Government Bond Risk and Return in the US and China*

Jennifer N. Carpenter
NYU Stern School of Business

Fangzhou Lu
HKU Business School, The University of Hong Kong

Robert F. Whitelaw
NYU Stern School of Business and NBER

February 26, 2025

Abstract

We propose a new approach to modeling bond risk and risk premia, inspired by the equity risk-return literature, which does not impose the tight restrictions found in models that generate closed-form bond prices. We estimate the joint dynamics of the volatility and Sharpe ratio of principal-component bond-factor portfolios for the US and China. Predictors include yield curve variables and, for the US, VIX. We document complex time-varying relations between the price and quantity of interest rate risk inconsistent with the frameworks in existing studies. Interesting differences between the US and China further highlight the need for our more flexible approach.

JEL Codes: G12, G15

Keywords: bond risk premia, bond Sharpe ratios, interest rate volatility, US Treasury bonds, Chinese government bonds, no arbitrage, principal components analysis

*Many thanks for helpful comments and suggestions from an anonymous referee, the editor, Nikolai Roussanov, Darrell Duffie, Rob Engle, Matt Richardson, Alexi Savov, Yu Xu, and seminar participants at New York University and the China International Risk Forum. Please direct correspondence to jnc2@stern.nyu.edu, lufz@hku.hk, rw2@stern.nyu.edu.

Government Bond Risk and Return in the US and China

Abstract

We propose a new approach to modeling bond risk and risk premia, inspired by the equity risk-return literature, which does not impose the tight restrictions found in models that generate closed-form bond prices. We estimate the joint dynamics of the volatility and Sharpe ratio of principal-component bond-factor portfolios for the US and China. Predictors include yield curve variables and, for the US, VIX. We document complex time-varying relations between the price and quantity of interest rate risk inconsistent with the frameworks in existing studies. Interesting differences between the US and China further highlight the need for our more flexible approach.

JEL Codes: G12, G15

Keywords: bond risk premia, bond Sharpe ratios, interest rate volatility, US Treasury bonds, Chinese government bonds, no arbitrage, principal components analysis

1 Introduction

Understanding the risk and return of major asset classes is essential for optimal portfolio choice and the calibration of reasonable equilibrium models. A vast literature studies risk and return in the equity markets. The fixed income markets are even larger than the equity markets, but the literature on bond risk and return is still developing. Most existing approaches build on the term structure literature that develops models designed to match bond prices and to price interest rate derivatives. However, the constraints imposed in these models to generate closed-form bond prices greatly reduce their ability to describe the return data. By contrast, the equity return literature does not limit itself to models that match observed equity prices. This paper provides new evidence on the dynamics of government bond risk and risk premia by taking a more flexible approach to the modeling of bond volatility and Sharpe ratios, while still imposing the no-arbitrage condition.

In spirit, our paper is most closely related to the empirical stock market risk-return literature. Merton (1980) is perhaps the first paper in this literature, and he highlights the challenging nature of the econometric environment. Later papers, such as French, Schwert, and Stambaugh (1987), Glosten, Jagannathan, and Runkle (1993), and Whitelaw (1994), document a weak, or even negative, relation between conditional expected returns and volatility, despite the large unconditional equity risk premium. Our paper examines the link between risk and return in the bond market, which may be a more natural starting point given the absence of the complexities associated with cash flow risks inherent in stock returns.

We study government bond markets in both the US and China, which are, respectively, the largest and second largest bond markets in the world.[1] Our paper is one of the first to provide evidence on the pricing of Chinese government bonds (CGB). Although the time series of CGB pricing data is still relatively short, China's bond market is growing explosively, as shown in Figure 1, and already has a total market value of 21 trillion USD at the end of 2023, compared with 59 trillion USD for the US bond market. CGB constitute a smaller fraction of this market than US Treasuries (UST) represent as a fraction of the total US bond market, 21% vs. 42%, respectively, but they still represent an important benchmark for pricing. Of equal importance, CGB returns have low correlation with UST returns, thus providing important independent evidence. In fact, some of the dynamics of risk and return in China are quite different from those in the US. Of course, as an emerging market, the CGB market is very likely subject to political risk, at least from the perspective of foreign investors. Moreover, participation in this market by global central banks and other foreign

[1] See Amstad and He (2018) for a description of China's bond markets and Clayton, Dos Santos, Maggiori, and Schreger (2022) for an analysis of the internationalization of the Renminbi.

investors has increased over our sample period, raising concerns about non-stationarity. Both of these factors may affect our results, and we make some effort to quantify their potential influence.

One of the barriers to empirical work on bond returns is the absence of natural bond portfolio return series. Our paper begins by constructing principal-component bond-factor portfolio returns. We use data on key-maturity UST and CGB par rates to construct monthly excess returns on zero-coupon bonds with annual maturities from one to ten years. We then use a principal components analysis (PCA) of the standardized excess returns on these bonds to reduce each bond market to two factor portfolios, which together explain most of the variation in the zero returns. For example, in the US bond market over the post-Volcker period, Factor 1 explains 91% of this variation and Factor 2 explains 7%. In China these proportions are 82% and 14%, respectively. Consistent with Litterman and Scheinkman (1991), movements in the first and second factor portfolios correspond to movements in the level and steepness of the yield curve. Interestingly, this is true in both the US and China.

Next, we lay out a continuous-time model of nominal bond returns, in which we incorporate the two-factor structure of bond returns as well as the no-arbitrage condition that risk premia are solely compensation for risk. The discrete-time analogue of this model guides our empirical specification of monthly bond-factor returns, in which we take conditional factor volatilities and Sharpe ratios to be functions of a set of predictor variables.

Finally, as our main analysis, we perform a simultaneous generalized method of moments (GMM) estimation of the joint dynamics of the conditional volatility and Sharpe ratio processes for each bond factor. For both the US and China, we use traditional yield curve variables and lagged realized volatility to forecast the volatilities and Sharpe ratios of the bond-factor portfolios. For the US bond returns, we also introduce VIX as an important predictor variable, unspanned by yields. In addition, in the US, we introduce an indicator variable that allows the coefficients in our specification to differ in the period during which short-term interest rates were at or close to the zero lower bound (ZLB). This estimation allows us to test hypotheses about the relation between bond risk and risk premia and uncover the underlying dynamic structure of bond returns.

We have three main findings. First, we identify an important second factor in bond risk premia, which accounts for the fact that unconditional Sharpe ratios of bonds decline in maturity in both the US and China. The similarity between the factor structure of bond returns in the US and China is striking given that these are two effectively segmented markets whose returns are relatively uncorrelated. Thus, this structure may, in fact, be an inherent feature of default-free bond returns rather than something characteristic of only the US market.

Second, for each bond-factor portfolio, both the conditional volatility and the conditional Sharpe ratio vary stochastically. However, the nature of this stochastic variation differs markedly across the two factors within a country, and across the two countries. For the first and dominant factor in the US, the conditional Sharpe ratio and conditional volatility are highly correlated, as equilibrium models would predict for risk factors that are correlated with aggregate consumption. This strong positive correlation aligns well with the tight restrictions implied by many of the models that generate closed-form bond prices, which explains why these models have had success in matching the US data. However, the second US factor, while significantly smaller, is critical for matching certain important features of bond risk premia, as argued above. While the correlation between the conditional volatility and Sharpe ratio of this second factor is also positive and quite high in our baseline model without the ZLB indicator, it becomes negative once we include this indicator. Interestingly, this negative unconditional correlation is the product of a highly negative correlation during the ZLB period and a positive correlation during the non-ZLB period. This feature of the data is difficult, if not impossible, to accommodate in existing theoretical models that generate closed-form bond prices.

The CGB results also highlight the importance of modeling flexibility. In China, both factors exhibit negative correlation between the conditional volatility and Sharpe ratio over the full sample, but substantial variation in this correlation over shorter periods, spanning both positive and negative values. Thus, the ability of existing theoretical models to fit the US data should not be interpreted as evidence that these models are adequate for understanding default-free bond returns more generally. For the first and largest factor in CGB returns, this negative correlation is driven primarily by large declines in volatility and increases in Sharpe ratios during aggressive monetary policy interventions associated with two crisis periods: the financial crisis of 2008 and the stock market crash of 2015. This result in some ways parallels the ZLB result in the US. Aggressive monetary policy interventions appear to be associated with a decoupling of the prices and quantities of risk in both markets.

The stochastic variations in the factor volatilities and Sharpe ratios combine to generate interesting risk and return dynamics for two-year and ten-year zero-coupon bonds. Risk premia in the US exhibit both interesting cyclical patterns and time trends. Specifically, the term structure of bond risk premia is steeply upward sloping at the beginning of expansions, but declines over the cycle to the point where it is flat. However, we show that these variations in risk premia are difficult to understand fully without our decomposition into the price and quantity of risk. This interpretation is made more difficult by the fact that these components are strongly positively correlated in the US. Of some interest, over time, the Sharpe ratios of both factors have declined to the point where risk premia are close to zero across bonds of all maturities. In China, volatility has declined, but there is no evidence of a decline in the price of risk of either factor, so Sharpe ratios remain higher on shorter-term zeroes in recent years.

Third, and finally, we find that bond risk premia are solely compensation for bond risk in both countries, as no-arbitrage theory predicts. That is, bond risk premia go to zero as bond volatility goes to zero. The fact that we are unable to reject this hypothesis provides some evidence that our empirical specification is reasonable, i.e., that both the choice of the predictor variables and the functional form of conditional volatility and Sharpe ratio specifications are adequate to capture the key features of the data. Moreover, imposing this restriction improves the power of the estimation and sharpens the results. Note that though this no-arbitrage result might seem unsurprising, it is in contrast to the finding in the equity literature of a weak or even negative relation between risk and return.

While our approach is motivated by the equity return literature, our results are related to, and extend those of, the term structure literature. One strand of this literature, which includes Fama and Bliss (1987), Cochrane and Piazzesi (2005), and Ludvigson and Ng (2009), focuses on uncovering violations of the Expectations Hypothesis, documenting time variation in bond risk premia, and identifying key predictor variables such as forward rates and macro factors.[2] Another strand, which includes Ang and Piazzesi (2003), Duffee (2010), Duffee (2011), Wright (2011), Lettau and Wachter (2011), Joslin, Singleton, and Zhu (2011), Adrian, Crump, and Moench (2013), Joslin, Le, and Singleton (2013), Joslin, Priebsch, and Singleton (2014), Greenwood and Vayanos (2014), and Cieslak and Povala (2015), works with Gaussian term structure models, in which bond prices have deterministic volatility. These models force bond Sharpe ratios to do all the work of accommodating stochastic variation in risk premia. All of these papers are largely silent on the corresponding dynamics of bond risk; however, a significant body of research, inspired by Engle, Lilien, and Robins (1987) provides evidence of systematic time variation in interest rate volatility.[3] Thus, while they may fit risk premia well, since these premia can be levered arbitrarily, they are not very informative without an understanding of the true decomposition into the quantity and price of risk. For example, the portfolio response to an increase in the conditional risk premium driven by an increase in volatility is very different than if the increase in the conditional risk premium is driven by an increase in Sharpe ratio.

The broader class of affine term structure models (Duffie and Kan, 1996), which can accommodate stochastic volatility, has been a popular framework for modeling the dynamics of bond returns. Unfortunately, affine models of bond risk premia can only incorporate

[2]See also Campbell and Shiller (1991).

[3]See, e.g., Boudoukh, Downing, Richardson, Stanton, and Whitelaw (2010) and references cited therein.

stochastic volatility in bond returns by imposing a tight link between the functional forms of the price and quantity of risk (Dai and Singleton, 2000; Duffee, 2002; Cheridito, Filipović, and Kimmel, 2007; Cieslak and Povala, 2016; Joslin and Le, 2021). In addition, affine models typically imply that bond yields span all relevant information about bond risk premia, except in knife-edge cases (Duffee, 2011; Joslin et al., 2014). Thus they generically rule out unspanned stochastic volatility, such as that documented by Collin-Dufresne and Goldstein (2002), and unspanned macro predictors of bond risk and return.[4]

More recently, Filipović, Larsson, and Trolle (2017) develop a set of linear-rational term structure models that allow for unspanned volatility while still delivering closed-form bond prices. Creal and Wu (2020) develop a consumption-based equilibrium model that also accommodates stochastic volatility and risk price while delivering closed-form bond prices. However, these models still impose tight, albeit different, restrictions on the forms of the price and quantity of interest rate risk. Others, such as Ghysels, Le, Park, and Zhu (2014), Creal and Wu (2017), and Li, Sarno, and Zinna (2021), attempt to accommodate stochastic volatility under the true probability measure while preserving the exponential-affine form of bond prices by assuming that volatility under the risk-neutral measure is constant. These models are inconsistent with the evidence in Figure 2 that option-implied bond volatility varies stochastically.

This paper goes beyond the confines of models that deliver closed-form bond prices in order to let the data speak more freely about the dynamics of bond returns. The disadvantage of this approach is that it does not model the underlying riskless rate process, and it does not exploit information from matching observed price levels in the estimation. A fully specified term structure model will obviously have more power, but only if the model is correct. If the model is misspecified, then any empirical results are extremely difficult to interpret. Given the empirical evidence that we document, especially for our ZLB-augmented model in the US and for China, model misspecification is a clear concern. The advantage of our approach is that we estimate a more flexible model, which better captures key empirical features of bond risk and risk premia, and is therefore more useful for investors and risk managers.

The paper proceeds as follows. To lay the groundwork for our main analysis, Section 2 presents preliminary evidence on the performance and factor structure of government bond excess returns in both the US and China. Section 3 lays out our theoretical model of nominal bond returns, the corresponding empirical specification, and our estimation strategy. Section 4 presents the estimation results for US Treasury bonds and Section 5 presents the

[4]There is an unresolved debate in the literature about whether macro factors have predictive power incremental to that contained in the yield curve (Bauer and Hamilton, 2018). The resolution of this debate is tangential to the main points we make in this paper.

estimation results for Chinese government bonds. Section 6 concludes.

2 Preliminary Evidence on Bond Returns

To lay the groundwork for our model of conditional bond return volatility and price of risk in Section 3, this section presents the results of PCAs of implied zero-coupon bond excess returns in the US and China. Much of the existing empirical literature, going back to Fama and Bliss (1987), forecasts bond risk premia maturity by maturity. For a number of reasons, we focus on forecasting the risk premia of the first two principal components of bond returns, i.e., the risk premia on portfolios of these bonds. First, using returns on portfolios rather than on individual bonds avoids many of the measurement error issues that have been discussed extensively in the prior literature. Specifically, in regressions that use maturity-matched term structure variables as predictors, the same bond price shows up in both the return on the left-hand side of the forecasting regression and the yield or forward rate for the same maturity on the right-hand side. Thus, the same measurement error in this price also potentially shows up on both sides of the regression equation. We also use yields as predictors, but there are returns of bonds with many different maturities in the portfolio return we are trying to predict, so the problem of common measurement error is much less severe. Second, the PCA dramatically reduces the dimensionality of the problem, so we can present results for only two factors rather than for multiple different maturities, making the results easier to analyze and interpret. Third, more recent papers, such as Cochrane and Piazzesi (2005) and Cieslak and Povala (2015), emphasize the existence of a single dominant factor in expected returns. In a no-arbitrage framework, a single factor structure would imply that all bonds have the approximately the same Sharpe ratio, assuming little idiosyncratic risk, which is inconsistent with the strongly declining Sharpe ratio pattern in the data that we discuss below. However, given the low dimensionality of the bond return data, two factors are likely to pick up much of the time variation in returns and thus of risk premia. This two-factor structure is consistent with the empirical results of Duffee (2010) who finds this feature of the data to be robust in a variety of estimated Gaussian models.

The results of these PCAs are strikingly consistent across the US and China. They also explain the pattern of declining Sharpe ratios with maturity, documented by Duffee (2010) and Frazzini and Pedersen (2014), in terms of an important second priced factor, on which short-term bonds load positively and long-term bonds load negatively.

2.1 Priced Factors in Bond Returns

In the spirit of the analysis of Litterman and Scheinkman (1991) for UST implied zeroes over the period 1984–1988, Panel A of Table 1 presents the results of PCAs of the standardized excess returns of the implied zeroes. To construct the monthly returns on implied zero-coupon bonds with annual maturities 1, 2, ..., 10 years, we first fit a cubic exponential spline function through the key-maturity par rates from FRED for the US or from WIND for China. Then we back out the implied zero rates for semi-annual maturities, fit another spline through these implied zero rates, and compute monthly prices and returns for zero-coupon bonds with monthly maturities. The columns on the left-hand side of Table 1 are for UST implied zeroes for two subperiods, 7/1976–12/1989 and 1/1990–12/2022. These correspond roughly to the Volcker period and the post-Volcker period.[5] The columns on the right are for CGB implied zeroes.

In each subsample, we standardize each zero's excess return series by its monthly volatility so that the PCA is not dominated by the longer-maturity, higher-volatility zero returns.[6] Thus, in the ten-maturity zero PCAs, the sum of the ten annualized variances, and thus the sum of the ten resulting principal-component factor-portfolio variances, is 120. Panel A of Table 1 contains the results for the first three principal-component factor portfolios. The first row shows the percent of total variation explained by each of these portfolios. The first factor explains most of the total variation of the standardized zero returns, while the second factor also explains a material portion. In the more recent subperiods, the second factor becomes more important. For the UST implied zeroes during the post-Volcker period, Factor 1 explains 91% of the total variance of the standardized zero returns, while Factor 2 explains 7%. Factor 3 explains an additional 1% of the variation and the remaining factors are negligible. For the CGB implied zeroes, the second factor is even more important; Factor 1 explains 82% of total variance and Factor 2 explains 14%. Panel A of Table 1 also shows the annualized Sharpe ratios of each of the factor portfolios. We sign the factors so that they have positive Sharpe ratios. The Sharpe ratios of Factors 1 and 2 are fairly large, especially in the UST zeroes in the post-Volcker period, where the Factor 1 portfolio has a Sharpe ratio of 0.64 and the Factor 2 portfolio has a Sharpe ratio of 0.73.[7]

The column-vector of zero loadings under each factor in Panel A of Table 1 is the factor eigenvector. It simultaneously shows the loadings of the different standardized zero returns

[5]Paul Volcker was Chairman of the Federal Reserve from August 1979 to August 1987. The precise start of the second subperiod is dictated by the availability of VIX data as we discuss later.

[6]The results with unstandardized excess returns are qualitatively similar.

[7]Balduzzi, Connolly, and Marcus (2021) obtain an essentially equivalent rotation of US Factors 1 and 2, with similar implications for bond returns, using an alternative construction based on cross-sectional regressions of zero excess returns on duration and convexity.

on the factor portfolio return and the holdings of standardized zeroes in the factor portfolio. The compositions of the three factor portfolios are similar across subperiods and across markets. Factor 1 is a roughly equal-weighted portfolio of standardized zeroes. Factor 2 is long short-maturity zeroes and short long-maturity zeroes. For the UST implied zeroes, Factor 3 is long extreme-maturity zeroes and short middle-maturity zeroes. For the CGB implied zeroes, Factor 3 is short extreme-maturity zeroes and long middle-maturity zeroes, but the sign of this factor is not stable since its Sharpe ratio is very close to zero. As in Litterman and Scheinkman (1991), movements in the three factors correspond roughly to shifts in the level, steepness, and curvature of the yield curve, respectively, for all subsamples.

2.2 The "Betting-Against-Duration" Pattern in Sharpe Ratios

Duffee (2010) and Frazzini and Pedersen (2014) document a "betting-against-duration" pattern in the Sharpe ratios of Treasury bonds: Sharpe ratios are declining with bond maturity. We verify that this pattern is robust across two US subsamples and in China. Panel B of Table 1 presents unconditional annualized mean monthly excess returns, volatilities, and Sharpe ratios for the ten constant-maturity zeroes. In both subperiods, the means and volatilities of the UST implied zero returns are increasing with zero maturity, while their Sharpe ratios are decreasing with maturity. The patterns of the performance measures for the CGB implied zeroes are qualitatively very similar. In particular, the Sharpe ratios of CGB implied zeroes are also mostly declining in maturity. This is somewhat surprising, given that the Chinese securities markets are largely segmented from other global financial markets, with limited ownership by foreign investors, and given that CGB bond-factor portfolio returns have low correlation with the UST bond-factor portfolio returns. The highest correlation is 22%, between CGB Factor 1 and UST Factor 1 returns.

Frazzini and Pedersen (2014) attribute the "betting-against-beta" pattern in asset prices to leverage-constrained investors bidding up high-beta assets for their high returns. However, this explanation is less plausible in the bond markets, where the repo market facilitates the use of leverage. The declining pattern of bond Sharpe ratios with maturity is better explained through the presence of the important second priced factor in bond returns, on which short bonds load positively and long bonds load negatively.

2.3 The Factor Structure and Performance of UST ETFs

Table 2 verifies that the bond factor structure and performance patterns presented in Table 1 are not simply artifacts of our implied zeroes construction by demonstrating the same patterns in the excess returns of UST exchange-traded funds (ETFs). These ETFs are traded assets, in contrast to our synthetic zeroes, and therefore their returns are free from any measurement error that might be induced by our splining procedure, for example. The data, from the Center for Research in Security Prices for the period 2/2007 to 12/2022, are for returns net of fees. Table 2 shows results for excess returns augmented with the 15-basis point annual management fee charged by Blackrock iShares. Gross of these fees, the Sharpe ratios on the ETFs decline sharply with the maturity of the underlying bonds. Panel A of Table 2 verifies that the factor structure of UST ETF returns mirrors that of the UST implied zeroes. The large Sharpe ratio on Factor 2 explains the declining pattern of ETF Sharpe ratios with maturity that we document in Panel B.

3 A Model of Nominal Bond Returns

This section develops a model of nominal bond returns that positions the bond market within the broader financial market, formalizes our assumptions about the factor structure of bond returns, derives a testable no-arbitrage relation between bond risk and return, and motivates the empirical specification of bond returns in the estimation that follows. Suppose real asset prices are Itô processes with respect to a standard d-dimensional Brownian motion B_t. In particular, there is a riskless real money market account with instantaneous riskless rate r_t and there are n risky assets with real cum-dividend prices $S_{i,t}$ that follow

$$\frac{dS_{i,t}}{S_{i,t}} = \mu_{i,t}\, dt + \sigma_{i,t}\, dB_t \ , \tag{1}$$

where r_t, the $\mu_{i,t}$, and the d-dimensional row vectors $\sigma_{i,t}$ are stochastic processes that are measurable with respect to the information generated by the Brownian motion and satisfy standard integrability conditions that ensure the processes $S_{i,t}$ are well-defined. The value W_t of a self-financing portfolio that invests value $\pi_{i,t}$ in risky asset i, for $i = 1, \ldots, n$, follows

$$dW_t = (r_t W_t + \pi_t(\mu_t - r_t 1))\, dt + \pi_t \sigma_t\, dB_t \ , \tag{2}$$

where π_t is the n-dimensional row vector with elements $\pi_{i,t}$, μ_t is the n-dimensional column vector with elements $\mu_{i,t}$, 1 is the n-dimensional column vector of 1's, and σ_t is the $n \times d$-dimensional matrix with rows equal to the $\sigma_{i,t}$. Assume that π_t is such that $\pi_t(\mu_t - r_t 1)$ and $\pi_t \sigma_t$ satisfy the integrability conditions that ensure W_t is well-defined.

3.1 The No-Arbitrage Condition

In the absence of arbitrage, the real price processes $S_{i,t}$ must satisfy the condition that if π_t is such that $\pi_t \sigma_t = 0$, then $\pi_t(\mu_t - r_t 1) = 0$. That is, a portfolio with zero risk must have a zero risk premium. Otherwise, it would be possible to generate a locally riskless portfolio that appreciates at a rate greater than r_t. This condition is algebraically equivalent to the condition that there exists a d-dimensional vector θ_t such that

$$\sigma_t \theta_t = \mu_t - r_t 1 \ . \tag{3}$$

It follows that there exists a d-dimensional vector process θ_t satisfying Equation (3), as well as suitable measurability and integrability conditions.[8] This process is typically called a "market price of risk" or simply a "price of risk." Therefore, in the absence of arbitrage, we can re-write Equation (1) as

$$\frac{dS_{i,t}}{S_{i,t}} - r_t \, dt = \sigma_{i,t} \theta_t \, dt + \sigma_{i,t} \, dB_t \ , \tag{4}$$

for any market price of risk process θ_t. Moreover, together with the riskless rate r_t, any such market price of risk process θ_t determines the dynamics of a stochastic discount factor

$$M_t = e^{-\int_0^t r_s \, ds - \int_0^t \theta_s' \, dB_s - \frac{1}{2} \int_0^t |\theta_s|^2 \, ds} \tag{5}$$

such that

$$S_{i,t} = \mathrm{E}_t \{ \frac{M_u}{M_t} S_{i,u} \} \text{ for all } 0 < t < u \text{ and } i = 1, \ldots, n \ . \tag{6}$$

In many equilibrium models, the equilibrium stochastic discount factor is equal to the marginal utility of consumption of the representative agent, and the equilibrium market price of risk on the claim to aggregate consumption is

$$\theta_t = R_t \sigma_{c,t} \tag{7}$$

where R_t is the coefficient of relative risk aversion of the representative agent, and $\sigma_{c,t}$ is the volatility vector of aggregate consumption.[9]

[8] See Karatzas and Shreve (1998), Theorem 4.2.

[9] See, for example, Karatzas and Shreve (1998), Eqn. (6.21).

3.2 Nominal Asset Prices with Locally Riskless Inflation

Suppose the price level q_t is locally riskless, i.e.,

$$\frac{dq_t}{q_t} = i_t\, dt, \tag{8}$$

where the expected inflation rate i_t is suitably integrable and measurable with respect to the information generated by the d Brownian motions. Then the nominal riskless rate, that is, the rate on a nominally riskless money market account, is $r_t + i_t$ and nominal asset prices, $P_{i,t} = q_t S_{i,t}$ satisfy

$$\frac{dP_{i,t}}{P_{i,t}} - (r_t + i_t)\, dt = \frac{dq_t}{q_t} + \frac{dS_{i,t}}{S_{i,t}} - (r_t + i_t)\, dt = \frac{dS_{i,t}}{S_{i,t}} - r_t\, dt = \sigma_{i,t}\theta_t\, dt + \sigma_{i,t}\, dB_t\ . \tag{9}$$

Thus, nominal returns in excess of the nominal riskless rate are the same as real returns in excess of the real riskless rate, and can shed light on the real price of risk θ_t.[10]

Note that the nominal stochastic discount factor for nominal asset prices is

$$M_t/q_t = e^{-\int_0^t (r_s+i_s)\, ds - \int_0^t \theta_s'\, dB_s - \frac{1}{2}\int_0^t |\theta_s|^2\, ds}\ , \tag{10}$$

and the nominal price of a nominal zero-coupon bond with maturity T is

$$P_t^T = \mathrm{E}_t\{e^{-\int_t^T (r_s+i_s)\, ds - \int_t^T \theta_s'\, dB_s - \frac{1}{2}\int_t^T |\theta_s|^2\, ds}\}\ . \tag{11}$$

Therefore, the volatilities of nominal bond returns will in general reflect exposure to shocks to the inflation rate i_t, and the risk premia on nominal bonds will contain compensation for this exposure. I.e., there will in general be an inflation risk premium in both the real and nominal excess returns of nominal bonds.

3.3 Bond Market Factors and Implied Zero Excess Returns

Motivated by the evidence from Section 2.1 of two important, orthogonal factor portfolios, which together explain virtually all of the variation in nominal bond returns, we identify the excess return of Factor 1 with the first Brownian motion and the excess return of Factor 2 with the second Brownian motion. This identification is without loss of generality, since we can always rotate the original Brownian motions to achieve this representation. Thus, for

[10] Cochrane and Piazzesi (2005) and Cieslak and Povala (2015) effectively make this assumption as well.

$j = 1, 2$, we write the excess return on Factor j, dF_j as

$$dF_{j,t} = \sigma_{j,t}\theta_{j,t}\,dt + \sigma_{j,t}\,dB_{j,t} \ , \tag{12}$$

where for $j = 1, 2$, $\sigma_{j,t}$ is now the scalar conditional volatility process for Factor j and θ_j is now the uniquely defined Sharpe ratio for Factor j. A natural interpretation is that Factors 1 and 2 from the bond market are correlated with important latent risk factors in aggregate consumption, and their Sharpe ratios thus shed light on the prices of those dimensions of consumption risk.

Next, taking the ten annual maturity nominal implied zeroes to be the first ten risky assets in the market, we write the nominal implied zero excess returns as

$$\frac{dP_{i,t}}{P_{i,t}} - (r_t + i_t)\,dt = \beta_{i,1}dF_{1,t} + \beta_{i,2}dF_{2,t} \ , \ \text{for } i = 1, \ldots, 10 \ , \tag{13}$$

where $\beta_{i,1}$ and $\beta_{i,2}$ are the components of the eigenvectors associated with Factors 1 and 2, respectively. In particular, in light of evidence that the risk associated with the third and higher principal components is economically negligible, we treat the zero-cost constant-maturity implied-zero portfolios as constant-beta portfolios of the Factors 1 and 2 only. Note that we are not restricting the conditional Factor 1 and Factor 2 volatilities and Sharpe ratios $\sigma_{j,t}$ and $\theta_{j,t}$ to depend only on the information generated by the first two Brownian motions. In general, these can depend on the information generated by the full set of d Brownian motions, which justifies the possibility of a large set of predictor variables for these conditional moments, not limited to bond yields. In particular, this flexible model can accommodate unspanned stochastic volatility, such as that documented by Collin-Dufresne and Goldstein (2002), and unspanned macro risks, such as in Joslin et al. (2014), among others.

Once we empirically characterize the conditional factor volatilities and Sharpe ratios $\sigma_{j,t}$ and $\theta_{j,t}$, then we can recover the conditional volatility of each implied zero i as the two-dimensional vector $(\beta_{i,1}\sigma_{1,t}, \beta_{i,2}\sigma_{2,t})$ and the risk premium on implied zero i as $\beta_{i,1}\sigma_{1,t}\theta_{1,t} + \beta_{i,2}\sigma_{2,t}\theta_{j,t}$. In particular, the risk premia on the two factors, $\sigma_{1,t}\theta_{1,t}$ and $\sigma_{2,t}\theta_{2,t}$, will drive the risk premia on all ten zeroes, simply as a consequence of the two-factor structure of bond returns. To the extent that the first bond factor's risk premium, $\sigma_{1,t}\theta_{1,t}$, is dominant, as the evidence in Table 1 suggests, it will appear as though this single forecasting variable drives returns on all zeroes, with the individual zero loadings given by the $\beta_{i,1}$. For the ordinary unstandardized zero returns, each zero's loading is its element in the Factor 1 eigenvector in Panel A of Table 1 times its volatility from Panel B of Table 1. These loadings are monotonic in the maturity of the zeroes. Thus, the presence of a dominant first bond factor with time-varying risk premia will produce the finding of Cochrane and Piazzesi (2005) that a single forecasting factor drives returns on all bonds, with loadings monotonic in maturity.

3.4 Empirical Specification and GMM Estimation

To take the continuous-time model to monthly time-series data, we work with a discrete-time analogue of Equation (12),

$$R_{j,t+1} = \sigma_{j,t}\theta_{j,t} + \sigma_{j,t}\varepsilon_{j,t+1} \text{ for } j = 1, 2 \ , \tag{14}$$

where R_j is the monthly excess return on Factor j, the $\varepsilon_{j,t}$ are i.i.d. standard normal, and we assume that the volatilities and prices of risk satisfy

$$\sigma_{j,t} = X_t \beta_j^{\sigma} \tag{15}$$

and

$$\theta_{j,t} = X_t \beta_j^{\theta} \tag{16}$$

for a row-vector of predictor variables, X_t, which includes a constant.

3.4.1 Predictor Variables

A large literature going back to Fama (1986) uses yield curve variables to forecast bond risk premia, while another literature going back to Chan, Karolyi, Longstaff, and Sanders (1992) uses yield curve variables to forecast interest rate volatility. To capture the information about future bond return volatility and risk premia in the yield curve, X_t includes three variables that describe the yield curve level, slope, and curvature, namely, the two-year zero-coupon yield, $Y_{2,t}$, the ten-year yield minus the two-year yield, $Y_{2,t} - Y_{10,t}$, and the six-year yield minus the average of the two- and ten-year yields, $Y_{6,t} - \frac{Y_{2,t}+Y_{10,t}}{2}$.[11] As with the return data, we make a conscious choice to reduce the dimensionality of the yield data used as predictors for a number of reasons. First, and most important, we want to reduce the possibility of overfitting. Second, the structure of yields looks similar to the structure of returns in that there are a few factors that capture the vast majority of the time variation in these series. While it is theoretically possible that a yield factor that explains a very small fraction of the variation in yields explains a large fraction of the variation in risk premia, this possibility seems economically implausible. Third, the goal of the paper is not to maximize the R^2's

[11] We use the two-year yield rather than the one-year yield to avoid any distortions in the short end of the yield curve associated with monetary policy, although using the latter instead of the former produces qualitatively similar results.

of our regressions. Rather we are trying to illuminate the underlying economic structure of bond risk premia in as simple and parsimonious a specification as possible. We leave a detailed specification search intended to maximize forecasting power to future research.

For the UST factors, X_t also includes VIX, which is an index of the implied volatility of the 30-day return on the S&P 500 derived from S&P 500 index options.[12] In theory, this implied volatility measure contains both a forecast of market volatility and information about risk aversion, so it may be relevant for predicting both bond return volatility and the price of risk.

We also tried the MOVE Index, which tracks the U.S. Treasury yield volatility implied by current prices of one-month over-the-counter options on two-year, five-year, ten-year and thirty-year Treasuries. MOVE is highly correlated with VIX and is subsumed by VIX in our empirical specifications. This result is perhaps surprising, since one might expect that a bond market volatility measure such as MOVE would do better than a stock market measure such as VIX. However, the latter is based on a much more liquid and widely traded set of instruments, especially in the early part of the sample, which may explain the result.

In addition, in an effort to decompose implied volatility into information about future volatility and information about risk aversion, we estimate a GARCH(1,1) model on the S&P 500 monthly return series, on a rolling basis. This model produces a monthly series of volatility forecasts. The difference between VIX and this forecast is an estimate of the price of volatility risk. The conclusions from this analysis are twofold. First, when included as a substitute for VIX, the GARCH volatility forecast plays a very similar role in the specifications, showing up with coefficients of the same sign and magnitude, albeit less statistically significant in most cases. Second, the difference between VIX and the GARCH volatility forecast, when included with the GARCH volatility forecast, shows no incremental explanatory power.[13] For the CGB factors, the analogous GARCH forecast of stock market volatility in China does not enter significantly in any of the specifications and does not qualitatively alter any of the results.

For both the UST and CGB factors, we also include the lagged value of the realized volatility of each bond factor return, approximated as $\sqrt{\frac{\pi}{2}}|R_{j,t}|$, as an alternative predictor for bond market volatility. This variable is motivated by the well known empirical stylized fact that return volatility in financial markets is persistent.

In the U.S., we introduce a final indicator variable in an augmented specification to

[12]The VIX data are available from the CBOE going back to January 1990, which dictates the precise start date of the sample period for our GMM estimation. This date also coincides roughly with the end of the Volcker period.

[13]This result is consistent with Cieslak and Povala (2016) who look directly at the role of the interest rate variance risk premium in explaining bond risk premia and find that it is negligible.

address the question of whether the periods in which the short-term interest rate approached the ZLB are meaningfully different in terms of the properties of the price and quantity of interest rate risk. Specifically, this variable equals one if the effective Fed funds rate, from FRED, at the end of the prior month was less than 25 basis points. This definition picks up two ZLB subperiods, the post-crisis period 1/2009-12/2015 and the COVID pandemic period 4/2020-3/2022. A number of papers attempt to tackle the question of the effects of the ZLB in the context of a term structure model (see, for example, Wu and Xia (2016) and Filipović et al. (2017)). While the results are somewhat mixed, evidence of changes in term structure dynamics near the ZLB do not necessarily imply different functional forms for volatilities or Sharpe ratios. Differential dynamics of the various predictor variables during the ZLB periods could in principle capture the ZLB effects. However, it is an interesting empirical question.

Of course, the aggressive monetary policy responses to the financial crisis and the COVID pandemic were not limited to very low interest rates, but also included forward guidance and quantitative easing. Introducing additional variables in our specification, such as the quantity of asset purchases or the size of the Fed balance sheet is conceptually simple. However, there are concerns with overfitting, especially given the lack of theoretical guidance as to which specific variables should be included. Moreover, as shown in the recent paper Haddad, Moreira, and Muir (2025), it is not just the specific actions of the Fed, but the market's perception of their future state-contingent policy actions and their impact that influence prices and risk. One further argument against attempting to include these additional factors is that the variables that are already in the specification, i.e., yields and implied volatility, may be sufficient to capture the relevant effects. For example, some monetary policy acts directly through yields, which, in turn, also reflect the market's required returns.[14] While this argument might also apply to the ZLB indicator, as noted above, periods when interest rates are close to their lower bound may be fundamentally different. This constraint on yields and their movements may fundamentally alter the dynamics of yields and the way in which they reflect market expectations. Thus, we would argue that a ZLB indicator is perhaps a more natural first variable to include.

Finally, in China, we also introduce an indicator variable to check for non-stationarity of the specification over time. This variable divides the full sample in two equal subsamples in order to test if the estimated parameters change significantly from the first to the second half of the period. In addition, we introduce two other predictor variables, the total amount

[14] This argument is closely related to the argument made in Bauer and Hamilton (2018) for why one should not necessarily expect macro variables to provide additional predictive power for bond risk premia over and above the information contained in the yield curve.

of government bonds issued in the month and the total trading volume in all government bonds during the month. In general, these two quantities increase over our sample. If these increases are associated with changes in the price or quantity of risk, perhaps as liquidity improves or the participation of different investors alters the pricing of risk, these variables should proxy for these changes.

A number of other variables have been used to predict bond excess returns in the literature. Fama and Bliss (1987) use matching-maturity forward rates to forecast excess returns on zeroes with annual maturities one through five years. Cochrane and Piazzesi (2005) use all five forward rates to forecast the excess returns on individual zeroes with annual maturities one through five years. In our setting, we work with factor portfolios of zeroes with annual maturities up to ten years. To include all ten forward rates seems likely to lead to over-fitting, so we prefer the more parsimonious summary of yield curve information contained in our Level, Slope, and Curvature variables, which correspond roughly to the first three principal components of yields. Ang and Piazzesi (2003) and Joslin et al. (2014) use measures of economic growth and inflation, Ludvigson and Ng (2009) use PCs from 132 macro variables, Greenwood and Vayanos (2014) use measures of Treasury bond supply, Cieslak and Povala (2015) use residuals from regressions of yields on an average of past inflation, and Brooks and Moskowitz (2017) use measures of value, momentum, and carry. We limit our predictor variables to our three yield curve variables, VIX, lagged realized volatility, and the country-specific variables described above, which seem natural and well-motivated.

3.4.2 GMM Estimation Equations and Diagnostics

In our baseline specification, for each factor $j = 1, 2$, we perform a simultaneous GMM estimation of β_j^σ and β_j^θ from the following two equations:

$$R_{j,t+1} = \alpha_j + (X_t\beta_j^\sigma)(X_t\beta_j^\theta) + u_{j,t+1} \ , \tag{17}$$

$$\sqrt{\frac{\pi}{2}}|u_{j,t+1}| = X_t\beta_j^\sigma + v_{j,t+1} \ , \tag{18}$$

where we use $\mathrm{E}\{\sqrt{\frac{\pi}{2}}|u_{j,t}|\} = \mathrm{E}\{\sqrt{\frac{\pi}{2}}|\sigma_{j,t-1}\varepsilon_{j,t}|\} = \sigma_{j,t-1}$. We refer to Equation (17) as the "return equation" and Equation (18) as the "volatility equation." The "return constant" α_j in Equation (17) should be zero in theory by no arbitrage.[15] We include this constant in preliminary specifications to check for possible mis-specification in Equations (15) and (16).

[15] Other papers that have made this point in the context of bond pricing include Cox, Ingersoll, and Ross (1985) and Stanton (1997).

Unless otherwise specified, the set of moment conditions we use in the estimations are

$$\mathrm{E}\{u_{j,t+1}Z_t\} = \mathrm{E}\{[R_{j,t+1} - [\alpha_j + (X_t\beta_j^{\sigma})(X_t\beta_j^{\theta})]]Z_t\} = 0 \ , \quad (19)$$

$$\mathrm{E}\{v_{j,t+1}X_t'\} = \mathrm{E}\{[\sqrt{\frac{\pi}{2}}|R_{j,t+1} - [\alpha_j + (X_t\beta_j^{\sigma})(X_t\beta_j^{\theta})]| - X_t\beta_j^{\sigma}]X_t'\} = 0 \ , \quad (20)$$

where the vector Z_t includes all of the unique elements of the matrix $X_t'X_t$. These moment conditions allow us to test the restrictions on the coefficients on the square and cross-product terms in $X_t'X_t$ imposed by Equations (15) and (16) using the standard J-statistic over-identifying restrictions test.

In our augmented specification, we introduce the indicator variable L_t, described in the previous subsection, which takes the value one if the short-term interest rate at time t is close to the ZLB and zero otherwise. Then we augment equations (17) and (18) as follows:

$$R_{j,t+1} = \alpha_j + \delta_j L_t + (X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma})(X_t\beta_j^{\theta} + L_tX_t\gamma_j^{\theta}) + u_{j,t+1} \ , \quad (21)$$

$$\sqrt{\frac{\pi}{2}}|u_{j,t+1}| = X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma} + v_{j,t+1} \ . \quad (22)$$

In this specification, the coefficients β_j^{σ} and β_j^{θ} capture estimates for the non-ZLB periods, and the coefficients γ_j^{σ} and γ_j^{θ} are the estimated differences in the coefficients between these periods and the ZLB periods. Unless otherwise specified, the set of moment conditions we use in the estimations are

$$\mathrm{E}\{u_{j,t+1}Z_t\} = \mathrm{E}\{[R_{j,t+1} - [\alpha_j + \delta_j L_t + (X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma})(X_t\beta_j^{\theta} + L_tX_t\gamma_j^{\theta})]Z_t\} = 0 \ , \quad (23)$$

$$\mathrm{E}\{u_{j,t+1}L_tX_t'\} = \mathrm{E}\{[R_{j,t+1} - [\alpha_j + \delta_j L_t + (X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma})(X_t\beta_j^{\theta} + L_tX_t\gamma_j^{\theta})]L_tX_t'\} = 0 \ , \quad (24)$$

$$\mathrm{E}\{v_{j,t+1}X_t'\} = \mathrm{E}\{[\sqrt{\frac{\pi}{2}}|R_{j,t+1} - [\alpha_j + \delta_j L_t + (X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma})(X_t\beta_j^{\theta} + L_tX_t\gamma_j^{\theta})]| - (X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma})]X_t'\} = 0 \ , \quad (25)$$

$$\mathrm{E}\{v_{j,t+1}L_tX_t'\} = \mathrm{E}\{[\sqrt{\frac{\pi}{2}}|R_{j,t+1} - [\alpha_j + \delta_j L_t + (X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma})(X_t\beta_j^{\theta} + L_tX_t\gamma_j^{\theta})]| - (X_t\beta_j^{\sigma} + L_tX_t\gamma_j^{\sigma})]L_tX_t'\} = 0 \ . \quad (26)$$

This augmented specification provides a good illustration of the flexibility of our empirical approach. With our empirical methodology, the conditional volatility and the conditional

Sharpe ratio over the period t to $t+1$ can be modeled as functions of any variable that is in the information set at time t. We estimate this augmented specification using the full sample, and the methodology easily accommodates non-contiguous ZLB periods; we do not estimate a separate model for subsamples of the data, but a single model with time-varying coefficients. This estimation strategy also allows for simple tests of whether the coefficients are statistically different in the ZLB period.

We report goodness-of-fit measures for the two estimated equations, defined as

$$\text{Goodness-of-fit (1)} = 1 - \frac{\sum_t v_{j,t}^2}{\frac{\pi}{2}\sum_t(|u_{j,t}| - \bar{|u_j|})^2} \ , \tag{27}$$

$$\text{Goodness-of-fit (2)} = 1 - \frac{\sum_t u_{j,t}^2}{\sum_t(R_{j,t} - \bar{R}_j)^2} \ . \tag{28}$$

These are similar to ordinary-least-squares (OLS) regression R^2's. The difference is that an OLS regression chooses coefficients to maximize R^2, while the GMM estimation chooses coefficients to minimize the weighted sum of the squares and cross-products of the sample moments.

In addition, in our baseline specification, we formally test three null hypotheses about the dynamics of the bond factor returns. The first null hypothesis, based on the no-arbitrage theory, is that bond factor risk premia are solely compensation for bond risk, that is,

$$H_{0,0} : \alpha_j = 0 \ .$$

We test this with the standard z-test. The second null hypothesis is that bond factor volatility is constant, that is,

$$H_{0,1} : \beta^{\sigma}_{j,1} = \beta^{\sigma}_{j,2} = \cdots = \beta^{\sigma}_{j,k} = 0 \ ,$$

where the $\beta^{\sigma}_{j,1}, \ldots, \beta^{\sigma}_{j,k}$ are the volatility coefficients on the k non-constant elements of X. We test this joint hypothesis with a standard Wald test. The third null hypothesis is that the price of bond factor risk is constant, that is,

$$H_{0,2} : \beta^{\theta}_{j,1} = \cdots = \beta^{\theta}_{j,k} = 0 \ ,$$

where the $\beta^{\theta}_{j,1}, \ldots, \beta^{\theta}_{j,k}$ are the Sharpe ratio coefficients on the k non-constant elements of X. We also test this joint hypothesis with a standard Wald test.

Finally, in the augmented specification, we formally test two additional null hypotheses, i.e., that bond factor volatility coefficients are the same in the ZLB and non-ZLB periods:

$$H_{0,1} : \gamma^{\sigma}_{j,1} = \gamma^{\sigma}_{j,2} = \cdots = \gamma^{\sigma}_{j,k} = 0 \ ,$$

and that the bond factor Sharpe ratio coefficients are the same in the two periods:

$$H_{0,2} : \gamma^{\theta}_{j,1} = \cdots = \gamma^{\theta}_{j,k} = 0 \ ,$$

Again, these hypotheses are tested using a standard Wald test.

4 Results for US Treasury Bonds

This section first presents the results of the GMM estimation of UST factor volatility and Sharpe ratio dynamics using data from FRED for the period 1990 to 2022. Then we analyze the time series of fitted volatility and Sharpe ratio values of the two factors, with a particular focus on business cycle variation and the effects of the ZLB period. To shed additional light on the dynamics of return premia and highlight the advantage of decomposing these premia into their components, we examine more closely the fitted expected excess returns on two-year and ten-year zero coupon bonds. This exercise also allows us to contrast our results with empirical approaches that cannot accommodate this decomposition. Next, we provide evidence of the robustness of our results to alternative specifications. Finally, we provide evidence on the effect of the length of the return horizon, monthly or annual, on the OLS R^2's of excess return regressions, and we show that our goodness-of-fit measures for the return equation are comparable to R^2's in bond return regressions documented in the previous literature.

4.1 GMM Estimation Results for the UST Factors

The top panel of Table 3 presents GMM estimates of α_j, β^{σ}_j, β^{θ}_j, and their robust z-statistics for alternative specifications of Equations (17) and (18) for the UST factors. The bottom panel indicates the number of moment conditions used in the estimation, the p-value of the J-statistic over-identifying restrictions test, p-values for the Wald tests of null hypotheses $H_{0,1}$ and $H_{0,2}$ described above, and the goodness-of-fit measures. The left side of Table 3 reports results for UST Factor 1 and the right side reports results for UST Factor 2. For convenience, the yield curve variables are multiplied by 10 and VIX is divided by 100 to give their coefficients comparable magnitude.

The first specification for UST Factor 1, Specification (1a), includes all the predictor variables linearly, as well as the "return constant" α_1. The z-statistic for the estimate of the return constant is insignificant, as predicted by theory. The p-value of the J-statistic test for mis-specification is large, suggesting that we are not omitting any important higher-order terms in our specification. The p-values for the Wald tests indicate that we can easily reject Hypothesis $H_{0,1}$ that Factor 1 volatility is constant but we cannot yet reject Hypothesis $H_{0,2}$ that the Factor 1 Sharpe ratio is constant. However, when we impose the no-arbitrage restriction that $\alpha_1 = 0$ in Specification (1b), we increase power.[16] In particular, while the estimates of the volatility and Sharpe ratio coefficients β_j^σ and β_j^θ in Specification (1b) remain similar to those in (1a), we are now not only able to reject $H_{0,1}$ easily but we are also able to reject $H_{0,2}$ at close to the 10% level. The Curvature and Realized Vol variables are insignificant in both the volatility and return equations, so to further increase power, we exclude these variables in Specification (1c). More generally, we exclude any variable that is insignificant in both equations, but we keep any variable in both equations if it is significant in either equation. The exclusion of the insignificant variables boosts the significance levels of most of the coefficients on the other predictor variables. In particular, in Specification (1c), both the volatility and the Sharpe ratio of UST Factor 1 are significantly positive functions of Level and Slope, consistent with previous studies forecasting bond risk premia and interest rate volatility, although our analysis is the first to decompose these effects into the price and quantity of interest rate risk in bond returns. We also find that the volatility of Factor 1 is a significantly positive function of VIX. The p-value of the J-statistic remains large, suggesting this model is well-specified, and the p-values of the Wald tests are 0.0% and 5.4%, so we reject that volatility and the price of risk are constant.

For UST Factor 2, Specifications (2a) and (2b) are analogous to (1a) and (1b) for Factor 1. The p-values of the J-statistics are still well above 10%, suggesting that the linear specifications are adequate. The estimate of the return constant α_2 in Specification (2a) is insignificant, so we impose the no-arbitrage restriction $\alpha_2 = 0$ in Specification (2b). This again boosts power, and brings the p-values for the Wald tests down below 1%. Thus, we strongly reject the hypotheses that Factor 2 volatility is constant and that the price of Factor 2 risk is constant. Factor 2 volatility is a significantly positive function of Level, Slope, VIX, and Realized Vol and a significantly negative function of Curvature. Factor 2 price of risk is a significantly positive function of Level and VIX.

The result that expected returns in the bond market are solely compensation for risk, i.e., that bond risk premia go to zero as bond risk goes to zero, is consistent with the no-

[16] The decision about whether or not to impose this restriction involves the usual tradeoff between efficiency and robustness, as noted in a slightly different asset pricing context by Cochrane (2005) (see p. 236). We follow the natural recommendation of Lewellen, Nagel, and Shanken (2010) to both test the restriction and impose it ex ante (see the discussion of their Prescription 2).

arbitrage restriction in our model of Section 3. However, this result is in stark contrast to much of the literature on the risk-return relation in the stock market. Starting with French et al. (1987), this literature has often failed to find a statistically significant or even positive relation between expected returns and the conditional volatility of stock returns, despite the evidence of a large unconditional equity risk premium.

One important question is whether our estimates change during the two subperiods during which short-term interest rates were at or close to the ZLB. To address this question we estimate the augmented specification from Equations (21) and (22). The results from the estimation of this augmented specification are provided in Table 4. First, and perhaps counter to intuition, the statistical evidence that the coefficients for the first and dominant factor in returns are different in the ZLB period is relatively weak. Specifically, the Wald tests that the coefficients on the predictor variables are different in the ZLB period have p-values of 57% and 13% for the volatility and Sharpe ratio functions, respectively. These results do not imply that the properties of returns were the same in the ZLB and non-ZLB periods, but rather that these differences are potentially picked up by the differences in the levels of the state variables that forecast the price and quantity of risk.

Second, in spite of this weak statistical evidence of a ZLB effect for the first factor, from an economic perspective, there are interesting differences in the coefficient estimates across the two subperiods. Specifically, return volatility seems to be much more strongly positively related to the level of interest rates in the ZLB period, although the coefficients are statistically indistinguishable, while the Sharpe ratio seems to be negatively related to the level in this same period. In addition, the Sharpe ratio appears to be much more strongly positively related to the slope of the term structure during the ZLB period, and this difference is marginally statistically significant.

Interestingly, for Factor 2, the evidence for a ZLB effect is both statistically and economically strong. Specifically, the hypothesis that the coefficients are the same in the ZLB period can be rejected at the 0% and 2% levels for the volatility and Sharpe ratio, respectively. These rejections are driven by estimated coefficients that appear to vary dramatically in magnitude across the two periods. Focusing on the estimated coefficients with the strongest statistical evidence of differences in the ZLB period, it is again the Slope variable that stands out. In contrast to the results for Factor 1, for Factor 2, the Sharpe ratio is strongly negatively related to the slope of the term structure during the ZLB period, while the coefficient is statistically indistinguishable from zero in the non-ZLB period. In other words, upward sloping term structures suggest high prices of risk for Factor 1 and low prices of risk for Factor 2 in the ZLB period. However, for Factor 2, but not for Factor 1, there is also strong evidence that these same upward sloping term structures are associated with higher volatility during the ZLB period. One other notable result is that the effect of VIX on the volatility appears to disappear in the ZLB period. The estimated coefficient on this variable is positive and statistically significant for both the full sample and for the non-ZLB period, but it is close to zero in the ZLB period. Interestingly, Realized Vol exhibits the same behavior, i.e., it is a significantly positive predictor of volatility for both the full sample and the non-ZLB period, but the effect disappears completely, or even reverses during the ZLB period, although the evidence is not statistically strong.

Given the evidence that the ZLB effect may be economically important for both factors, though it is statistically weak for Factor 1, we report results for both specifications for the remainder of this section. The potential benefit of including the ZLB effect is obvious–if it is truly economically important, then including it generates a more meaningful picture of bond risk and return. The cost of inclusion is that it significantly increases the number of parameters that need to be estimated and thus, in most cases, the standard errors associated with these estimates. The ZLB period is relatively short, therefore the differences between the coefficient estimates in this period and those in the longer non-ZLB period will be difficult to pin down and subject to significant estimation error.

4.2 Fitted UST Factor Volatilities and Sharpe Ratios

Figures 3 plots the time series of annualized fitted values of UST Factor 1 Sharpe ratios and volatilities. The top panel, Panel A, is based on the GMM estimates from Table 3, Specification (1c), and the bottom panel, Panel B, plots the same quantities using the GMM estimates from the model augmented with ZLB effects in Table 4. We focus first on the baseline specification without ZLB effects. As is clear from a cursory examination of Panel A, the correlation between the Sharpe ratio (price of risk) and the volatility (quantity of risk) is strongly positive–the time-series correlation between these series is 93%. The positive relation between the factor prices and quantities of risk is consistent with the predictions of equilibrium models for the pricing of risk factors that are correlated with aggregate consumption.[17] At the same time, the fitted Sharpe ratios for Factor 1 change sign over the sample period, which cannot be accommodated by affine models with stochastic variation in volatility (Duffee, 2002).

The high correlation between the Sharpe ratio and volatility for the first, and empirically dominant, factor may explain why the literature has focused almost exclusively on theoretical models that generate closed-form solutions for bond prices but also impose tight restrictions on the functional relation between the price and quantity of risk. Specifically, this high

[17] See, for example, Campbell (1987).

correlation is broadly consistent with these tight restrictions and thus explains why these models may do a decent job of fitting the return data.

Panel A also shows that the quantity of risk appears to rise during NBER recessions. Increases in volatility during recessions are also a feature seen in other financial and economic series. There is, perhaps, some evidence of a similar rise in the price of risk during recessions. This effect is similar to the cyclical pattern of the US stock market Sharpe ratio demonstrated by Tang and Whitelaw (2011), and it is consistent with increasing risk aversion in bad economic times. More generally, the components of the risk premia on Factors 1 and 2, discussed below, exhibit significant correlations with various aggregate macroeconomic variables such as GDP growth and the unemployment rate. That said, while statistically significant, these variables do not explain an economically large fraction of variation in the price and quantity of risk, and the relations are sensitive to the specification. Moreover, we find little or no evidence that these macroeconomic variables have additional predictive power on top of those that we already include in our specification. This result is consistent with Bauer and Hamilton (2018), who question the evidence that macroeconomic variables provide additional information for risk premia over and above that contained in the yield curve. To summarize, we would argue that there is clearly variation in the price and quantity of risk that is related to the business cycle, but the relations are more complex than should be characterized simply by words such as "countercyclical," consistent with the conclusion of Duffee (2011), again in the context of bond risk premia.

In addition to these cyclical patterns, there is also evidence of a notable decline in the volatility over the sample period. Fitted volatilities are approximately two thirds as large at the end of the sample as they are at the beginning. However, this latter part of the sample is dominated by the ZLB period, so one might legitimately wonder if this effect is primarily a result of this unusual period. The Sharpe ratio exhibits a similar time series pattern, exceeding one in the early 1990s but declining to values that average closer to zero from 2012 onwards. We address the question of whether this is a trend or a ZLB effect with the augmented specification discussed next.

Turning to Panel B, it is clear that the inclusion of a ZLB indicator does alter some of the conclusions from the baseline model. First, while the unconditional correlation between the Factor 1 price and quantity of risk is still high at 71% in the ZLB-augmented specification, this unconditional correlation is the result of a very high 95% correlation in the non-ZLB period and a much lower 56% correlation in the ZLB period. In other words, the ZLB period seems to be associated with a partial decoupling of the price and quantity of risk for the first factor. Second, inferences about time trends in the Sharpe ratio and volatility do seem to be sensitive to the inclusion of an indicator variable for the ZLB period. Specifically, the evidence that there has been a large decline in risk is much weaker. The decrease in volatility appears to be primarily attributable to low volatility in the ZLB period, particularly the subperiod associated with the COVID pandemic. Both between the two ZLB subperiods, and after the end of the most recent episode, volatility seems to have returned to a level much closer to that of earlier in the sample. However, the evidence for a decline in the price of risk appears more robust to the inclusion of the ZLB indicator. Finally, there is distinct evidence that the two ZLB subperiods were, themselves, quite different. Specifically, the pandemic period is associated with both a much lower volatility and much lower price of risk than the earlier subperiod after the financial crisis. Of course, all of these results need to be interpreted with some caution given the relatively weak statistical evidence for a ZLB effect in the first factor.

We now turn to Figure 4, which plots the same time series of annualized fitted Sharpe ratio and volatility for Factor 2. Again, the top panel, Panel A, is based on the baseline specification, in this case the GMM estimates from Table 3, Specification (2b), and the bottom panel, Panel B, plots the same quantities using the GMM estimates from the model augmented with ZLB effects in Table 4. As above, we focus first on the baseline specification without ZLB effects. The correlation between the Sharpe ratio and the volatility is again strongly positive, albeit less so than for Factor 1, with a time-series correlation of 62%. Similar to Factor 1, Factor 2 risk, and perhaps to a lesser extent the price of this risk, appear to rise during recessions. There is also an apparent decline in the levels of both series over time, although the decline in volatility is less marked, and both the volatility and Sharpe ratio do show rebounds at the end of the sample.

However, for the second factor, the inclusion of a ZLB effect has more dramatic effects, which is not surprising given the strong statistical evidence in Table 4. First, the positive unconditional correlation reverses sign to a value of -54% for the ZLB-augmented model, which again is a product of two very different correlations in the non-ZLB and ZLB periods of 33% and -93%, respectively. Note that this strong negative correlation in the ZLB period shows up in both subperiods, post-crisis and during the pandemic, and it is not simply a function of declines in volatility and increases in Sharpe ratios during these periods. Rather, the post-crisis period shows episodes of both low volatility and high Sharpe ratios and higher volatility and low Sharpe ratios. Only the former are apparent in the pandemic period. Moreover, the fitted Sharpe ratios, when they are high, are much higher than in the non-ZLB period. In other words, the ZLB period appears to exhibit, at least at some points, very high prices for taking on slope risk. Of some interest, these prices of risk are particularly high at both the beginning and ends of these periods, perhaps when there is larger uncertainty about the future course of monetary policy. During the heart of the post-crisis period, in contrast, while the volatility of Factor 2 is quite high, the price of risk appears to be low or even negative. Together, these results highlight both the flexibility of our approach, and also the dangers of imposing tight restrictions on the relation between the price and quantity of interest rate risk, which is typical of existing approaches.

Given that we have factored conditional risk premia into conditional volatilities and conditional Sharpe ratios, a natural question to ask is, what is the relative contribution to the time variation in risk premia of each of these two component factors? To address this question, we start with the decomposition $\Delta(\hat{\sigma}\hat{\theta}) = \hat{\sigma}(\Delta\hat{\theta}) + \hat{\theta}(\Delta\hat{\sigma}) + (\Delta\hat{\sigma})(\Delta\hat{\theta})$. Then, using a first-order approximation, we drop the higher-order term $(\Delta\hat{\sigma})(\Delta\hat{\theta})$ and approximate the squared change in the risk premium as

$$[\Delta(\hat{\sigma}\hat{\theta})]^2 \approx \hat{\sigma}^2(\Delta\hat{\theta})^2 + \hat{\theta}^2(\Delta\hat{\sigma})^2 + 2\hat{\sigma}\hat{\theta}(\Delta\hat{\sigma})(\Delta\hat{\theta}) \ . \tag{29}$$

Summing Equation (29) over the observations in our sample and dividing by the left-hand side, we get a sample variance decomposition under the assumption that the mean of the risk premium is zero. In the baseline specification, for UST Factor 1, the components of this decomposition are $\frac{\sum_{t=1}^{T-1}\hat{\sigma}_t^2(\hat{\theta}_{t+1}-\hat{\theta}_t)^2}{\sum_{t=1}^{T-1}(\hat{\sigma}_{t+1}\hat{\theta}_{t+1}-\hat{\sigma}_t\hat{\theta}_t)^2} = 57\%$, $\frac{\sum_{t=1}^{T-1}\hat{\theta}_t^2(\hat{\sigma}_{t+1}-\hat{\sigma}_t)^2}{\sum_{t=1}^{T-1}(\hat{\sigma}_{t+1}\hat{\theta}_{t+1}-\hat{\sigma}_t\hat{\theta}_t)^2} = 8\%$, and $2\frac{\sum_{t=1}^{T-1}\hat{\sigma}_t\hat{\theta}_t(\hat{\sigma}_{t+1}-\hat{\sigma}_t)(\hat{\theta}_{t+1}-\hat{\theta}_t)}{\sum_{t=1}^{T-1}(\hat{\sigma}_{t+1}\hat{\theta}_{t+1}-\hat{\sigma}_t\hat{\theta}_t)^2} = 33\%$.[18] Thus, neglecting variation in volatility or the price of risk will effectively misattribute a large fraction of the variation in risk premia. Similarly, for UST Factor 2 these components are $\frac{\sum_{t=1}^{T-1}\hat{\sigma}_t^2(\hat{\theta}_{t+1}-\hat{\theta}_t)^2}{\sum_{t=1}^{T-1}(\hat{\sigma}_{t+1}\hat{\theta}_{t+1}-\hat{\sigma}_t\hat{\theta}_t)^2} = 34\%$, $\frac{\sum_{t=1}^{T-1}\hat{\theta}_t^2(\hat{\sigma}_{t+1}-\hat{\sigma}_t)^2}{\sum_{t=1}^{T-1}(\hat{\sigma}_{t+1}\hat{\theta}_{t+1}-\hat{\sigma}_t\hat{\theta}_t)^2} = 22\%$, and $2\frac{\sum_{t=1}^{T-1}\hat{\sigma}_t\hat{\theta}_t(\hat{\sigma}_{t+1}-\hat{\sigma}_t)(\hat{\theta}_{t+1}-\hat{\theta}_t)}{\sum_{t=1}^{T-1}(\hat{\sigma}_{t+1}\hat{\theta}_{t+1}-\hat{\sigma}_t\hat{\theta}_t)^2} = 47\%$. As in the case of UST Factor 1, both components are important for explaining the variation in risk premia. Our approximation is much less accurate for the augmented model, but the same basic result holds. These results suggest that neglecting either component is a mistake, and, more specifically, empirical studies motivated by constant volatility models, where all variation in risk premia is attributable to movements in the price of risk, are missing an important part of the story.

4.3 Fitted UST Bond Volatilities, Sharpe Ratios, and Risk Premia

The factor volatilities and prices of risk are interesting in their own right, but the underlying securities in this market, US Treasury bonds, load on both factors, with these loadings varying across maturity. Thus, the returns on these securities may exhibit other interesting features. Moreover, the existing empirical literature tends to focus on estimating bond return premia, so a closer examination of these premia allows us to illustrate the advantages of our decomposition of premia into the price and quantity of risk.

[18]These components do not sum to exactly one because we dropped the higher-order terms.

As discussed in Section 3.3, we can recover the risk and return dynamics of the zero-coupon bonds from the dynamics of the factors together with the zero volatilities and the loadings of the standardized zeroes on the factors from the principal components analysis in Table 1. For simplicity, we assume that just the first two principal component factors are driving the zero returns and we ignore Factors 3 through 10 since their combined explanatory power is small. Thus, the standardized monthly excess return on the zero with maturity i is the loading-weighted sum of the monthly excess returns on Factors 1 and 2:

$$sz_{i,t} = \beta_{i1} R_{1,t} + \beta_{i2} R_{2,t} \ , \tag{30}$$

where β_{ij} are the loadings of standardized zero i on factor j from Table 1, Panel A. Letting v_i denote the unconditional monthly volatility of the i-year zero, based on Table 1, Panel B, the unstandardized monthly excess return on zero i is $z_{i,t} = v_i sz_{i,t}$. It follows that the annualized fitted conditional volatility of the unstandardized excess return on zero i is

$$vol_t(z_i) = \sqrt{12} v_i \sqrt{\beta_{i1}^2 \hat{\sigma}_{1,t}^2 + \beta_{i2}^2 \hat{\sigma}_{2,t}^2} \ . \tag{31}$$

Similarly, the annualized fitted conditional risk premium of the unstandardized excess return on zero i is

$$rp_t(z_i) = 12 v_i (\beta_{i1} \hat{\sigma}_{1,t} \hat{\theta}_{1,t} + \beta_{i2} \hat{\sigma}_{2,t} \hat{\theta}_{2,t}) \ , \tag{32}$$

and the annualized fitted conditional Sharpe ratio of the unstandardized excess return on zero i is

$$sr_t(z_i) = rp_t(z_i) / vol_t(z_i) \ . \tag{33}$$

Figure 5 illustrates the time series of the annualized fitted conditional volatilities, Sharpe ratios, and risk premia of the unstandardized excess returns on the UST two-year and ten-year zeroes in the baseline model. The loadings of the standardized returns on these two bonds on Factors 1 and 2, β_{i1} and β_{i2}, are 0.31 and 0.41 for the two-year and 0.31 and -0.34 for the ten-year. The bonds have the same loading on the first and empirically dominant level factor, and loadings with approximately the same magnitude but opposite signs on the slope factor. The plots illustrate a number of interesting features of the data.

We start by examining the bond risk premia in Panel C because these results can be most easily compared to those in the existing literature. These risk premia exhibit both a marked business cycle pattern and a trend over time. The patterns for both bonds are similar, with a strong positive correlation of 0.81 between the two series, but they are more visible for the ten-year simply because the returns on this bond are approximately five times more volatile due to the longer duration of this security. In terms of business cycle behavior, although there are only four recessions in the sample, marked by shaded bars in the figures, the general pattern is clear. Risk premia increase dramatically during recessions, peaking at the beginning of the subsequent expansion, and then decline over the course of the expansion. The time trend is also readily apparent–on average risk premia have declined substantially over the course of our sample. Both the local maxima at the beginning of each expansion and the local minima at the end of each expansion decline monotonically over the four cycles.

Neither of these results is new to our paper. The cyclical behavior of risk premia is noted in numerous papers, including Cochrane and Piazzesi (2005), Ludvigson and Ng (2009), Adrian et al. (2013) and Joslin et al. (2014). The idea that risk premia have declined over time is less prevalent in the literature, perhaps because this decline is especially evident in the more recent data, but it is noted by both Wright (2011) and Joslin et al. (2014), with the former paper documenting this phenomenon across a number of developed government bond markets. Adrian et al. (2013) develop perhaps the most flexible methodology for estimating bond risk premia, and their estimates are very similar to ours during the period over which our samples overlap.[19]

However, what none of these papers can deliver is a flexible yet theoretically motivated decomposition of risk premia into their two components, the prices and quantities of risk. For example, Adrian et al. (2013) assume homoscedasticity, as is true of all Gaussian models. Thus, if one interprets their model literally, all of the variation in risk premia must come from variation in the Sharpe ratio, i.e., the price of risk, since the quantity of risk, i.e., the volatility, is constant. Of course, one could examine the estimated risk premia and attempt to correlate it with other variables. In fact, in Figure 10 of their paper, they plot the term premium against the MOVE Index, a reasonable measure of conditional bond volatility. They interpret the strikingly strong positive correlation between these series "as evidence that [their] term premium estimate reflects the risk of holding Treasury securities." Our decomposition discussed below illustrates the danger of this type of ad hoc empirical approach. At the same time, we acknowledge the advantage of an approach, as in Adrian et al. (2013), that starts from a model with closed form bond prices. This feature allows the econometrician to use the information in the yield curve to estimate the model, with the possibility that this additional information will generate better risk premia estimates.

We next turn to our risk premia decomposition, with the conditional volatilities of the returns on the two bonds plotted in Panel A and the corresponding Sharpe ratios in Panel B. With regard to volatility, it is no surprise that the two-year and ten-year zeroes exhibit very

[19] Conveniently, Adrian et al. (2013) also graph their estimated risk premia for two-year and ten-year bonds. See the top two graphs in their Figures 1 and 5 for these estimates from two different model specifications. Moreover, the first graph in Figure 7 shows that these estimates for the ten-year bond are similar across four different models.

similar time series behavior, with the ten-year volatility scaled by a factor of approximately five, i.e., its relative duration.[20] Both volatility series exhibit both cyclical fluctuations and approximately the same decline in volatility exhibited by the underlying factors as discussed in Section 4.2.[21]

Panel B shows that the Sharpe ratios of both zeroes also exhibit similar cyclical fluctuations and a similar time series trend. There are major declines in fitted Sharpe ratios over the sample, but the decline is larger in magnitude for the two-year zero. For much of the sample, the Sharpe ratio on the shorter-term security exceeds that on the longer-term security, consistent with the unconditional evidence in Table 1. However, by the end of the sample these Sharpe ratios have converged, as the fitted Sharpe ratio on the second bond market factor, which determines Sharpe ratio differences across the term structure, hovers close to zero. As we noted earlier, in a world with a single priced factor on which all bonds load positively, the Sharpe ratios on all bonds are equal.

So, what do we learn from this decomposition? One important insight is that the fluctuations in risk premia are not attributable solely to movements in either the price or quantity of risk, but rather to an interaction of both components. For example, there is a substantial rise in the risk premium on the ten-year zero during the 2008 recession associated with the financial crisis. Panel B shows that the Sharpe ratio tracks this increase closely, but there is also a corresponding increase in volatility. Risk is approximately 20% higher at the end of the recession than at the beginning. Thus, the risk premium would be 20% lower at the end of the recession had volatility not changed. This phenomenon has obvious implications for risk management and investment.

Second, and related, the two components, the price and quantity of risk, are strongly positively correlated. For two-year zeroes this correlation is 0.97, while it is a somewhat lower but still large, 0.72, for the ten-year. Of course, this correlation implies that the risk premia are also highly correlated with both components, for both maturities, which means that any ad hoc empirical analysis of estimated risk premia is dangerous. Specifically, proxies for either the price or quantity of risk will appear to explain fluctuations in risk premia well when, in fact, our estimation indicates that both components contribute meaningfully to these fluctuations. Thus, correlations between risk premia with variables such as the MOVE Index, as in Adrian et al. (2013), or measures of inflation uncertainty, as in Wright (2011), are difficult to interpret correctly without the decomposition that we provide.

Finally, the variation in the price and quantity of risk provide a guide for building equi-

[20]The correlation between these two estimated volatility series is almost perfect, 0.999.

[21]An examination of the specification augmented with ZLB effects presents a more nuanced view of the decline in factor volatility, and we will look at the results from this specification below.

librium models that explain bond returns. Specifically, the correlation between the two components suggests a world in which the marginal investor in US government bonds has limited risk bearing capacity, so when the amount of risk increases, these investors must be compensated with higher Sharpe ratios in order to induce them to hold this increased quantity of risk. Interestingly, this effect is somewhat weaker for ten-year bonds due to their negative loading on Factor 2, the slope factor.

Figure 6 illustrates the same bond return dynamics, but using the estimated parameters from the augmented specification with ZLB effects. Note the same axis scaling is used for the two figures in order for ease of comparison. It is clear that including the ZLB effects induces significantly more variation in these fitted quantities, although this increase is heavily concentrated in the ZLB period, particularly in the Sharpe ratios and thus also in the bond risk premia. Estimation error is an obvious concern, but it is still interesting to examine the evidence given that there do appear to be statistically and economically significant ZLB effects.

First, as with the factors, the apparent downward trend in bond volatility is much less pronounced when controlling for the ZLB effect. In other words, although there does seem to have been a small decline in volatility over time, the lower volatility at the end of the sample is due in large part to the ZLB period, particularly the pandemic.

Second, the decline in fitted Sharpe ratios appears robust to the inclusion of the ZLB indicator. However, there is large variation in the Sharpe ratio particularly for the ten-year bond during the post-crisis ZLB subperiod. Interestingly, in the non-ZLB subperiod in the second half of the sample, Sharpe ratios on the two bonds appear very similar and very close to zero, immediately before the financial crisis, between the post-crisis ZLB subperiod and the pandemic, and after the pandemic. That is, there is little recent evidence of either the effect of the second factor on Sharpe ratios that we see over the full sample or of a significant price of interest rate risk at all.

Finally, this Sharpe ratio variation is the dominant factor in the risk premia variation, again particularly for the ten-year bond in the post-crisis ZLB subperiod. There appear to be times during this period of aggressive monetary policy, when the market was charging very large premia for holding long-term securities. Of course, the sample is small, but the evidence is interesting nonetheless. The other interesting phenomenon is that risk premia on both short- and long-term bonds have converged to values close to zero for the later non-ZLB subperiods. While there may be uncertainty about the true magnitude of some of these effects in the ZLB period, controlling for this unique period appears to tighten the inference for the non-ZLB period.

The augmented specification further highlights the importance of our decomposition relative to estimations that target only bond risk premia. Specifically, it appears that this decomposition is not stable in that the periods of aggressive monetary policy associated with the zero lower bound have very different properties than the non-ZLB period. Monetary policy risks may require compensation that is not explained by total bond return volatility in a ZLB environment.

4.4 Robustness

One might be concerned that our results discussed above are sensitive to the particular model specification that we use. For example, GARCH models are often used to estimate volatility dynamics as alternatives to our "stochastic volatility" model specified in equations (17)-(18). To address this concern, we estimate a GARCH model that is similar to our stochastic volatility model. Specifically, we estimate a GARCH(1,1)-M specification of the following form:[22]

$$R_{j,t} = (X_t\beta_j^{\theta})\sigma_{j,t-1} + \epsilon_{j,t} \qquad \epsilon_{j,t} \sim N(0, \sigma_{j,t-1}) \tag{34}$$

$$\sigma_{j,t}^2 = e^{X_t\beta_j^{\sigma}} + \rho\sigma_{j,t-1}^2 + \lambda\epsilon_{j,t}^2 \ . \tag{35}$$

Just as in our stochastic volatility model in equations (17)-(18), the expected return is modeled as a product of conditional volatility and the Sharpe ratio, both of which are allowed to depend on the full set of predictor variables. However, it is important to note that the precise specification of the expected return has very little effect on the estimation of volatility because the variation in returns is primarily due to unexpected returns, not time-variation in expected returns. The primary content of the GARCH model, relative to our stochastic volatility model, is the inclusion of the latent lagged conditional variance, $\sigma_{j,t-1}^2$, in the variance equation (35). We estimate this specification via maximum likelihood for both factors. The signs and statistical significance of the coefficient estimates are very much in line with those reported in Table 3 for our stochastic volatility model estimated via GMM. While the magnitudes are more difficult to compare due to the exponential function in the GARCH specification, the fitted conditional volatility series from the GARCH estimations have very high correlations with the corresponding fitted conditional volatility series from the stochastic volatility model estimations. We conclude that our key results are robust to the precise model specified.

A second question is how important is the inclusion of VIX as a predictor variable to our results. This question is potentially interesting for two reasons. First, as far as we know, we are the first paper to use VIX to predict bond market volatility. Second, VIX is not

[22]For brevity, we do not report the results.

available in China, so it is important to know whether this difference in the availability of predictors is likely to explain any differences between the results in the two countries. To answer this question we re-estimate Specifications (1c) and (2b) in Table 3 excluding VIX.[23] The term structure variables, Level, Slope and Curvature, have very low correlations with VIX, while Realized Vol for Factor 2 does have a somewhat higher correlation, at about 0.3. Our GMM estimation is not exactly equivalent to running linear regressions, but the intuition for the effect of excluding a predictor that has relatively low correlation with the other explanatory variables does carry through for most of the results. Specifically, there is no meaningful change in either the statistical significance of the remaining predictors or in the economic magnitudes of the coefficient estimates on these predictors. However, in some cases, the goodness of fit does show a marked decline of more than 50% when we omit VIX. Perhaps not surprisingly, these declines are larger for the volatility equation. In other words, VIX is clearly an important variable for modeling time-variation in bond risk. That said, conclusions about the general patterns in the time-variation in the price and quantity of risk are not extremely sensitive to the exclusion of VIX. The fitted values of these quantities under the two models, with and without VIX, have correlations that are never less than 0.8 and go as high as 0.95 for the Factor 2 price of risk. In summary, an examination of the time series properties of the price and quantity of risk appears to benefit significantly from the inclusion of VIX as a predictor, but the more general patterns in these series are not dependent on this inclusion.

4.5 Monthly versus Annual R^2's in Bond Return Regressions

While the empirical results in Table 3 are both economically and statistically significant, and we document significant predictable variation in UST returns, the goodness-of-fit measures in the return equation look small relative to those in the existing literature. Specifically, it is not unusual to see R^2's in linear regressions of maturity-specific bond returns on various predictor variables of 30% or more.[24] Why then are our goodness-of-fit measures so much lower than the R^2's reported in earlier papers? The simple answer is that, for the most part, the existing literature uses monthly overlapping annual returns as the dependent variable in these regressions, whereas we use non-overlapping monthly returns. As we illustrate below, the use of overlapping annual returns instead of monthly returns mechanically boosts R^2's.

However, there is one clear benefit of using monthly returns when the predictor variables are persistent: higher frequency non-overlapping returns generate larger effective sample sizes, which increases confidence in the validity of asymptotic inference and reduces concerns

[23]For brevity, we do not report the results.

[24]See, for example, Cochrane and Piazzesi (2005) and Cieslak and Povala (2015).

about small sample biases. This issue has been discussed extensively in the stock-return predictability literature, with Boudoukh and Richardson (1994) and Boudoukh, Israel, and Richardson (2021) providing a comprehensive analysis of the properties of long-horizon return regressions. In the context of bond-return predictability, Bauer and Hamilton (2018) show that there are substantial biases in the standard errors and regression R^2's in studies with overlapping annual returns due to their poor small sample properties.

We illustrate the effect of using overlapping annual returns instead of monthly, and put the goodness-of-fit measures presented in Table 3 into perspective, as follows. We estimate regressions of UST Factor 1 returns on a fitted volatility measure and contrast the R^2's from regressions of monthly returns with the R^2's from regressions of overlapping annual returns. For ease of comparison to existing papers, we do not use the simultaneous GMM estimation of Table 3, but rather a two-stage OLS approach.

Table 5 presents the full set of results in five steps. Panel A shows the first-stage regression of realized Factor 1 monthly return volatility on the three predictor variables in our preferred Specification (1c) in Table 3. In addition to the fact that this volatility regression is not estimated simultaneously with the return equation, the other difference from our previous econometric strategy is that the independent variable uses the total Factor 1 return rather than the fitted unexpected return for the obvious reason that we have not yet estimated the expected component of this return. Nevertheless, the results are very consistent with the earlier estimation. All three predictors are statistically and economically significant, and the magnitudes of the coefficients are similar.

The fitted monthly volatility from this first-stage regression will be the predictor variable in the second-stage return equation. However, before we get to this estimation, it is important to understand the time-series properties of this predictor. Therefore, Panel B shows the results from a simple first-order autoregression (AR(1)) of fitted volatility. There are two related results of note. Fitted volatility is extremely persistent, with an autoregression coefficient exceeding 0.9, and this simple AR(1) model seems to provide a reasonably good description of the data. The high serial correlation is of particular importance, because it is this feature together with the overlap in annual returns that boosts the R^2 of the annual-return regression and also creates small-sample biases.

In Panel C we run the second-stage predictive regression for monthly Factor 1 returns. This regression is likely mis-specified, given the evidence in Table 3 of a time-varying price of risk, but it is sufficient to illustrate the point. Fitted volatility predicts returns with a positive and significant coefficient and an R^2 of just over 4%, which is slightly below the goodness-of-fit from our GMM specification. Up to this point in Table 5, we have only reported simple OLS t-statistics in parentheses but we now also report Newey-West t-statistics in square brackets, calculated using twelve lags. At the monthly frequency, the Newey-West adjustment makes little difference because there is little, if any, serial correlation in the monthly returns.

Panel D illustrates what happens to this predictive regression when the returns are aggregated to the annual level. The same fitted volatility is used as the lone predictor variable, and the regression uses monthly overlapping annual returns. The results are striking. The R^2 increases by a factor of approximately five and the coefficient increases by even more. In many ways, these results look much more impressive than their monthly counterparts, but are they really? Not surprisingly, the OLS t-statistic is deceptively high. Once we adjust for serial correlation in the residuals, the t-statistic returns to the level from the monthly regression. Moreover, even this t-statistic is likely overstated because, while the Newey-West methodology has good asymptotic properties, it underweights the correlations in small samples in the context of overlapping data in order to ensure positive definiteness.

Boudoukh, Richardson, and Whitelaw (2008) show analytically how the regression coefficient and the R^2 should scale up as the data are aggregated. Specifically, even under the null hypothesis that there is no true predictability, if the predictor is sufficiently highly autocorrelated, these estimated quantities increase dramatically with the return horizon. Panel E shows the annual-return regression coefficient and R^2 that the econometrician should expect to see under the assumption that fitted volatility follows an AR(1).[25] In particular, even when the annual-return regression provides no incremental information about return predictability relative to the monthly return regression, the econometrician should expect to see an R^2 an order of magnitude higher with the annual regression. This phenomenon is what Boudoukh et al. (2008) call the myth of long-horizon predictability. The annual R^2 of 27%, while seemingly very large, provides no more evidence of predictability than the monthly R^2 closer to 4%. In this particular instance, the implied annual numbers actually exceed the estimates generated using annual returns, so the idea that running annual return regressions provides incremental information is difficult to support.

Putting these results together, our conclusion is that there is no good reason to use annual returns in our analyses. While the goodness-of-fit measures using monthly returns may look less impressive, statistically and economically they support the same conclusions without the econometric problems associated with using long-horizon, overlapping return data.

[25] See equations (6) and (7) in Boudoukh et al. (2008).

5 Results for Chinese Government Bonds

This section first presents the results of the GMM estimation of CGB factor volatility and Sharpe ratio dynamics using data from Wind for the period 5/2004 to 12/2022. Then we analyze the time series of fitted volatilities and Sharpe ratios for bond-factor portfolios and individual bonds in China. Finally, we check the robustness of our results to the inclusion of additional predictor variables that might pick up non-stationarity due to changes in the CGB market over time.

These results are important for three reasons. First, the size of the CGB market and its increasing global importance make the market inherently worthy of study. Second, since for most of the sample the CGB market was effectively segmented from the UST market, the CGB market provides independent evidence on the pricing of interest rate risk. Third, the structure of the CGB market is quite different from the UST market, therefore these results shed some light on the extent to which market structure affects the pricing of risk.

5.1 GMM Estimation Results for the CGB Factors

The top panel of Table 6 presents GMM estimates of α_j, β_j^σ, β_j^θ, and their robust z-statistics for alternative specifications of Equations (17) and (18) for the CGB factors. The bottom panel indicates the number of moment conditions used in the estimation, the p-value of the J-statistic over-identifying restrictions test, p-values for the Wald tests of null hypotheses $H_{0,1}$ and $H_{0,2}$, and the goodness-of-fit measures. The left side of Table 6 reports results for CGB Factor 1 and the right side reports results for CGB Factor 2.

For each CGB factor, the table reports results for specifications that include all three yield curve variables and the lagged realized volatility in the volatility and Sharpe ratio functions. The p-values of the J-statistic tests for Factor 2 are uniformly high, suggesting that linear functions of the predictor variables are adequate for modeling the factor volatilities and Sharpe Ratios for this factor. These statistics are lower for Factor 1, but, given the relatively short sample, we decide not to explore the inclusion of higher order functions of the predictor variables. For CGB Factor 1, Column (1a) of Table 6 reports estimation results for the specification that includes the return constant α_1. As the table shows, the estimate of α_1 is insignificant, as no-arbitrage theory predicts, so in Specification (1b), we impose the theoretical restriction $\alpha_1 = 0$. This has little effect on the estimates of the volatility coefficients, but imposing the theoretical restriction $\alpha_1 = 0$ appears to have a larger effect on the estimation of the Sharpe ratio coefficients. In addition, the p-values for the Wald tests for the volatility and Sharpe ratio equations fall below 1% and 2%, respectively. We easily reject the hypotheses that CGB Factor 1 volatility is constant and that the price of CGB

Factor 1 risk is constant.

These results display a striking similarity to those for UST Factor 1 in Table 3. The signs of the coefficients on the three term structure variables in both the volatility and Sharpe ratio functions are identical across markets. The difference is in the importance of curvature. While we dropped this variable from the UST specifications because of its statistical insignificance, in China it is by far the most significant variable in the volatility function and it also shows up significantly in the Sharpe ratio. Moreover, the magnitude of the curvature coefficient, both in an absolute sense and relative to the coefficients on level and slope, is much bigger in China. We will return to this feature of the data when examining the prices and quantities of interest rate risk below. In contrast, it is the slope of the term structure that does not appear to have significant explanatory power, and we drop this variable in Specification (1c).

For CGB Factor 2, Specification (2a) in Table 6 includes the return constant α_2 and the estimate of α_2 is again insignificant, as no-arbitrage theory predicts. In Specification (2b), we impose the theoretical restriction $\alpha_2 = 0$. The Wald test p-value is again less than 2% in testing the null hypothesis that the price of CGB Factor 2 risk is constant. For Factor 2, while the signs of the coefficients in the volatility function are the same as those in the US, the same is not true of the Sharpe ratio. For Factor 2, it is curvature rather than slope that is statistically insignificant in both equations and we drop this variable in Specification (2c).

Most importantly, we conclude that, as in the case of the UST factors, the risk premia in the CGB factors are solely compensation for risk, and both the quantities and prices of these risks vary stochastically. This confirmatory evidence from China indicates that modeling these components of bond risk premia separately, as the theory would suggest, is important for understanding the economic underpinnings of time variation in these premia.

5.2 Fitted CGB Factor Volatilities and Sharpe Ratios

Figure 7 plots the time series of fitted values of CGB Factor 1 and Factor 2 Sharpe ratios and volatilities based on GMM estimates from Table 6. Panel A plots Factor 1 fitted values from Specification (1c) of Table 6, and Panel B plots Factor 2 fitted values from Specification (2c) in the same table. In contrast to the results for the UST factors, the CGB factors exhibit negative correlations between their prices and quantities of risk. In particular, the time-series correlation between the Sharpe ratio of Factor 1 and the volatility of Factor 1 is -27%, and this same correlation for Factor 2 is -56%.

For Factor 1, these negative correlations appear to be driven by the dynamics around two periods with heavy government interventions, that of the massive post-crisis stimulus starting in 2009, and that following the stock market crash in the summer of 2015. During each of these periods, the People's Bank of China (PBoC) conducted major monetary policy interventions involving five reductions of the benchmark bank deposit and lending rates and four reductions of the bank deposit reserve requirement ratio. These interventions may have led bond market participants to anticipate significant stabilization of prices, reflected in the drop in expected volatility. At the same time, an increase in risk aversion during these periods of economic and stock market crisis may have led to an increase in the price of risk. Interestingly, it is the curvature variable, which has opposite signs in the volatility and Sharpe ratio equations, that appears to pick up this phenomenon. Of some note, the decoupling of the price and quantity of risk in both China and the US does seem to coincide with aggressive monetary policy. It is conceivable that these interventions associated with dramatic monetary easing have similar effects in both markets.

The variation in this correlation over time can be seen by computing 36-month rolling correlations between the volatility and the Sharpe ratio. In this case, this correlation varies from a high of 20% to a low of -52%. Thus, the feature of the post-Volcker period bond returns in the US that they are, at least to a first order, largely consistent with theoretical models that permit closed-form solutions for bond prices, is clearly not the case in China. In other words, the reasonable fit of these models in recent US data is not a sign that their tight restrictions are somehow a universal feature of default-free bond returns. Rather, the Chinese data strongly suggest that we need new models that accommodate the features of the data uncovered by our flexible empirical approach.

The two bond factor volatilities appear to follow a time trend broadly similar to that in the US. Specifically, both series exhibit significant declines in magnitudes over the sample period. In the US, most of this decline occurs in the latter half of the sample, which corresponds to the sample period over which we have Chinese data. By contrast, there is little or no evidence of a decline in the price of risk in China. A full exploration of the economic underpinnings of this empirical evidence is beyond the scope of this paper, but the results do show the potential of our theoretically motivated decomposition of bond risk premia to highlight important economic phenomena.

5.3 CGB Bond Volatilities, Sharpe Ratios, and Risk Premia

Following the method described in Section 4.3, we recover the annualized fitted conditional volatilities, Sharpe ratios, and risk premia of the unstandardized excess returns on the CGB two-year and ten-year zero-coupon bonds from the fitted values of the conditional volatilities and Sharpe ratios of CGB Factors 1 and 2. Figure 8 illustrates their time series.

The decline in volatility over time is perhaps not surprising given the results from the section above. The Sharpe ratios do not exhibit an obvious time trend, but they do exhibit substantial time variation. For China, the higher unconditional Sharpe ratio for shorter maturity bonds seems to be attributable to the latter part of the sample, in contrast to the result from the US. In fact, the post-crisis stimulus appears to coincide with a period when the Sharpe ratio of the ten-year zero greatly exceeded that of the two-year zero. Putting these components together, the gap between the two bonds' risk premia shows interesting variation. There are apparently substantial periods of time when the risk premia on longer-term bonds are very high compared to those on shorter-term bonds. However, this difference has all but disappeared in recent years as the higher Sharpe ratio on the two-year bond offsets its lower volatility.

5.4 Robustness

We believe we are one of the first papers to look in detail at this time variation, and, as argued earlier, we think the evidence from China brings an important element to the study of the price and quantity of risk in government bonds. The low correlation between the US and China government bond returns means that the Chinese market brings close to independent evidence to bear on potential patterns in default-free bond returns. That said, this important new evidence does bring with it some potential concerns. More specifically, are there features of the Chinese government bond market that make interpretation of the results less straightforward than in the US setting?

One concern might be that Chinese government bond returns reflect default or political risk or changes in these risks. Given that China's bond ratings have been relatively stable in our sample period, ranging from A to AA for much of the sample on the S&P scale, for example, default risk does not seem to be a large concern. Credit default swap spreads tell a similar story. With the exception of occasional spikes associated with global events such as the financial crisis, these spreads have been relatively low and stable. It is important to note that our theoretical model, and the associated empirical implementation, do not require the assets of interest to be default risk free, i.e., the model and, more specifically, the no-arbitrage condition, apply regardless of the existence of default risk. That said, such default risk, and variations therein, would clearly affect any interpretation of the empirical results.

A second and perhaps greater concern is non-stationarity in the prices and quantities of risk in these bonds as the market for Chinese government bonds has evolved over our sample period. For example, liquidity, and compensation for liquidity risk, may have changed as the market and associated trading has grown over time, or the arrival of more international investors in the latter part of the sample period may have altered pricing. It is obviously not possible to definitively reject all possible hypotheses along these lines, but the robustness checks that we did perform all suggest that these are not major concerns. First, we checked for general instability over the sample period using an indicator variable that took on the value 1 in the second half of the sample, analogous to our ZLB indicator in the US. Using this variable, there is no statistically significant evidence that the functions we estimate differ in the latter half of the sample. Second, we added two additional variables as predictors, government bond issuance and trading volume. These variables could pick up liquidity effects, supply effects more generally, or increases in trading and associated pricing effects associated with the entrance of new participants. Neither of these variables shows statistically significant predictive power for the volatility or Sharpe ratio of either Factor 1 or 2. To summarize, we can find no evidence of instability in the specifications that we estimate.

At the same time, stability of the parameters does not imply that political risk or the evolution of the market are unimportant for the time series properties of the price and quantity of risk in the CGB market. As we argued above in a slightly different context, as long as the relation between yields and these quantities is stable, then the model can accommodate changes in the economic environment. In other words, we cannot reject that the joint dynamics of volatilities and Sharpe ratios in the CGB market are determined, at least in part, by the special features of this market. Nevertheless, we still believe that evidence from the CGB market is of substantial interest.

6 Conclusion

Government bonds are a critical component of many investors' portfolios, in some cases even more critical than equities. However, the academic literature on the risk and return of these bonds has not evolved to answer a number of key questions. Some of these studies neglect consideration of risk altogether, which is a significant omission in fixed income markets where expected returns can be levered almost arbitrarily. Other studies impose restrictive functional forms on the relation between the price and quantity of risk.

Our paper advances the literature by providing critical empirical insights in a more flexible framework, while still imposing no arbitrage by restricting risk premia to be linear in risk. We decompose risk premia into two components: the quantity of risk (volatility) and the price of that risk (the Sharpe ratio). Our focus on Sharpe ratios reveals the existence of two important factors in government bond returns in both the US and China. For both factors and in both countries, the quantity and price of risk vary over time in important ways.

The factor structure of risk premia in the Chinese government bond market is broadly similar to that in the US Treasury market, despite the fact that for much of the sample the bond market in China was effectively segregated from the bond market in the US. This independent evidence lends credence to the argument that we have uncovered fundamental structural components of bond risk premia. However, the quantity and price of risk exhibit a negative unconditional correlation for the second factor in the US and for both factors in China. Moreover, these correlations vary significantly over time. For example, the ZLB period in the US and periods of significant government intervention in China, associated with the 2008 financial crisis and the 2015 stock market meltdown, exhibit a strong negative correlation between the price and quantity of interest rate risk, but these components are positively correlated in other periods. This time variation is difficult, if not impossible, to accommodate in the existing theoretical models that generate closed-form solutions for bond prices. Thus, the fact that these models seem to fit the post-Volcker US data well should not be construed as indicating that they are sufficiently flexible to fit default-free bond returns in general.

References

Adrian, Tobias, Richard K Crump, and Emanuel Moench, 2013, Pricing the term structure with linear regressions, *Journal of Financial Economics* 110, 110–138.

Amstad, Marlene, and Zhiguo He, 2018, Handbook of China's Financial System Chapter 6: Chinese bond market and interbank market.

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

Balduzzi, Pierluigi, Mike Connolly, and Alan Marcus, 2021, Greek-based bond factors, Unpublished working paper. Boston College.

Bauer, Michael D, and James D Hamilton, 2018, Robust bond risk premia, *Review of Financial Studies* 31, 399–448.

Boudoukh, Jacob, Christopher Downing, Matthew Richardson, Richard Stanton, and Robert Whitelaw, 2010, A multifactor, nonlinear, continuous-time model of interest rate volatility, *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle* 296–322.

Boudoukh, Jacob, Ronen Israel, and Matthew Richardson, 2021, Biases in long-horizon predictive regressions, *Journal of Financial Economics* forthcoming.

Boudoukh, Jacob, and Matthew Richardson, 1994, The statistics of long-horizon regressions revisited, *Mathematical Finance* 4, 103–119.

Boudoukh, Jacob, Matthew Richardson, and Robert F Whitelaw, 2008, The myth of long-horizon predictability, *Review of Financial Studies* 21, 1577–1605.

Brooks, Jordan, and Tobias J Moskowitz, 2017, Yield curve premia, Unpublished working paper. Yale University.

Campbell, John Y, 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373–399.

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

Chan, Kalok C, G Andrew Karolyi, Francis A Longstaff, and Anthony B Sanders, 1992, An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance* 47, 1209–1227.

Cheridito, Patrick, Damir Filipović, and Robert L Kimmel, 2007, Market price of risk specifications for affine models: Theory and evidence, *Journal of Financial Economics* 83, 123–170.

Cieslak, Anna, and Pavol Povala, 2015, Expected returns in Treasury bonds, *Review of Financial Studies* 28, 2859–2901.

Cieslak, Anna, and Pavol Povala, 2016, Information in the term structure of yield curve volatility, *Journal of Finance* 71, 1393–1436.

Clayton, Christopher, Amanda Dos Santos, Matteo Maggiori, and Jesse Schreger, 2022, Internationalizing like China, *Available at SSRN* .

Cochrane, John H, 2005, *Asset Pricing*, revised edition (Princeton University Press, Princeton, New Jersey).

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

Collin-Dufresne, Pierre, and Robert Goldstein, 2002, Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility, *Journal of Finance* 57, 1685–1730.

Cox, John C, Jonathan E Ingersoll, and Stephen A Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Creal, Drew D, and Jing Cynthia Wu, 2017, Monetary policy uncertainty and economic fluctuations, *International Economic Review* 58, 1317–1354.

Creal, Drew D, and Jing Cynthia Wu, 2020, Bond risk premia in consumption-based models, *Quantitative Economics* 11, 1461–1484.

Dai, Qiang, and Kenneth J Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

Duffee, Gregory R, 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 405–443.

Duffee, Gregory R, 2010, Sharpe ratios in term structure models, Unpublished working paper. Johns Hopkins University.

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

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Engle, Robert F, David M Lilien, and Russell P Robins, 1987, Estimating time varying risk premia in the term structure: The ARCH-M model, *Econometrica* 55, 391–407.

Fama, Eugene F, 1986, Term premiums and default premiums in money markets, *Journal of Financial Economics* 17, 175–196.

Fama, Eugene F, and Robert R Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680–692.

Filipović, Damir, Martin Larsson, and Anders B Trolle, 2017, Linear-rational term structure models, *Journal of Finance* 72, 655–704.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

French, Kenneth R, G William Schwert, and Robert F Stambaugh, 1987, Expected stock returns and volatility, *Journal of Financial Economics* 19, 3–29.

Ghysels, Eric, Anh Le, Sunjin Park, and Haoxiang Zhu, 2014, Risk and return trade-off in the US Treasury market, *Available at SSRN 2403483* .

Glosten, Lawrence R, Ravi Jagannathan, and David E Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779–1801.

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

Haddad, Valentin, Alan Moreira, and Tyler Muir, 2025, Whatever it takes? the impact of conditional policy promises, *American Economic Review* 115, 295–329.

Joslin, Scott, and Anh Le, 2021, Interest rate volatility and no-arbitrage affine term structure models, *Management Science* 67, 7391–7416.

Joslin, Scott, Anh Le, and Kenneth J Singleton, 2013, Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs, *Journal of Financial Economics* 109, 604–622.

Joslin, Scott, Marcel Priebsch, and Kenneth J Singleton, 2014, Risk premiums in dynamic term structure models with unspanned macro risks, *Journal of Finance* 69, 1197–1233.

Joslin, Scott, Kenneth J Singleton, and Haoxiang Zhu, 2011, A new perspective on Gaussian dynamic term structure models, *Review of Financial Studies* 24, 926–970.

Karatzas, Ioannis, and Steven E Shreve, 1998, *Methods of Mathematical Finance*, volume 39 (Springer).

Lettau, Martin, and Jessica A Wachter, 2011, The term structures of equity and interest rates, *Journal of Financial Economics* 101, 90–113.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* 96, 175–194.

Li, Junye, Lucio Sarno, and Gabriele Zinna, 2021, Risks and risk premia in the US Treasury market, *Available at SSRN 3640341* .

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

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

Merton, Robert C, 1980, On estimating the expected return on the market, *Journal of Financial Economics* 8, 323–361.

Stanton, Richard, 1997, A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance* 52, 1973–2002.

Tang, Yi, and Robert F Whitelaw, 2011, Time-varying Sharpe ratios and market timing, *Quarterly Journal of Finance* 1, 465–493.

Whitelaw, Robert F, 1994, Time variations and covariations in the expectation and volatility of stock market returns, *Journal of Finance* 49, 515–541.

Wright, Jonathan H, 2011, Term premia and inflation uncertainty: Empirical evidence from an international panel dataset, *American Economic Review* 101, 1514–34.

Wu, Jing Cynthia, and Fan Dora Xia, 2016, Measuring the macroeconomic impact of monetary policy at the zero lower bound, *Journal of Money, Credit and Banking* 48, 253–291.

Table 1: Factor Structure and Performance of UST and CGB Implied Zero Excess Returns
The factor structure of US Treasury and Chinese Government Bond implied zero excess returns in Panel A, and their unconditional means, volatilities, and Sharpe ratios in Panel B. All quantities are annualized. Means and volatilities are in percent. Panel A shows the factor structure of the standardized excess zero returns based on PCAs of their 10×10 correlation matrix for each subperiod and market. For each subperiod and market, Panel A contains results for the first three principal components, F1, F2, and F3. Factor Var. as % of Tot. is the factor's eigenvalue expressed as a percent of the sum of all ten eigenvalues from the PCA. Factor Vol and SR are the volatility and Sharpe ratio of each factor portfolio, constructed with holdings in the standardized zeroes given by the eigenvector for the factor. The column-vector of standardized zero loadings under each factor is the factor eigenvector.

	UST Implied Zeroes						CGB Implied Zeroes		
	7/1976–12/1989			1/1990–12/2022			5/2004–12/2022		
A. Factor Structure	F1	F2	F3	F1	F2	F3	F1	F2	F3
Factor Var. as % of Tot.	94.69	4.07	0.73	90.85	7.18	1.42	82.36	13.54	2.46
Factor Vol	10.66	2.21	0.93	10.44	2.93	1.31	9.94	4.03	1.72
Factor SR	0.27	0.45	0.45	0.64	0.73	0.51	0.51	0.14	0.02
1-year zero loadings	0.30	0.57	0.43	0.26	0.66	0.65	0.25	0.54	-0.60
2-year zero loadings	0.31	0.40	0.11	0.31	0.41	-0.24	0.29	0.45	-0.11
3-year zero loadings	0.32	0.27	-0.14	0.32	0.24	-0.37	0.32	0.30	0.25
4-year zero loadings	0.32	0.15	-0.29	0.33	0.09	-0.33	0.34	0.16	0.38
5-year zero loadings	0.32	0.02	-0.39	0.33	-0.03	-0.23	0.34	0.02	0.39
6-year zero loadings	0.32	-0.14	-0.33	0.33	-0.12	-0.11	0.34	-0.11	0.23
7-year zero loadings	0.32	-0.26	-0.19	0.33	-0.20	0.01	0.33	-0.22	0.00
8-year zero loadings	0.32	-0.31	0.02	0.32	-0.26	0.13	0.33	-0.28	-0.14
9-year zero loadings	0.32	-0.34	0.27	0.32	-0.31	0.25	0.32	-0.33	-0.26
10-year zero loadings	0.31	-0.34	0.58	0.31	-0.34	0.36	0.30	-0.36	-0.36
B. Performance Measures	Mean	Vol	SR	Mean	Vol	SR	Mean	Vol	SR
1-year zero	1.40	2.51	0.56	0.69	0.66	1.05	0.37	0.88	0.42
2-year zero	1.56	4.70	0.33	1.30	1.66	0.78	0.81	1.51	0.54
3-year zero	1.68	6.34	0.26	1.76	2.70	0.65	1.08	2.09	0.52
4-year zero	1.94	8.07	0.24	2.33	3.71	0.63	1.33	2.69	0.49
5-year zero	2.26	9.71	0.23	2.76	4.70	0.59	1.61	3.35	0.48
6-year zero	2.61	11.11	0.23	3.24	5.62	0.58	2.02	3.87	0.52
7-year zero	2.60	12.43	0.21	3.38	6.51	0.52	1.95	4.45	0.44
8-year zero	2.73	13.55	0.20	3.63	7.37	0.49	2.11	5.03	0.42
9-year zero	2.83	14.46	0.20	3.79	8.24	0.46	2.30	5.64	0.41
10-year zero	2.84	15.25	0.19	3.77	9.11	0.41	2.43	6.29	0.39

Table 2: Factor Structure and Performance of UST ETF Excess Returns

The factor structure of US Treasury ETF excess returns gross of 15-basis-point annual fees in Panel A, and their unconditional means, volatilities, and Sharpe ratios in Panel B. The sample period is 2/2007–12/2022. All quantities are annualized. Means and volatilities are in percent. Panel A shows the factor structure of the standardized excess ETF returns based on PCAs of their 6×6 correlation matrix. Panel A contains results for the first three principal components, F1, F2, and F3. Factor Var. as % of Tot. is the factor's eigenvalue expressed as a percent of the sum of all six eigenvalues from the PCA. Factor Vol and SR are the volatility and Sharpe ratio of each factor portfolio, constructed with holdings in the standardized ETFs given by the eigenvector for the factor. The column-vector of standardized ETF loadings under each factor is the factor eigenvector.

A. Factor Structure	F1	F2	F3
Factor Var. as % of Tot.	78.53	14.79	5.17
Factor Vol	7.52	3.26	1.93
Factor SR	0.58	0.79	0.44
0-1-year ETF	0.30	0.74	0.59
1-3-year ETF	0.40	0.39	-0.51
3-7-year ETF	0.44	0.05	-0.43
7-10-year ETF	0.45	-0.20	-0.06
10-20-year ETF	0.43	-0.33	0.21
>20-year ETF	0.41	-0.40	0.39
B. Performance Measures	**Mean**	**Vol**	**SR**
0-1-year ETF	0.26	0.24	1.08
1-3-year ETF	0.79	1.33	0.59
3-7-year ETF	2.07	3.77	0.55
7-10-year ETF	2.87	6.57	0.44
10-20-year ETF	3.01	9.58	0.31
>20-year ETF	4.07	14.02	0.29

Table 3: GMM Estimates of UST Factor Dynamics

GMM estimates of α_j, β_j^σ, β_j^θ, and their robust z-statistics for alternative specifications of the system

$$R_{j,t+1} = \alpha_j + (X_t\beta_j^\sigma)(X_t\beta_j^\theta) + u_{j,t+1} \ ,$$

$$\sqrt{\frac{\pi}{2}}|u_{j,t+1}| = X_t\beta_j^\sigma + v_{j,t+1} \ .$$

The sample period is 1/1990–12/2022. R_1 and R_2 are the monthly returns on the first and second principal-component UST factor portfolios. Results for Factor 1 are on the left, results for Factor 2 are on the right. X_t is the vector of (lagged) predictor variables indicated by the row titles. Level = $10Y_2$, Slope = $10(Y_{10} - Y_2)$, and Curvature = $10(Y_6 - \frac{Y_2+Y_{10}}{2})$, where Y_T is the yield on the T-year zero, for T = 2, 6, and 10. VIX is an index of the implied volatility of the 30-day return on the S&P 500 derived from S&P 500 index options. Realized Vol is the absolute value of the factor return times $\sqrt{12\frac{\pi}{2}}$. Wald test (1) tests the null hypothesis that factor volatility is constant, i.e., $\beta_{j,1}^\sigma = \beta_{j,2}^\sigma = \cdots = \beta_{j,k}^\sigma = 0$. Wald test (2) tests the null hypothesis that the factor price of risk is constant, i.e., $\beta_{j,1}^\theta = \cdots = \beta_{j,k}^\theta = 0$. Goodness-of-fit (1) = 1-$\frac{\sum_t v_{j,t}^2}{\frac{\pi}{2}\sum_t(|u_{j,t}|-\bar{|u_j|})^2}$. Goodness-of-fit (2) = 1-$\frac{\sum_t u_{j,t}^2}{\sum_t(R_{j,t}-\bar{R_j})^2}$.

	UST Factor 1			UST Factor 2	
	(1a)	(1b)	(1c)	(2a)	(2b)
Volatility Coefficients (β_j^σ)					
Constant	0.51	0.26	0.74	0.13	0.04
	(1.13)	(0.57)	(1.99)	(1.00)	(0.32)
Level	2.11	2.50	2.21	0.41	0.42
	(3.40)	(3.80)	(4.18)	(2.47)	(2.46)
Slope	4.99	5.66	4.36	2.10	1.89
	(2.21)	(2.28)	(2.84)	(2.51)	(2.38)
Curvature	-3.99	-4.55		-6.72	-6.50
	(-0.50)	(-0.49)		(-1.89)	(-2.05)
VIX/100	5.13	5.18	4.27	1.47	1.81
	(3.44)	(2.88)	(3.14)	(2.83)	(3.60)
Realized Vol	0.44	0.55		2.89	3.97
	(0.35)	(0.37)		(1.90)	(2.51)
Sharpe Ratio Coefficients (β_j^θ)					
Constant	-3.07	-0.46	-0.42	2.13	-0.39
	(-1.19)	(-2.04)	(-2.11)	(1.11)	(-1.86)
Level	1.91	0.73	0.71	-0.10	0.83
	(1.74)	(2.40)	(2.98)	(-0.12)	(2.62)
Slope	4.62	1.87	1.91	-3.85	0.70
	(1.60)	(1.51)	(2.61)	(-1.09)	(0.53)
Curvature	-1.76	0.59		12.92	-1.63
	(-0.27)	(0.13)		(1.08)	(-0.31)
VIX/100	3.59	1.26	0.71	-0.77	1.39
	(1.47)	(1.69)	(1.03)	(-0.46)	(2.19)
Realized Vol	-0.39	-0.71		-3.21	1.18
	(-0.51)	(-1.06)		(-0.71)	(0.61)
Return Constant (α_j)	3.87			-1.01	
	(0.95)			(-1.22)	
No. Moment Conditions	27	27	14	27	27
J-stat p-value (in %)	54.71	45.61	81.47	34.30	12.87
Wald test (1) p-value (in %)	0.01	0.00	0.00	0.04	0.00
Wald test (2) p-value (in %)	43.21	2.73	0.99	41.58	0.97
Goodness-of-fit (1) (in %)	5.45	6.32	6.53	8.85	11.64
Goodness-of-fit (2) (in %)	7.09	6.41	5.03	5.57	1.87

Table 4: GMM Estimates of UST Factor Dynamics with ZLB Effects

GMM estimates of α_j, β_j^σ, β_j^θ, and their robust z-statistics for alternative specifications of the system

$$R_{j,t+1} = \alpha_j + \delta_j L_t + (X_t\beta_j^\sigma + L_t X_t \gamma_j^\sigma)(X_t\beta_j^\theta + L_t X_t \gamma_j^\theta) + u_{j,t+1} \ ,$$

$$\sqrt{\frac{\pi}{2}}|u_{j,t+1}| = X_t\beta_j^\sigma + L_t X_t\gamma_j^\sigma + v_{j,t+1} \ .$$

The sample period is 1/1990–12/2022. R_1 and R_2 are the monthly returns on the first and second principal-component UST factor portfolios. Results for Factor 1 are on the left, results for Factor 2 are on the right. X_t is the vector of predictor variables indicated by the row titles. Level = $10Y_2$, Slope = $10(Y_{10} - Y_2)$, and Curvature = $10(Y_6 - \frac{Y_2+Y_{10}}{2})$, where Y_T is the yield on the T-year zero, for T = 2, 6, and 10. VIX is an index of the implied volatility of the 30-day return on the S&P 500 derived from S&P 500 index options. Realized Vol is the absolute value of the factor return times $\sqrt{12\frac{\pi}{2}}$. L_t is an indicator variable that takes the value one if the effective Fed funds rate at time t is below 25 basis points and zero otherwise. Wald test (1) tests the null hypothesis that factor volatility is constant during the non-ZLB period. Wald test (2) tests the null hypothesis that the factor price of risk is constant during the non-ZLB period. Wald test (3) tests the null hypothesis that factor volatility is different in the ZLB period. Wald test (4) tests the null hypothesis that the factor price of risk is different in the ZLB period. Goodness-of-fit (1) = 1-$\frac{\sum_t v_{j,t}^2}{\frac{\pi}{2}\sum_t(|u_{j,t}|-\overline{|u_j|})^2}$. Goodness-of-fit (2) = 1-$\frac{\sum_t u_{j,t}^2}{\sum_t(R_{j,t}-\bar{R}_j)^2}$.

	UST Factor 1		UST Factor 2	
	Non-ZLB	ZLB Diff	Non-ZLB	ZLB Diff
Volatility Coefficients (β_j^σ and γ_j^σ)				
Constant	1.25	-1.28	-0.04	-0.11
	(2.74)	(-1.69)	(-0.25)	(-0.38)
Level	0.79	7.85	0.30	-4.45
	(1.12)	(1.17)	(1.30)	(-1.26)
Slope	4.76	0.35	2.55	4.92
	(2.64)	(0.11)	(1.79)	(2.66)
Curvature			-9.50	-5.79
			(-1.66)	(-0.68)
VIX/100	5.14	-1.85	2.66	-2.77
	(2.65)	(-0.71)	(3.38)	(-2.59)
Realized Vol			3.81	-4.64
			(1.97)	(-1.27)
Sharpe Ratio Coefficients (β_j^θ and γ_j^θ)				
Constant	-0.47	-0.29	-0.29	1.96
	(-2.11)	(-0.62)	(-1.29)	(2.14)
Level	0.55	-3.63	0.80	17.87
	(1.68)	(-0.96)	(2.03)	(1.76)
Slope	1.64	3.77	-0.41	-11.48
	(2.00)	(1.88)	(-0.22)	(-2.24)
Curvature			5.46	-0.16
			(0.74)	(-0.01)
VIX/100	1.53	-1.61	0.62	-0.32
	(1.71)	(-1.13)	(0.81)	(-0.11)
Realized Vol			1.66	-0.21
			(0.82)	(-0.03)
No. Moment Conditions	22		39	
J-stat p-value (in %)	47.81		38.10	
Wald test (1) p-value (in %)	0.07		0.11	
Wald test (2) p-value (in %)	1.86		3.57	
Wald test (3) p-value (in %)		57.30		0.00
Wald test (4) p-value (in %)		12.58		2.49
Goodness-of-fit (1) (in %)	9.60		4.57	
Goodness-of-fit (2) (in %)	7.29		2.84	

Table 5: R^2's in Monthly and Annual Return Regressions

The table compares regression results using monthly returns with those using overlapping annual returns. Panel A shows the coefficients, t-statistics, and R^2 from the first-stage regression of realized volatility, measured as $\sqrt{\frac{\pi}{2}}|R_{1,t+1}|$, on the indicated predictor variables. R_1 is the return on UST Factor 1. Level = $10Y_2$ and Slope = $10(Y_{10} - Y_2)$, where Y_T is the yield on the T-year zero for $T = 2$ and 10. Panel B shows the coefficients, t-statistics, and R^2 from the autoregression of fitted volatility values, Volhat, from the first-stage regression. Panel C shows the coefficients, t-statistics, and R^2 from the second-stage regression of UST Factor 1 monthly returns on Volhat. Panel D shows the coefficients, t-statistics, and R^2 from the second-stage regression of UST Factor 1 annual, overlapping returns on Volhat. Panel E shows the coefficient and R^2 for the second-stage annual, overlapping return regression implied by the model of Boudoukh et al. (2008). Ordinary-least-squares t-statistics are in parenthesis, Newey-West t-statistics are in brackets, and R^2's are in percent.

A. First-stage volatility regression				
Constant	Level	Slope	VIX/100	R^2
-0.19	3.08	7.72	6.19	11.82
(-0.40)	(4.80)	(4.78)	(3.79)	
B. Autoregression of fitted volatility				
Constant	Volhat	R^2		
0.18	0.94	87.19		
(3.03)	(49.29)			
C. Second-stage monthly return regression				
Constant	Volhat	R^2		
-1.55	0.74	4.32		
(-2.71)	(4.02)			
[-3.00]	[4.25]			
D. Second-stage annual return regression				
Constant	Volhat	R^2		
-10.28	6.13	19.12		
(-4.84)	(9.06)			
[-2.23]	[4.22]			
E. BRW-implied annual return regression				
	Volhat	R^2		
	6.43	27.01		

Table 6: GMM Estimates of CGB Factor Dynamics

GMM estimates of α_j, β_j^σ, β_j^θ, and their robust z-statistics for alternative specifications of the system

$$R_{j,t+1} = \alpha_j + (X_t\beta_j^\sigma)(X_t\beta_j^\theta) + u_{j,t+1} \ ,$$

$$\sqrt{\frac{\pi}{2}}|u_{j,t+1}| = X_t\beta_j^\sigma + v_{j,t+1} \ .$$

The sample period is 5/2004–12/2022. R_1 and R_2 are the monthly returns on the first and second principal-component CGB factor portfolios. Results for Factor 1 are on the left, results for Factor 2 are on the right. X_t is the vector of predictor variables indicated by the row titles. Level = $10Y_2$, Slope = $10(Y_{10} - Y_2)$, and Curvature = $10(Y_6 - \frac{Y_2+Y_{10}}{2})$, where Y_T is the yield on the T-year zero, for T = 2, 6, and 10. Realized Vol is the absolute value of the factor return times $\sqrt{12\frac{\pi}{2}}$. Wald test (1) tests the null hypothesis that factor volatility is constant, i.e., $\beta_{j,1}^\sigma = \beta_{j,2}^\sigma = \cdots = \beta_{j,k}^\sigma = 0$. Wald test (2) tests the null hypothesis that the factor price of risk is constant, i.e., $\beta_{j,1}^\theta = \cdots = \beta_{j,k}^\theta = 0$. Goodness-of-fit (1) = 1-$\frac{\sum_t v_{j,t}^2}{\frac{\pi}{2}\sum_t(|u_{j,t}|-\overline{|u_j|})^2}$. Goodness-of-fit (2) = 1-$\frac{\sum_t u_{j,t}^2}{\sum_t(R_{j,t}-\bar{R}_j)^2}$.

	CGB Factor 1			CGB Factor 2		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
Volatility Coefficients (β_j^σ)						
Constant	2.24	2.79	3.34	-0.14	-0.04	-0.31
	(2.15)	(2.68)	(4.20)	(-0.35)	(-0.09)	(-0.84)
Level	-0.25	-0.99	-2.29	2.06	1.95	2.57
	(-0.08)	(-0.35)	(-0.99)	(1.88)	(1.70)	(2.28)
Slope	7.22	4.31		5.66	5.34	5.28
	(1.71)	(0.86)		(4.05)	(3.92)	(3.91)
Curvature	-44.88	-50.98	-46.48	-7.02	-7.35	
	(-2.85)	(-3.18)	(-3.07)	(-1.04)	(-1.01)	
Realized Vol	3.42	2.95	4.06	6.42	4.80	5.14
	(1.90)	(1.52)	(2.29)	(3.18)	(2.13)	(2.61)
Sharpe Ratio Coefficients (β_j^θ)						
Constant	-2.89	-1.34	-1.15	1.26	0.83	0.45
	(-1.81)	(-2.32)	(-2.67)	(1.05)	(1.68)	(0.95)
Level	3.87	3.88	3.53	-1.90	-1.19	-0.27
	(1.94)	(2.46)	(2.72)	(-1.12)	(-0.84)	(-0.20)
Slope	3.49	0.73		-4.80	-3.93	-4.25
	(0.92)	(0.32)		(-1.50)	(-2.15)	(-2.32)
Curvature	-5.09	21.28	21.10	-13.61	-11.15	
	(-0.22)	(2.00)	(1.95)	(-1.74)	(-1.53)	
Realized Vol	2.81	1.24	1.22	1.47	2.30	1.26
	(2.04)	(1.26)	(1.35)	(0.49)	(1.25)	(0.83)
Return Constant (α_j)	3.55			-0.12		
	(1.18)			(-0.22)		
No. Moment Conditions	20	20	14	20	20	14
J-stat p-value (in %)	14.34	24.72	8.99	57.51	59.58	66.19
Wald test (1) p-value (in %)	1.99	0.96	0.31	0.00	0.00	0.00
Wald test (2) p-value (in %)	3.50	1.43	1.39	0.49	1.99	4.67
Goodness-of-fit (1) (in %)	8.68	5.79	4.91	16.20	15.26	14.08
Goodness-of-fit (2) (in %)	7.05	6.88	6.32	6.14	6.46	4.26

Figure 1: China Bond Market Capitalization in Trillion CNY

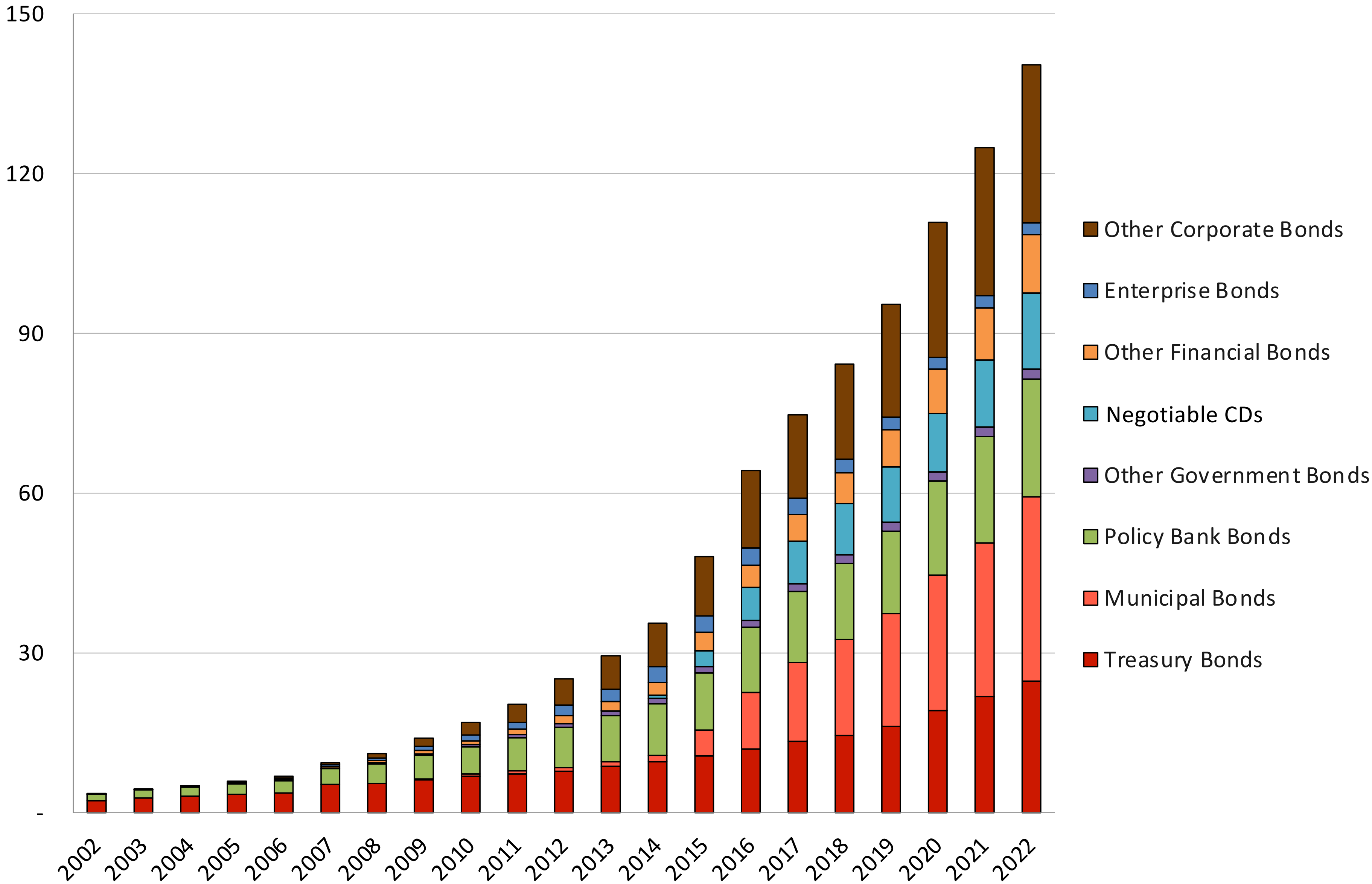
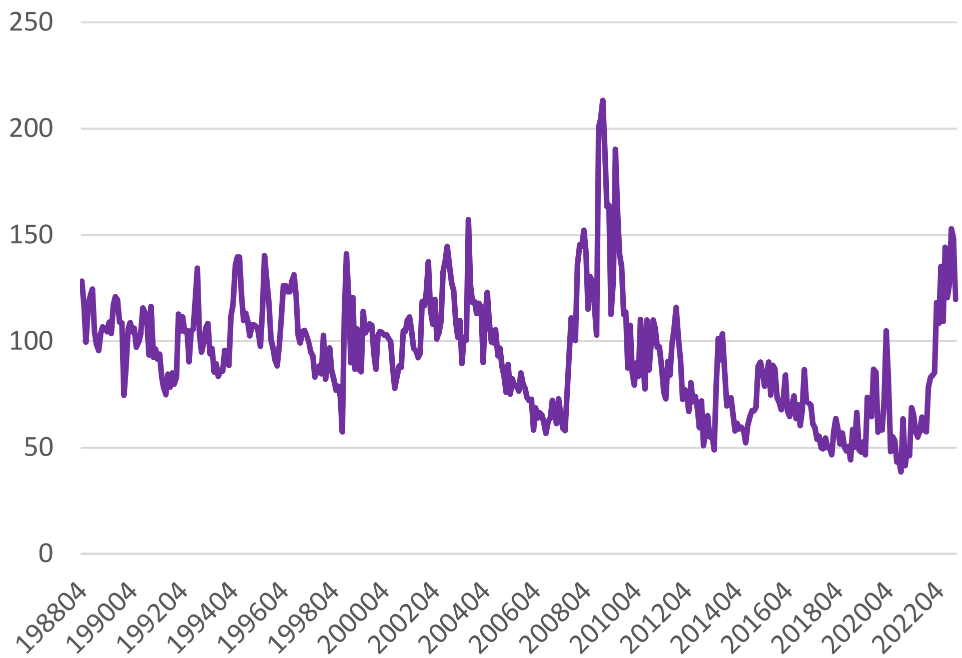
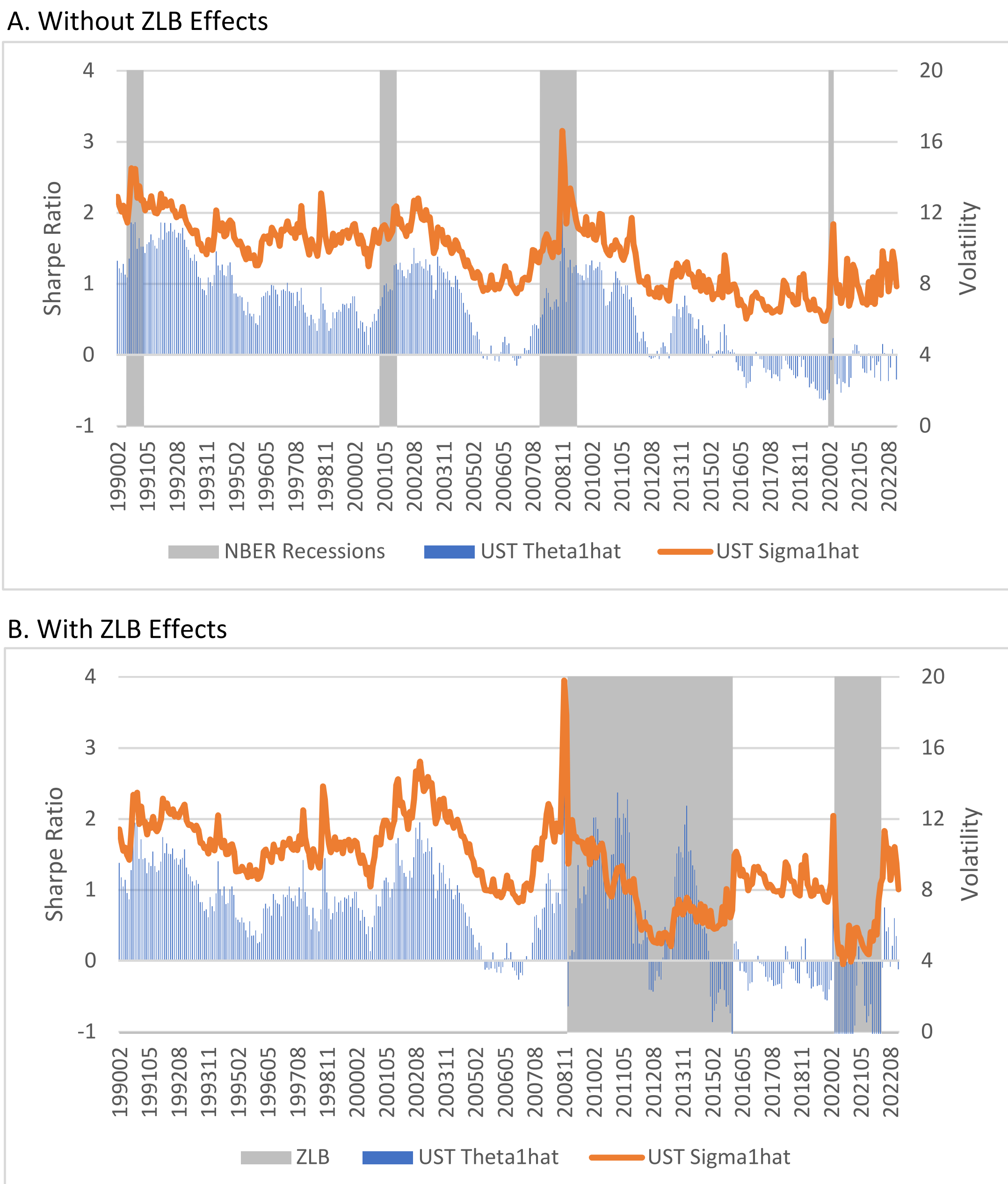


Figure 2: MOVE Index of Implied Yield Volatility from One-Month Treasury Bond Options



MOVE measures the implied yield volatility of a basket of one-month over-the-counter options on 2-year, 5-year, 10-year, and 30-year Treasuries.

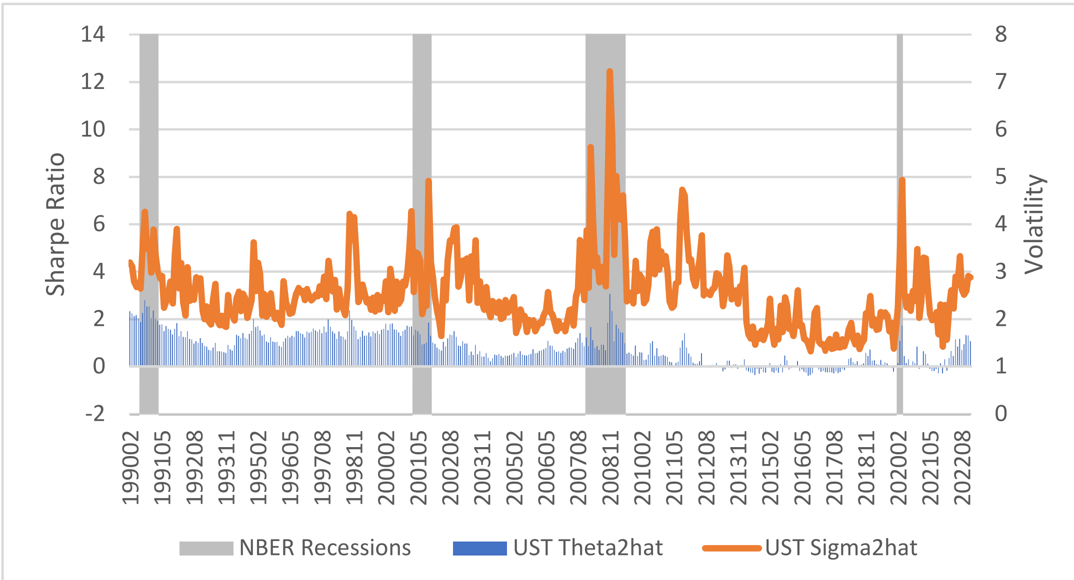
Figure 3: UST Factor 1 Dynamics



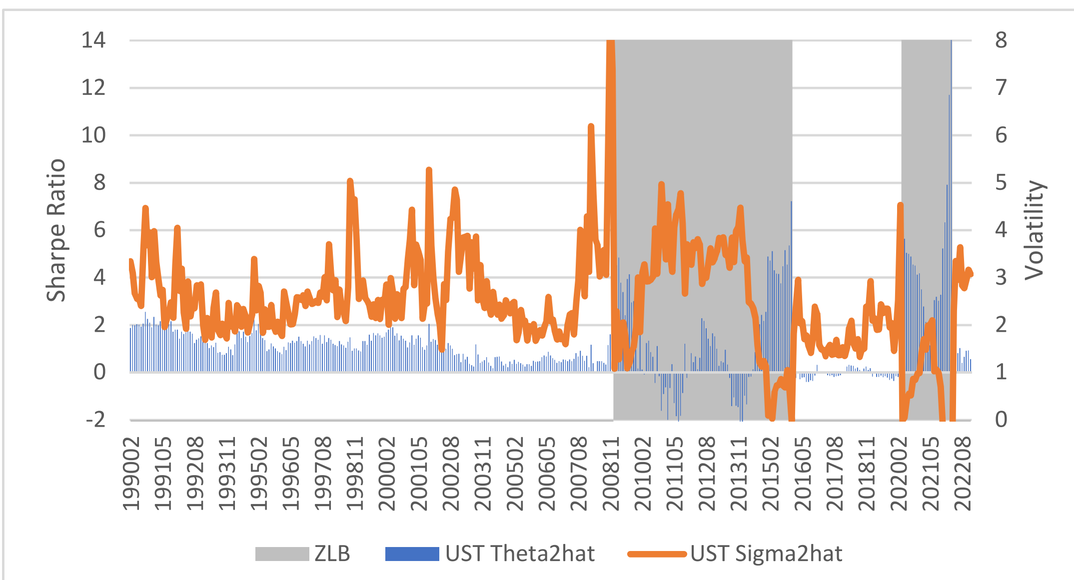
Time series of annualized fitted values of UST Factor 1 Sharpe ratios and volatilities based on GMM estimates of factor dynamics from Specifications (1c) of Table 3 and Table 4, respectively.

Figure 4: UST Factor 2 Dynamics

A. Without ZLB Effects

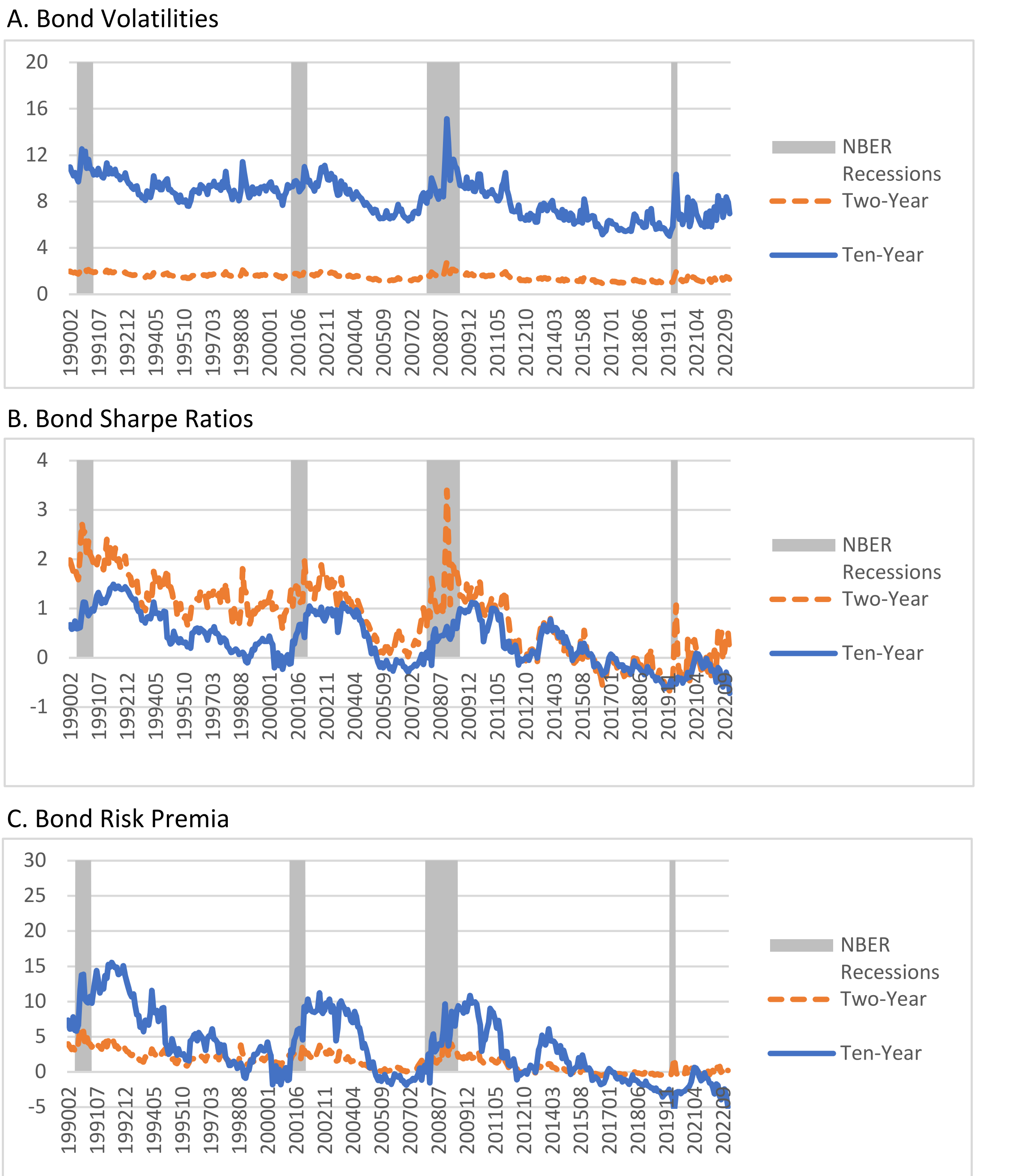


B. With ZLB Effects



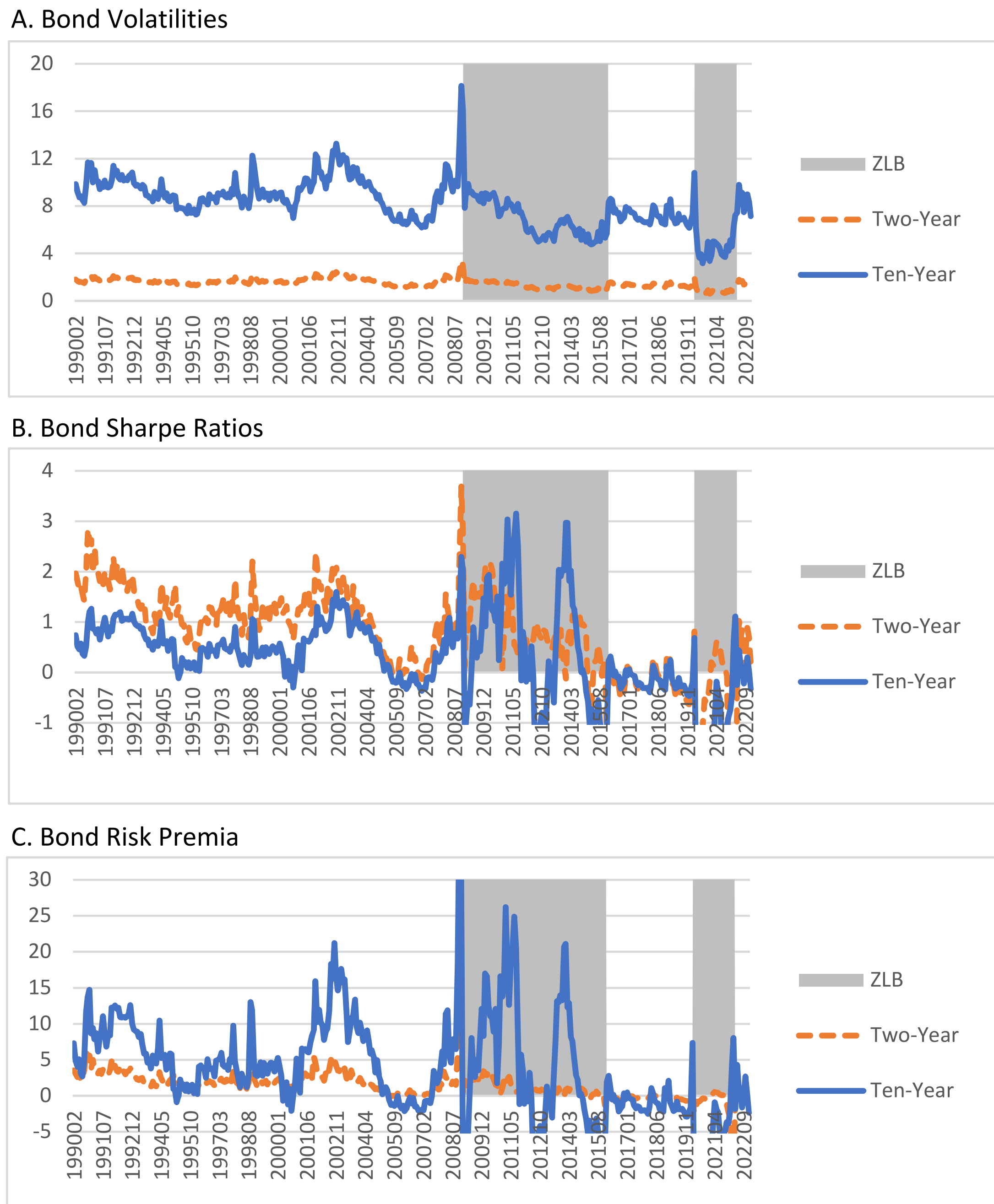
Time series of annualized fitted values of UST Factor 2 Sharpe ratios and volatilities based on GMM estimates of factor dynamics from Specifications (2b) of Table 3 and Table 4, respectively.

Figure 5: UST Bond Dynamics



Time series of annualized fitted volatilities, Sharpe ratios, and risk premia of UST implied 2-year and 10-year zero-coupon bonds, based on the fitted values of UST Factor 1 and Factor 2 Sharpe ratios and volatilities together with the loadings of the standardized excess returns of the zeroes on the factors and the unconditional volatilities of the zero excess returns from Table 1.

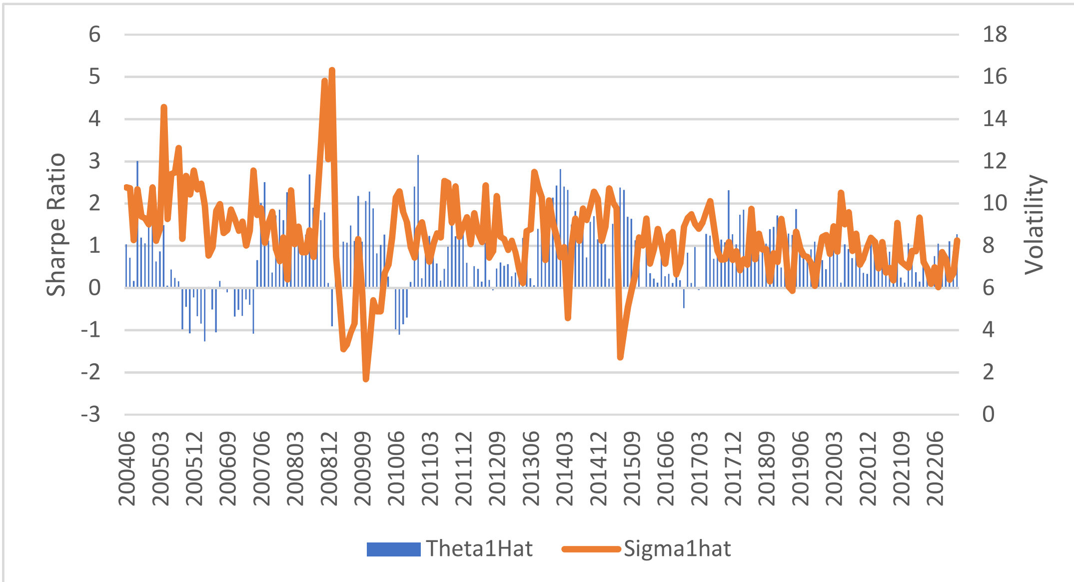
Figure 6: UST Bond Dynamics with ZLB Effects



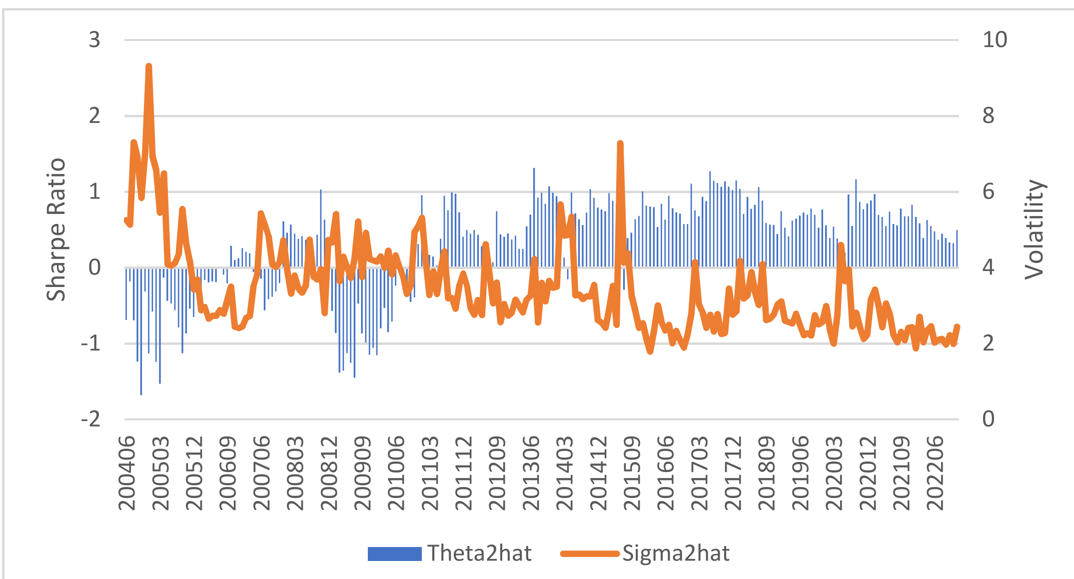
Time series of annualized fitted volatilities, Sharpe ratios, and risk premia of UST implied 2-year and 10-year zero-coupon bonds, based on the fitted values of UST Factor 1 and Factor 2 Sharpe ratios and volatilities, with ZLB effects, together with the loadings of the standardized excess returns of the zeroes on the factors and the unconditional volatilities of the zero excess returns from Table 1.

Figure 7: CGB Factor Dynamics

A. Factor 1

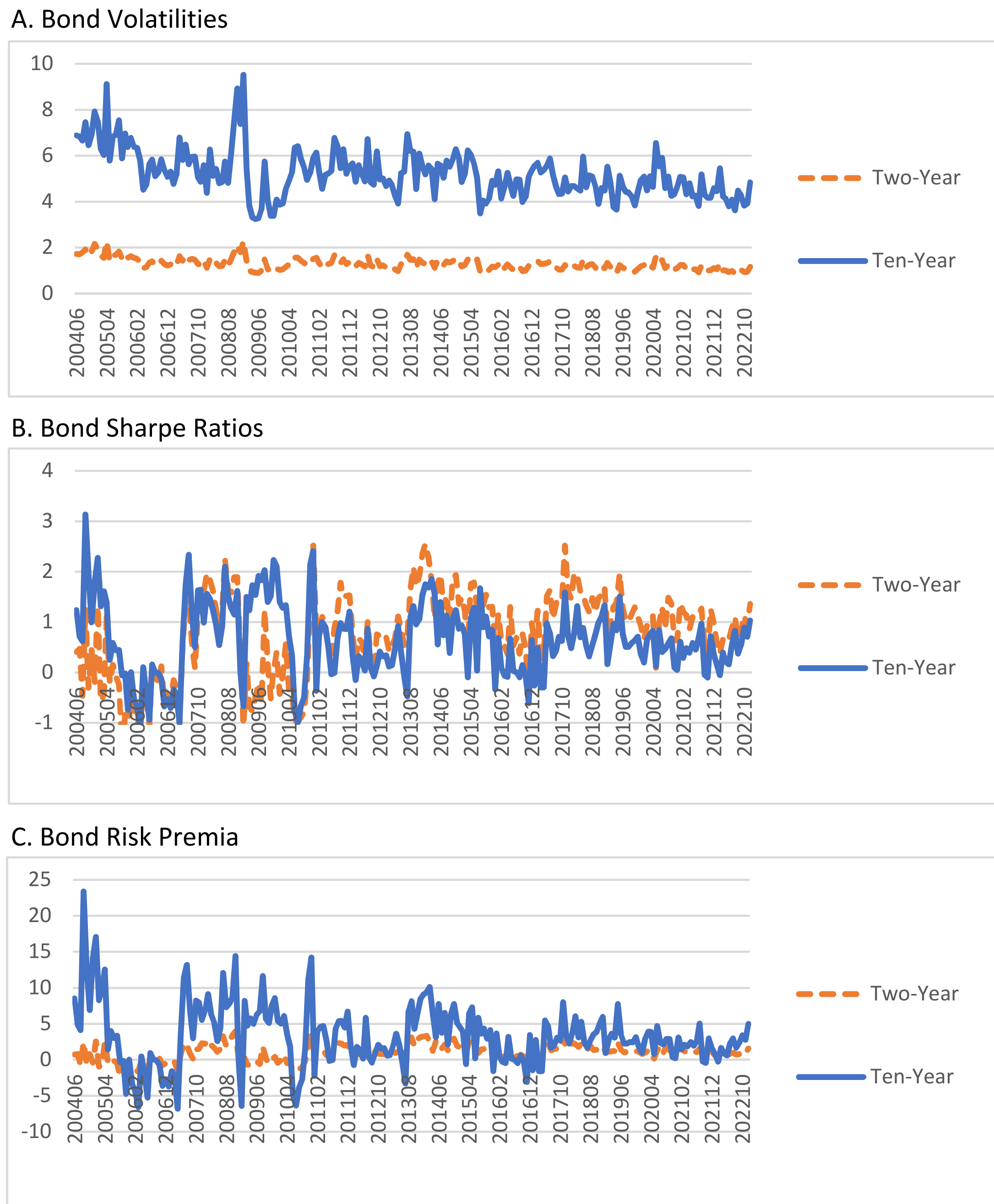


B. Factor 2



Time series of annualized fitted values of CGB Factor 1 and Factor 2 Sharpe ratios and volatilities based on GMM estimates of factor dynamics from Specifications (1b) and (2b) of Table 6, respectively.

Figure 8: CGB Bond Dynamics



Time series of annualized fitted volatilities, Sharpe ratios, and risk premia of CGB implied 2-year and 10-year zero-coupon bonds, based on the fitted values of CGB Factor 1 and Factor 2 Sharpe ratios and volatilities together with the loadings of the standardized excess returns of the zeroes on the factors and the unconditional volatilities of the zero excess returns from Table 1.