Determinants of Bond Risk Premia*

Jing-zhi Huang and Zhan Shi

Penn State University

First draft: November 2009
This version: January 2011

Abstract

In this paper, we provide new and robust evidence on the power of macro variables for forecasting bond risk premia by using a recently developed model selection method–the supervised adaptive group "least absolute shrinkage and selection operator" (lasso) approach. We identify a single macro factor that can not only subsume the macro factors documented in the existing literature but also can substantially raise their forecasting power for future bond excess returns. Specifically, we find that the new macro factor, a linear combination of four group factors (including employment, housing, and price indices), can explain the variation in excess returns on bonds with maturities ranging from 2 to 5 years up to 43%. The new factor is countercyclical and furthermore picks up unspanned predictability in bond excess returns. Namely, the new macro factor contains substantial information on expected excess returns (as well as expected future short rates) but has negligible impact on the cross section of bond yields.

*We thank Yakov Amihud, Charles Cao, Long Chen, Peter Feldhütter (AFA discussant), Laura Field, Raymond Kan, Anh Le, Hong Liu, Marco Rossi, Ilona Shiller, Joel Vanden, Hong Yan, Weina Zhang, Hao Zhou, and seminar participants at National University of Singapore, Penn State, Singapore Management University, Wisconsin-Milwaukee, the 20th Annual Derivatives Securities and Risk Management Conference at FDIC, the 2010 CICF, the 2010 FMA, the 2010 SAIF and CKGSB Summer Research Conference, and the 2011 AFA for helpful comments and suggestions. Both authors are at the Smeal College of Business, Penn State University, University Park, PA 16802, USA. Email addresses: jxh56@psu.edu (Huang) and zus116@psu.edu (Shi).

1 Introduction

Recent empirical evidence has documented that some financial and macroeconomic variables can be used to predict the excess returns of the U.S. Treasury bonds. For instance, financial variables found to have such predictive power include forward rates or spreads (Fama and Bliss, 1987; Stambaugh, 1988; Cochrane and Piazzesi, 2005) and yield spreads (Campbell and Shiller, 1991). In particular, Cochrane and Piazzesi show that a tent-shaped linear combination of five forward rates can explain between 30 and 35 percent of the variation in one-year excess returns on bonds with 2-5 years to maturity. On the other hand, Ludvigson and Ng (2009b) obtain a macro factor (extracted from a monthly panel of 131 macroeconomic variables using dynamic factor analysis) that has forecasting power for bond excess returns, above and beyond the power contained in the aforementioned financial variables. Specially, Ludvigson and Ng find that their factor alone can explain 21-26 percent of the one-year excess returns and can raise it to 42-45 percent when augmented with the Cochrane-Piazzesi factor. Interestingly, Cochrane and Piazzesi, Ludvigson and Ng, and Duffee (2008) all document empirically the presence of a so-called "hidden" factor, namely, a factor that contains substantial information about expected excess returns but has negligible impact on the cross section of bond yields. These findings have generated important insights and implications for term structure modeling, and spawned a fast growing literature on the determinants of bond risk premia. Nonetheless, some recent studies have raised concerns about the robustness of the documented power of those financial and macro variables for predicting bond risk premia (see, e.g., Duffee, 2007).

In this paper, we reexamine the potential power of macro variables for forecasting bond risk premia using a recently developed model selection method, namely, the supervised adaptive group "least absolute shrinkage and selection operator" (lasso) approach (referred to as the SAGLasso approach, hereafter). We first extract a new macro factor from a standard monthly panel of macro variables–the same set of macro variables used in Ludvigson and Ng (2009b)–using the SAGLasso approach. We then examine the intuition of the new macro factor and, in particular, investigate whether the new factor has any forecasting power for bond risk premia above and beyond the predictive power contained in those financial and macro factors identified in the literature. Finally, as a robustness check, we address two issues raised recently by Duffee (2007, 2008) about the empirical literature on the prediction of bond excess returns.

The new macro factor that we obtain is a linear combination of four non-overlapping group factors, each of which itself is a linear combination of a small number of closely related macro variables (a subset of the original 131 macro variables) and thus has a clear interpretation. More specifically, the four group factors represent employment, housing, price indices, and financial, respectively. As such, our new macro factor is easy to interpret.

We find that the new macro factor can predict excess returns on 2- to 5-year maturity bonds with (in sample) R^2 up to 43 percent. This is significantly higher than that found by either Cochrane and Piazzesi (2005, CP hereafter) or Ludvigson and Ng (2009b, LN hereafter). Furthermore, our new macro factor is found to subsume the LN factor. However, like the LN factor, our factor does not subsume the CP factor and contains information about bond risk premia that is not contained in the CP factor. Augmenting our factor with the latter can increase the R^2 of the forecasting regression to 47 percent. Like the CP and LN factors, the new macro factor is found to be countercyclical. We also find that our new factor has strong out-of-sample forecasting power as well and moreover has significantly incremental predictive power beyond that in the LN factor. Overall, results from both in-sample and out-of-sample analysis indicate that our new macro factor contains information about future bond excess returns beyond what captured by the CP and LN factors.

We also find that our employment group factor can subsume the output gap factor found by Cooper and Priestley (2009) that can predict excess returns on 2- to 5-year maturity bonds with R^2 equal to 2 percent. As such, our macro factor goes beyond output gap and inflation (two main macro variables considered in existing studies) and, in particular, includes a component of macro risk tied to economic measures in the housing sector, that is consistent with the implication of the Piazzesi, Schneider, and Tuzel (2007) model.[1]

To explore further the information content in the new macro factor, we include the realized jump-mean factor of Wright and Zhou (2009) in predictive regressions of the bond risk premium on the macro factor and the Duffee hidden factor, both jointly and separately. Regression results indicate that these three factors are all significant and jointly can predict excess returns on 2- to 5-year maturity bonds with R^2 up to 66 percent (where the sample period used is 1984-2007, a sub sample period over which the jump factor can be constructed).

Finally, we conduct a robustness test of the empirically documented predictability of our new macro factor by addressing two issues raised by Duffee (2007, 2008). He argues in the former study that all existing predictive regression studies actually test a restrictive null hypothesis that excess bond returns are unforecastable, whereas the more relevant null hypothesis should be that "expected excess returns are stochastic, persistent, and independent of the macroeconomy." We attempt to distinguish between these two nulls by documenting the strong forecasting power of the lagged value of excess returns themselves. We construct tests for both restrictive and general null hypotheses and find that the general null, which Duffee cannot reject in the finite sample analysis, is rejected regardless whether asymptotic or simulated critical values are used. The other issue is whether the evidence for predictability of excess returns is simply a symptom of small-sample biases in estimated t-statistics or R^2. We find that all the evidence of return predictability (based on regressions with asymptotic

[1]Piazzesi et al. (2007) focus on excess stock returns. But the same mechanism applies to excess bond returns because risk premia on bonds and stocks are largely driven by the same business-cycle factors (Fama and French, 1989).

theories) in our new macro factor persists even after we adjust the estimated test statistics for their finite-sample properties.

To sum, we provide new and robust evidence on the link between expected excess bond returns and macroeconomic variables. The new macro factor identified in our analysis is intuitive, includes a housing component, subsumes both the Ludvigson-Ng macro factor and the output gap identified in Cooper and Priestley, and contains the information about future bond excess returns that is not contained in the Cochrane-Piazzesi forward rate factor, the Duffee hidden factor (referred to as an expectation factor by some researchers), and the Wright-Zhou realized jump-mean factor. Furthermore, our analysis indicates that sources of bond risk premium predictability include macro variables, jumps, and an expectation factor.

The SAGLasso approach, used to extract our bond return forecasting factor from a large set of 131 macro variables using, has some advantages over the standard principal component analysis (PCA) or factor analysis. First, the SAGLasso approach selects predictors based on their association with the dependent variable (the bond risk premium in our case), whereas principal components may contain most information with respect to the data matrix of independent variables, but this information may not be most correlated with the dependent variable to be forecasted.[2] Second, the SAGLasso approach picks only a few most important ones (out of those 131 macro variables) as explanatory variables by shrinkage, whereas principal components or factors estimated using the PCA method are linear combinations of all 131 macro variables. In particular, due to cluster structure of macroeconomic data, we can divide 131 macro variables into groups and then apply the SAGLasso approach at the group level to help us select group factors (which are informative and easy to interpret) and thus identify underlying economic determinants of bond risk premia. Finally, predictive regressions of excess bond returns tend to exhibit autocorrelation (due to both high serial and cross-sectional correlations of bond prices) and the SAGLasso approach provides a robust way to correct for autocorrelated disturbances with an unspecified structure in such regressions.

Our study builds directly on the insightful studies by Cochrane and Piazzesi (2005) and, in particular, Ludvigson and Ng (2009b), respectively, the latter of which documents among other things that macro factors have important forecasting power for future bond excess returns, above and beyond the predictive power contained in yield curve factors (identified by the former study). We extend LN in several directions. First, we extract macro factors using the SAGLasso approach instead of dynamic factor analysis, and identify a new factor– the housing factor. Secondly, we identify more sources of bond risk premium predictability. Finally, we address the concerns raised by Duffee (2007, 2008) on the robustness of such a predictability. Overall, we provide new and robust evidence to support LN's findings.

[2]However, the information on the dependent variable can be used to help select a particular combination of those principal components (and/or their higher order terms) as a predictor in predictive regressions. See Ludvigson and Ng (2009b).

Our study is also closely related to the macro finance literature (Ang and Piazzesi, 2003). Several recent studies of dynamic term structure models (DTSM) document that factors unspanned by bond yields have predictive content for bond excess returns. For instance, Cochrane and Piazzesi (2009) and Duffee (2008) focus on unspanned "yield-curve" risks by allowing yield factors, other than the traditional "level", "slope" and "curvature" ones to drive variation in expected excess returns. Joslin, Priebsch, and Singleton (2009) develop a model that incorporates macro factors but allows for components of macroeconomic risks orthogonal to the yield curve. Our empirical analysis sheds more light on the nature of unspanned predictability documented in the aforementioned studies. Specifically, we identify macroeconomic risk over and beyond that associated with variations in output gap and inflation, the focus of current literature. Also, our regression results highlight the importance of incorporating both yield-curve evolution and macroeconomic fundamentals in extending the conventional three-factor DTSMs, because risk premia on unspanned predictors are not identified otherwise.

The organization of the paper is as follows: The next section lays out the econometric framework and introduces the Supervised Adaptive Group Lasso (SAGLasso) method. Section 3 reports empirical results. In particular, we extract first those macro factors with significant predictive power for excess bond returns and then conduct both in-sample and out-of-sample forecasting regression analysis. Section 4 presents a bootstrap analysis for finite-sample inference. Section 5 summarizes the results of our investigation. The appendix provides a list of macroeconomic variables used in the analysis and also describes the dynamic term structure model used in the bootstrap analysis.

2 The Empirical Method

This section introduces the Supervised Adaptive Group Lasso method and illustrates how to use it to select macroeconomic factors that can forecast excess bond returns. Below we first describe the penalized least squares, lasso, adaptive lasso, and group lasso. We then propose our SAGLasso procedure based on the latter.

2.1 Motivation

There are two types of excess returns used in the literature on predicting excess bond returns. In this paper, we follow Fama and Bliss (1987) by using continuously compounded annual log returns on a n-year zero-coupon Treasury bond in excess of the annualized yield on a 1-year zero-coupon Treasury bond. For $t = 1, \cdots T$, excess returns are defined $rx_{t+1}^{(n)} = r_{t+1}^{(n)} - y_t^{(1)} = ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - y_t^{(1)}$, where $r_{t+1}^{(n)}$ is the one-year log holding-period return on an n-year bond purchased at time t and sold at time $t+1$, and $y_t^{(n)}$ is the log yield on the n-year bond.

To examine if predictable variation in excess bond returns is specifically related to the macroeconomic state, researchers often run the following predictive regression:

$$rx_{t+1}^{(n)} = \gamma' Z_t + e_{t+1} \tag{1}$$

The predictors Z_t are ususally based on a few predetermined macroeconomic measures such as GDP growth, NAPM price index and personal consumption expenditure etc. Nevertheless, the decision as to which predetermined macro variables to use in the econometric analysis can substantially influence the estimated predictability of excess returns. Moreover, as pointed out by LN, there is a potential degree-of-freedom problem here if the number of predictors used is large. In fact, LN emphasize that it is infeasible to follow standard econometric procedure with mass information contained in the $T \times N$ panel "unless we have a way of ordering the importance of the N series in forming conditional expectation."

To get around these difficulties, LN employ the dynamic factor analysis to estimate several linear combinations from a large panel of macroeconomic series. Such estimated (static) factors can then be used in predictive regressions. One big advantage of the LN methodology is that it allows us to summarize the information from a large number of time series using only a few factors as follows

$$rx_{t+1}^{(n)} = \beta' F_t + \gamma' Z_t + e_{t+1} \tag{2}$$

Although very insightful, the LN method has some limitations. First, as mentioned earlier principal component analysis (PCA) is an effective tool for finding linear combinations of features that exhibit largest variation in a data set, but the resulting factors are not necessarily most related with variations in the outcome variable to be predicted. In particular, the first component does not necessarily correlate strongly with the dependent variable in factor-augmented regressions. For instance, the fifth and sixth components $\hat{F}_{5t}$ and $\hat{F}_{6t}$ identified in LN are found to have little forecasting power for excess bond returns.

To summarize, asymptotic principal component analysis is a standard method for modeling correlation but the factors extracted using this method are not necessarily those most correlated with the dependent variables that we want to forecast. The SAGLasso approach used in this study singles out subsets of input features (and groups) associated with a particular response/dependent variable and thus drives the estimated factor toward it. As a result, under this approach, it is possible to extract a sparse loading vector from a large panel of noisy macroeconomic series and identify the association between bond risk premia and different economic sectors.

2.2 Lasso, Adaptive Lasso, and Group Lasso

We assume that there are N macroeconomic measures observed for T time periods. Let $\mathbf{X}$ be the $T \times N$ panel of macroeconomic data with elements x_{it}, $i = 1, \cdots N$, $t = 1, \cdots T$. As in dynamic factor analysis, the cross-sectional dimension here, N, is large and possibly greater than the number of observations, T. Throughout this paper, we define

$$||\eta||_P = (\eta' P \eta)^{1/2}$$

for a vector $\eta \in \mathbb{R}^k$ and a symmetric $k \times k$ positive definite matrix P. And we write $||\eta|| = ||\eta||_{I_k}$ for brevity. For a $T \times 1$ response vector $\mathbf{y}$, the penalized least squares function is defined to be

$$||\mathbf{y} - \mathbf{X}\beta||^2 + \lambda \sum_{i=1}^{N} |\beta_i|, \tag{3}$$

where $\lambda \geq 0$ is a tuning parameter. Note that the penalty functions $p_i(\cdot)$ are not necessarily the same for all i. The ℓ_1-norm penalty $|\beta_i|$ used here induces sparsity in the solution, and defines the lasso method (Tibshirani, 1996). Bai and Ng (2008) pioneer the use of the novel shrinkage method in macroeconomic forecasting by setting the soft thresholding to select a subset of variables from which factors are extracted.

We now discuss why LASSO regression is preferred to using a linear combination of all macro variables available as the fitted value estimated from OLS. First, in terms of forecasting accuracy, OLS estimates usually exhibit low bias but large variance but shrinkage methods can sacrifice a little bias to reduce the variance of the predicted value and thus improve the overall forecasting performance. Next, it is well known that OLS has poor finite sample property when the dimension of parameters to be estimated is comparable with the number of observations.[3] The lasso approach is developed to handle such problems.

Zou (2006) modifies the lasso penalty as the following

$$||\mathbf{y} - \mathbf{X}\beta||^2 + \sum_{i=1}^{N} \lambda_i |\beta_i| \tag{4}$$

such that different amounts of shrinkage are allowed for different regression coefficients. It has been shown that this "Adaptive Lasso" has the oracle property if the weights λ_i are data-dependent and appropriately chosen.

Despite its popularity, lasso has limitations. For instance, if each explanatory factor may be represented by a group of derived input variables, then the lasso tends to select only one variable from each group and does not care which one is selected, especially when there is a group of variables among which the pairwise correlations are very high. In another word, the solution for lasso depends on how the factors are represented,[4] and this is undesirable in economic forecasting. Another issue is that when $N > T$, the lasso selects at most T variables before it saturates, because of the nature of the convex optimization problem. As

[3]For our empirical study, there are 131 macroeconomic variables with only 528 observations.

[4]That is, if any factor is reparameterized through a different set of orthonormal contrasts, we may obtain a different set of factors in the solution.

shown below, the first problem can be fixed by the Group Lasso of Yuan and Lin (2006). The second concern is automatically removed by the SAGLasso procedure.

We begin with the following liner model

$$\mathbf{Y} = \mathbf{X}\beta^{\mathbf{0}} + e \tag{5}$$

with the assumption that e is a T-dimensional vector of *iid* errors, and we well relax this assumption later. The central modeling assumption of Group Lasso is that some subvectors of the true coefficients $\beta^{\mathbf{0}}$ are zero. And we denote by $h \in \mathcal{H}_1 = \{h : \beta^{\mathbf{0}}_{\mathbf{h}} \neq 0\}$ the unknown index set of non-zero subvectors of $\beta^{\mathbf{0}}$. Hence, the Group Lasso involves identifying $\mathcal{H}_1$ and estimating $\beta^{\mathbf{0}}$. This can be naturally formulated in the framework of penalized least squares.

The general form of Group Lasso estimate is defined in Yuan and Lin (2006) as the solution to the following problem:

$$\min_{\beta \in \mathbb{R}^N} \left\{ ||\mathbf{y} - \mathbf{X}\beta||^2 + \lambda \sum_h ||\beta_h||_{W_h} \right\}, \tag{6}$$

given $N_h \times N_h$ positive definite matrices W_h. However, in practice, the Group Lasso is usually implemented by estimating the following restrictive form

$$\min_{\beta \in \mathbb{R}^N} \left\{ ||\mathbf{y} - \mathbf{X}\beta||^2 + \lambda \sum_h ||\beta_h|| \right\}. \tag{7}$$

Note that expression (7) reduces to the Lasso when $|\mathcal{H}| = N$ and each h corresponds to the 1-dimensional subspace of $\mathbb{R}^T$ spanned by the corresponding column of the design matrix $\mathbf{X}$. Instead, we stick to the general Group Lasso in this paper and mimic the Adaptive Lasso by setting $W_h = w_h I_{N_h}$. The resulted penalty function is given by

$$\min_{\beta \in \mathbb{R}^N} \left\{ ||\mathbf{y} - \mathbf{X}\beta||^2 + \lambda \sum_h w_h ||\beta_h|| \right\} \tag{8}$$

For a large-scale macroeconomic data set, series are usually organized in a hierarchical manner, making the Group Lasso perfectly suitable for factor selection. However, even within a group, different variables may represent certain quantitative measurements of different economic sectors. In our context, it is natural to conjecture that Industrial Production (IP) Index for consumer goods might have connection with the bond risk premia in a different way from the connection between IP Index of materials and the bond risk premia. Thus, it seems necessary to also consider variable selection at the within-cluster level so that the irrelevant individual series can be screened out. Furthermore, the capability of selecting informative economic measures within the selected groups is especially valuable for practitioners that require parsimonious models with specific input variables.

2.3 Supervised Adaptive Group Lasso

Essentially, the Supervised Adaptive Group Lasso (SAGLasso) consists of two steps. In the first step, we identify informative individual macro series within each cluster using the Adaptive Lasso method. In the second step, we select important clusters using the Group Lasso. The details of the SAGLasso procedure are provided in the appendix. To our knowledge, the SAGLasso is the first to consider penalized time series selection at both the cluster level and the within cluster level.

Compared to individual variable selection methods, the SAGLasso is capable of taking cluster information into consideration. This makes it possible to reveal the associations between term premia and macrocosmic fundamentals. With the proposed approach, we can identify macroeconomic measures which are jointly significantly associated with risk premia in bond returns. Compared to simple cluster based methods such as Group Lasso, SAGLasso carries out the additional within cluster selection. This leads to a small number of variables within each cluster. So beyond identifying influential factors, the proposed approach can also identify the economic sectors that actually cause the association.

Besides the SAGLasso method, another two-step supervised learning approach in the literature is the supervised principal component analysis (SPCA) proposed by Bair et al. (2006) in a biological setting. Bai and Ng (2008) incorporate SPCA into the confines of Diffusion Index framework and apply it to inflation forecasts. Significant differences exist between the SAGLasso and other two-step approaches like SPCA. In other two-step approaches, the first supervised screening step considers all candidate variables simultaneously and the cluster structure is ignored, whereas the main merit of the SAGLasso is the usage of the information of underlying cluster structure. Moreover, in SPCA, the selected features are the principal components. Although they may have satisfactory prediction performance, economic interpretations may not be clear. As a comparison, for SAGLasso clear economic interpretations of macro series identification results are available as shown in the next section. Finally, the asymptotic theory for SPCA, established by Bair et al. (2006), is based on the assumption of *iid* disturbance, which is usually not the case for financial time series. In contrast, penalized least squares (or more generally, penalized likelihood) can be applied to general linear models with autocorrelated disturbance, as long as the objective function is well defined and locally differentiable. In our context, suppose the disturbances are homoscedastic but correlated across observations

$$rx_{t+1}^{(n)} = \beta' X_t + \epsilon_{t+1}, A(L)\epsilon_{t+1} = \nu_{t+1}, \tag{9}$$

where ν_{t+1} is nonautocorrelated. If the lag operator has the order of p, then the AR(p) disturbance model is

$$A(L)rx_{t+1}^{(n)} = \beta' A(L) X_t + \nu_{t+1}. \tag{10}$$

Therefore, any model with an AR(p) disturbance can be written as a more general autoregressive distributed lag (ARDL) model with both p autoregressive terms and p distributed lag terms, namely, the following ARDL(p, p) model

$$B(L) rx_{t+1}^{(n)} = \beta' C(L) X_t + \nu_{t+1} \tag{11}$$

with a nonlinear restriction imposed $B(L) = C(L)$. Thus, upon finding evidence of autocorrelation on the basis of LN's and CP's model, we proceed with relaxing the nonlinear restrictions on the ARDL model, instead of seeking serial correlation in omitted variables. This interpretation is largely supported by our empirical result: adding additional lags of the dependent variable boosts the explanatory power of the model, with R^2 up to nearly 90%, leaving little room for other explanatory variables.

3 Empirical Results

In this section we extract macro factors from a monthly panel of 131 measures of economic activity over the period 1964-2007 and then examine their power for forecasting excess bond returns. Section 3.1 describes the data and discuss two different null hypotheses to infer the relationship between term premia and the macroeconomy. Section 3.2 presents some preliminary results from CP and LN's analysis, which motivates our use of the general ARDL model. The empirical results obtained from our SAGLasso are summarized in Sections 3.3 and 3.4. Sections 3.5 and 3.6 discuss the economic interpretation of the estimated factor by characterizing its countercyclic pattern and unspanned predictive ability, respectively. Section 3.7 examines the role of jumps. Finally, Section 3.8 reports some robustness checks and assess the performance in subsample and out-of-sample analysis.

3.1 Data and Null Hypotheses

Monthly prices for 1-year through 5-year zero coupon U.S. Treasury bonds from CRSP are used to construct annual excess returns, as specified at the beginning of Section 2. To follow the literature, we construct annual returns by continuously compounding monthly return observations, rather than constructing monthly excess returns. In spite of the well-known statistical problem associated with regressions involving overlapping observations, there may truly be more information on predictability of excess returns using the annual excess returns because they subtract the 1-year yield instead of the 1-month yield.

Our macroeconomic data set consists of a balanced panel of 131 monthly macroeconomic times series, each spanning the period from January 1964 to September 2007. The same data are used in most dynamic factor model studies, such as Stock and Watson (2002,2005) and Ludvigson and Ng (2009a, 2009b). These series are initially transformed to induce stationarity. To provide a basis for comparison, we include in our data set as many economic series as used in Ludvigson and Ng (2009a), and the sample periods are exactly the same. These series are roughly identified into 15 broad categories: real output and income; employment and hours; real retail, manufacturing and trade sales; consumption; housing starts and sales; real inventories; orders; commercial credit; stock indexes; exchange rates; interest rates and spreads; money and credit quantity aggregates; inflation indexes; average hourly earnings; and miscellaneous. The complete list of series and their transformation is given in the appendix.

Duffee (2007) argues that the existing literature does not test the relevant null hypothesis that expected excess bond returns are persistent and uncorrelated with macroeconomic measures. Instead, previous studies uses the restrictive null that excess returns are unforecastable. For simplicity, we follow Duffee by referring to the former null hypothesis as the general null and the latter null hypothesis as the restrictive null. Statistically, the well-known spurious regression problem is exemplification of their difference and the general null can be typically incorporated into the linear regressions by adjusting the covariance matrix of parameter estimates. And if we follow the ARDL methodology to correct for the autocorrelation, the issue will be fully clarified. However, in predictive regression the critical issue associated with different null hypotheses is whether the small-sample properties of test statistics are close to standard asymptotic properties. Therefore, when analyzing the finite-sample properties of their techniques, existing studies usually makes bootstrap inference based on the restrictive null hypothesis. Put differently, the model used to generate simulation data offer only a choice between term premia that covaries with macroeconomic variables and term premia that are serially uncorrelated.

But the most significant effect brought about by different null hypotheses may lie in the economic aspect. If we focus on predictors from the financial sector, using restrictive null may not yield any conclusive result on whether expected excess returns are correlated with the macroeconomy. As stressed by Duffee (2007), least square regression detects partial correlations instead of unconditional correlations. Therefore, if the macroeconomic series are correlated with the noise in financial variables derived from prices of risky securities, they would exhibit forecasting power in regressions even if these series are independent of excess returns. To assess the independent predictive power of macroeconomic inputs, in this paper we test both restrictive and general null hypothesis.

3.2 Some Preliminary Results

We first construct the return-forecasting factors used in CP and LN, completely following their methodology. The explanatory variables used by CP are the annualized forward rates $F_t^{(n)} = p_t^{(n)} - p_t^{(n+1)}$, $n = 0, \ldots 4$. Stack them in the vector For_t. The specification in LN's forecasting regression is slightly more complicated. Using eight static factors ($\hat{f}_{it}$) estimated from asymptotic PCA, they perform best-subset selection among different subsets of $\hat{f}_{it}$ as well as their quadratic and cubic functions. A nine-factor subsect given by $\hat{F}9_t = \left(\hat{f}_{1t}, \ldots \hat{f}_{8t}, \hat{f}_{1t}^3\right)$ is found to minimize the in-sample and the out-of-sample BIC. The CP and LN single predictive factors are then formed as the fitted value from following regression

$$
\begin{aligned}
arx_{t+1} &= \gamma_0 + \gamma' For_t, && (12) \\
arx_{t+1} &= \delta_0 + \delta' \hat{F}9_t, && (13)
\end{aligned}
$$

where $arx_{t+1} = \frac{1}{4}\sum_{n=2}^{5} rx_{t+1}^{(n)}$. Table 1 shows coefficient estimates, associated t-statistics and R^2 value for Eq. (13). We denote the two single factors $\widehat{CP}_t = \hat{\gamma}' For_t$ and $\widehat{LN}_t = \hat{\delta}' \hat{F}9_t$, respectively.

Panel A of Table 2 reports results of univariate regressions of 2-5 year excess returns on CP and LN factors. As the regression for annual excess returns use overlapping observations, we compute standard errors using the Hansen-Hodrick (1980) GMM correction for overlap (the first row in parentheses). Also, the Newey-West (1987) asymptotic standard t-statistics (with 18 lags) are reported to correct serial correlation. Note that these two correction methods correspond to different null hypotheses for the Wald tests. Because the restrictive null states that forecast errors are serially uncorrelated and the general null is that forecast errors contain persistent components independent of the macroeconomy, the robust Hansen-Hodrick method is used to estimate the covariance matrix of parameter estimates for the restrictive null, and the Newey-West method is used for the general null. Under both null hypotheses the test is asymptotically distributed as a $\chi^2(1)$.

The regression with $\widehat{CP}_t$ have a R^2 of $26 \sim 30$ percent, slightly lower than reported by CP, who use data through 2003. And the result of regressing excess returns on $\widehat{LN}_t$ is remarkably similar to what is obtained by LN. Result of multivariate regressions, in which both CP and LN factors are included, is shown in Panel B. This specification serves as the starting point of our analysis. First of all, $\widehat{LN}_t$ alone explains $0.26 \sim 0.28$ variance in excess bound returns with maturities of 2-5 years. Adding CP_t to the regression increases R^2 to about 0.4. Our SAGLasso methodology is intended to investigate whether these return-forecasting factors have fully captured forecastable variations in excess returns, or equivalently, whether we can construct a macroeconomic factor from the same penal data with predictive power above and beyond the LN factor. Moreover, according to Panel C, the null hypothesis that the first-order autocorrelation in the error term is zero is overwhelmingly rejected, using Ljung-Box refined Q test. Note that Q test is even too conservative when the null hypothesis is false, because it does not conditional independent variables. Autocorrelated disturbance may occur due to misspecification, such as omitting relevant variables, choosing too low a lag order for dependent or independent variables, or using inappropriate transformed variables. But the documented evidence seems too strong to be purely explained by omitted explanatory variables. Hence we proceed with relaxing the nonlinear restriction imposed on essentially general ARDL model, and perform the SAGLasso algorithm on the basis of following model

$$C(L)arx_{t+1} = \beta B(L) X_t + \varepsilon_t \tag{14}$$

The dimension of $B(L)$ and $C(L)$ is set to be 7 (with 6 lags). This original specification is arbitrary and aims to make sure that the model contains more than the true lagged values. Our conjecture is confirmed by the final estimation result (reported in Table 3) that the resulted SAGLasso regression model does not exhibit significant serial correlation.

3.3 SAGLasso: In-Sample Analysis

Following Ludvigson and Ng (2009a), we divide the data into 8 blocks. These are (1) output, (2) labor market, (3) housing sector, (4) orders and inventories, (5) money and credit (6) bond and FX, (7) prices and (8) stock market. Then Adaptive Lasso is conducted with each group, along with lagged value of excess bond returns. The use of Adaptive Lasso provides the flexibility to not penalize the coefficients associated with lagged dependent variables. One merit of estimating Group Lasso is that the factor estimates are easy to interpret. We notice that after performing the first step of SAGLasso, the dimension of the exogenous explanatory variables has been greatly reduced. Only 38 macroeconomic variables have non-zero coefficients on their contemporaneous and/or lagged values. For instance, the largest group at the cluster level, the "labor market group", contains 32 series and thus 224 candidate explanatory variables,[5] but only 11 of them have non-zero coefficients after the group lasso is applied.

In the second step, we select important clusters using the Group Lasso. Yuan and Lin (2006) show that the solution to the Group Lasso problem (6) can be obtained efficiently by using a modified least angle regression selection (LARS) algorithm of Efron et al. (2004). With tuning parameters determined by 2-fold cross validation, coefficients of 4 clusters/groups, (1), (4), (5) and (8), are shrunk to exactly zero. The estimates indicate that macro factors associated with labor market, housing, interest rate and prices show strongest connection with bond risk premia. We will revert back to the implication of this shrinkage result in Section 3.4. Along the lines of CP and LN, the single SAGLasso factor, denoted $\widehat{G}_t$, is defined as $\widehat{\beta B(L)} X_t$.

Table 3 presents results from in-sample predictive regressions of 2- through 5-year long excess bond returns on the SAGLasso factor along with lags of excess returns, in the form of (13). Estimates of regressions with maturities of 3-5 years should be a preferable focus of our interpretation, since a reparametrization of Eq. (11) reveals that for the regression with

[5] For each series, we include its 6 lagged variables; so the total number of variables equals $32 \times 7 = 224$.

the 2-year bond, some of the log prices appear on the both sides of the following equations

$$rx_{t+1}^{(2)} = A_0 rx_t^{(2)} + \beta B(L) X_t + \sum_{i=1}^{5} A_i rx_{t-i}^{(2)}, \tag{15}$$

$$p_{t+1}^{(1)} - p_t^{(2)} + p_t^{(1)} = A_0 \left(p_t^{(1)} - p_{t-1}^{(2)} + p_{t-1}^{(1)} \right) + \beta B(L) X_t + \sum_{i=1}^{5} A_i rx_{t-i}^{(2)}. \tag{16}$$

Note that $\widehat{G}_t$ is basically a linear combination of macroeconomic series as well as their lagged values. The essential difference between the SAGLasso factor and the LN factor is that the former (1) exerts shrinkage of sufficient magnitude to give an economically interpretable model as well as to substantially reduce the forecast variance and (2) takes into account the dynamic respondence of risk premia to macroeconomic innovations. Panel A (and the following Table 4) gives prominence to the first feature by showing that the SAGLasso factor have statistically and economically significant predictive power conditional on lagged returns. $\widehat{G}_t$ is highly significant in forecasting regression of all maturities, implying the estimated factors contain information about future returns that is not contained in its own historical path. This test directly corresponds to the general null hypothesis, as it demonstrates that the lagged returns per se, especially the first lag, have considerable forecasting power for the future ones. More strikingly, the SAGLasso factor, together with lagged excess returns, explains nearly 90 percent of next year's returns of 2- to 5-year Treasury bonds. It seems safe to conclude that this specification has captured nearly all forecastbale variations in excess bond returns and thus leaves marginal scope for other explanatory variable, in favor of using the ARDL framework.

Another support for this specification comes from the result presented in Panel B, which highlights the second feature of our methodology. The Ljung-Box Q test does not detect the presence of AR(1) serial correlation in our model, in sharp contrast to extant studies. It is well known that Q test is less powerful when the null hypothesis does not hold. Breusch and Pagan (1980) also point out that Q test is inappropriate when the regressors include both lagged dependent variables and exogenous variables. Hence, we also perform Lagrangian multiplier test for robustness,[6] but it cannot reject the null hypothesis either. Testing for higher order serial correlation does not alter the result.[7]

Next, we examine whether the SAGLasso factor has unconditional predictive power for excess returns; this amounts to estimating the restricted version of (13), where $A(L)$ is restricted to zero except for the contemporaneous one. Table 4 presents the estimation results of univariate predictive regression on $\widehat{G}_t$. The most striking finding is that the SAGLasso factor could explain more than 43% of the variations in excess returns on 2-5 year maturity bonds. This R^2 statistics, to the best of our knowledge, is the highest in

[6]See Breusch (1978) and Godfrey (1978) for details.
[7]Therefore, we do not report the Newey-West estimator for this specification.

the literature to investigate the predictability of excess returns. Results not reported here show that the lagged values along explains about 62% of the variance in the 3-5 year excess bond returns. In other words, adding $\widehat{G}_t$ to the corresponding AR(6) model increases R^2 by nearly 30%. Combining these numbers with results reported in Table 4, we can see that the SAGLasso factor contains macroeconomic information on bond risk premia that is unspanned by past bond prices. This implication is not surprising, though, because the shrinkage path of our SAGLasso regression is based on partial correlations controlling for lagged dependent variables (lagged values of excess returns are chosen to not be penalized). Economic interpretation of this SAGLasso factor will be further discussed in Section 3.5.

These estimates also confirm that the unsupervised procedures adopted by LN underestimate the predictive power of macro variable (with R^2 of 0.28 versus 0.43), as much useful information is missed in the process of unsupervised factor analysis. This point appears more prominent in Panel B, which presents the results of regressions with both $\widehat{G}_t$ and $\widehat{LN}_t$ as predictors. We find adding $\widehat{LN}_t$ into the regression make little improvement in the predictability, measured by R^2. Moreover, whether HH or NW t-statistics indicate that the LN factor does not have the conditional predictive power for excess bond returns, as it becomes completely insignificant once $\widehat{G}_t$ is included in the regression. Hence we conclude that our SAGLasso factor contains most macroeconomic information on term premia and thus absorbs the role of LN factor in forecasting excess bond returns.

3.4 SAGLasso: the Group Level

A major advantage of SAGLasso is that it allows for using a priori information to organize the numerous time series into several cluster. Unlike data obtained in other field, such as microarray data, macroeconomic data usually has natural hierarchical structure that may enhance our interpretation of empirical results. And the Group Lasso in the second step produces accurate prediction while selecting a subset of important factors (clusters). This subsection aims to evaluate the individual predictive power of each group factor and characterize their connections with the bond risk premia.

Only 21 macroeconomic variable have non-zero coefficients associated with their contemporaneous and/or lagged values after two-step screening. See the column labeled "$\widehat{G}_t$" in Table A.1 for details. A quick review of these macro variables with non-zero coefficient validates our use of ARDL model, as many series have lagged effect on risk premia of Treasury bonds. Especially, shocks to consumer prices require a long lag to manifest their impact on the bond market.

To provide a clearer picture of the relationship between excess bond returns and each macroeconomic cluster, we first form 4 group factors using estimates from the Group Lasso

$$\hat{g}_{ht} = \tilde{X}_h \tilde{\beta}_h, h = 2, 3, 6, 7. \tag{17}$$

Recall that these four supervised macroeconomic factors correspond to employment, housing, interest spread and inflation, according to the result yielded in the second step of SAGLasso. We then examine the predictive power of these group factors.

Table 5 presents results from in-sample predictive regressions of 2- through 5-year long excess bond returns on these group factors (relabeled as $\hat{g}_{ht}, h = 1, \ldots, 4$ in tables thereafter). Panel A reports the estimates of univariate regression. We find that all group factors exhibit significant unconditional predictive power. Each of them alone can explain 20 to 30 percent of variance in excess returns on bonds of various maturities. Results reported in Panel B are based on regressions that include four group factors as predictors. This specification can be viewed as the counterpart of Table 1 using unsupervised factors. An inspection of the results in Tables 2 and 3 reveals that, using the same panel of macro data, our SAGLasso approach uncovers 4 cluster factors with the R^2 significantly higher than that in the LN regression (0.43 versus 0.29). This evidence confirms the first two major conjectures discussed in Section 2. In the first place, the unsupervised PCA does yield some useless factors. Even if subset selection based on BIC is performed to select predictors, we still find some insignificant factors in the forecasting regression ($\hat{f}_{3t}$ and $\hat{f}_{5t}$), largely due to the discrete nature of subset selection methods. In the second place, some important information on term premia is missed in unsupervised factors. Although the SAGLasso regression only yields 4 group factors with non-zero coefficients, they can explain more of the variations in excess bond returns. Note that unlike the estimates from large-scale factor analysis, our group factors have clear economic interpretations.

Finally, we investigate whether these group factors have predictive power conditional on Cochrane-Piazzesi return forecasting factor $\widehat{CP}_t$, a common benchmark in the literature. The estimation result shows $\hat{g}_{ht}$s have statistically significant and economically important predictive power beyond that contained in the forward-rate factor $\widehat{CP}_t$. For the regression of excess return for 2-year bonds, the R^2 statistic rises from 0.26 to 0.48 once $\hat{g}_{ht}$s are included in the regression. Unlike the $\hat{f}_{2t}$ in LN's regression, our group factor $\hat{g}_{6t}$, which is estimated from the "bond spreads" block, does not lose its marginal predictive power when $\widehat{CP}_t$ is added as a predictor. On the other hand, $\widehat{CP}_t$ remains significant in the multivariate regressions, implying that the macroeconomic factors do not subsume its role in predicting excess bond returns.

Cooper and Priestley (2009) document that the output gap can predict excess returns on 2- to 5-year maturity bonds with R^2 equal to 2 percent, where the output gap is measured as the deviations of the log of industrial production index from a quadratic and linear trend. The top two panels in Table 6 replicate the results on the predictive power reported in Cooper and Priestley, where $\widehat{CP}_t^{\perp}$ is the Cochrane-Piazzesi (2005) factor orthogonalized relative to *gap*. Results reported in the bottom panel of Table 6 show that our employment group factor can subsume the output gap factor.

3.5 SAGLasso: a Countercyclical Component

Given the special statistical relevance of the SAGLasso factor in predicting excess returns, it seems interesting to investigate its economic implication. We begin with characterizing its countercyclic pattern, as predicted by economic theory. In the next subsection, we interpret its role in yield dynamic and show that it contains information about risk premia that is not already embedded in bond market data.

Wachter (2006) generalizes the habit formation model of Campbell and Cochrane (1999) and shows that bond risk premia covary with consumption surplus, which is driven by shocks to aggregate consumption. As documented in LN about their macro factor, our SAGLasso also captures the countercyclic component in risk premia. Figure 1 plots the 6 month moving average of $\widehat{G}_t$ and the growth rate of industrial production (GIP). Shaded areas indicate the periods designated by the National Bureau of Economic Research (NBER) as recession periods, where are characterized by low growth rate of IP and high values of $\widehat{G}_t$. The figure shows that $\widehat{G}_t$ is strongly negatively correlated with GIP, with a correlation coefficient of -0.78. The SAGLasso factor falls to troughs in the mid-to-late stage of economic expansions and reaches its peaks at the end of recessions.

To illustrate that it captures more countercyclical variation in real activity than that documented by existing studies, we follow LNs methodology by including different sets of state variables in a VAR system to calculate multiperiod forecasts of excess returns. Specifically, we consider two benchmarks, which we compare to our specification including the SAGLasso factor $Z_t^{SL} = [rx_t^{(5)}, rx_t^{(4)}, rx_t^{(3)}, rx_t^{(2)}, \widehat{CP}_t, \hat{G}_t]'$. The first benchmark amounts to a restricted VAR model that excludes our SAGLasso factor $Z_t^{CP} = [rx_t^{(5)}, rx_t^{(4)}, rx_t^{(3)}, rx_t^{(2)}, \widehat{CP}_t]'$. For comparison, we also construct bond forecasts with a LN benchmark version, which contains their single forecasting factor as well as all variables in the first benchmark $Z_t^{LN} = [rx_t^{(5)}, rx_t^{(4)}, rx_t^{(3)}, rx_t^{(2)}, \widehat{CP}_t, \widehat{LN}_t]'$. And a VAR(12) model will be estimated separately with each of the three state vectors.

Note that all three model specifications involve $\widehat{CP}_t$, as neither our SAGLasso factor nor LN's factor subsume the information in CP's forward rate factor, as implied in Table 2 and Table 5. Table 7 reports similar results obtained from the following regression

$$rx_{t+1}^{(n)} = \beta_0 + \beta_1 \widehat{G}_t + \beta_2' \widehat{CP}_t + e_{t+1}. \tag{18}$$

Both HH and NW t-tests overwhelmingly reject the hypotheses that each factor can be excluded from the joint regression, indicating that $\widehat{G}_t$ and $\widehat{CP}_t$ are picking up different sources of predictability. Especially, when both factor are included in the regression, the R^2s are much higher than they are when only one is included. Moreover, even if lagged excess returns are incorporated into the regression, the CP factor still has significant predictive power for the 4-year bond. Its significance in predicting excess returns across different maturities is curved at the long end, closely resembling the pattern documented in CP. This suggests that the CP factor does summarized information on risk premia at the long end.

Figure 2 plots estimates of five-year bond premium $E[rx_{t+1}^{(5)}]$ verses time over the sample period, where $\hat{E}[rx_{t+1}^{(5)}]$ is obtained by solving the VAR forward to create forecasts of future at monthly horizons. Panel A shows the estimated risk premium using Z^{SL}, while Panel B displays the same estimates with Z^{CP}. LN has shown that the risk premia exhibits greater countercyclicality than in the absence of macroeconomic factors. Here we confirm that this result is robust regardless of the methodology used to extract macroeconomic information. Careful contrast of the two panels implies that the difference mainly arises in recession periods, when return premia estimated with SAGLasso factor increases dramatically. Their correlations with IP growth are -0.263 and 0.034, respectively.

We also form an estimate of the term premia

$$\widehat{TP}_t^{(n)} = \frac{1}{n} \sum_{i=1}^{n-1} \hat{E}_t \left(rx_{t+i}^{(n+1-i)} \right) \tag{19}$$

as the average of obtained estimates of return risk premia of declining maturity. This time we compare our SAGLasso-based estimates of term premia to the corresponding values generated from the LN benchmark. The six-month moving average of two estimated premium components in the five-year bond yield are plotted over time in Figure 3. We find both series exhibit a similar pattern in the sense that the yield risk premium tends to rise over the course of a recession and peak just after the recession period. However, term premia are more countercyclical and reach greater values in recessions when macroeconomic information is summarized by the SAGLasso factor. For example, in the recession of early 1980's the difference in term premia reached a level of 0.83% per annum. Indeed, the former estimate has a contemporaneous correlation of -42.8% with the growth of IP and the latter has a correlation of -37.2%. Generally, these findings are consistent with the implications of the general equilibrium model developed in Wachter (2006).

3.6 Unspanned Predictability

Besides CP, empirical research in dynamic term structure models has revealed the presence of another predictor that is unspanned by cross section of yields. For example, Duffee (2008) uses Kalman filtering estimation and finds evidence of a "hidden" factor that has an imperceptible affect on yields but nevertheless has substantial forecasting power for future yields and returns; Barillas (2009) estimated a macro-finance model with a unspanned risk factor, which appears to add substantial predictability to beyond what is already contained in the term structure. In this subsection, we demonstrate that our SAGLasso factor shares the same property and further show its leverage over macroeconomic indicators.

We firstly examine the cross sectional relation between the SAGLasso factor and bond yields. Consider the following regression

$$\Delta y_t^{(n)} = \theta_0 + \theta_1 \widehat{G}_t + \varepsilon_t, \quad n = 1, \ldots, 5. \tag{20}$$

The percentage of yield changes variance explained by the SAGLasso factor is defined as $100 \times trace(cov(\theta_1 \widehat{G}_t))/trace(cov(\Delta y_t))$. We find $\widehat{G}_t$ explains only 0.09% of the variance of yield changes, indistinguishable from noise in yields. As expected, although the SAGLasso factor contains substantial information about future excess bond returns, its contribution to the overall volatility of the cross section of bond yields is imperceptible. Put differently, our empirical evidence supplements Duffee (2008)'s argument by uncovering another "hidden factor" that comes from the macroeconomy.

Indeed, we can decompose variation in bond yields into expectations and term premium components

$$y_t^{(n)} = \frac{1}{n} E_t \left(\sum_{i=0}^{n-1} y_{t+i}^{(1)} \right) + TP_t^{(n)}. \tag{21}$$

If aggregate risk aversion is time varying in response to both news about aggregate consumption growth and news about inflation, as indicated by Campbell and Cochrane (1999), shocks to consumption growth (or an unexpected increase in inflation) temporarily raises risk aversion and cause agents to demand greater premia on risky long-term bonds. On the other hand, investors believe that the Fed will attempt to offset these short-lived macroeconomic shocks with monetary policy actions. It in turn drives down the expectations of future short rate. Thus the net effect of the macro shocks on current yields becomes insignificant because the expected change in short rates and the change in risk premia have opposite effects. In fact, our SAGLasso factor is also found (results not reported) to have a significant forecasting power for 12-month-ahead 3-month annualized bill yield, with a NW t-statistic of -2.54.

To demonstrate that our SAGLasso factor truly captures market expectations, we plot the impulse response functions (IRF) obtained from factor-augmented vector autoregressive models (FAVARs). A generalization of dynamic factor models, FAVAR models the joint dynamics of r unobservable factors (F_t) and a small number of observable economic variable (Y_t) of our ultimate interest.

$$Z_t = \Phi(L) Z_{t-1} + u_t \tag{22}$$

where $Z_t = (F_t' Y_t')$ is of dimension $r + m$. As inspired by Bernanke et al. (2005), F_t can be interrelated as theoretically motivated economic concepts that is hardly measurable. In this context, the presence of dynamic factors in the model is to capture information set of policy-makers and the private sector, which is not contained in a few predetermined economic indicators (Y_t). Therefore, for our application the interpretation of F_t is the focus of our analysis. Instead, we are interested in uncovering the structural relationships among economic variables in Y_t. To extract F_t, we make use a large information set, i.e. the 131 macroeconomic time series X_t. Specifically, we related X_t to Z_t by an observation equation of the form

$$X_t = \beta Z_t + \epsilon_t. \tag{23}$$

The model is estimated using likelihood-based Gibbs sampling, as outlined by Bernanke et al. (2005). As the Bai-NG information criteria indicate that 8 factors are need to capture the majority information in X_t, we set the number of unobservable factors as $8 - m$, where m is the dimension of Y_t.

Suppose an econometrician has extracted the variable that reflects investor's expectation of Fed's response to macroeconomic shocks, following our SAGLasso procedures. In our first specification, Y_t consists of IP growth and the SAGLasso factor. If the the SAGLasso factor does denote financial market participants' risk aversion in the presence of macroeconomic release (or equivalently, their belief of Federal Reserve's reaction), a negative shock to output growth will lead to an immediate increase in the SAGLasso factor. And the Panel A in Figure 4 matches this expectation indeed: the SAGLasso factor dramatically drops following a positive innovation in IP growth.[8]

Another attractive feature of the FAVAR framework is that IRF can be constructed for any observable economic variables included in X_t. In order to highlight the effect of the market expectation on short rate and term premia, we estimate another FAVAR specification in which SAGLasso is the only variable contained in Y_t. Panel B reports the IRFs, along with 90 percent confidence intervals, of a selection of economic indicators to a one-standard-deviation shock in SAGLasso.

First of all, let us focus on the response of the 3-month Treasury bill yield. The SAGLasso factor has zero effect on the short rate the Month 0, but it results in a nearly 40-basis-points drop in short rate after 18 months. Moreover, the response of the short rate is quite persistent, remaining 25 basis points 4 years later. In contrast, the response of 5-year bond yield seems insignificant throughout all periods, especially for the first 18 months in which the IRF is flat. By the definition of excess returns, responses of short rate and long-maturity bond yield jointly imply an increase in the risk premium. To sum up, the estimated IRF make it glaringly apparent that the SAGLasso factor has no contemporaneous effect on the term structure but contains substantial information about expected excess bond returns.

Another direct evidence for the expectation nature of SAGLasso comes from the last subpanel, which plots the IRF of (UMichigan) Consumer Expectation Index. The immediate effect of SAGLasso on the index confirms that the unspanned predictability of SAGLass is associated with market expectations. Finally, the responses of other variables are generally of the expected sign and magnitude. Particularly, variables that typically exhibit stickiness,

[8] This inverse relation between IP and the hidden factor is also documented by Duffee (2008), who employs a simple regression analysis.

such as CPI, investment and Unemployment, do have slow-moving responses to SAGLasso, while housing start index and dividend, which are sensitive to future changes in short rate, initially jump and eventually wear off.

As such, both our SAGLasso factor and Duffee (2008)'s "hidden" factor represent components of risks borne by investors which are orthogonal to the yield curve. However, the two factors do not subsume each other, as indicated by the results reported in Table 8. Panel A of the table confirms Duffee's finding that his five latent factors extracted from the yield data contain information about future bond excess returns, even more than the forward rates do (see panel A of Table 2). The R^2s here ranging from 0.32 to 0.35 are slightly lower than those documented in Duffee (2008), who use data through 2006. Comparing these R^2s with those reported in panel A of Table 4, we can see that the SAGLasso factor has stronger predictive power than Duffee's five factors do. Results reported in panel B show that when both sets of factors are used in regression, the R^2s rise to nearly 50 percent, much higher than the values when only one set of factors is used. Moreover both SAGLasso and the "hidden" yield factors remain highly significant, implying they are not measuring the same component of bond risk premia. The implication of this result is that macroeconomic risk underlies, but does not perfectly capture, the variation in excess returns that is not contained in the first three yield principal components. Therefore, both unspanned "yield-curve" risk and unspanned "macro" risk are priced and bond risk premia are not fully identified in absence of either information set. This evidence also sheds light on DTSMs in determining the dimension of risk factors driving expected excess returns, which is not necessarily the same as the dimension of state vector that prices the relevant universe of bonds.

3.7 Realized Jump Risk and Subsample Analysis

Wright and Zhou (2009) show that some realized jump measures explain a nontrivial fraction of the variation in post-1984 excess bond returns. In particular, they find that inclusion of a rolling realized jump mean into the benchmark predictive regression on forward rates nearly doubles the R^2. In this section, we aim to identify the source of realized jump risk and to what extent the information in jump measures has been contained in known macro and yield predictors.

Consider the following baseline regression in Wright and Zhou (hereafter WZ):

$$rx_{t+1}^{(n)} = \beta_0 + \beta_1 RV_t^1 + \beta_2 JI_t^{24} + \beta_3 JM_t^{24} + \beta_4 JV_t^{24} + \beta_5 F_t^{(1)} + \beta_6 F_t^{(3)} + \beta_7 F_t^{(5)} + \varepsilon_{t+1}. \quad (24)$$

where $F_t^{(n)}$ is the n-year forward rate used in CP, RV_t^1 monthly realized volatility, JI_t^{24}, JM_t^{24}, and JV_t^{24} denote 24-month rolling average realized jump intensity, jump mean, and jump volatility, respectively. The latter four measures are constructed using data on 30-year Treasury bond futures at the five-minute frequency from CBOT. However, as the high-frequency Treasury bond future data are not available until July 1982, the sample period for the 24-month rolling jump mean starts only from August 1984.

Table 9 reports coefficient estimates, associated HH and NW t-statistics and (adjusted) R^2 values for several specifications of the form of Eq. (24). Panel A, the counterpart of WZ's Table 2, shows that the jump mean is the most significant predictor among various volatility and jump risk measures, with the R^2 ranging from 14 to 16 percent. The negative coefficients imply that a downward realized jumps in long bond would cause short term bond prices to appreciate over the next year. Other results are almost identical to those documented by WZ.[9] Panel B replicates their main result that augmenting CP's regression model with jump mean nearly doubles the predictability, with the adjusted R^2 increased to about 60 percent. WZ argue that it implies that forward rates and jump mean are not picking up the same predictability, and that the latter complements the information content of the former.

Before we conduct a comparative analysis of return predictability using jump mean and other known instruments, it would be helpful to provide some insight about the source of realized jumps in bond markets. Intuitively, most of them are associated with macroeconomic announcements. Using jump-diffusion term structure models, for instance, Das (2002) and Johannes (2004) discover direct connection between model-implied jumps and macroeconomic shocks. As the high-frequency data based method employed by WZ should result in more timely jump filtering, we could reestablish this connection by identify the events that caused the detected jumps.

Table 10 compiles a list of the 20 biggest realized jumps during our sample period, and major news events on the jump dates. We can observe that each of these jumps coincides with unexpected macroeconomic news arrivals. As documented by Johannes (2004), there are three major source of jumps: (1) leading economic indicators releases such as unemployment announcements, (2) official announcements on monetary policy such as Federal Open Market Committee target, and (3) exogenous political events regarding the nation's vital interests, e.g. the breakout of Gulf War. Judging from the frequency, most jumps are generated equally by regularly scheduled announcements, consistent with the findings of Fleming and Remolona (1997) and Balduzzi et al. (2001). Judging from the absolute jump size, unemployment announcements produce those largest moves, implying surprises regarding the current state of the real economy significant affect long-term bond prices.

These results give an economic interpretation of realized jumps that provide the mechanism through which macroeconomic shocks enter the Treasury bond market. Therefore, it is of interest to examine whether the information in jump mean has been subsumed by the SAGLasso factor or other known predictors. Coincidentally, the high-frequency Treasury bond future data availability results in a subsample that exactly covers the post monetary experiment period. Since there is considerable evidence of a regime switch during the late 1970s and early 1980s, we firstly investigate the subsample performance of our SAGLasso

[9] Our sample period is slightly (16 months) longer than that used in WZ.

factor and filtered yield factors.

Table 11 reports a breakdown by subsamples of regressions of bond excess returns with different maturities rx_{t+1} on the SAGLasso factor and Duffee's five yield factors. The SAGLasso factor appears to have less explanatory power for bond risk premia, and the univariate R^2 drops to about 30 percent from more than 40 percent from the full sample. In contrast, the filtered factors account for a larger proportion of term premia than they do in the full sample. The enhanced predictive power of the yield factors mainly comes from the fourth factor, which becomes particularly significant especially in regressions of long-term bond returns.

These results point toward an important indication that after the Fed changed operating procedures, interest rates were less volatile, so macroeconomic instruments, especially inflation measures, became less correlated with bond risk premia. Consequently, latent factors that require filtering may capture an more important component of risk premia in the post experiment period. We also run forecasting regressions using group factors, the result (not reported here) also implies that the fourth (inflation) factor turns out significant in the subsample regressions, consistent with our conjecture. However, the combined predictive capacity of these two sets of predictors remains almost unchanged during the subsample period, with R^2 of 45-50 percent. These results contrast with those reported by some recent work that most of the predictability came in the 1970s and 1980s and there is little if any significant predictability in the post-1985 period.

The more important question here is whether realized jump mean and the SAGLasso factor are picking up the same predictability. The evidence in the Panel A of Table 10 says that they are not. Augmenting our SAGLasso regression with WZ's estimates of jump mean raises the R^2 to 39 percent. Moreover, the coefficients on SAGLasso factor and jump mean do not change greatly when they enter multivariate forecasting regressions. It implies that unlike our macroeconomic factor the realized jump measures capture a high-frequency relation between macroeconomic variables and bond yields. On the other hand, the information content of the filtered yield factors seems to be orthogonal to that of jump mean in that the R^2 (0.62) of the regression on both jump mean and forward rates is larger than the sum of the R^2s on each set of the variables separately. For completeness, the Panel B reports the regression results using all of these instruments shown to have significant predictive power. The SAGlasso factor, Duffee's hidden factor and WZ's jump mean remain significant and jointly explain 65% of the variation in excess bond return.

What is the implication of our finding for affine term structure model? The results reported in Table 12 indicate that there are two primary conduits through which information about the macroeconomic enters the term structure. One is captured by our SAGLasso factor which describes a low-frequency relation between macroeconomic variables and yields, while realized jumps measure the other one which how yields directly respond to unexpected shocks from real economy and monetary policy. Absence of any one of them in term structure models may induce misspecification. These findings are not taken into consideration by commonly employed affine term structure models where the forecastability of bond returns is completely summarized by the cross-section of yields. The recent macro-finance models take into account the first conduit by allowing macroeconomic factors to enter state variables. But a further extension is required to deal with the jump-induced misspecification.

However, even a jump-diffusion macro-finance model may not be able to account for all of these empirical regularities. As discussed by WZ, the implication that jump risk factors are not spanned by the current yield curve reminds may have something to do with unspanned stochastic volatility documented in Collin-Dufresne and Goldstein (2002).

3.8 SAGLasso: Out of Sample Analysis

A common concern in tests of predictability is that significant in-sample evidence of predictability does not guarantee significant out-of-sample predictability. This evidence is often interpreted as an indication that in-sample evidence is likely to be spurious and should be discounted. In this subsection we report results on out-of-sample forecasting performance of factors estimated in the previous section. To avoid involving future information, we carry on fully recursive factor estimation and parameter estimation with data only through time t for prediction at time $t + 1$. The following description of the procedure applies to forecasts of annual excess returns with factors estimated using SAGLasso. The notation of Clark and McCracken (2001) is used here.

The in-sample observations span 1 to R. That is, observations 1 through R of the macro variables and observations $rx^{(n)}_{1,13}$ through $rx^{(n)}_{R,R+12}$ of annual excess returns are used to estimate the SAGLasso factors $\widehat{G}_t$ and the return-forecasting regression. Given the estimated parameters, forecast $rx^{(n)}_{R+12,R+24}$ using $\widehat{G}_{R+12}$ as the linear combinations of the macro variables for observation R+12. Denote the realized forecast error by u(n) un,1, where the first subscript refers to a forecast error from an unrestricted regression. Collect the realized forecast error $u^{(n)}_{u,1}$, where the first subscript refers to a forecast error from an unrestricted regression. Then repeat this exercise using an additional observation, recompute the supervised factors and estimate the new regression using observations 1 through R + 1, and so on. Letting P denote the number of 1-step ahead predictions, the out-of-sample observations span R + 1 through R + P. To account for the 12-period overlap induced from continuously compounding monthly returns to obtain annual returns, the length of the resulted time series of forecast errors $u^{(n)}_{u,t}$ is $P = T - R - 11$, where T is the total number of observations of macroeconomic variables.

The time series of restricted forecast errors $u^{(n)}_{r,t}$ can be constructed using the same methodology, where the forecasting regression uses a constant term as a benchmark, apart from an MA(12) error term. The first subscript refers to a forecast error from a restricted regression. But the restrictive hypothesis has already been strongly rejected even in out-of-sample analysis. In order to investigate whether the SAGLasso factor has additional predictive power conditional on information contained in past returns, we compare the out-of-sample forecasting performance of our shrunk ARDL model to a simple AR(6) specification. This second specification can be viewed as the out-of-sample counterpart of the specification (11). To assess the incremental predictive power of the SAGLasso factor above and beyond the predictive power in $\widehat{LN}_t$, we conduct another model comparison. We compare the out-of-sample forecasting performance of a specification that includes the SAGLasso factor plus to the LN factor to a benchmark model that includes just the LN factor.

For the full sample 1964:1 through 2007:12, the initial estimation period span 1964:1 to 1984:12 (in terms of independent variables) so that R=252, thus $P = 265$ for annual excess returns. The choice of the length of in-sample portion is arbitrary, but alternative choices do not lead to qualitatively different results. Table 13 reports test statistics for the ENC-REG test of Ericsson (1992) and the ENC-NEW test of Clark and McCracken (2001). Both tests examine the null hypothesis that the benchmark model encompasses the unrestricted model with additional predictors. The motive for using two encompassing tests is for that Ericsson test critical values from a standard normal distribution are conservative if $\pi = \lim_{P,R\to\inf} P/R > 0$.[10] As the asymptotic ratio of P/R is unknown for our case, the Ericsson test is used as a robust check.

The column labeled "Ericsson" in Table 13 reports the ENC-REG test statistic and its 95 percent critical value is 1.645. Similarly, column "Clark-McCracken" presents the ENC-NEW test statistic and its 95 percent critical value is 1.584. The results show that the forecasting model including the SAGLasso factor improves remarkably over the constant expected returns benchmark. Specifically, when the supervised factors are incorporated, the model is shown to have a forecast error variance that is only 72 percent of the constant expected returns benchmark for $rx^{(2)}_{t+1}$. Tests for the forecasting model with the SAGLasso factor versus a simple AR(6) benchmark, presented in Panel B, rejects the general null hypothesis as well. More strikingly, incorporating the SAGLasso factor into the original AR(6) model produces a mean-squared error that is anywhere form 62 to 64 percent of the autoregressive benchmark mean-squared error. Overall, both null hypotheses are completely rejected no matter which test is conducted. But note that the Ericsson test may not appropriate to be applied to the general null because the forecasts are not truly out of sample.

Panel C indicates that the model including the SAGLasso factor and $\widehat{LN}_t$ improves substantially over a benchmark that includes a constant and $\widehat{LN}_t$, consistent with Table 4. Both test statistics indicate that the improvement in forecast power is strongly statistically significant, at the one percent or better level. Moreover, the models have a mean-squared

[10]The precise asymptotic distribution of the test statistics for both Ericsson and Clark-McCracken tests depends on the π, the asymptotic ratio of P/R.

error that is anywhere from 87 to 93 percent of the LN expected returns benchmark mean-squared error. To sum up, this result reinforces the conclusion form the in-sample analysis, i.e. the SAGLasso factor contains information about future returns that is not captured by the LN factor.

4 Finite-Sample Property

In this section, we proceed with a finite-sample analysis by generating bootstrap samples of the yields as well as of the exogenous predictors, under both restrictive and general null hypotheses. Only regressions on four group factors estimated from SAGLasso are examined here because the Monte Carlo simulations involves a term structure model to specify the dynamics of yields and state variables. For the term structure model used to generate simulation data under the general null, the number of latent term premia factors must equal the number of macro factor. Thus only one macro factor and one latent factor may not fully capture the joint variation in the macro variables and yield data.

4.1 Data-Generating Processes

Stambaugh (1999) and Ferson, Sarkissian, and Simin (2003) discuss why small-sample inference is essentially important in our context. The two reasons identified by them for the deviation of finite-sample properties from asymptotic properties are relevant to our forecasting regressions. First, if the standard instruments employed as predictors are highly persistent and/or contemporaneously correlated with the idiosyncratic noise in returns, serious over-rejection could result. For our factors estimated with SAGLasso, some display high persistence, with first-order autoregressive coefficient up to 0.9282. Second, when overlapping observations are used in constructing regressands, estimates of standard errors for regression coefficients show strong bias. This concern also applies to our regression analysis, as computing annual excess returns involves overlapping yields data.

In the literature, there are two major methods used to generate simulation data for finite-sample inference. One approach is to construct a time-series model for the log yields and rerun the predictive regressions.[11] Residuals are bootstrapped to form the empirical distribution. Note that for this method the independent variables (the macroeconomic factors) are not simulated and enter the predictive regression as their actual values. The other one is based on a dynamic term structure model that satisfies either the restrictive or the general null hypothesis. For this data-generating process, state variables are also simulated from the term structure model using parameters estimated from the actual sample data. Duffee (2007) constructs a class of discrete-time term structure models to make finite-sample

[11] For example, CP (2005) run a unconstrained VAR(12) model of all 5 yields, and LN (2008) use a MA(12) process to test the restrictive null.

inference for both restrictive and general null hypotheses.

In our case, the two simulation procedures yield almost indistinguishable empirical distributions for our test statistics, as shown in Table 14. For the general null hypothesis, along the spirit of the first approach we run a VAR(12) for the yield process that imposes a single unit root (one common trend). With the second approach we construct a term structure model in which macro factors determine the cross section of short rate but the price of risk does not depend on macro factors (c.f. the appendix).

Though in different forms, both data-generating processes satisfy the general null hypothesis that the the excess returns may not be constant but their time variation has nothing to do with the macroeconomic factors. A close scrutiny of our simulation data suggests the reason from this consistent result. First of all, as the group factors estimated from actual data are highly persistent variables (the most persistent factor has a monthly AR(1) coefficient of 0.9282), the Gaussian VAR(1) process, specified in the term structure model, may be a good approximation for their joint dynamics. More important, as both data-generating processes satisfy the general null of independence, what is relevant for the test statistics is the time series properties of the excess returns, instead of how the returns are generated (VAR or term structure model), because that is how the distribution of our test statistics are determined. Indeed, in the two different settings that we use to generate excess bond returns, they give similar time series properties (when we plot the average ACFs for the simulated excess returns generated by two simulation model, they look very close).

Though not reported here, the finite-sample distributions generating by these two approach under the restrictive null are also qualitatively identical. Therefore, for our small-sample inference we employ a vector moving average (VMA) model of order 12, namely, a VMA(12) model, to form the bootstrap samples under the restrictive null hypothesis, and use an eight-factor term structure model, specified in the appendix, for the general null. The use of VMA model is because the monthly bond price data used to construct continuously compounded annual returns induces an MA(12) error structure in the annual excess returns. For the term structure model-based simulations, an initial draw of the state variables is taken from their unconditional multivariate normal distribution. Subsequent draws use their conditional multivariate normal distribution. The finite-sample distributions are constructed based on 50,000 Monte Carlo simulations, and the length of each simulation is 528 months, the same as the length of the full sample used in our empirical analysis.

4.2 Bootstrapped Results

The finite-sample properties based on simulation are reported in Table 15. The first three columns specify the details of simulation: the type of regression (in-sample or out-of-sample), the test used, and the bond maturity. The column labeled "Rejection Rates" presents finite-sample rejection rates of tests of the null hypothesis when using the asymptotic five percent critical value, which is 9.49 for Wald test, 1.645 for Ericsson test and 3.007 for Clark-McCracken test. The "5% CV" column reports true finite-sample critical values at a five percent rejection rate. To test the restrictive null hypothesis we drawing random samples from the empirical distribution of the residuals from a VMA (12) model, so that annual returns are forecastable up to MA (12) error structure. Note that for the restrictive null we pre-estimate the supervised factors by re-sampling the $T \times N$ panel of data. This procedure creates bootstrapped samples of the predictors themselves.

The main conclusion is that the results based on bootstrap inference are broadly consistent with those based on asymptotic inference in Tables 4 and 12. Our results support Duffee's finding that small-sample distributions of test statistics associated with the general null markedly diverge from their counterparts under the restrictive null, as well as the asymptotic distributions of these test statistics. However, even after adjusting the estimated test statistics for their finite-sample properties, all regression evidence of return predictability presented in Section 4 remains robust. For example, the actual values of Wald, Ericsson and Clark-McCracken test statistics are 67.5, 3.36, and 129, respectively, all of which are much higher than the 95% small-sample critical values even under the general null, namely, 19.2, 2.23 and 26.9. Indeed, both null hypotheses are rejected regardless of the type of regression or the test used.

Table 16, the counterpart of Table 5 based on finite-sample distributions, reports the evidence for the in-sample analysis. The magnitude of predictability found in historical data, measured in R^2s and χ^2 tests, is too large to be accounted for by sampling error in samples of the size we currently have. And P-values computed with the empirical distributions of 50,000 bootstrapped samples are almost zero as well.

5 Conclusion

Although it is now believed that expected excess bond returns are time-varying and forecastable, the empirical evidence on the correlation between term premia and their macroeconomic underpinnings is mixed so far. In this paper, we reassess the predictive power of macroeconomic indicators using the supervised adaptive group lasso (SAGLasso) approach, a new model selection approach that allows us to explore the underlying structure of macroeconomic variables with respect to risk premium in bond markets. Our empirical analysis provides new and robust evidence on the explanatory power of macroeconomic fundamentals for variations in excess bond returns, which is even stronger than previously documented in the literature. Furthermore, we find evidence of an unspanned predictor extracted from macroeconomic variables. Overall, our study provides further support for the implication from Ludvigson and Ng (2009b) that we should look beyond observable bond yields when building term structure models, as well as predicting future returns.

References

Ang, A. and M. Piazzesi (2003), "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables," *Journal of Monetary Economics*, vol. 50, 745–787.

Bai, J. and S. Ng (2008), "Forecasting economic time series using targeted predictors," *Journal of Econometrics*, vol. 146, 304–317.

Bair, E., T. Hastie, D. Paul, and R. Tibshirani (2006), "Prediction by supervised principal components," *Journal of the American Statistical Association*, vol. 101, 119–137.

Balduzzi, P., E.J. Elton, and T.C. Green (2001), "Economic news and bond prices: Evidence from the US Treasury market," *Journal of Financial and Quantitative Analysis*, vol. 36, 523–543.

Barillas, F. (2009), "Macroeconomic Releases and the Term Structure of Interest Rates," *Working Paper, New York University*.

Bernanke, B.S., J. Boivin, and P. Eliasz (2005), "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach," *Quarterly Journal of Economics*, vol. 120, 387–422.

Breusch, T.S. (1978), "Testing for autocorrelation in dynamic linear models," *Australian Economic Papers*, vol. 17, 334–335.

Breusch, TS and AR Pagan (1980), "The Lagrange multiplier test and its applications to model specification in econometrics," *The Review of Economic Studies*, vol. 47, 239–253.

Campbell, J.Y. and J.H. Cochrane (1999), "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of political Economy*, vol. 107, 205–251.

Campbell, J.Y. and R.J. Shiller (1991), "Yield spreads and interest rate movements: A bird's eye view," *The Review of Economic Studies*, vol. 58, 495–514.

Clark, T.E. and M.W. McCracken (2001), "Tests of equal forecast accuracy and encompassing for nested models," *Journal of Econometrics*, vol. 105, 85–110.

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

Cochrane, John H. and Monika Piazzesi (2009), "Decomposing the Yield Curve," *SSRN eLibrary, http://ssrn.com/paper=1333274*.

Collin-Dufresne, P. and R.S. Goldstein (2002), "Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility," *The Journal of Finance*, vol. 57, 1685–1730.

Cooper, I. and R. Priestley (2009), "Time-varying risk premiums and the output gap," *Review of Financial Studies*, vol. 22, 2801.

Das, S.R. (2002), "The surprise element: Jumps in interest rates," *Journal of Econometrics*, vol. 106, 27–65.

Duffee, G.R. (2002), "Term premia and interest rate forecasts in affine models," *The Journal of Finance*, vol. 57, 405–443.

Duffee, G.R. (2007), "Are variations in term premia related to the macroeconomy," *WP, John Hopkins*.

Duffee, G.R. (2008), "Information in (and not in) the term structure duffee," *WP, John Hopkins University*.

Duffee, G.R. and R. Stanton (2004), "Estimation of dynamic term structure models," *WP, UC Berkeley*.

Duffie, D. and R. Kan (1996), "A yield-factor model of interest rates," *Mathematical Finance*, vol. 6, 379–406.

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani (2004), "Least angle regression," *Annals of statistics*, pages 407–451.

Ericsson, N.R. (1992), "Parameter constancy, mean square forecast errors, and measuring forecast performance: An exposition, extensions, and illustration," *Journal of Policy Modeling*, vol. 14, 465–495.

Fama, E.F. and R.R. Bliss (1987), "The information in long-maturity forward rates," *The American Economic Review*, vol. 77, 680–692.

Fama, E.F. and K.R. French (1989), "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, vol. 25, 23–49.

Ferson, W.E., S. Sarkissian, and T.T. Simin (2003), "Spurious regressions in financial economics?" *The Journal of Finance*, vol. 58, 1393–1414.

Fleming, M.J. and E.M. Remolona (1997), "What moves the bond market?" *Research Paper-Federal Reserve Bank Of New York*.

Godfrey, LG (1978), "Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables," *Econometrica*, vol. 46, 1293–1301.

Hansen, L.P. and R.J. Hodrick (1980), "Forward exchange rates as optimal predictors of future spot rates: An econometric analysis," *Journal of Political Economy*, vol. 88, 829–853.

Johannes, M. (2004), "The statistical and economic role of jumps in continuous-time interest rate models," *Journal of Finance*, vol. 59, 227–260.

Joslin, S., M. Priebsch, and K.J. Singleton (2009), "Risk premiums in dynamic term structure models with unspanned macro risks," *Working Paper, MIT and Stanford*.

Ludvigson, S.C. and S. Ng (2009a), "A factor analysis of bond risk premia," *working paper, New York University*.

Ludvigson, S.C. and S. Ng (2009b), "Macro factors in bond risk premia," *Review of Financial Studies*, vol. 22, 5027–5067.

Newey, Whitney K. and Kenneth D. West (1987), "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, vol. 55, 703–708.

Piazzesi, M., M. Schneider, and S. Tuzel (2007), "Housing, consumption and asset pricing," *Journal of Financial Economics*, vol. 83, 531–569.

Stambaugh, R.F. (1988), "The information in forward rates: Implications for models of the term structure," *Journal of Financial Economics*, vol. 21, 41–70.

Stambaugh, R.F. (1999), "Predictive regressions," *Journal of Financial Economics*, vol. 54, 375–421.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288.

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

Wright, J.H. and H. Zhou (2009), "Bond risk premia and realized jump risk," *Journal of Banking and Finance*, vol. 33, 2333–2345.

Yuan, M. and Y. Lin (2006), "Model selection and estimation in regression with grouped variables," *Journal of Royal Statistical Society Series B (Statistical Methodology)*, vol. 68, 49–67.

Zou, H. (2006), "The adaptive lasso and its oracle properties," *Journal of the American Statistical Association*, vol. 101, 1418–1429.

A Macroeconomic Series Used in the Analysis

Table A.1 lists macroeconomic series used in our empirical analysis. Following Ludvigson and Ng (2009b), we provide the short name of each series, its mnemonic (the series label used in the source database), its transformation code, and a brief data description. The transformation codes are defined as follows. Code 1: no transformation applied to the series; 2: the first difference applied; 3: the second difference; 4: the logarithm; 5: the first difference of logarithm; and 6: the second difference of logarithm. The "$\widehat{G}_t$" column specifies whether the macroeconomic variable has a non-zero coefficient for contemporaneous and/or lagged value in the SAGLasso regression. The value of "0" under $\widehat{G}_t$ corresponds to the contemporaneous variable and the value of 1 through 6 denotes corresponding lagged values.

Table A.1 Data Description

Series No.	Group	Mnemonic	Description	Short Name	tran	$\widehat{G}_t$
1	1	a0m052	Personal income (AR, bil. chain 2000 $)	PI	5	
2	1	A0M051	Personal income less transfer payments (AR, bil. chain 2000 $)	PI less transfers	5	
3	4	A0M224R	Real Consumption (AC) A0m224/gmdc	Consumption	5	
4	4	A0M057	Manufacturing and trade sales (mil. Chain 1996 $)	M & T sales	5	
5	4	A0M059	Sales of retail stores (mil. Chain 2000 $)	Retail sales	5	
6	1	IPS10	INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX	IP: total	5	
7	1	IPS11	INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL	IP: products	5	
8	1	IPS299	INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS	IP: final prod	5	
9	1	IPS12	INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS	IP: cons gds	5	
10	1	IPS13	INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS	IP: cons dble	5	
11	1	IPS18	INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS	iIP:cons nondble	5	
12	1	IPS25	INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT	IP:bus eqpt	5	
13	1	IPS32	INDUSTRIAL PRODUCTION INDEX - MATERIALS	IP: matls	5	
14	1	IPS34	INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS	IP: dble mats	5	
15	1	IPS38	INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS	IP:nondble mats	5	
16	1	IPS43	INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)	IP: mfg	5	
17	1	IPS307	INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES	IP: res util	5	
18	1	IPS306	INDUSTRIAL PRODUCTION INDEX - FUELS	IP: fuels	5	
19	1	PMP	NAPM PRODUCTION INDEX (PERCENT)	NAPM prodn	1	
20	1	A0m082	Capacity Utilization (Mfg)	Cap util	2	
21	2	LHEL	INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)	Help wanted indx	2	
22	2	LHELX	EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF	Help wanted/emp	2	
23	2	LHEM	CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)	Emp CPS total	5	
24	2	LHNAG	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)	Emp CPS nonag	5	
25	2	LHUR	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%,SA)	U: all	2	
26	2	LHU680	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)	U: mean duration	2	
27	2	LHU5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)	U ¡ 5 wks	5	
28	2	LHU14	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)	U 5-14 wks	5	
29	2	LHU15	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)	U 15+ wks	5	
30	2	LHU26	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)	U 15-26 wks	5	
31	2	LHU27	UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS,SA)	U 27+ wks	5	
32	2	A0M005	Average weekly initial claims, unemploy. insurance (thous.)	UI claims	5	
33	2	CES002	EMPLOYEES ON NONFARM PAYROLLS - TOTAL PRIVATE	Emp: total	5	
34	2	CES003	EMPLOYEES ON NONFARM PAYROLLS - GOODS-PRODUCING	Emp: gds prod	5	
35	2	CES006	EMPLOYEES ON NONFARM PAYROLLS - MINING	Emp: mining	5	
36	2	CES011	EMPLOYEES ON NONFARM PAYROLLS - CONSTRUCTION	Emp: const	5	
37	2	CES015	EMPLOYEES ON NONFARM PAYROLLS - MANUFACTURING	Emp: mfg	5	
38	2	CES017	EMPLOYEES ON NONFARM PAYROLLS - DURABLE GOODS	Emp: dble gds	5	
39	2	CES033	EMPLOYEES ON NONFARM PAYROLLS - NONDURABLE GOODS	Emp: nondbles	5	
40	2	CES046	EMPLOYEES ON NONFARM PAYROLLS - SERVICE-PROVIDING	Emp: services	5	
41	2	CES048	EMPLOYEES ON NONFARM PAYROLLS - TRADE, TRANSPORTATION, AND UTILITIES	Emp: TTU	5	
42	2	CES049	EMPLOYEES ON NONFARM PAYROLLS - WHOLESALE TRADE	Emp: wholesale	5	0
43	2	CES053	EMPLOYEES ON NONFARM PAYROLLS - RETAIL TRADE	Emp: retail	5	
44	2	CES088	EMPLOYEES ON NONFARM PAYROLLS - FINANCIAL ACTIVITIES	Emp: FIRE	5	0,1,2,3,4,6
45	2	CES140	EMPLOYEES ON NONFARM PAYROLLS - GOVERNMENT	Emp: Govt	5	4
46	2	CES151	AVERAGE WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NONFAR	Avg hrs	1	
47	2	CES155	AVERAGE WEEKLY HOURS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NONFAR	Overtime: mfg	2	
48	2	aom001	Average weekly hours, mfg. (hours)	Avg hrs: mfg	1	
49	2	PMEMP	NAPM EMPLOYMENT INDEX (PERCENT)	NAPM empl	1	0,3,6
50	3	HSFR	HOUSING STARTS:NONFARM(1947-58);TOTAL FARM & NONFARM(1959-)(THOUS.,SA	HStarts: Total	4	
51	3	HSNE	HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.	HStarts: NE	4	5,6
52	3	HSMW	HOUSING STARTS:MIDWEST(THOUS.U.)S.A.	HStarts: MW	4	2,3,4,5,6
53	3	HSSOU	HOUSING STARTS:SOUTH (THOUS.U.)S.A.	HStarts: South	4	0,2
54	3	HSWST	HOUSING STARTS:WEST (THOUS.U.)S.A.	HStarts: West	4	
55	3	HSBR	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)	BP: total	4	
56	3	HSBNE	HOUSES AUTHORIZED BY BUILD. PERMITS:NORTHEAST(THOU.U.)S.A	BP: NE	4	
57	3	HSBMW	HOUSES AUTHORIZED BY BUILD. PERMITS:MIDWEST(THOU.U.)S.A.	BP: MW	4	
58	3	HSBSOU	HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH(THOU.U.)S.A.	BP: South	4	5
59	3	HSBWST	HOUSES AUTHORIZED BY BUILD. PERMITS:WEST(THOU.U.)S.A.	BP: West	4	4
60	4	PMI	PURCHASING MANAGERS' INDEX (SA)	PMI	1	
61	4	PMNO	NAPM NEW ORDERS INDEX (PERCENT)	NAPM new ordrs	1	
62	4	PMDEL	NAPM VENDOR DELIVERIES INDEX (PERCENT)	NAPM vendor del	1	
63	4	PMNV	NAPM INVENTORIES INDEX (PERCENT)	NAPM Invent	1	
64	4	A0M008	Mfrs' new orders, consumer goods and materials (bil. chain 1982 $)	Orders: cons gds	5	
65	4	A0M007	Mfrs' new orders, durable goods industries (bil. chain 2000 $)	Orders: dble gds	5	
66	4	A0M027	Mfrs' new orders, nondefense capital goods (mil. chain 1982 $)	Orders: cap gds	5	

67	4	A1M092	Mfrs' unfilled orders, durable goods indus. (bil. chain 2000 $)	Unf orders: dble	5	
68	4	A0M070	Manufacturing and trade inventories (bil. chain 2000 $)	M & T invent	5	
69	4	A0M077	Ratio, mfg. and trade inventories to sales (based on chain 2000 $)	M & T invent/sales	2	
70	5	FM1	MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA)	M1	6	
71	5	FM2	MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM TIME DEP(BIL$,	M2	6	
72	5	FM3	MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL$,SA)	M3	6	
73	5	FM2DQ	MONEY SUPPLY - M2 IN 1996 DOLLARS (BCI)	M2 (real)	5	
74	5	FMFBA	MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA)	MB	6	
75	5	FMRRA	DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL$,SA)	Reserves tot	6	
76	5	FMRNBA	DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL$,SA)	Reserves nonbor	6	
77	5	FCLNQ	COMMERCIAL & INDUSTRIAL LOANS OUSTANDING IN 1996 DOLLARS (BCI)	C&I loans	6	
78	5	FCLBMC	WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL$,SAAR)	C&I loans	1	
79	5	CCINRV	CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)	Cons credit-Nonrevolving	6	
80	5	A0M095	Ratio, consumer installment credit to personal income (pct.)	Inst cred/PI	2	
81	8	FSPCOM	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)	S&P 500	5	
82	8	FSPIN	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)	S&P: indust	5	
83	8	FSDXP	S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)	S&P div yield	2	
84	8	FSPXE	S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%,NSA)	S&P PE ratio	5	
85	6	FYFF	INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)	FedFunds	2	6
86	6	CP90	Cmmercial Paper Rate (AC)	Commpaper	2	
87	6	FYGM3	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)	3 mo T-bill	2	
88	6	FYGM6	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)	6 mo T-bill	2	
89	6	FYGT1	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)	1 yr T-bond	2	
90	6	FYGT5	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)	5 yr T-bond	2	
91	6	FYGT10	INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)	10 yr T-bond	2	
92	6	FYAAAC	BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)	Aaabond	2	
93	6	FYBAAC	BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)	Baa bond	2	
94	6	scp90	cp90-fyff	CP-FF spread	1	6
95	6	sfygm3	fygm3-fyff	3 mo-FF spread	1	
96	6	sFYGM6	fygm6-fyff	6 mo-FF spread	1	6
97	6	sFYGT1	fygt1-fyff	1 yr-FF spread	1	2
98	6	sFYGT5	fygt5-fyff	5 yr-FFspread	1	0
99	6	sFYGT10	fygt10-fyff	10yr-FF spread	1	4
100	6	sFYAAAC	fyaaac-fyff	Aaa-FF spread	1	
101	6	sFYBAAC	fybaac-fyff	Baa-FF spread	1	0
102	6	EXRUS	UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)	Ex rate: avg	5	5,6
103	6	EXRSW	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$)	Ex rate: Switz	5	
104	6	EXRJAN	FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$)	Ex rate: Japan	5	
105	6	EXRUK	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)	Ex rate: UK	5	
106	6	EXRCAN	FOREIGN EXCHANGE RATE: CANADA (CANADIAN *PERU.S.*)	EX rate: Canada	5	
107	7	PWFSA	PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)	PPI: fin gds	6	
108	7	PWFCSA	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)	PPI: cons gds	6	
109	7	PWIMSA	PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)	PPI: int matls	6	
110	7	PWCMSA	PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)	PPI: crude matls	6	
111	7	PSCCOM	SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100)	Commod: spot price	6	
112	7	PSM99Q	INDEX OF SENSITIVE MATERIALS PRICES (1990=100)(BCI-99A)	Sens matls price	6	0
113	7	PMCP	NAPM COMMODITY PRICES INDEX (PERCENT)	NAPM com price	1	0,4,5,6
114	7	PUNEW	CPI-U: ALL ITEMS (82-84=100,SA)	CPI-U: all	6	
115	7	PU83	CPI-U: APPAREL & UPKEEP (82-84=100,SA)	CPI-U: apparel	6	
116	7	PU84	CPI-U: TRANSPORTATION (82-84=100,SA)	CPI-U: transp	6	
117	7	PU85	CPI-U: MEDICAL CARE (82-84=100,SA)	CPI-U: medical	6	
118	7	PUC	CPI-U: COMMODITIES (82-84=100,SA)	CPI-U: comm.	6	
119	7	PUCD	CPI-U: DURABLES (82-84=100,SA)	CPI-U: dbles	6	
120	7	PUS	CPI-U: SERVICES (82-84=100,SA)	CPI-U: services	6	
121	7	PUXF	CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)	CPI-U: ex food	6	
122	7	PUXHS	CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)	CPI-U: ex shelter	6	6
123	7	PUXM	CPI-U: ALL ITEMS LESS MEDICAL CARE (82-84=100,SA)	CPI-U: ex med	6	
124	7	GMDC	PCE,IMPL PR DEFL:PCE (1987=100)	PCE defl	6	6
125	7	GMDCD	PCE,IMPL PR DEFL:PCE; DURABLES (1987=100)	PCE defl: dlbes	6	
126	7	GMDCN	PCE,IMPL PR DEFL:PCE; NONDURABLES (1996=100)	PCE defl: nondble	6	
127	7	GMDCS	PCE,IMPL PR DEFL:PCE; SERVICES (1987=100)	PCE defl: services	6	
128	2	CES275	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NO	AHE: goods	6	
129	2	CES277	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NO	AHE: const	6	
130	2	CES278	AVERAGE HOURLY EARNINGS OF PRODUCTION OR NONSUPERVISORY WORKERS ON PRIVATE NO	AHE: mfg	6	
131	4	HHSNTN	U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)	Consumer expect	2	

B The Term Structure Model Used in the Bootstrap Analysis

This appendix describes in detail the dynamic term structure model used to test the general null hypothesis. As in Ang and Piazzesi (2003), we specify the short rate as an affine function of a vector of both macroeconomic and latent factors as follows: $f_t = \left(m_t', \, x_t'\right)'$, where m_t is a 4×1 vector of macroeconomic factors (to be estimated by supervised principle components), and x_t a 4×1 vector of latent factors. Furthermore, following Duffee (2007), we restrict the latent factors from driving the dynamics of short rate r, and their only role is to drive the risk compensation for corresponding macro factors. Namely,

$$r_t = \delta_0 + \delta_1' f_t, \text{ where } \delta_1' = (\delta_m', \, 0_{1\times 4}) \tag{25}$$

Under this specification, the vector f_t fully reflects all available information on the state of the economy at time t; hence, for example, one need not consider lags of f_t.

We assume that the state vector follows a Gaussian VAR(1) process

$$\begin{aligned} f_t &= \mu + \Phi f_{t-1} + \Sigma \epsilon_t \\ &= \begin{bmatrix} \mu_m \\ 0_{4\times 1} \end{bmatrix} + \begin{bmatrix} \Phi_m & 0_{4\times 4} \\ 0_{4\times 4} & \Phi_x \end{bmatrix} \begin{bmatrix} m_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \Sigma_m & 0_{4\times 4} \\ 0_{4\times 4} & \Sigma_x \end{bmatrix} \begin{bmatrix} \epsilon_t^m \\ \epsilon_t^x \end{bmatrix} \end{aligned} \tag{26}$$

where shocks $\epsilon_t \sim N(0,1)$. With the restriction imposed on Φ and Σ, the evolution of the latent term premia factors depends only on the latent factors. Moreover, innovations in m_t, and thus innovations in the short rate, are by construction orthogonal to the latent state vector. In the DTSM proposed by Joslin et al. (2009), the information embodied in macroeconomic variables is not spanned by latent factors as well. However, their latent factors underlie the variation in short rate and are still correlated (but not perfectly) with macro factors. Intuitively, their model allows for more flexibility and therefore should fit better the comovement of economic indicators and bond yields. To generate data conforming with the general null hypothesis, we need macro factors that do not covary with latent ones, as can be seen later in the market price of risk.

It follows from the no-arbitrage restriction of Duffie and Kan (1996) that the period-t price of any asset with valuation P_{t+1} at the end of period $t+1$ satisfies

$$P_t = E_t^Q[\exp(-r_t) P_{t+1}] \; = \; \frac{1}{\xi_t} E_t^P [\xi_{t+1} \exp(-r_t) P_{t+1}] \tag{27}$$

where Q and P denote the risk-neutral and the physical measures, respectively, and ξ_t is the Radon-Nikodym derivative that follows the log-normal process

$$\xi_{t+1} = \xi_t \, \exp(-\lambda_t' \lambda_t / 2 - \lambda_t' \epsilon_{t+1}), \tag{28}$$

where the market prices of risk follow the essentially-affine specification (Duffee, 2002):

$$\Sigma \lambda_t = \lambda_0 + \lambda_1 f_t. \tag{29}$$

If we use lowercase letters to indicate logs, then we have

$$p_t = E_t^P[p_{t+1} + m_{t+1}] \tag{30}$$

where the pricing kernel m_{t+1} is given as follows

$$m_{t+1} = -r_t + \log\frac{\xi_{t+1}}{\xi_t} = -\delta_0 - \delta_1' f_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}. \tag{31}$$

If the matrix λ_1 contains 8×8 free parameters, our model exactly follows the details of Ang-Piazzesi macro-finance model: the price of risk relates shocks in the underlying state variables (macro and latent factors) to ξ and investors require both compensations to face uncertainty in macro and latent factors. Instead we parameterize λ_0 and λ_1 in the following way

$$\lambda_0 = \begin{bmatrix}\lambda_{0m}\\ 0_{4\times1}\end{bmatrix} \text{ and } \lambda_1 = \begin{bmatrix}\lambda_{1m} & I_{4\times4}\\ 0_{4\times4} & 0_{4\times4}\end{bmatrix} \tag{32}$$

such that the compensation required by investors depends on latent factors that evolve independently of the macro factors but affect the risk compensation for macro factors through 1-to-1 mappings.[12] Note that the specification of risk premia in Eq. (32) distinguishes between macro and observable influences on term premia and thus embodies the general null hypothesis. To see this, we write the functional form of λ_t as

$$\Sigma\lambda_t = \begin{bmatrix}\lambda_{0m} + \lambda_{1m}m_t + x_t\\ 0_{4\times1}\end{bmatrix}. \tag{33}$$

The risk compensation depends on the macro factors only through λ_{1m}. Therefore, with the restriction $\lambda_{1m} = 0$, shocks to the macroeconomic factors have no impact on expected excess returns at all leads and lags, and thus the model corresponds to the general null that excess bond returns are stochastic and persistent, but independent of the macroeconomy. Otherwise, we cannot recover latent factors solely from the market price of risk.

Finally, bond prices are exponential affine functions of the state variables

$$p_t^{(n)} = A_n + B_n' f_t \tag{34}$$

where A_n and B_n can be computed recursively as the following:

$$A_n = A_{n-1} + B_{n-1}'(\mu - \lambda_0) + \frac{1}{2}B_{n-1}'\Sigma\Sigma' B_{n-1}' - \delta_0 \tag{35}$$

$$B_n = B_{n-1}'(\Phi - \lambda_1) - \delta_1' \tag{36}$$

with the initial values $A_1 = \delta_0$ and $B_1 = -\delta_1'$.

We estimate the term structure model using the Kalman filter, following Duffee and Stanton (2004). The parametrization of expected excess returns for the model under general null requires 61 free parameters, including five measurement error parameters. Some of the initial values (those for Φ_m and Σ_m) are set to those from OLS estimation of the VAR(1), and analytic derivatives are used in the derivative-based optimization routine.

[12] Latent factor i affects only the risk compensation for macro factor i.

C Supervised Adaptive Group Lasso Method

The Supervised Adaptive Group Lasso (SAGLasso) method for predicting excess returns that we propose consists of the following steps:

1. For cluster $h \in \mathcal{H}$, compute $\hat{\beta}^h$–the cluster-wise Adaptive Lasso estimate of β^h. Namely,

$$\hat{\beta}^h = \underset{\beta^h}{\text{argmin}} \left\{ ||\mathbf{arx} - \mathbf{X}_h \beta^h||^2 + \sum_j \lambda_h * \hat{w}_{hj} |\beta_j^h| \right\}, \tag{37}$$

   where $\mathbf{arx}$ is a vector of average excess bond returns across maturity and $\hat{w}_{hj}$ the j-th component of $\hat{\mathbf{w}}_h$, the vector of the (adaptive) weights. Zou (2006) recommends using $\hat{\beta}^{OLS}$ to construct $\hat{\mathbf{w}}_h$. As collinearity is a concern in our case, we set $\hat{\mathbf{w}}_h = 1/|\hat{\beta}_h^{RID}|^{\gamma_h}$, where $\hat{\beta}_h^{RID}$ is the best ridge regression fit of $\mathbf{arx}$ on $\mathbf{X}_h$. That is, for cluster h we only use macroeconomic variables within that cluster to construct predictive models. The optimal pairs of (γ_h, λ_h) are determined using two-dimensional cross-validations. It is worth noting that tuning parameters λ_h are selected for each cluster separately in order to have different degrees of regularization for different clusters. This flexibility allows us to uncover subtle structures that otherwise will be missed when applying the (adaptive) lasso method to all the series/clusters at the same time.

   Notice that for each cluster $h \in \mathcal{H}$, the adaptive lasso $\hat{\beta}^h$ has only a small number of nonzero components. Let $\tilde{\beta}^h = \hat{\beta}^h \setminus \mathbf{0}$, the vector of nonzero estimated components of $\hat{\beta}^h$ given by the cluster-wise model (37), and denote the corresponding part of $\mathbf{X}_h$ by $\tilde{X}_h$. In our case, a typical cluster size $(\dim(\mathbf{X}_h))$ of 80 variables may reduce to a $\dim(\tilde{X}_h)$ of $8 \sim 10$. Namely, the number of macroeconomic measures selected in Step 1 is significantly smaller than the original number to begin with.

2. Construct the joint predictive model under the Group Lasso constraint as the following:

$$\hat{\beta} = \underset{\beta}{\text{argmin}} \left\{ ||\mathbf{arx} - \tilde{\mathbf{X}}\beta||^2 + \lambda \sum_{h \in \mathcal{H}} w_h ||\beta_h|| \right\}, \tag{38}$$

   where $\tilde{X}$ is formed by concatenating the design matrices $\tilde{X}_h$. λ is also chosen by cross validation. With $\lambda \to \infty$, estimates of some components of $\tilde{\beta}_h$s can be exactly zero.

Table 1: Regressions of Annual Excess Bond Returns on LN Static Factors

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\hat{f}_{it}$s, the macroeconomic factors estimated by the method of unsupervised factor analysis, and their quadratic or cubic functions. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The column labeled "Joint Test" reports Wald tests of the hypothesis that all coefficients equal zero. Asymptotic p-values, based on a $\chi^2(8)$ distribution, are in brackets. The sample spans the period January 1964 to December 2007.

maturity (yr)	$\hat{f}_{1t}$	$\hat{f}_{2t}$	$\hat{f}_{3t}$	$\hat{f}_{4t}$	$\hat{f}_{5t}$	$\hat{f}_{6t}$	$\hat{f}_{7t}$	$\hat{f}_{8t}$	$\hat{f}^3_{1t}$	$\bar{R}^2$	Joint Test	P-val
2	0.951	0.319	-0.044	-0.402	-0.037	-0.274	-0.146	0.303	-0.058	0.296		
HH	(5.505)	(2.463)	(-1.159)	(-2.970)	(-0.403)	(-1.721)	(-1.419)	(3.697)	(-3.346)		97.979	[0.000]
NW	(5.779)	(2.605)	(-1.165)	(-3.174)	(-0.419)	(-1.891)	(-1.510)	(3.570)	(-3.254)		94.722	[0.000]
3	1.621	0.689	0.001	-0.601	-0.146	-0.577	-0.349	0.556	-0.107	0.281		
HH	(5.345)	(2.686)	(0.011)	(-2.393)	(-0.916)	(-2.028)	(-1.905)	(3.957)	(-3.924)		110.328	[0.000]
NW	(5.533)	(2.899)	(0.011)	(-2.552)	(-0.927)	(-2.234)	(-2.011)	(3.789)	(-3.803)		100.170	[0.000]
4	2.086	1.048	0.033	-0.697	-0.288	-0.872	-0.552	0.828	-0.148	0.275		
HH	(5.215)	(2.741)	(0.297)	(-1.865)	(-1.305)	(-2.218)	(-2.139)	(4.192)	(-4.195)		194.215	[0.000]
NW	(5.334)	(3.003)	(0.319)	(-2.017)	(-1.344)	(-2.436)	(-2.252)	(3.975)	(-4.073)		108.917	[0.000]
5	2.315	1.328	0.074	-0.831	-0.474	-1.101	-0.716	0.999	-0.169	0.264		
HH	(4.702)	(2.826)	(0.518)	(-1.784)	(-1.760)	(-2.268)	(-2.274)	(4.239)	(-4.144)		352.734	[0.000]
NW	(4.824)	(3.090)	(0.570)	(-1.927)	(-1.819)	(-2.484)	(-2.383)	(4.007)	(-4.036)		116.916	[0.000]

Table 2: Regressions of Annual Excess Bond Returns on Single Return-Forecasting Factor

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\widehat{CP}_t$, the Cochrane-Piazzesi (2005) predictor (a linear combination of forward rates), and/or $\widehat{LN}_t$, the Lugvigson-Ng (2009b) factor (a linear combination of static factors estimated as the fitted values from OLS). The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The column labeled "Joint Test" reports Wald tests of the hypothesis that all coefficients equal zero. Asymptotic p-values, based on a $\chi^2(1)$ distribution, are in brackets. Ljung-Box Q statistic is used to test autocorrelation in the error term in the multivariate regressions. The sample spans the period January 1964 to December 2007.

Panel A: Univariate predictive regressions

maturity (yr)	$\widehat{CP}_t$	R^2	Joint Test	P-val	$\widehat{LN}_t$	R^2	Joint Test	P-val
2	0.453	0.256			0.482	0.282		
HH	(5.851)		34.237	[0.000]	(6.833)		46.686	[0.000]
NW	(6.402)		40.983	[0.000]	(7.349)		54.003	[0.000]
3	0.854	0.272			0.877	0.279		
HH	(5.439)		29.588	[0.000]	(6.805)		46.311	[0.000]
NW	(6.011)		36.134	[0.000]	(7.295)		53.211	[0.000]
4	1.242	0.300			1.204	0.275		
HH	(5.419)		29.367	[0.000]	(6.949)		48.284	[0.000]
NW	(6.043)		36.522	[0.000]	(7.295)		53.211	[0.000]
5	1.451	0.273			1.437	0.261		
HH	(5.065)		25.655	[0.000]	(7.015)		49.203	[0.000]
NW	(5.638)		31.787	[0.000]	(7.426)		55.143	[0.000]

Panel B: Multivariate predictive regressions

maturity (yr)	$\widehat{CP}_t$	$\widehat{LN}_t$	R^2	Joint Test	P-val
2	0.359	0.317	0.390		
HH	(4.277)	(3.080)		101.273	[0.000]
NW	(4.697)	(3.474)		110.466	[0.000]
3	0.639	0.613	0.399		
HH	(4.393)	(3.137)		91.274	[0.000]
NW	(4.830)	(3.548)		99.461	[0.000]
4	0.847	0.922	0.416		
HH	(4.335)	(3.226)		100.256	[0.000]
NW	(4.776)	(3.669)		108.752	[0.000]
5	1.025	1.065	0.387		
HH	(4.367)	(2.959)		102.795	[0.000]
NW	(4.792)	(3.361)		107.259	[0.000]

Panel C: Testing for Serial Correlation in the multivariate regression models

maturity (yr)	2	3	4	5
Q	348.89	353.30	362.57	332.23
P-val	0.000	0.000	0.000	0.000

Table 3: SAGLasso Regressions of Annual Excess Bond Returns on Macroeconomic Variables and Lagged Returns.

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\widehat{G}_t$, the single predictor estimated by SAGLasso, and lagged bond excess returns. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "Q" reports Ljung-Box test statistic for first order autocorrelation in the error term in the SAGLasso regressions. The row labeled "LM" reports Breusch-Godfrey LM test statistic for first order autocorrelated disturbance. The sample spans the period January 1964 to December 2007. The sample spans the period January 1964 to December 2007.

Panel A: Results from the ARDL model estimated from SAGLasso

maturity (yr)	$\widehat{G}_t$	$rx_t^{(n)}$	$rx_{t-1}^{(n)}$	$rx_{t-2}^{(n)}$	$rx_{t-3}^{(n)}$	$rx_{t-4}^{(n)}$	$rx_{t-5}^{(n)}$	$\bar{R}^2$
2	0.1898	0.9890	-0.1661	0.0866	-0.1205	0.1780	-0.1255	0.880
HH	(3.4912)	(16.1747)	(-1.5009)	(0.9021)	(-1.1587)	(1.7512)	(-2.2185)	
3	0.3373	1.0254	-0.2133	0.1042	-0.0892	0.0905	-0.0699	0.884
HH	(3.4641)	(24.1105)	(-2.4438)	(1.1947)	(-0.8803)	(0.8919)	(-1.3797)	
4	0.4698	1.0045	-0.1773	0.0933	-0.1000	0.1247	-0.0994	0.883
HH	(4.5464)	(24.7101)	(-2.3040)	(1.2554)	(-1.2874)	(1.3258)	(-1.5297)	
5	0.6338	0.9369	-0.1147	0.1095	-0.1246	0.1054	-0.0791	0.866
HH	(4.1879)	(20.6022)	(-1.6537)	(1.3886)	(-1.3624)	(0.9904)	(-1.1378)	

Panel B: Results from a two-regressor model

maturity (yr)	$\widehat{G}_t$	$rx_t^{(n)}$	$\bar{R}^2$
2	0.180	0.861	0.876
HH	(2.709)	(23.487)	
3	0.326	0.867	0.879
HH	(2.756)	(28.548)	
4	0.457	0.867	0.879
HH	(3.531)	(33.819)	
5	0.617	0.853	0.863
HH	(3.445)	(30.140)	

Panel C: Testing for serial correlation in the SAGLasso regression model.

maturity (yr)	2	3	4	5
Q test	0.508	0.521	0.556	0.585
P-val	0.476	0.471	0.456	0.446
LM test	1.347	2.201	1.860	2.247
P-val	0.247	0.138	0.178	0.134

Table 4: Regressions of Annual Excess Bond Returns on Single Predictive Factor

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\widehat{G}_t$, the single predictor estimated by SAGLasso, and/or $\widehat{LN}_t$, the Lugvigson-Ng (2009b) factor (a linear combination of static factors estimated as the fitted values from OLS). The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The column labeled "Joint Test" reports Wald tests of the hypothesis that all coefficients equal zero. Asymptotic p-values, based on a $\chi^2(1)$ distribution, are in brackets. The sample spans the period January 1964 to December 2007.

Panel A: Univariate predictive regressions

maturity (yr)	$\widehat{G}_t$	R^2	Joint Test	P-val
2	1.064	0.437		
HH	(10.797)		116.573	[0.000]
NW	(11.519)		132.680	[0.000]
3	1.897	0.414		
HH	(9.566)		91.513	[0.000]
NW	(10.210)		104.240	[0.000]
4	2.591	0.404		
HH	(9.070)		82.261	[0.000]
NW	(9.665)		93.405	[0.000]
5	3.122	0.391		
HH	(8.909)		79.362	[0.000]
NW	(9.441)		89.138	[0.000]

Panel B: Multivariate predictive regression of excess returns on $\widehat{G}_t$ and $\widehat{LN}_t$

maturity (yr)	$\widehat{G}_t$	$\widehat{LN}_t$	R^2	Joint Test	P-val
2	0.903	0.132	0.448		
HH	(4.544)	(1.230)		153.297	[0.000]
NW	(5.111)	(1.340)		154.803	[0.000]
3	1.569	0.268	0.428		
HH	(4.392)	(1.405)		107.831	[0.000]
NW	(4.918)	(1.535)		114.888	[0.000]
4	2.129	0.378	0.418		
HH	(4.057)	(1.356)		100.149	[0.000]
NW	(4.535)	(1.491)		107.005	[0.000]
5	2.596	0.431	0.403		
HH	(4.089)	(1.291)		96.097	[0.000]
NW	(4.564)	(1.422)		101.814	[0.000]

Table 5: Regressions of Annual Excess Bond Returns on Group Factors

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\hat{g}_{it}$s, the four macroeconomic group factors estimated by the Group Lasso method. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The column labeled "Joint Test" reports Wald tests of the hypothesis that all coefficients equal zero. Asymptotic p-values, based on a $\chi^2(4)$ distribution, are in brackets. For brevity, panel B does not report test statistics based on the "HH" correction. The sample spans the period January 1964 to December 2007.

Panel A: Univariate predictive regressions on each $\hat{g}_{it}$

maturity (yr)	$\hat{g}_{1t}$	$\hat{g}_{2t}$	$\hat{g}_{3t}$	$\hat{g}_{4t}$
2	0.493	0.569	0.396	0.489
HH	(5.897)	(3.966)	(2.702)	(2.741)
NW	(6.456)	(4.376)	(2.975)	(3.031)
R^2	0.255	0.236	0.075	0.128
3	0.886	0.914	0.794	0.885
HH	(5.808)	(3.302)	(2.925)	(2.562)
NW	(6.320)	(3.667)	(3.347)	(2.843)
R^2	0.246	0.181	0.009	0.126
4	1.190	1.166	1.219	1.211
HH	(5.808)	(2.953)	(3.240)	(2.484)
NW	(6.269)	(3.272)	(3.603)	(2.759)
R^2	0.232	0.154	0.111	0.123
5	1.430	1.351	1.591	1.415
HH	(5.931)	(2.836)	(3.589)	(2.345)
NW	(6.349)	(3.129)	(3.980)	(2.599)
R^2	0.223	0.138	0.126	0.112

Panel B: Multivariate predictive regression of excess returns on group factors $\{\hat{g}_{it}\}$

	$rx^{(2)}_{t+1}$	$rx^{(3)}_{t+1}$	$rx^{(4)}_{t+1}$	$rx^{(5)}_{t+1}$	$rx^{(2)}_{t+1}$	$rx^{(3)}_{t+1}$	$rx^{(4)}_{t+1}$	$rx^{(5)}_{t+1}$	$rx^{(2)}_{t+1}$	$rx^{(3)}_{t+1}$	$rx^{(4)}_{t+1}$	$rx^{(5)}_{t+1}$
$\hat{g}_{1t}$	0.297	0.561	0.735	0.880					0.270	0.525	0.710	0.885
NW	(2.545)	(2.560)	(2.341)	(2.303)					(2.256)	(2.323)	(2.217)	(2.245)
$\hat{g}_{2t}$	0.358	0.496	0.591	0.672					0.362	0.561	0.740	0.895
NW	(2.171)	(1.600)	(1.333)	(1.248)					(2.115)	(1.698)	(1.542)	(1.522)
$\hat{g}_{3t}$	0.127	0.310	0.590	0.881					0.180	0.240	0.455	0.697
NW	(0.986)	(1.342)	(1.852)	(2.309)					(0.801)	(0.614)	(0.867)	(1.113)
$\hat{g}_{4t}$	0.075	0.168	0.231	0.192					0.192	0.340	0.486	0.504
NW	(0.470)	(0.544)	(0.522)	(0.349)					(1.049)	(0.975)	(0.995)	(0.832)
$\hat{f}_{1t}$					0.576	0.928	1.128	1.224	-0.023	-0.118	-0.284	-0.491
NW					(2.850)	(2.476)	(2.283)	(2.088)	(-0.103)	(-0.277)	(-0.501)	(-0.718)
$\hat{f}_{2t}$					0.310	0.674	1.027	1.304	-0.296	-0.333	-0.472	-0.599
NW					(2.435)	(2.720)	(2.836)	(2.935)	(-1.230)	(-0.811)	(-0.848)	(-0.892)
$\hat{f}_{3t}$					-0.042	0.004	0.038	0.080	-0.071	-0.041	-0.048	-0.040
NW					(-1.110)	(0.063)	(0.368)	(0.614)	(-1.587)	(-0.496)	(-0.382)	(-0.263)
$\hat{f}_{4t}$					-0.495	-0.773	-0.936	-1.102	-0.070	-0.057	0.041	0.076
NW					(-3.929)	(-3.326)	(-2.842)	(-2.697)	(-0.413)	(-0.184)	(0.093)	(0.139)
$\hat{f}_{5t}$					-0.089	-0.243	-0.422	-0.627	-0.070	-0.206	-0.309	-0.429
NW					(-0.756)	(-1.132)	(-1.500)	(-1.876)	(-0.546)	(-0.871)	(-1.005)	(-1.144)
$\hat{f}_{6t}$					-0.268	-0.566	-0.856	-1.083	-0.195	-0.457	-0.677	-0.827
NW					(-1.816)	(-2.134)	(-2.331)	(-2.394)	(-1.225)	(-1.592)	(-1.697)	(-1.694)
$\hat{f}_{7t}$					-0.116	-0.294	-0.476	-0.629	-0.124	-0.302	-0.466	-0.587
NW					(-1.312)	(-1.840)	(-2.090)	(-2.234)	(-1.243)	(-1.669)	(-1.835)	(-1.904)
$\hat{f}_{8t}$					0.298	0.547	0.815	0.984	0.170	0.321	0.476	0.530
NW					(3.689)	(3.844)	(4.041)	(4.126)	(2.345)	(2.617)	(2.887)	(2.757)
R^2	0.346	0.313	0.298	0.291	0.246	0.230	0.224	0.220	0.390	0.357	0.348	0.342

Table 6: Regressions of Annual Excess Bond Returns on the Output Gap and Other Predictors.

The dependent variable is the return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond. The independent variables include the output gap *gap*–measured as the deviations of the log of industrial production index from a quadratic and linear trend as in Cooper and Priestley (2009), $\widehat{CP}_t^{\perp}$–the Cochrane-Piazzesi (2005) factor orthogonalized relative to *gap*, $\widehat{G}_t^{\perp}$–the SAGLasso orthogonalized relative to *gap*, and $\hat{g}_{1t}$–the first macroeconomic group factor identified using the Group Lasso. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. Results based on the Hansen-Hodrick GMM correction for overlap are similar and not reported here for brevity. The sample spans the period January 1964 to December 2007.

maturity	gap_{t-1}	$\widehat{CP}_t^{\perp}$	$\widehat{G}_t^{\perp}$	$\hat{g}_{1t}$	$\bar{R}^2$
2	-8.798				0.094
NW	(-2.480)				
3	-13.924				0.070
NW	(-2.142)				
4	-19.676				0.073
NW	(-2.211)				
5	-23.702				0.071
NW	(-2.224)				
2	-8.798	0.433			0.254
NW	(-2.812)	(5.241)			
3	-13.924	0.884			0.270
NW	(-2.482)	(5.630)			
4	-19.676	1.299			0.299
NW	(-2.594)	(5.832)			
5	-23.702	1.501			0.271
NW	(-2.606)	(5.409)			
2	-8.798		1.213		0.451
NW	(-3.681)		(11.356)		
3	-13.924		2.258		0.440
NW	(-3.159)		(10.539)		
4	-19.676		3.049		0.425
NW	(-3.127)		(10.309)		
5	-23.702		3.674		0.412
NW	(-3.094)		(10.193)		
2	-4.623			3.887	0.268
NW	(-1.309)			(3.741)	
3	-6.238			7.155	0.246
NW	(-0.959)			(3.695)	
4	-9.564			9.414	0.232
NW	(-1.048)			(3.519)	
5	-11.760			11.117	0.219
NW	(-1.062)			(3.442)	

Table 7: Comparison between the SAGLasso factor and CP's Return Forecasting Factor

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\widehat{CP}_t$, the Cochrane-Piazzesi (2005) return-forecasting factor, alone with $\widehat{G}_t$, the SAGLasso factor. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The column labeled "Joint Test" reports Wald tests of the hypothesis that all coefficients equal zero. Asymptotic p-values, based on a $\chi^2(1)$ distribution, are in brackets. The sample spans the period January 1964 to December 2007.

Panel A: Predictive regressions of excess returns on $\widehat{G}_t$ and $\widehat{CP}_t$.

maturity	$\widehat{CP}_t$	$\widehat{G}_t$	R^2	Joint Test	P-val
2	0.227	0.616	0.395		
HH	(2.580)	(3.983)		58.296	[0.000]
NW	(2.879)	(4.458)		70.817	[0.000]
3	0.491	0.994	0.380		
HH	(2.700)	(3.398)		48.675	[0.000]
NW	(3.030)	(3.820)		59.883	[0.000]
4	0.784	1.251	0.389		
HH	(2.896)	(3.148)		47.028	[0.000]
NW	(3.271)	(3.529)		58.438	[0.000]
5	0.886	1.542	0.364		
HH	(2.639)	(3.344)		43.698	[0.000]
NW	(2.974)	(3.719)		53.665	[0.000]

Panel B: Predictive regression including $\widehat{CP}_t$ into the ARDL model estimated from SAGLasoo.

maturity	$\widehat{G}_t$	$\widehat{CP}_t$	$rx_t^{(n)}$	$rx_{t-1}^{(n)}$	$rx_{t-2}^{(n)}$	$rx_{t-3}^{(n)}$	$rx_{t-4}^{(n)}$	$rx_{t-5}^{(n)}$	$\bar{R}^2$	P-val
2	0.1743	0.0253	0.9865	-0.1673	0.0865	-0.1214	0.1798	-0.1328	0.881	
HH	3.0392	1.2539	16.0496	-1.4767	0.8955	-1.1459	1.7419	-2.4486		[0.000]
3	0.2867	0.0781	1.0167	-0.2109	0.1042	-0.0966	0.0995	-0.0813	0.885	
HH	2.7547	1.8312	23.4257	-2.3027	1.1565	-0.9260	0.9523	-1.6752		[0.000]
4	0.3557	0.1766	0.9913	-0.1772	0.0909	-0.1025	0.1291	-0.1130	0.887	
HH	3.0553	3.0471	22.6590	-2.1409	1.1641	-1.2807	1.3166	-1.7877		[0.000]
5	0.5577	0.1162	0.9297	-0.1157	0.1096	-0.1259	0.1057	-0.0847	0.867	
HH	3.3432	1.4921	20.2872	-1.6142	1.3235	-1.3486	0.9792	-1.2778		[0.000]

Table 8: Comparison between the SAGLasso factor and Duffee's Yield Factors

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\widetilde{H}_t^i, i = 1, \ldots, 5$, Duffee (2008)'s five latent yield factors (the 5th factor being the hidden factor) estimated using Kalman filtering, alone with $\widehat{G}_t$, the single predictor estimated by SAGLasso. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The column labeled "Joint Test" reports Wald tests of the hypothesis that all coefficients equal zero. Asymptotic p-values, based on a $\chi^2(1)$ distribution, are in brackets. The sample spans the period January 1964 to December 2007.

Panel A: Predictive regressions of excess returns on filtered state variable estimated from the yield curve

maturity	$\widetilde{H}_t^1$	$\widetilde{H}_t^2$	$\widetilde{H}_t^3$	$\widetilde{H}_t^4$	$\widetilde{H}_t^5$	R^2	Joint Test	P-val
2	0.913	8.884	-11.174	50.994	204.850	0.312		
HH	(2.228)	(3.764)	(-1.309)	(2.570)	(3.689)		78.077	[0.000]
NW	(2.488)	(4.099)	(-1.461)	(2.627)	(3.652)		70.183	[0.000]
3	1.187	16.800	-14.034	80.784	441.654	0.325		
HH	(1.608)	(3.655)	(-0.965)	(2.329)	(4.699)		68.985	[0.000]
NW	(1.802)	(4.052)	(-1.075)	(2.368)	(4.543)		62.991	[0.000]
4	1.385	25.432	-17.551	99.892	613.020	0.344		
HH	(1.346)	(3.842)	(-0.961)	(2.151)	(5.017)		63.075	[0.000]
NW	(1.513)	(4.297)	(-1.059)	(2.193)	(4.797)		60.085	[0.000]
5	1.453	32.377	-16.186	103.049	680.247	0.324		
HH	(1.145)	(3.961)	(-0.750)	(1.745)	(4.582)		48.579	[0.000]
NW	(1.287)	(4.446)	(-0.821)	(1.794)	(4.354)		50.095	[0.000]

Panel B: Predictive regression of excess returns on both $\hat{G}_t$ and $\widetilde{H}_t^i$s.

maturity	$\widehat{G}_t$	$\widetilde{H}_t^1$	$\widetilde{H}_t^2$	$\widetilde{H}_t^3$	$\widetilde{H}_t^4$	$\widetilde{H}_t^5$	$\bar{R}^2$	Joint Test	P-val
2	0.888	0.619	1.726	-3.146	27.656	134.905	0.501		
HH	(6.999)	(1.606)	(0.677)	(-0.755)	(1.494)	(2.480)		181.579	[0.000]
NW	(7.633)	(1.785)	(0.747)	(-0.717)	(1.567)	(2.608)		180.158	[0.000]
3	1.536	0.678	4.417	-0.147	40.414	320.660	0.495		
HH	(6.421)	(0.973)	(0.881)	(-0.019)	(1.274)	(3.457)		164.723	[0.000]
NW	(6.997)	(1.080)	(0.981)	(-0.018)	(1.312)	(3.560)		161.920	[0.000]
4	1.981	0.729	9.469	0.352	47.850	457.045	0.491		
HH	(5.771)	(0.739)	(1.291)	(0.035)	(1.126)	(3.611)		154.765	[0.000]
NW	(6.234)	(0.823)	(1.444)	(0.031)	(1.139)	(3.717)		152.962	[0.000]
5	2.381	0.664	13.190	5.334	40.493	492.761	0.465		
HH	(5.730)	(0.555)	(1.441)	(0.420)	(0.757)	(3.197)		142.390	[0.000]
NW	(6.152)	(0.617)	(1.615)	(0.378)	(0.766)	(3.282)		140.217	[0.000]

Table 9: Baseline Regressions in Wright and Zhou (2008)

The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on different sets of lagged predictor variables consisting of realized volatility RV_t, jump intensity JI_t, jump mean JM_t, jump volatility JV_t and forward rates. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The column labeled "Joint Test" reports Wald tests of the hypothesis that all coefficients equal zero. Asymptotic p-values, based on a $\chi^2(1)$ distribution, are in brackets. The sample spans the period August 1984 to December 2007.

Panel A: Univariate regressions of excess returns on realized volatility and jump measures.

maturity	RV_t	JI_t	JM_t	JV_t	R^2
2	1.135 (0.671) [0.751]				0.010
		0.445 (0.045) [0.051]			0.000
			-6.794 (-4.282) [-4.359]		0.143
				-6.984 (-1.946) [-2.201]	0.106
3	2.054 (0.649) [0.729]				0.009
		1.236 (0.074) [0.083]			0.000
			-13.662 (-5.611) [-5.225]		0.160
				-13.040 (-1.916) [-2.145]	0.102
4	3.148 (0.693) [0.779]				0.010
		1.703 (0.075) [0.084]			0.000
			-18.206 (-5.986) [-5.233]		0.141
				-16.529 (-1.732) [-1.927]	0.081
5	3.791 (0.681) [0.765]				0.010
		-0.957 (-0.037) [-0.041]			0.000
			-22.090 (-6.093) [-5.136]		0.137
				-17.990 (-1.539) [-1.702]	0.064

Panel B: Predictive regression of excess returns on realized jump means and the term structure of forward rates.

maturity	JM_t	$F_t^{(1)}$	$F_t^{(3)}$	$F_t^{(5)}$	$\bar{R}^2$	Joint Test	P-val
2	-9.870	-0.288	0.974	-0.199	0.587		
HH	(-5.172)	(-1.047)	(1.916)	(-0.484)		323.013	[0.000]
NW	(-5.420)	(-1.184)	(2.120)	(-0.538)		251.541	[0.000]
3	-18.807	-0.750	2.191	-0.576	0.591		
HH	(-5.560)	(-1.532)	(2.377)	(-0.771)		279.582	[0.000]
NW	(-5.745)	(-1.723)	(2.611)	(-0.853)		211.971	[0.000]
4	-25.017	-1.289	3.229	-0.713	0.588		
HH	(-5.308)	(-1.987)	(2.612)	(-0.705)		207.637	[0.000]
NW	(-5.508)	(-2.216)	(2.830)	(-0.772)		169.130	[0.000]
5	-30.206	-1.572	3.476	-0.395	0.567		
HH	(-5.078)	(-2.032)	(2.327)	(-0.325)		162.591	[0.000]
NW	(-5.280)	(-2.253)	(2.511)	(-0.354)		136.028	[0.000]

Table 10: Realized Jumps and Macroeconomic News

This table reports the jump dates and jump size of the 20 biggest jumps during our sample period, identified by model-free realized jump tests, and associated events on jump dates. The sample period is July 1982 through December 2007.

Jump Dates	Jump Size	Events
07-Oct-1982	1.0859	FOMC Target Announcement*
12-Oct-1982	-0.9370	Advanced Retail Sales Announcement
12-May-1989	1.2003	Advanced Retail Sales Announcement
09-Jan-1991	-1.3131	the Outbreak of the Gulf War
13-Nov-1991	-0.8426	CPI Announcement
04-Sep-1992	0.9330	Unemployment Rate Announcement
02-Jun-1995	1.1363	Unemployment Rate Announcement
08-Mar-1996	-1.3810	the Third Taiwan Strait Crisis
07-Jun-1996	-1.7188	Unemployment Rate Announcement
05-Jul-1996	-1.4057	Unemployment Rate Announcement
30-Aug-1996	-0.7399	GDP Announcement
29-Apr-1997	1.2346	Consumer Confidence Announcement
05-Mar-1999	1.0107	Unemployment Rate Announcement
14-May-1999	-0.8258	Industrial Production Announcement
06-Dec-2002	0.9711	Unemployment Rate Announcement
03-Oct-2003	-0.7913	Unemployment Rate Announcement
09-Jan-2004	1.2738	Unemployment Rate Announcement
28-Jan-2004	-1.0643	FOMC Target Announcement**
02-Jul-2004	1.2727	Unemployment Rate Announcement
06-Aug-2004	1.6548	Unemployment Rate Announcement

*Federal Reserve announced a reduction in the discount rate from 10 percent to 9-1/2 percent.

**In contrast with pre-announcement forecasts by economists, Federal Reserve decided to keep its target for the federal funds rate at 1 percent.

Table 11: Subsample Analysis

This table examines the period after the momentous period 1979-1982. The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on $\widetilde{H}_t^i, i = 1, \ldots, 5$, Duffee (2008)'s five latent yield factors (the 5th factor being the hidden factor) estimated using Kalman filtering, alone with $\widehat{G}_t$, the single predictor estimated by SAGLasso. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The sample spans the period August 1984 to December 2007.

Panel A: Predictive regression of excess returns on SAGLasso factor or filtered yield factors alone.

maturity	$\widehat{G}_t$	$\bar{R}^2$	$\widetilde{H}_t^1$	$\widetilde{H}_t^2$	$\widetilde{H}_t^3$	$\widetilde{H}_t^4$	$\widetilde{H}_t^5$	$\bar{R}^2$
2	1.109	0.308	-8671.185	8606.115	3305.392	649.437	183.304	0.434
HH	(5.438)		(-1.093)	(3.702)	(2.882)	(1.991)	(3.622)	
NW	(5.791)		(-1.118)	(3.033)	(3.081)	(2.204)	(4.049)	
3	2.088	0.302	-21232.067	14990.519	5222.633	1441.950	296.596	0.404
HH	(5.688)		(-1.347)	(2.923)	(2.322)	(2.353)	(3.112)	
NW	(5.785)		(-1.381)	(2.546)	(2.493)	(2.600)	(3.491)	
4	3.019	0.313	-29410.600	20208.495	6587.152	2482.737	394.956	0.421
HH	(6.107)		(-1.370)	(2.676)	(2.091)	(3.052)	(3.074)	
NW	(5.963)		(-1.387)	(2.439)	(2.245)	(3.355)	(3.440)	
5	3.585	0.292	-27898.433	24005.974	7001.084	3346.762	445.140	0.394
HH	(5.652)		(-1.009)	(2.525)	(1.797)	(3.356)	(2.909)	
NW	(5.481)		(-1.025)	(2.350)	(1.925)	(3.672)	(3.246)	

Panel B: Predictive regression of excess returns on SAGLasso factor and filtered yield factors.

maturity	$\widehat{G}_t$	$\widetilde{H}_t^1$	$\widetilde{H}_t^2$	$\widetilde{H}_t^3$	$\widetilde{H}_t^4$	$\widetilde{H}_t^5$	$\bar{R}^2$
2	0.658	-7289.019	6299.917	2421.811	289.536	156.468	0.493
HH	(3.661)	(-1.090)	(3.246)	(2.530)	(0.805)	(3.002)	
NW	(3.483)	(-1.047)	(2.456)	(2.511)	(0.870)	(3.251)	
3	1.215	-18680.646	10733.372	3591.580	777.587	247.058	0.458
HH	(3.509)	(-1.385)	(2.609)	(1.908)	(1.157)	(2.456)	
NW	(3.222)	(-1.350)	(2.069)	(1.893)	(1.250)	(2.674)	
4	1.650	-25946.328	14428.220	4372.539	1580.678	327.694	0.469
HH	(3.333)	(-1.403)	(2.393)	(1.678)	(1.708)	(2.352)	
NW	(2.982)	(-1.352)	(1.999)	(1.641)	(1.839)	(2.553)	
5	1.882	-23946.247	17411.594	4474.560	2317.655	368.405	0.433
HH	(2.917)	(-0.981)	(2.298)	(1.383)	(2.025)	(2.173)	
NW	(2.596)	(-0.961)	(1.958)	(1.340)	(2.171)	(2.355)	

Table 12: Regressions of Annual Excess Bond Returns on Jump Mean and Other Predictors.

This table examines the period after the momentous period 1979-1982. The return to an n-year zero-coupon Treasury bond from month t to month $t+12$ less the month-t yield on a one-year Treasury bond is regressed on JM_t realized jump mean, along with $\widetilde{H}_t^i, i = 1, \ldots, 5$, Duffee (2008)'s five latent yield factors (the 5th factor being the hidden factor) estimated using Kalman filtering, and/or $\widehat{G}_t$, the single predictor estimated by SAGLasso. The row labeled "HH" reports test statistics computed using standard errors with the Hansen-Hodrick GMM correction for overlap. The row labeled "NW" reports test statistics computed using standard errors with 18 Newey-West lags to correct serial correlation. The sample spans the period August 1984 to December 2007.

Panel A: Predictive regression of excess returns on realized jump mean along with SAGLasso factor/filtered factors.

maturity	JM_t	$\widehat{G}_t$	$\bar{R}^2$	JM_t	$\widetilde{H}_t^1$	$\widetilde{H}_t^2$	$\widetilde{H}_t^3$	$\widetilde{H}_t^4$	$\widetilde{H}_t^5$	$\bar{R}^2$
2	-5.324	1.014	0.389	-9.087	-6570.428	4928.972	1763.101	683.984	236.420	0.622
HH	(-2.784)	(4.353)		(-4.526)	(-1.130)	(2.218)	(1.539)	(3.135)	(4.057)	
NW	(-2.945)	(4.737)		(-4.709)	(-1.148)	(1.929)	(1.639)	(3.398)	(4.605)	
3	-10.917	1.892	0.397	-18.171	-17031.024	7637.060	2138.399	1511.034	402.815	0.613
HH	(-4.007)	(5.009)		(-5.196)	(-1.405)	(1.750)	(0.982)	(3.830)	(3.874)	
NW	(-4.011)	(5.198)		(-5.253)	(-1.448)	(1.538)	(1.045)	(4.123)	(4.401)	
4	-14.195	2.765	0.391	-24.791	-23679.120	10176.176	2379.333	2576.990	539.870	0.612
HH	(-5.044)	(5.831)		(-5.312)	(-1.324)	(1.636)	(0.789)	(4.795)	(3.879)	
NW	(-4.628)	(5.733)		(-5.337)	(-1.372)	(1.480)	(0.839)	(5.122)	(4.408)	
5	-17.339	3.275	0.369	-30.834	-20769.858	11528.194	1767.574	3463.989	625.379	0.590
HH	(-5.712)	(5.779)		(-5.537)	(-0.878)	(1.513)	(0.489)	(5.241)	(3.794)	
NW	(-4.953)	(5.525)		(-5.553)	(-0.916)	(1.381)	(0.517)	(5.552)	(4.302)	

Panel B: Predictive regression of excess returns on jump mean, SAGLasso factor and filtered yield factors.

maturity	JM_t	$\widehat{G}_t$	$\widetilde{H}_t^1$	$\widetilde{H}_t^2$	$\widetilde{H}_t^3$	$\widetilde{H}_t^4$	$\widetilde{H}_t^5$	$\bar{R}^2$
2	-8.373	0.472	-5744.607	3564.557	1250.825	423.294	213.014	0.656
HH	(-3.949)	(2.024)	(-1.095)	(1.510)	(1.379)	(1.659)	(3.601)	
NW	(-4.203)	(2.194)	(-1.078)	(1.355)	(1.382)	(1.779)	(4.015)	
3	-16.903	0.839	-15562.952	5211.520	1227.719	1047.603	361.206	0.643
HH	(-4.781)	(2.123)	(-1.404)	(1.184)	(0.680)	(2.294)	(3.360)	
NW	(-4.950)	(2.238)	(-1.418)	(1.051)	(0.685)	(2.451)	(3.758)	
4	-23.073	1.136	-21690.575	6890.714	1145.792	1949.259	483.510	0.639
HH	(-5.166)	(2.223)	(-1.313)	(1.114)	(0.454)	(3.011)	(3.299)	
NW	(-5.291)	(2.248)	(-1.337)	(1.011)	(0.453)	(3.207)	(3.687)	
5	-28.963	1.237	-18604.122	7949.978	424.117	2780.324	563.997	0.610
HH	(-5.530)	(2.017)	(-0.837)	(1.071)	(0.134)	(3.487)	(3.201)	
NW	(-5.658)	(1.985)	(-0.866)	(0.971)	(0.134)	(3.686)	(3.574)	

Table 13: Out-of-Sample Predictive Power of Supervised and Unsupervised Macro Factors

This table reports results from one-year-ahead out-of-sample forecast comparisons of n-period log excess bond returns, $rx_{t+1}^{(n)}$. Panel A reports forecast comparisons of an unrestricted model with the SAGLasso Factor $\widehat{G}_t$ as the predictor, against a restricted, constant expected returns benchmark. Panel B reports forecast comparisons of an unrestricted model that includes $\widehat{G}_t$ and $AR(6)$ as predictors, against an $AR(6)$ benchmark model. Panel C reports forecast comparisons of an unrestricted model, which includes $\widehat{G}_t$ and $\widehat{LN}_t$ as predictors, with a restricted benchmark model that includes a constant and $\widehat{LN}_t$. The column labeled "MSE_u/MSE_r" reports the ratio of the mean-squared forecasting error of the unrestricted model to the mean-squared forecasting error of the restricted benchmark model that excludes additional forecasting variables. The column labeled "Ericsson" reports the ENC-REG test statistics of Ericsson (1992), and its 95th percentile of the asymptotic distribution is $\Phi^{-1} = 1.645$. The column labeled "Clark-McCracken" reports the ENC-NEW test statistics of Clark and McCracken (2001), and its 95% critic value is 1.584 for testing one additional predictor. The two tests share the same null hypothesis that the benchmark model encompasses the unrestricted model with excess parameter.

Panel A: SAGLasso Factor $\widehat{G}_t$ v.s. constant

maturity (yr)	Ericsson	MSE_u/MSE_r	Clark-McCracken
2	3.364	0.7217	129.6905
3	3.159	0.7580	111.3578
4	3.201	0.7520	111.0675
5	3.370	0.7706	104.6911

Panel B: $\widehat{G}_t$ + AR(6) v.s. AR(6)

maturity (yr)	Ericsson	MSE_u/MSE_r	Clark-McCracken
2	6.728	0.6446	177.076
3	6.835	0.6421	160.268
4	7.083	0.6197	162.507
5	6.850	0.6384	149.888

Panel C: $\widehat{G}_t + \widehat{LN}_t$ v.s. $\widehat{LN}_t$ + constant

maturity (yr)	Ericsson	MSE_u/MSE_r	Clark-McCracken
2	2.730	0.9384	42.988
3	2.838	0.9064	40.440
4	3.139	0.8749	43.355
5	3.179	0.8716	41.994

Table 14: Finite-Sample Distributions Based on Different Data Generating Processes

This table summarizes results from 50000 Monte Carlo simulations based on a VAR (12) process or a macro-finance term structure model. Both data-generating processes satisfy the general null hypothesis that expected excess bond returns are time-varying but independent of the macroeconomy. And 528 months of data are generated for each simulation. Excess returns are regressed on the month-t values of 4 group macroeconomic factors. The in-sample test statistic is a Wald test, for which the New-West procedure is used with 18 lags, of the hypothesis that the coefficients are jointly zero. The table reports the empirical rejection rate using the five percent critical value for a $\chi^2(4)$ distribution, as well as the finite sample five percent critical value. Similar statistics are reported for the out-of-sample ENC-REG test of Ericsson (1992) and ENC-NEW test of Clark and McCracken (2001). The ENC-REG test has an asymptotic N(0,1) distribution and the 95% asymptotic critic value is 3.007 for the ENC-NEW test statistic.

			VAR (12)		Term Structure Model	
Type of regression	Test Statistic	Maturity (yr)	Rejection rate	True 5% CV	Rejection rate	True 5% CV
In-sample	Wald	2	0.2686	21.1813	0.2444	19.2377
	..	3	0.2703	20.8757	0.2428	18.7567
	..	4	0.2678	20.7292	0.2320	17.9501
	..	5	0.2715	20.6988	0.2300	17.5332
Out-of-Sample	Ericsson	2	0.0931	2.1600	0.0984	2.2282
	..	3	0.0902	2.1146	0.0934	2.2018
	..	4	0.0928	2.1527	0.0896	2.1203
	..	5	0.0942	2.1632	0.0874	2.0641
	Clark-McCracken	2	0.3670	28.7888	0.3454	26.8619
	..	3	0.3588	27.7468	0.3444	27.8161
	..	4	0.3593	27.6246	0.3360	25.8401
	..	5	0.3581	26.9972	0.3284	24.8980

Table 15: Finite-Sample Properties of Forecasting Regressions

This table summarizes results from 50000 Monte Carlo simulations based on a VMA(12) process or a macro-finance term structure model. Yield data generated from the VMA process satisfies the restrictive null hypothesis of no predictability. The term structure model satisfies the general null hypothesis that expected excess bond returns are time-varying but independent of the macroeconomy. And 528 months of data are generated for each simulation. Excess returns are regressed on the month-t values of four group macroeconomic factors. The in-sample test statistic is a Wald test of the hypothesis that the coefficients are jointly zero. To mimic the choices used in Table 2, the covariance matrix of the parameter estimates is computed using the robust Hansen-Hodrick approach for the restrictive null, and the Newey-West procedure is used with 18 lags for the general null. The table reports the empirical rejection rate using the five percent critical value for a $\chi^2(4)$ distribution, as well as the finite sample five percent critical value. Similar statistics are reported for the out-of-sample ENC-REG test of Ericsson (1992) and ENC-NEW test of Clark and McCracken (2001). The ENC-REG test has an asymptotic N(0,1) distribution and the 95% asymptotic critic value is 3.007 for the ENC-NEW test statistic.

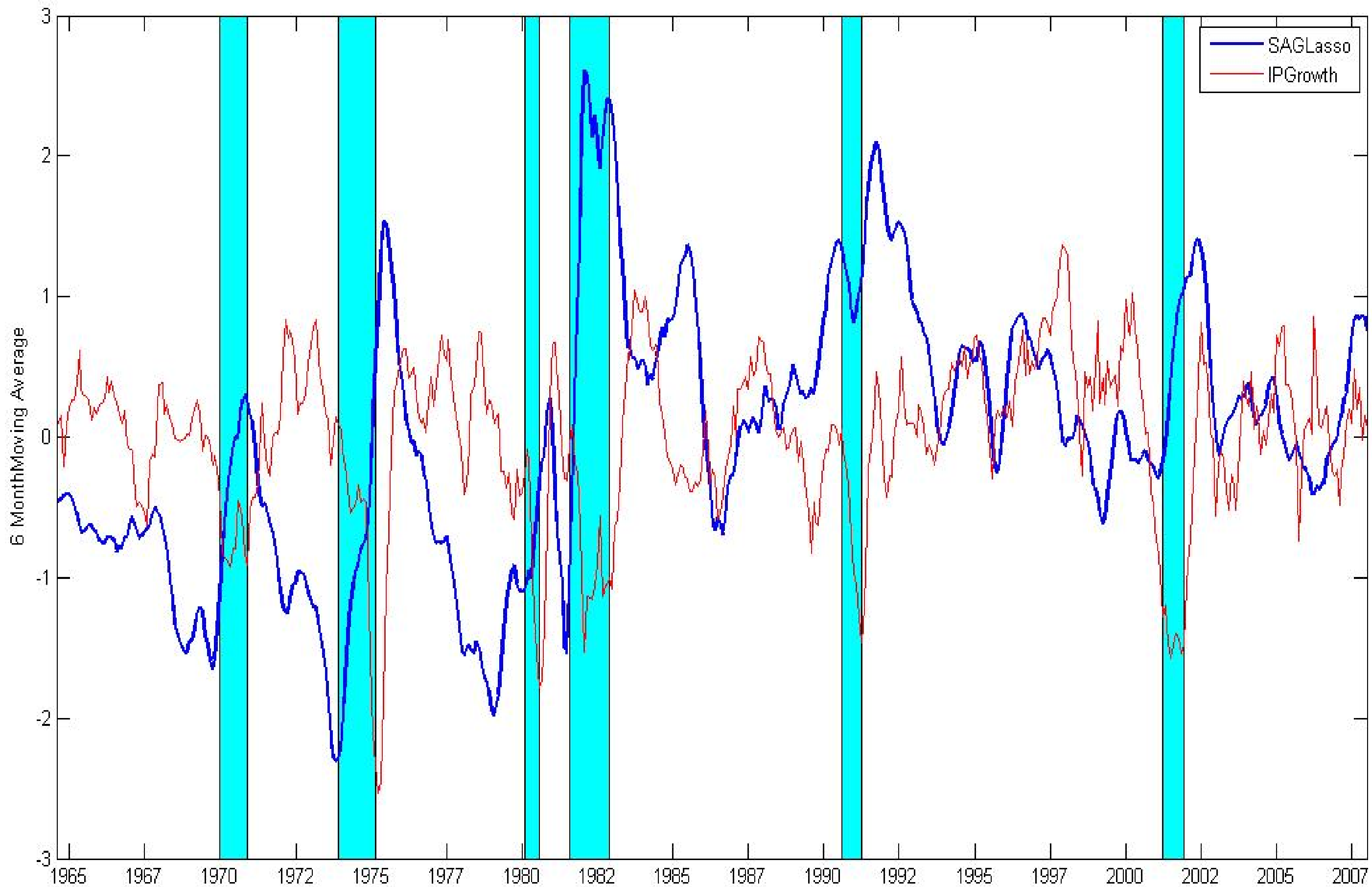
			Restrictive Null		General Null	
Type of regression	Type of Test	Maturity	Rejection rate	True 5% CV	Rejection rate	True 5% CV
In-Sample	Wald	2	0.1430	11.6890	0.2444	19.2377
	Wald	3	0.1307	11.3191	0.2428	18.7567
	Wald	4	0.1506	11.9882	0.2320	17.9501
	Wald	5	0.1328	11.3777	0.2300	17.5332
Out-of-Sample	Ericsson	2	0.0711	1.7230	0.0984	2.2282
	Ericsson	3	0.0743	1.7676	0.0934	2.2018
	Ericsson	4	0.0928	2.1527	0.0896	2.1203
	Ericsson	5	0.0942	2.1632	0.0874	2.0641
	Clark-McCracken	2	0.1345	10.5535	0.3454	26.8619
	Clark-McCracken	3	0.1024	9.2635	0.3444	27.8161
	Clark-McCracken	4	0.1482	11.7716	0.3360	25.8401
	Clark-McCracken	5	0.1456	11.8012	0.3284	24.8980

Table 16: Small Sample Inference for the Predictability of Excess Bond Returns

This table is the counterpart of Table 5 based on finite-sample distributions for test statistics. 50000 Monte Carlo simulations are run based on a VMA(12) process or a macro-finance term structure model. Yield data generated from the VMA process satisfies the restrictive null hypothesis of no predictability. The term structure model satisfies the general null hypothesis that expected excess bond returns are time-varying but independent of the macroeconomy. And 528 months of data are generated for each simulation. 95-percent confidence intervals for R^2, under each null hypothesis, are reported in square brackets. The column labeled "Joint Test" reports the Wald test statistics computed with the actual data, but the P-values are based on the empirical distributions of 50,000 bootstrapped samples. Both the restrictive and general hypotheses are defined as in Duffee (2008).

maturity (yr)	Null Hypothesis	R^2	Joint Test	P-val
2		0.436		
	Restrictive	[0.0025,0.0559]	56.977	[0.0000]
	General	[0.0072,0.1272]	67.484	[0.0002]
3		0.405		
	Restrictive	[0.0021,0.0441]	50.032	[0.0000]
	General	[0.0074,0.1277]	57.658	[0.0003]
4		0.392		
	Restrictive	[0.0028,0.0547]	52.324	[0.0000]
	General	[0.0071,0.1220]	57.995	[0.0010]
5		0.379		
	Restrictive	[0.0037,0.0543]	55.818	[0.0000]
	General	[0.0071,0.1190]	61.473	[0.0004]

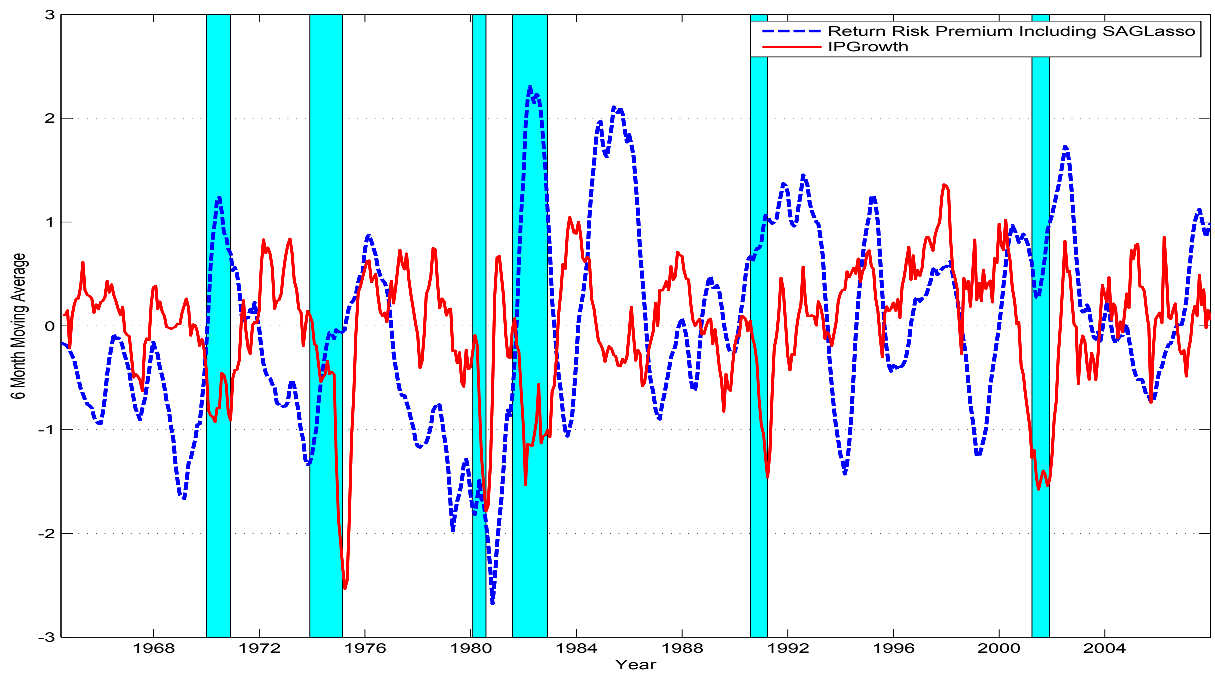
Figure 1: Time Variations of the SAGLasso Factor and the IP Growth



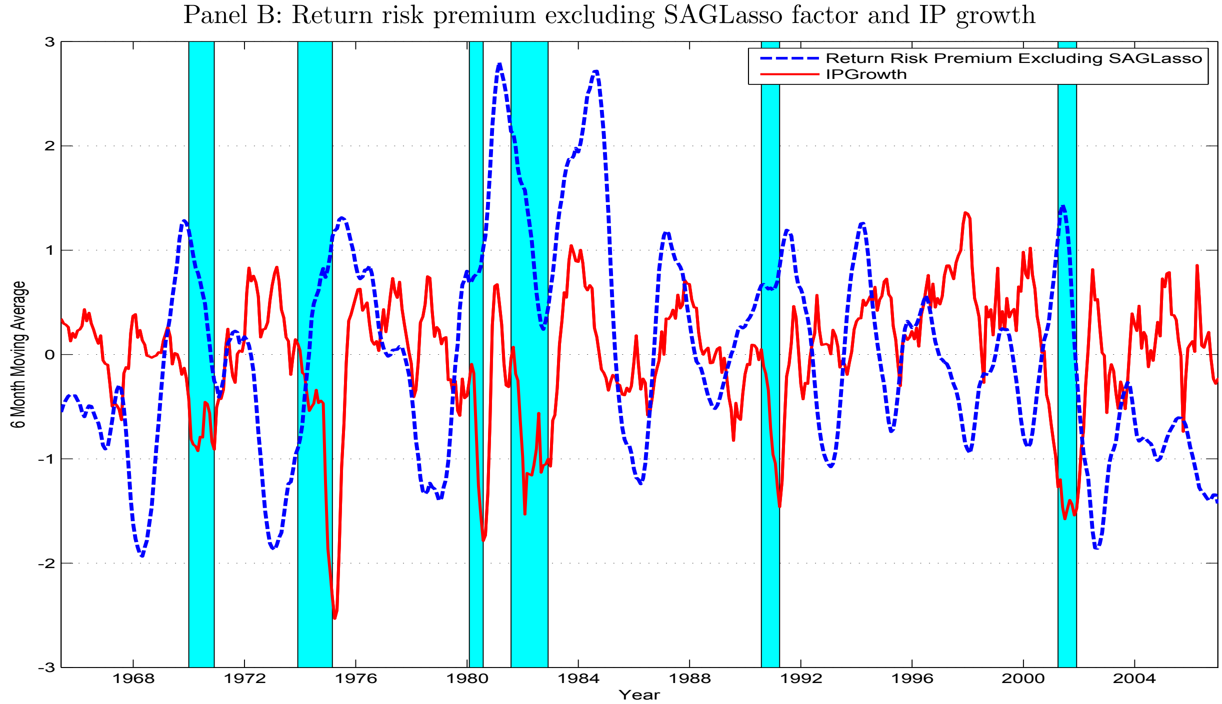
This figure plots the normalized SAGLasso Factor and the IP Growth over time. Shaded bars denote months designated as recessions by the National Bureau of Economic Research.

Figure 2: Return Risk Premium of the 5-year Bond Estimated from Different VAR Specifications

Panel A: Return risk premium including SAGLasso factor and IP growth.

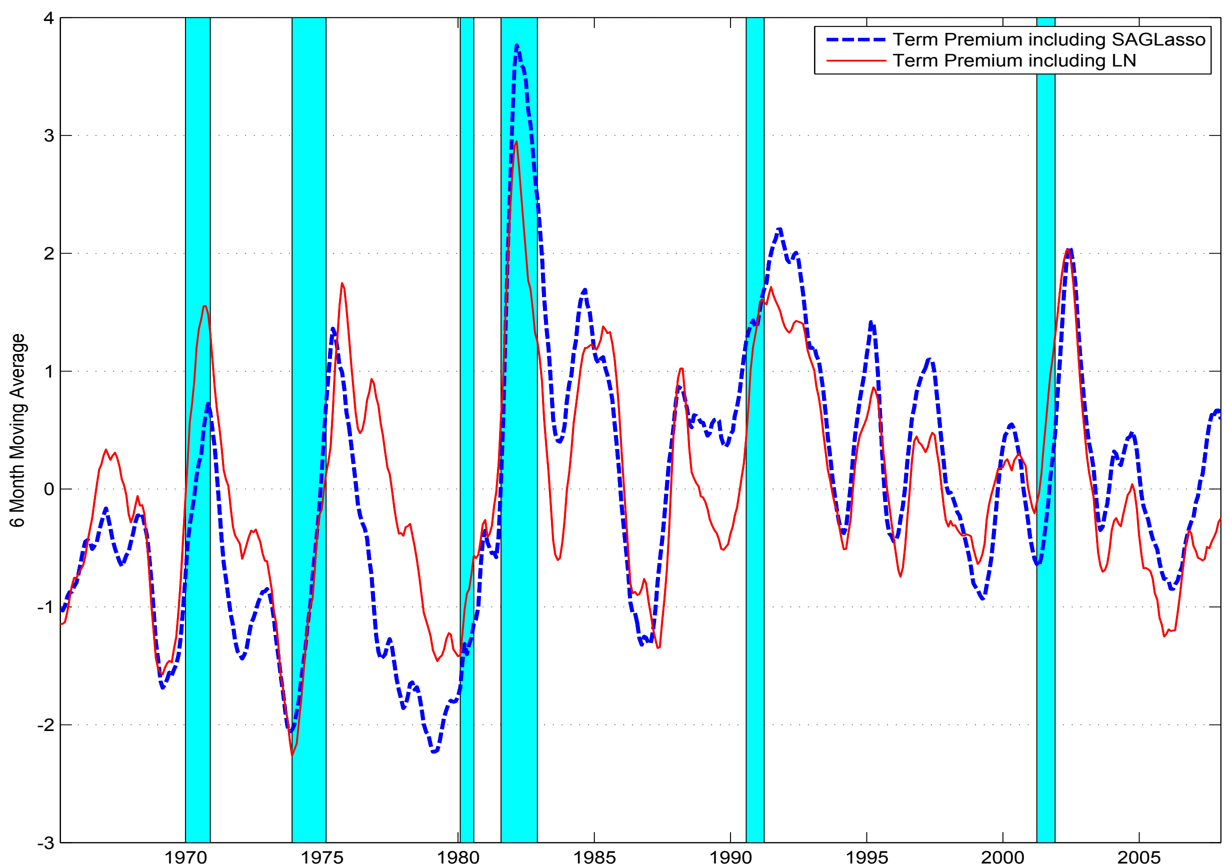


Panel B: Return risk premium excluding SAGLasso factor and IP growth



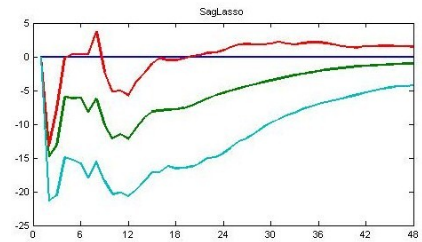
This figure plots the return risk premium of the 5-year bond that is estimated using two alternative VAR models, along with the growth of industrial production. The first VAR model involves state variables including both the Cochrane-Piazzesi (2005) and SAGLasso factors and the second specification excludes the SAGLasso factor. Shaded bars denote months designated as recessions by the National Bureau of Economic Research.

Figure 3: Term Premium of the 5-year Bond Estimated Different VAR Specifications

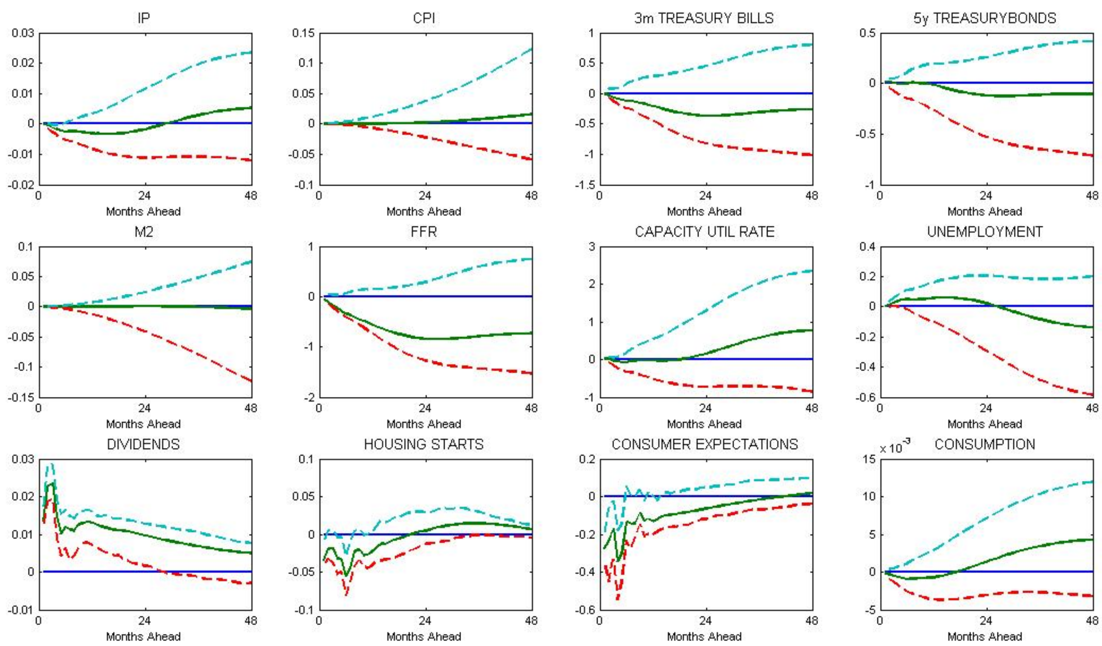


This figure plots the term premium of the 5-year bond that is estimated using two alternative VAR models. The first VAR model involves state variables including both the Cochrane-Piazzesi (2005) and SAGLasso factors and the second state vector contains the Cochrane-Piazzesi and the Ludvigson-Ng (2009b) factors. Shaded bars denote months designated as recessions by the National Bureau of Economic Research.

Figure 4: Impulse Response Generated from alternative FAVAR Specifications



Panel A: This figure plots the estimated impulse response, with 90 percent confidence interval, of the SAGLasso factor to one-standard-deviation shocks in logarithm IP growth.



Panel B: This figure plots the estimated impulse responses, with 90 percent confidence intervals, of key macroeconomic indicators to one-standard-deviation shocks in the SAGLasso factor.