



DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY



Expected Business Conditions and Bond Risk Premia

Jonas Nygaard Eriksen

CREATES Research Paper 2015-44

Expected Business Conditions and Bond Risk Premia*

Jonas Nygaard Eriksen[†]

Abstract

This paper studies the predictability of bond risk premia by means of expectations to future business conditions using survey forecasts from the Survey of Professional Forecasters. We show that expected business conditions consistently affect excess bond returns and that the inclusion of expected business conditions in standard predictive regressions improve forecast performance relative to models using information derived from the current term structure or macroeconomic variables. The results are confirmed in a real-time out-of-sample exercise, where the predictive accuracy of the models is evaluated both statistically and from the perspective of a mean-variance investor that trades in the bond market.

Keywords: Bond risk premia, expected business conditions, predictability, economic value, expectations hypothesis, time-varying risk premia

JEL Classification: E43, E44, E47, G11, G12.

This version: September 24, 2015

* I am grateful to Stig Vinther Møller, Maik Schmeling, an anonymous referee, and seminar participants at Aarhus University and CREATES and the 2015 Nordic Econometric Meeting in Helsinki for helpful comments and suggestions. The author acknowledges support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation

[†] CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. Tel.: +45 87165321. E-mail: jeriksen@econ.au.dk.

1 Introduction

There is mounting evidence that the risk premia required by investors for holding Treasury bonds contain a time-varying and predictable component. One strand of research relates such variations to forward spreads (Fama and Bliss, 1987; Fama, 2006), yield spreads (Keim and Stambaugh, 1986; Campbell and Shiller, 1991), and forward rates (Cochrane and Piazzesi, 2005; Dahlquist and Hasseltoft, 2013; Zhu, 2015), whereas more recent studies link the predictable component to factors whose variations lie outside the span of current yields. Existing research within this strand has uncovered a wide array of factors including jump risk (Wright and Zhou, 2009), option prices (Almeida, Graveline, and Joslin, 2011), and macroeconomic variables (Ilmanen, 1995; Moench, 2008; Cooper and Priestley, 2009; Ludvigson and Ng, 2009; Wright, 2011; Favero, Niu, and Sala, 2012; Joslin, Priebisch, and Singleton, 2014; Zhou and Zhu, 2015).

While existing studies mainly rely on information in the current term structure and business environment to explain variations in the predictable component of bond risk premia, our study takes a forward-looking perspective by studying the link between expected business conditions and bond risk premia using survey forecasts from the *Survey of Professional Forecasters* (SPF). Albeit empirical studies frequently conclude that macroeconomic fundamentals carry information about bond risk premia not already captured by the yield curve, they rarely account for issues with publication lags and data revisions in macroeconomic time series. As recently demonstrated by Ghysels, Horan, and Moench (2014), such features may drive a wedge between the factor spaces spanned by latent common factors extracted from a panel of revised and non-revised data, respectively, implying that a sizable fraction of the predictability documented by Ludvigson and Ng (2009) may largely be driven by revision and publication lag components that are unavailable to an investor in real-time. Survey forecasts, as a result, provide an ideal data source for studying the predictability of bond risk premia as they are model-free, available in real-time, and not subject to subsequent revisions (Amato and Swanson, 2001; Croushore and Stark, 2001; Pesaran and Timmermann, 2005). Moreover, the use of survey forecasts allows us to investigate if macroeconomic expectations influence future bond risk premia as suggested by, among others, Chen, Roll, and Ross (1986), Fama and French (1989), Barro (1990), Fama (1990), and Ferson and Harvey (1991).

Using the SPF survey forecasts, we investigate empirically whether expected business conditions are related to one-year ahead variations in bond risk premia and, secondly, whether expected business conditions contain information orthogonal to current business conditions and yield curve information. Our findings suggest that excess bond returns are indeed predictable by expected business conditions. In particular, we find that expected business conditions are able explain about 20% of the one-year ahead variation in bond risk premia. Consistent with recent research, we find the majority of the informational content

in expected business conditions to lie outside the span of contemporaneous yields (Duffee, 2011; Joslin et al., 2014). From a modeling perspective, this has important implications for leading affine term structure models, where it is normally assumed that all macroeconomic state variables can be fully spanned by the current term structure.¹ More importantly, we find that expected business conditions contain information about future bond risk premia not already embedded in either the CP_t factor from Cochrane and Piazzesi (2005) or the LN_t factor from Ludvigson and Ng (2009). In both cases, we see sizable improvements in the in-sample fit from augmenting the existing models with expected business conditions. This indicates that expected business conditions are not simply a repackaged version of the standard factors, but constitute an independent and important source of risk.

While empirical studies frequently report that bond risk premia are predictable in-sample, there is less systematic evidence for bond risk premia being predictable out-of-sample and, in particular, that an investor’s economic utility can be improved by exploiting this predictability to guide real-time asset allocation decisions. Notably, Thornton and Valente (2012) find that predictive models using long-term forward rates are unable to generate economic value over a simple expectations hypothesis (EH) no-predictability benchmark. A similar result is found in Sarno, Schneider, and Wagner (2014). One possible explanation, as suggested by Andreasen, Engsted, Møller, and Sander (2015), is that bond return predictability by term structure variables is confined to periods of economic expansion, whereas predictability is largely absent in recessions.² Gargano, Pettenuzzo, and Timmermann (2014), on the other hand, using models that allow for time-varying parameters and stochastic volatility in the predictive regressions, find that bond predictability by factors such as LN_t is concentrated in recession periods. Although Ludvigson and Ng (2009) find that latent factors from a large panel of macroeconomic variables forecast excess bond returns out-of-sample, Ghysels et al. (2014) argue that the majority of this predictability derives from using revised data. To address such concerns directly, we construct a panel of 67 macroeconomic time series from vintage data. Using vintages of data consisting of historically available information allows us to formulate a variant of the LN_t factor that is available in real-time and free from publication lag and revision issues. The panel broadly covers the same categories of variables used in Ludvigson and Ng (2009), although we exclude financial variables to focus on the relative forecasting performance of current and expected business conditions, respectively.

To assess the statistical and economic value of bond excess return predictability, we consider an out-of-sample analysis in which we generate excess bond return forecasts for the

¹Examples include Ang and Piazzesi (2003), Ang, Dong, and Piazzesi (2007b) Rudebusch and Wu (2008), Bikbov and Chernov (2010), and Bansal and Shaliastovich (2013). See Andersen and Benzoni (2010), Chernov and Mueller (2012), and Joslin et al. (2014) for recent applications using unspanned risk factors.

²Interestingly, this is the opposite pattern often found for stock returns (Henkel, Martin, and Nardari, 2011; Dangl and Halling, 2012; Rapach and Zhou, 2013; Møller and Sander, 2014).

period 1990-2014 at a quarterly frequency. The forecasts are generated recursively using an expanding window of observations, where model parameters and predictor variables are updated recursively prior to each forecast as well. Importantly, we rely on historically available information only to mimic a real-time forecasting environment and avoid concerns of look-ahead bias induced by full sample parameters. Compared to the benchmark EH model, we find that expected business conditions produce significantly positive out-of-sample R^2 values. Moreover, expected business conditions provide equally or more accurate predictions than the CP_t and the LN_t factors for most bond maturities and the predictive power is increasing with maturity. Likewise, augmenting the existing models with expected business conditions results in near universal improvement in forecasting accuracy, where the best forecasts, in a statistical sense, are achieved from a three-factor specification including CP_t , LN_t , and expected business conditions. The economic value of using expected business conditions to guide portfolio allocations is evaluated from the perspective of an investor with mean-variance preferences who allocates capital between a risky bond with $n = 2, \dots, 5$ years to maturity and a risk-free Treasury bond that matures in one year.³ Our results demonstrate that an investor, who relies on the investment signals generated by expected business conditions, is able to improve upon the economic utility realized by an investor whose portfolio allocations are based on the recursively updated EH model. Moreover, we find sizable utility gains from using expected business conditions across the entire maturity spectrum relative to using CP_t and LN_t and, additionally, that augmenting the standard models with expected business conditions results in universally improved portfolio performance.

Our approach complements an extensive literature that studies bond market dynamics using survey information. [Froot \(1989\)](#), in an early paper, uses survey data on interest rates to document that the severity of expectations hypothesis (EH) violations depends on bond maturity. Similarly, [Chun \(2011\)](#), using survey forecasts for output, inflation, and the federal funds rate, finds that the term structure moves with monetary policy expectations. A related result is established in [Altavilla, Giacomini, and Constantini \(2014\)](#) whose monetary policy expectations, however, are extracted from federal funds rate futures. [Dick, Schmeling, and Schrimpf \(2013\)](#) use forecasters' subjective yield expectations to construct a real-time proxy for expected term premia that predicts future bond risk premia. A similar approach is found in [Piazzesi, Salomao, and Schneider \(2013\)](#). Moving away from yield curve information, [Wright \(2011\)](#) considers survey forecasts on macroeconomic fundamentals to improve term premia estimates. Similarly, [Chernov and Mueller \(2012\)](#) use information from the term structure of survey-based inflation expectations to uncover a hidden factor in the yield curve and [Joslin et al. \(2014\)](#) use economic growth and expected inflation as unspanned factors to obtain more precise term premia estimates. Using the

³Similar approaches are found in [Campbell and Thompson \(2008\)](#), [Dangl and Halling \(2012\)](#), [Thornton and Valente \(2012\)](#), [Rapach and Zhou \(2013\)](#), [Gargano et al. \(2014\)](#), [Sarno et al. \(2014\)](#), and [Zhu \(2015\)](#).

cross-sectional dispersion in survey forecasts, [Bansal and Shaliastovich \(2013\)](#) find that bond risk premia rise with uncertainty about expected inflation and fall with uncertainty about expected economic growth. Our approach is also related to a large literature that investigates the predictive ability of survey forecasts for macroeconomic forecasting. For instance, [Fama and Gibbons \(1984\)](#), [Thomas \(1999\)](#), and [Ang, Bekaert, and Wei \(2007a\)](#) show that survey forecasts are robust predictors of future inflation. Finally, our study contributes to the literature that studies the predictability of bond risk premia directly in a predictive regression framework by demonstrating that forward-looking macro expectations are important for understanding the time-varying component of bond risk premia and, importantly, that the predictive information in expected business conditions is not already captured by standard factors.⁴

The rest of the paper is organized as follows. Section 2 lays out the basic bond pricing terminology and the econometric framework used to assess bond excess return predictability. Section 3 outlines our proxy for expected business conditions. Section 4 details the data and provides descriptive statistics. Section 5 presents the empirical results. Finally, Section 6 concludes.

2 Predicting bond risk premia

Our study focus on zero-coupon US Treasury bonds with maturities ranging from two to five years. Let $p_t^{(n)} = \log P_t^{(n)}$ denote the time t log nominal price of a bond with n -years left to expiry and a terminal payoff of a dollar. The continuously compounded (log) yield of an n -year bond is given by

$$y_t^{(n)} = -(1/n) p_t^{(n)} \quad (1)$$

where we use parentheses to distinguish the remaining time to maturity from exponentiation in the superscript. The log forward rate at time t for loans between time $t + n - 1$ and $t + n$ is defined as

$$f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)} \quad (2)$$

which, since $p_t^{(0)} = 0$, implies the relation $f_t^{(1)} = y_t^{(1)} = -p_t^{(1)}$ that throughout will be considered the risk-free rate of return. The return on a risky n -year Treasury bond less the risk-free rate of return is then given by

$$rx_{t+4}^{(n)} = p_{t+4}^{(n-1)} - p_t^{(n)} - y_t^{(1)} \quad (3)$$

where $p_{t+4}^{(n-1)} - p_t^{(n)} = r_{t+4}^{(n)}$ is the log holding period return from buying an n -year bond at time t and selling it as an $n - 1$ year bond after four quarters. The average excess return

⁴See also [Campbell and Diebold \(2009\)](#) and [Schmeling and Schrimpf \(2011\)](#) for a similar approach.

across the bond maturity spectrum is given as

$$\overline{rx}_{t+4} = 4^{-1} \sum_{n=2}^5 rx_{t+4}^{(n)} \quad (4)$$

which can be thought of as the annual excess log holding period return to an equally-weighted portfolio of risky Treasury bonds with maturities ranging from two through five years.

The prevailing practice in the literature is to assess the predictive ability of time t observable covariates for bond risk premia by means of a predictive regression nested within the general specification

$$rx_{t+4}^{(n)} = \alpha + \mathbf{Z}'_t \beta + \varepsilon_{t+4}^{(n)} \quad (5)$$

where $rx_{t+4}^{(n)}$ is the one-year ahead excess return on a bond with n -years left to expiry from (3) and \mathbf{Z}_t is a vector of time t observable covariates. We suppress the maturity dependence in the parameters for notational simplicity throughout. Note that the expression in (5) for $\beta = 0$ reduces to an expectation hypothesis (EH) equivalent under which bond risk premia are unpredictable and equal to the constant $\alpha^{(n)}$ (maturity dependence emphasized), which can vary across maturities, but not in time.

The previous empirical literature has uncovered a wide array of candidates for \mathbf{Z}_t that explains part of the one-year ahead variation in bond risk premia, cf. Section 1. The most prominent examples include the [Cochrane and Piazzesi \(2005\)](#) forward rate-based factor, which is a tent-shaped linear combination of forward rates, and the [Ludvigson and Ng \(2009\)](#) factor, which is a linear combination of latent factors extracted from a large panel of macroeconomic information. While [Ludvigson and Ng \(2009\)](#) use a panel of revised data series, we consider a variant of their factor that permits the use of real-time macroeconomic data for factor construction in the out-of-sample analysis. In particular, since we are ultimately interested in predicting bond risk premia in real-time for asset allocation purposes, we consider a panel of purely real-time observable macroeconomic variables from the Archival Federal Reserve Economic Data (ALFRED) database maintained by the Federal Reserve Bank of St. Louis. As in [Ghysels et al. \(2014\)](#), we exclude financial variables to focus on the distinction between current and expected business conditions. We estimate the latent common factors using the dynamic factor model of [Stock and Watson \(2002a,b\)](#) and determine the optimal number of factors to extract using the panel information criterion from [Bai and Ng \(2002\)](#). Following [Ludvigson and Ng \(2009\)](#), we then select the optimal subset of the factors and their nonlinear transformations using the Schwartz-Bayesian information criterion (SIC).

While these factors are designed to exploit the information in the current term structure and business environment, respectively, this paper takes a forward-looking perspective by

constructing a proxy for expected business conditions from survey forecasts to serve as a candidate predictor. The motivation for this choice originates, in part, from the seminal papers of [Chen et al. \(1986\)](#), [Fama and French \(1989\)](#), [Barro \(1990\)](#), [Fama \(1990\)](#), and [Ferson and Harvey \(1991\)](#), who argue that expected business conditions are fundamental drivers of fluctuations in expected risk premia and interpret the predictive power of financial ratios and variables through that lens.⁵ Secondly, expected business conditions are, from a theoretical point of view, likely to contain information about time-variations in investment opportunities in the sense of [Merton \(1973\)](#), suggesting that macro expectations should be informative about the time-varying component in bond risk premia. Such a view is recently explored by [Rossi and Timmermann \(2015\)](#), who show that conditional covariance risk, measured as the realized covariance between daily economic news about the future state of the economy and daily returns, plays an important role in explaining time-variations in risk premia. Moreover, modern asset pricing models, such as those discussed in [Campbell and Cochrane \(1999\)](#), [Bansal and Yaron \(2004\)](#), [Wachter \(2006\)](#), [Buraschi and Jiltsov \(2007\)](#), and [Bansal and Shaliastovich \(2013\)](#), argue that investors are concerned about the future state of the real economy and require compensation for taking on business cycle risk. We detail the construction of our proxy for expected business conditions in the next section.

3 Measuring expected business conditions

Although several model-based procedures exist for extracting expectations about the future course of the real economy, we rely on survey forecasts from the *Survey of Professional Forecasters* (SPF) to measure expected business conditions. As mentioned in the introduction, such an approach complements a large literature relying on survey information for forecasting and understanding bond market dynamics. While studies such as [Chun \(2011\)](#), [Wright \(2011\)](#), [Chernov and Mueller \(2012\)](#), [Dick et al. \(2013\)](#), [Piazzesi et al. \(2013\)](#), and [Joslin et al. \(2014\)](#) rely on survey forecasts to understand yield dynamics and improve term premia estimates, we follow the tradition of, among others, [Cochrane and Piazzesi \(2005\)](#) and [Ludvigson and Ng \(2009\)](#) and study the predictive ability of survey forecasts for bond risk premia directly in a predictive regression framework.⁶

The use of survey forecasts in a study like ours, where the aim is to investigate the link between expected business conditions and bond risk premia, seems ideal for several reasons. First, and most notably, survey forecasts are available in real-time and not subject to revisions following the initial release ([Amato and Swanson, 2001](#); [Croushore and Stark,](#)

⁵[Campbell and Diebold \(2009\)](#) provide empirical evidence in support of this interpretation for stock returns using survey forecasts for GDP as a proxy for expected business conditions.

⁶Our study can, in some sense, be viewed as a generalization of the approach in [Chernov and Mueller \(2012\)](#) as we are using the term structure of survey expectations for a broad set of fundamentals rather than the term structure of inflation expectations alone.

2001). The notion of model-free survey forecasts that are naturally constructed in real-time further mitigates any concerns about potential look-ahead biases in forecasts constructed from an ex post econometric model (Pesaran and Timmermann, 2005; Capistrán and Timmermann, 2009). In addition, survey forecasts are often found to be robust predictors of the surveyed macroeconomic fundamentals (Fama and Gibbons, 1984; Thomas, 1999; Ang et al., 2007a). The use of survey forecasts is, however, also associated with a few shortcomings. For instance, as suggested in Ottaviani and Sørensen (2006), survey participants may give strategic forecasts. Survey forecasts are, in such circumstances, unlikely to reveal the true expectations of market participants, but such incentives should be significantly reduced given the anonymity of the forecasters in the SPF. Moreover, as will be detailed shortly, using a median aggregate across participating forecasters is likely to further iron out the impact of any such outliers.

The SPF has been running since 1968 and is conducted on a quarterly basis. The questionnaires are sent out at the end of the first month of each quarter with response deadline in the middle of the second month of the quarter. The SPF solicits a broad variety of participants covering financial firms, banks, consulting firms, and research centers concerning their views on economic variables today, a so-called nowcast, and one- through four-quarters ahead. That respondents are asked to provide a nowcast is an important advantage of the SPF relative to similar surveys as it allows for the construction of one-through four-quarter-ahead growth forecasts based on the forecasters own expectation for the current level. Using the SPF, as a result, allows us to measure the term structure of growth expectations over the annual bond holding period usually assumed in empirical work.⁷

We collect survey forecasts for six macroeconomic fundamentals that are available at a quarterly frequency from 1968:Q4 to 2014:Q4, where 1968:Q4 marks the initiation of the SPF. The macroeconomic fundamentals include 1) GDP ($\text{gdp}_t^{\mathbb{E}}$), 2) the GDP price index ($\text{inf}_t^{\mathbb{E}}$), 3) the unemployment rate ($\text{unemp}_t^{\mathbb{E}}$), 4) corporate profits after tax ($\text{cprof}_t^{\mathbb{E}}$), 5) industrial production ($\text{ip}_t^{\mathbb{E}}$), and 6) housing starts ($\text{hous}_t^{\mathbb{E}}$). For each variable and survey respondent, we compute log expected growth rates one- through four-quarters ahead relative to the forecaster's own nowcast, with the exception of housing starts for which we compute expected changes. The individual growth forecasts are then aggregated into median forecasts such that we obtain for each series and survey date a median growth forecast one- through four-quarters ahead. We use the median forecast due to its robustness against outliers. The average number of participating forecasters per quarter is about 36 in our sample.

[Table 1 around here]

⁷See, among others, Cochrane and Piazzesi (2005), Ludvigson and Ng (2009), and Thornton and Valente (2012).

Table 1 provides descriptive statistics for our set of survey forecasts. A quick glance reveals that mean expectations are monotonically increasing with forecast horizon for most variables, the exception being unemployment rates ($\text{unemp}_t^{\mathbb{E}}$), and that the variability of the forecasts scales with the forecast horizon. A similar observation emerges from investigating the time series of survey forecasts as they tend to be extrapolative in the forecast horizon, i.e. the median forecasts at a given point in time are usually either increasing or decreasing in forecast horizon.

Considering the information in the term structure of survey forecasts over the holding period of the bond provides us with a total of 24 economic time series to use for forecasting. Although possible, using all variables at once quickly becomes impractical as the predictive regression would likely suffer from degrees of freedom problems and model selection would require the evaluation of all 2^{24} possible predictor combinations. Instead, we follow the line of thought from Ludvigson and Ng (2009), and to some extent Litterman and Scheinkman (1991), and summarize the information in the term structure of expected business conditions by a small number of latent factors, or principal components, denoted $\mathcal{P}_t^{\mathbb{E}}$, which are distinct linear combinations of the underlying survey forecasts. We concentrate our analysis on the first three principal components ($\mathcal{P}_{1,t}^{\mathbb{E}}, \mathcal{P}_{2,t}^{\mathbb{E}}, \mathcal{P}_{3,t}^{\mathbb{E}}$), which explain 53%, 25%, and 14% of the total variation in the panel, respectively.⁸

The loadings from the principal component analysis, which can be used to give an informal interpretation of the full sample components, are reported in the bottom half of each panel of Table 1. We see that the first component, $\mathcal{P}_{1,t}^{\mathbb{E}}$, generally loads on real activity related variables such as GDP, corporate profitability, unemployment rates, and industrial production, implying that this component can be interpreted as a real activity factor. The second component, $\mathcal{P}_{2,t}^{\mathbb{E}}$, tends to load highly on inflation, making it an inflation factor. Finally, we label the third component, $\mathcal{P}_{3,t}^{\mathbb{E}}$, a housing factor as it loads strongly on expected housing starts. Figure 1 plots the time series dynamics of the full sample components against NBER defined recession periods. The real activity factor tends to decline quickly in recession periods, as defined by the NBER, where economic growth is usually low, and the inflation factor displays the long-term decline exhibited by realized inflation.

[Figure 1 around here]

Cochrane and Piazzesi (2005) and Ludvigson and Ng (2009) show that a single forecasting factor, which they make observable through a linear combination of forward rates and latent macro factors, respectively, forecasts excess holding period returns on Treasury bonds with maturities ranging from two to five years. It seems natural to investigate

⁸Although the components that are pervasive in explaining the panel of survey-based growth forecasts may not similarly be important for bond risk premia (Bai and Ng, 2008), we find no significant gains from including later components in any of our analyses.

whether the informational content in the survey forecast components can be similarly condensed to recover a single mean-reverting factor that forecasts bond excess returns for all maturities. To answer this question, we begin by investigating the predictive ability of the latent factors for each bond maturity by projecting the annual holding period excess returns from (3) onto the principal components extracted from the term structure of expected business conditions, i.e.

$$rx_{t+4}^{(n)} = \alpha + \beta_1 \mathcal{P}_{1,t}^{\mathbb{E}} + \beta_2 \mathcal{P}_{2,t}^{\mathbb{E}} + \beta_3 \mathcal{P}_{3,t}^{\mathbb{E}} + \varepsilon_{t+4}^{(n)} \quad (6)$$

where the superscript (n) denoting bond maturity is suppressed in the parameters for notational simplicity. The results from estimating (6) are presented in rows denoted (a) in each panel of Table 2. The real activity factor, $\mathcal{P}_{1,t}^{\mathbb{E}}$, and the housing factor, $\mathcal{P}_{3,t}^{\mathbb{E}}$, are strongly significant for all maturities, with the exception of $\mathcal{P}_{1,t}^{\mathbb{E}}$ for the two-year bond. The inflation factor, $\mathcal{P}_{2,t}^{\mathbb{E}}$, however, is significant only for the long end of the bond maturity spectrum. The components are jointly able to explain between 19% and 20% of the one-year ahead variation in bond risk premia, indicating that expected business conditions play a non-negligible role in explaining time-variations in bond risk premia. $\mathcal{P}_{1,t}^{\mathbb{E}}$ and $\mathcal{P}_{3,t}^{\mathbb{E}}$ enter with positive coefficients, implying that increases in forward-looking expectations about the real economy and housing market, respectively, are associated with higher bond risk premia going forward. $\mathcal{P}_{2,t}^{\mathbb{E}}$, on the other hand, enters with a negative sign for all bond maturities, implying that lower inflation expectations predict increasing risk premia. This is consistent with prior studies on expected inflation and asset returns (Fama and Schwert, 1977; Campbell and Ammer, 1993; Cieslak and Povala, 2015).

[Table 2 around here]

Similarly to the findings in Cochrane and Piazzesi (2005) and Ludvigson and Ng (2009), we see that the same function of covariates, where the absolute value of the loadings is increasing in maturity, forecasts bond risk premia across the maturity spectrum. To exploit this pattern, we construct a single factor by first regressing the mean excess return across the maturity spectrum upon the survey factors

$$\overline{rx}_{t+4} = \delta + \gamma_1 \mathcal{P}_{1,t}^{\mathbb{E}} + \gamma_2 \mathcal{P}_{2,t}^{\mathbb{E}} + \gamma_3 \mathcal{P}_{3,t}^{\mathbb{E}} + \nu_{t+4} \quad (7)$$

where \overline{rx}_{t+4} denotes the average excess bond return from (4). We then compute the single factor, which we denote by ME_t for macroeconomic expectations, as the fitted values from (7), i.e. $\text{ME}_t = \widehat{\delta} + \widehat{\gamma} \mathcal{P}_t^{\mathbb{E}}$.⁹ We then regress ME_t on excess bond returns and present the results in rows (b) of Table 2. The coefficients are all positive and increasing in bond maturity, which indicates that an increase in expected business conditions predicts higher

⁹The estimated values of γ_1 , γ_2 , and γ_3 are 0.16, -0.33, and 0.72, respectively. Hansen and Hodrick (1980) t-statistics are equal to 2 or higher in absolute size for all parameters.

future excess bond returns. Notably, the adjusted R^2 values remain comparable to the full model, suggesting that no efficiency is lost from using the single factor over the full model.

[Figure 2 around here]

Figure 2 plots the ME_t factor against both the Chicago Fed National Activity Index (CFNAI) and NBER defined recession periods.¹⁰ ME_t tends to increase during periods of economic contraction and peak at the end of a recession, whereas it tends to decline during expansion with its low points being reached in the mid-to-late expansion periods. The CFNAI, in contrast, tends to move in the opposite direction. It tends to peak mid-to-late in expansion periods and bottom out mid-to-late in recessionary periods. This suggests that ME_t captures cyclical variation related to the business cycle. In particular, since ME_t enters with a positive sign in the excess bond return regressions in Table 2 and is negatively correlated with the CFNAI (-0.19 to be exact), this implies that excess bond returns are predicted to be high when ME_t (CFNAI) is high (low) and, as a result, exhibit a countercyclical behavior. These observations are consistent with economic theories and models in which investors require compensation for bearing business cycle risk, e.g. Campbell and Cochrane (1999), Wachter (2006), and Buraschi and Jiltsov (2007).

4 Data sources

Treasury bond data covering the period from 1968:Q4 to 2014:Q4 are obtained from the Fama-Bliss dataset available from the Center for Research in Security Prices (CRSP), which contains observations on one- through five-year zero-coupon US Treasury bond prices. We construct bond risk premia and forward rates for the Cochrane and Piazzesi (2005) factor as described in Section 2. Due to the odd timing of the SPF responses mentioned in Section 3, we follow, among others, Bansal and Shaliastovich (2013) and sample bond prices at the end of the second month of every quarter to align survey responses and bond prices. As mentioned in Section 3, survey forecasts are obtained from the *Survey of Professional Forecasters* (SPF), which is available at the Federal Reserve Bank of Philadelphia, to measure forecasters' subjective expectation to future growth in our set of macroeconomic fundamentals.¹¹

The data for our variant of the LN_t factor, as mentioned in Section 2, are obtained from the Archival Federal Reserve Economic Data (ALFRED) database maintained by the Federal Reserve Bank of St. Louis.¹² Specifically, we construct a panel consisting of 67

¹⁰The Federal Reserve Bank of Chicago constructs the CFNAI from inflation adjusted economic indicators such as: Production and income (23 series), employment and hours (24 series), personal consumption and housing (15 series), and sales, order, and inventories (23 series). The estimation methodology follows Stock and Watson (1999).

¹¹The survey is available at <http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/>.

¹²The data are available at <http://alfred.stlouisfed.org>.

macroeconomic time series, where any variable is required to fulfill the following criteria for inclusion. First, each variable is required to have a full time series of observations from 1968 or earlier available in all of its vintages and, secondly, each variable must have real-time vintages available from 1989 onwards. Appendix A provides a detailed account of this real-time dataset along with the transformations that we apply in order to ensure covariance stationarity. The data broadly cover the same economic categories used in Ludvigson and Ng (2009). In particular, the series include output and labor market variables, housing indicators, income series, price indexes, and the money stock. The macroeconomic series and vintages are recorded at a monthly frequency, so we follow the approach from above and sample from the second month of each quarter. Although the real-time vintages are mostly recorded at a monthly frequency, there are cases where vintages are updated multiple times within a month or not at all. To account for this, we use the following generic set of rules. In months with multiple vintages, we pick out the latest observable vintage. Formally, this is based on an assumption that forecasters usually make their predictions and asset allocation decisions at month’s end. If a month has no recorded vintage, we use the previous month’s vintage to avoid the introduction of look-ahead biases. Also obtained from the Federal Reserve Bank of St. Louis are the Chicago Fed National Activity Index (CFNAI) and the National Bureau of Economic Research (NBER) recession dates.

[Table 3 around here]

Table 3 presents summary statistics for excess bond returns, $rx_{t+4}^{(n)}$, $n = 2, \dots, 5$, and predictors in Panel A along with their contemporaneous correlations in Panel B. Mean excess bond returns range from 0.60% for the two-year bond to 1.63% for the five-year bond and corresponding standard deviations range from 1.79 to 5.52, suggesting that investors require a higher premium for investing in longer maturity (riskier) Treasury bonds. The return and volatility values correspond to Sharpe ratios of around 0.30-0.34 over our sample period. Excess bond returns are negatively skewed and leptokurtic. All bonds display strong serial correlation with first-order autocorrelation coefficients of about 0.75. The serial dependence, however, reduces drastically when we instead consider the fourth-order autocorrelation coefficient due to the overlapping nature of the holding period returns. Here autocorrelation coefficients are only in the 0.1-0.2 range and insignificant (highest t-stat is around 1.6). As is evident in Panel B, excess bond returns are highly cross-sectionally correlated with coefficients well above 0.90.

Turning our attention to the predictor variables, we see that CP_t , LN_t , and ME_t are positively correlated with one-year ahead bond risk premia. As in Ludvigson and Ng (2009), we note that CP_t and LN_t are almost unrelated with a correlation of only 0.11 in the full sample. Similarly, we see that the informational content in our macroeconomic expectations factor, ME_t , appears to be largely unrelated to CP_t and LN_t with correlations

of 0.42 and 0.33, respectively, which are far from perfect. As such, ME_t appears from the outset to capture a relatively large informational component orthogonal to the existing forecasting factors from the literature. Whether this information is useful for predicting excess bond returns is, of course, an important empirical question that this paper seeks to address.

5 Empirical results

This section presents the results from our empirical analyses. We begin by considering results based on full sample estimates to remain comparable with the existing literature on bond risk premia. Specifically, Section 5.1 examines whether expected business conditions can be thought of as an unspanned risk factor in the sense of [Joslin et al. \(2014\)](#), whereas Section 5.2 compares the predictive ability of expected business conditions to competing forecasting factors representing information in the current term structure and business environment. However, a good in-sample fit does not necessarily translate into positive out-of-sample performance ([Inoue and Kilian, 2005](#); [Thornton and Valente, 2012](#)). To investigate the real-time predictive ability of expected business conditions, Section 5.3 considers an out-of-sample exercise that mimics a real-time forecasting environment in which model parameters and forecasting factors are re-estimated recursively using information available at the time of the forecast only. Notably, we base the estimation of our variant of the LN_t factor on vintage data to void concerns of look-ahead bias induced by ignoring publication lag and revision issues. The forecasts are evaluated using both conventional statistical measures (Section 5.3.1) and from the perspective of a mean-variance investor (Section 5.3.2). Section 5.4 investigates the gains from combining model forecasts and Section 5.5 investigates the link between forecast performance and the real economy. Finally, Section 5.6, as a robustness check, studies the predictive ability of expected business conditions for quarterly holding period returns rather than the annual holding period used for the main results.

5.1 Expected business conditions as an unspanned risk factor

Several recent papers have considered the possibility that some factors in- and outside the current term structure are hidden in the sense that while they are largely irrelevant for explaining the cross-sectional variation in yields, they are important for forecasting bond risk premia. In particular, [Duffee \(2011\)](#) uses filtering to recover a hidden factor in yields, whereas [Joslin et al. \(2014\)](#) recover two hidden factors from economic activity and expected inflation that are largely unspanned by the current term structure. Similarly, [Chernov and Mueller \(2012\)](#) uncover a hidden factor in the yield curve using the term structure of survey-based inflation expectations. While Section 3 finds that expected

business conditions explain a sizable fraction of the one-year variation in bond risk premia, this section examines whether this predictive information is unspanned by current yields.

As is standard in the literature, we summarize the term structure by a small set of linear combination of yields, e.g. principal components, as they virtually capture all cross-sectional variation (Litterman and Scheinkman, 1991). The first three PCs, say $\mathcal{Y}_{1,t}$, $\mathcal{Y}_{2,t}$, and $\mathcal{Y}_{3,t}$, are commonly known as level, slope, and curvature, respectively. To gauge the degree to which ME_t can be thought of as an unspanned risk factor in the sense of Joslin et al. (2014), we consider the macro-spanning condition

$$\mathcal{M}_{it} = \alpha + \theta \mathcal{Y}_t + \varepsilon_{it} \quad (8)$$

where \mathcal{M}_t denotes the vector of variables used to construct ME_t and \mathcal{Y}_t denotes the vector of either the first three (\mathcal{Y}_t^3) or the first five (\mathcal{Y}_t^5) principal components of the yield curve covariance matrix.¹³

[Table 4 around here]

If a variable is fully spanned by current yields, then the error term in (8) should be identically zero and the associated R^2 should equal unity. Table 4 demonstrates that this is not the case. In particular, a projection of $\mathcal{P}_{1,t}^{\mathbb{E}}$ upon \mathcal{Y}_t^3 (\mathcal{Y}_t^5) gives an adjusted R^2 of 28% (28%), implying that around 72% of the variation in $\mathcal{P}_{1,t}^{\mathbb{E}}$ is unspanned by the current term structure. The corresponding numbers for $\mathcal{P}_{2,t}^{\mathbb{E}}$ and $\mathcal{P}_{3,t}^{\mathbb{E}}$ are, respectively, 61% (63%) and 11% (18%), revealing a large fraction of unspanned risk, even for the inflation factor. This is in line with Joslin et al. (2014), who document similar results for economic growth measured as the three-month moving average of the CFNAI and one-year ahead expected inflation from the SPF. Projecting ME_t on \mathcal{Y}_t^3 (\mathcal{Y}_t^5) reveals that 58% (55%) of the variation in expected business conditions is unspanned by the current term structure. In unreported results, we find that ME_t survives the inclusion of \mathcal{Y}_t^5 in predictive regressions such as (5), implying that the projection errors from (8) retain most of the predictive power from the variables in the vector \mathcal{M}_t . Such results have significant implications for affine term structure models, where all macroeconomic risk is assumed to be fully spanned by yields. Moreover, it tells us that we need to look beyond the term structure to fully understand the observed time-variations in bond risk premia.

5.2 In-sample comparisons with standard predictors

This section compares the predictive ability of expected business conditions to standard factors identified in the literature. Table 5 presents results from estimating predictive regressions over the full range of available observations for different combination of

¹³See also Andersen and Benzoni (2010).

predictors. Although such estimates are not available to an investor in real-time, they are useful for understanding the basic mechanisms of the models. We report slope estimates, t-statistics based on Hansen and Hodrick (1980) standard errors implemented with four lags, and adjusted R^2 values.

[Table 5 around here]

We begin with the results for the CP_t factor presented in rows denoted (a) in Panels A-D in Table 5. We see that CP_t is able to explain 13-18% of the one-year ahead variation in bond risk premia across the maturity spectrum. While the explanatory power is lower than in Cochrane and Piazzesi (2005), this can be attributed to differences in sample period and frequency. Similarly, however, we obtain significant coefficients that are monotonically increasing with maturity. If we augment their model with our ME_t factor in a two-factor specification in rows (b) of Panels A-D, we find that the coefficient on ME_t remains highly significant across all maturities, whereas the magnitude and statistical significance of CP_t reduce substantially. The inclusion of expected business conditions raises for all bond maturities the adjusted R^2 with about ten percentage points, indicating that expected business conditions provide an important complimentary information source to the term structure of forward rates, despite the forward-looking nature of forward rates.

Next, we turn to our variant of the macro-factor from Ludvigson and Ng (2009) in rows (c). LN_t is able to explain between 15% and 18% of the one-year ahead variation in bond risk premia, where the largest proportion is explained for the tree- and four-year bonds. Note that these results are obtained using the latest available vintages of ALFRED data.¹⁴ An important question to ask at this point is whether expected business conditions contain information about bond risk premia that is distinguishably different from that contained in current business conditions. To answer this, we consider a two-factor specification in which we augment the LN_t model with ME_t in rows denoted (d). As it turns out, both are individually strongly significant for all bond maturities, which implies that they capture quite distinct aspects of the potentially high-dimensional set of risks that governs the time-variations in bond risk premia. The effect of augmenting the predictive regression with ME_t is economically important as well with increases in the explanatory power of about eleven percentage point across the maturity spectrum.

As a last step, we first consider a two-factor specification using CP_t and LN_t and, secondly, a three-factor specification that investigates the effects of augmenting the model with ME_t . Analogously to Ludvigson and Ng (2009), we find CP_t and LN_t to contain complementary information as both factors remain significant and jointly produce adjusted R^2 values larger than their individual values for all maturities. Augmenting the model with ME_t results in increases in adjusted R^2 s of about four percentage points on average across the maturity spectrum and ME_t remains significant at the 5% level for most maturities.

¹⁴In our case, this means the November 2014 vintages.

Summing up the results so far, we find that while both current term structure information and current business conditions, in the form of CP_t and LN_t , respectively, are important for explaining the variations in one-year ahead bond risk premia, there is strong evidence for a non-negligible role for expected business conditions as a complementary source of predictive information, implying that expectations to the future course of the real economy directly influence the risk premia demanded by investors for holding risky Treasury bonds.

5.3 Out-of-sample forecasting and asset allocation

In this section, we evaluate the ability of the bond return prediction models from Section 5.2 to accurately forecast bond risk premia in an out-of-sample setting using information available at the time of the forecast only. The forecasts are generated recursively using an expanding window of observations, where the period 1968:Q4-1989:Q4 constitutes the initial estimation period and 1990:Q1-2014:Q4 the forecast evaluation period.¹⁵ As in the in-sample analysis, we set $n = 2, 3, 4, 5$ and predict bond risk premia for each bond maturity separately. To mimic a real-time forecasting exercise, we estimate model parameters and forecasting factors recursively prior to each forecast using information restricted to match the information set available at time t . Importantly, we address issues with publication lags and data revisions in macroeconomic time series directly by using the vintages of data available at the time of the forecast only in the construction of our variant of the LN_t factor.

5.3.1 Statistical significance

As a first way to gauge the value of bond return prediction models, we consider a statistical evaluation of the predictive accuracy of the out-of-sample forecasts relative to a recursively updated expectations hypothesis (EH) benchmark computed as a recursively updated projection of bond excess returns upon a constant. We consider two measures of statistical significance well-known to the literature. First, for each bond maturity and model, we follow [Campbell and Thompson \(2008\)](#) and provide an out-of-sample R^2 relative to the EH benchmark model given as

$$R_{\text{oos}}^2 = 1 - \frac{\sum_{t=1}^P \left(rx_{t+4}^{(n)} - \widehat{rx}_{t+4,i}^{(n)} \right)^2}{\sum_{t=1}^P \left(rx_{t+4}^{(n)} - \widehat{rx}_{t+4,EH}^{(n)} \right)^2} \quad (9)$$

where $\widehat{rx}_{t+4,i}^{(n)}$ and $\widehat{rx}_{t+4,EH}^{(n)}$ denote the forecast from the i th candidate model and the EH benchmark model, respectively, and P is the number of out-of-sample forecasts. The R_{oos}^2

¹⁵The use of an expanding window of observations seems to be the prevailing standard in the literature ([Ludvigson and Ng, 2009](#); [Thornton and Valente, 2012](#); [Gargano et al., 2014](#)).

in (9) can be understood as the ratio of mean squared prediction errors (MSPE) for the i th candidate model relative to, in this case, the EH benchmark model. A $R_{\text{OOS}}^2 > 0$ implies that the MSPE of the i th candidate model is lower than that of the EH benchmark model, indicating a higher predictive accuracy.

While the R_{OOS}^2 metric measures the relative improvement in MSPE for the i th bond excess return forecasting model over the EH model, it does not tell us whether these differences are large in a statistical sense. As a result, we follow, among others, [Rapach and Zhou \(2013\)](#), [Gargano et al. \(2014\)](#), and [Andreasen et al. \(2015\)](#) and conduct more formal tests for the differences in predictive accuracy using the [Clark and West \(2007\)](#) test for equal predictive accuracy. In this setup, we interpret the EH model as a no predictability benchmark and test the null of no predictability ($R_{\text{OOS}}^2 \leq 0$) against the one-sided alternative of predictability by the i th candidate forecasting model ($R_{\text{OOS}}^2 > 0$). The [Clark and West \(2007\)](#) test is ideally suited for an environment like ours, where the competing forecast is generated by a nested model.¹⁶ The test is computed by first defining

$$\psi_{t+4} = \left(rx_{t+4}^{(n)} - \widehat{rx}_{t+4,EH}^{(n)} \right)^2 - \left[\left(rx_{t+4}^{(n)} - \widehat{rx}_{t+4,i}^{(n)} \right)^2 - \left(\widehat{rx}_{t+4,EH}^{(n)} - \widehat{rx}_{t+4,i}^{(n)} \right)^2 \right] \quad (10)$$

where ψ_{t+4} can be interpreted as a bias adjusted difference in MSPE between the EH model and the i th candidate model, where the adjustment takes into account the upward bias in MSPE induced by having to estimate parameters that are zero under the null of no predictability. Carrying out the test amounts to projecting ψ_{t+4} upon a constant and evaluating the t-statistic using the one-sided upper-tail p-value from a normal distribution, where the t-statistic is computed using [Newey and West \(1987\)](#) standard errors implemented with six lags.

[Table 6 around here]

The results from the statistical evaluation of the models against the EH benchmark are presented Panel A of Table 6. We see in the first column that CP_t consistently delivers a negative R_{OOS}^2 in the range of -16.26% to -20.27% over the spectrum of bond maturities, signaling a forecasting performance inferior to the simple EH benchmark of constant expected returns. This is corroborated by the failure to reject the null of no predictability for all maturities. The results for our real-time variant of the LN_t factor is presented in the second column. We see, even when taking explicitly into account issues with publication lags and data revisions using vintage data corresponding to the time of the forecasts, that LN_t is able to forecast bond risk premia more accurately than the EH model. The out-of-sample R^2 implies statistically significant reductions in MSPE of about 8-9% on average. This stands in contrast to the general conclusion in [Ghysels et al. \(2014\)](#) and shows that past macroeconomic variables do contain important information about future

¹⁶As mentioned in Section 2, the forecasting models reduce to the EH model for $\beta = 0$.

bond risk premia, at least in a statistical sense. Considering a two-factor model including CP_t and LN_t results in a positive and statistically significant R_{oos}^2 value for the two-year bond, but negative R_{oos}^2 values for the remainder of the maturity spectrum.

Our main interest centers on the out-of-sample predictive ability of expected business conditions relative to both a recursively updated EH model and the existing forecasting factors. Starting with the first point, we see that ME_t performs poorly at the very short end of the maturity spectrum, where it realizes a negative R_{oos}^2 of -7.17%. It does, however, produce highly positive and significant MSPE reductions for the remainder of the maturity spectrum from 2.08% to 13.08%. Augmenting the CP_t specification with ME_t results in positive and significant R_{oos}^2 values for all bond maturities ranging from a low of 6.53% for the two-year bond to a high of 16.12% for the five-year bond. Consequently, ME_t and CP_t appear to contain complementary information that results in significant out-of-sample forecasting gains. Similarly, augmenting the LN_t specification with ME_t results in statistically significant, and mostly positive, improvements over the EH model, although the LN_t model performs better on its own for the two and three-year bonds.

Lastly, we consider a three-factor specification that includes all forecasting factors. In this case, we see positive and highly significant R_{oos}^2 values across the entire bond maturity spectrum. The R_{oos}^2 ranges from a low of 12.31% for the two-year bond to a high of 17.86% for the five-year bond, which are the highest attained values for any maturity. This sustains the real-time importance of ME_t as a reliable source of information to consider when forecasting bond excess returns out-of-sample, especially for the longer maturity bonds. It also supports the notion that all three factors contain independent information important for correctly assessing one-year ahead variations in bond risk premia.

[Table 7 around here]

In Table 7, we consider the results from testing nested excess bond return prediction models against each other to verify, in a statistical sense, the observed patterns from Panel A in Table 6. In the first row of each panel, we test whether augmenting the CP_t model with ME_t results in significant improvements in predictive accuracy. This is indeed the case with high certainty for all maturities with R_{oos}^2 values in the range of 19-30% (p-values are virtually zero). As mentioned above, we see that augmenting the LN_t model with ME_t only leads to significant improvements for the four- and five-year bonds, whereas LN_t is better on its own at the shorter end. Lastly, we note that augmenting the two-factor CP_t and LN_t specification with ME_t for all maturities results in improved forecasting performance with positive and significant R_{oos}^2 values in the range of about 12-20% (p-values of 0.02 or less).

As a supplement to the results in Table 6, we follow, among others, [Goyal and Welch \(2008\)](#) and [Rapach and Zhou \(2013\)](#) and compute and plot in Figure 3 the differences in cumulative squared prediction errors (DCSPE) between the EH benchmark and the i th

forecasting model given by

$$\text{DCSPE}_{i,t} = \sum_{t=1}^P \left(rx_{t+4}^{(n)} - \widehat{rx}_{t+4,EH}^{(n)} \right)^2 - \sum_{t=1}^P \left(rx_{t+4}^{(n)} - \widehat{rx}_{t+4,i}^{(n)} \right)^2 \quad (11)$$

for $t = 1, \dots, P$, with P being the number of out-of-sample forecasts. The purpose of this exercise is to assess the consistency of the model's forecasting performance over time. That is, to check that forecasting gains are not solely derived from a specific event or period of time.

[Figure 3 around here]

The graphical output in Figure 3 is largely supportive of the findings so far. In particular, we note that the predictive performance of ME_t , and the models in which it appears, tends to increase steadily over time, implying consistent improvements in forecasting accuracy over the EH model. Notably, this suggests that the predictive ability is not derived from a particular event or time period, although clear improvements are visible around recession periods. A similar pattern is observed for the LN_t factor. The CP_t factor, on the other hand, while doing well in the beginning of the sample, sees a sharp decline in predictive performance during the latest part of the sample, which is especially pronounced for the longer maturity bonds.

5.3.2 Economic significance

The empirical analysis has insofar concentrated on statistical measures of predictive accuracy. An investor, however, is likely to care more about the portfolio performance attainable from using return prediction models to guide investment decisions. While empirical studies frequently find statistical support for predictability, they fail to document economic value to a mean-variance investor that trades in the Treasury bond market (Thornton and Valente, 2012; Sarno et al., 2014).

To examine the economic value of using expected business conditions, we consider the asset allocation problem of an investor with mean-variance preferences that chooses the weight $\omega_t^{(n)}$ to invest in a risky bond with n -years to maturity versus the one-year risk-free yield $y_t^{(1)}$.¹⁷ The resulting portfolio return $r_{p,t+4}^{(n)}$ is given by

$$r_{p,t+4}^{(n)} = y_t^{(1)} + \omega_t^{(n)} rx_{t+4}^{(n)} \quad (12)$$

where $y_t^{(1)}$ and $rx_{t+4}^{(n)}$ are defined in (1) and (3), respectively. In order to determine the optimal weights, we assume that the mean-variance investor trades off expected returns

¹⁷See, among others, Campbell and Thompson (2008), Dangl and Halling (2012), Thornton and Valente (2012), Rapach and Zhou (2013), Gargano et al. (2014), and Sarno et al. (2014), for similar analyses.

and risk by solving the expected utility maximization problem

$$\max_{\omega_t^{(n)}} E_t \left[r_{\mathcal{P},t+4}^{(n)} \right] - \frac{\varphi}{2} \text{Var}_t \left[r_{\mathcal{P},t+4}^{(n)} \right] \quad (13)$$

where φ measures the investor's level of relative risk aversion, which we set equal to ten. While this value is somewhat higher than what is normally considered in empirical studies, this choice, as in [Gargano et al. \(2014\)](#), reflects the fact that lower values of φ causes the weights placed on the risky bonds to hit the upper bounds of our investment constraints more often than not for the EH benchmark and most candidate models for the two-year bond, which makes it hard to reliably distinguish the competing hypotheses.¹⁸ The solution to the maximization problem in (13) delivers the following weight on the risky n -year bond

$$\omega_t^{(n)} = \frac{1}{\varphi} \frac{E_t \left[r_{t+4}^{(n)} \right]}{\text{Var}_t \left[r_{t+4}^{(n)} \right]} \quad (14)$$

where $E_t \left[r_{t+4}^{(n)} \right]$ and $\text{Var}_t \left[r_{t+4}^{(n)} \right]$ denote the conditional expectation and variance of $r_{t+4}^{(n)}$, respectively. For each $t \in 1, \dots, P$ in the out-of-sample period, we estimate the conditional mean using either the i th bond excess return prediction model or the EH no-predictability benchmark model. The conditional variance is computed using a rolling window estimator over the past twenty years of observations. To prevent the investor from taking extreme positions, we restrict the weight $\omega_t^{(n)}$ on the risky bond to lie in the interval $[0, 1.5]$. This corresponds to imposing a maximum leverage of 50% and no short-selling of the risky bond.

The average utility realized by the investor from using the sequence of realized weights $\hat{w}_t^{(n)}$, $t = 1, \dots, P$, is given by

$$\bar{U}^{(n)} = P^{-1} \sum_{t=1}^P \left[r_{\mathcal{P},t+4}^{(n)} - \frac{\varphi}{2} \left(r_{\mathcal{P},t+4}^{(n)} - \mu_{\mathcal{P}}^{(n)} \right)^2 \right] \quad (15)$$

where $\mu_{\mathcal{P}}^{(n)}$ denotes the average portfolio return attained by following the trading strategy. For each excess return prediction model, we compute the annualized percentage utility gain, denoted Δ (%), relative to the simple EH no-predictability benchmark. This average utility gain, or certainty equivalent return, can be interpreted as the portfolio management fee (in annualized percentage returns) that an investor would be willing to pay in order to switch from the EH model to the i th predictive model. As a way to test statistically the differences in utility gains, we compute [Diebold and Mariano \(1995\)](#) t-tests for the equality of realized utilities from the i th predictive regression model and the EH model.

¹⁸[Campbell and Thompson \(2008\)](#), [Thornton and Valente \(2012\)](#), and [Rapach and Zhou \(2013\)](#) use a coefficient of relative risk aversion equal to five, whereas [Sarno et al. \(2014\)](#) use a value of 3.

A similar approach is used in [Gargano et al. \(2014\)](#). To construct the test, we compute time series of realized utilities using (15) and test whether the difference between the EH benchmark and the i th candidate model is distinguishably different from zero.¹⁹

Another frequently used measure to assess portfolio performance is the Sharpe ratio (SR). Alas, despite the wide use of the SR, it is subject to a series of drawbacks. Most importantly, it can be manipulated in various ways ([Goetzmann, Ingersoll, Spiegel, and Welch, 2007](#)). We therefore follow [Thornton and Valente \(2012\)](#) and report a manipulation-proof performance measure for the i th predictor model relative to the EH model defined as

$$\Theta = \frac{1}{1-\varphi} \left[\ln \left(P^{-1} \sum_{t=1}^P \left[\frac{R_{p,t+4,i}^{(n)}}{1+y_{t+4}^{(1)}} \right]^{1-\varphi} \right) - \ln \left(P^{-1} \sum_{t=1}^P \left[\frac{R_{p,t+4,EH}^{(n)}}{1+y_{t+4}^{(1)}} \right]^{1-\varphi} \right) \right] \quad (16)$$

where $R_{p,t+4,i}^{(n)}$ denotes the gross portfolio return attained from following the strategy implied by the i th forecasting model and $R_{p,t+4,EH}^{(n)}$ the analogous gross portfolio return for the EH model.²⁰ We report the annualized percentage measure Θ (%) for each candidate model and bond maturity.

Table 6 contains the results from the economic evaluation of the forecasting models. Panel B presents the results for the average annualized percentage utility gains, Δ (%), attained relative to the EH benchmark and Panel C presents results for the manipulation-proof performance measure, Θ (%). Guiding investment decisions using CP_t results in negative utility gains across the entire maturity spectrum ranging from -0.84 to -1.95, implying that CP_t does not provide any economic benefit over simply using a constant expected return forecast. This echoes the results from [Thornton and Valente \(2012\)](#), the statistical evaluation herein, and the manipulation-proof performance measure, Θ (%), in Panel C, which takes on negative values for the entire maturity spectrum as well. The LN_t factor delivers negative utility gains for the two-year bond, but positive gains from 0.43 to 1.93 for the remainder of the maturity spectrum, where the gains are significant for the three- and four-year bonds. Θ (%) tells a similar story. This supports the argument in [Thornton and Valente \(2012\)](#) that statistical gains does not mechanically translate into economic gains. Considering a joint two-factor specification of LN_t and CP_t results in negative utility gains for all maturities. The same is true for Θ (%), except for the five-year bond, where it is just on the positive side.

Expected business conditions, on the other hand, generate positive utility gains ranging monotonically from 0.32 for the two-year bond to 3.85 for the five-year bond. These gains are significant at the 5% level or less for all maturities. Moreover, despite the statistical

¹⁹To be more specific, we estimate the regression $u_{i,t}^{(n)} - u_{EH,t}^{(n)} = \delta^{(n)} + \varepsilon_t^{(n)}$, where $u_{i,t}^{(n)}$ is based on (15) and test whether $\delta^{(n)}$ is equal to zero using a HAC standard error estimator ([Gargano et al., 2014](#)).

²⁰We refer to [Goetzmann et al. \(2007\)](#) for detailed information about the performance metric.

evidence suggesting a weaker forecasting performance at the short end of the maturity spectrum relative to LN_t , we see that the economic gains from using ME_t are higher for all maturities. Similar to the statistical evidence, however, we see that the largest utility gains are attained at the long end of the maturity spectrum. These results are mirrored by Θ (%), which increases monotonically from 0.47 to 4.01. Augmenting the CP_t model with ME_t yields positive utility gains ranging from 0.16 to 2.38 for the three- to five-year bonds, although the gains are only significant for the four- and five-year bonds. A similar story is told by the Θ (%) measure, which ranges from -0.27 to 2.51. On the other hand, we see that the two-factor specification including ME_t and the real-time LN_t factor delivers positive utility gains for all maturities of a magnitude slightly larger than for ME_t alone. Δ (%) ranges from a low of 0.31 for the two-year bond to a high of 3.95 for the five-year bond and the average utility gains are all significant at a 5% level or lower. As for the statistical evaluation, we see the largest gains at the longer end of the maturity spectrum. Interestingly, this makes a compelling case for trading based on the current and expected state of the real economy. Finally, while the three-factor specification delivered the highest statistical significance, it fails to provide superior economic value.

In the remaining columns of Table 7, we present the results from computing realized utility gains and risk-adjusted returns from augmenting the existing forecasting models from the literature with our measure of expected business conditions. In contrast to the statistical evaluations, we see universal improvements from adding ME_t as a covariate in all cases. As a whole, the utility gains range from a low of 0.23 to a high of 3.95 across model specifications and bond maturities, where almost all the gains are significant at a 10% level or lower. The same story is told by Θ (%), which ranges between 0.27 and 3.58. Thus, from an economic perspective, there is ample evidence supporting the use of expected business conditions to guide portfolio decisions. In fact, an investor's economic utility could have been significantly improved in all instances had the investor exploited the information in expected business conditions.

[Figure 4 around here]

Figure 4 plots the cumulative differences between the realized utility from relying on the i th forecasting model and the EH benchmark, respectively, to gauge whether utility gains are realized consistently over the out-of-sample period or attributable to specific periods in time only. We compute time series of realized utilities as in (15) and the cumulative differences analogously to (11). The figure, for most parts, mirrors the results from Figure 3. Most utility gains do not seem to be driven by a specific period in time, i.e. most slopes are on average in the same direction. In particular, models including ME_t seem to deliver particularly stable utility gains over the out-of-sample forecasting period. The exception being CP_t , which fails to provide useful signals during and after the latest financial crisis. We return briefly to this point in Section 5.5. Overall, both

the statistical and the economic evaluations of the out-of-sample forecasting performance suggest that ME_t contains information useful for forecasting bond risk premia and making asset allocation decisions in real-time. In particular, not only does ME_t predict excess bond returns on its own, but combining it with existing forecasting factors results in increased statistical significance, particular at the long end of the maturity spectrum, and universal improvements in portfolio performance. As such, this suggests that not only do we need to look beyond the term structure to fully understand bond risk premia as suggested in [Ludvigson and Ng \(2009\)](#), [Cooper and Priestley \(2009\)](#), [Duffee \(2011\)](#), and [Joslin et al. \(2014\)](#), we also need to look beyond current and past values of macroeconomic fundamentals.

5.4 Model combinations

Since the seminal paper of [Bates and Granger \(1969\)](#), it has been widely acknowledged that combining forecasts across models may produce a forecast that performs better than any of the individual forecasts ([Rapach, Strauss, and Zhou, 2010](#); [Dangl and Halling, 2012](#)). [Timmermann \(2006\)](#) points out that forecast combination can be viewed as a diversification strategy in the sense that combined forecasts are less volatile than individual forecasts, thereby dealing with model uncertainty and parameter instability in a simple manner. Model uncertainty would as such not be a concern if all models were able to beat the EH benchmark consistently. However, this is not the case as the evaluations in [Section 5.3](#) revealed significant variation in model performances. To investigate whether an investor could improve upon the individual models by combining forecasts, we consider two different combination schemes. The first scheme diversifies across the three forecasts generated by CP_t and LN_t models, whereas the second scheme combines the forecasts from all seven models considered in the in-sample and out-of-sample analyses. For both cases, we consider the simple equal-weighted combination of forecasts.

[Table 8 around here]

The results from the forecast combination schemes are presented in [Table 8](#). The first combination scheme seems to mirror the impressions from the individual model evaluations ([Panel A](#)). While the reductions in MSPE relative to the EH benchmark are statistically significant at a 6% significance level or lower for all maturities, we see that both Δ (%) and Θ (%) turns out negative, with the exception of Θ (%) for the five-year bond. In the second combination scheme, where we combine forecasts from all seven models, we again see positive and statistically significant R_{oos}^2 values ([Panel B](#)). The MSPE reductions, which range from 15.81% to 19.60%, are of sizable magnitude and, in fact, surpasses all individual models from [Table 6](#). Unlike the first scheme, the second scheme delivers positive utility gains with Δ (%) ranging from 0.13 to 2.11 and Θ (%) ranging from 0.16

to 2.17. The utility gains are significant for the four- and five-year bonds. As was the case for the individual model evaluations, we see the largest gains, both statistically and economically, being concentrated in the longer end of the bond maturity spectrum.

5.5 Links to the real economy

Recent studies argue that return predictability itself may be time-varying and not necessarily present in all states of the economy (Henkel et al., 2011; Dangl and Halling, 2012; Rapach and Zhou, 2013; Møller and Sander, 2014; Andreasen et al., 2015). A standard way to assess state-dependent predictability is to split the out-of-sample forecast errors, ex post, into periods of expansions and recessions using the NBER recession indicator. However, due to the use of annual excess holding period returns and the quarterly data frequency, such an approach seems infeasible as it would be difficult to disentangle the effects for holding periods returns realized over multiple economic states. Examining how forecast performance is related to the real economy, however, still seems a worthwhile exercise. To do so, we consider the contemporaneous correlations between the differences in cumulative squared prediction errors and realized utilities, respectively, and the CFNAI in Panels A and B of Table 9. We start by noticing that correlations vary markedly across forecasting models. In particular, the forecasting performance of CP_t relative to the EH benchmark appears to be positively correlated with the CFNAI, indicating that CP_t performs well during good times when the CFNAI is high and poorly during recessions when CFNAI is low. This is consistent with the findings in Andreasen et al. (2015), who document that bond risk premia are only predictable by term structure variables in expansion periods.²¹ In fact, all models containing CP_t tend to perform poorly in recessions relative to the EH benchmark model, something that is also visible in Figure 3. The relative forecast performance for the remaining models, which are based on LN_t , ME_t , or a combination of the two, is negatively correlated with the CFNAI, implying that these models perform comparatively better during periods of economic recession in which the CFNAI is low. Notice that this suggests different roles for term structure variables and business conditions in expansions vis-a-vis recessions. A similar picture is observed for realized utilities in Panel B, although all correlations are positive for the two-year bond.

[Table 9 around here]

To provide an informal investigation of the forecasting mechanisms of the models, we consider the correlation between the out-of-sample forecasts and the CFNAI and macroeconomic uncertainty ($\mathbb{U}_t^{\text{Macro}}$) in Panels C and D, respectively, where we measure macroeconomic uncertainty using the index constructed in Jurado, Ludvigson, and Ng

²¹Note that this is the opposite pattern found for stock returns in, among other, Henkel et al. (2011), Dangl and Halling (2012), Rapach and Zhou (2013), and Møller and Sander (2014).

(2015).²² Starting with Panel C, we see that almost all model forecasts are negatively correlated with the CFNAI, which is consistent with risk-based explanations for time-variations in bond risk premia. That is, the models predict high (low) excess return when economic conditions are bad (good). The exception being the CP_t model whose predictions are, albeit the magnitudes are small, on average positively correlated with the real economy, accounting for part of its poor out-of-sample performance. Looking at the correlation between the out-of-sample forecasts and macro uncertainty in Panel D, we see that the traditional models are all negatively correlated with U_t^{Macro} , implying that bond risk premia are predicted to be high when macroeconomic uncertainty is low. ME_t , on the other hand, is highly positively correlated with U_t^{Macro} , indicating that expected business conditions predicts bond risk premia to be high when macro uncertainty is high as well. The same is true for the two-factor model including LN_t and ME_t , whereas the remaining models are either unrelated or negatively correlated with macroeconomic uncertainty.

5.6 Forecasting quarterly excess bond returns

As a complement, and robustness check, to the main empirical results presented so far for one-year ahead bond risk premia, this section follows a recent strand of the literature that studies bond excess returns over shorter holding periods (Gargano et al., 2014; Andreasen et al., 2015). To be specific, we consider quarterly excess bond returns, which is the shortest holding period permitted by the frequency of the SPF responses, and explore whether the main conclusions are robust to a change in the holding period length.

To obtain quarterly bond risk premia, we start by reconstructing the yield curve at a daily frequency relying on the methods developed in Nelson and Siegel (1987) and Svensson (1994) and the model parameters estimated in Gürkaynak, Sack, and Wright (2007). The time t log yield of an n -year zero-coupon Treasury bond is, in this setup, given by

$$y_\tau^{(n)} = \beta_0 + \beta_1 \frac{1 - \exp\left(-\frac{n}{\kappa_1}\right)}{\frac{n}{\kappa_1}} + \beta_2 \left[\frac{1 - \exp\left(-\frac{n}{\kappa_1}\right)}{\frac{n}{\kappa_1}} - \exp\left(-\frac{n}{\kappa_1}\right) \right] + \beta_3 \left[\frac{1 - \exp\left(-\frac{n}{\kappa_2}\right)}{\frac{n}{\kappa_2}} - \exp\left(-\frac{n}{\kappa_2}\right) \right] \quad (17)$$

where τ denotes the daily time index. The parameters $(\beta_0, \beta_1, \beta_2, \beta_3, \kappa_1, \kappa_2)$ are provided by Gürkaynak et al. (2007), who report estimates of the yield curve from June 1961 and onwards at a daily frequency for the entire maturity spectrum spanned by outstanding Treasury securities.²³ The authors rely on the Nelson and Siegel (1987) specification for the

²²The macro uncertainty index is available at <http://www.econ.nyu.edu/user/ludvigsons/>.

²³The data is available at <http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html>.

period up until 1980, i.e. restrict β_3 to equal zero, to obtain a better identified yield curve without compromising fit (Gürkaynak et al., 2007, p. 2297). To keep results comparable, we focus on bonds with the same maturities as above. Yields are converted to a quarterly frequency by sampling from the last day of the second month in each quarter. Quarterly bond risk premia are then computed analogously to (1)-(3) in Section 2 by buying an n -year bond today and selling it after one quarter as an $n - 1/4$ year bond. The risk-free rate is a three-month Treasury bill (Secondary Market Rate).

[Table 10 around here]

Table 10 presents full sample results from regressing one-step ahead quarterly bond risk premia on a subset of the model specifications from Table 5. We reconstruct all predictors using the quarterly bond price data analogously to their annual holding period return counterparts and add the superscripted suffix *GSW* to avoid confusion. For expected business conditions, we rely on one-quarter ahead survey forecasts only to match the holding period. The results in Table 10 largely mirror the ones from Sections 3 and 5.2. In particular, we see that ME_t^{GSW} is highly significant across the maturity spectrum and, by itself, is able to explain between 4% and 6% of the variation in one-quarter ahead bond risk premia.²⁴ Running horse races against the standard factors, CP_t^{GSW} and LN_t^{GSW} , reveals that ME_t^{GSW} contains relevant information about future risk premia not already captured by the current term structure or business environment. In all cases, we see that all factors are highly significant and that the addition of ME_t increases the explanatory power, supporting the conclusion that ME_t can be viewed as a complementary source of information for explaining one-step ahead variations in bond risk premia.

6 Concluding remarks

This paper contributes to the literature that examines the source of variations in the risk premia required by investors for holding US Treasury bonds. While existing studies mainly rely on information in the current term structure and business environment, our study takes a forward-looking perspective by constructing a real-time proxy for expected business conditions using survey forecasts from the *Survey of Professional Forecasters* (SPF).

Our findings indeed suggest that a sizable fraction of the risk premia on two- to five-year zero-coupon US Treasury bonds can be accounted for by expected business conditions. Consistent with recent research, we find that the majority of the predictive information lies outside the span of contemporaneous yields. Moreover, direct comparisons with the Cochrane and Piazzesi (2005) and Ludvigson and Ng (2009) factors reveal a non-negligible role for expected business conditions in explaining one-year ahead variations in bond risk

²⁴The R^2 values are naturally lower in this setting due to the shorter holding periods, see also Gargano et al. (2014).

premia. As a result, we view expected business conditions as a complementary source of information not already embedded in the standard predictors. Our results are confirmed in a real-time out-of-sample exercise, where the predictive accuracy of expected business conditions relative to the existing predictors is evaluated using both conventional statistical measures and from the perspective of a mean-variance investor. Our results suggest that expected business conditions are able to generate significantly better forecasts than a recursively updated EH model and that augmenting existing models with expected business conditions improves forecasting performance in most cases. Finally, we find that expected business conditions significantly improve upon an investor's economic utility when used to guide real-time investment decisions.

References

- Almeida, C., Graveline, J. J., Joslin, S., 2011. Do interest rate options contain information about excess returns? *Journal of Econometrics* 164.
- Altavilla, C., Giacomini, R., Constantini, R., 2014. Bond returns and market expectations. *Journal of Financial Econometrics* 12 (4), 708–729.
- Amato, J. D., Swanson, N. R., 2001. The real-time predictive content of money for output. *Journal of Monetary Economics* 48, 3–24.
- Andersen, T. G., Benzoni, L., 2010. Do bonds span volatility risk in the U.S. treasury market? A specification test for affine term structure models. *Journal of Finance* 65 (2), 603–653.
- Andreasen, M. M., Engsted, T., Møller, S. V., Sander, M., 2015. Bond return predictability in expansions and recessions, Working Paper.
- Ang, A., Bekaert, G., Wei, M., 2007a. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics* 54, 1163–1212.
- Ang, A., Dong, S., Piazzesi, M., 2007b. No-arbitrage Taylor rules, Working Paper, Columbia University.
- Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of the term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50, 745–787.
- Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. *Econometrica* 70 (1), 135–172.
- Bai, J., Ng, S., 2008. Forecasting economic time series using targeted predictors. *Journal of Econometrics* 146, 304–317.
- Bansal, R., Shaliastovich, I., 2013. A long-run risks explanation of predictability puzzles in bond and currency markets. *The Review of Financial Studies* 26 (1), 1–33.
- Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance* 59 (4), 1481–1509.
- Barro, R. J., 1990. The stock market and investment. *Review of Financial Studies* 3, 115–131.
- Bates, J. M., Granger, C. W. J., 1969. The combination of forecasts. *Operational Research Quarterly* 20 (4), 451–468.
- Bikbov, R., Chernov, M., 2010. No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics* 159, 166–182.
- Buraschi, A., Jiltsov, A., 2007. Habit formation and macroeconomic models of the term structure of interest rates. *The Journal of Finance* 62 (6), 3009–3063.
- Campbell, J. Y., Ammer, J., 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns. *The Journal of Finance* 48 (1), 3–37.

- Campbell, J. Y., Cochrane, J. H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *The Journal of Political Economy* 107, 205–251.
- Campbell, J. Y., Shiller, R. J., 1991. Yield spreads and interest rate movements: A bird's eye view. *The Review of Economic Studies* 58 (3), 495–514.
- Campbell, J. Y., Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies* 21 (4), 1509–1531.
- Campbell, S. D., Diebold, F. X., 2009. Stock returns and expected business conditions: Half a century of direct evidence. *Journal of Business and Economic Statistics* 27 (2), 266–278.
- Capistrán, C., Timmermann, A., 2009. Forecast combination with entry and exist of experts. *Journal of Business & Economic Statistics* 27 (4), 428–440.
- Chen, N.-F., Roll, R., Ross, S. A., 1986. Economic forces and the stock market. *Journal of Business* 59 (3), 383–403.
- Chernov, M., Mueller, P., 2012. The term structure of inflation expectations. *Journal of Financial Economics* 106, 367–394.
- Chun, A. L., 2011. Expectations, bond yields, and monetary policy. *The Review of Financial Studies* 24 (1), 208–247.
- Cieslak, A., Povala, P., 2015. Expected returns in treasury bonds. *Review of Financial Studies* 28 (10), 2859–2901.
- Clark, T. E., West, K. D., 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138, 291–311.
- Cochrane, J. H., Piazzesi, M., 2005. Bond risk premia. *American Economic Review* 95 (1), 138–160.
- Cooper, I., Priestley, R., 2009. Time-varying risk premiums and the output gap. *The Review of Financial Studies* 22 (7), 2801–283.
- Croushore, D., Stark, T., 2001. A real-time data set for macroeconomists. *Journal of Econometrics* 105, 111–130.
- Dahlquist, M., Hasseltoft, H., 2013. International bond risk premia. *Journal of International Economics* 90, 17–32.
- Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. *Journal of Financial Economics* 106, 157–181.
- Dick, C. D., Schmeling, M., Schrimpf, A., 2013. Macro-expectations, aggregate uncertainty, and expected term premia. *European Economic Review* 58, 58–80.
- Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–265.
- Duffee, G. R., 2011. Information in (and not in) the term structure. *The Review of Financial Studies* 24 (9), 2895–2934.

- Fama, E., 1990. Stock returns, expected returns, and real activity. *The Journal of Finance* 45, 1089–1108.
- Fama, E., 2006. The behavior of interest rates. *Review of Financial Studies* 19 (2), 359–379.
- Fama, E., French, K., 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25, 23–49.
- Fama, E., Schwert, G. W., 1977. Asset returns and inflation. *Journal of Financial Economics* 5, 115–146.
- Fama, E. F., Bliss, R. R., 1987. The information in long-maturity forward rates. *The American Economic Review* 77 (4), 680–692.
- Fama, E. F., Gibbons, M. R., 1984. A comparison of inflation forecasts. *Journal of Monetary Economics* 13, 327–348.
- Favero, C. A., Niu, L., Sala, L., 2012. Term structure forecasting: No-arbitrage restriction versus large information set. *Journal of Forecasting* 31, 124–156.
- Ferson, W. E., Harvey, C. R., 1991. The variation of economic risk premiums. *Journal of Political Economy* 99 (2), 385–415.
- Froot, K. A., 1989. New hope for the expectations hypothesis of the term structure of interest rates. *Journal of Finance* 44 (2), 283–305.
- Gargano, A., Pettenuzzo, D., Timmermann, A., 2014. Bond return predictability: Economic value and links to the macroeconomy, Working Paper.
- Ghysels, E., Horan, C., Moench, E., 2014. Forecasting through the rear-view mirror: Data revisions and bond return predictability, Federal Reserve Bank of New York Staff Reports No. 581.
- Goetzmann, W., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures. *The Review of Financial Studies* 20 (5), 1503–1546.
- Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies* 21 (4), 1455–1508.
- Gürkaynak, R. S., Sack, B., Wright, J. H., 2007. The U.S. treasury yield curve: 1961 to the present. *Journal of Monetary Economics* 54 (8), 2291–2304.
- Hansen, L. P., Hodrick, R. J., 1980. Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *The Journal of Political Economy* 88 (5), 829–853.
- Henkel, S. J., Martin, J. S., Nardari, F., 2011. Time-varying short-horizon predictability. *Journal of Financial Economics* 99, 560–580.
- Ilmanen, A., 1995. Time-varying expected returns in international bond markets. *Journal of Finance* 50 (2), 481–506.
- Inoue, A., Kilian, L., 2005. In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews* 23 (4), 371–402.
- Joslin, S., Pribsch, M., Singleton, K. J., 2014. Risk premiums in dynamic term structure models with unspanned macro risks. *The Journal of Finance* 69 (3), 1197–1233.

- Jurado, K., Ludvigson, S. C., Ng, S., 2015. Measuring uncertainty. *American Economic Review* 105 (3), 1177–1216.
- Keim, D. B., Stambaugh, R. F., 1986. Predicting returns in the stock and bond markets. *Journal of Financial Economics* 17, 357–390.
- Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1 (1), 54–61.
- Ludvigson, S. C., Ng, S., 2009. Macro factors in bond risk premia. *The Review of Financial Studies* 22, 5027–5067.
- McCracken, M. W., Ng, S., 2015. FRED-MD: A monthly database for macroeconomic research, Federal Reserve Bank of St. Louis Working Paper Series.
- Merton, R. C., 1973. An intertemporal capital asset pricing model. *Econometrica* 41 (5), 867–887.
- Moench, E., 2008. Forecasting the yield curve and a data-rich environment: A no-arbitrage factor-augmented VAR approach. *Journal of Econometrics* 146, 26–43.
- Møller, S. V., Sander, M., 2014. Stock return and dividend growth predictability across the business cycle, Working Paper.
- Nelson, C. R., Siegel, A. F., 1987. Parsimonious modeling of yield curves. *Journal of Business* 60 (4), 473–489.
- Newey, W. K., West, K. D., 1987. A simple positive-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (3), 703–708.
- Ottaviani, M., Sørensen, P. N., 2006. The strategy of professional forecasting. *Journal of Financial Economics* 81, 441–466.
- Pesaran, M. H., Timmermann, A., 2005. Real time econometrics. *Econometric Theory* 21, 212–231.
- Piazzesi, M., Salomao, J., Schneider, M., 2013. Trend and cycle in bond premia, Working Paper.
- Rapach, D., Strauss, J. K., Zhou, G., 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies* 23 (2), 821–862.
- Rapach, D., Zhou, G., 2013. Forecasting stock returns. In: Elliott, G., Timmermann, A. (Eds.), *Handbook of Economic Forecasting*. Vol. 2, Part A of *Handbook of Economic Forecasting*. Elsevier, Ch. 6, pp. 328–383.
- Rossi, A. G., Timmermann, A., 2015. Modeling covariance risk in merton’s ICAMP. *Review of Financial Studies* 28 (5), 1428–1461.
- Rudebusch, G. D., Wu, T., 2008. A macro-finance model of the term structure, monetary policy and the economy. *The Economic Journal* 118, 906–926.
- Sarno, L., Schneider, P., Wagner, C., 2014. The economic value of predicting bond risk premia: Can anything beat the expectations hypothesis?, Working Paper, Cass Business School.

- Schmeling, M., Schrimpf, A., 2011. Expected inflation, expected stock returns, and money illusion: What can we learn from survey expectations? *European Economic Review* 55, 702–719.
- Stock, J. H., Watson, M. W., 1999. Forecasting inflation. *Journal of Monetary Economics* 44, 293–335.
- Stock, J. H., Watson, M. W., 2002a. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167–1179.
- Stock, J. H., Watson, M. W., 2002b. Macroeconomic forecasting using diffusion indexes. *American Statistical Association* 20 (2), 147–162.
- Svensson, L. E. O., 1994. Estimating and interpreting forward interest rates: Sweden 1992-1994, NBER Working Paper No. 4871.
- Thomas, L. B., 1999. Survey measures of expected U.S. inflation. *Journal of Economic Perspectives* 13 (4), 125–144.
- Thornton, D. L., Valente, G., 2012. Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. *The Review of Financial Studies* 25 (10), 3141–3168.
- Timmermann, A., 2006. Forecast combination. In: Elliot, G., Granger, C., Timmermann, A. (Eds.), *Handbook of Economic Forecasting*. Vol. 1. North-Holland: Elsevier, pp. 135–196.
- Wachter, J. A., 2006. A consumption-based model of the term structure of interest rates. *Journal of Financial Economics* 79, 365–399.
- Wright, J. H., 2011. Term premia and inflation uncertainty: Empirical evidence from an international panel dataset. *American Economic Review* 101, 1514–1534.
- Wright, J. H., Zhou, H., 2009. Bond risk premia and realized jump risk. *Journal of Banking and Finance* 33, 2333–2345.
- Zhou, G., Zhu, X., 2015. What drives the international bond risk premia?, Working Paper.
- Zhu, X., 2015. Out-of-sample bond premium predictions: A global common factor. *Journal of International Money and Finance* 51, 155–173.

A Data appendix

Below we list the 67 macroeconomic time series used to construct the latent common factors underlying the recursive construction of our variant of the LN_t factor. The data are obtained from the Archival Federal Reserve Economic database (ALFRED) hosted at the Federal Reserve Bank of St. Louis. We refer to Section 4 for detailed information on the construction of the LN_t factor, our treatment of the data, and dealing with multiple or missing vintages in a particular month. For each variable, we report FRED mnemonics, a full variable description, and the transformation code (tcode) used to ensure covariance stationarity of the underlying data series.

To fix notation for the transformations, let $x_{i,t}$ and $x_{i,t}^{\text{raw}}$ denote the i th variable observed at time t after transformation and the actual (untransformed) input series, respectively, and let $\Delta = (1 - L)$, with L being a lag operator such that $Lx_{i,t}^{\text{raw}} = x_{i,t-1}^{\text{raw}}$. We then apply one of six possible transformations:

1. lvl: $x_{i,t} = x_{i,t}^{\text{raw}}$
2. Δ lvl: $x_{i,t} = x_{i,t}^{\text{raw}} - x_{i,t-1}^{\text{raw}}$
3. Δ^2 lvl: $x_{i,t} = \Delta^2 x_{i,t}^{\text{raw}}$
4. ln: $x_{i,t} = \ln(x_{i,t}^{\text{raw}})$
5. Δ ln: $x_{i,t} = \ln(x_{i,t}^{\text{raw}}) - \ln(x_{i,t-1}^{\text{raw}})$
6. Δ^2 ln: $x_{i,t} = \Delta^2 \ln(x_{i,t}^{\text{raw}})$

Where applicable, we use the same transformation codes for variable transformation as in the FRED-MD dataset discussed in McCracken and Ng (2015), which relies on the same underlying data source.²⁵

²⁵The FRED-MD dataset is available from <https://research.stlouisfed.org/econ/mccracken/fred-databases/>.

No.	Mnemonic	Variable description	tcode
1	AWHMAN	Average Weekly Hours of Production and Nonsupervisory Employees: Manufacturing	1
2	AWHNONAG	Average Weekly Hours Of Production And Nonsupervisory Employees: Total private	2
3	AWOTMAN	Average Weekly Overtime Hours of Production and Nonsupervisory Employees: Manufacturing	2
4	CE16OV	Civilian Employment	5
5	CLF16OV	Civilian Labor Force	5
6	CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items	6
7	CURRDD	Currency Component of M1 Plus Demand Deposits	6
8	CURRSL	Currency Component of M1	5
9	DEMDEPSL	Demand Deposits at Commercial Banks	6
10	DMANEMP	All Employees: Durable Goods	5
11	DSPI	Disposable Personal Income	5
12	DSPIC96	Real Disposable Personal Income	5
13	HOUST	Housing Starts: Total: New Privately Owned Housing Units Started	4
14	HOUST1F	Privately Owned Housing Starts: 1-Unit Structures	4
15	HOUST2F	Housing Starts: 2-4 Units	4
16	HOUSTMW	Housing Starts in Midwest Census Region	4
17	HOUSTNE	Housing Starts in Northeast Census Region	4
18	HOUSTS	Housing Starts in South Census Region	4
19	HOUSTW	Housing Starts in West Census Region	4
20	INDPRO	Industrial Production Index	5
21	M1SL	M1 Money Stock	6
22	M2SL	M2 Money Stock	6
23	MANEMP	All Employees: Manufacturing	5
24	NDMANEMP	All Employees: Nondurable goods	5
25	OCDSL	Other Checkable Deposits	6
26	PAYEMS	All Employees: Total nonfarm	5
27	PCE	Personal Consumption Expenditures	5
28	PCEDG	Personal Consumption Expenditures: Durable Goods	5
29	PCEND	Personal Consumption Expenditures: Nondurable Goods	5
30	PCES	Personal Consumption Expenditures: Services	5
31	PCGGEF	Producer Price Index: Finished Consumer Goods Excluding Foods	6
32	PI	Personal Income	5
33	PPICPE	Producer Price Index: Finished Goods: Capital Equipment	6
34	PPICRM	Producer Price Index: Crude Materials for Further Processing	6
35	PPIFCF	Producer Price Index: Finished Consumer Foods	6
36	PPIFGS	Producer Price Index: Finished Goods	6
37	PPIIFF	Producer Price Index: Intermediate Foods & Feeds	6
38	PPIITM	Producer Price Index: Intermediate Materials: Supplies & Components	6
39	SAVINGSL	Savings Deposits - Total	6
40	SRVPRD	All Employees: Service-Providing Industries	5
41	STDCBSL	Small Time Deposits at Commercial Banks	6
42	STDSL	Small Time Deposits - Total	6
43	STDTI	Small Time Deposits at Thrift Institutions	6
44	SVGCBSL	Savings Deposits at Commercial Banks	6
45	SVGTI	Savings Deposits at Thrift Institutions	6
46	SVSTCBSL	Savings and Small Time Deposits at Commercial Banks	6
47	SVSTSL	Savings and Small Time Deposits - Total	6
48	TCDSL	Total Checkable Deposits	6
49	UEMP5TO14	Number of Civilians Unemployed for 5 to 14 Weeks	5
50	UEMP15OV	Number of Civilians Unemployed for 15 Weeks & Over	5
51	UEMP15T26	Number of Civilians Unemployed for 15 to 26 Weeks	5
52	UEMP27OV	Number of Civilians Unemployed for 27 Weeks and Over	5
53	UEMPLT5	Number of Civilians Unemployed - Less Than 5 Weeks	5
54	UEMPMEAN	Average (Mean) Duration of Unemployment	2
55	UEMPMED	Median Duration of Unemployment	2
56	UNEMPLOY	Unemployed	5
57	UNRATE	Civilian Unemployment Rate	2
58	USCONS	All Employees: Construction	5
59	USFIRE	All Employees: Financial Activities	5
60	USGOOD	All Employees: Goods-Producing Industries	5
61	USGOVT	All Employees: Government	5
62	USMINE	All Employees: Mining and logging	5
63	USPRIV	All Employees: Total Private Industries	5
64	USSERV	All Employees: Other Services	5
65	USTPU	All Employees: Trade, Transportation & Utilities	5
66	USTRADE	All Employees: Retail Trade	5
67	USWTRADE	All Employees: Wholesale Trade	5

Figure 1: Estimated survey factors.

This figure illustrates the time series dynamics of the survey factors estimated from the term structure of survey forecasts. The factors are plotted against National Bureau of Economic Research (NBER) recession periods marked in gray shading. The principal components are estimated using the full range of available observations covering the period 1968:Q4 to 2014:Q4.

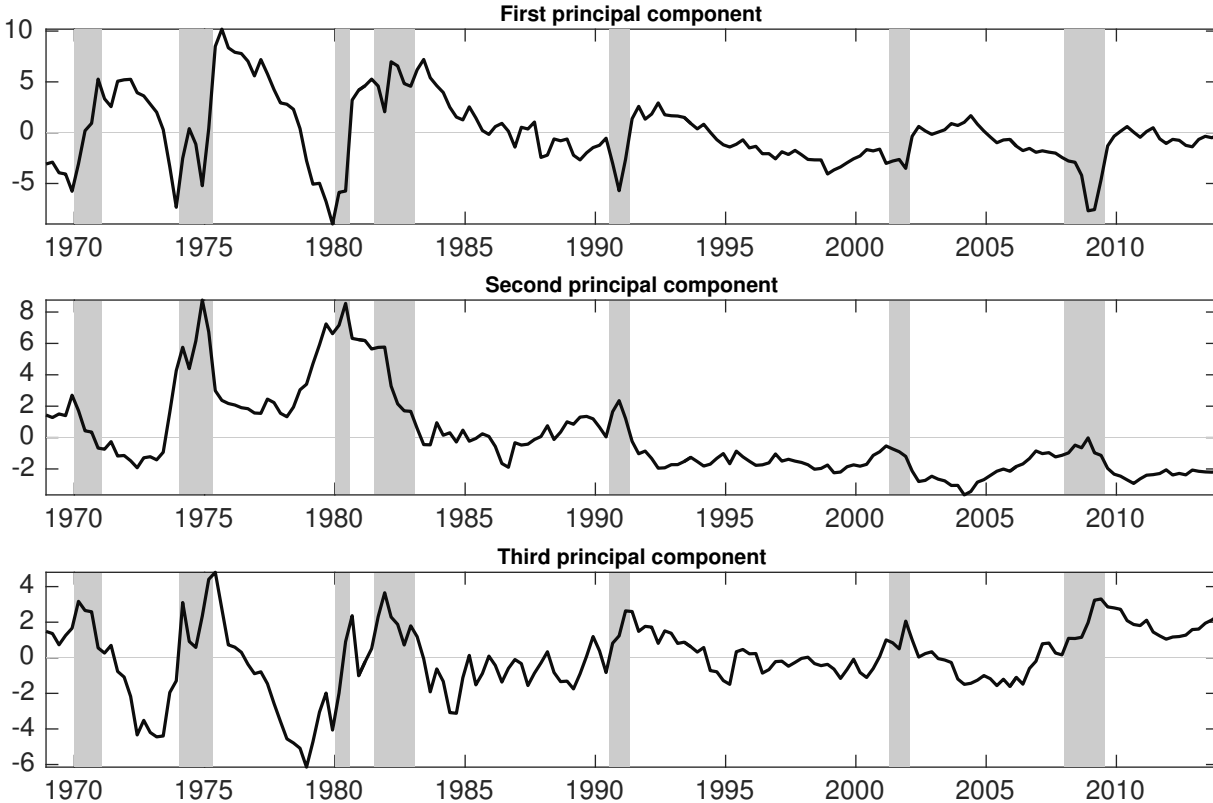


Figure 2: Time series dynamics of the ME_t factor.

This figure illustrates the relation between expected business conditions (ME_t) and the real economy as measured by the Chicago National Activity Index ($CFNAI_t$) and National Bureau of Economic Research (NBER) defined recession periods in gray shading over the sample period 1968:Q4-2014:Q4. ME_t and $CFNAI_t$ are plotted in standardized units for convenience.

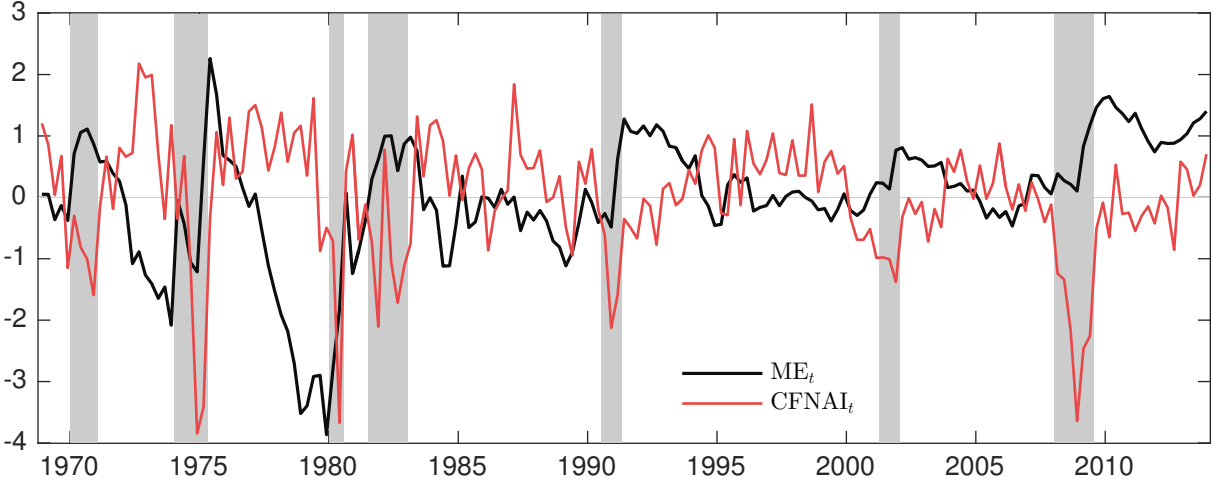


Figure 3: Differences in cumulative squared prediction errors.

This figure illustrates the relative forecasting performance by plotting the difference in cumulative squared prediction errors from (11) between the candidate forecasting model and the expectations hypothesis (EH) model over the out-of-sample evaluation period, which covers the period from 1990:Q1 to 2014:Q4. National Bureau of Economic Research (NBER) recession periods are marked in gray shading.



Figure 4: Differences in cumulative realized utility.

This figure illustrates the relative portfolio performance by plotting the difference in cumulative realized utility from (15) between the candidate forecasting model and the expectations hypothesis (EH) model over the out-of-sample evaluation period, which covers the period from 1990:Q1 to 2014:Q4. National Bureau of Economic Research (NBER) recession periods are marked in gray shading.

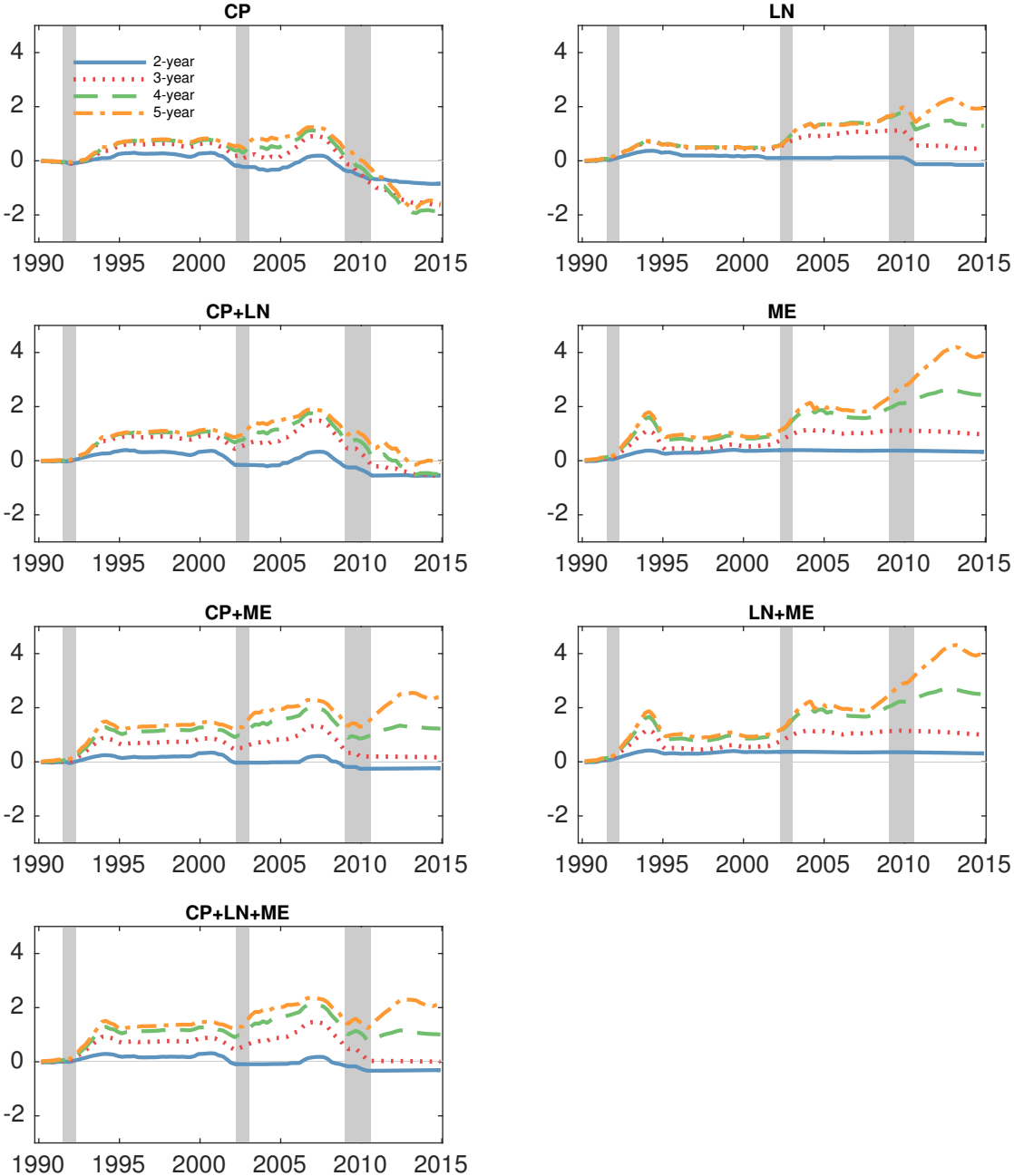


Table 1: Descriptive statistics: Survey variables.

This table reports descriptive statistics for the median expected growth rates for the macroeconomic fundamentals collected from the Survey of Professional Forecasters (SPF). The variables include: 1) GDP ($\text{gdp}_t^{\mathbb{E}}$), 2) the GDP price index ($\text{inf}_t^{\mathbb{E}}$), 3) the unemployment rate ($\text{unemp}_t^{\mathbb{E}}$), 4) corporate profits after tax ($\text{cprof}_t^{\mathbb{E}}$), 5) industrial production ($\text{ip}_t^{\mathbb{E}}$), and 6) housing starts ($\text{hous}_t^{\mathbb{E}}$). We report means, standard deviations, skewness, and kurtosis for each fundamental and forecast horizon. The last three rows in each panel reports the loadings from the principal component estimation. The sample covers the period from 1968:Q4 to 2014:Q4.

	$\text{gdp}_t^{\mathbb{E}}$	$\text{inf}_t^{\mathbb{E}}$	$\text{cprof}_t^{\mathbb{E}}$	$\text{unemp}_t^{\mathbb{E}}$	$\text{ip}_t^{\mathbb{E}}$	$\text{hous}_t^{\mathbb{E}}$
Panel A: One quarter ahead						
Mean	1.53	0.88	1.42	0.51	0.77	0.23
Std	0.57	0.50	2.11	2.74	0.79	5.67
Skewness	0.63	1.18	-0.70	1.01	-0.67	-0.05
Kurtosis	3.18	3.55	5.46	3.31	5.79	4.42
Load $\mathcal{P}_{1,t}^{\mathbb{E}}$	0.22	0.07	0.24	-0.20	0.24	0.14
Load $\mathcal{P}_{2,t}^{\mathbb{E}}$	0.18	0.36	-0.16	0.20	-0.15	0.06
Load $\mathcal{P}_{3,t}^{\mathbb{E}}$	-0.21	-0.10	-0.03	0.16	-0.10	0.43
Panel B: Two quarter ahead						
Mean	3.13	1.77	3.19	0.52	1.66	1.67
Std	1.11	0.97	3.50	4.78	1.32	10.12
Skewness	0.73	1.15	-0.37	1.10	-0.66	0.20
Kurtosis	2.95	3.49	4.95	3.74	5.87	3.79
Load $\mathcal{P}_{1,t}^{\mathbb{E}}$	0.22	0.08	0.26	-0.23	0.26	0.12
Load $\mathcal{P}_{2,t}^{\mathbb{E}}$	0.21	0.36	-0.12	0.19	-0.12	0.11
Load $\mathcal{P}_{3,t}^{\mathbb{E}}$	-0.17	-0.10	0.04	0.09	-0.03	0.45
Panel C: Three quarter ahead						
Mean	4.76	2.65	5.19	-0.08	2.63	3.73
Std	1.62	1.42	4.45	5.81	1.64	14.32
Skewness	0.78	1.09	0.23	0.89	-0.06	0.48
Kurtosis	2.83	3.36	4.02	3.29	4.38	3.69
Load $\mathcal{P}_{1,t}^{\mathbb{E}}$	0.21	0.08	0.26	-0.24	0.27	0.09
Load $\mathcal{P}_{2,t}^{\mathbb{E}}$	0.23	0.36	-0.07	0.17	-0.08	0.16
Load $\mathcal{P}_{3,t}^{\mathbb{E}}$	-0.14	-0.11	0.09	0.02	0.01	0.45
Panel D: Four quarter ahead						
Mean	6.44	3.54	7.06	-0.97	3.61	5.79
Std	2.15	1.86	5.09	6.55	1.91	17.78
Skewness	0.83	1.08	0.72	0.64	0.48	0.56
Kurtosis	2.78	3.34	3.56	2.79	3.88	3.60
Load $\mathcal{P}_{1,t}^{\mathbb{E}}$	0.20	0.09	0.25	-0.25	0.28	0.06
Load $\mathcal{P}_{2,t}^{\mathbb{E}}$	0.26	0.35	-0.01	0.14	-0.04	0.19
Load $\mathcal{P}_{3,t}^{\mathbb{E}}$	-0.12	-0.11	0.13	-0.03	0.04	0.43

Table 2: Estimating the macroeconomic expectations factor.

This table reports slope estimates from regressing one-year ahead bond risk premia upon the first three principal components ($\mathcal{P}_{1,t}^{\mathbb{E}}, \mathcal{P}_{2,t}^{\mathbb{E}}, \mathcal{P}_{3,t}^{\mathbb{E}}$) of the Survey of Professional Forecasters (SPF) forecasts and expected business conditions (ME_t) in rows (a) and (b), respectively. Although not reported, all regressions contain an intercept. Hansen and Hodrick (1980) t-statistics implemented with four lags are presented in parentheses. Adj. R^2 (%) denotes the full sample adjusted coefficient of determination in percentage. The sample period starts in 1968:Q4 and ends in 2014:Q4.

	$\mathcal{P}_{1,t}^{\mathbb{E}}$	$\mathcal{P}_{2,t}^{\mathbb{E}}$	$\mathcal{P}_{3,t}^{\mathbb{E}}$	ME_t	adj R^2 (%)
Panel A: Two-year bond					
(a)	0.08 (2.03)	-0.10 (-1.18)	0.38 (3.32)		18.87
(b)				0.46 (4.56)	18.86
Panel B: Three-year bond					
(a)	0.13 (1.79)	-0.24 (-1.58)	0.67 (3.49)		19.22
(b)				0.86 (4.71)	19.85
Panel C: Four-year bond					
(a)	0.20 (2.03)	-0.43 (-2.10)	0.83 (3.38)		18.99
(b)				1.20 (4.80)	19.84
Panel D: Five-year bond					
(a)	0.24 (2.00)	-0.56 (-2.42)	0.99 (3.57)		19.56
(b)				1.47 (4.90)	20.29

Table 3: Descriptive statistics: Bond risk premia and forecasting factors.

This table reports descriptive statistics for excess bond returns $rx_{t+4}^{(n)}$, $n = 2, \dots, 5$ and predictor variables used in the empirical analyses (Panel A) and their contemporaneous correlations (Panel B). CP_t is the forward rate-based factor from [Cochrane and Piazzesi \(2005\)](#), LN_t is the macro-based factor from [Ludvigson and Ng \(2009\)](#), and ME_t represents our proxy for expected business conditions described in Section 3. For each variable, we report means, standard deviations, skewness, and kurtosis as well as first- and second-order autocorrelations. In addition, we report Sharpe ratios (SR) for each of the Treasury bonds. The sample covers the period from 1968:Q4 to 2014:Q4.

	$rx_{t+4}^{(2)}$	$rx_{t+4}^{(3)}$	$rx_{t+4}^{(4)}$	$rx_{t+4}^{(5)}$	CP_t	LN_t	ME_t
Panel A: Descriptive statistics							
Mean	0.60	1.06	1.48	1.63	1.19	1.19	1.19
Std	1.79	3.27	4.54	5.52	1.54	1.56	1.70
Skewness	-0.24	-0.29	-0.28	-0.20	-0.10	-0.72	-1.22
Kurtosis	3.73	3.72	3.76	3.44	3.80	5.79	5.48
AC(1)	0.75	0.74	0.75	0.74	0.68	0.57	0.89
AC(4)	0.20	0.15	0.14	0.10	0.42	0.13	0.62
SR	0.34	0.33	0.33	0.30	-	-	-
Panel B: Correlation matrix							
$rx_{t+4}^{(2)}$	1.00						
$rx_{t+4}^{(3)}$	0.98	1.00					
$rx_{t+4}^{(4)}$	0.96	0.99	1.00				
$rx_{t+4}^{(5)}$	0.93	0.97	0.99	1.00			
CP_t	0.37	0.40	0.43	0.40	1.00		
LN_t	0.41	0.43	0.42	0.40	0.11	1.00	
ME_t	0.44	0.45	0.45	0.46	0.42	0.33	1.00

Table 4: Macro-spanning condition for expected business conditions.

This table reports slope estimates from the Joslin et al. (2014) macro-spanning condition in (8). Although not reported, all regressions contain an intercept. $\mathcal{P}_{1,t}^E$, $\mathcal{P}_{2,t}^E$, and $\mathcal{P}_{3,t}^E$ denote the first three principal components estimated from survey forecasts and level_t , slope_t , curv_t , $\mathcal{Y}_{4,t}$, and $\mathcal{Y}_{5,t}$ are the principal components of the yield covariance matrix. Newey and West (1987) t-statistics implemented with six lags are presented in parentheses. Adj. R^2 (%) denotes the full sample adjusted coefficient of determination in percentage. The sample period starts in 1968:Q4 and ends in 2014:Q4.

	Variable	level_t	slope_t	curv_t	$\mathcal{Y}_{4,t}$	$\mathcal{Y}_{5,t}$	adj R^2 (%)
(a)	$\mathcal{P}_{1,t}^E$	0.12 (1.82)	2.51 (4.35)	-0.55 (-0.16)			28.15
(b)		0.12 (1.88)	2.51 (4.46)	-0.55 (-0.16)	-0.51 (-0.10)	4.28 (0.77)	27.87
(a)	$\mathcal{P}_{2,t}^E$	0.25 (8.36)	-1.16 (-3.46)	4.97 (2.75)			60.67
(b)		0.26 (8.66)	-1.16 (-3.55)	4.97 (3.07)	0.27 (0.09)	7.42 (3.54)	62.97
(a)	$\mathcal{P}_{3,t}^E$	-0.07 (-1.42)	0.54 (1.53)	2.22 (1.72)			10.57
(b)		-0.07 (-1.84)	0.54 (1.94)	2.24 (1.78)	6.17 (2.52)	5.31 (1.98)	17.75
(a)	ME_t	-0.12 (-2.72)	1.19 (3.46)	-0.14 (-0.20)			42.17
(b)		-0.12 (-3.20)	1.19 (3.93)	-0.13 (-0.14)	4.27 (1.68)	2.05 (1.13)	45.23

Table 5: In-sample results.

This table reports slope estimates from regressing one-year ahead bond risk premia upon various combinations of predictors. CP_t is the forward rate-based factor from [Cochrane and Piazzesi \(2005\)](#), LN_t is the macro-based factor from [Ludvigson and Ng \(2009\)](#), and ME_t represents our proxy for expected business conditions described in Section 3. Although not reported, all regressions contain an intercept. [Hansen and Hodrick \(1980\)](#) t-statistics implemented with four lags are presented in parentheses. Adj. R^2 (%) denotes the full sample adjusted coefficient of determination in percentage. The sample period starts in 1968:Q4 and ends in 2014:Q4.

	CP_t	LN_t	ME_t	adj R^2 (%)
Panel A: Two-year bond				
(a)	0.43 (3.44)			13.17
(b)	0.26 (1.63)		0.36 (2.80)	22.55
(c)		0.47 (3.58)		16.47
(d)		0.34 (3.98)	0.36 (3.78)	26.52
(e)	0.38 (2.94)	0.43 (3.95)		26.82
(f)	0.27 (1.65)	0.35 (4.69)	0.25 (2.06)	30.72
Panel B: Three-year bond				
(a)	0.84 (3.40)			15.16
(b)	0.53 (1.90)		0.66 (3.06)	24.56
(c)		0.90 (4.21)		18.06
(d)		0.66 (4.74)	0.66 (3.57)	28.46
(e)	0.75 (3.08)	0.82 (5.00)		30.07
(f)	0.56 (1.95)	0.68 (6.03)	0.45 (2.09)	33.77
Panel C: Four-year bond				
(a)	1.28 (3.59)			18.39
(b)	0.87 (2.13)		0.87 (2.95)	26.69
(c)		1.22 (4.25)		17.10
(d)		0.89 (4.53)	0.93 (3.62)	27.74
(e)	1.16 (3.24)	1.10 (5.15)		32.12
(f)	0.91 (2.15)	0.92 (5.79)	0.58 (1.92)	35.28
Panel D: Five-year bond				
(a)	1.45 (3.16)			15.93
(b)	0.93 (1.78)		1.12 (3.01)	25.38
(c)		1.40 (4.48)		15.25
(d)		0.98 (4.56)	1.18 (3.67)	26.78
(e)	1.31 (2.85)	1.26 (5.18)		28.21
(f)	0.97 (1.79)	1.01 (5.33)	0.80 (2.04)	32.41

Table 6: Out-sample-results.

This table reports the out-of-sample results from forecasting one-year ahead bond risk premia using various combinations of predictors. CP_t is the forward rate-based factor from [Cochrane and Piazzesi \(2005\)](#), LN_t is the macro-based factor from [Ludvigson and Ng \(2009\)](#), and ME_t represents our proxy for expected business conditions described in Section 3. Panel A reports the [Campbell and Thompson \(2008\)](#) R_{oos}^2 statistic from (9) relative to the expectations hypothesis (EH) benchmark accompanied by p-values from [Clark and West \(2007\)](#) tests of equal predictive ability in square brackets. Panel B reports annualized percentage utility gain, Δ (%), relative to the EH model accompanied by p-valued from [Diebold and Mariano \(1995\)](#) tests of equal utilities in square brackets. Finally, Panel C reports annualized [Goetzmann et al. \(2007\)](#) manipulation-proof performance measures, Θ (%). The out-of-sample evaluation period starts in 1990:Q1 and ends in 2014:Q4.

n	CP	LN	CP+LN	ME	CP+ME	LN+ME	CP+LN+ME
Panel A: R_{oos}^2							
2	-16.26 [0.17]	7.57 [0.01]	0.80 [0.02]	-7.17 [0.07]	6.53 [0.04]	-1.13 [0.05]	12.31 [0.03]
3	-20.10 [0.24]	9.08 [0.01]	-2.05 [0.04]	2.08 [0.03]	10.15 [0.03]	6.99 [0.02]	14.54 [0.02]
4	-20.27 [0.20]	8.77 [0.01]	-3.98 [0.05]	8.77 [0.02]	14.06 [0.02]	11.51 [0.02]	16.09 [0.02]
5	-18.62 [0.28]	8.82 [0.01]	-3.54 [0.07]	13.08 [0.01]	16.12 [0.02]	15.05 [0.01]	17.86 [0.02]
Panel B: Δ (%)							
2	-0.84 [0.97]	-0.15 [0.75]	-0.54 [0.88]	0.32 [0.01]	-0.23 [0.75]	0.31 [0.02]	-0.31 [0.82]
3	-1.64 [0.98]	0.43 [0.17]	-0.57 [0.77]	0.96 [0.04]	0.16 [0.40]	0.99 [0.04]	-0.00 [0.50]
4	-1.95 [0.99]	1.28 [0.02]	-0.56 [0.74]	2.39 [0.00]	1.21 [0.06]	2.46 [0.00]	0.99 [0.11]
5	-1.57 [0.97]	1.93 [0.00]	-0.09 [0.53]	3.85 [0.00]	2.38 [0.00]	3.95 [0.00]	2.10 [0.01]
Panel C: Θ (%)							
2	-0.90	-0.03	-0.61	0.47	-0.27	0.46	-0.34
3	-1.40	0.76	-0.37	1.34	0.27	1.39	0.18
4	-1.42	1.56	-0.14	2.86	1.41	2.97	1.30
5	-1.06	2.06	0.40	4.01	2.51	4.19	2.39

Table 7: Model Comparisons.

This table reports the out-of-sample results from forecasting one-year ahead bond risk premia using nested models with and without expected business conditions (ME_t). R_{oos}^2 is the out-of-sample R^2 suggested in [Campbell and Thompson \(2008\)](#). For each R_{oos}^2 , we report p-values from the [Clark and West \(2007\)](#) test of equal predictive ability in square brackets. Δ (%) denotes the annualized percentage utility gain between the model including expected business conditions and the nested benchmark omitting expected business conditions. Θ (%) is the annualized manipulation-proof measure from [Goetzmann et al. \(2007\)](#). For each Δ (%), we report p-values from the [Diebold and Mariano \(1995\)](#) test in square brackets. The out-of-sample evaluation period starts in 1990:Q1 and ends in 2014:Q4.

	R_{oos}^2	p-val	Δ (%)	p-val	Θ (%)
Panel A: Two-year bond					
CP+ME vs. CP	19.60	[0.00]	0.61	[0.03]	0.63
LN+ME vs. LN	-9.41	[0.16]	0.46	[0.01]	0.49
CP+LN+ME vs. CP+LN	11.60	[0.02]	0.23	[0.19]	0.27
Panel B: Three-year bond					
CP+ME vs. CP	25.19	[0.00]	1.80	[0.00]	1.67
LN+ME vs. LN	-2.30	[0.11]	0.55	[0.13]	0.62
CP+LN+ME vs. CP+LN	16.25	[0.01]	0.57	[0.07]	0.56
Panel C: Four-year bond					
CP+ME vs. CP	28.54	[0.00]	3.16	[0.00]	2.83
LN+ME vs. LN	3.00	[0.07]	1.18	[0.05]	1.41
CP+LN+ME vs. CP+LN	19.30	[0.00]	1.56	[0.00]	1.44
Panel D: Five-year bond					
CP+ME vs. CP	29.28	[0.00]	3.95	[0.00]	3.58
LN+ME vs. LN	6.83	[0.05]	2.02	[0.01]	2.13
CP+LN+ME vs. CP+LN	20.67	[0.00]	2.19	[0.00]	2.00

Table 8: Forecast combination.

This table reports the out-of-sample results from forecasting one-year ahead bond risk premia using combinations of forecasts. Panel A presents the results from combining forecasts from the three CP_t and LN_t models and Panel B presents the results from combining forecasts for all seven model. R_{oos}^2 is the out-of-sample R^2 suggested in [Campbell and Thompson \(2008\)](#). For each R_{oos}^2 , we report p-values from the [Clark and West \(2007\)](#) test of equal predictive ability in square brackets. Δ (%) denotes the annualized percentage utility gain between the model including expected business conditions and the nested benchmark omitting expected business conditions. Θ (%) is the annualized manipulation-proof measure from [Goetzmann et al. \(2007\)](#). For each Δ (%), we report p-values from the [Diebold and Mariano \(1995\)](#) test in square brackets. The out-of-sample evaluation period starts in 1990:Q1 and ends in 2014:Q4.

n	R_{oos}^2	p-val	Δ (%)	p-val	Θ (%)
Panel A: Forecast combination using CP_t and LN_t models					
2	8.00	[0.02]	-0.41	[0.85]	-0.46
3	5.98	[0.04]	-0.50	[0.76]	-0.39
4	5.05	[0.04]	-0.50	[0.74]	-0.22
5	3.85	[0.06]	-0.07	[0.53]	0.21
Panel B: Forecast combination using all models					
2	15.81	[0.02]	0.13	[0.28]	0.16
3	17.71	[0.02]	0.42	[0.23]	0.61
4	19.60	[0.01]	1.22	[0.04]	1.39
5	19.41	[0.01]	2.11	[0.00]	2.17

Table 9: Forecast performance, expected returns, and economic variables.

This table reports contemporaneous correlations between relative forecast and portfolio performance and the real economy as measured by the Chicago Fed National Activity Index (CFNAI_t) in Panels A and B. Relative forecast performance is defined as the difference in cumulative squared prediction error (DCSPE_t) from (11) and portfolio performance is the cumulative difference in realized utilities (DCRU_t) from (15). Panels C and D report contemporaneous correlations between expected bond risk premia and the CFNAI_t and macroeconomic uncertainty ($\mathbb{U}_t^{\text{Macro}}$), respectively. Macro uncertainty is the uncertainty index constructed in Jurado et al. (2015). The out-of-sample evaluation period starts in 1990:Q1 and ends in 2014:Q4.

	2-year	3-year	4-year	5-year	2-year	3-year	4-year	5-year
	Panel A: $\rho(\text{DCSPE}_t, \text{CFNAI}_t)$				Panel B: $\rho(\text{DCRU}_t, \text{CFNAI}_t)$			
CP	0.53	0.49	0.47	0.44	0.48	0.37	0.30	0.26
LN	-0.25	-0.25	-0.22	-0.25	0.24	-0.14	-0.27	-0.27
CP+LN	0.42	0.42	0.44	0.41	0.50	0.38	0.29	0.23
ME	-0.31	-0.42	-0.32	-0.26	0.02	-0.28	-0.27	-0.24
CP+ME	0.16	0.24	0.29	0.23	0.54	0.45	0.31	0.14
LN+ME	-0.41	-0.42	-0.32	-0.27	0.04	-0.28	-0.27	-0.25
CP+LN+ME	0.03	0.12	0.22	0.17	0.49	0.41	0.29	0.16
FC1	0.25	0.30	0.33	0.32	0.50	0.38	0.31	0.24
FC2	-0.08	-0.03	0.02	0.02	0.30	0.27	0.18	0.07
	Panel C: $\rho(\mathbb{E}_t r_{t+4}^{(n)}, \text{CFNAI}_t)$				Panel D: $\rho(\mathbb{E}_t r_{t+4}^{(n)}, \mathbb{U}_t^{\text{Macro}})$			
CP	0.03	0.02	0.03	0.02	-0.24	-0.23	-0.24	-0.23
LN	-0.09	-0.09	-0.07	-0.07	-0.14	-0.13	-0.13	-0.11
CP+LN	-0.02	-0.02	0.01	0.00	-0.29	-0.28	-0.30	-0.28
ME	-0.43	-0.43	-0.42	-0.42	0.18	0.20	0.20	0.20
CP+ME	-0.29	-0.27	-0.23	-0.26	0.02	0.02	-0.03	0.01
LN+ME	-0.37	-0.37	-0.34	-0.36	0.06	0.07	0.06	0.10
CP+LN+ME	-0.23	-0.21	-0.15	-0.19	-0.10	-0.11	-0.16	-0.10
FC1	-0.03	-0.03	-0.00	-0.01	-0.28	-0.27	-0.28	-0.26
FC2	-0.24	-0.23	-0.20	-0.22	-0.09	-0.09	-0.11	-0.07

Table 10: In-sample results for quarterly bond risk premia.

This table reports slope estimates from regressing one-quarter ahead bond risk premia upon various combinations of predictors. CP_t^{GSW} is the forward rate-based factor from [Cochrane and Piazzesi \(2005\)](#), LN_t^{GSW} is the macro-based factor from [Ludvigson and Ng \(2009\)](#), and ME_t^{GSW} represents our proxy for expected business conditions described in [Section 3](#). Although not reported, all regressions contain an intercept. [Newey and West \(1987\)](#) t-statistics implemented with six lags are presented in parentheses. Adj. R^2 (%) denotes the full sample adjusted coefficient of determination in percentage. The sample period starts in 1968:Q4 and ends in 2014:Q4.

	CP_t^{GSW}	LN_t^{GSW}	ME_t^{GSW}	adj R^2 (%)
Panel A: Two-year bond				
(a)			0.55 (2.44)	3.88
(b)	0.65 (3.20)		0.57 (2.21)	8.90
(c)		0.64 (5.04)	0.53 (2.62)	9.29
(d)	0.57 (2.67)	0.60 (4.12)		10.43
(e)	0.59 (2.99)	0.58 (3.84)	0.54 (2.36)	13.72
Panel B: Three-year bond				
(a)			0.88 (2.90)	4.91
(b)	0.92 (3.40)		0.91 (2.59)	9.91
(c)		0.89 (5.35)	0.84 (3.20)	10.24
(d)	0.80 (2.81)	0.84 (4.28)		10.33
(e)	0.83 (3.21)	0.81 (3.83)	0.87 (2.82)	14.64
Panel C: Four-year bond				
(a)			1.16 (3.21)	5.43
(b)	1.16 (3.51)		1.19 (2.83)	10.53
(c)		1.09 (5.29)	1.11 (3.60)	10.47
(d)	1.01 (2.91)	1.03 (4.15)		10.14
(e)	1.05 (3.35)	0.98 (3.63)	1.14 (3.13)	14.99
Panel D: Five-year bond				
(a)			1.41 (3.40)	5.71
(b)	1.38 (3.55)		1.44 (2.99)	10.86
(c)		1.25 (5.01)	1.35 (3.86)	10.39
(d)	1.21 (2.94)	1.18 (3.88)		9.85
(e)	1.25 (3.39)	1.12 (3.36)	1.39 (3.32)	14.99

Internet Appendix for
Expected Business Conditions and Bond Risk Premia
(not intended for publication)

IA.A Predictive ability across forecast horizons

To better understand the predictive ability of survey forecasts for different forecast horizons, we construct horizon specific versions of the first three principal components $(\mathcal{P}_{1,t}^E, \mathcal{P}_{2,t}^E, \mathcal{P}_{3,t}^E)$ and the macroeconomic expectations factor, ME_t , using data from the Survey of Professional Forecasters (SPF). We estimate the horizon specific principal components and ME_t factors analogously to their full horizon counterparts in Section 3. The results, which are presented in Table IA.1, show that the predictive ability of survey forecasts is not confined to a particular forecast horizon. In fact, albeit some heterogeneity is present in the loading sizes and significance, we see that the overall results are remarkably similar across forecast horizons. The pattern of loadings being an increasing function of bond maturity in absolute value emerges again and the same functions of horizon specific components forecast bond risk premia for all maturities. As such, we view our use of the full term structure of survey expectations as a way to average, in a sense, over the heterogeneities in the predictive abilities, thereby avoiding having to evaluate all possible combinations of components and forecast horizons in the SPF data.

IA.B The term structure of survey forecasts

The descriptive statistics in Section 3 suggested that survey forecasts are extrapolative in the forecast horizon. As a verification, we plot in Figure IA.1 one- through four-quarter ahead forecasts from the SPF. As suggested by the descriptive statistics, we see that the majority of the forecasts are indeed displaying extrapolative behavior.

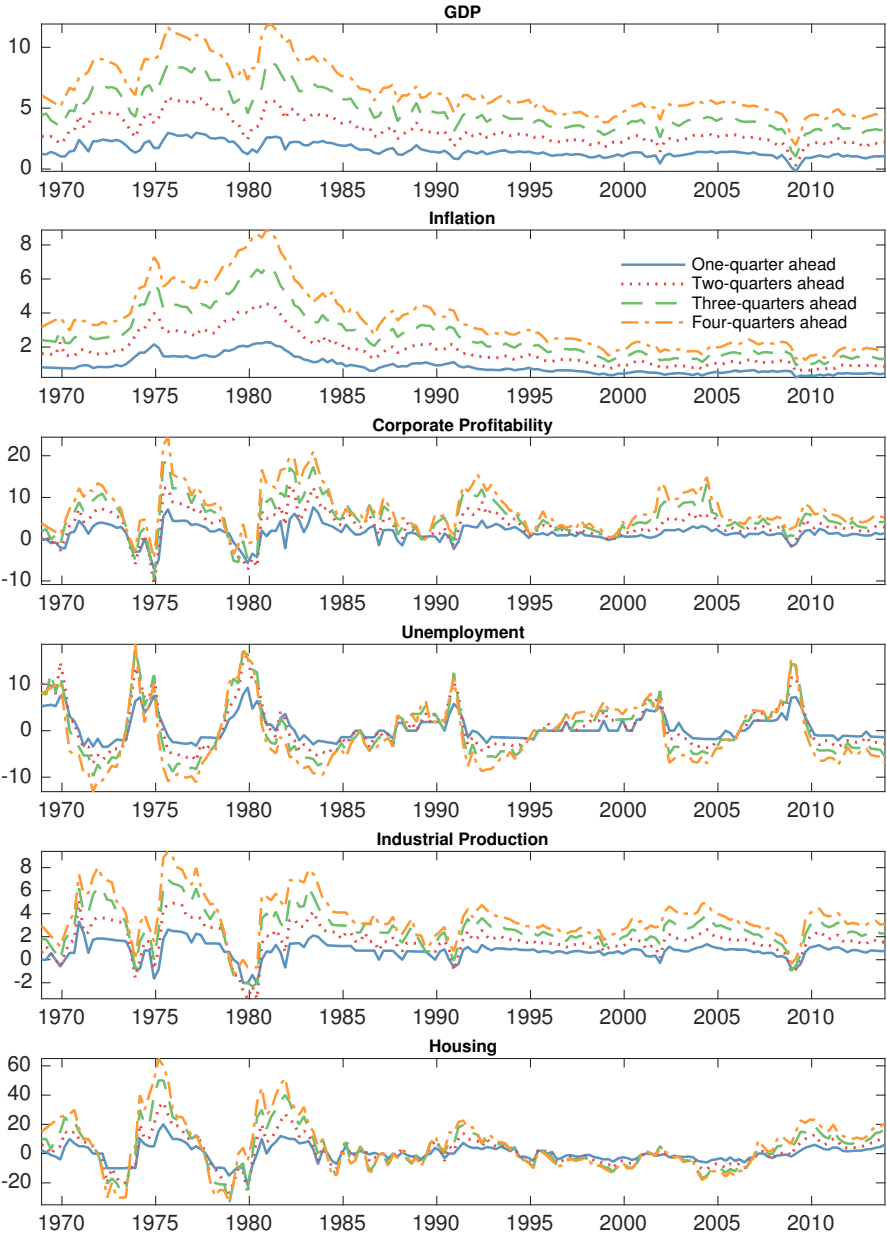
Table IA.1: Predictive ability across forecast horizons.

This table reports slope estimates from regressing one-year ahead bond risk premia upon horizon specific versions of the first three principal components ($\mathcal{P}_{1,t}^E, \mathcal{P}_{2,t}^E, \mathcal{P}_{3,t}^E$) of the Survey of Professional Forecasters (SPF) forecasts and expected business conditions (ME_t). Although not reported, all regressions contain an intercept. Hansen and Hodrick (1980) t-statistics implemented with four lags are presented in parentheses. Adj. R² (%) denotes the full sample adjusted coefficient of determination in percentage. The sample period starts in 1968:Q4 and ends in 2014:Q4.

	$\mathcal{P}_{1,t}^E$	$\mathcal{P}_{2,t}^E$	$\mathcal{P}_{3,t}^E$	adj R ² (%)	ME_t	adj R ² (%)
Panel A: Two-year bond						
One-quarter ahead	0.13 (1.44)	0.21 (1.42)	-0.77 (-3.94)	20.08	0.47 (5.29)	19.97
Two-quarters ahead	0.19 (2.36)	0.19 (1.22)	-0.77 (-3.62)	20.35	0.46 (5.35)	20.39
Three-quarters ahead	0.16 (2.27)	0.22 (1.31)	-0.75 (-3.12)	18.58	0.46 (4.37)	18.63
Four-quarters ahead	0.13 (1.60)	0.24 (1.22)	-0.69 (-2.58)	15.33	0.45 (3.23)	15.37
Panel B: Three-year bond						
One-quarter ahead	0.22 (1.46)	0.51 (1.82)	-1.31 (-3.88)	19.43	0.86 (5.26)	20.05
Two-quarters ahead	0.32 (2.23)	0.46 (1.66)	-1.37 (-3.83)	20.70	0.86 (5.51)	21.32
Three-quarters ahead	0.27 (1.91)	0.51 (1.75)	-1.35 (-3.35)	19.31	0.86 (4.60)	19.95
Four-quarters ahead	0.19 (1.21)	0.55 (1.60)	-1.24 (-2.73)	16.03	0.86 (3.40)	16.72
Panel C: Four-year bond						
One-quarter ahead	0.42 (1.95)	0.87 (2.29)	-1.61 (-3.82)	19.24	1.20 (5.39)	20.08
Two-quarters ahead	0.51 (2.58)	0.83 (2.18)	-1.69 (-3.79)	20.67	1.20 (5.68)	21.50
Three-quarters ahead	0.38 (2.05)	0.91 (2.25)	-1.66 (-3.23)	19.03	1.20 (4.70)	19.88
Four-quarters ahead	0.23 (1.14)	0.98 (2.03)	-1.50 (-2.52)	15.81	1.20 (3.44)	16.69
Panel D: Five-year bond						
One-quarter ahead	0.52 (1.99)	1.12 (2.58)	-1.90 (-4.02)	19.43	1.46 (5.35)	20.15
Two-quarters ahead	0.63 (2.57)	1.08 (2.52)	-2.02 (-4.03)	21.12	1.47 (5.76)	21.84
Three-quarters ahead	0.46 (2.03)	1.19 (2.58)	-1.99 (-3.42)	19.72	1.48 (4.84)	20.46
Four-quarters ahead	0.27 (1.13)	1.27 (2.30)	-1.82 (-2.66)	16.69	1.49 (3.58)	17.46

Figure IA.1: Term structure of expected business conditions.

This figure plots one- through four-quarter ahead forecasts from the Survey of Professional Forecasters (SPF). The sample period starts in 1968:Q4 and ends in 2014:Q4.

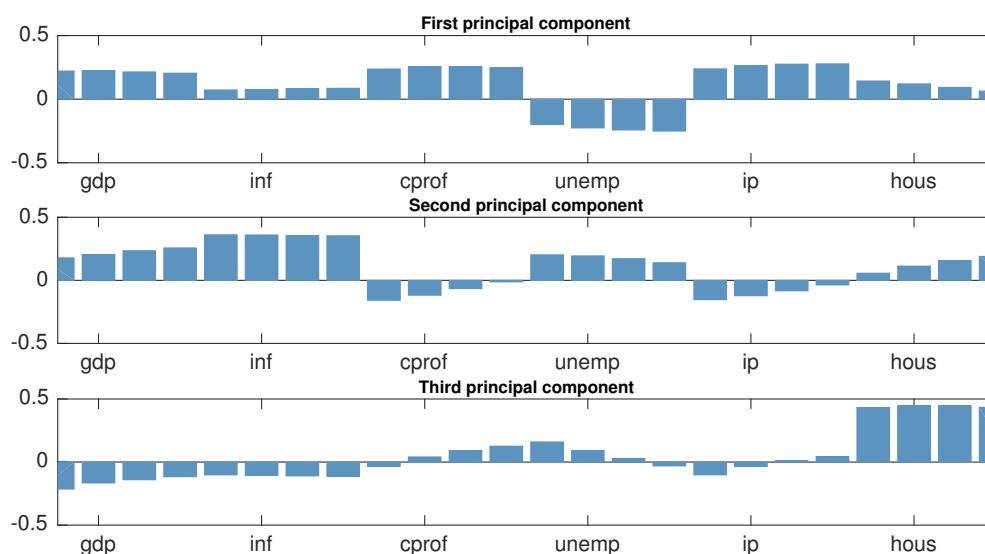


IA.C Visualizing principal component loadings

While the loadings from the principal component analysis are tabulated in Table 1, this section provides a visual interpretation by plotting the loadings in bar form in Figure IA.2. The interpretation is naturally identical to the one obtained from the tabulated loadings,

Figure IA.2: Principal component loadings.

This figure illustrates the loadings from the principal component analysis in Section 3. Loadings are grouped by macroeconomic fundamentals and ordered from shortest to longest forecast horizon.



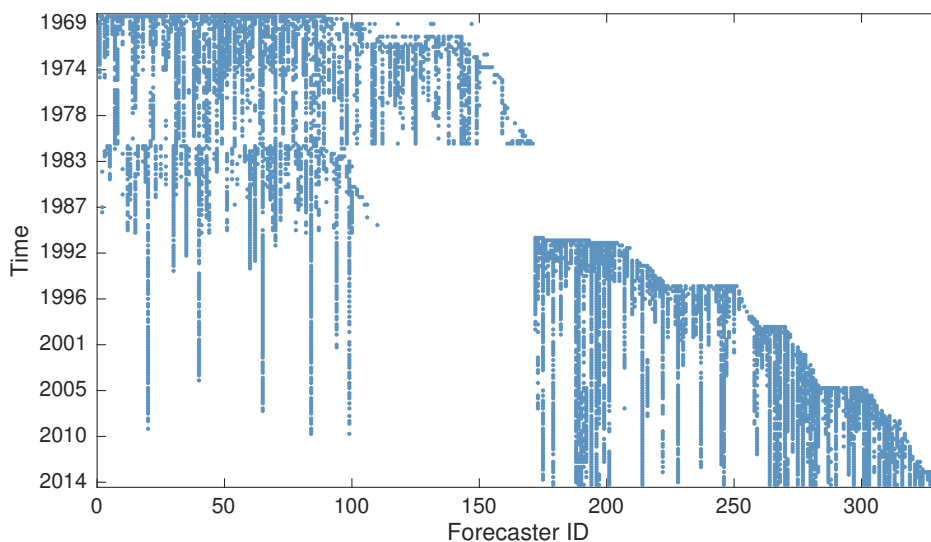
but the visualization may be more informative in some aspects. In particular, the graphical illustration naturally reveals that loadings are similar in size for a particular fundamental across forecast horizons, albeit some heterogeneity in loading sizes are present. However, no loading is placed entirely on a particular forecast horizon for any of the macroeconomic fundamentals.

IA.D Participation of individual forecasters

As mentioned in Section 3, the average number of participating forecasters in our sample is about 36 per quarter. However, this number fluctuates over time as individual forecasters exit and enter the sample at various points in time.²⁶ To illustrate the participation, we consider in Figure IA.3 the one-quarter ahead forecasts provided for the GDP variable over our sample period. Each dot represents a response by forecaster i , who are represented by

Figure IA.3: Participation of individual forecasters.

This figure illustrates the participation of individual forecasters over our sample period. The blue dots represents a submitted forecasts from forecaster i at time t . Forecaster IDs with no forecast submissions have been removed for easier readability of the figure. As a result, our notation of Forecaster ID need not correspond to the ID from the Survey of Professional Forecasters (SPF). The sample period starts in 1968:Q4 and ends in 2014:Q4.



a unique, but anonymous, forecaster ID, at time t . We see that the composition of the panel of professional forecasters vary over time and that individual forecasters frequently enter, exit, and re-enter the sample at various points in time.

²⁶See [Capistrán and Timmermann \(2009\)](#) for a discussion on how best to deal with these exits and entries from a forecasting perspective.

Research Papers 2013



- 2015-27: Rasmus Søndergaard Pedersen and Anders Rahbek: Nonstationary ARCH and GARCH with t-distributed Innovations
- 2015-28: Tommaso Proietti and Eric Hillebrand: Seasonal Changes in Central England Temperatures
- 2015-29: Laurent Callot and Johannes Tang Kristensen: Regularized Estimation of Structural Instability in Factor Models: The US Macroeconomy and the Great Moderation
- 2015-30: Davide Delle Monache, Stefano Grassi and Paolo Santucci de Magistris: Testing for Level Shifts in Fractionally Integrated Processes: a State Space Approach
- 2015-31: Matias D. Cattaneo, Michael Jansson and Whitney K. Newey: Treatment Effects with Many Covariates and Heteroskedasticity
- 2015-32: Jean-Guy Simonato and Lars Stentoft: Which pricing approach for options under GARCH with non-normal innovations?
- 2015-33: Nina Munkholt Jakobsen and Michael Sørensen: Efficient Estimation for Diffusions Sampled at High Frequency Over a Fixed Time Interval
- 2015-34: Wei Wei and Denis Pelletier: A Jump-Diffusion Model with Stochastic Volatility and Durations
- 2015-35: Yunus Emre Ergemen and Carlos Velasco: Estimation of Fractionally Integrated Panels with Fixed Effects and Cross-Section Dependence
- 2015-36: Markku Lanne and Henri Nyberg: Nonlinear dynamic interrelationships between real activity and stock returns
- 2015-37: Markku Lanne and Jani Luoto: Estimation of DSGE Models under Diffuse Priors and Data-Driven Identification Constraints
- 2015-38: Lorenzo Boldrini and Eric Hillebrand: Supervision in Factor Models Using a Large Number of Predictors
- 2015-39: Lorenzo Boldrini and Eric Hillebrand: The Forecasting Power of the Yield Curve, a Supervised Factor Model Approach
- 2015-40: Lorenzo Boldrini: Forecasting the Global Mean Sea Level, a Continuous-Time State-Space Approach
- 2015-41: Yunus Emre Ergemen and Abderrahim Taamouti: Parametric Portfolio Policies with Common Volatility Dynamics
- 2015-42: Mikkel Bennedsen: Rough electricity: a new fractal multi-factor model of electricity spot prices
- 2015-43: Mikkel Bennedsen, Asger Lunde and Mikko S. Pakkanen: Hybrid scheme for Brownian semistationary processes
- 2015-44: Jonas Nygaard Eriksen: Expected Business Conditions and Bond Risk Premia