10-year bond, and the b_i^n become significantly negative also for the 5-year bond across all quantiles and for the 2-year bond between the 30^{th} and the 60^{th} percentiles.

Overall, we find weak but supportive evidence of extrapolative behaviour in the formation of expectations in bond markets. When we run a regression of expected returns (instead of expected *excess* returns) on realized returns we find that the slope coefficient is significantly positive and the R-squared is around 35%, but only for the short term bond. This is consistent with agents extrapolating that the trajectory of short term interest rates will continue in the same direction. Figure 13 shows the slope coefficients and the R-squares of this regression for all the deciles of the cross-sectional distribution of forecasts. When we run regression for bond risk premia, we find a negative relation between realized excess returns and expected long-term bond excess returns. The result is consistent with an increase in risk aversion after agents experience losses due to a drop in bond prices, as argued by Malmendier and Nagel (2011).

B.2. Information rigidities

A recent literature on rational expectations models with information frictions (see e.g. Mankiw and Reis (2002) and Woodford (2002)) emphasizes how information rigidities can account for otherwise puzzling empirical findings, and Coibion and Gorodnichenko (2015) rely on these theoretical models to guide their choice of the relevant regressors in a new set of tests of full-information rational expectations (FIRE). Namely, Coibion and Gorodnichenko (2015) show that in sticky-information models (Mankiw and Reis (2002)) and noisy-information models (Woodford (2002)) the cross-sectional average forecast error should be predictable by the average forecast revision, and their empirical results support the presence of significant information rigidities in the formation of expectation about macroeconomic variables such as inflation and GDP growth.

The baseline test can be performed by running a regression of the consensus forecast error on the forecast revisions:[18]

$$fe^n_{c,t+1} = a^n + b^n \Delta erx^n_{c,t} + \epsilon^n_{t+1}. \tag{11}$$

where $fe^n_{c,t+1}$ denotes the forecast error of the consensus and $\Delta erx^n_{c,t}$ is its forecast revision. In order to compute the forecast revision we need forecasts for the same future period at two consecutive times, i.e. forecasts at two different horizons. All our analysis so far focused on the expected excess returns at a fixed one year horizon, but these 1-year forecasts were constructed initially by combining the agents' forecasts for the average yield 4 and 5 quarters ahead. Now we obtain forecast revisions by going back to the original data and computing the difference between the 4 quarters ahead forecasts and the 5 quarters ahead forecasts three months before. Since the horizon of the forecasts changes slightly depending on the month we are considering, instead of using a single monthly series we have to consider three different quarterly series, corresponding to the first, second and third month in each quarter. For all three quarterly series, the estimation of regression (11), for $n = 2, 5, 10$ years, yields

[18] The original test in Coibion and Gorodnichenko (2015) holds for the cross-sectional *average*, while we consider the forecast error and forecast and forecast revision of the *median* forecaster to be consistent with previous results. However, the results are almost identical using the mean.

positive and significant intercepts, a^n, negative and significant slope coefficients b^n, and R-squares between about 5% and 14%, all increasing in absolute value with n. The estimates for the three series are very similar and Panel A of Table XII reports results for the first series (corresponding to the first month in each quarter), for which the horizon is closer to the 1-year horizon we considered so far.[19] Models of information frictions imply a positive relation between forecast updates and forecast errors (see Coibion and Gorodnichenko (2015)), which is inconsistent with our finding of a negative and significant slope coefficient in regression (11). Moreover, in models of information frictions, the predictability in forecast errors follows from the aggregation of forecasts across agents, even if no such predictability exists at the individual level, while we find a strong negative relation between forecast errors and forecast revisions also for the individual forecasters and across the different quantiles of the cross-section of forecasts.

We check the robustness of the results in Panel B of Table XII, which reports estimation results for an augmented version of regression (11), which includes the time-t level and slope of the term structure of interest rates as additional regressors:

$$fe^n_{t+1} = a^n + b^n_1 \Delta erx^n_{c,t} + b^n_2 Level_t + b^n_2 Slope_t + \epsilon^n_{t+1}. \tag{12}$$

The forecast revision remains significantly negative and the additional variables are not significant, but the explanatory power increases, which adjusted R-squares up to 25.8% for the 10-year bond. These results suggest that, in contrast to the case of macroeconomic variables reported in other papers, information rigidities do not play a role in the formation of expectation about excess bond returns.

B.3. Heterogeneity in loss aversion and forecast smoothing

Coibion and Gorodnichenko (2015) also derive testable implications for a number of potential alternative theoretical explanations for the state dependence of forecast errors. Our finding of a negative relation between forecast errors and forecast revisions in the cross-section of forecasts is potentially consistent with models in which forecasters have asymmetric loss func-

[19] The first month each quarter the agents forecast the average excess return realized over the 4^{th} quarter, i.e. the average of the excess returns 9, 10 and 11 months ahead. Therefore, the average horizon for our first quarterly series is equal to 10 months.

tions with heterogeneity in the degree of loss-aversion (Capistran and Timmermann (2009)), with forecasters having identical and complete information. An alternative explanation for predictable forecast errors that is consistent with their negative relation with forecast revision is that the forecasters engage in forecast smoothing for reputational considerations. Namely, forecasters may want to avoid drastic short-run changes in their forecasts. As well as heterogeneity in loss aversion, forecast smoothing could make forecast errors predictable also in the absence of information frictions.

Therefore, models with full information and rational expectations such as heterogeneity in loss aversion or forecast smoothing could potentially explain the state dependence of the forecast errors, even if this is probably not the only source of forecast error predictability in the data. It would be interesting to disentangle these two alternative theoretical explanations by investigating empirically additional model predictions. Preliminary results indicate that a simple model with asymmetries in the forecasters' cost of over- and under-predictions on the lines of Capistran and Timmermann (2009) is consistent with a number of our empirical results, such as the systematic bias in forecast errors and the persistence in the ranking of forecast errors across the different percentiles of the cross-sectional distribution.

VI. Conclusion

This paper studies the expectations of bond returns taken directly from survey data and compare them to standard measures of bond risk premia, in order to investigate the expectation formation process in fixed income markets. Our analysis reveals a number of interesting results. First, we find that individual risk premia are largely heterogeneous and the consensus does not subsume the information contained in the distribution of forecasts. We find a significant amount of persistence in agents beliefs on bond excess returns and in the degree of optimism/pessimism relative to consensus. Forecasts about macroeconomic fundamentals are only weakly correlated with bond return forecasts, except at the extreme of the distribution. This suggests that the drivers of beliefs about the yield curve and macro fundamentals are significantly different.

Second, we find strong evidence of time-varying expections of bond risk premia. Moreover, expectations of bond risk premia are largely spanned by the current term structure of bonds prices and the degree of spanning is substantially larger than when using sample averages of future excess returns as proxies of bond risk premia. Even more importantly, the degree of spanning greatly differs in the cross-section of agents beliefs. Indeed, there is a strong positive relation between spanning and forecasting accuracy in the cross-section: the beliefs of agents who have been more accurate in their forecasts in the preceeding months are more spanned by the term structure of bond yields. This is consistent with the predictions of general equilibrium heterogenous agents models with speculative trading and no frictions. In these models, the pricing kernel is a stochastic weighted average of agents beliefs, where relative weights depends on the wealth accumulation generated by belief-based trading. When optimists happen to be more accurate, their beliefs are more than three times more spanned than pessimists'.

Third, the predictive power of the consensus forecast for future realized excess bond returns is low (2%-5%) but the relation between survey expectation and realized returns is positive across the distribution of agents' beliefs. The accuracy of survey forecasts is much higher in the right tail of the cross-sectional distribution of subjective bond risk premia. In the last 27 years of monthly data, optimistic agents have significantly lower RMSE of forecasts, and this is due especially to their better performance during bad times, identified as periods of decreasing short-term interest rates. Compared to standard statistical models, surveys are less accurate in-sample but more accurate out-of-sample at all maturities if we focus on the optimistic forecasters.

In fact, we provide evidence that survey-implied bond risk premia are highly competitive in forecasting future realised excess returns relative to some popular reduced form models. Therefore, surveys can indeed be used to build reliable measures of bond risk premia, and thus also to evaluate models in real time, avoiding the issues related to in-sample and out-of-sample model fitting. In fact, if surveys provide a good measure of bond risk premia, the (absolute) correlation between this measure and each factor would provide us with a simple and fully nonparametric way to separate good models from the bad, as long as the sign of the correlation is consistent with the economic intuition behind the factor.

Indeed, contrary to the findings of Greenwood and Schleifer (2014) for the stock market, we find significant positive correlation between proxies of expected excess returns obtained from some of the rational expectation models and subjective expectations of bond excess returns, suggesting that, at least in the context of bond markets, rational expectation models cannot be dismissed so quickly.

Finally, we propose and test a number of alternative theoretical explanations for the observed state-dependence in the survey forecast errors. Information rigidities do not seem to play an important role in the formation of expectation about excess bond returns, while models with full information, rational expectations and heterogeneity in the degree loss aversion is potentially consistent with a number of our empirical results, such as the systematic bias in forecast errors and the persistence in the ranking of forecast errors across the different percentiles of the cross-sectional distribution.

References

Aioli, M., C. Capistran, and A. Timmermann, 2011, *The Oxford handbook of economic forecasting* . chap. Forecast combinations (Oxford: Oxford University Press).

Akima, Hiroshi, 1970, A new method of interpolation and smooth curve fitting based on local procedures, *J. ACM* 17, 589–602.

Alchian, A., 1950, Uncertainty, evolution and economic theory, *Journal of Political Economy* 58, 211–221.

Ang, A., G. Bekaert, and M. Wei, 2007, Do macro variables, asset markets, or surveys forecast inflation better?, *Journal of Monetary Economics* 54, 1163–1212.

Bansal, R., and I. Shaliastovich, 2013, A long-run risks explanation of predictability puzzles in bond and currency markets, *Review of Financial Studies* 26, 1–33.

Bansal, R., and A. Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *The Journal of Finance* 59, 1481–1509.

Basak, S., 2005, Asset pricing with heterogeneous beliefs, *Journal of Banking and Finance,* 29, 2849–2881.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Buraschi, A., and A. Jiltsov, 2006, Model uncertainty and option markets with heterogeneous beliefs, *The Journal of Finance* 61, 2841–2897.

———, 2007, Habit formation and macroeconomic models of the term structure of interest rates, *Journal of Finance* 62, 3009 – 3063.

Buraschi, A., F. Trojani, and A. Vedolin, 2014, Economic Uncertainty, Disagreement, and Credit Markets, *Management Science* 60, 1281–1296.

Buraschi, A., and P. Whelan, 2012, Term structure models and differences in belief, *Imperial College, Working Paper.*

Buraschi, Andrea, and Paul Whelan, 2016, Hedging, speculation, and interest rates, *Working paper.*

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* 107, 205–251.

Campbell, John Y, and Robert J Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495–514.

Capistran, C., and A. Timmermann, 2009, Disagreement and biases in inflation expectations, *Journal of Money, Credit and Banking* 41, 365–396.

Chen, H., S. Joslin, and N. Tran, 2012, Rare disasters and risk sharing with heterogeneous beliefs, *Review of Financial Studies* 25, 2189–2224.

Cieslak, A., and P. Povala, 2012, Expecting the Fed, *Working paper, Kellogg School of Management.*

Cochrane, J.H., and M. Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Coibion, O., and Y. Gorodnichenko, 2015, Information rigidity and the expectation formation process: A simple framework and new facts, *American Economic Review* 105, 2644–2678.

Dai, Q., and K.J. Singleton, 2000, Specification analysis of term structure of interest rates, *Journal of Finance* 55, p. 1943–78.

Dai, Qiang, and Kenneth Singleton, 2003, Term structure dynamics in theory and reality, *Review of Financial Studies* 16, 631–678.

D'Amico, S., and A. Orphanides, 2008, Uncertainty and Disagreement in Eeconomic Forecasting, Discussion paper, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.

Ehling, Paul, Michael Gallmeyer, Christian Heyerdahl-Larsen, and Philipp Illeditsch, 2013, Disagreement about inflation and the yield curve, *Working paper. University of Pennsylvania.*

Fama, Eugene F, and Robert R Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680–92.

Fontaine, Jean-Sébastien, and René Garcia, 2012, Bond liquidity premia, *Review of Financial Studies* 25, 1207–1254.

Giordani, P., and P. Soderlind, 2003, Inflation Forecast Uncertainty, *European Economic Review* 47, 1037–1059.

Goyal, A., and I. Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Greenwood, R., and A. Schleifer, 2014, Expectations of returns and expected returns, *Review of Financial Studies* 27, 714–746.

Greenwood, Robin, and Dimitri Vayanos, 2014, Bond supply and excess bond returns, *Review of Financial Studies* 27, 663–713.

Gürkaynak, Sack, and Wright, 2006, The u.s. treasury yield curve: 1961 to the present, *Federal Reserve Board Working Paper Series*.

Hong, H., D. Sraer, and J. Yu, 2013, Reaching for maturity, Discussion paper, Princeton University.

Jitmaneeroj, B., and A. Wood, 2013, The expectations hypothesis: New hope or illusory support?, *Journal of Banking & Finance* 37, 1084–1092.

Joslin, S., K. J. Singleton, and H. Zhu, 2011, A new perspective on gaussian dynamic term structure models, *Review of Financial Studies* 24, 926–970.

Jouini, E., and C. Napp, 2006, Heterogeneous beliefs and asset pricing in discrete time: An analysis of pessimism and doubt, *Journal of Economic Dynamics and Control* 30, 1233–1260.

Koijen, R., M. Schmeling, and E. Vrugt, 2015, Survey expectations of returns and asset pricing puzzles, Discussion paper, London Business School, Cass Business School and VU University Amsterdam.

Le, A., and K.J. Singleton, 2013, The structure of risks in equilibrium affine models of bond yields, Discussion paper, Working Paper, UNC.

Litterman, R., and J. Scheinkman, 1991, Common factors affecting bond returns, *The Journal of Fixed Income* 1, 54–61.

Ludvigson, Sydney C., and Serena Ng, 2009, Macro factors in bond risk premia, *Review of Financial Studies* 22, 5027–5067.

Malkhozov, Aytek, Philippe Mueller, Andrea Vedolin, and Gyuri Venter, 2016, Mortgage risk and the yield curve*, *Review of Financial Studies* p. hhw003.

Malmendier, U., and S. Nagel, 2011, Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?, *Quarterly Journal of Economics* 126, 373–416.

Mankiw, N.G., and R. Reis, 2002, Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve, *Quarterly Journal of Economics* 117, 1295–1328.

Mueller, P, A Vedolin, P Sabtchevsky, and P Whelan, 2016, Variance risk premia on stocks and bonds, *Working Paper*.

Patton, A. J., and A. Timmermann, 2010, Why do Forecasters Disagree? Lessons from the Term Structure of Cross-Sectional Dispersion, *Journal of Monetary Economics* 57, 803–820.

Piatti, Ilaria, 2014, Heterogeneous beliefs about rare event risk in the lucas orchard, *Working Paper, Oxford University.*

Piazzesi, M., 2005, Bond yields and the federal reserve, *Journal of Political Economy* 113, 311–344.

———, J. Salomao, and M. Schneider, 2015, Trend and cycle in bond premia, Discussion paper, Stanford, NBER and University of Minnesota.

Vayanos, Dimitri, and Jean-Luc Vila, 2009, A preferred-habitat model of the term structure of interest rates, Discussion paper, National Bureau of Economic Research.

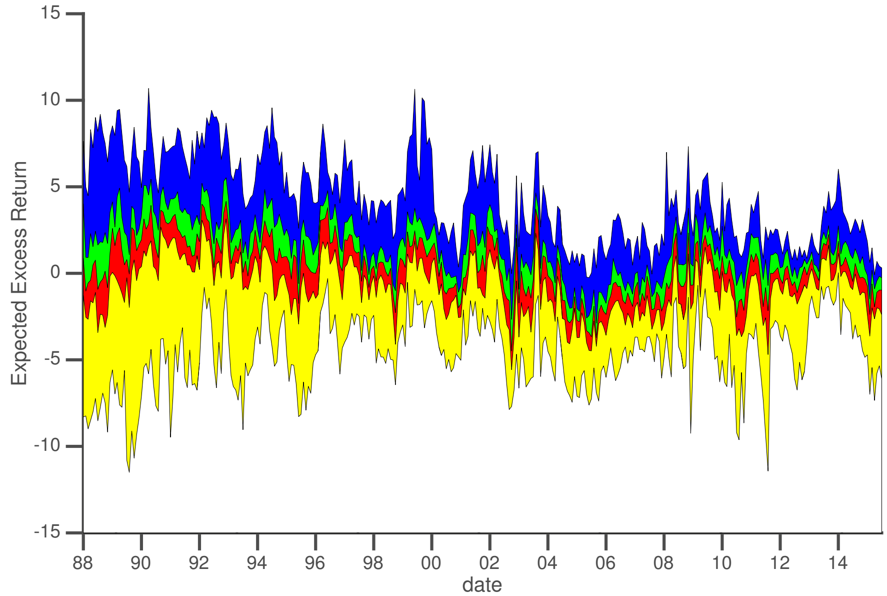
Wachter, J.A., 2006, A consumption-based model of the term structure of interest rates, *Journal of Financial Economics* 79, 365–399.

Woodford, M., 2002, *Knowledge, Information, and Expectations in Modern MacroeconMacro: In Honor of Edmund S. Phelps* . chap. Imperfect Common Knowledge and the Effects of Monetary Policy (Princeton University Press).

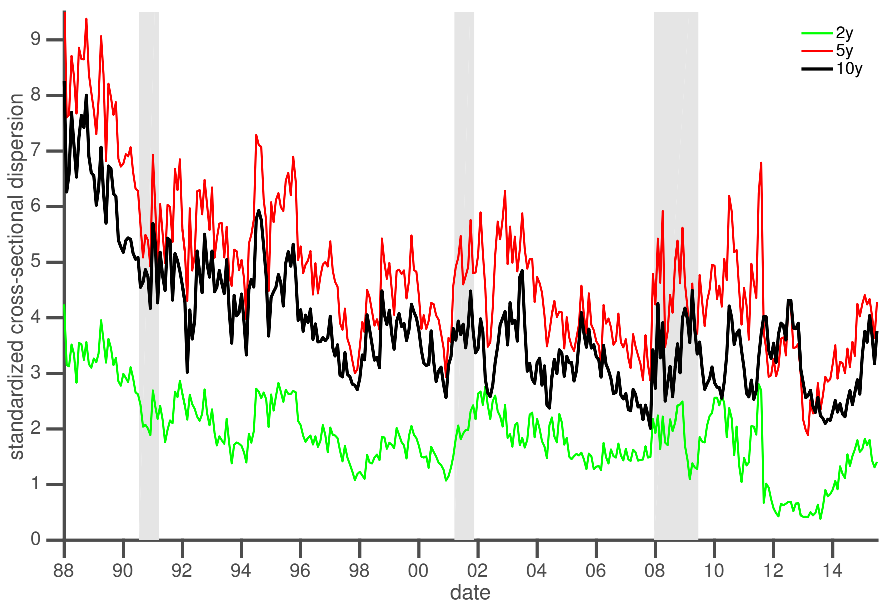
Wright, Jonathan H, and Hao Zhou, 2009, Bond risk premia and realized jump risk, *Journal of Banking & Finance* 33, 2333–2345.

Xiong, W., and H. Yan, 2010, Heterogeneous expectations and bond markets, *Review of Financial Studies* 23, 1433–1466.

VII. Figures



(a)



(b)

Figure 1. Cross-Sectional Heterogeneity

The top panel plots the Min, Q1, median, Q3 and max of the cross-sectional distribution of EBR for 5-year maturity bonds. The bottom panel plots the cross-sectional standard deviation of EBR standardized by the full-sample mean EBR. Shaded areas denote NBER recessions.

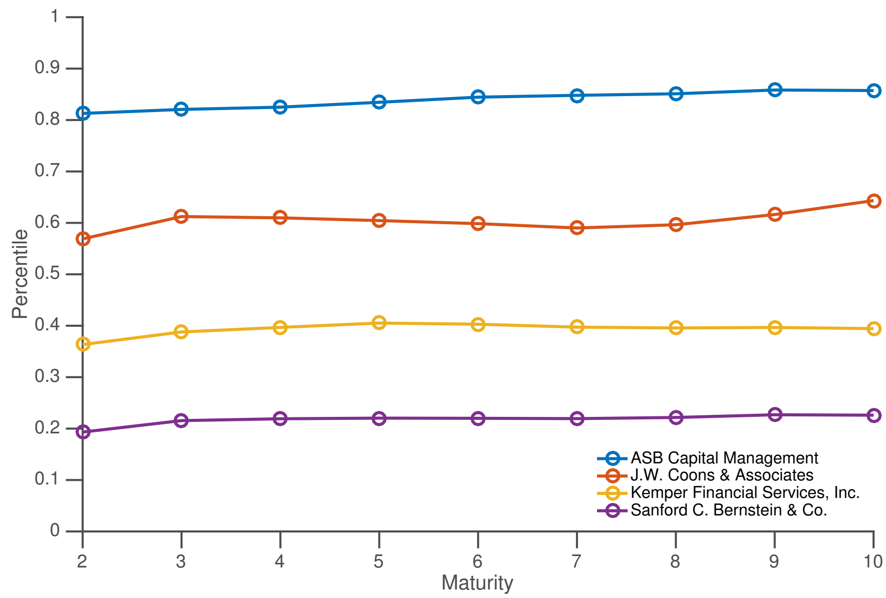
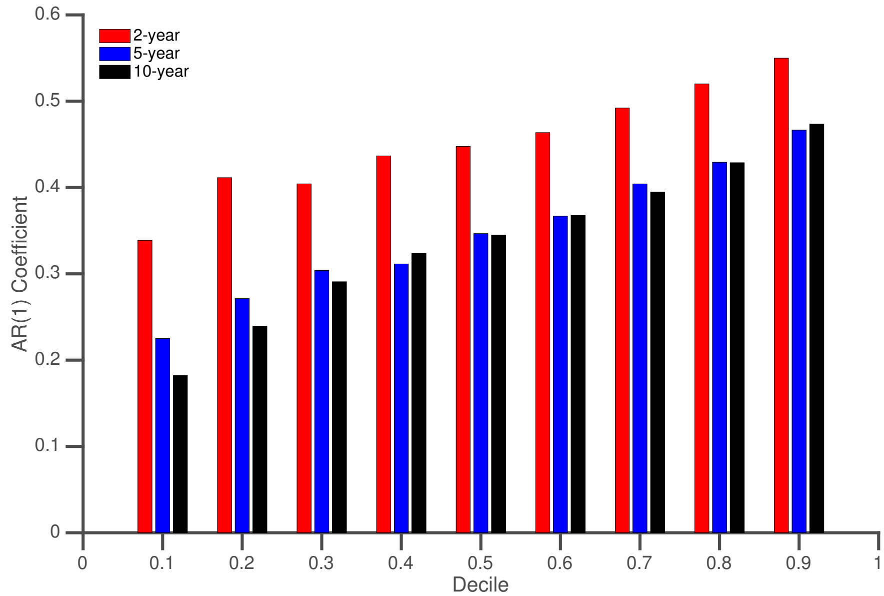
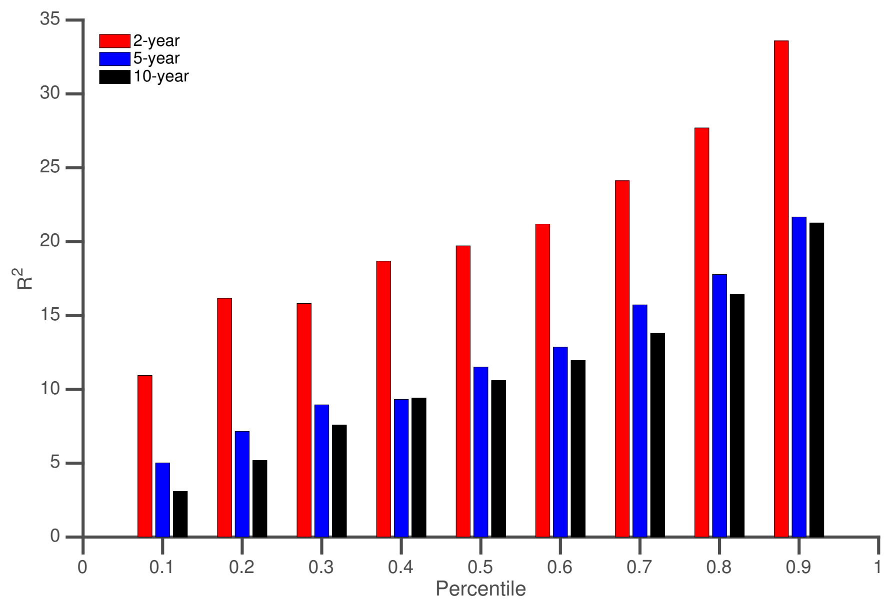


Figure 2. Selected Forecasters' Average Positions

Average position in the cross-sectional distribution of forecasters of four selected forecasters, for bond maturities between 2 and 10 years.



(a) ρ_i^n

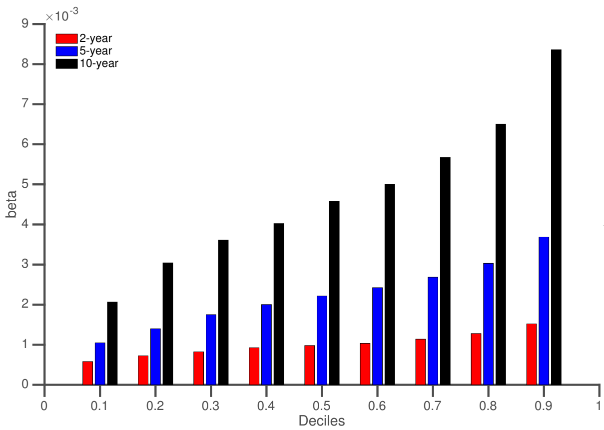


(b) R^2

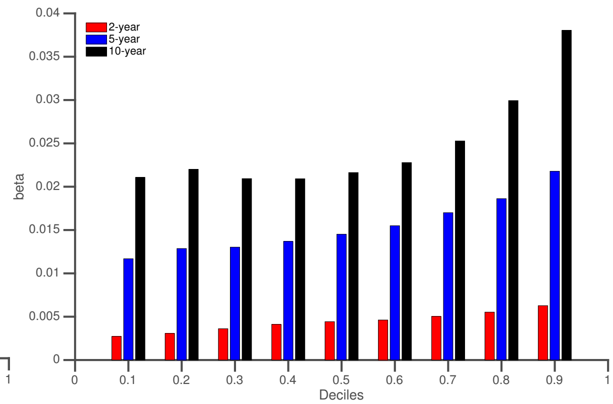
Figure 3. Autocorrelation Functions

Autocorrelation coefficients (top panel) and R^2 (bottom panel) for $n = 2, 5, 10$-year expected excess bond returns for percentile i of the cross-sectional distribution of expectations

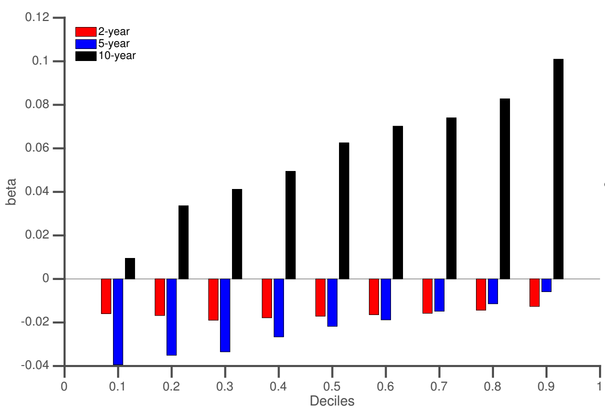
$$erx_{i,t+1}^n = \alpha_i^n + \rho_i^n erx_{i,t}^n + \epsilon_{i,t+1}^n.$$



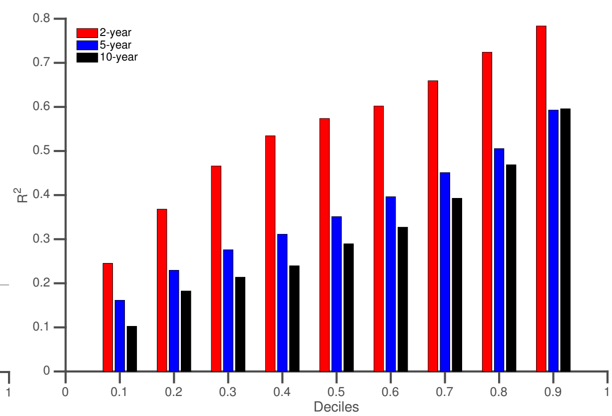
(a) $\beta^n_{i,PC1}$



(b) $\beta^n_{i,PC2}$



(c) $\beta^n_{i,PC3}$

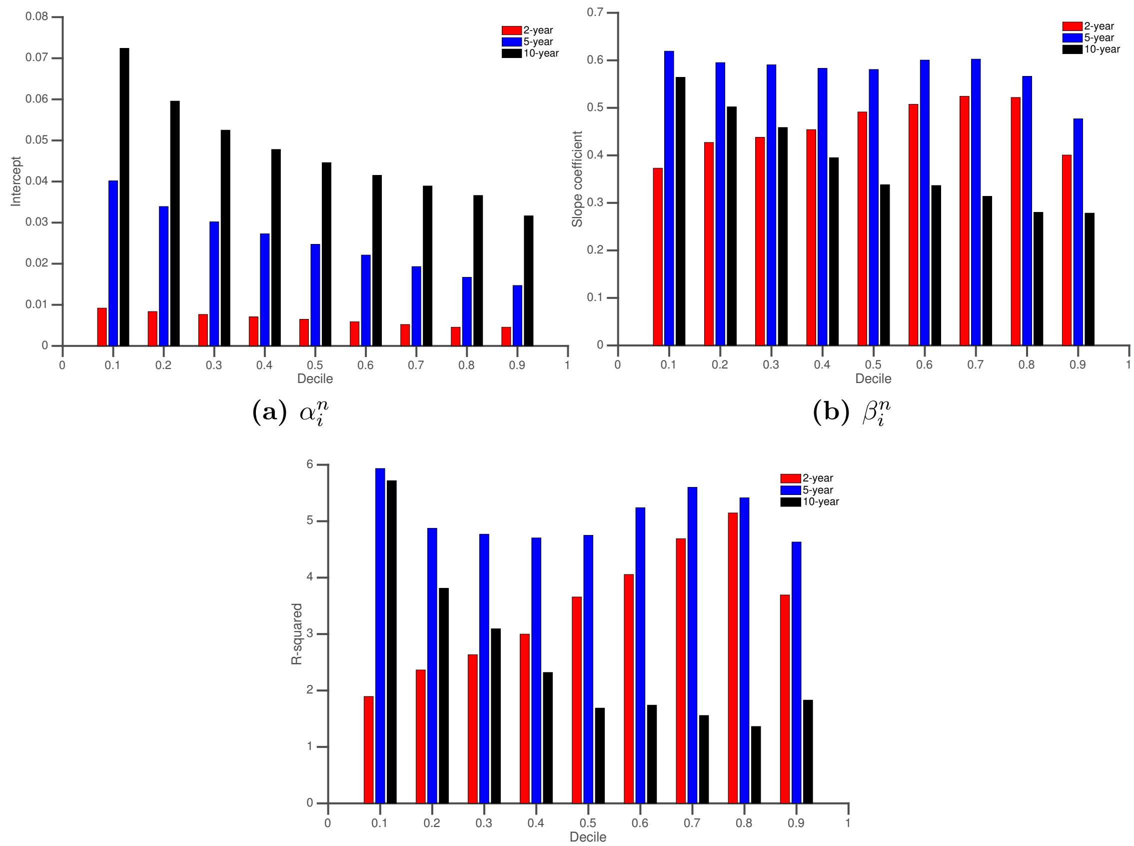


(d) R^2

Figure 4. EBR on Level, Slope, and Curvature

Betas and adjusted R^2 from regressions of the cross-sectional percentiles of EBRs on level (PC1), slope (PC2) and curvature (PC3) of the yield curve:

$$erx^n_{i,t} = \beta^n_{i,0} + \beta^n_{i,1} Level_t + \beta^n_{i,2} Slope_t + \beta^n_{i,3} Curv_t + \epsilon^n_{i,t}.$$



(a) α_i^n

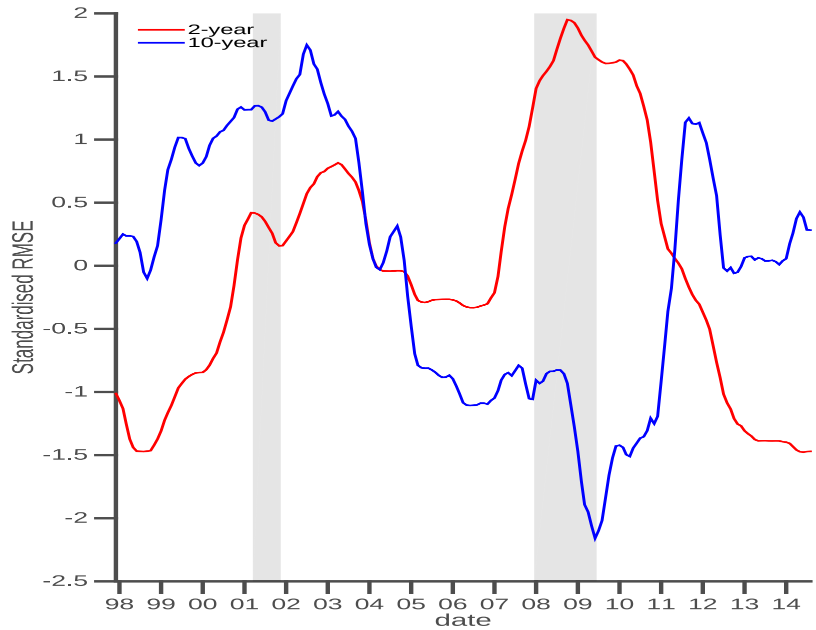
(b) β_i^n

(c) R^2

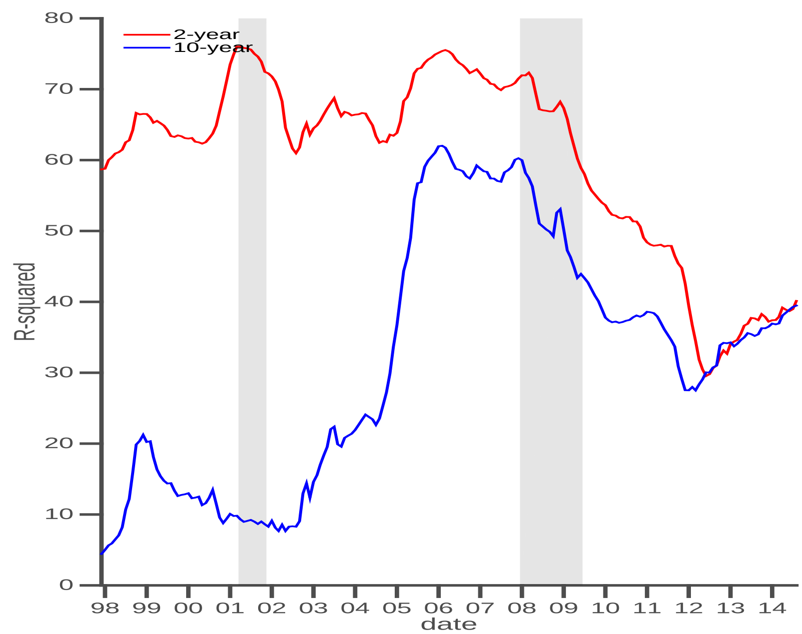
Figure 5. Cross-Section of Predictive Regressions

Estimated regression coefficients and adjusted R^2 of regressions of the realized excess n-year bond returns, for $n = 2, 5$ and 10, on the expected excess bond returns for percentile i of the cross-sectional distribution of expectations:

$$rx_{t+1}^n = \alpha_i^n + \beta_i^n erx_{i,t}^n + \epsilon_{i,t+1}^n.$$



(a)



(b)

Figure 6.

Rolling standardised *RMSE*s and spanning regression R-squared of the consensus
The top panel shows the time series of estimated *RMSE*s of predictive regressions of consensus expected excess bond returns for realized excess bond return, obtained from rolling windows of 120 months and bond maturities of 2 and 10 years. The bottom panel shows the R-squared of regressions of consensus EBR on the three principal components of the current term structure of interest rates, using rolling windows of 120 months. The x-axis shows the time in which the forecast (and the error) is realized in the top panel and the time of the forecast in the bottom panel. Shaded areas denote NBER recessions.

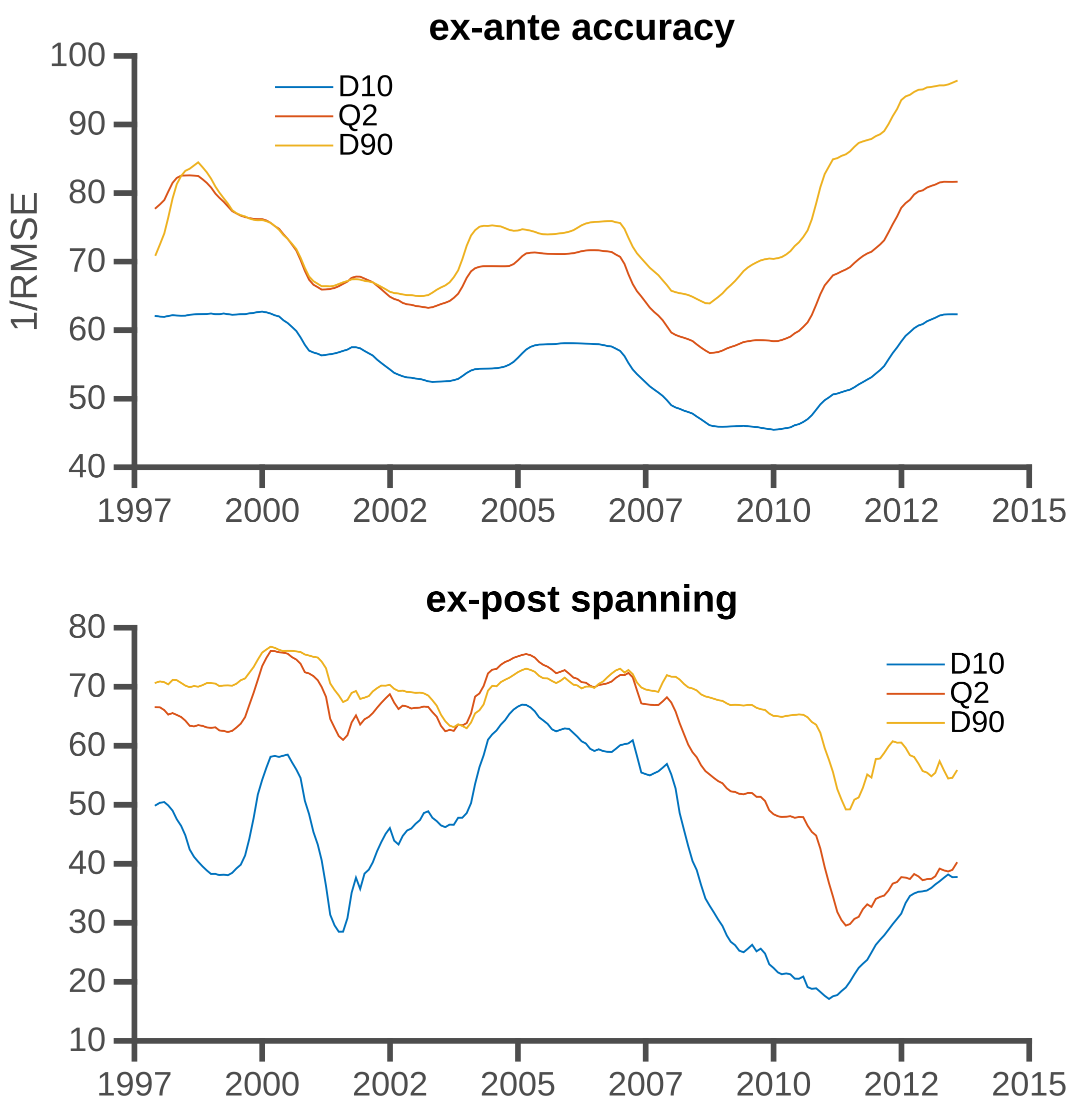


Figure 7. Time Series of Spanning and Forecast Accuracy

Time series of forecast accuracy (upper panel) for the 10^{th} (D10), 50^{th} (Q2) and 90^{th} (D90) percentile of the cross-sectional distribution of 2-year expected excess bond returns, and their ex-post spanning (bottom panel). Spanning is defined as the R-squared of regressions of EBR on the three principal components of the current term structure of interest rates, using rolling windows of 120 months. The forecast accuracy is defined as the inverse of the RMSE of the EBR, over the same rolling windows, but with a lag of 1 year, which corresponds to the horizon of the forecast.

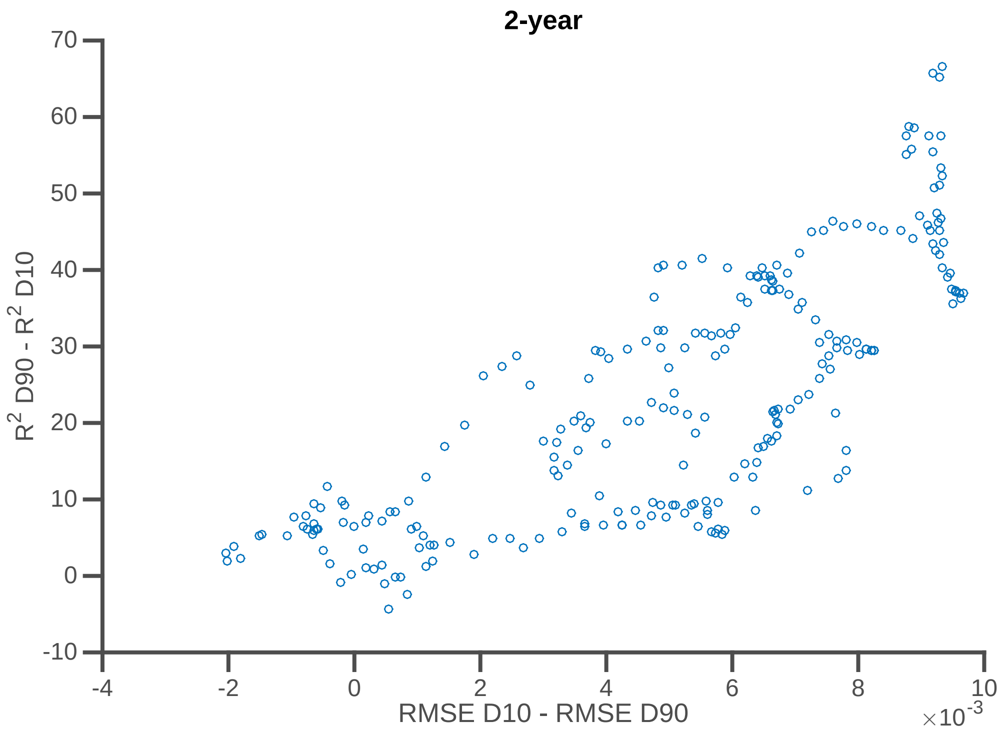
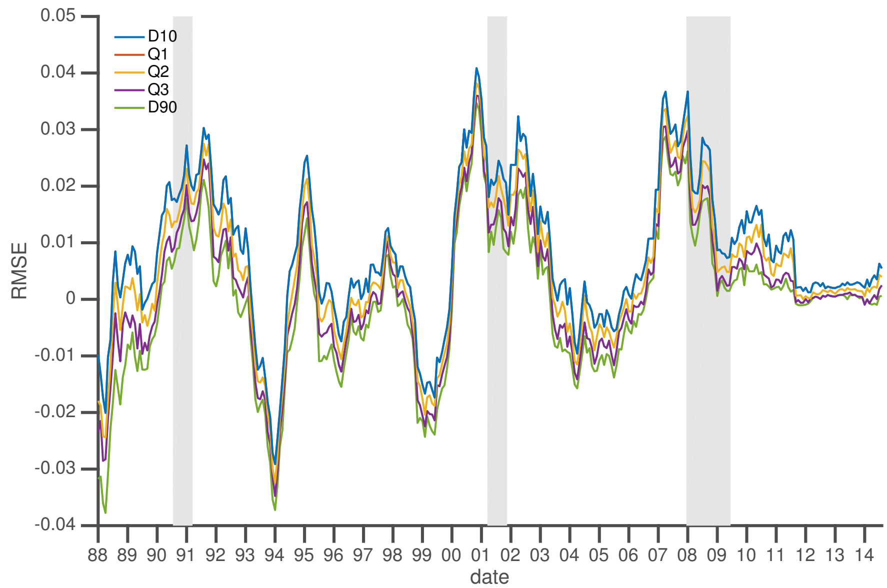
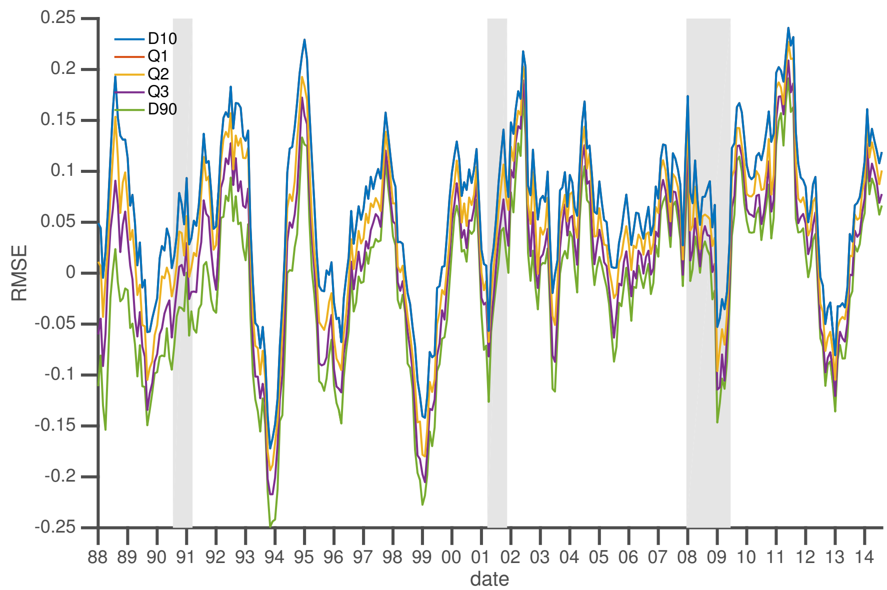


Figure 8. Relative Spanning vs Forecast Accuracy
Scatter plot of the relative spanning of optimist and pessimists against their ex-ante relative accuracy, for the 2-year bond excess returns. The relative spanning is defined as the difference in the R-squared of regressions of EBR on the three principal component of the term structure for the 90^{th} and 10^{th} percentile, using rolling windows of 60 months. The relative accuracy is defined as the difference in RMSE of the 10^{th} and 90^{th} percentile, over the same rolling windows, but with a lag of 1 year, which corresponds to the horizon of the forecast.



(a)



(b)

Figure 9. Time Series of Forecast Errors

Time series of forecast errors in excess bond returns, implied by different quantiles of the cross-sectional distribution of survey forecasts, for bond maturities of 2, 5 and 10 years. Shaded areas denote NBER recessions.

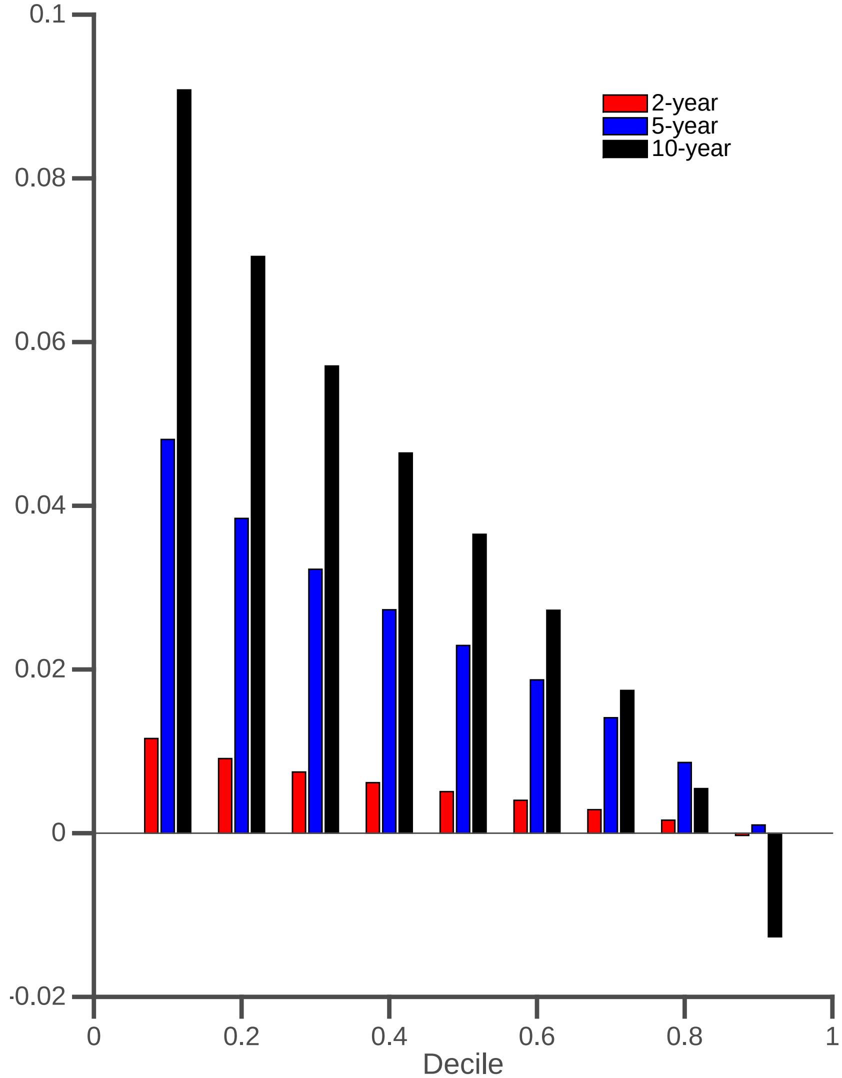


Figure 10. Cross-Section of Average Forecast Errors
Average forecast errors, $fe^n_{i,t+1} = rx^n_{t+1} - erx^n_{i,t}$, of different deciles i of the cross-sectional distribution of expected excess returns, for bond maturities of 2,5, and 10 years.

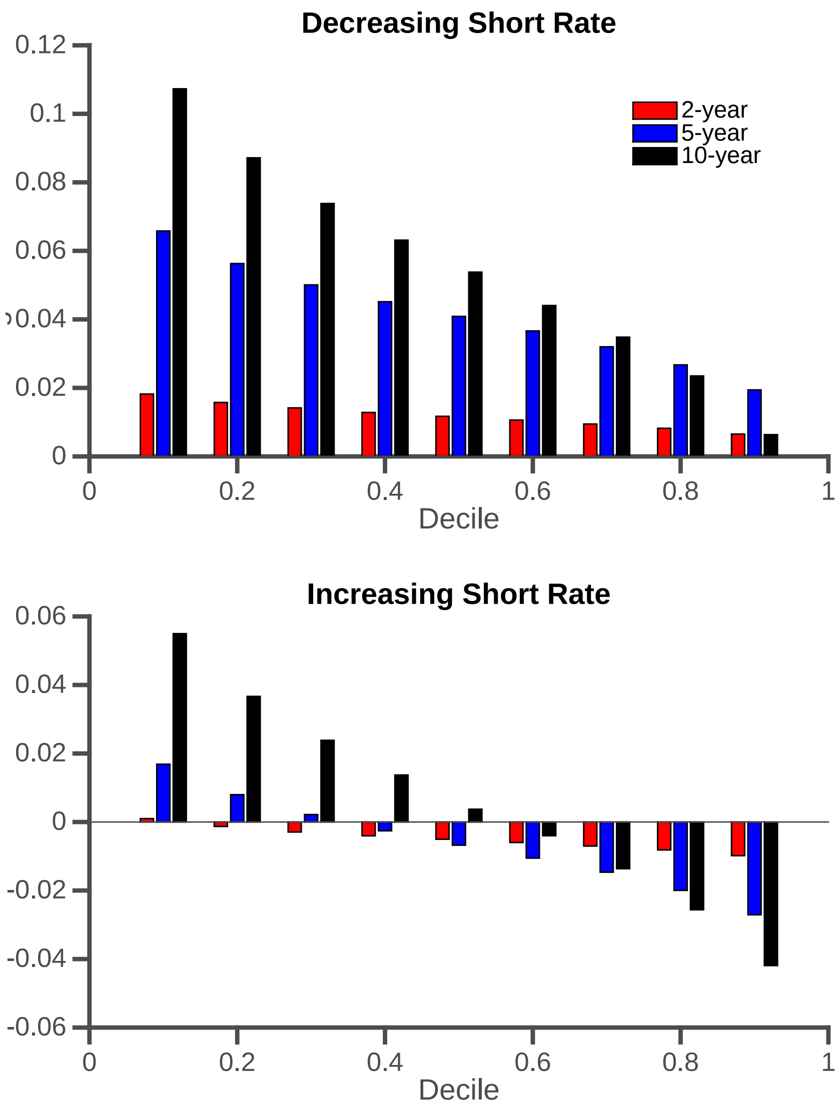


Figure 11. Cross-Section of Average Forecast Errors in Good and Bad Times
Average forecast errors of different deciles of the cross-sectional distribution of expected excess returns, for bond maturities of 2,5, and 10 years, in periods characterized by decreasing (top panel) and increasing (bottom panel) 1-year yield. We consider the 1-year rate to be decreasing (increasing) if at the date of the forecast, the exponential moving average of the monthly changes in the 1-year yield over the past 12 months is negative (positive).

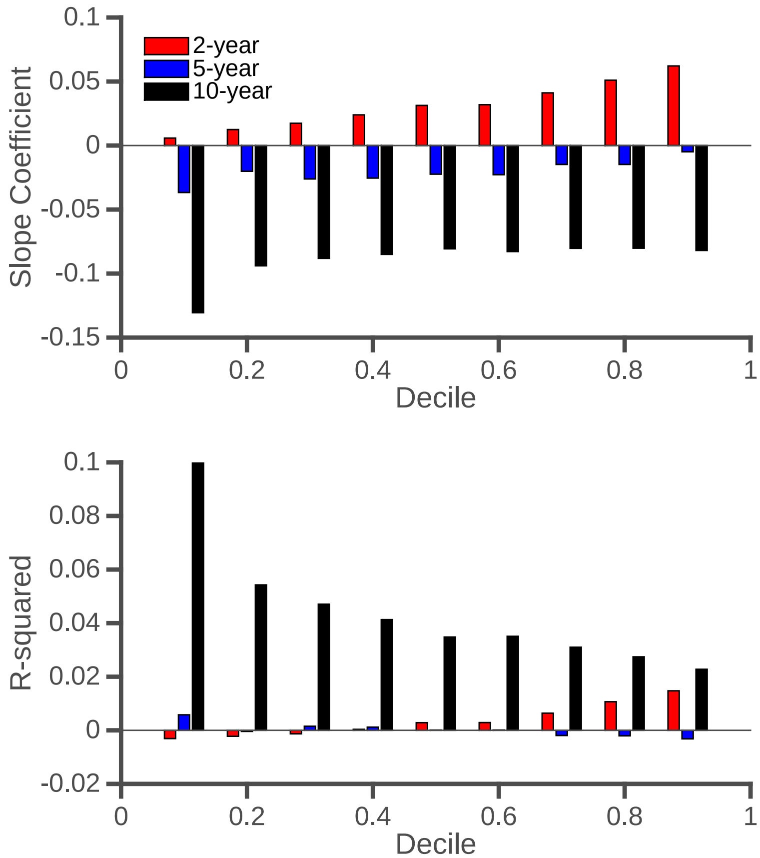


Figure 12. Cross-Section of Extrapolation in Excess Returns
Slope coefficient (top panel) and R-squared (bottom panel) of regression of expected excess returns on lagged realized excess returns, for different percentiles of the cross-sectional distribution of expected excess returns and bond maturities of 2, 5 and 10 years.

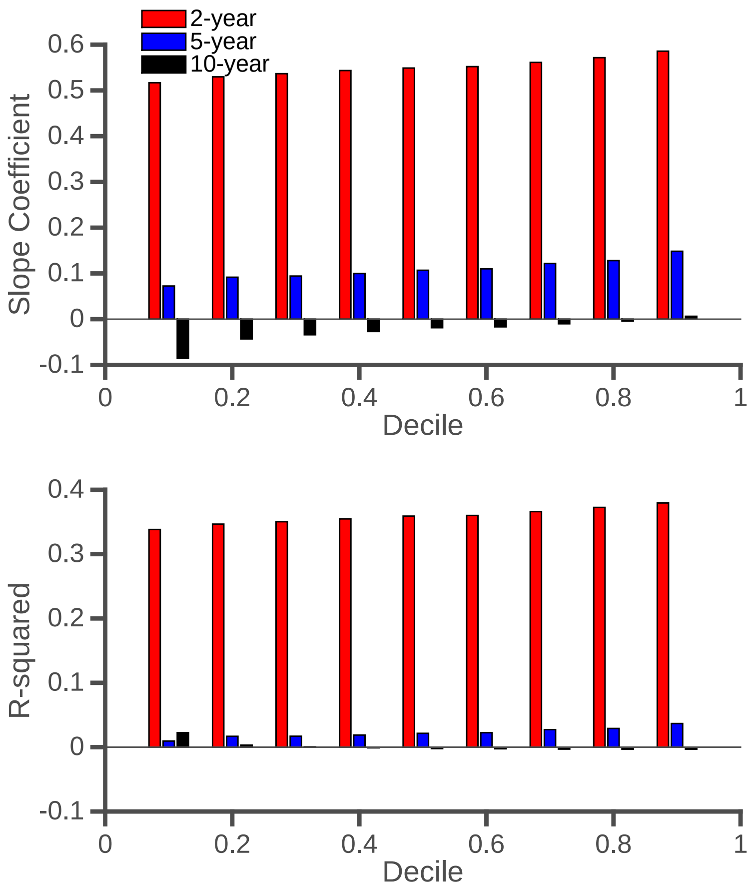


Figure 13. Cross-Section of Extrapolation in Returns
Slope coefficient (top panel) and R-squared (bottom panel) of regression of expected returns on lagged realized returns, for different percentiles of the cross-sectional distribution of expected returns and bond maturities of 2, 5 and 10 years.

VIII. Tables

Table I. Summary Statistics

Summary statistics of the first (Q1), second (Q2) and third (Q3) quartiles of the distribution of subjective expected excess bond returns, for maturities of 2, 5 and 10 years, and forecast horizon of 1 year. Sample period is January 1988 to July 2015 (331 observations).

Q1	**2 Year**	**5 Year**	**10 Year**
Mean	−0.06%	−0.83%	−1.56%
Std Dev	0.49%	1.59%	3.21%
Min	−1.38%	−5.57%	−10.14%
Max	1.18%	3.14%	10.15%
Skew	−0.0776	−0.1040	0.0153
Kurtosis	2.4555	2.5930	3.0395
AR(1)	0.8036	0.7486	0.7351
Q2	**2 Year**	**5 Year**	**10 Year**
Mean	0.26%	0.40%	1.08%
Std Dev	0.52%	1.61%	3.32%
Min	−1.08%	−4.48%	−7.67%
Max	1.65%	4.16%	11.97%
Skew	0.0369	−0.0770	0.0036
Kurtosis	2.3396	2.6674	2.8716
AR(1)	0.8277	0.7585	0.7570
Q3	**2 Year**	**5 Year**	**10 Year**
Mean	0.53%	1.54%	3.55%
Std Dev	0.55%	1.74%	3.55%
Min	−0.79%	−2.88%	−5.15%
Max	1.95%	5.52%	14.53%
Skew	0.1790	0.0015	0.0659
Kurtosis	2.2051	2.4911	2.5866
AR(1)	0.8568	0.7937	0.7827

Table II. Transition Probabilities Returns

This table presents the probability of a forecaster transitioning from a given quartile of the cross-sectional distribution of forecasts to another quartile in the following month, for bond maturities of 2 and 10 years.

	2-year bond					10-year bond			
	Q1	**Q2**	**Q3**	**Q4**		**Q1**	**Q2**	**Q3**	**Q4**
Q1	75%	19%	5%	2%	**Q1**	73%	19%	5%	3%
Q2	21%	50%	23%	6%	**Q2**	21%	50%	23%	5%
Q3	5%	23%	54%	19%	**Q3**	5%	23%	52%	20%
Q4	1%	6%	25%	67%	**Q4**	2%	6%	23%	70%

Table III. Transition Probabilities Macro

This table presents the probability of a forecaster transitioning from a given quartile of the cross-sectional distribution of GDP (left) and CPI (right) forecasts to another quartile in the following month.

	GDP					CPI			
	Q1	**Q2**	**Q3**	**Q4**		**Q1**	**Q2**	**Q3**	**Q4**
Q1	71%	20%	6%	3%	**Q1**	76%	17%	4%	2%
Q2	20%	51%	22%	7%	**Q2**	17%	59%	20%	4%
Q3	7%	21%	52%	20%	**Q3**	4%	20%	58%	18%
Q4	3%	6%	17%	74%	**Q4**	1%	5%	15%	79%

Table IV. Conditional Probabilities Returns vs Macro

This table presents the probability of a forecaster being in a given quartile of the cross-sectional distribution of Macro forecasts (GDP in top panels and CPI in the bottom panels), given that the forecaster is in a particular quartile of the cross-sectional distribution of EBR forecasts, for bond maturities of 2 (left panels) and 10 years (right panels).

		2-year bond					10-year bond			
		Q1	**Q2**	**Q3**	**Q4**		**Q1**	**Q2**	**Q3**	**Q4**
GDP	**Q1**	20%	21%	29%	30%	**Q1**	20%	23%	27%	30%
	Q2	22%	28%	28%	22%	**Q2**	23%	26%	27%	24%
	Q3	26%	28%	28%	19%	**Q3**	25%	27%	27%	22%
	Q4	37%	24%	23%	16%	**Q4**	36%	22%	23%	18%
		Q1	**Q2**	**Q3**	**Q4**		**Q1**	**Q2**	**Q3**	**Q4**
CPI	**Q1**	13%	22%	29%	35%	**Q1**	12%	20%	29%	39%
	Q2	21%	27%	30%	22%	**Q2**	21%	26%	29%	24%
	Q3	29%	28%	25%	17%	**Q3**	27%	30%	27%	17%
	Q4	41%	23%	22%	13%	**Q4**	42%	23%	19%	16%

Table V. Autoregressive Regression

Regressions on the consensus (Q2) subjective excess returns of 2, 5, and 10-year zero-coupon bonds on a constant and their own lag at the 1-year horizon. t-statistics, reported in parentheses below the point estimates, are Newey-West corrected. The adjusted R-squared of the regressions is also reported.

Maturity	**const**	$\mathbf{AR(1)}$	$\mathbf{R^2}$
2-year	0.00	0.46	20.44%
	(1.85)	(4.14)	
5-year	0.00	0.35	11.45%
	(1.21)	(2.92)	
10-year	0.01	0.36	11.30%
	(1.31)	(3.82)	

Table VI. Expected Bond Risk Premia and Term Structure Factors

Regressions of the median 1-year subjective excess return (Panel A) and realized 1-year excess return (Panel B) for 2, 5 and 10-year zero-coupon bonds on the level (PC1), slope (PC2), and curvature (PC3) term structure factors. t-statistics, reported in parentheses below the point estimates, are Newey-West corrected. The adjusted R-squared of the regressions is also reported. The sample period is from January 1988 to July 2015.

Maturity	**Level**	**Slope**	**Curv**	$\mathbf{R}^2$
Panel A: Subjective Excess Returns				
2 years	0.59	0.39	−0.29	57.47%
	(5.59)	(4.48)	(−4.02)	
5 years	0.41	0.41	−0.13	34.36%
	(3.32)	(4.38)	(−1.25)	
10 years	0.38	0.29	0.07	22.44%
	(3.11)	(2.00)	(0.51)	
Panel B: Realized Excess Returns				
2 years	0.24	0.13	−0.05	6.97%
	(2.02)	(0.84)	(−0.33)	
5 years	0.11	0.30	−0.08	9.82%
	(1.17)	(1.98)	(−0.62)	
10 years	0.06	0.44	−0.05	19.11%
	(0.61)	(3.63)	(−0.46)	

Table VII. Predictive Regression

Regressions of the realized excess bond return on the consensus (Q2) expected excess bond returns, for maturities of 2, 5 and 10 years: $rx^n_{t+1} = \alpha^n + \beta^n erx^n_{c,t} + \epsilon^n_{t+1}$. t-statistics, reported in parentheses below the point estimates, are Newey-West corrected. The adjusted R-squared of the regressions is also reported. The sample period is from January 1988 to July 2015.

Maturity	**const**	**EBR**	$\mathbf{R^2}$
2y	0.007	0.488	3.67%
	(3.672)	(1.618)	
5y	0.025	0.581	4.74%
	(4.878)	(2.318)	
10y	0.045	0.343	1.78%
	(4.865)	(1.639)	

Table VIII. RMSE: Surveys vs Models

This table reports the in-sample RMSE of surveys (Panel A) for percentiles $i = 0.10, \ldots, 0.90$ and statistical models (Panel B) for bond maturities of 2, 5 and 10 years. The statistical models we consider and from Ludvigson and Ng (2009) and Cochrane and Piazzesi (2005). Panels C and D report the out-of-sample RMSE of these models relative to the RMSE of the consensus and optimistic forecaster, respectively, over the same period. The sample period for the in-sample forecasts is from January 1988 to June 2014. The out-of-sample period starts in January 1998.

	n = 2	**n = 5**	**n = 10**
Panel A: $RMSE_i^n(Surv)$			
0.10	1.7432	6.3415	11.7933
0.20	1.5812	5.6572	10.3727
0.30	1.4909	5.2546	9.5627
0.40	1.4267	4.9672	9.0484
0.50	1.3753	4.7419	8.6599
0.60	1.3362	4.5407	8.3320
0.70	1.3020	4.3675	8.1192
0.80	1.2797	4.2447	8.0490
0.90	1.3067	4.2301	8.2530
Panel B: $RMSE^n(Model)$			
LN	1.2117	4.1604	7.7758
CP	1.2340	4.0644	7.3857
LN+CP	1.1735	4.0003	7.3855
Panel C: $RMSE^n(Model)/RMSE_{0.5}^n(Surv)$			
LN	0.9166	0.7848	0.8169
CP	0.9207	0.8527	0.8617
LN+CP	0.9467	0.8391	0.8720
Panel D: $RMSE^n(Model)/RMSE_{0.9}^n(Surv)$			
LN	1.1218	1.0421	1.0351
CP	1.1267	1.1321	1.0919
LN+CP	1.1586	1.1141	1.1049

Table IX. Determinants of Subjective 2-year Bond Returns

Table reports estimates from regressions of the subjective expected excess returns on 2-year bonds for optimistic (D90) forecasters on a set of explanatory variables. These factors are discussed in detail in the main body of the paper. t-statistics, reported in parentheses below the point estimates, are Newey-West corrected. Adjusted R-squared of the regressions are reported in the last column. The sample period is from January 1990 to July 2013.

	$DiB(g)$	$DiB(\pi)$	Liq	$Dura$	$Surp$	$LRR(g)$	$LRR(\pi)$	$TVRP$	$Jump$	$\sigma_y(3m)$	LN	CP	$\overline{R}^2$
(i)	0.26	−0.33											0.13
	(4.71)	(−4.75)											
(ii)			−0.35										0.12
			(−4.55)										
(iii)				0.38									0.14
				(6.94)									
(iv)					0.05								−0.00
					(0.82)								
(v)						0.27	0.27						0.22
						(3.91)	(1.43)						
(vi)	0.12	−0.40	−0.17	0.21	−0.09	0.22	0.33						0.49
	(2.76)	(−6.10)	(−2.99)	(3.65)	(−0.92)	(2.02)	(2.56)						
(vii)								0.12	0.18	−0.12			0.04
								(2.04)	(2.05)	(−1.71)			
$(viii)$											−0.08	0.69	0.47
											(−1.72)	(15.28)	

Table X. Determinants of Subjective 10-year Bond Returns

Table reports estimates from regressions of the subjective expected excess returns on 10-year bonds for optimistic (D90) forecasters on a set of explanatory variables. These factors are discussed in detail in the main body of the paper. t-statistics, reported in parentheses below the point estimates, are Newey-West corrected. Adjusted R-squared of the regressions are reported in the last column. The sample period is from January 1990 to July 2013.

	$DiB(g)$	$DiB(\pi)$	Liq	$Dura$	$Surp$	$LRR(g)$	$LRR(\pi)$	$TVRP$	$Jump$	$\sigma_y(3m)$	LN	CP	$\overline{R}^2$
(i)	0.30	−0.08											0.08
	(5.20)	(−0.69)											
(ii)			−0.17										0.02
			(−2.15)										
(iii)				0.19									0.03
				(2.44)									
(iv)					−0.09								0.00
					(−1.14)								
(v)						0.22	0.41						0.31
						(3.05)	(3.62)						
(vi)	0.22	−0.24	−0.01	0.03	−0.04	0.21	0.43						0.39
	(4.40)	(−3.55)	(−0.22)	(0.43)	(−0.34)	(2.11)	(4.28)						
(vii)								0.13	0.15	0.00			0.03
								(2.16)	(1.80)	(0.04)			
$(viii)$											0.25	0.43	0.27
											(3.77)	(6.94)	

Table XI. Forecast Errors and Term Structure Factors

Regressions of the forecast errors of all the deciles of the cross-sectional distribution of expected excess returns, for 2 and 10-year zero-coupon bonds, on the level (PC1), slope (PC2), and curvature (PC3) term structure factors. t-statistics, reported in parentheses below the point estimates, are Newey-West corrected. The adjusted R-squared of the regressions is also reported. The sample period is from January 1988 to July 2015.

Decile	**Level**	**Slope**	**Curv**	$\mathbf{R^2}$
		Panel A: $n = 2$ **years**		
0.10	0.11	0.04	0.08	1.01%
	(0.95)	(0.24)	(0.56)	
0.20	0.07	0.03	0.08	0.18%
	(0.60)	(0.16)	(0.54)	
0.30	0.04	0.01	0.08	−0.05%
	(0.35)	(0.03)	(0.61)	
0.40	0.02	−0.01	0.08	−0.33%
	(0.14)	(−0.07)	(0.54)	
0.50	0.00	−0.02	0.07	−0.40%
	(0.03)	(−0.14)	(0.50)	
0.60	−0.01	−0.03	0.07	−0.42%
	(−0.08)	(−0.18)	(0.47)	
0.70	−0.03	−0.05	0.06	−0.26%
	(−0.28)	(−0.27)	(0.43)	
0.80	−0.07	−0.06	0.05	0.16%
	(−0.55)	(−0.38)	(0.33)	
0.90	−0.13	−0.09	0.03	1.53%
	(−0.96)	(−0.53)	(0.20)	
		Panel B: $n = 10$ **years**		
0.10	0.01	0.33	−0.00	10.04%
	(0.10)	(2.74)	(−0.01)	
0.20	−0.03	0.32	−0.04	9.53%
	(−0.26)	(2.61)	(−0.38)	
0.30	−0.05	0.32	−0.05	10.06%
	(−0.45)	(2.60)	(−0.46)	
0.40	−0.07	0.32	−0.06	10.14%
	(−0.60)	(2.54)	(−0.57)	
0.50	−0.09	0.31	−0.08	10.23%
	(−0.80)	(2.43)	(−0.74)	
0.60	−0.11	0.30	−0.09	10.27%
	(−0.93)	(2.37)	(−0.83)	
0.70	−0.13	0.29	−0.09	10.05%
	(−1.17)	(2.22)	(−0.89)	
0.80	−0.16	0.26	−0.10	9.55%
	(−1.44)	(1.99)	(−0.99)	
0.90	−0.23	0.21	−0.13	10.50%
	(−2.09)	(1.58)	(−1.26)	

Table XII. Test of Information Rigidities

Panel A reports the results of regressions of the realized excess bond return forecast errors on the forecast revision for the consensus forecaster, for maturities of 2, 5 and 10 years:

$$fe^n_{t+1} = a^n + b^n \Delta erx^n_{c,t} + \epsilon^n_{t+1}.$$

t-statistics, reported in parentheses below the point estimates, are Newey-West corrected. The adjusted R-squared of the regressions is also reported. The sample is quarterly, from January 1988 to October 2014.

Panel B reports the same regression including the time-t level and slope of the term structure of interest rates on the right hand side:

$$fe^n_{t+1} = a^n + b^n_1 \Delta erx^n_{c,t} + b^n_2 Level_t + b^n_3 Slope_t + \epsilon^n_{t+1},$$

where *Level* and *Slope* are the first two principal components of the term structure of yields.

Maturity	**const**	**Revision**	**Level**	**Slope**	$\bar{\mathbf{R}}^2$
		Panel A: Baseline Regression			
2y	0.006	−0.770			5.58%
	(2.444)	(−3.336)			
5y	0.025	−0.906			8.63%
	(3.661)	(−3.834)			
10y	0.041	−1.142			13.61%
	(3.607)	(−5.254)			
		Panel B: Additional Regressors			
2y	0.006	−0.963	−0.001	0.003	6.24%
	(2.638)	(−4.050)	(−1.004)	(0.668)	
5y	0.026	−1.188	−0.003	0.019	14.23%
	(4.109)	(−5.429)	(−1.693)	(1.392)	
10y	0.043	−1.385	−0.005	0.054	25.79%
	(4.409)	(−7.211)	(−1.748)	(2.697)	