Understanding Term Premia on Real Bonds*

Jing-Zhi Huang
Penn State University

Zhan Shi
Ohio State University

March 18, 2016

Abstract

This paper studies the dynamic behavior and determinants of risk premia on real bonds, using GDTSMs. We find that the real term structure itself contains a component that drives risk premia but is undetectable from cross section of bond yields. In addition, we present evidence on the link between real bond premia and macroeconomic variables. More specifically, we find that macro factors associated with real estate and consumer income and expenditure can capture a large portion of forecastable variation in excess returns on real bonds.

*We are very grateful to David Brown (MFA discussant), Stefania D'Amico (AFA discussant), Gregory Duffee, Olesya Grishchenko, Robert Kimmel, Anh Le, Yan Li (SICF discussant), Nick Pan (NFA discussant), Joel Vanden, Wei Yang, and seminar participants at the 2013 AFA Annual Meetings (San Diego), 2012 MFA Annual Meetings (New Orleans), 2012 NFA Annual Meetings (Niagara, Canada), 2012 Singapore International Conference on Finance (SICF), and 2012 Penn State Workshop on the Mathematics of Risk Management for helpful comments and suggestions. Huang is at the Smeal College of Business, Penn State University, University Park, PA 16802, USA. Email address: jxh56@psu.edu. Shi is at the Fisher College of Business, Ohio State University, Columbus, OH 43210, USA. Email address: shi.777@osu.edu.

1 Introduction

Many studies have examined the behavior and potential economic determinants of risk premia on nominal Treasury bonds. One stylized fact documented in this literature is that excess returns on nominal bonds are predictable (Fama and Bliss, 1987; Campbell and Shiller, 1991; Cochrane and Piazzesi, 2005), which invalidates the expectations hypothesis that bond risk premia are constant. Another stylized fact is that some factors found to have predictive power are not spanned by the current term structure of interest rates. These factors are uncovered and labeled as "hidden factors" by Duffee (2011), as they are hidden from cross-sectional yields despite their nontrivial effect on yield dynamics. On the other hand, the literature on risk premia of real bonds is rather thin, even though they represent a major asset class that can significantly expand one's investment opportunities (Campbell, Shiller, and Viceira, 2009; Bekaert and Wang, 2010).[1] In particular, while a few studies have documented the failure of expectations hypothesis in the real bond market (Evans, 1998; Pflueger and Viceira, 2011), evidence on the presence of hidden factors is still missing.

In this paper, we shed light on this important issue, arguing that the use of smoothed yields on zero-coupon inflation-indexed bonds in the literature is the main reason for this problem. In the TIPS market, the dominant source of zero-coupon yields used in academic studies is a panel produced by Gürkaynak, Sack, and Wright (2010),[2] who fit Svensson (1995)'s smooth function to the term structure. We find that although they are very useful in many aspects, such constructed smoothed real yield curves make the hidden component in term premia hard to detect. In contrast, based on a new set of zero-coupon real yields constructed using Fama and Bliss (1987)'s bootstrap method, we show that hidden factors play a key role in the dynamics of real short rates and excess bond returns.

More specifically, we estimate a four-factor Gaussian term structure model using Fama-Bliss type of unsmoothed real yields from 2004 through 2014. The Kalman filter allows us to infer the hidden component in risk premia. Based on the model's point estimates, about 40% of variations in expected monthly excess returns are attributable to this hidden component.

[1] For instance, TIPS can provide notable diversification benefits for investors (see, e.g., Campbell, Shiller, and Viceira, 2009; Huang and Zhong, 2013).

[2] Extant studies based on this dataset include, but are not limited to, DAmico et al. (2012), Grishchenko and Huang (2013), Pflueger and Viceira (2010, 2011), Christensen et al. (2010).

A one-standard-deviation change in this hidden factor raises the expected monthly return to a ten-year bond by 30 basis points (bps). In contrast, if the same model is fitted with smooth yields, information in the yield curve accounts for about 90% of variations in term premia, leaving little room for the hidden factor.

This finding highlights the differences between the unsmoothed and smoothed yield data in the inference of term-premium factors, which have been pointed out by Dai et al. (2004), Cochrane and Piazzesi (2008), and Le and Singleton (2013). In particular, an early version of Duffee (2011) suggests that "the smoothing process can artificially erase" a hidden factor, such as the fifth principal component (PC) of nominal yields that he demonstrates to have substantial forecasting power. However, to the best of our knowledge this paper is the first to formally assess the veracity of this conjecture. Our results show that, by retaining the information in hidden factors, the Fama-Bliss yield curve implies much more predictability in excess returns than do smoothed yield curves.

We document that return predictability implied by the estimated model differs sharply from that suggested by the sample. For example, the model's population properties imply that about 17% of the annual excess return on a five-year real bonds are forecastable by filtered state factors, but the R^2 in the sample ranges from 25% to 32%. For annual excess returns, we find that this divergence is driven by both the (model-suggested) small-sample bias and sampling error,[3] as the finite-sample R^2s produced by the model lie in between. However, this discrepancy does not affect our conclusion that information other that in the cross section of yields is very useful in forming return forecasts. Indeed, the sample R^2s for predictive regressions of annual returns would improve by $5 \sim 20\%$, depending on the bond maturity, if the filtered state vector replaces the yield curve PCs as the predictors.

Our model also implies that a substantial fraction of long-run risk premia on real bonds are orthogonal to the short-run premia. While Duffee (2010) has presented similar evidence with respect to nominal bonds, our results for the first time distinguish between the hidden components of monthly and annual expected excess returns, and more importantly, link

[3]In a recent study, Bauer and Rudebusch (2015) also find that both issues contribute to regression-based evidence for risk premium factors unspanned by the yield curve. Our main results on hidden factors in the real bond market are not derived from regression analysis. Moreover, while their study focuses on regressions of annual bond returns, we also present results on hidden factors in monthly excess returns, which show a much lower degree of persistence than annual returns and therefore are less subject to small-sample distortions.

them to different macroeconomic sectors. Specifically, we find that the "monthly" hidden factor shows significant covariation with general measures of real activity, and that a housing market indicator explains a substantial fraction of variation in the "annual" hidden factor.

The clear economic interpretation of our hidden factors disentangles themselves from hidden factors extracted from the nominal Treasury market, which are shown to have limited correlation with standard economic variables considered in the macro-finance literature (Duffee, 2011; Chernov and Mueller, 2012).[4] Owing to the lack of priori information on what macroeconomic variables are important to modeling term premia on real bonds, we do not incorporate macro data in our inference of yield-unspanned predictability as Joslin, Priebsch, and Singleton (2014). Nonetheless, our findings provide helpful guidance for future works on macro-finance models of real term structure.

Our analysis by no means indicates that parametric models underperform the Fama-Bliss method in estimating the term structure. Rather, we show that the Fama-Bliss and Svensson yield curves deliver comparable performance in pricing individual TIPS out of sample. As discussed in Gürkaynak et al. (2007, 2010), the choice about yield-curve estimation method depends on the purpose that the curve is intended to serve. While a smoothed yield curve may lose the information in higher-order principal components and is thus less attractive to practitioners, it could be very useful in macroeconomic interpretation and policy analysis.

The organization of the paper is as follows: The next sections set out the Fama-Bliss method of yield curve estimation and compares its out-of-sample goodness-of-fit with that of the Svensson method. Section 3 presents the term structure model and summarizes properties of the estimated model. In Section 4 we link the extracted hidden factors to expectations of future real rates and to macroeconomic variables. Section 5 concludes.

[4]Chernov and Mueller (2012) uncover a hidden factor whose presence is detected only if survey forecasts of inflation are included in the model estimation. But they find that this factor cannot be linked to any macro factors considered in their analysis.

2 Data

2.1 Data on TIPS and Real Yields

Information on historical TIPS issuance, including issuance dates, maturity dates and coupon rates, is obtained from TreasuryDirect. For each TIPS issue, we retrieve month-end price quotes from Thomson Reuters. The averages of bid and ask quotes are used as the true price. We exclude issues with less than six months to maturity duo to their low liquidity. Given the liquidity problems in the early years of TIPS program, our sample period extends from January 2004 to December 2014.[5] As such, there are 13 TIPS bonds and 39 TIPS notes contained in our sample.

The TIPS zero yield data most widely used in the literature is provided by Gürkaynak, Sack and Wright (2010, GSW hereafter), who estimate the Svensson (1995) parametric zero-curve model, an extension of Nelson and Siegel (1987). The resulting zero curve brings several desirable characteristics for monetary policy analysis: the smoothness and asymptotical flatness. However, the smoothness is attained by "evening out" the idiosyncratic movements in individual yields. As demonstrated in Section 3, it is difficult to distinguish shocks to higher-order principal components of the yield curve from idiosyncratic yield shocks. Consequently, if those higher-order factors contain non-trivial information on expected excess returns, the smoothed yield curve could lose this information. As such, the GSW data may not be suitable for this study given our purpose of forecasting bond excess returns. Below we construct TIPS zero curves using the non-parametric Fama-Bliss method.

2.1.1 Construction of the Fama-Bliss Real Yield Curve

Table 1 presents the term structures of means and volatilities of zero yields constructed using the Fama-Bliss (columns 1 and 2) and Svensson (columns 3 and 4) methods. For comparison, we also report summary statistics of the yield curve estimated by GSW[6] (columns 5 and 6), who utilize the price data provided by Barclays Capital Markets. Note that in our estimates

[5] See, e.g., Roll (2004); Shen (2006); D'Amico, Kim, and Wei (2014). In particular, Shen (2006) finds that before 2004 liquidity risk premia in TIPS were too large to be ignored.

[6] Their zero yield data is available through http://www.federalreserve.gov/econresdata/researchdata/feds200805.xls.

of both Fama-Bliss and Svensson yield curves, the one-year yield is available over a shorter sample, which starts from January 2006, while GSW's zero yields do not include maturities below two years.

Columns 1 through 4 indicate that the Fama-Bliss and Svensson methods, implemented using the same underlying quotes, produce subtle difference in the mean term structure. For maturities of one through eight years, the difference is no more than 1.8 bps. However, the Fama-Bliss curve tends to diverge from the Svensson one at the long end owing to its higher degree of flexibility. On the other hand, the Svensson and GSW yield curves used in Table 1 are constructed using the same method but different data. Columns 3 through 6 in the table indicate that different underlying data sources do not lead to any substantial discrepancy in the first two moments of the yield curves.

It is worth noting from the table that regardless of the underlying data and the estimation method used, all three yield curves show the same pattern across maturities: the average term structure of TIPS yields is upward-sloping and slightly concave,[7] while the term structure of yield volatilities is downward-sloping. These features are generally consistent with the evidence documented in Ang et al. (2008) and D'Amico et al. (2014), both of which estimate the dynamics of real rates using data on observed inflation and nominal yields (as well as their survey forecasts in the latter study).

2.1.2 From TIPS Yields to Real Yields

If TIPS are perfectly indexed, the estimated yield curves can be used to represent the term structure of real interest rates. However, in practice TIPS coupon and principal payments are linked to the three-month lagged inflation index, rather than the current level of the index. We correct for this indexation lag to obtain real yields that will be used in the empirical analysis of real risk premia. See Appendix A.1 for the details of the method used to estimate the correction due to the indexation lag.

The implementation of this estimation method requires data on both the nominal yields and TIPS yields. For the latter, we use the Fama-Bliss TIPS zero yields with maturities of one through ten years constructed in Section 2.1.1. To be consistent, we need to use nominal

[7]Note that in Table 1 statistics for the one-year bond are based on a shorter sample. The mean term structure is monotonically deceasing if the average for other maturities is taken over the same sample period.

yields constructed using the same method. However, Fama-Bliss nominal (zero) yields are available for maturities up to five years only. As such, we need to extend the Fama-Bliss nominal data to longer maturities. Following Le and Singleton (2013), we construct a set of monthly Fama-Bliss nominal zero yields with maturities of six through ten years over our sample period, using data on the quotes of individual Treasury coupon bonds from the CRSP Monthly Treasury Master File.

The last two columns of Table 1 present summary statistics of real zero yields estimated using the Fama-Bliss nominal and TIPS zero yields. Comparing with the summary statistics of Fama-Bliss TIPS yields shown in columns 3 and 4, we see that on average the real yield curve is below the TIPS curve as expected but the effect of the three-month indexation lag on the zero curve is small in our sample, especially for long-term zero yields. Untabulated results indicate the magnitude of our estimates of indexation-lag effect is generally in line with the results obtained in other studies (e.g., Grishchenko and Huang 2013).

Figure 1 plots estimated real, TIPS, and nominal yields for ten-year zero-coupon bonds over the period 2004–2014. Consistent with Table 1, the 10-year TIPS and real yields behave similarly in our sample. Although there is a notable difference between the two yields during the financial crisis period, the magnitude of the difference is generally small. Also, the long-term real interest rate seems highly correlated with the long-term nominate rate in the pre-crisis and the post-crisis periods.

2.2 Survey Data

One important part of our empirical analysis is to provide economic interpretations of extracted hidden factors. Section 4.3 examines whether these factors capture investor's expectations by comparing the hidden factors to survey forecasts of future real rates. We construct quarterly forecasts of the three-month real rate at various horizons using survey data from Blue Chip Financial Forecasts.

The Blue Chip conducts monthly surveys that ask approximately 45 financial market professionals for their projections of a set of economic fundamentals covering real, nominal and monetary variables. However, these variables do not include real yields but do cover nominal bond yields only. In order to estimate implied forecasts of real yields, we do the

following: First, in quarter i we extract every analyst' forecasts of the three-month nominal yield realized at the end of quarter i, as well as the three-month inflation rate realized at the end of quarter $i+1$. Then, we calculate each analyst's expectation of the three-month real rate (prevailing quarter $i+1$) by assuming a zero inflation premium at the three-month horizon.[8]

Note that the Blue Chip forecasts are made for a specific calender quarter and thus the forecast horizon varies each month. For example, the 3-month ahead forecast is only observed every third month in January, April, July and October, whereas the 1-month ahead forecast is only observed in March, June, September and December. As a result, we construct *quarterly* forecast of the three-month real rate at horizons ranging from one to twelve months.

Let y_t^3 denote the three-month real rate at time (month) t, and $\hat{E}_t(\Delta y_{t+j}^3)$ the mean survey forecast of the three-month real rate estimated at month t, where $j = 1, \ldots, 12$.[9] As these estimates are persistent, we focus on expected *changes* in real yields in regressions of real rate forecasts on hidden factors in Section 4.3. The expected change is defined as follows:

$$\hat{E}_t(\Delta y_{t+j}^3) = \hat{E}_t(y_{t+j}^3) - \hat{y}_t^3. \tag{1}$$

where $\hat{y}_t^3$ is the time-t estimate of y_t^3.

How to estimate short-term real yields such as $\hat{y}_t^3$? Note that neither three-month real T-bills nor equivalent short-term instruments have been issued in the market. Also, short-term real yields are hard to be effectively estimated from near-term TIPS because the effect of the indexation lag makes the prices of these securities erratic. We follow Campbell and Shiller (1996) and obtain a hypothetical real yield series by estimating a VAR model with monthly U.S. data, as elaborated in Appendix A.2.

[8]The assumption of a zero inflation risk premium at short horizons is also adopted in, e.g. Ang et al. (2008).

[9]Each month's report is released at the beginning of the month, thus the month-t forecasts of Treasury yields and inflation rates are achieved from month-$(t+1)$'s report.

3 Dynamic Term Structure Models with Real Yields

By definition, hidden factors are difficult to extract via a principal component analysis of yield curves; however, their presence can be inferred through the application of some filtering techniques. In this section we present empirical evidence of hidden factors in the real bond market, obtained by first constructing term structure models with real rates and then estimating the models via filtering with unsmoothed zero yields. We also show that the same filtering technique with the same term structure models fails to reveal the the presence of hidden factors when smoothed yield data are used instead.

To obtain maximum flexibility in the risk-premium specification, we rely on a four-factor Gaussian dynamic term structure model (GDTSM) to draw inferences about hidden factors. The choice of model dimension is based on three considerations. First, we find that the first two principal components can explain 99% of the cross-sectional variance of real zero-coupon yields in our sample, consistent with Gürkaynak, Sack, and Wright (2010) who use their estimated zero-curves of TIPS. Second, we follow Duffee (2011) by adding two more dimensions into a DTSM including only pricing factors. The resulting four-factor model gives hidden factor(s) a reasonable opportunity to manifest themselves. Finally, a model with five or more factor contains at least 29 free parameters, even with the restrictions imposed in Section 3.2. Extracting information about each of these parameters could be beyond the limits placed by our 11-year data sample.

3.1 The Modeling Framework

We begin with the canonical representation of GDTSMs developed by Joslin, Singleton, and Zhu (2011; JSZ hereafter), which defines the most general admissible GDTSM for a given dimension of the state vector. JSZ show that any canonical GDTSM can be transformed to a unique GDTSM parameterized by Θ_X, a parameter set to be specified below.

Let r denote the short rate and X the state vector of length N whose dynamics follow a Gaussian process under both the physical measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$. The

discrete-time $\mathbb{Q}$-dynamics of X_t and the resulting bond pricing formula are given as follows:

$$
\begin{aligned}
r_t &= \imath \cdot X_t, && (2)\\
\Delta X_t &= \mu_X^{\mathbb{Q}} + K_X^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}, && (3)\\
y_t^m &= A_m(\Theta_X^{\mathbb{Q}}) + B_m(\Theta_X^{\mathbb{Q}})' X_t && (4)
\end{aligned}
$$

where $\imath$ is a vector of ones; $\mu_X^{\mathbb{Q}} = [u_\infty^{\mathbb{Q}} \quad 0_{1\times(N-1)}]'$, with $u_\infty^{\mathbb{Q}}$ determining the long-run mean of short rate; $K_X^{\mathbb{Q}}$ has the real Jordan form determined by the eigenvalue vector $\gamma^{\mathbb{Q}}$; Σ_X is lower triangular and $\epsilon_t^{\mathbb{Q}} \sim N(0, I_N)$. In addition, y_t^m is the time-t yield of a zero-coupon bond maturing at m; $\Theta_X^{\mathbb{Q}} = (\gamma^{\mathbb{Q}}, u_\infty^{\mathbb{Q}}, \Sigma_X)$, a subset of Θ_X, governs X_t's $\mathbb{Q}$-dynamics and thus fully determines bond pricing; coefficients A_m and B_m satisfy a form of what is known as a Riccati equation.

Given the latent nature of the state vector X, we can obtain equivalent Gaussian models using invariant affine transformations. As we aim to identify sources of risk compensation and, in particular, to infer the hidden component of real bond risk premia, we consider a factor rotation that leads to orthogonal *factor shocks* instead of orthogonal factors. More specifically, we construct a rotation matrix using the principal component (PC) decomposition of *yield innovations* rather than that of yield levels.

Denote such a rotation matrix by L and the resulting new state vector by $\mathcal{P}$. Let Y_t be stacked time-t yields on k zero-coupon bonds with maturities $\mathcal{M} = \{m_1, \ldots, m_k\}$ as follows:

$$
Y_t = \mathcal{A}_x + \mathcal{B}_x X_t, \tag{5}
$$

where the $k \times N$ matrix $\mathcal{B}_x$ contains rows $\{B_m' | m \in \mathcal{M}\}$ for each bond.

It follows from Duffee (2011) that:

$$
\begin{aligned}
L &= W' \mathcal{B}_x(\Theta_X^{\mathbb{Q}}), && (6)\\
\mathcal{P}_t &\equiv L\left(X_t + (K_X^{\mathbb{Q}})^{-1} \mu_X^{\mathbb{Q}}\right), && (7)
\end{aligned}
$$

where rows of the columns of W are eigenvectors of the covariance matrix of shocks to Y_t and thus L diagonalizes the covariance matrix $\Sigma_X \Sigma_X'$ such that $\Omega_{\mathcal{P}} \equiv L\Sigma_X \Sigma_X' L'$ contains the corresponding eigenvalues. As a result, the $\mathbb{Q}$-dynamics of $\mathcal{P}_t$ are given by

$$\begin{aligned} \Delta\mathcal{P}_t &= K_{\mathcal{P}}^{\mathbb{Q}}\mathcal{P}_{t-1} + \Omega_{\mathcal{P}}^{1/2}\epsilon_t^{\mathbb{Q}}, & (8) \\ K_{\mathcal{P}}^{\mathbb{Q}} &= LK_X^{\mathbb{Q}}L^{-1}, & (9) \end{aligned}$$

where importantly, innovations among new state factors become independent of each other. In other words, each new factor has risk exposure to its shocks only. For instance, the constructed "level" factor has risk exposure to "level" shocks only. In addition to uncorrelated PC risks, another advantage of rotation L is that it depends solely on $\Theta_X^{\mathbb{Q}}$. As emphasized in Duffee (2011), this irrelevance of physical measure dynamics simplifies the model estimation.

Under the physical measure $\mathbb{P}$, the rotated state vector $\mathcal{P}$ has the following dynamics:

$$\begin{aligned} \Delta\mathcal{P}_t &= \mu_{\mathcal{P}}^{\mathbb{P}} + K_{\mathcal{P}}^{\mathbb{P}}\mathcal{P}_{t-1} + \Omega_{\mathcal{P}}^{1/2}\epsilon_t^{\mathbb{P}}, & (10) \\ \mu_{\mathcal{P}}^{\mathbb{P}} &= L\left(\mu_X^{\mathbb{P}} - K_X^{\mathbb{P}}(K_X^{\mathbb{Q}})^{-1}\mu_X^{\mathbb{Q}}\right), & (11) \\ K_{\mathcal{P}}^{\mathbb{P}} &= LK_X^{\mathbb{P}}L^{-1}. & (12) \end{aligned}$$

where $\mu_X^{\mathbb{P}}$ and $K_X^{\mathbb{P}}$ are the $\mathbb{P}$-counterparts of $\mu_X^{\mathbb{Q}}$ and $K_X^{\mathbb{Q}}$, respectively.

3.2 The Structure of Expected Excess Returns

If there are bond portfolios on which excess returns track the priced risks $\mathcal{P}_t$, their one-period risk premia equal the drift of $\mathcal{P}_t$ under $\mathbb{P}$ minus its drift under $\mathbb{Q}$. That is

$$\Omega_{\mathcal{P}}^{1/2}\Lambda_t = \mu^{\mathbb{P}} + (K_{\mathcal{P}}^{\mathbb{P}} - K_{\mathcal{P}}^{\mathbb{Q}})\mathcal{P}_t = \lambda_0 + \lambda_1\mathcal{P}_t, \quad (13)$$

where the first two elements of Λ_t represent the market price of "level" risk and that of "slope" risk, respectively.

An important question here is what restrictions should be placed on λ_0 and λ_1. Specifically, econometricians focus on the following two question on term premia specification: (1)

which of the PC risks are priced? (2) what is the dimension of "risk-premium factors" that drive expected excess returns? While several studies have looked into these questions in nominal bonds (see, e.g., Cochrane and Piazzesi 2005, 2008; Duffee 2010; Joslin, Priebsch, and Singleton 2014), evidence from the real bond market is scarce and variations in real risk premia are often assumed to be driven by a three- or four-dimensional state vector in the literature (see, e.g., Chernov and Mueller 2012; D'Amico et al. 2014; Haubrich et al. 2012). In this subsection, we conduct an extensive analysis in order to shed more light on these important issues. Below we summarize the main results from the analysis and leave the details of the analysis to Appendix B.

To address the first question above, we follow Duffee (2010) to focus on specifications that lead to a reasonable magnitude of model-implied Sharpe ratios. Using a four-factor GDTSM for real bonds, we find evidence that while investors generally demand compensation for exposure to both the risk of "level" shocks and that of "slope" shocks, the market prices of the third- and fourth-PC risks are fairly small in magnitude. As such, it seems reasonable to set the last two rows of λ_0 and λ_1 to zero.

Regarding the number of factors that drive expected excess returns (the second aforementioned question), we conduct first a principal component analysis of model-implied expected returns. This approach share the same spirit as Cochrane and Piazzesi (2008), with the differences that (1) we estimate expected excess returns using an (unconstrained) GDTSM instead of a return-forecasting regression and (2) we consider not only annual returns but also monthly returns. We find that the covariance matrix of annual returns is dominated by the first PC, which captures 98.5% of the variance. On the other hand, for monthly excess returns, the first PC explains about 95.7% of their variations and the second PC accounts for almost all the remaining variance.[10] These results suggest that variations in monthly (yearly) excess returns are largely driven by two factors (one factor).

Next, we conduct a rigourous test in the case of monthly excess returns. Specifically, following Joslin et al. (2011) we perform a likelihood-ratio test with the null hypotheses that

[10]This result may have something to do with the use of one-month yield in the construction of monthly excess returns. As inflation-indexed bills are never issued, no information about the short end of real term structure is available in our model estimation. As such, the filtered, noise-free one-month real yield depends on both the model and the information of longer-maturity bonds, and thus may differ from the "true" real short rate.

the rank of λ_1 is one. Test results indicate that we cannot reject the null that monthly expected excess returns are driven by a single linear combination of pricing factors at the conventional 5% significance level. This result along with our earlier finding that only the level and slope are priced factors suggests the following functional form for risk premia, where a single factor $\lambda_{1l}'\mathcal{P}_t$ drives expected excess returns:

$$\Omega_{\mathcal{P}}^{1/2}\Lambda_t = \begin{bmatrix} \lambda_{0l} \\ \lambda_{0s} \\ 0_{2\times 1} \end{bmatrix} + \begin{bmatrix} \lambda_{1l}' \\ \\ 0_{3\times 4} \\ \\ \end{bmatrix} \mathcal{P}_t, \tag{14}$$

where λ_{1l} is a vector of length 4. One important advantage of this functional form is that we can easily identify the hidden component of the single risk-premium factor $\lambda_{1l}'\mathcal{P}_t$.

In the analysis that follows, we consider mainly the configuration (14) but also report the results from the unconstrained model when necessary.

3.3 Model Estimation

When estimating an N-factor term structure model, econometricians often assume that N zero-coupon bond yields or N linear combinations of such yields are priced perfectly by the model. Unfortunately, this assumption is not suitable for the purpose of detecting hidden factors because with such factors, the length-N state vector cannot be inferred solely from properties of the cross-section of yields. For instance, if X_t in Eq. (5) contains hidden factors, $\mathcal{B}_x$ has a rank less than N. However, as noted in Duffee (2011), the presence of hidden factors can be inferred using filtering techniques. Furthermore, if the underlying model fits into the Gaussian term structure framework, the Kalman filter can produce correct conditional mean and covariance parameters.

As such, we assume that all zero yields are measured with error as the following:

$$\widetilde{Y}_t = \mathcal{A}_p + \mathcal{B}_p\mathcal{P}_t + \eta_t, \tag{15}$$

where $\widetilde{Y}_t$ denotes noise-contaminated yields; $\mathcal{B}_p$ their loadings on the rotated factors $\mathcal{P}_t$; the noise η_t is assumed to be independent across maturities and all yields share an identical standard deviation σ_η^2.

With the principal components rotation, there are 32 non-zero parameters in the constrained four-factor model,

$$\psi_{\mathcal{P}} = (\Phi_{\mathcal{P}}^{\mathbb{Q}}, \Omega_{\mathcal{P}}, \delta_{0\mathcal{P}}, \delta_{1\mathcal{P}}, \lambda_{0l}, \lambda_{0s}, \lambda_{1l}, \sigma_\eta),$$

where $\Phi_{\mathcal{P}}^{\mathbb{Q}} = K_{\mathcal{P}}^{\mathbb{Q}} + I$. However, not all risk-neutral parameters are identifiable under this factor rotation. According to the JSZ canonical form, there are only 22 free parameters after the 14 zero restrictions are imposed on the risk premium dynamics

$$\Theta_X = (\gamma^{\mathbb{Q}}, u_\infty^{\mathbb{Q}}, \Sigma_X, \lambda_{0l}, \lambda_{0s}, \lambda_{1l}, \sigma_\eta).$$

Given the transition equation (10) and the observation equation (15), it is straightforward to implement the Kalman filter and perform the maximum likelihood estimation. All models are estimated using a monthly panel of all ten constructed zero yields on real bonds. One-year bond yields are not available for all dates, which introduces missing data in the observation equation and are handled in the standard way by allowing the dimensions of the matrices $\mathcal{A}_p$ and $\mathcal{B}_p$ in Eq. (15) to be time-dependent.

To facilitate our following discussion, we report the estimate of $\psi_{\mathcal{P}}$ rather than Θ_X in Table 2, where standard errors from Monte Carlo simulations are in parentheses. The estimate of $\Omega_{\mathcal{P}}$ indicates that the first two principal components explain 99.4% of conditional variations in the term structure, consistent with a principal component analysis of the unconditional covariance matrix of bond yields.

To illustrate the mapping from the PCs to observed real yields, Figure 2 plots model-implied yield loadings (solid lines) on each of the four factor in four separate panels. The top two panels show that the first two factors play a noticeable role in the term structure. For example, a one-standard-deviation increase in the "level" factor (the first factor) raises the two-year yield by about 17 bps (the top left panel), a significant impact given that the sample mean of two-year yields is merely 1.3 bps. Note that the first factor here, $\mathcal{P}_{1t}$, does

not move all yields together by approximately the same amount because, by construction, $\mathcal{P}_{1t}$ is the "level" component for *innovations* to bond yields, not the conventional "level" factor extracted from the unconditional covariance matrix of yields.[11] In contrast, the yield curve's responses to one-standard-deviation shocks to the third and fourth factors are less than 3 bps regardless of maturities. This result should not be taken for granted as the state factors are PCs of yield innovations instead of yield levels. This also explains why the yield loading on the third factor $\mathcal{P}_{3t}$ does not have the familiar hump shape (the bottom left panel).

Any credible term structure model should produce principal components that match principal components in the data. Along with the model-implied loadings in (15), corresponding sample values are also displayed as triangle markers in each panel. They are estimated by regressing yields on the filtered factors. One key observation from the figure is that the model does a good job reproducing the sample loading on each factor. All sample loadings, except the one-year yield's on the third factor, are within two-sided 95% confidence bounds calculated from Monte Carlo simulations.[12] Moreover, the distance between the upper and lower bounds is economically small, no more than 5 bps in any panel. Overall, we find that the fitted loadings closely resemble the actual loadings at least for those maturities involved in model estimation.

The main conclusion drawn from Figure 2 is that only the "level" and "slope" factors make noticeable contributions to the cross section of real yields. Due to the small contribution of the higher-order factors (the 3rd and 4th ones), it might be difficult to disentangle them from noise in the real yield curve. Note that the estimated standard deviation of measurement error is about 1.51 bps, which is enough to obscure their effects on the term structure. As such, one may expect that these two factors do no play an important role in determining term premia on real bonds, given the conventional wisdom that the current term structure should contain any information investors have about future excess returns. However, results in the next subsection show that these higher-order factors capture a substantial portion of forecastable variations in excess bond returns.

[11] The (untabulated) estimate of W, defined in Eq. (6), confirms that $\mathcal{P}_{1t}$'s effect on yield shocks are much more uniform across maturities.

[12] Since the sample of the one-year yield is shorter than other maturities, it generally has the largest divergence between model-implied and sample loadings.

3.4 Population Properties of Expected Excess Returns

This subsection examines expected excess returns implied from the four-factor term-structure model estimated in Section 3.3.

Table 3 reports the main features of excess returns to a five-year bond (Panel A) and a 10-year bond (Panel B), based on both monthly returns (columns 2 through 6) and annual returns (columns 7 through 11). These two bonds are chosen because a five-year (10-year) bond is considered to be a long bond in studies of nominal (real) bond risk premia. However, as risk premia are driven by a single factor in the model, the (untabulated) results for other maturities are similar.

Consider the five-year bond first (Panel A). Note that the unconditional mean of monthly excess return is 9 bps (column 2) or 1.07% annualized and, as expected, higher than the mean annual excess return of 0.94% (column 7), due to the sightly convex upward real yield curve. However, there is substantial statistical uncertainty in these point estimates as can be seen from their 95% confidence intervals shown in columns 2 and 8 (based on the same 1,000 simulations used in Table 2). For instance, the mean of monthly excess returns ranges from -6.7 to 29 bps, implying a nontrivial probability that the ML estimation with 132 monthly observations produces a negative point estimate. The unconditional variance of returns (uncontaminated by measurement errors) is 10.6 bps for monthly excess returns (column 3) and 73.2 bps for annual ones (column 8), suggesting that the *realized* excess return is highly volatile.

The variation in the predicted excess returns—the variable of interest here—depends on the information set used in return prediction. Under the assumption that econometricians perfectly observe the current state factors, the predictable variance for the monthly returns is 1.73 bps (column 4), accounting for 16% of the total return variance of 10.6 bps (column 3), and as expected, the annual return is more predictable, with a model-implied R^2 of about 23% (16.9/73.2). In reality, econometricians are unlikely to have access to this information (unless they conduct surveys on investors' expectation) and instead, either reply on the current term structure or perform filtering analysis to infer the state vector. We consider the former approach below and the latter in Section 3.5.

Without loss of generality, the time-t yield curve information used to calculate the expected excess return $E_t(rx^m_{t+j})$ is assumed to be summarized by the first four PCs of the zero curve with noise.[13] Let $\widetilde{\mathcal{P}}$ denote the vector of these four PCs. By the law of iterated conditioning, forecasting returns with $\widetilde{\mathcal{P}}$ boils down to an inference of the true state vector $\mathcal{P}$ based on $\widetilde{\mathcal{P}}$. For example, in the case of monthly returns we have

$$E(rx^m_{t+1}|\widetilde{\mathcal{P}}_t) = E(E(rx^m_{t+1}|\mathcal{P}_t)|\widetilde{\mathcal{P}}_t) \;=\; \alpha^1_m + \beta^1_m E(\mathcal{P}_t|\widetilde{\mathcal{P}}_t), \tag{16a}$$

$$\alpha^1_m = -(m-1)B'_{\mathcal{P},m-1}\lambda_0 - \frac{1}{2}(m-1)^2 B'_{\mathcal{P},m-1}\Omega_{\mathcal{P}}B_{\mathcal{P},m-1}, \tag{16b}$$

$$\beta^1_m = -(m-1)B'_{\mathcal{P},m-1}\lambda_1. \tag{16c}$$

With partial information about $\mathcal{P}$ econometricians tend to make less accurate forecasts than $E(rx^m_{t+j}|\mathcal{P}_t)$. Indeed, as indicated by columns 5 and 10 of Table 3, the $\widetilde{\mathcal{P}}_t$-based return prediction yields an R^2 of 9.43% (1.0/10.6) for monthly returns and that of 13.95% (10.2/73.2) for annual returns. As such, the spread between $\mathcal{P}_t$-based (full-information) R^2 and $\widetilde{\mathcal{P}}_t$-based (partial-information) R^2 is 6.57% (=16-9.43) for monthly returns and 9.05% (=23-13.95) for annual returns. This spread between $\mathcal{P}_t$- and $\widetilde{\mathcal{P}}_t$-based R^2s can be traced back to measurement errors introduced in Eq. (15). While these errors hardly affect the first two PCs due to their dominant role in the cross section, measurement errors make it difficult to extract the true 3rd and 4th PCs ($\mathcal{P}_{3t}$ and $\mathcal{P}_{4t}$) from the observed term structure. If these higher-order factors contain substantial information about future bond returns, then it is missing in $\widetilde{\mathcal{P}}_{3t}$ and $\widetilde{\mathcal{P}}_{4t}$ and thus in $\widetilde{\mathcal{P}}_t$. Indeed, replacing $\widetilde{\mathcal{P}}_t$ with $[\widetilde{\mathcal{P}}_{1t}\ \widetilde{\mathcal{P}}_{2t}]'$ makes little difference in the variance of return forecasts.[14]

Another way to measure the gap between the information contained in $\mathcal{P}_t$ and that in $\widetilde{\mathcal{P}}_t$ is to calculate the ratio of variances of these two forecasts, denoted $VR_{\widetilde{\mathcal{P}}}$, as suggested by

[13] Alternatively, we can use four bond yields themselves as a proxy for state variables. Untabulated results indicate that using 2-, 5-, 7- and 10-year yields leads to spanning ratios even lower than those reported in Table 3.

[14] As implied in Figure 2, different factor rotation schemes do not significantly alter the information content in the first two factors. We confirm this by analysing the expected excess returns conditioning on $[\mathcal{P}_{1t}\ \mathcal{P}_{2t}]$. The resulting variance is comparable to $Var(rx^m_{t+j}|[\widetilde{\mathcal{P}}_{1t}\ \widetilde{\mathcal{P}}_{2t}])$.

Duffee (2011). In the case of monthly excess returns, the ratio equals

$$VR_{\widetilde{\mathcal{P}}} = \frac{\beta_m^1 \text{Var}(\mathcal{P}_t|\widetilde{\mathcal{P}}_t)\beta_m^{1'}}{\beta_m^1 \text{Var}(\mathcal{P}_t)\beta_m^{1'}}. \tag{17}$$

Given the restrictions imposed on λ_1, it follows from Eq. (16c) that β_m^1 is essentially determined by the product of the first element in $B_{\mathcal{P},m-1}$ and λ_{1l}. Hence, $VR_{\widetilde{\mathcal{P}}}$ is close to one if the last two elements in λ_{1l} (corresponding to the third and fourth factors) are zeros. This leads to the special case in which higher-order factors do not affect the risk compensation for the "level" factor (see Eq. (14)). Put differently, when variations in term premia are driven by only the first two factors, econometricians have pretty much the same information set as possessed by market participants.

The point estimate of $VR_{\widetilde{\mathcal{P}}}$, reported in Panel A of Table 3, is 59.1% for monthly returns (column 6) and 60.5% for annual returns (column 11). In other words, 40% of the information in the true state vector is lost if we rely on cross-sectional yields to forecast excess returns. Even if we focus on the upper bound of $\widehat{VR}_{\widetilde{\mathcal{P}}}$'s confidence intervals reported in the table, at least 20% of the information contained in $\mathcal{P}_t$ is hidden in the yield curve.

Consider next the 10-year bond. As expected, results for this maturity (reported in Panel B) are similar to those for the five-year bond. In particular, as shown in column 6 (11), only 59.3% (66.9%) of variance in expected excess monthly (annual) returns is captured by the cross section of bond yields. Also, the unconditional mean of the excess return is higher than that of five-year bonds (column 2), reflecting the slightly upward-sloping term structure of real bond yields. Nevertheless, the model-implied Sharpe ratio is lower for ten-year bonds. This observation is consistent with the generally inverse relation between the TIPS maturity and Sharpe ratio as shown in Panel A of Table 9.

The analysis with the Fama-Bliss data has provided strong evidence for hidden factors in the real bond market. Do similar results obtain with the GSW data? We estimate the same four-factor (constrained) GDTSM using the GSW zero yields. Results reported in rows labeled "SV_4" in Table 3 show that not much of the variations in expected returns is hidden from the cross-section of yields. More specifically, four PCs of the yield curve capture around 86–89% (90%) of the predictability of the state vector in the case of monthly

(annual) excess returns. These results are especially worth noting given that the model-implied unconditional moments here under "SV_4" are very close to those under "FB_4". They indicate that the risk premium factor in Model "SV_4" loads thinly on the third and fourth factors. Accordingly, untabulated results confirm that with the GSW data the last two elements in $\hat{\lambda}_{1l}$ are barely significant. The above evidence indicates that the smooth Svensson function has, to a great extent, washed out the hidden component in term premia.

The results presented so far are based on the four-factor GDTSM. A relevant question is whether a lower-dimensional GDTSM can reproduce the evidence of hidden factors. As reported in rows labeled "FB_2" in Table 3, variance ratios implied by the two-factor GDTSM are greater than 94.5% (columns 6 and 11), regardless of bond maturities or return horizons considered. This result is expected as the role of the first two factors in the cross section is too important to be curtained by measurement errors.

Surprisingly, the three-factor GDTSM does not imply a considerable degree of unspanned predictability either, as shown in rows labeled "FB_3" in the table. For instance, the model-implied variance ratio is over 93% for both monthly and annual returns to the 10-year bond (Panel B) and, in general, much closer to its counterpart under the two-factor model than to the four-factor model. These results suggest that the "curvature" factor does not drive a substantial fraction of the predictability of excess returns, at least compared to the fourth factor. Indeed, unreported variance decomposition analysis (based on Model "FB_4") shows that a one-standard-deviation shock to the fourth factor moves the expected monthly return on a five-year bond by about three times as an equivalent "curvature" shock does.

3.5 Inferring the State Vector from Yield Dynamics

Given the evidence from the four-factor GDTSM that the cross section of yields contains less information about risk premia than is in the state vector, we need to look beyond the yield curve when making return forecasts. However, the true state vector is not directly observable. This subsection examines the question that how much information econometricians can recover from yield dynamics by using Kalman filter.

We first conduct a simulation exercise to assess the effectiveness of Kalman filter in inferring the state vector in the real bond market. The details of the exercise are provided

in Appendix C. We summarize the main results of the exercise here: First, Kalman filter really makes difference in its accurate estimates of the third and fourth factors (higher-order factors). Interestingly, Kalman filtering does a better job in inferring the fourth factor than the third factor. Second, Regressions of filtered estimates on observed yields produce an R^2 ranging from 82% to about 86% for the third and fourth factors; still, the 18-24% unexplained portions of these estimates are clear evidence for the unspanned nature of the third and fourth factors. These results imply that when model parameters are unknown, econometricians can still effectively infer the state vector by estimating the model with a reasonable, eleven-year sample.

Next, we evaluate the effectiveness of filtered (and smoothed) estimates of factors in forecasting excess returns and especially, relative to that of the true state vector containing all information on expected returns. Below we focus first on the population R^2 of predictive regressions implied by the estimated model and then on the finite-sample R^2.

3.5.1 Population R^2

Consider the j-month-ahead return on an m-period bond. For regressions on the true state vector, their R^2s values can be calculated analytically. The population R^2 is given by

$$R^2(\mathcal{P}_t) = \frac{\beta_m^j \text{Var}(\mathcal{P}_t)\beta_m^{j'}}{\beta_m^j \text{Var}(\mathcal{P}_t)\beta_m^{j'} + (m-j)^2 B_{\mathcal{P},m-j}\Omega B'_{\mathcal{P},m-1} + 2(m^2 - jm + j^2)\sigma_\eta^2}, \tag{18a}$$

$$\beta_m^j = mB'_{\mathcal{P},m} - (m-j)B'_{\mathcal{P},m-j}(\Phi_{\mathcal{P}}^{\mathbb{Q}})^j - jB'_{\mathcal{P},j}. \tag{18b}$$

Unlike the R^2 for "true" returns, (18a) takes measurement errors into account, as reflected by the last term in the denominator. For this reason, $R^2(\mathcal{P}_t)$ is slightly smaller than the corresponding R^2 implied by Table 3.

Similarly, the population R^2 using the filtered (smoothed) estimate can be obtained by replacing $\text{Var}(\mathcal{P}_t)$ in the numerator in Eq. (18a) with the unconditional variance of $\widehat{\mathcal{P}}_{t|t}$ ($\widehat{\mathcal{P}}_{t|T}$). For instance, the R^2 using the filtered estimate is given by

$$R^2(\widehat{\mathcal{P}}_{t|t}) = \frac{\beta_m^j \text{Var}(\widehat{\mathcal{P}}_{t|t})\beta_m^{j'}}{\beta_m^j \text{Var}(\mathcal{P}_t)\beta_m^{j'} + (m-j)^2 B_{\mathcal{P},m-j}\Omega B'_{\mathcal{P},m-j} + 2(m^2 - jm + j^2)\sigma_\eta^2}. \tag{19}$$

Since there is no analytic expression for $\text{Var}(\widehat{\mathcal{P}}_{t|t})$ or $\text{Var}(\widehat{\mathcal{P}}_{t|T})$, they are again constructed by simulating 10,000 years of data.

The first three rows in Panel A of Table 4 report results respectively for (the full-information) $R^2(\mathcal{P}_t)$, $R^2(\widehat{\mathcal{P}}_{t|t})$, and $R^2(\widehat{\mathcal{P}}_{t|T})$, based on both monthly and annual returns. We make three observations from these results. First, there is a greater predictability for annual returns than for monthly returns. Second, as expected, the Kalman filter cannot recover all information in the true factors about risk premia. Specifically, the gap between $R^2(\mathcal{P}_t)$ and $R^2(\widehat{\mathcal{P}}_{t|t})$ ranges from 3.4% to 5.8%, depending on the return horizon and bond maturity. Nevertheless, $R^2(\widehat{\mathcal{P}}_{t|t})$ is greater than the yield-curve based R^2. For example, the Kalman filter captures 10.8% of variation in monthly excess return on a five-year bond, while the yield curve captures only 8.37%.[15] Also, as is the case with the true state vector, forecasting with filtered estimates also leads to greater predictability for annual returns than for monthly returns. Third, $R^2(\widehat{\mathcal{P}}_{t|T})$ is greater than $R^2(\widehat{\mathcal{P}}_{t|t})$, regardless of the return horizon or bond maturity considered. Due to the look-ahead feature of smoothed estimates, they contain more information on future bond returns than filtered estimates. Also, although $R^2(\widehat{\mathcal{P}}_{t|T})$ is less than $R^2(\mathcal{P}_t)$ in our estimated model, $R^2(\widehat{\mathcal{P}}_{t|T})$ can be higher than $R^2(\mathcal{P}_t)$ theoretically as the Kalman smoother may contain future information beyond the current state vector.

An alternative approach to assessing the predictive power of estimated factors is to construct OLS estimates. Calculation of $R^2(\widehat{\mathcal{P}}_{t|t})$ in (19) is essentially inferring the market expectation of future excess returns, under the null that the estimated model is true. But we can also lift the functional restriction and directly run a predictive regression on filtered factor estimates. The resulting R^2 is given by

$$R^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t}) = \frac{\gamma_m^j \text{Var}(\widehat{\mathcal{P}}_{t|t})\gamma_m^{j'}}{\beta_m^j \text{Var}(\mathcal{P}_t)\beta_m^{j'} + (m-j)^2 B_{\mathcal{P},m-j}\Omega B'_{\mathcal{P},m-j} + 2(m^2 - jm + j^2)\sigma_\eta^2}, \tag{20}$$

where γ_m denotes the OLS coefficients. The row labeled "OLS Filtering" in Panel A of Table 4 shows that this $R^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t})$ is less than (very close to) $R^2(\widehat{\mathcal{P}}_{t|t})$ for monthly (annual) returns. Hence, relying on OLS to form the return predictor does not produce better results.

[15] It is slightly lower than the R^2 implied by Table 3, in which bond returns are uncontaminated by measurement errors.

This implication is not surprising in the current setting, since the data is generated by the estimated model. Given a sufficiently long simulated sample, we expect β_m to pick a more accurate linear combination of factors than γ_m does. Note that the latter forces the orthogonality between the resulting predictor $\gamma_m \widehat{\mathcal{P}}_{t|t}$ and the residual, which is not true because $\widehat{\mathcal{P}}_{t|t}$ does not capture all forecastable variations in excess returns. For completeness, row 5 in the panel reports the results for $R^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t})$; however, it is not particularly interesting as $\widehat{\mathcal{P}}_{t|T}$ contains information available after t.

3.5.2 Finite-Sample R^2

With an extremely long sample, it is reasonable to assume that econometricians would ultimately learn the true model parameters, as long as their estimation method is consistent. In practice, only a limited time series of data is available. Consequently, the fraction of forecastable variation that econometricians can capture depends on the estimator's efficiency as well as the model's finite-sample properties. This subsection examines such fractions of forecastable variation.

As before, we use 1,000 small-sample simulations to examine the accuracy of forecasts produced by the Kalman filter in this setting. To facilitate the computation of $R^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t})$, we assume that (noise-contaminated) zero yields are observable for all maturities, but only bonds with maturities of one through ten years are used to estimate the model. Note that once the model estimation and Kalman filtering is done, we also obtain filtered innovations to the state vector $\hat{\epsilon}_{t|t}$ as a by-product.

As a result, the model-implied finite-sample R^2 can be expressed as follows:

$$\widehat{R}^2(\widehat{\mathcal{P}}_{t|t}) = \frac{\hat{\beta}^j_m \widehat{\text{Var}}(\widehat{\mathcal{P}}_{t|t}) \hat{\beta}^{j'}_m}{\widehat{\text{Var}}\left(\hat{\beta}^j_m \mathcal{P}_{t|t} - (m-j)\hat{B}_{\mathcal{P},m-j}\hat{\Omega}^{1/2}\hat{\epsilon}_{t+1|t+1}\right) + 2(m^2 - jm + j^2)\hat{\sigma}^2_\eta}, \tag{21}$$

where the hat in $\widehat{R}^2(\cdot)$ denotes a finite-sample R^2. Similarly, the finite-sample equivalent to the ratio (20) is given by

$$\widehat{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t}) = \frac{\hat{\gamma}^j_m \widehat{\text{Var}}(\widehat{\mathcal{P}}_{t|t}) \hat{\gamma}^{j'}_m}{\widehat{\text{Var}}\left(rx^m_{t+j}\right)}. \tag{22}$$

The smoothed estimates-based $\widehat{R}^2(\widehat{\mathcal{P}}_{t|T})$ and $\widehat{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|T})$ can be obtained similarly.

We calculate the above four ratios using simulations and report their average values of 1,000 simulations in Panel B of Table 4. Due to the uncertainty in sampling and estimation, most ratios in Panel B deviate from their counterparts in Panel A by at least 1%. However, notice that the deviation is smaller for $\widehat{R}^2(\widehat{\mathcal{P}}_{t|t})$ and $\widehat{R}^2(\widehat{\mathcal{P}}_{t|T})$. Therefore, if econometricians fully trust the model estimated with maximum likelihood, their measurements of R^2s would be more in line with their "true" values; if they only appreciate the ability of the Kalman filter (without the model) to extract information from yield dynamics, they tend to seriously underestimate the predictability of monthly excess returns and overestimate predictability of annual returns. Despite the clear small-sample bias, filtered factors alone still appear more informative about the expected excess returns than the cross section of yields. For example, regressing the monthly return to a five-year bond on these factors yields an R^2 of 4.7%, while the first four PCs of the yield curve capture only 3.2% of return variations (untabulated).

When the bond data is truly generated by a GDTSM, the (true) model-suggested optimal predictor $\beta_m^j \widehat{\mathcal{P}}_{t|t}$ is expected to have the best small-sample performance. As demonstrated in Panel B, even if parameters constituting β_m^j are unknown and need to be estimated, $\hat{\beta}_m^j \widehat{\mathcal{P}}_{t|t}$ still outperform its OLS counterpart (in terms of the closeness to the population R^2). In reality the yield dynamics may not be consistent with the Gaussian assumption. It is thus important to make inferences about expected excess returns in our real sample, especially, to explore the difference between the sample R^2 and the model-implied finite-sample R^2.

3.5.3 R^2 in the sample

The first row in Panel C of Table 4 reports the sample version of $\widehat{R}^2(\widehat{\mathcal{P}}_{t|t})$ given in (21), denoted $\widetilde{R}^2(\widehat{\mathcal{P}}_{t|t})$. Clearly $\widetilde{R}^2(\widehat{\mathcal{P}}_{t|t})$ is greater than $\widehat{R}^2(\widehat{\mathcal{P}}_{t|t})$ (the average of simulated samples), regardless of the return horizon and bond maturity; the difference is substantial for annual excess returns, around 10%. However, the sample $\widetilde{R}^2(\widehat{\mathcal{P}}_{t|t})$ is still covered by the 90% confidence bounds constructed from the same simulations, as shown in brackets underneath the mean ratios in Panel B.[16] This implies that the filtered estimate of the state factor does

[16]Since all ratios in Panel C are greater than their corresponding simulated ratios in Panel B, it becomes more like a one-sided hypothesis test. Therefore, we report the 90% confidence interval where the upper bound is the 50th largest ratio among the 1,000 simulations.

not offer compelling evidence for model misspecification. The same conclusion can be drawn from $\widetilde{R}^2(\widehat{\mathcal{P}}_{t|T})$ (the sample version of $\widehat{R}^2(\widehat{\mathcal{P}}_{t|T})$) reported in the second row.[17]

We implement the sample version of (22), denoted $\widetilde{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t})$, for annual returns (over the 2006-2014 period) only, due to the unavailability of one-month and one-year yields. As can be seen from Panels B and C, for both 5- and 10-year bonds $\widetilde{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t})$ is higher than $\widehat{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t})$ and, in fact, near the upper bound of the latter's 90% confidence interval. An adjustment for the difference in the number of observations (between the regression in the sample and the regression in simulations) narrows the gap slightly, but the sample ratio $\widetilde{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t|t})$ is still substantially higher.

Overall, the simulation evidence indicates that the R^2 realized in our data sample seems a bit too high compared to the expected values implied by the model. On the other hand, there is a serious possibility that the estimated model truly describes the data-generating process but we somehow picked a particular sample with an R^2 well above the average.

We find that both the finite-sample bias (the differences between Panel A and B) and sampling uncertainty (the differences between Panel B and C) contribute to the inflated R^2s in our data sample. The magnitude of this inflation is consistent with the evidence about nominal Treasury bonds, as documented in an early version of Duffee (2011): in the case of using filtering factors to predict five-year bond returns, the real sample R^2 (38.1%) is about 20.6% higher than the population R^2 (17.5%); and the gap reported by Duffee is around 19.5% (37.7% against 18.2%).[18] Despite this discrepancy, filtered factors are shown more informative in predicting excess returns than cross-sectional yields, in both the population and the finite-sample settings. A comparison between their performances in the real sample is presented in the next subsection.

[17] Another potential mismatch between the model and sample properties, as noted in Duffee (2011), is about the sample correlation between filtered month-t conditional expectations of excess returns and month-$t+1$ return shocks. According to Eq. (21), a highly positive correlation would elevate the sample variance of "true" returns, leading to a relatively low R^2, and vice versa. This issue does not arise in our sample, as the highest (absolute) correlation coefficient among all returns is less than 15%. Accordingly, the average correlation in our small-sample simulation ranges from 3% to 10%, depending on the bond maturity.

[18] In our study, the gap for smoothed factors is remarkably smaller than Duffee's estimate, as the short sample for real bonds does not offer substantial benefit by using future information.

3.6 Sample Properties of Alternative GDTSMs

Section 3.4 presents evidence, based on population properties of alternative GDTSMs, that the two- and three-factor GDTSMs underestimate the importance of the "hidden" component in risk premia, and thus imply a lower degree of return predictability. Given the substantial differences between the sample and population properties of the four-factor model, it is essential to look at the sample R^2 suggested by alternative models as well. The top panel of Figure 3 plots the R^2 for predictive regressions of annual excess returns on the state factor filtered using different models. The 95% confidence bounds are plotted only for the four-factor model, as represented by dashed lines. Complementary to the results in Table 4, the point estimate of the sample R^2 under the four-factor model is close to the upper bound of the confidence interval for five-year or longer maturity bonds. For shorter-maturity bonds, the sample R^2 lies near the middle of the confidence interval.

A brief glance at those solid lines reveals that a model with higher dimension is more useful in forecasting returns. The dominance of five-factor model is most clear with the four-year bond, with a difference in R^2 of at least 16% from three-factor and two-factor models. It also outperforms for almost all other maturities, except for the ten-year bond. If we rely on a principal component analysis to determine the model dimension and thus choose the two-factor model, a lot of information would be lost in our inference of term premia. Compared to the four-factor model, the loss in state variables' predictive power is greatest for those intermediate-term bonds. This pattern is consistent with results in Table 3, where the "FB_2"-implied population R^2 is 8% lower for the five-year bond and 3% lower for the ten-year bond. As expected, the two-dimensional state vector produces R^2 statistics subtly different from those based on principal components of bond yields.

Consistent with the results obtained in Section 3.4, we find that the two-factor and three-factor models have difficulty in fully uncovering information in hidden factors over the sample period. But all these models are constrained in the sense that expected excess returns are restricted to being driven by a single factor. To complete our analysis, we draw a comparison of R^2s between constrained and unconstrained four-factor models. As illustrated by the lower panel in Figure 3, the green line representing the unconstrained model exhibits the same shape across maturities as the blue line, but at a higher level. In other words, the

restriction (14) to some extent inhibited us from extracting more information on expected excess returns, though it leads to more reasonable Sharpe ratios.

To further assess the validity of the restriction imposed on term premia, we plot in the same figure the sample R^2s based on the annual return predictor, $\beta_m^{12'}\mathcal{P}_t$, implied by the constrained model,[19]

$$R^2 = \frac{\hat{\zeta}_m^2 \widehat{\text{Var}}(\hat{\beta}_m^{12'}\widehat{\mathcal{P}}_{t|t})}{\widehat{\text{Var}}\left(rx_{t+12}^m\right)},$$

where $\hat{\zeta}_m^2$ denotes the univariate regression coefficient. We find that the resulting cyan line generally runs below the blue line, but within the 95% confidence bounds for (22). Therefore, the difference in the predictive power between $\hat{\beta}_m^{12'}\widehat{\mathcal{P}}_{t|t}$ and $\widehat{\mathcal{P}}_{t|t}$ can be roughly considered statistically insignificant. When we take a closer look at the multivariate regression coefficients $\hat{\gamma}_m^{12}$, we find that they are largely proportional to, but substantially greater than, model-implied coefficients $\beta_m^{12'}$. Accordingly, the estimate of ζ_m^2 is greater than one for all maturities. Therefore, the constrained model is picking the right linear combination of state variables, but underestimating the total amount of predictability.

4 Extracting Hidden Factors

All point estimates in the last section point to an economically important component of bond risk premia that is hidden from the yield curve. However, there is nontrivial sampling uncertainty and model uncertainty in these estimates, judging from the wide confidence intervals and the gap between return predictability in the sample and the model-implied predictability. To ensure that the evidence for hidden factors is not spurious, we need to explicitly extract them from the sample and link them to other sources of information.

Results in Section 3.2 suggest a divergence between factors driving monthly excess returns and those driving annual excess returns. We confirm this finding in this section and thus look for a "monthly" hidden factor and an "annual" hidden factor. The next subsection presents our inference of the hidden component in monthly bond premia, which is facilitated by the model restriction imposed on the market price of risk. As to annual returns, we follow

[19] To take account of the difference in the number of predictors, all lines in the lower panel depict adjusted R^2s.

an early version of Duffee (2011) by investigating the characteristics of each filtered state factor. We will demonstrate in Section 4.2 that the fourth factor is closest to the definition of hidden factor – it makes little contribution to the cross section of yields but plays an important role in yield dynamics and annual excess returns.

4.1 The Hidden Factor Driving Monthly Excess Returns

The constrained four-factor model presented in Section 3.2 posits that the one-period expected excess returns on bond portfolios be driven by $\lambda_{1l}'\mathcal{P}_t$, a single line combination of the state factors. It follows that the hidden component of this risk-premium factor can be defined as the part unspanned by the first four PCs of observed yields as follows:

$$H_t \equiv \lambda_{1l}'\mathcal{P}_t - E(\lambda_{1l}'\mathcal{P}_t|\widetilde{\mathcal{P}}_t). \tag{23}$$

By construction, since $\widetilde{\mathcal{P}}_t$ summarizes almost all information on cross-sectional yields, H_t supposedly has a minimal effect on the contemporaneous term structure and thus is a "hidden factor." This is illustrated in Panel A of Figure 4, which plots the yield loadings on H_t as implied by the estimated model. The impact of a one-standard-deviation change in H_t is no more than 2 bps regardless of the bond maturities, too small to be distinguished from the noise η_t in (15).[20]

Despite the negligible effect on the cross section, the evidence in Table 3 implies that H_t should account for a substantial fraction of predictable variations in monthly excess returns. Panel B depicts the following projection of expected monthly excess returns onto H_t

$$E(rx_{t+1}^m|H_t) = -(m-1)B_{\mathcal{P},m-1}\lambda_1\text{Cov}(\mathcal{P}_t, H_t)\text{Var}(H_t)^{-1}H_t,$$

where the expressions for $\text{Cov}(\mathcal{P}_t, H_t)$ and $\text{Var}(H_t)$ are given in Duffee (2011). Note from the figure that the effect of a one-standard-deviation shock to H_t on the expected excess return ranges from 21 to 32 bps for bonds with maturities of two through ten years. For instance, one such negative shock can lower the expected excess return of a five-year bond to -13 bps

[20]Since the signs of latent factors are not important, for interpretation purpose we reverse the sign of H_t such that it positively co-varies with expected excess returns, as shown in Panel B of Figure 4.

if the expected value currently is assumed to be about 9 bps per month (its unconditional mean).

If H_t drives up expected excess returns while keeping the yield level constant, it has to lower the expectation of future short rates, according to the following yield curve decomposition:

$$y_t^m = \frac{1}{m}\sum_{i=0}^{m-1} E_t\left(y_{t+i}^1\right) + \frac{1}{m}\sum_{i=1}^{m-1} E_t\left(rx_{t+i}^{m-i+1}\right) \tag{24}$$

The impulse responses illustrated in Panel C of Figure 4 supports this implication. While a one-standard-deviation increase in H_t has little effect on the one-month rate at time zero, it induces a decline of about 8.5 bps three months later; however, half of the decline would die away in another five months and the short rate no longer exhibits much response one year after the original shock. Such quickly disappearing effect implies that H_t's predictive power for monthly excess returns may be limited to short forecast horizons. For instance, consider a five-year zero at time t. While a one-standard-deviation shock to H_t increases this bond's excess return over the next month ($rx_{t,t+1}^{60}$) by 22 bps (Panel B of Figure 4), the impact of this shock on the same bond (the projection of $rx_{t+12,t+13}^{48}$ on H_t) would shrink to merely 1.6 bps in one year.

The evidence in Panel C also suggest that H_t plays a much less important role in expected annual excess returns. Applying Eq. (24) to a two-year bond yields

$$\begin{aligned}
y_t^{24} &= \frac{1}{24}\sum_{i=0}^{23} E_t\left(y_{t+i}^1\right) + \frac{1}{24}\sum_{i=1}^{23} E_t\left(rx_{t+i}^{25-i}\right) \\
&= \frac{1}{24}\sum_{i=0}^{11} E_t\left(y_{t+i}^1\right) + \frac{1}{24}\sum_{i=12}^{23} E_t\left(E_{t+12}(y_{t+i}^1)\right) + \frac{1}{24}\sum_{i=1}^{12} E_t\left(rx_{t+i}^{25-i}\right) + \frac{1}{24}\sum_{i=13}^{23} E_t\left(rx_{t+i}^{25-i}\right) && \text{(25a)} \\
&= \frac{1}{2}y_t^{12} - \frac{1}{24}\sum_{i=1}^{11} E_t\left(rx_{t+i}^{13-i}\right) + \frac{1}{24}\sum_{i=1}^{12} E_t\left(rx_{t+i}^{25-i}\right) + \frac{1}{2}E_t\left(y_{t+12}^{12}\right) && \text{(25b)} \\
&= \frac{1}{2}y_t^{12} + \frac{1}{2}E_t\left(y_{t+12}^{12}\right) + \frac{1}{2}E_t\left(rx_{t+12}^{24}\right). && \text{(25c)}
\end{aligned}$$

Within the expectation component, expected short rates less than 12 months ahead constitute the term (the first term in (25a)) showing significant reactions to the hidden factor. In other words, H_t exerts very limited influence on second term, which turns into the expected one-year rate $\frac{1}{2}E_t\left(y_{t+12}^{12}\right)$ in (25b) and (25c) by combining the last term in (25a). Therefore, H_t's impact on the future one-year rate is expected be much less impressive than that on

one-month rate. Since H_t also has little effect on the current one-year rate, the annualized version of yield decomposition (25c) implies that the hidden component in monthly bond premia is not a prominent driver of annual excess returns.[21]

Impulse responses of the one-year yield to one-standard-deviation changes in H_t, shown in Panel D of Figure 4, reinforce the above finding. Note that the peak of the impulse response function has a magnitude of only 4 bps (in 3-4 months), much lower than the largest response of one-month rate to the same shock (more than 8.5 bps, as shown in Panel C). This contrast is more sharp considering that the model-implied unconditional mean of one-year rate is 6 basis points greater than that of one-month rate. Like the case of one-month rate, H_t also has almost no immediate effect on one-year rate (0.23 basis point), hence a lesser negative effect on the the expected one-year rate means a lesser positive effect on the annual risk premia.

Overall, we find that the hidden component of the risk-premium factor drives monthly risk premia and expected one-month rates in opposite directions. On the other hand, it carries less important implications about the yield composition with a one-year holding period, as its effects on both annual risk premia and expected one-year rates are much weaker. This finding implies a significant departure of main determinants of short-run term premia from those driving expected returns at longer horizons. Indeed, the single risk-premium factor $\lambda_{1l}'\mathcal{P}_t$, which is supposed to summarize all information on expected monthly returns, explains a much smaller portion of than indicated by Figure 3, with the lowest R^2 of 9.2%.

4.2 The Fourth Factor Being An "Annual" Hidden Factor

Given the evidence in Section 4.1 that neither $\lambda_{1l}'\mathcal{P}_t$ nor its hidden component shows particularly strong predictive power for annual excess returns, the next question is whether there are hidden factors that are important in determining investors' expectations of future yields at a relatively long horizon. Following an early version of Duffee (2011), we explore the possibility that certain state factors behave exactly as a hidden factor.

[21] From a different angle, we can see that H_t's impact on the term premium component is mainly through the third term in (25a). But it is to a great extent offset by the second term in (25b). As a result, H_t's net effect on the combination of these two terms, which is the expected annual returns $1/2E_t(rx_{t+12}^{24})$ as represented by the last term in (25c), is more like a mixed bag.

We first look at the importance of each state factor in forecasting annual excess returns in the data sample. Specifically, given a n-year bond with n ranging from 2 to 10, we run a predictive regression of annual excess returns of the bond on each of the four filtered state factors separately; for every regression, the covariance matrix of the coefficient estimates is computed using either the robust Hansen-Hodrick approach or the Newey-West procedure. Panel A of Figure 5 plots the Hansen-Hodrick t-statistic (solid line) of these regressions, along with its 95% confidence bounds (dotted lines) for the finite-sample t-statistics, versus the bond maturity. Since using Newey-West standard errors leads to similar t-statistics (Panel B), we focus on Panel A only in the discussion that follows.

We make three observations from Panel A. First, the "level" factor shows strong predictive power for short-maturity bonds and, in particular, the t-statistic decreases in the bond maturity for bonds longer than five years and becomes statistically insignificant for the ten-year bond. Second, the "slope" factor shows a different pattern of the forecast power across maturities: It is not significant in forecasting returns on the two-year and three-year bonds, but sees its t-statistic generally increasing with the maturity. Third, while the "curvature" factor (the third factor) shows no significant forecast power for bond excess returns, the fourth factor does exhibit such forecast power, regardless of the bond maturities (the entire curve of the t-statistic for the fourth factor lies within the confidence bounds, indicating a statistical significance at the conventional 5% level). This result is expected: Given the evidence documented earlier on the importance of unspanned factors in driving expected excess returns, one of the higher order state factors must explain a great portion of predictable variance.

As such, together with the fact that the four factor has trivial effect on the yield curve, the evidence in Figure 5 suggests that the fourth factor is a hidden factor.

To gain more evidence on the above interpretation, we examine the model-implied properties of the fourth factor and see whether it drives risk premia and expected one-year rates in opposite directions. Panel E of Figure 4 illustrates the impulse responses of the one-year yield to a one-standard-deviation shock to $\mathcal{P}_{4,t}$ (again with its sign flipped). Despite its zero effect on the month-t yield, the shock leads to a 18-basis-point drop in month $t+10$. And the yield remains 10 bps below its mean, 20 months ahead. On the other hand, Panel F of

the figure shows that $\mathcal{P}_{4,t}$ has significant, positive impact on annual excess returns rx^m_{t+12} and that the resulting increase in expected excess returns ranges from 129 to 248 bps.

The evidence so far indicates that the fourth factor shows particular importance in determining investors' expectations of future yields. We next investigate its role in term structure modeling. Figure 2 shows that the third factor is curved at about the three-year maturity, with the heaviest loading on the three-one year spread. Instead, the fourth factor is curved at the five-year maturity (the median point among all maturities used in estimation); it loads most strongly on the two-five year spread, unlike the standard "curvature" factor that is usually centered at the short end of the yield curve. Theoretically, the additional curvature captured by the fourth factor permits greater flexibility in fitting the yield curve. But does this advantage translate into better modeling performance?

To answer this question, we examine the ability of the benchmark model (the four-factor GDTSM) and its two special cases (the two- and three-factor models) to predict real yields at different horizons. Specifically, for a given forecast horizon j (month) at time t, we focus on each model's root mean squared error (RMSE), where the pricing error for any given bond is defined as the time-$(t+j)$ model yield implied by the filtered time-t state vector $(\mathcal{P}_{t|t})$ minus the time-$(t+j)$ observed yield of the same bond. Recall that $\mathcal{P}_{t|t}$ available at month t can be used to construct forecasts of month $t+j$ bond yields.

Table 5 reports each GDTSM's RMSE for each of the bonds with maturities of one through ten years, for four different forecast horizons that include $j = 0$ (Panel A), 1 month (Panel B), 6 months (Panel C), and 12 months (Panel D). Note that when $j = 0$, the pricing error is also referred to as the fitting error. We make three observations from the table. First, the four-factor model clearly dominates the two- and three-factor models, regardless of the bond maturities and forecast horizons considered. This result is consistent with the implication of the model population properties summarized in Tables 3 and 4—namely, the four-factor model does better than a lower-dimensional model at predicting bond excess returns or equivalently, predicting future yields. Second, the four-factor model improves the performance especially for bonds with maturities of four through six years, likely due to the fact that the fourth factor is an additional curvature factor centered around the five-year maturity. The resulting convexity effect introduced into the middle section of the yield

curve also helps improve the model performance in the long end of the curve. Third, the improvement in forecast accuracy generally increases with the forecast horizon. For instance, the average of the ratios of the four-factor model's RMSE to the three-factor one's decreases from 94% to 91% and further to 85% when the forecast horizon increases from one month to six and then twelve months.

4.3 Economic Interpretation of Hidden Factors

A main bugbear for studies on hidden factors is a direct link to fundamental macroeconomic forces (Duffee, 2011; Chernov and Mueller, 2012). This subsection aims to provide an economic interpretation of hidden factors extracted in the real bond market. To this end, we first show that these factors truly capture investor's expectations of future short rates. We then explore the relation between hidden factors and various macroeconomic instruments. We find that while the "monthly" hidden factor H_t covaries with aggregate economic activities, the "annual" hidden factor is closely related with the housing sector. Lastly, we demonstrate that neither hidden factor is a "liquidity" factor.

4.3.1 Evidence from Surveys

Consider $\hat{E}_t(\Delta y^3_{t+j})$, the expected change in the three-month real rate constructed using Eq. (1) for a given forecast horizon of j (month), where $j = 1, \ldots, 12$. To see how much of the expected change can be explained by the two hidden factors (H_t and $\mathcal{P}_{4,t}$), we regress $\hat{E}_t(\Delta y^3_{t+j})$ on filtered estimates of each hidden factors separately, and report the regression results in Panel A of Table 6. Note that both the hidden factors are normalized to have a unity standard deviation over the sample period 2004Q1–2012Q4, slight shorter than the one used to estimate the term structure model due to the availability of survey data. Also, as the constructed quarterly forecasts are serially correlated, t-statistics shown (in parentheses) are calculated using Newey-West standard errors for four lags of moving average residuals.

The slope coefficient on the "monthly" hidden factor (H_t) is significantly negative, regardless of the forecast horizons considered. A one-standard-deviation increase in H_t corresponds to an expected drop of about 10 bps in the three-month rate. Both the sign and magnitude of coefficients are consistent with the model-implied impulse responses (which are similar to

those plotted in Panel C of Figure 4). However, the pattern of variation in the magnitude of this coefficient across forecast horizons is different from the model's implication: instead of quickly dying away after the first quarter, the magnitude of the coefficient gradually declines (from around 10.6 for $j = 1$ to 5.8 for $j = 12$) and remains significant even twelve months ahead. The regression R^2 decreases from 0.25 to 0.10 as the forecast horizon (j) increases from one to 12 months, indicating that H_t has significant explanatory power for $\hat{E}_t(\Delta y^3_{t+j})\,\forall j$.

Like the coefficient on H_t, the coefficient on $\mathcal{P}_{4,t}$ (the "annual" hidden factor) is also significantly negative (consistent with the model due to the reverse sign of $\mathcal{P}_{4,t}$), regardless of the forecast horizons considered. However, the magnitudes of the latter coefficient and its t-value are notably lower than those for the former coefficient except for $j = 12$. This is not surprising as $\mathcal{P}_{4,t}$, supposed to have more significant effects on expected one-year yields, draws less strong response of the three-month yields. The pattern of variation in the magnitude of the coefficient on $\mathcal{P}_{4,t}$ across forecast horizons is in line with the model-implied properties of $\mathcal{P}_{4,t}$: the magnitude is does not decline much as the forecast horizon increases and the response to $\mathcal{P}_{4,t}$ is more persistent over long horizons,

Overall, the regression results provide evidence that the filtered hidden factors are not spurious and capture investor's expectations of future short rates.

4.3.2 Hidden Factors and Macroeconomic Variables

Given the evidence presented in Section 4.3.1, a natural question then is through what economic mechanisms the hidden factors affect expected real yields. To answer this question, we consider economic variables that may have significant explanatory power for variations in these factors.

We focus on two measures of real economic activity first. One is the three-month moving average of the Chicago Fed National Activity Index (CFNAI-MA3), a weighted average of 85 economic indicators that is often used as a proxy for the risk associated with real economic conditions (e.g., Joslin et al. 2014).[22] The other is the first PC of 131 macroeconomic

[22]CFNAI-MA3 is designed to gauge overall economic activity and related inflationary pressure, and similar to the index of economic activity developed by Stock and Watson (1999). Generally, a positive CFNAI-MA3 indicates that the national economy is expanding at an above-historical-average growth rate, and vice versa.

variables used in Ludvigson and Ng (2008).[23] This measure, denoted by $\hat{f}_1$, is shown by Duffee (2011) to have significant explanatory power for his *nominal* hidden factor.

The first two columns in Panel B of Table 6 report the results from univariate regressions on CFNAI-MA3 and $\hat{f}_1$, respectively, for each of the two hidden factors. We can see that both CFNAI-MA3 and $\hat{f}_1$ are negatively associated with H_t (the "monthly" hidden factor) and have significant explanatory power for this factor. More specifically, CFNAI-MA3 and $\hat{f}_1$ have respectively an R^2 of 8.9% and 31%. This difference in R^2 is mainly due to the recent financial crisis, during which H_t remained stable in Q1 and Q2 2008 but plumbed in Q3, mirroring the variation in the Industrial Production Index (IP). This macro shock is captured by $\hat{f}_1$ that loads heavily on IP and related indices, but less so by CFNAI-MA3 whose moving average feature smoothes out the macro shock. On the other hand, neither CFNAI-MA3 nor $\hat{f}_1$ shows any explanatory power for $\mathcal{P}_{4,t}$, the hidden factor driving expected annual returns.

Next we consider economic variables beyond those of real activity. Specifically, we divide the 131 series used to construct $\hat{f}_1$ into the following eight different groups as defined in Ludvigson and Ng (2011): (i) output, (ii) labor market, (iii) housing sector, (iv) orders and inventories, (v) money and credit, (vi) nominal bond and FX, (vii) prices or price indices, and (viii) stock market. Then, for each of these groups we estimate a dynamic factor model to construct a single "group" factor, denoted $\hat{g}_i$, $i = 1, \ldots, 8$, and examine its explanatory power for the hidden factors.

Columns 3 through 7 in Panel B report the results for those five group factors found to have significant explanatory power for either of the two hidden factors. Among these five group factors, the output ($\hat{g}_1$), labor market ($\hat{g}_2$), nominal bond market and FX ($\hat{g}_6$), and stock market ($\hat{g}_8$) factors each negatively co-varies with H_t with an R^2 of 9%, 9.3%, 10.3%, and 11.6%, respectively. The housing market factor ($\hat{g}_3$) is found to have significant explanatory power for $\mathcal{P}_{4,t}$ (the "annual" hidden factor) with a t-value of -9.26 and an R^2 of 27.5%.[24] Untabulated results indicate that the eight group factors jointly explain 44%

[23]The original data set consists of 132 time series. However, the variable "Employee Hours In Nonagricultural Establishments" has been discontinued.

[24]This group factor is constructed from 10 different measures of aggregate housing starts and housing authorization and may be a proxy for consumers' expenditure share on housing. The latter's effect on risk premia is theorized by Piazzesi et al. (2007). Huang and Shi (2010) find that a different factor constructed using the same 10 series possesses significant forecasting power for excess returns on nominal Treasury bonds.

of variation in H_t and that not surprisingly, the inflation factor (g_7) is insignificant in both univariate and multivariate regressions.

To summarize, Panel B provides evidence that hidden factors in the real bond market have clear economic interpretations. In particular, while the hidden component in monthly real risk premia strongly co-varies with general measures of macroeconomic conditions, the "annual" hidden factor is strongly correlated with the dynamic factor retrieved from the group of housing market variables.

4.3.3 Hidden Factors and Illiquidity

This subsection addresses the concern that our estimates of hidden factors may simply capture the liquidity premium in the TIPS market rather than the pure risk compensation for receiving long-horizon cash flows. In particular, even though our sample does not cover the early years of the TIPS market that are subject to poor liquidity, it does include the recent financial crisis during which a flight-to-liquidity occurred. As such, we examine whether these hidden factors are related to standard measures of the liquidity conditions in the TIPS market.

We consider the following seven liquidity measures commonly used in the literature:[25] (i) the TIPS transaction volume relative to that of nominal Treasury notes and bonds; (ii) the bid-ask spread of ten-year TIPS; (iii) the average TIPS curve fitting errors based on the Svensson model; (iv) the difference between the asset swap spread (ASW) for TIPS and that for (off-the-run) nominal Treasuries; (v) the spread between the ten-year inflation swap rate and the TIPS breakeven rate; (vi) the spread between the on-the-run and off-the-run ten-year nominal Treasury yields; (vii) the spread between the 20-year Resolution Funding Corporation (Refcorp) STRIPS and Treasury STRIPS. Among these liquidity proxies, the first three represent the "quantities" of liquidity risk in the TIPS market, the next two supposedly capture the risk premium for taking such liquidity risk,[26] and the last two are indicators of liquidity premia in the general Treasury market.

[25] Studies using at least one of these measures include DAmico et al. (2012), Roush (2008), Campbell et al. (2009), Gürkaynak et al. (2010), and Pflueger and Viceira (2010).

[26] These two variables are not available over the entire sample period. We have data on TIPS asset swap spreads (from Barclays Capital) available since July 2007 and the ICAP inflation swap data since April 2004.

The top panel of Table 7 shows the evidence on the performance of the seven liquidity proxies in explaining H_t. Only "Fitting error" and "Refcorp spread" are marginally significant in univariate regressions, but the significance disappears once regressions are augmented with the other proxies. The spread between on- and off-the-run nominal yields appears the only significant variable in the multivariate regression, albeit with an unexpected negative sign, where the seven measures jointly explain 12.1% of variation in H_t. This R^2 level is well below the R^2 generated by a single macro factor $\hat{f}_1$.

Similarly results obtain when $\mathcal{P}_{4,t}$ is considered, as indicated in the bottom panel of the table. For instance, while only "InfSwap-BEI spread" and "TIPS-nominal ASW spread" are significant respectively in the univariate and multivariate regressions, the signs of their coefficients are inconsistent with the economic intuition. Furthermore, the (adjusted) R^2 of the multivariate regression is only 14.3%, about half of what is obtained with the housing factor $\hat{g}_3$ (Panel B of Table 6).

To summarize, we find more convincing evidence that hidden factors reflect changes in the macroeconomic fundamental rather than the bond illiquidity.

5 Conclusion

This paper examines the predictability of returns in real bonds, using information from the TIPS market. Specifically, we first construct unsmoothed real yields of zero-coupon bonds with maturities of one through ten years using TIPS data. We then conduct our analysis within the framework of Gaussian dynamic term structure models (GDTSMs). Our main finding is that a significant portion of variations in real bond risk premia are captured by predictors that hidden from the cross-section of real yields. Importantly, these so-called hidden factors can be extracted from unsmoothed real yield curves only as the process of smoothing yield curves artificially erases hidden factors. We find evidence that our extracted hidden factors are linked to measures of real activity and housing factors.

Specifically, we estimate a four-factor GDTSM with unsmoothed real yield curves and infer hidden factors using the Kalman filter. We find evidence of two hidden factors, one for monthly real bond excess returns and one for yearly excess returns. Based on the model's

point estimates, about 40% of variations in monthly real risk premia are attributable to the "monthly" hidden factor. In the case of annual excess returns, depending the bond maturity, the R^2 of predictive regressions of such returns improves by 5 to 20% when the cross-section of yields is replaced by the filtered state vector as the return predictor. Furthermore, we find that while the "monthly" hidden factor significantly co-varies with general measures of real activity, the "annual" hidden factor is strongly correlated with the housing factor. Additionally, bond illiquidity cannot explain variations in the hidden factors.

In conclusion, this study provides robust evidence that information other than what is contained in the cross section of real yields is very useful for predicting excess returns on real bonds and that the use of unsmoothed real yields is necessary for the construction of such (hidden) return predictors.

A Construction of Real Zero-Yield Curves

A.1 Estimation of TIPS Indexation Lag Correction

We consider zero-coupon bonds with a unity par only in this subsection. Let y_t^m, $y_t^{TIPS,m}$, and $p_t^{\$}(m)$ denote respectively the time-t real yield, TIPS yield, and log nominal price with maturity m. Also, let $f_t^{\$}(m \to h)$ be the time-t nominal forward rate over the period $[m, h]$.

To obtain an estimate of real yields, we begin with the following equation that links them with nominal and TIPS yields:

$$y_t^m = \frac{h}{m} y_t^{TIPS,h} - \frac{l}{m} f_t^{\$}(m \to h) + \frac{1}{m} \gamma_t(m) \tag{26}$$

where $\ell > 0$ is the TIPS indexation lag and $h = m + \ell$; $\gamma_t(m) \equiv \text{cov}_t(p_{t+m}^{\$}(\ell), i_{t+m} - i_t)$, denoting the conditional covariance between the future inflation and nominal bond log prices. Note that the last term on the right-hand side (RHS) of Eq. (26) represents an adjustment made to the TIPS yield due to the indexation lag. As shown in Evans (1998), Eq. (26) provides an arbitrage-free and relatively model-free approach to estimating real yields by correcting for the indexation lag.

We estimate the correction $\gamma_t(m)$ using an approach in the spirit of Evans (1998). The idea is to include all relevant variables in a vector autoregression (VAR) system to generate multi-period forecasts of $\gamma_t(m)$. Consider the following first-order VAR model:

$$x_{t+1} = A x_t + e_{t+1}, \tag{27}$$

where $x_t' = [\Delta i_t, p_t^{\$}(\ell), z_t]$, z_t is a vector of conditional variables that may affect $\gamma_t(m)$, and A the coefficient matrix. Once A and innovation variances $V(e_{t+j}|x_t)$, $\forall j > 0$ are known, $\gamma_t(m)$ can be calculated using the following formula:

$$\gamma_t(m) = \iota_1' \left[\sum_{i=1}^{m} A^{\tau - i} \left(\sum_{j=1}^{i} A^{i-j} V(e_{t+j}|x_t) A^{i-j,} \right) \right] \iota_2, \tag{28}$$

where ι_1 and ι_2 are the selection vector such that $\Delta i_t = \iota_1' x_t$ and $p_t^{\$}(\ell) = \iota_2' x_t$.

Note that given the purpose of this study, we need to estimate $\gamma_t(m)$ recursively, via Eq. (28), using information up to time t only. Doing a one-time estimation (of $\gamma_t(m)$) with the whole sample (the conventional approach in the literature) is not suitable here as it involves the use of future information. To have sufficient observations to run the recursive estimation, we extend our sample period back to January 1997 (when TIPS were initiated) in this exercise. Since January 2004, $\gamma_t(m)$ are estimated with an expanding window (using all data available up to month t).

The first four columns of Table 8 present annualized estimates of $\gamma_t(m)$ based on the first- to fourth-order VAR systems, where the estimates in bps are obtained by multiplying monthly estimates of $\gamma_t(m)$ by $-120000/m$, for $m = 1, \ldots, 10$ years. For comparison, the $\gamma_t(m)$ estimates obtained non-recursively are reported in the last four columns. Note that the recursive estimation results in a smaller magnitude of $\gamma_t(m)$, regardless of the VAR models and bond maturities considered. Nonetheless, the difference is not substantial as $\gamma_t(m)$ itself contributes merely 0.5-1.7 bps to the annualized real yields.[27] Such a small correction, especially for long-term yields, is due to the fact that the covariance term in Eq. (26) contributes little to the correction for indexation lag. The (negative) difference between real and TIPS yields is mainly because the nominal term structure is more upward-sloping than the real term structure. As a result, the second term on the RHS of Eq. (26) tends to overweigh the difference between $(h/m)y_t^{TIPS,h}$ and $y_t^{TIPS,m}$.

Among the four sets of estimates of $\gamma_t(m)$ obtained recursively, we choose those based on the VAR(2) model (ranked the first under the Hannan-Quinn and Schwarz information criteria) in the construction of real yields that are used in the empirical analysis of this study. Note that the size of the covariance term under this VAR model generally increases with the maturity.

[27] These estimates of $\gamma_t(m)$ are consistent with those obtained by Grishchenko and Huang (2013) using TIPS data for the 2000–2007 period, and also with Evans (1998)'s based on the UK's data.

A.2 Estimation of the Three-Month Real Rate

This section describes how to construct the three-month real rate (y_t^3) using the Campbell and Shiller (1996) VAR-based approach, under the assumption of no inflation premium at short horizons.

Variables considered in the Campbell and Shiller (1996) study include the ex post real return on a 3-month nominal Treasury bill, the nominal bill yield, and the once-lagged annual inflation rate. Our VAR system includes all these variables as well as long-maturity nominal bond yields. We solve the VAR forward to generate forecasts of future three-month real interest rates. Note that this estimation of real short rates does not presume that the expectation hypothesis describes the real term structure, as we do not aggregate the forecasts to generate implied long-term real bond yields. Rather, this method is implicitly based on the assumption of zero inflation premium—whose magnitude is shown to be small for shortest maturity bonds (Buraschi and Jiltsov, 2005).

Figure 6 displays the three-month-ahead VAR forecast, which serves as our fitted yield on a hypothetical three-month real bill. For comparison, we also plot realized real returns on nominal 3-month T-bills, the variable to be forecasted. We note that the real rate appears to be more volatile when calculated ex post. It indicates that completely indexed Treasury bills, if they exist, do offer a considerable degree of inflation protection. The overall level of short real rates is fairly low over our sample period, averaged at -76 basis points. As expected, it is overwhelmingly negative after the financial crisis.

B Restrictions on Risk Premia Dynamics

In a four-factor GDTSM, there are twenty free parameters governing the $\mathbb{P}$-distribution of risk factors. Faced with such a large number of free parameters, we explore the best set of restrictions on Λ_t to make the model more parsimonious and economically plausible. In this appendix, we aim to gain some basic understanding of how PC risks are reflected in excess returns on real bonds.

Our analysis on the source of real bond premia is inspired by Duffee (2010), who finds that when the model-implied conditional Sharpe ratios are constrained to reasonable values,

risk premia are earned only as compensation for exposure to "level" and "slope" shocks. We firstly investigate whether real term structure models suffer the same problem on Sharpe ratios. Panel A of Table 9 reports sample means, standard deviations and Sharpe ratios of monthly excess returns on TIPS. Theoretically, real (inflation-adjusted) returns should be the variable of interest. However, while one-month (four-week) nominal T-bills are issued in the US, there exists no equivalent short-term instrument with fixed real payoffs. As discussed in Campbell and Shiller (1996), the ex post real bill return is a poor proxy for the ex ante real rate, even for very short maturities. This issue is tackle in two ways: on one hand, we follow Campbell and Shiller (1996) by constructing an ex-ante measure of one-month real interest rate, based on a VAR system; on the other hand, we also consider the nominal returns on TIPS.

Consistent with the finding in previous studies on nominal Treasuries (Fama and French, 1993; Campbell and Viceira, 2001; Duffee, 2010), we find that unconditional Sharpe ratios are highest for short-maturity TIPS. This result is robust to the use of real returns or nominal returns. Nevertheless, the relation between maturity and Sharpe ratio is not monotonic. While the five-year to ten-year portfolio does have a much lower Sharpe ratio compared to the one-year to five-year portfolio, its numbers are also a little lower than the bucket with longer maturities. This slight "U"-shape relation is undiscovered in the literature because previous studies focus on maturities no longer than ten years. Given the magnitude of Sharpe ratios reported in the table, we set 0.18–0.23 as the benchmark for unconditional maximum Sharpe ratios in our model diagnosis.

Panel B exhibits Sharpe ratios implied from different models. Unconditional maximum ratios are reported only for simple returns, assuming a complete bond market. For the unconstrained model, the number is 0.759, about three times the benchmark range. This model also produces conditional maximum Sharpe ratios averaged at 0.92–1.70, depending on whether the ratios are calculated using simple returns or log returns. The simple-return-based conditional Sharpe ratios are substantially higher, ranged to 22.67.

As the results indicate that the four-factor real model overfits the data, we impose a Sharpe ratio constraint in the model estimation. Following Duffee (2010), we specify that the sample mean of model-fitted conditional maximum Sharpe ratios $\hat{\theta}_t$ cannot be greater

than a scalar c

$$\overline{\hat{\theta}_t} = \sum_{t=1}^{T} \sqrt{(\lambda_0 + \lambda_1 \hat{\mathcal{P}}_t)' \Omega^{-1} (\lambda_0 + \lambda_1 \hat{\mathcal{P}}_t)} \leq c. \tag{29}$$

Starting at $c = 0.10$, we estimate the model 11 times; each time the constraint is tightened by 0.2 until it reaches $c = 0.60$. We find that when c is fixed at 0.30, the model-implied unconditional maximum ratio is aligned with the benchmark, between 0.18 and 0.23. To look at the properties of this constrained model, we decompose population means of excess (log) returns into compensations for exposure to PC risks. As discussed in Section 3.1, our principal component decomposition is based on the covariance matrix of shocks to yields. As such, it provides a consistent way for understanding which PC risk is priced. In this sense, it is slightly different from Duffee (2010)'s decomposition, which is based on the unconditional covariance of yield levels, but is consistent with Duffee (2011)'s approach.

The decomposition result is reported in Panel C, where all entries are scaled by the unconditional standard deviation of excess returns. In this way, they are naturally linked to the unconditional Sharpe ratios for each bond. Since factors here are rotated to PCs of yield shocks, instead of PCs of yield shocks, the effect of level factor on expected returns varies considerably with maturity. It ranges from 0.196 for one-year bonds to 0.057 for ten-year bonds. As expected, the effect of slope factor rises almost linearly with maturity. While investors also require compensation to face the risk that the intermediate-term yields increases relatively to long-term and short-term yields, its magnitude is very small. Overall, our results support Duffee (2010)'s and Joslin et al. (2014)'s finding that expected returns mainly move in response to changes in level and slope factors.

Next, we state investigating whether some constraints should be imposed on the dimensionality of bond risk premia. If we take the earlier result that there is little compensation earned for exposure to shocks to the third and fourth factors, the dimension of risk-premium vector cannot exceed two. We firstly form a factor decomposition of model-implied expected returns by eigenvalue-decomposing their covariance matrix. The result on monthly excess returns, as shown by the first two numbers in Panel A of Table 10, supports a specification with two risk-premium variables. In particular, the second "monthly" factor explains 3.9% of the variance of expected returns. However, the last two entries suggest that annual excess returns seem captured by a single factor. As discussed in Sec. 3.2, we might tend to

downweight the monthly result relative to the annual result, since one-monthly real rates are not included in our model estimation. What makes the results more complicated is that the leading "monthly return" factor appears far from perfectly correlated to the "annual" factor. More than 62% of the predictable variation in annual excess returns is orthogonal to the first "monthly" principal component. When we use the "annual" factor to explain expected monthly returns, the amount is almost the same.

To further pin down the dimensionality of monthly excess returns, we conduct a likelihood-ratio test to examine the rank of λ_1. The test results are shown in Panel B. The test statistic lies in the neighborhood of 10% critical value. If the significance level is set to 5%, we cannot reject the null hypothesis that a one-dimensional state variable is sufficient to describe variations in Λ_t. If we adopt the 10% significance level, statistical significance is not attained.

C Effectiveness of Kalman Filter for Inferring the State Vector

This appendix describes the simulation exercise conducted to assess the effectiveness of Kalman filter in inferring the true state vector. We simulate 10,000 years of monthly yields from the estimated model with maturities ranging from one through ten years, and then apply the Kalman filter to the simulated data (treating the model parameters as known). In addition to standard filtered estimates $\widehat{\mathcal{P}}_{t|t}$, we also obtain smoothed estimates $\widehat{\mathcal{P}}_{t|T} = E(\mathcal{P}_t|\{\widetilde{Y_1}\ldots\widetilde{Y_T}\})$ using the Kalman smoother (Rauch, Striebel, and Tung 1965). It is known that while the (Kalman) filter performs a real-time inference, the smoother essentially makes use of the whole sample and thus is expected to produce more accurate estimates of state factors.

Column 1 (2) in Panel A of Table 11 reports the correlations between the true state factor and its filtered (smoothed) estimates, for each of the four factors. As expected, correlations with the filtered estimates are lower than those with the smoothed ones. Importantly, while correlations for the first and second factors are all higher than 95%, Kalman filter really makes difference in its accurate estimates of higher-order factors, as indicated by the correlations

of over 89% (94%) for the third (fourth) factors.[28] Interestingly, both Kalman filtering and smoothing do a better job in inferring the fourth factor than inferring the third factor.

Columns 3 through 5 in Panel A report the explanatory power of the yield curve for the true state, its filtered estimate, and its smoothed estimate, respectively, for each of the four factors. Note that the regression R^2 is almost 100% for the "level" factor and greater than 91% for the "slope" factor, regardless of the dependent variables used. On the other hand, the R^2 is only about 74% (77%) for the true third (fourth) factor. Being more closely related to observed yields, both filtered and smoothed estimates produce higher R^2s, ranging from 82% to about 86% for the third and fourth factors. Still the 18-24% unexplained portions of these estimates are clear evidence for the unspanned nature of the third and fourth factors.

Next, we investigate the accuracy of finite-sample estimates produced by the Kalman filter because in practice, econometricians neither observe thousands years of bond yields nor perfectly know the true model parameters. We follow a procedure similar to the one used to generate standard errors of parameter estimates reported in Table 2. In each simulation trial, after estimating model parameters, Kalman filter and smoother are then applied to get estimates of the state vector. We can then calculate the correlations between these filtered (or smoothed) estimates and the simulated, true state factors.

Column 1 (2) in Panel B reports the correlation coefficients averaged over 1,000 simulations for filtered (smoothed) estimates. Note that the magnitudes of these correlations are generally comparable to those of population-based correlations shown in Panel A. This result implies that when model parameters are unknown, econometricians can still effectively infer the state vector by estimating the model with a reasonable, eleven-year sample. Columns 3 and 4 in the panel suggest that in a small sample, a greater portion of higher-order PC risks are unspanned by bond yields.

[28] High correlations for the first two factors can be achieved even with a simple singular value decomposition of the covariance matrix, if the factor rotation is based on principal components of yields.

Table 1: Summary Statistics on Estimated TIPS and Real Term Structures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Maturity (yr)	TIPS Zero Yields						Real Zero Yields Fama-Bliss	
	Fama-Bliss		Svensson		GSW			
	Mean	Stdev	Mean	Stdev	Mean	Stdev	Mean	Stdev
1	0.212	1.758	0.224	1.720			−0.116	1.847
2	0.193	1.544	0.209	1.540	0.187	1.529	0.013	1.573
3	0.334	1.433	0.331	1.416	0.321	1.409	0.171	1.405
4	0.473	1.275	0.489	1.329	0.483	1.324	0.342	1.272
5	0.665	1.277	0.656	1.254	0.645	1.253	0.556	1.305
6	0.821	1.210	0.808	1.185	0.797	1.190	0.717	1.198
7	0.952	1.152	0.946	1.123	0.935	1.132	0.862	1.141
8	1.064	1.078	1.069	1.067	1.058	1.078	0.981	1.083
9	1.146	1.006	1.177	1.017	1.167	1.028	1.070	1.008
10	1.213	0.939	1.271	0.973	1.263	0.982	1.121	0.899

This table reports the means and standard deviations of zero yields estimated from quotes on individual TIPS trading on the last business day of the month from January 2004 to December 2014. Columns under the "Fama-Bliss" and "Svensson" labels are based on data taken from Thomson Reuters, while statistics in "GSW" columns are computed by directly using estimates from Gürkaynak, Sack, and Wright (2010). The last two columns present the results on real yields implied from the Fama-Bliss zero yields on TIPS. Statistics of the one-year yield are reported for the 2006-2014 subperiod. All reported values are annualized and in percentages.

Table 2: Maximum Likelihood Estimates of a Four-Factor GDTSM

Parameter vector	Factor (or PC) Number			
	1	2	3	4
$\Phi_{\mathcal{P}}^{\mathbb{Q}}$	0.858	0.366	1.890	1.669
	(0.040)	(0.010)	(0.080)	(0.086)
	0.022	0.864	−0.550	−0.546
	(0.002)	(0.026)	(0.028)	(0.020)
	−0.001	−0.033	0.799	0.017
	(0.000)	(0.003)	(0.034)	(0.001)
	−0.008	0.016	0.037	1.069
	(0.000)	(0.000)	(0.002)	(0.039)
$\text{diag}(\Omega_{\mathcal{P}}^{1/2})(\times 10^2)$	1.170	0.330	0.094	0.013
	(0.055)	(0.028)	(0.004)	(0.000)
$\lambda_0(\times 10^2)$	−5.578	−1.644	0.000	0.000
	(2.579)	(1.013)	(0.000)	(0.000)
λ_{1l}	−0.342	−0.238	−1.227	9.946
	(0.071)	(0.051)	(0.224)	(1.818)
$\delta_{1\mathcal{P}}$	0.538	−1.244	−6.369	−4.637
	(0.024)	(0.043)	(0.732)	(0.191)
$\delta_{0\mathcal{P}}(\times 10^2)$		0.412		
		(0.078)		
$\sigma_\eta(\times 10^2)$		0.022		
		(0.000)		

This table reports the estimate of the parameter vector of a constrained four-factor GDTSM, whose state vector $\mathcal{P}_t$ is normalized to the first four principal components (PCs) of *shocks* to the real yield curve. The short rate $r_t = \delta_{0\mathcal{P}} + \delta_{1\mathcal{P}}'\mathcal{P}_t$. The $\mathbb{Q}$- and $\mathbb{P}$-dynamics of $\mathcal{P}_t$ are specified respectively in Eqs. (8) and (10). This GDTSM is constrained in the sense that the risk premium is specified by Eq. (14) such that only the first two elements of λ_0 and the first row ($\lambda_{1\ell}$) of λ_1 are free parameters. The model is estimated with maximum likelihood and Kalman filter using month-end real yields on zero-coupon bonds with maturities of one through ten years from 2004 through 2014. These real yields are assumed to have iid measure error with a standard deviation of σ_η. Quantities in parentheses are standard errors from 1,000 Monte Carlo simulations, under the null hypothesis that the estimated model is true. Each simulated data sample is 132 monthly observations of ten bond yields, with the same maturities as used in model estimation.

Table 3: Model-Implied Population Moments of Excess Bond Returns

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Model	Monthly Returns					Annual Returns				
	Mean	True Var.	Conditional Variance			Mean	True Var.	Conditional Variance		
			full info	4 PCs	Ratio			full info	4 PCs	Ratio
Panel A: Five-year bonds										
FB_4	0.090	0.106	0.017	0.010	0.591	0.936	0.732	0.169	0.102	0.605
	[-0.07 0.29]	[0.08 0.13]			[0.39 0.76]	[-0.58 2.27]	[0.54 1.10]			[0.42 0.81]
SV_4	0.087	0.106	0.011	0.010	0.895	0.898	0.779	0.117	0.105	0.901
FB_3	0.122	0.094	0.010	0.009	0.904	1.065	0.622	0.100	0.081	0.817
FB_2	0.149	0.096	0.009	0.009	0.968	1.449	0.566	0.084	0.080	0.945
Panel B: Ten-year bonds										
FB_4	0.102	0.231	0.052	0.031	0.593	1.155	1.041	0.257	0.172	0.669
	[-0.05 0.33]	[0.19 0.30]			[0.40 0.74]	[-0.44 2.35]	[0.60 1.41]			[0.47 0.85]
SV_4	0.099	0.225	0.036	0.031	0.864	1.260	1.013	0.195	0.174	0.891
FB_3	0.127	0.228	0.031	0.029	0.932	1.820	0.803	0.163	0.153	0.942
FB_2	0.175	0.209	0.030	0.028	0.959	1.650	0.730	0.144	0.140	0.975

The constrained four-factor GDTSM is estimated with maximum likelihood and Kalman filter using month-end real yields from 2004 through 2014. Parameter estimates are used to derive population properties of excess returns on a five-year bond (Panel A) and on a ten-year bond (Panel B). Yield curve fitting models used (column 1) include the Fama and Bliss (1987) and Svensson (1995) models. Columns under the title "Monthly Returns" report numerical characteristics of one-month log returns obtained by borrowing at the one-month rate, buying a long-term bond, and selling it in one month. Columns under the title "Annual Returns" show one-year log returns on a five (ten)-year bond in excess of the log return on a one-year bond. Columns 3 and 8 report the conditional variances of *ex post* excess returns uncontaminated by measurement errors. The "Full Info" and "4 PCs" columns quantify the volatility of true conditional expected excess returns attributable to time-variation in the true state vector and to time-variation in the principal components of cross-sectional yields. All entries in the table are expressed in percentage points except for variance ratios.

Table 4: Accuracy of Excess Return Forecasts Based on the Kalman Filter

	Monthly Returns		Annual Returns	
	5-Year	10-Year	5-Year	10-Year
Panel A: Population R^2s				
$R^2(\mathcal{P}_t)$ - Full Information	0.142	0.182	0.226	0.234
$R^2(\widehat{\mathcal{P}}_{t\|t})$ - Optimal Filtering	0.108	0.137	0.171	0.176
$R^2(\widehat{\mathcal{P}}_{t\|T})$ - Optimal Smoothing	0.116	0.153	0.182	0.192
$R^2_{\text{ols}}(\widehat{\mathcal{P}}_{t\|t})$ - OLS with Filtered	0.098	0.101	0.175	0.177
$R^2_{\text{ols}}(\widehat{\mathcal{P}}_{t\|T})$ - OLS with Smoothed	0.133	0.128	0.286	0.247
Panel B: Finite-sample R^2s				
$\widehat{R}^2(\widehat{\mathcal{P}}_{t\|t})$ - Optimal Filtering	0.098	0.118	0.180	0.167
	[0.06 0.15]	[0.08 0.19]	[0.10 0.29]	[0.07 0.28]
$\widehat{R}^2(\widehat{\mathcal{P}}_{t\|T})$ - Optimal Smoothing	0.146	0.167	0.217	0.198
	[0.10 0.21]	[0.10 0.22]	[0.12 0.34]	[0.07 0.32]
$\widehat{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t\|t})$ - OLS with Filtered	0.047	0.044	0.290	0.241
			[0.05 0.38]	[0.04 0.35]
$\widehat{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t\|T})$ - OLS with Smoothed	0.052	0.048	0.288	0.230
			[0.05 0.38]	[0.04 0.34]
Panel C: R^2s in the sample				
$\widetilde{R}^2(\widehat{\mathcal{P}}_{t\|t})$ - Optimal Filtering	0.118	0.132	0.289	0.252
$\widetilde{R}^2(\widehat{\mathcal{P}}_{t\|T})$ - Optimal Smoothing	0.144	0.178	0.331	0.279
$\widetilde{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t\|t})$ - OLS with Filtered			0.381	0.324
$\widetilde{R}^2_{\text{ols}}(\widehat{\mathcal{P}}_{t\|T})$ - OLS with Smoothed			0.383	0.335

This table shows the accuracy of return forecasts conditioning on factor estimated with the Kalman filter. Rows labeled "Optimal Filtering" and "Optimal Smoothing" present the results based on model-implied expectations about future excess returns based on filtered and smoothed states. Rows labeled "OLS with Filtered" and "OLS with Smoothed" report the R^2s from predictive regressions of realized excess returns on estimated state factors. Population results in Panel A are based on a 10,000-year simulation of state factors and bond yields, where the "true" model parameters are the ones tabulated in Table 2 and assumed known by econometricians. Results in Panel B are obtained by simulating 1,000 samples with the same length as the real sample. Maximum likelihood estimation and Kalman filter are performed for each sample to obtain the finite-sample R^2. The R^2 ratios shown in each row are the average over the 1,000 samples, with the 951th and 50th largest ones reported in brackets. Panel C presents the same type of ratios as in Panel B, which are obtained from the 2004-2014 data sample.

Table 5: Fitting and Forecasting Performance of Term Structure Models

Model	1-year	2-year	3-year	4-year	5-year	6-year	7-year	8-year	9-year	10-year
Panel A: Cross-sectional										
FB_4	3.12	7.85	7.03	12.29	7.47	5.69	5.30	4.64	3.73	6.14
FB_3	3.73	8.12	6.88	19.36	8.56	7.15	6.65	5.02	3.90	9.84
FB_2	3.86	9.54	8.09	20.24	8.97	7.66	7.08	5.09	4.18	10.35
Panel B: One month ahead										
FB_4	60.20	44.47	36.49	32.77	30.43	28.60	28.17	25.68	23.54	22.81
FB_3	61.91	46.92	39.43	37.10	30.96	29.41	29.23	27.07	25.49	25.73
FB_2	63.93	48.38	40.77	38.05	32.10	30.49	30.26	28.12	26.30	26.36
Panel C: Six months ahead										
FB_4	112.10	82.93	72.46	64.36	65.89	60.04	58.31	54.32	51.14	49.32
FB_3	126.34	97.46	86.15	69.44	72.58	65.79	63.14	58.99	54.40	50.59
FB_2	129.36	99.60	88.25	71.90	74.94	68.53	66.09	62.14	57.56	53.79
Panel D: Twelve months ahead										
FB_4	140.59	106.83	93.12	74.87	80.98	76.40	73.60	69.06	64.55	59.99
FB_3	157.54	126.45	113.32	92.60	97.61	91.83	87.53	81.49	74.58	68.06
FB_2	160.38	128.61	115.73	96.73	101.44	96.06	92.15	86.37	79.64	73.17

This table illustrates the pricing performance of three specifications of the constrained GDTSM given in Eqs. (8), (10), and (14): the two-, three-, and four-factor GDTSMs (denoted respectively FB_2, FB_3, and FB_4). All three models are estimated with maximum likelihood and Kalman filter, using monthly real bond yields on zero-coupon bonds with maturities of one through ten years from January 2004 through December 2014. Model performance is measured by the percentage change in root mean squared forecast errors (RMSE), which are calculated using filtered estimates of the state vector. The performance is examined for four different forecast horizons, including zero (Panel A), one (Panel B), six (Panel C), and 12 months (Panel D), for each bond maturity. The RMSE is reported in basis points of annualized yields.

Table 6: Linking Hidden Factors to the Macroeconomy

Panel A: Regressions of expected changes in the short rate on individual hidden factors

Hidden Factor	Months ahead (j)											
	1	2	3	4	5	6	7	8	9	10	11	12
H_t	-10.63	-10.63	-9.30	-10.37	-10.46	-9.12	-9.82	-9.63	-8.06	-8.63	-7.82	-5.83
	(-4.10)	(-3.44)	(-3.43)	(-3.99)	(-3.45)	(-3.48)	(-4.00)	(-3.39)	(-3.24)	(-3.74)	(-3.12)	(-2.53)
R^2	0.255	0.225	0.217	0.236	0.217	0.208	0.217	0.190	0.172	0.179	0.140	0.099
$\mathcal{P}_{4,t}$	-6.62	-6.38	-6.76	-6.65	-6.33	-6.66	-6.54	-6.24	-6.48	-6.34	-5.85	-6.16
	(-2.73)	(-2.37)	(-2.77)	(-2.56)	(-2.30)	(-2.62)	(-2.55)	(-2.35)	(-2.61)	(-2.51)	(-2.44)	(-2.94)
R^2	0.085	0.083	0.087	0.083	0.081	0.085	0.083	0.081	0.084	0.082	0.080	0.084

Panel B: Univariate regressions of individual hidden factors on macroeconomic variables

Hidden Factor	Macroeconomic variables							R^2
	CFNAI-MA3	$\hat{f}_1$	$\hat{g}_1$	$\hat{g}_2$	$\hat{g}_3$	$\hat{g}_6$	$\hat{g}_8$	
H_t	-0.34							0.089
	(-2.32)							
H_t		-0.61						0.314
		(-3.52)						
H_t			-0.34					0.090
			(-3.80)					
H_t				-0.35				0.093
				(-3.84)				
H_t						-0.36		0.103
						(-4.05)		
H_t							-0.38	0.116
							(-4.83)	
$\mathcal{P}_{4,t}$	0.05							0.003
	(0.36)							
$\mathcal{P}_{4,t}$		0.01						0.000
		(0.12)						
$\mathcal{P}_{4,t}$					-0.48			0.275
					(-9.26)			

Panel A reports the results from regressions of survey forecast of three-month real rates onto estimated hidden factors H_t and $\mathcal{P}_{4,t}$. Monthly observations of expectations of future T-bill yields and inflation rates are from Blue Chip Financial Forecasts. Regressions are run at a quarterly frequency to ensure time-invariant forecast horizon. The estimate of contemporaneous three-month yield is subtracted from the forecasts to produce forecasted changes in the yield. t-statistics based on Newey-West standard errors with a lag truncation parameter of 4 are shown in parentheses. Panel B displays the explanatory power of macroeconomic instruments for hidden factors. CFNAI-MA3 denotes the three-month moving average of the Chicago Fed National Activity Index. $\{\hat{f}_i, i = 1, \ldots, 8\}$ denotes static factors extracted by Ludvigson and Ng (2008) from 131 measures of economic activity and $\{\hat{g}_i, i = 1, \ldots, 8\}$ dynamic factors from the same dataset (Ludvigson and Ng, 2011). The signs of both hidden factors are reversed.

Table 7: Regressions of Hidden Factors on Observable Liquidity Measures

Hidden Factor	Relative Volume	Bid-Ask Spread	Fitting Error	TIPS-Off ASW	InfSwap-BEI Spread	Off-On Spread	Refcorp Spread	R^2	N
H_t	−1.05							0.075	132
	(−1.18)								
H_t		−0.09						−0.007	132
		(−0.28)							
H_t			0.09					0.081	132
			(2.11)						
H_t				−0.46				0.028	90
				(−0.61)					
H_t					−1.02			0.037	129
					(−1.54)				
H_t						−1.11		−0.009	132
						(−0.22)			
H_t							0.68	0.038	132
							(1.98)		
H_t	−0.49	−0.05	0.01	−0.34	0.11	−4.00	0.39	0.121	90
	(−1.19)	(−0.41)	(0.36)	(−1.06)	(0.31)	(−2.29)	(0.67)		
$\mathcal{P}_{4,t}$	0.87							0.049	132
	(1.24)								
$\mathcal{P}_{4,t}$		0.66						0.080	132
		(1.81)							
$\mathcal{P}_{4,t}$			−0.07					0.049	132
			(−1.43)						
$\mathcal{P}_{4,t}$				−0.75				0.077	90
				(−1.39)					
$\mathcal{P}_{4,t}$					−1.25			0.067	129
					(−2.08)				
$\mathcal{P}_{4,t}$						6.98		0.005	132
						(0.79)			
$\mathcal{P}_{4,t}$							0.96	0.088	132
							(1.10)		
$\mathcal{P}_{4,t}$	0.19	0.04	−0.03	−0.70	−0.02	1.97	0.27	0.143	90
	(0.57)	(0.37)	(−0.72)	(−2.41)	(−0.06)	(0.61)	(1.12)		

This table reports the results from regressions of estimated hidden factors H_t and $\mathcal{P}_{4,t}$ onto proxies for liquidity in the TIPS market. Liquidity proxies, from left to right in the table, include (1) the log ratio of TIPS trading volumes to nominal securities volume, (2) 10-year TIPS bid-ask spreads, (3) the average TIPS curve fitting errors, (4) the difference between 10-year TIPS and On-the-Run par nominal ASW spreads, (5) the difference between the 10-year inflation swap rate and breakeven inflation rate, (6) the spread between ten-year on- and off-the-run nominal Treasury yields, and (7) the spread between 20-year Refcorp and Treasury strips t-statistics based on Newey-West standard errors with a lag truncation parameter of 12 are shown in parentheses. The column labeled "N" reports the number of observations. The signs of both hidden factors are reversed.

Table 8: Correction for the Three-Month Indexation Lag of TIPS

Maturity (yr)	Recursive Estimation				Non-Recursive Estimation			
	VAR(1)	VAR(2)	VAR(3)	VAR(4)	VAR(1)	VAR(2)	VAR(3)	VAR(4)
1	0.58	0.79	1.20	1.32	0.59	0.86	1.29	1.34
2	0.58	0.82	1.26	1.38	0.61	0.96	1.52	1.66
3	0.63	0.84	1.20	1.26	0.68	1.04	1.57	1.71
4	0.70	0.87	1.14	1.14	0.76	1.11	1.58	1.69
5	0.76	0.89	1.07	1.04	0.84	1.17	1.56	1.63
6	0.81	0.90	1.01	0.95	0.91	1.21	1.53	1.57
7	0.87	0.91	0.95	0.88	0.96	1.25	1.49	1.50
8	0.91	0.91	0.90	0.81	1.01	1.27	1.44	1.43
9	0.95	0.91	0.85	0.76	1.05	1.28	1.40	1.36
10	0.98	0.91	0.81	0.71	1.08	1.28	1.35	1.29

This table reports corrections for the indexation lag of TIPS. They are based on estimates of $\gamma_t(m)$, defined as $Cov_t(i_{t+m} - i_t, p^{\$}_{t+m}(3))$, where $i_{t+m} - i_t$ is the log change in the price level from time t to time $t+m$, and $p^{\$}_{t+m}(3)$ is the log of prices at time $t+m$ for a nominal bond maturing in 3 months. The estimate of $\gamma_t(m)$ is computed as

$$\gamma_t(m) = \iota_1' \left[\sum_{i=1}^{\tau} A^{m-i} \left(\sum_{j=1}^{i} A^{i-j} V(e_{t+j}|x_t) A^{i-j,} \right) \right] \iota_2,$$

where ι_i is the selection vector such that $\Delta i_t = \iota_1' x_t$ and $p_t^{\$(l)} = \iota_2' x_t$; A is the coefficient matrix of a VAR model for x_t defined in Eq. (27); $V(e_{t+j}|x_t)$, $\forall j > 0$ are innovation variances of x_t. VAR models with lags of 1 through 4 are used in the estimation. Estimates of $\gamma_t(m)$ are obtained both recursively with expanding windows (the first four columns) and non-recursively using the entire sample (the last four columns). The first four columns report the mean of all estimates from January 2004 to December 2014. Multiplying $\gamma_t(m)$ by $-120000/m$ produces the correction to the real annual yields in basis points.

Table 9: Sample and Model-Implied Sharpe Ratios

Panel A: TIPS unconditional Sharpe ratios

Portfolio	Nominal Returns			Real Returns		
	Mean	Std	Sharpe Ratio	Mean	Std	Sharpe Ratio
1-5yr	0.198	0.853	0.232	0.133	0.483	0.276
5-10yr	0.301	1.708	0.177	0.315	1.767	0.178
$>$ 10yr	0.385	1.869	0.206	0.393	2.113	0.186

Panel B: Model-implied Sharpe ratios

Model	Bond Market Unconditional Maximum Ratios	Fixed-Income Market Conditional Maximum Ratios			
		Population		Sample	
		Simple Returns	Log Returns	Simple Returns	Log Returns
Unconstrained	0.759	1.619	0.918	1.701	0.924
Constrained	0.208	0.316	0.296	0.313	0.302
Preferred	0.198	0.374	0.306	0.384	0.271

Panel C: Decomposition of unconditional Sharpe ratios

Bond Maturity	Factor			
	1	2	3	4
1yr	0.196	−0.102	0.025	0.002
2yr	0.162	−0.034	−0.011	−0.003
5yr	0.110	0.036	−0.004	0.001
10yr	0.057	0.062	0.014	−0.005

Panel A shows unconditional moments and Sharpe ratios of excess TIPS returns at a monthly horizon. The sample period spans from Jananury 1997 to December 2014. Nominal returns are computed with the principal amount adjusted for changes in CPI. Model-implied Sharpe ratios are listed in Panel B. The second column shows the unconditional maximum Sharpe ratio in a hypothetical bond market that contains 120 zero-coupon bonds with maturities from 1 to 120 months. The remaining columns reports sample and population means of conditional maximum Sharpe ratios. Sample means are based on state variable values filtered from the estimated model, and population means are calculated using Monte Carlo simulations. Panel C shows the components of unconditional Sharpe ratios corresponding to the average compensations investors require to face PC risks.

Table 10: The Dimensionality of Risk Premia

Panel A: Eigenvalue-decomposition of the covariance matrix of expected excess returns

Expected Excess Returns	Fraction Explained by "Monthly Factors" 1st	"Monthly Factors" 2nd	"Annual Factors" 1st	"Annual Factors" 2nd
Monthly	0.957	0.039	0.372	0.160
Annual	0.374	0.237	0.985	0.015

Panel B: Likelihood Ratio Tests

	H_0: $\text{rk}(\lambda_1) = 1$			H_0: Eq. (14)		
$\log L_a$	$\log L_0$	stat $\chi^2(9)$	p-value	$\log L_0$	stat $\chi^2(14)$	p-value
49.437	49.372	14.868	0.095	49.264	39.517	3.0×10^{-4}

Panel A shows the properties of model-implied principal components of the covariance matrix of conditional expected excess monthly / annual returns on a set of bonds. The maturities of bonds involved in the principal component decomposition range from 2 years to 10 years. For each decomposition, only the first two principal components are extracted. Panel B reports the results of likelihood ratio tests. The null hypothesis in the first test is that the one-period expected excess returns lie in a one-dimensional space. The second test examines the fourteen zero restrictions imposed in Eq. (14).

Table 11: Model-Implied Properties of State Factors Estimated with the Kalman Filter

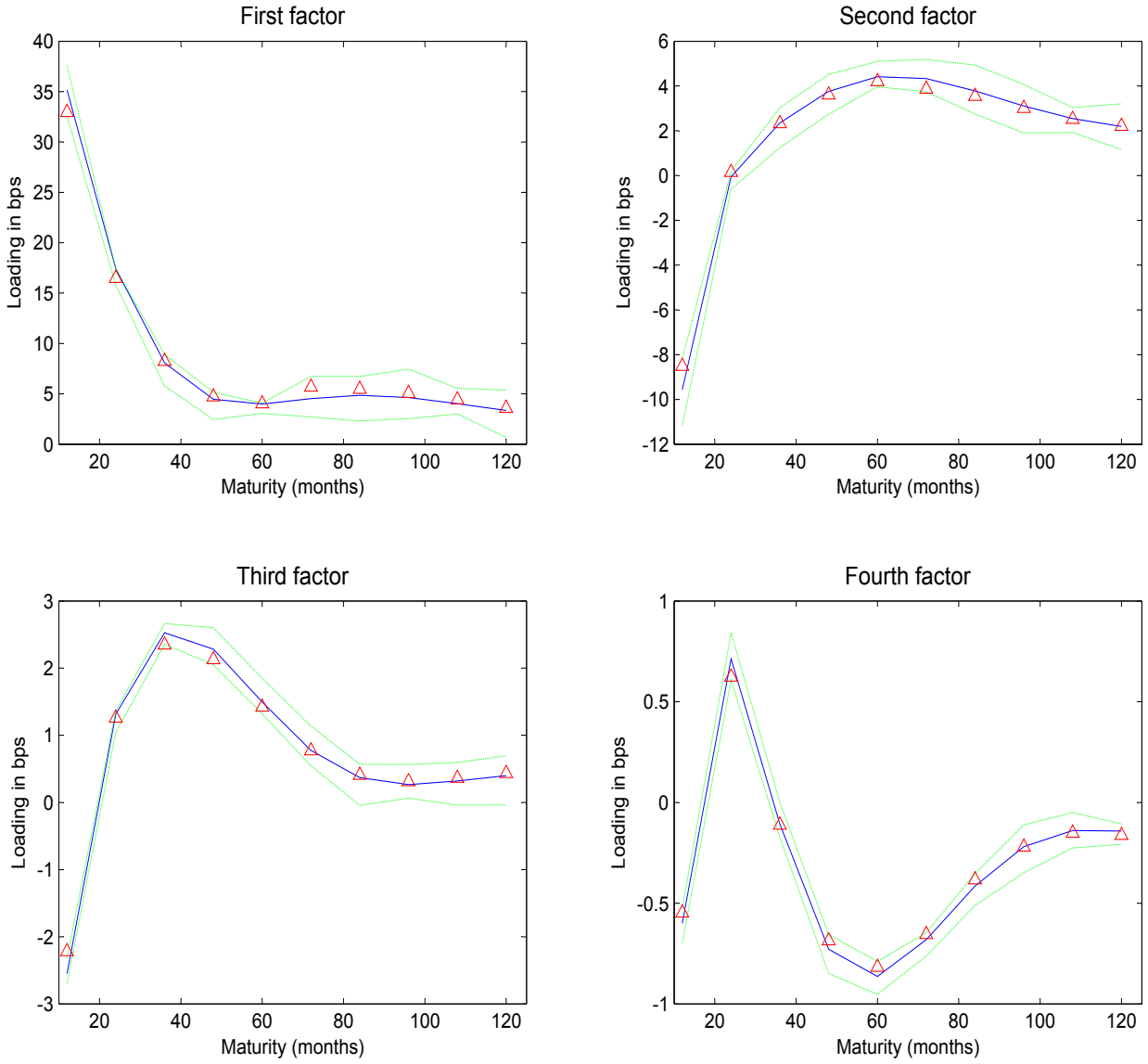
Factor No.	(1)	(2)	(3)	(4)	(5)
	Correlation of Individual True State Factors with		R^2 of Regressions on the Yield Curve Dependent variable used		
	Filtered Estimates	Smoothed Estimates	True Factor	Filtered Estimates	Smoothed Estimates
Panel A: Population properties of state factors					
1	0.993	0.993	0.995	0.999	0.999
2	0.958	0.974	0.928	0.942	0.911
3	0.891	0.920	0.766	0.858	0.830
4	0.943	0.967	0.738	0.834	0.821
Panel B: Finite-sample properties of state factors					
1	0.990	0.990	0.991	0.999	0.999
2	0.904	0.938	0.802	0.836	0.896
3	0.812	0.858	0.685	0.764	0.847
4	0.919	0.941	0.747	0.761	0.826

This table reports the population properties (Panel A) and finite-sample properties (Panel B) of filtered and smoothed estimates of state factors. Population properties of the Kalman filter are proxied by simulating 10,000 years of state factors and bond yields, where the "true" model is the model estimated with maximum likelihood. Finite-sample results are based on 1,000 samples generated by the estimated model. Each sample consists of 132 months of bond yields, with the maturities of one through ten years. Column 1 (2) shows the correlation of the true state vector with its filtered (smoothed) estimate. The last three columns report the R^2 of three sets of regressions on contemporaneous values of all ten noise-contaminated bond yields, where the dependent variable used is respectively the true state vector (column 3), its filtered estimate (column 4), and the smoothed estimate of the true state vector (column 5).

Figure 1: Real, Nominal and TIPS Yields of Ten-Year Zero-Coupon Bonds

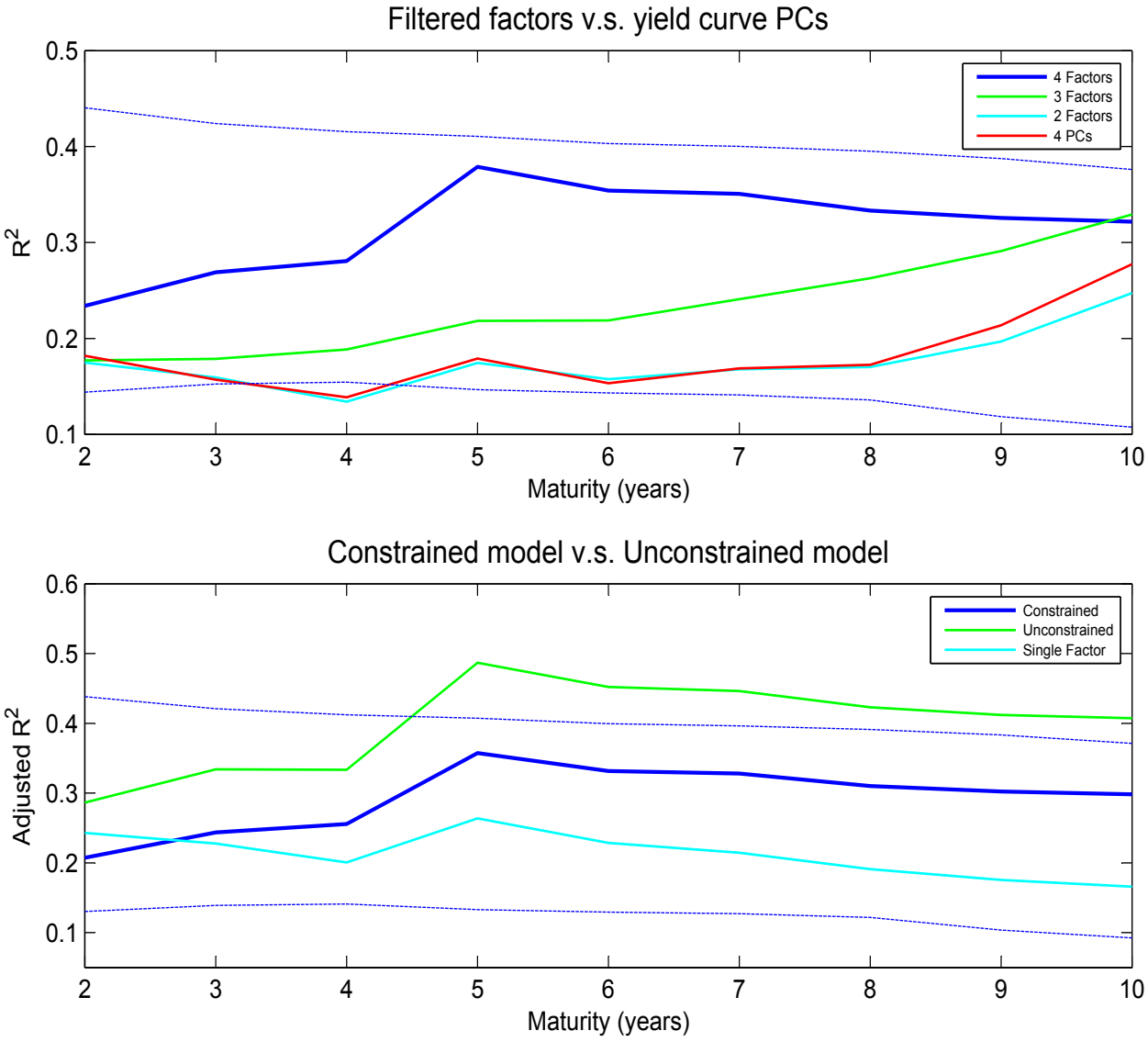
This figure plots the real, nominal, and TIPS yields for ten-year zero-coupon bonds over the period from January 2004 to December 2014. TIPS zero yields are constructed using the Fama-Bliss method. Real zero yields are TIPS zero yields corrected for the indexation lag.

Figure 2: Estimated Loadings of Annualized Real Yields on Real Term Structure Factors



The constrained four-factor GDTSM is estimated with maximum likelihood and Kalman filter, using month-end real yields from 2004 through 2014. The state vector is normalized to principal components of *shocks* to the yield curve. The triangles represent coefficients from regressions of the constructed real yields on filtered estimates of the state . The dotted lines are two-sided 95% confidence intervals calculated from 1,000 Monte Carlo simulations.

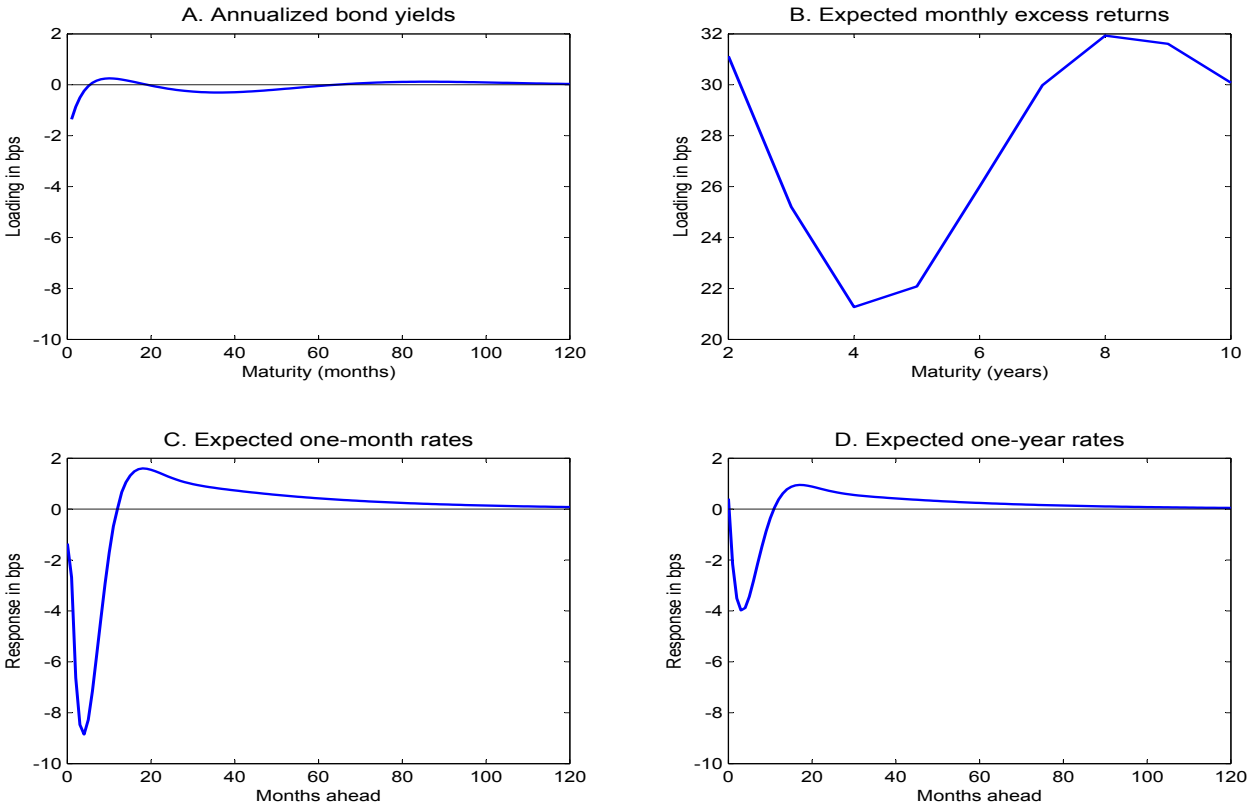
Figure 3: Predictability of Annual Excess Returns in the 2004-2014 Sample



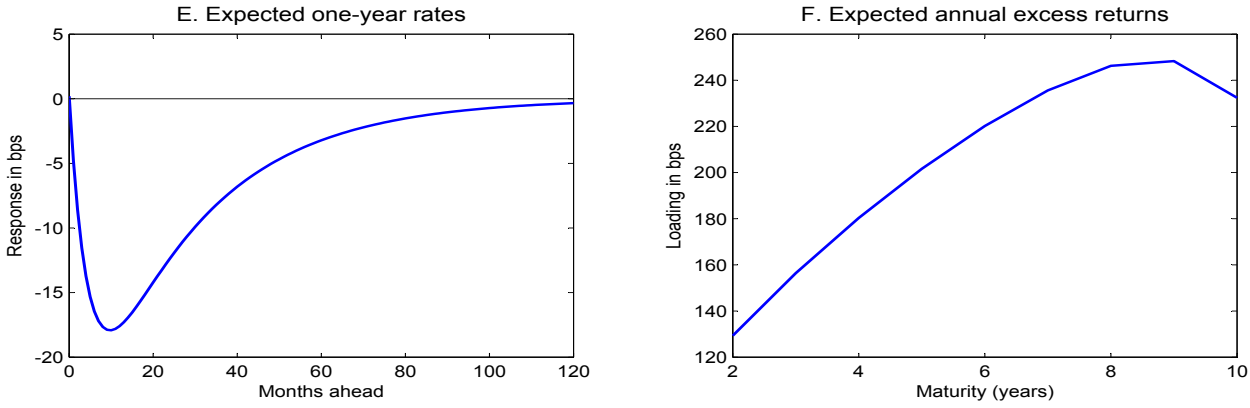
This figure exhibits the (adjusted) R^2s for predictive regressions of annual excess bond returns on risk factors extracted from estimated term structure models or from cross-sectional bond yields. Four different GDTSMs are estimated using monthly real bond yields from January 2004 through December 2014: a constrained four-factor model, a constrained three-factor model, a constrained two-factor model and an unconstrained four-factor model. The red line in the first panel shows the R^2s based on the first four principal components of the yield curve. The cyan line in the second panel shows the adjusted R^2s based on the single risk-premium factor, $\lambda_{1l}' P_t$, filtered from the constrained four-factor model. The blue dashed lines in both panels are two-sided 95% confidence intervals for regressions on four filtered factors extracted from the constrained model.

Figure 4: Model-Implied Effects of Hidden Factors

Effects of the "monthly" hidden factor H_t



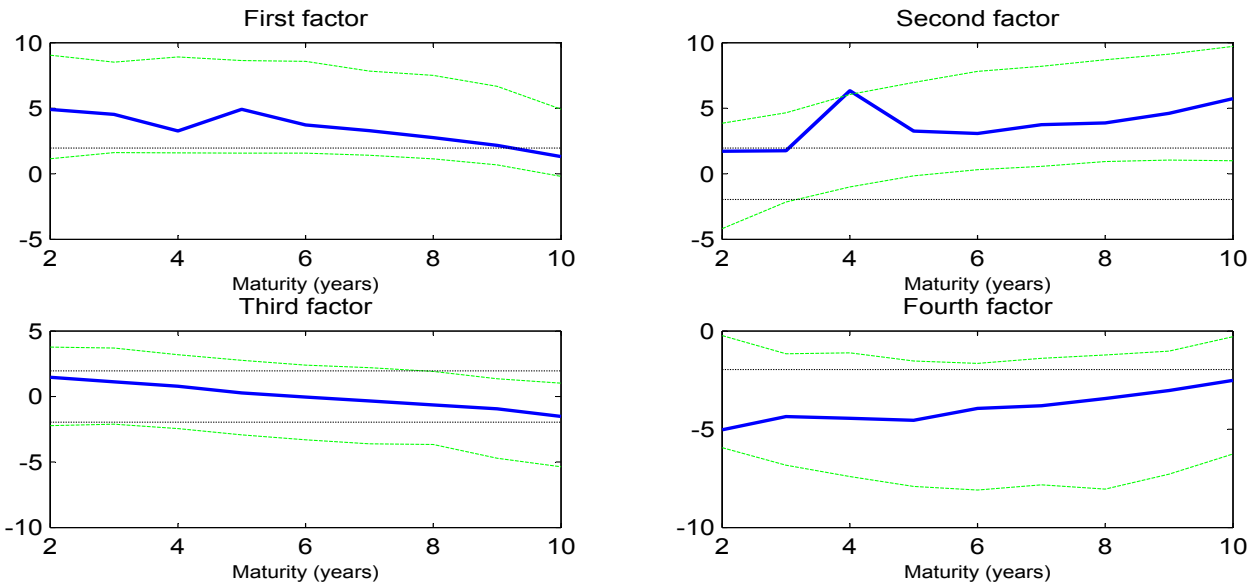
Effects of the "annual" hidden factor $\mathcal{P}_{4,t}$



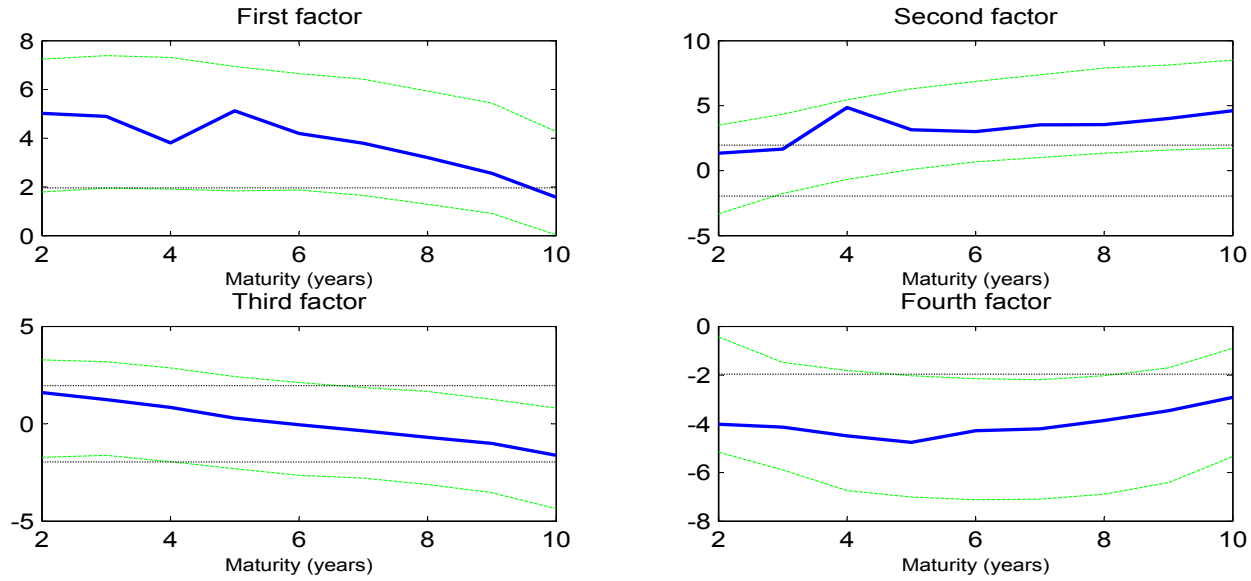
This figure exhibits the model-implied effects of hidden factors on the yield curve, on future short rates and on expected excess returns, scaled by their standard deviations. The model is estimated using monthly real bond yields from January 2004 through December 2014, with restrictions imposed on the market price of risk. H_t denotes the hidden part of the single risk-premium factor $\lambda_{1l}'\mathcal{P}_t$, which drives all variations in expected monthly excess returns. It is constructed as the residual from projecting $\lambda_{1l}'\mathcal{P}_t$ on the four principal components of observed yields $\widetilde{\mathcal{P}_t}$. $\mathcal{P}_{4,t}$ is the fourth state factor based on principal components analysis of the covariance matrix of shocks to yields. The signs of both hidden factors are reversed.

Figure 5: Significance of Filtered State Factors in Predictive Regressions of Annual Excess Returns

Panel A: t-statistics based on Hansen-Hodrick standard errors

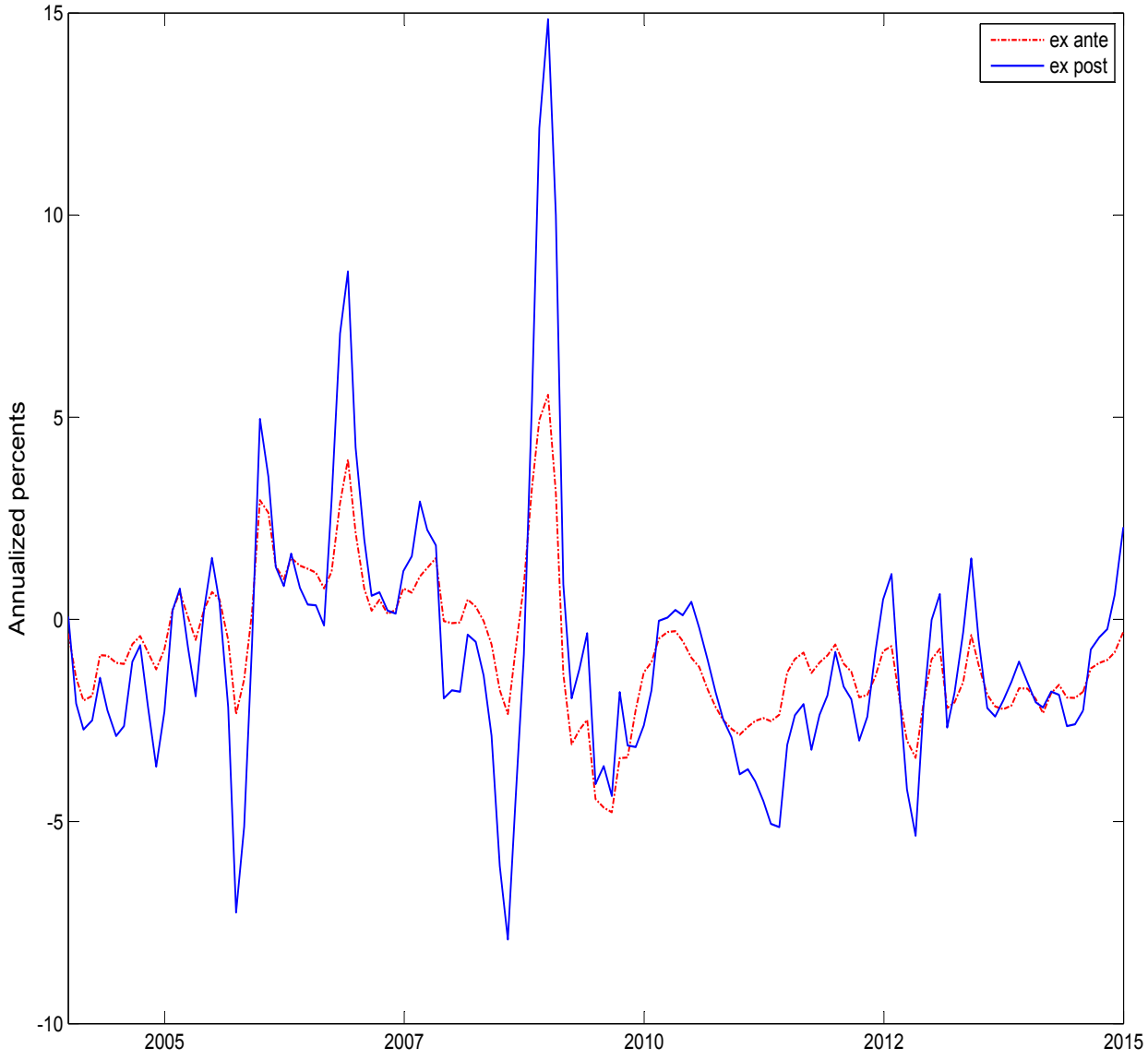


Panel B: t-statistics based on Newey-West standard errors



This figure exhibits the t-statistics in predictive regressions of annual excess bond returns on state factors filtered from estimated a four-factor GDTSM. The model is estimated using monthly real bond yields from January 2004 through December 2014. The factors are rotated into the principal components of the covariance matrix of yield innovations. The solid line in each panel shows the multivariate regression coefficients divided by Hansen-Hodrick standard errors or Newey-West standard errors. The dotted lines represent two-sided 95% confidence intervals calculated from 1,000 Monte Carlo simulations.

Figure 6: Estimated Three-Month Real Rates



This figure plots the ex post and ex ante (annualized) real returns on 3-month Treasury bills. The sample period extends from January 2004 to December 2014.

References

Ang, A., G. Bekaert, and M. Wei (2008). The term structure of real rates and expected inflation. *Journal of Finance 63*(2), 797–849.

Bauer, M. D. and G. D. Rudebusch (2015). Resolving the spanning puzzle in macro-finance term structure models. *Available at SSRN 2518037*.

Bekaert, G. and X. Wang (2010). Inflation risk and the inflation risk premium. *Economic Policy 25*(64), 755–806.

Buraschi, A. and A. Jiltsov (2005). Inflation risk premia and the expectations hypothesis. *Journal of Financial Economics 75*(2), 429–490.

Campbell, J. and R. Shiller (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies 58*(3), 495–514.

Campbell, J. and R. Shiller (1996). A scorecard for indexed government debt. *NBER Macroeconomics Annual 11*, 155–197.

Campbell, J., R. Shiller, and L. Viceira (2009). Understanding Inflation-Indexed Bond Markets. *Brookings Papers on Economic Activity 1*, 79–120.

Campbell, J. and L. Viceira (2001). Who should buy long-term bonds? *American Economic Review 91*(1), 99–127.

Chernov, M. and P. Mueller (2012). The term structure of inflation expectations. *Journal of Financial Economics 106*(2), 367–394.

Christensen, J., J. Lopez, and G. Rudebusch (2010). Inflation expectations and risk premiums in an arbitrage-free model of nominal and real bond yields. *Journal of Money, Credit and Banking 42*(6), 143.

Cochrane, J. and M. Piazzesi (2005). Bond risk premia. *American Economic Review 95*(1), 138–160.

Cochrane, J. and M. Piazzesi (2008). Decomposing the Yield Curve. *Working Paper, University of Chicago, Booth School of Business*.

Dai, Q., K. Singleton, and W. Yang (2004). Predictability of bond risk premia and affine term structure models. Working Paper, Stanford University.

DAmico, S., W. English, D. López-Salido, and E. Nelson (2012). The federal reserve's large-scale asset purchase programmes: Rationale and effects. *The Economic Journal 122*(564), F415–F446.

D'Amico, S., D. H. Kim, and M. Wei (2014). Tips from tips: the informational content of treasury inflation-protected security prices. *Working Paper, Federal Reserve*.

Duffee, G. (2010). Sharpe ratios in term structure models. Working paper, Johns Hopkins University.

Duffee, G. (2011). Information in (and not in) the term structure. *Review of Financial Studies 24*(9), 2895–2934.

Evans, M. (1998). Real rates, expected inflation, and inflation risk premia. *Journal of Finance 53*(1), 187–218.

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

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Grishchenko, O. and J.-Z. Huang (2013). Inflation Risk Premium: Evidence from the TIPS Market. *Journal of Fixed Income 22*(4), 5–30.

Gürkaynak, R., B. Sack, and J. Wright (2007). The US Treasury yield curve: 1961 to the present. *Journal of Monetary Economics 54*(8), 2291–2304.

Gürkaynak, R., B. Sack, and J. Wright (2010). The tips yield curve and inflation compensation. *American Economic Journal: Macroeconomics 2*(1), 70–92.

Haubrich, J., G. G. Pennacchi, and P. Ritchken (2012). Inflation expectations, real rates, and risk premia: Evidence from inflation swaps. *Review of Financial Studies 25*(5), 1588–1629.

Huang, J.-Z. and Z. Shi (2010). Determinants of Bond Risk Premia. *AFA 2011 Denver Meetings Paper*. Penn State University.

Huang, J.-Z. and Z. Zhong (2013). Time-variation in diversification benefits of commodity, reits, and tips. *Journal of Real Estate Finance and Economics 46*(1), 152–192.

Joslin, S., M. Priebsch, and K. J. Singleton (2014). Risk premiums in dynamic term structure models with unspanned macro risks. *Journal of Finance 69*(3), 1197–1233.

Joslin, S., K. J. Singleton, and H. Zhu (2011). A new perspective on gaussian dynamic term structure models. *Review of Financial Studies 24*(3), 926–970.

Le, A. and K. Singleton (2013). A robust analysis of the risk-structure of equilibrium term structures of bond yields. *Working Paper, Kenan-Flagler Business School, UNC*.

Ludvigson, S. and S. Ng (2008). Macro factors in bond risk premia. *Review of Financial Studies 22*, 5027–5067.

Ludvigson, S. and S. Ng (2011). A factor analysis of bond risk premia. In A. Ullah and D. E. A. Giles (Eds.), *Handbook of Empirical Economics and Finance*, pp. 313–372. CRC Press.

Nelson, C. and A. Siegel (1987). Parsimonious modeling of yield curves. *Journal of Business*, 473–489.

Pflueger, C. and L. Viceira (2010). An Empirical Decomposition of Risk and Liquidity in Nominal and Inflation-Indexed Government Bonds. *Working Paper 11-094, Harvard Business School*.

Pflueger, C. and L. Viceira (2011). Inflation-Indexed Bonds and the Expectations Hypothesis. *Annual Review of Financial Economics 3*, 139–158.

Piazzesi, M., M. Schneider, and S. Tuzel (2007). Housing, consumption and asset pricing. *Journal of Financial Economics 83*(3), 531–569.

Rauch, H. E., C. Striebel, and F. Tung (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA journal 3*(8), 1445–1450.

Roll, R. (2004). Empirical TIPS. *Financial Analysts Journal*, 31–53.

Roush, J. E. (2008). *The "growing pains" of TIPS issuance*. Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.

Shen, P. (2006). Liquidity risk premia and breakeven inflation rates. *Economic Review-Federal Reserve Bank of Kansas City 91*(2), 29.

Stock, J. H. and M. W. Watson (1999). Forecasting inflation. *Journal of Monetary Economics 44*(2), 293–335.

Svensson, L. E. (1995). Estimating forward interest rates with the extended nelson & siegel method. *Sveriges Riksbank Quarterly Review 3*(1), 13–26.