

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

March 28, 2022

Abstract

We propose a new approach to modeling bond risk and risk premia, inspired by the equity risk and return literature, which does not impose the tight restrictions found in models that generate closed-form bond prices. We study government bond markets in both the US and China. We estimate the joint dynamics of the volatility and Sharpe ratio of principal-component bond-factor portfolios as functions of yield curve variables and VIX. We uncover an important second factor in risk premia and document a complex time-varying relation between the price and quantity of interest rate risk inconsistent with the frameworks in existing studies.

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 Rob Engle, Matt Richardson, Alexi Savov, Yu Xu, and seminar participants at New York University. Please direct correspondence to carpen-ter@stern.nyu.edu, lufz@hku.hk, rwhitela@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 and return literature, which does not impose the tight restrictions found in models that generate closed-form bond prices. We study government bond markets in both the US and China. We estimate the joint dynamics of the volatility and Sharpe ratio of principal-component bond-factor portfolios as functions of yield curve variables and VIX. We uncover an important second factor in risk premia and document a complex time-varying relation between the price and quantity of interest rate risk inconsistent with the frameworks in existing studies.

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

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 flexibility 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. Early 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. Consistent with this negative risk-return relation, Fleming, Kirby, and Ostdiek (2001) and Moreira and Muir (2017) find that volatility timing strategies that increase equity exposure when volatility is low generate high average Sharpe ratios.¹ 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.² 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, Chinas bond market is growing explosively, as shown in Figure 1, and already has a total market value of 19 trillion USD at the end of 2021, compared with 47 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, 18% vs. 46%, 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 dramatically different from those in the US.

One of the barriers to empirical work on bond returns is the absence of natural bond

¹Lochstoer and Muir (2021) provide a theoretical explanation for this phenomenon based on excessive extrapolation of volatility.

²See 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.

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 zeroes 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 each bond factors conditional volatility and Sharpe ratio processes. For both the US and China, we use traditional yield-curve variables 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. 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 very highly correlated, as equilibrium models would predict for risk factors that are correlated with aggregate consumption, and peak around recessions. 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 full sample correlation between the conditional volatility and Sharpe ratio of this second factor is also positive and quite high, we see episodes of negative correlation between these quantities. 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 the 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.

The stochastic variation in volatilities and Sharpe ratios of the factors 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. Over time, volatilities and Sharpe ratios of both factors have declined to the point where risk premia are close to zero across bonds of all maturities. A similar decline over time shows up in volatility in China, but since there is no evidence of a decline in the price of risk of either factor, Sharpe ratios remain higher on shorter-term zeroes in recent years. Interestingly, variation in China appears to be policy driven rather than linked to economic fluctuations as it is in the US.

Third, and finally, we find that bond risk premia are solely compensation for bond risk in both countries, as no-arbitrage theory predicts. I.e., 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.

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 the 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.³ This literature is largely silent on the corresponding dynamics of bond risk. Since risk premia can be levered arbitrarily, they are not very informative without an understanding of their corresponding risk. Leverage-invariant Sharpe ratios are arguably more informative.

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 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, Priebisch, and Singleton, 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 China, model misspecification is a

³See also Campbell and Shiller (1991).

⁴There 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.

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 1 presents preliminary evidence on the performance and factor structure of government bond excess returns in both the US and China. Section 2 lays out our theoretical model of nominal bond returns, the corresponding empirical specification, and our estimation strategy. Section 3 presents the estimation results for US Treasury bonds. Section 4 presents the estimation results for Chinese government bonds, and Section 5 concludes.

1 Preliminary Evidence on Bond Returns

To lay the groundwork for our model of conditional bond return volatility and price of risk in Section 2, 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. 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, for a number of reasons. 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 possibility 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 will discuss later. 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.

1.1 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 1/1990–12/2019. 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 variance explained by each of these portfolios. The first factor explains most of the total variance 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,

⁵Paul 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.

⁶The results with unstandardized excess returns are qualitatively similar.

especially in the UST zeroes in the post-Volcker period, where the Factor 1 portfolio has a Sharpe ratio of 0.77 and the Factor 2 portfolio has a Sharpe ratio of 0.85.⁷

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 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. Factor 3 is long extreme-maturity zeroes and short middle-maturity zeroes.

Since the eigenvectors in Table 1 show the return responses of each implied zero to the returns on the factor portfolios, we can approximate the yield-curve shift associated with a one-annual-standard-deviation increase in each factor-portfolio return. Figure 3 plots the yield curve shifts associated with the three different factors. 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.

1.2 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

⁷Balduzzi, 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.

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.

1.3 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 12/2019, are for returns net of fees. The columns headed “Gross of 15-bp Fees” show results for excess returns augmented with the 15-basis point 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, and net of these fees, the Sharpe ratios decline with maturity for all but the shortest-maturity ETF. Panel A of Table 2 verifies that the factor structure of UST ETF returns mirrors that of the UST implied zeroes. The Sharpe ratios for Factor 1 and Factor 2 are even larger in the ETF market, gross of fees, perhaps reflecting some variance reduction associated with holding portfolios of bonds. The large Sharpe ratio on Factor 2 explains the declining pattern of ETF Sharpe ratios with maturity that we document in Panel B.

2 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

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

where r_t , the $\mu_{i,t}$, and the d -dimensional row vector $\sigma_{i,t}$ 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 $i = 1, \dots, n$, follows

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

where π_t is the n -dimensional row vector with elements $\pi_{i,t}$, μ_t is the n -dimensional column vector with elements $\mu_{i,t}$, $\mathbf{1}$ 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 $\pi_t(\mu_t - r_t \mathbf{1})$ and $\pi_t \sigma_t$ satisfy the integrability conditions that ensure W_t is well-defined.

2.1 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 $\pi_t(\mu_t - r_t \mathbf{1}) = 0$. 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

$$\sigma_t \theta_t = \mu_t - r_t \mathbf{1} . \quad (3)$$

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

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

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

$$M_t = e^{-\int_0^t r_s ds - \int_0^t \theta'_s dB_s - \frac{1}{2} \int_0^t |\theta_s|^2 ds} \quad (5)$$

such that

$$S_{i,t} = E_t \left\{ \frac{M_u}{M_t} S_{i,u} \right\} \text{ for all } 0 < t < u \text{ and } i = 1, \dots, n . \quad (6)$$

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

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

price of risk on the claim to aggregate consumption is

$$\theta_t = R_t \sigma_{c,t} \quad (7)$$

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

2.2 Nominal Asset Prices with Locally Riskless Inflation

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

$$\frac{dq_t}{q_t} = i_t dt, \quad (8)$$

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

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

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

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

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

$$P_t^T = \mathbf{E}_t \left\{ e^{-\int_t^T (r_s + i_s) ds - \int_t^T \theta'_s dB_s - \frac{1}{2} \int_t^T |\theta_s|^2 ds} \right\}. \quad (11)$$

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.

⁹See Karatzas and Shreve (1998), Eqn. (6.21).

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

2.3 Bond Market Factors and Implied Zero Excess Returns

Motivated by the evidence from Section 1.1 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 $j = 1, 2$, we write the excess return on Factor j , dF_j as

$$dF_{j,t} = \sigma_{j,t}\theta_{j,t} dt + \sigma_{j,t} dB_{j,t} , \quad (12)$$

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

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

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 $\beta_{i,1}\sigma_{1,t}\theta_{1,t} + \beta_{i,2}\sigma_{2,t}\theta_{2,t}$. 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.

2.4 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),

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

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

$$\sigma_{j,t} = X_t\beta_j^\sigma \quad (15)$$

and

$$\theta_{j,t} = X_t\beta_j^\theta \quad (16)$$

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

2.4.1 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

¹¹We 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.

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 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 should be relevant for predicting both bond return volatility and its 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 speculate 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, GARCH 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.

¹²The 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.

¹³This 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.

Finally, for both the UST and CGB factors, we also tried including the lagged value of the realized volatility of each bond factor return, approximated as $\sqrt{\frac{\pi}{2}}|R_{j,t}|$, as a predictor variable, but it is insignificant in all cases.

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 here, we are working with factor portfolios of zeroes with annual maturities up to ten years. To include all ten forward rates seems likely to lead to overfitting, 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. A number of other variables have been used to predict bond excess returns in the literature. 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 plus VIX, which seem natural and well-motivated.

2.4.2 GMM Estimation Equations and Diagnostics

For each $j = 1, 2$, we perform a simultaneous GMM estimation of β_j^σ and β_j^θ from the following two equations:

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

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

where we use $E\{\sqrt{\frac{\pi}{2}}|u_{j,t}|\} = E\{\sqrt{\frac{\pi}{2}}|\sigma_{j,t-1}\varepsilon_{j,t}|\} = \sigma_{j,t-1}$. 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.¹⁴ We include this constant in preliminary specifications to check for possible mis-specification in Equations (15) and (16).

¹⁴Other 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

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

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

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.

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

$$\text{Goodness-of-fit (1)} = 1 - \frac{\sum_t v_{j,t}^2}{\frac{\pi}{2} \sum_t (|u_{j,t}| - |\bar{u}_j|)^2} , \quad (21)$$

$$\text{Goodness-of-fit (2)} = 1 - \frac{\sum_t u_{j,t}^2}{\sum_t (R_{j,t} - \bar{R}_j)^2} . \quad (22)$$

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, 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,

$$H_{0,1} : \beta_{j,1}^\sigma = \beta_{j,2}^\sigma = \dots = \beta_{j,k}^\sigma = 0 ,$$

where the $\beta_{j,1}^\sigma, \dots, \beta_{j,k}^\sigma$ 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,

$$H_{0,2} : \beta_{j,1}^\theta = \dots = \beta_{j,k}^\theta = 0 ,$$

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

3 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 2019. Then 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. Finally, we analyze the time series of fitted volatility and Sharpe ratio values to shed additional light on the dynamics of return premia.

3.1 GMM Estimation Results for the UST Factors

The top panel of Table 3 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 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-hand side of Table 3 reports results for UST Factor 1 and the right-hand side reports results for UST Factor 2. For convenience, the yield-curve variables are divided 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 price of risk is constant. However, when we impose the no-arbitrage restriction that $\alpha_1 = 0$ in Specification (1b), we increase power.¹⁵ 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 variable is insignificant in both the volatility and return equations, so to further increase power, we exclude this variable in Specification (1c). This 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

¹⁵The 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).

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. 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, and VIX, 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-arbitrage restriction in our model of Section 2. 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.

3.2 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.¹⁶ 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.

¹⁶See, for example, Cochrane and Piazzesi (2005) and Cieslak and Povala (2015).

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 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 4 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 4, 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).¹⁷ 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.

¹⁷See equations (6) and (7) in Boudoukh et al. (2008).

3.3 Fitted UST Factor Volatilities and Sharpe Ratios

Figure 4 plots the time series of annualized fitted values of UST Factor-1 and Factor-2 Sharpe ratios and volatilities based on the GMM estimates from Table 3. Panel A plots Factor-1 fitted values from Specification (1c) of Table 3, and Panel B plots Factor-2 fitted values from Specification (2b) of Table 3. As the figure shows, the correlations between the Sharpe ratio (price of risk) and the volatility (quantity of risk) are significantly positive for both factors. More specifically, the time-series correlation between the Sharpe ratio of Factor 1 and the volatility of Factor 1 is 99.9% and this same correlation for Factor 2 is 55% with a Newey-West t -statistic of 5.51. The positive relation between the factor prices and quantities of risk are consistent with the predictions of equilibrium models of the pricing of risk factors that are correlated with aggregate consumption.¹⁸ At the same time, the fitted Sharpe ratios for Factor 1 and Factor 2 change sign over the sample period, which cannot be accommodated by affine models with stochastic variation in volatility (Duffee, 2002).

The near perfect correlation between the Sharpe ratios and volatilities 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 very high correlation is consistent with these tight restrictions and thus explains why these models do a good job of fitting the data. However, there are features of the data that are difficult, if not impossible, to accommodate in these models. While the full sample correlation between the Factor 2 price and quantity of risk is also positive, this correlation obscures significant time-variation. When computed over 36-month rolling periods, this correlation fluctuates between 91% and -40%.

Figure 4 also shows that factor prices and quantities of risk spike up during NBER recessions. This former 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. Increases in volatility during recessions are also a feature seen in other financial and economic series.

In addition to this cyclical pattern in volatility, there is also evidence of a notable decline in the volatility of both factors over the sample period. Fitted volatilities are approximately half as large at the end of the sample as they are at the beginning of the sample. Given the positive correlation between the volatilities and Sharpe ratios of both factors, it is not surprising that the Sharpe ratios exhibit a similar time series pattern. More specifically, fitted annualized Sharpe ratios that exceed one for both factors in the 1990s decline to

¹⁸See, for example, Campbell (1987).

values that average closer to zero from 2012 onwards.

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

$$[\Delta(\hat{\sigma}\hat{\theta})]^2 \approx \hat{\sigma}^2(\Delta\hat{\theta})^2 + \hat{\theta}^2(\Delta\hat{\sigma})^2 + 2\hat{\sigma}\hat{\theta}(\Delta\hat{\sigma})(\Delta\hat{\theta}) . \quad (23)$$

Summing Equation (23) over the observations in our sample and dividing by the left-hand side, we get a sample variance decomposition under the natural assumption that the mean of the risk premium is zero. For UST Factor 1, the components of this decomposition are $\frac{\sum_{t=1}^{T-1} \hat{\sigma}_t^2 (\hat{\theta}_{t+1} - \hat{\theta}_t)^2}{\sum_{t=1}^{T-1} (\hat{\sigma}_{t+1} \hat{\theta}_{t+1} - \hat{\sigma}_t \hat{\theta}_t)^2} = 48\%$, $\frac{\sum_{t=1}^{T-1} \hat{\theta}_t^2 (\hat{\sigma}_{t+1} - \hat{\sigma}_t)^2}{\sum_{t=1}^{T-1} (\hat{\sigma}_{t+1} \hat{\theta}_{t+1} - \hat{\sigma}_t \hat{\theta}_t)^2} = 9\%$, and $2 \frac{\sum_{t=1}^{T-1} \hat{\sigma}_t \hat{\theta}_t (\hat{\sigma}_{t+1} - \hat{\sigma}_t) (\hat{\theta}_{t+1} - \hat{\theta}_t)}{\sum_{t=1}^{T-1} (\hat{\sigma}_{t+1} \hat{\theta}_{t+1} - \hat{\sigma}_t \hat{\theta}_t)^2} = 38\%$.¹⁹ Thus, neglecting variation in volatility is effectively equivalent to misattributing about 50% of the variation in risk premia. Similarly, for UST Factor 2 these components are $\frac{\sum_{t=1}^{T-1} \hat{\sigma}_t^2 (\hat{\theta}_{t+1} - \hat{\theta}_t)^2}{\sum_{t=1}^{T-1} (\hat{\sigma}_{t+1} \hat{\theta}_{t+1} - \hat{\sigma}_t \hat{\theta}_t)^2} = 42\%$, $\frac{\sum_{t=1}^{T-1} \hat{\theta}_t^2 (\hat{\sigma}_{t+1} - \hat{\sigma}_t)^2}{\sum_{t=1}^{T-1} (\hat{\sigma}_{t+1} \hat{\theta}_{t+1} - \hat{\sigma}_t \hat{\theta}_t)^2} = 14\%$, and $2 \frac{\sum_{t=1}^{T-1} \hat{\sigma}_t \hat{\theta}_t (\hat{\sigma}_{t+1} - \hat{\sigma}_t) (\hat{\theta}_{t+1} - \hat{\theta}_t)}{\sum_{t=1}^{T-1} (\hat{\sigma}_{t+1} \hat{\theta}_{t+1} - \hat{\sigma}_t \hat{\theta}_t)^2} = 36\%$. As in the case of UST Factor 1, neglecting variation in volatility is equivalent to misattributing about 50% of the variation in risk premia. These results suggest that 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.

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

As discussed in Section 2.3, 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:

$$sz_{i,t} = \beta_{i1} R_{1,t} + \beta_{i2} R_{2,t} , \quad (24)$$

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

¹⁹These components do not sum to exactly one because we dropped the higher-order terms.

B, the unstandardized monthly excess return on zero i is $z_{i,t} = v_i s z_{i,t}$. It follows that the annualized fitted conditional volatility of the unstandardized excess return on zero i is

$$vol_t(z_i) = \sqrt{12} v_i \sqrt{\beta_{i1}^2 \hat{\sigma}_{1,t}^2 + \beta_{i2}^2 \hat{\sigma}_{2,t}^2} . \quad (25)$$

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

$$rp_t(z_i) = 12 v_i (\beta_{i1} \hat{\sigma}_{1,t} \hat{\theta}_{1,t} + \beta_{i2} \hat{\sigma}_{2,t} \hat{\theta}_{2,t}) , \quad (26)$$

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

$$sr_t(z_i) = rp_t(z_i) / vol_t(z_i) . \quad (27)$$

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. These plots illustrate a number of interesting features of the data.

First, with regard to volatility, it is no surprise that the two-year and ten-year zeroes exhibit similar time series behavior, with the ten-year volatility scaled by a factor of approximately five, i.e., its relative duration. Most notable is that both volatility series exhibit the same approximately 50% decline in volatility exhibited by the underlying factors as discussed in Section 3.3.

Second, while both zeroes also exhibit major declines in fitted Sharpe ratios over the sample, the decline is much 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.

Third, the joint dynamics of the volatilities and Sharpe ratios generate an interesting pattern in risk premia. Specifically, the gap between the risk premia on the ten-year and two-year zeroes shows both a marked business cycle pattern and a trend over time. This gap is minimal or even negative at the end of expansions and heading into recessions, but it spikes at the beginning of expansions. However, this gap has also apparently been declining over time, with small or no differences between the risk premia on the two bonds over the last five years of the sample.

4 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 12/2019. Then we analyze the times series of fitted volatilities and Sharpe ratios for bond-factor portfolios and individual bonds in China. 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.

4.1 GMM Estimation Results for the CGB Factors

The top panel of Table 5 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 5 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 in the volatility and Sharpe ratio functions. The p -values of the J -statistic tests are uniformly high, suggesting that linear functions of the predictor variables are adequate for modeling the factor volatilities and Sharpe Ratios. For CGB Factor 1, Column (1a) of Table 5 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 increase the power of the estimation of the Sharpe ratio coefficients. Three of the coefficient estimates become marginally to highly significant. In addition, the p -values for the Wald tests all fall below 1% or 5%. We 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 with at least marginal significance 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.

For CGB Factor 2, Specification (2a) in Table 5 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 $\alpha_2 = 0$. As with CGB Factor 1, this increases our power to reject the null hypothesis that the price of CGB Factor-2 risk is zero or constant. The Wald test p -value is about 1%. 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.

However, 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.

4.2 Fitted CGB Factor Volatilities and Sharpe Ratios

Figure 6 plots the time series of fitted values of CGB Factor-1 and Factor-2 Sharpe ratios and volatilities based on GMM estimates from Table 5. Panel A plots Factor-1 fitted values from Specification (1b) of Table 5, and Panel B plots Factor-2 fitted values from Specification (2b) of Table 5. 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 -46% with a Newey-West t -statistic of -2.68 and this same correlation for Factor 2 is -63% with a Newey-West t -statistic of -6.37.

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 lead 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 lead 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.

As for US Factor 2, 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 48% to a low of -92%. However, in contrast to the US results, this variation is in the first, and empirically dominant factor, in China. 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.

4.3 CGB Bond Volatilities, Sharpe Ratios, and Risk Premia

Following the method described in Section 3.4, 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 7 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 Conclusion

Government bonds are a critical component of many investors' portfolios, in some cases even more critical than equities. However, the associated 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 that still imposes no arbitrage by restricting risk premia to be functions of risk. We decompose 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.

Interestingly, these two components covary positively in the US Treasury market, with the components of the empirically dominant first factor exhibiting an almost perfect positive correlation. This result is in stark contrast to the evidence in the US equity markets, where the observed correlation is negative, leading to apparently Sharpe-ratio-maximizing, volatility-timing strategies that increase equity exposure when volatility is low (see, for example, Fleming et al. (2001) and Moreira and Muir (2017)). The reverse is true in the Treasury bond market. For example, a volatility-managed portfolio that holds UST Factor 1 in inverse proportion to its variance, as in Moreira and Muir (2017), has a Sharpe ratio only 67% as large as the original Factor 1 portfolio.

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, in China, the quantity and price of risk of both factors exhibit a negative unconditional correlation. Moreover, these correlations vary significantly over time. For example, periods of significant government intervention, associated with the 2008 financial crisis and the 2015 stock market meltdown, generate a strong negative correlation between the price and quantity of interest rate risk of the first factor, but these components are positively correlated in other periods. This time variation

is difficult, if not impossible, to accommodate in the theoretical models in the literature 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

- 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, 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.

- 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.
- Fleming, Jeff, Chris Kirby, and Barbara Ostdiek, 2001, The economic value of volatility timing, *Journal of Finance* 56, 329–352.
- 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.
- Joslin, Scott, and Anh Le, 2021, Interest rate volatility and no-arbitrage affine term structure models, *Management Science* 67, 7391–7416.
- Joslin, Scott, Marcel Pribsch, and Kenneth J Singleton, 2014, Risk premiums in dynamic term structure models with unspanned macro risks, *Journal of Finance* 69, 1197–1233.
- Karatzas, Ioannis, and Steven E Shreve, 1998, *Methods of Mathematical Finance*, volume 39 (Springer).
- 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.
- Lochstoer, Lars A, and Tyler Muir, 2021, Volatility expectations and returns, *Journal of Finance*, *forthcoming* .
- Ludvigson, Sydney C, and Serena Ng, 2009, Macro factors in bond risk premia, *Review of Financial Studies* 22, 5027–5067.
- Moreira, Alan, and Tyler Muir, 2017, Volatility-managed portfolios, *Journal of Finance* 72, 1611–1644.
- 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.

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 10x10 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			1/1990–12/2019			5/2004–12/2019		
A. Factor Structure	F1	F2	F3	F1	F2	F3	F1	F2	F3
Factor Var. as % of Tot.	94.69	4.07	0.73	90.64	7.35	1.42	82.05	13.91	2.37
Factor Vol	10.66	2.21	0.93	10.43	2.97	1.30	9.92	4.09	1.69
Factor SR	0.27	0.45	0.45	0.77	0.85	0.61	0.53	0.21	0.07
1-year zero loadings	0.30	0.57	0.43	0.26	0.66	0.66	0.24	0.56	0.59
2-year zero loadings	0.31	0.40	0.11	0.31	0.40	-0.24	0.29	0.46	0.09
3-year zero loadings	0.32	0.27	-0.14	0.32	0.24	-0.36	0.32	0.30	-0.27
4-year zero loadings	0.32	0.15	-0.29	0.33	0.10	-0.32	0.34	0.15	-0.39
5-year zero loadings	0.32	0.02	-0.39	0.33	-0.03	-0.23	0.34	0.01	-0.38
6-year zero loadings	0.32	-0.14	-0.33	0.33	-0.12	-0.11	0.34	-0.12	-0.21
7-year zero loadings	0.32	-0.26	-0.19	0.33	-0.20	0.01	0.33	-0.22	0.01
8-year zero loadings	0.32	-0.31	0.02	0.32	-0.26	0.13	0.33	-0.27	0.14
9-year zero loadings	0.32	-0.34	0.27	0.32	-0.31	0.25	0.32	-0.32	0.27
10-year zero loadings	0.31	-0.34	0.58	0.31	-0.35	0.37	0.31	-0.35	0.38
B. Performance Measures	Mean	Vol	SR	Mean	Vol	SR	Mean	Vol	SR
1-year zero	1.40	2.51	0.56	0.80	0.64	1.24	0.46	0.89	0.52
2-year zero	1.56	4.70	0.33	1.54	1.63	0.94	0.86	1.52	0.57
3-year zero	1.68	6.34	0.26	2.10	2.67	0.79	1.12	2.07	0.54
4-year zero	1.94	8.07	0.24	2.77	3.68	0.75	1.39	2.68	0.52
5-year zero	2.26	9.71	0.23	3.27	4.67	0.70	1.68	3.37	0.50
6-year zero	2.61	11.11	0.23	3.82	5.58	0.68	2.08	3.99	0.52
7-year zero	2.60	12.43	0.21	4.04	6.47	0.62	2.08	4.67	0.45
8-year zero	2.73	13.55	0.20	4.35	7.33	0.59	2.27	5.30	0.43
9-year zero	2.83	14.46	0.20	4.58	8.20	0.56	2.46	5.95	0.41
10-year zero	2.84	15.25	0.19	4.61	9.08	0.51	2.61	6.64	0.39

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

The factor structure of US Treasury ETF excess returns, gross and net of 15-basis-point annual fees, in Panel A, and their unconditional means, volatilities, and Sharpe ratios in Panel B. The sample period is 2/2007–12/2019. 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 6x6 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.

	Gross of 15-bp Fees			Net of Fees		
A. Factor Structure	F1	F2	F3	F1	F2	F3
Factor Var. as % of Tot.	76.19	16.30	5.92	76.19	16.30	5.92
Factor Vol	7.41	3.43	2.06	7.41	3.43	2.06
Factor SR	0.91	0.94	0.29	0.79	0.44	0.17
0-1-year ETF	0.27	0.75	0.59	0.27	0.75	-0.59
1-3-year ETF	0.40	0.39	-0.53	0.40	0.39	0.53
3-7-year ETF	0.45	0.06	-0.42	0.45	0.06	0.42
7-10-year ETF	0.46	-0.19	-0.02	0.46	-0.19	0.02
10-20-year ETF	0.44	-0.32	0.22	0.44	-0.32	-0.22
>20-year ETF	0.41	-0.38	0.38	0.41	-0.38	-0.38
B. Performance Measures	Mean	Vol	SR	Mean	Vol	SR
0-1-year ETF	0.33	0.25	1.32	0.18	0.25	0.72
1-3-year ETF	1.22	1.19	1.02	1.07	1.19	0.90
3-7-year ETF	3.11	3.54	0.88	2.96	3.54	0.83
7-10-year ETF	4.40	6.31	0.70	4.25	6.31	0.67
10-20-year ETF	5.33	8.76	0.61	5.18	8.76	0.59
>20-year ETF	6.85	13.71	0.50	6.70	13.71	0.49

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 1/1990–12/2019. 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 = $Y_2/10$, Slope = $(Y_{10} - Y_2)/10$, and Curvature = $(Y_6 - \frac{Y_2+Y_{10}}{2})/10$, 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. Wald test (1) tests the null hypothesis that factor volatility is constant, i.e., $\beta_{j,1}^\sigma = \beta_{j,2}^\sigma = \dots = \beta_{j,k}^\sigma = 0$. Wald test (2) tests the null hypothesis that the price of factor risk is constant, i.e., $\beta_{j,1}^\theta = \dots = \beta_{j,k}^\theta = 0$. Goodness-of-fit (1) = $1 - \frac{\sum_t v_{j,t}^2}{\frac{\pi}{2} \sum_t (|u_{j,t}| - |u_j|)^2}$. Goodness-of-fit (2) = $1 - \frac{\sum_t u_{j,t}^2}{\sum_t (R_{j,t} - R_j)^2}$.

	UST Factor 1			UST Factor 2	
	(1a)	(1b)	(1c)	(2a)	(2b)
Volatility Coefficients (β_j^σ)					
Constant	0.30 (0.65)	0.28 (0.73)	0.40 (1.10)	0.00 (0.03)	-0.02 (-0.14)
Level	2.90 (3.89)	2.93 (4.29)	2.60 (4.51)	0.61 (2.99)	0.62 (3.07)
Slope	7.44 (2.87)	7.49 (3.02)	6.16 (4.01)	2.79 (3.12)	2.83 (3.22)
Curvature	-6.72 (-0.71)	-6.90 (-0.74)		-8.41 (-2.29)	-8.59 (-2.41)
VIX/100	3.88 (2.81)	3.88 (2.90)	4.10 (3.03)	1.93 (3.74)	2.02 (3.71)
Sharpe Ratio Coefficients (β_j^θ)					
Constant	-0.48 (-0.27)	-0.25 (-1.11)	-0.36 (-1.69)	0.69 (0.36)	-0.36 (-1.64)
Level	0.44 (0.42)	0.30 (0.93)	0.57 (2.14)	0.12 (0.11)	0.66 (1.85)
Slope	0.76 (0.26)	0.41 (0.34)	1.44 (1.94)	-2.46 (-0.58)	-0.25 (-0.18)
Curvature	4.35 (0.80)	4.76 (1.04)		7.65 (0.61)	1.43 (0.27)
VIX/100	1.20 (0.78)	1.01 (1.29)	0.95 (1.24)	0.99 (0.49)	2.05 (3.14)
Return Constant (α_j)	0.32 (0.13)			-0.37 (-0.55)	
No. Moment Conditions	20	20	14	20	20
J -stat p -value (in %)	69.26	78.37	83.97	15.01	20.44
Wald test (1) p -value (in %)	0.00	0.00	0.00	0.00	0.00
Wald test (2) p -value (in %)	60.20	10.86	5.40	6.22	0.11
Goodness-of-fit (1) (in %)	8.08	8.19	8.17	10.22	11.30
Goodness-of-fit (2) (in %)	4.72	4.62	4.53	4.31	2.48

Table 4: R^2 's in Monthly and Annual Return Regressions

The table compares regression results using monthly returns with those using annual overlapping 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. 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 5: 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 5/2004–12/2019. 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 = $Y_2/10$, Slope = $(Y_{10} - Y_2)/10$, and Curvature = $(Y_6 - \frac{Y_2+Y_{10}}{2})/10$, where Y_T is the yield in percent on the T -year zero, for $T = 2, 6$, and 10. Wald test (1) tests the null hypothesis that factor volatility is constant, i.e., $\beta_{j,1}^\sigma = \beta_{j,2}^\sigma = \dots = \beta_{j,k}^\sigma = 0$. Wald test (2) tests the null hypothesis that the price of factor risk is constant, i.e., $\beta_{j,1}^\theta = \dots = \beta_{j,k}^\theta = 0$. Goodness-of-fit (1) = $1 - \frac{\sum_t v_{j,t}^2}{\frac{\pi}{2} \sum_t (|u_{j,t}| - |u_j|)^2}$. Goodness-of-fit (2) = $1 - \frac{\sum_t u_{j,t}^2}{\sum_t (R_{j,t} - R_j)^2}$.

	CGB Factor 1		CGB Factor 2	
	(1a)	(1b)	(2a)	(2b)
Volatility Coefficients (β_j^σ)				
Constant	2.19 (2.16)	1.96 (1.64)	-0.34 (-0.75)	-0.38 (-0.86)
Level	1.98 (0.73)	2.63 (0.85)	3.11 (2.56)	3.23 (2.71)
Slope	5.49 (1.17)	7.42 (1.42)	6.95 (4.24)	7.04 (4.46)
Curvature	-50.65 (-3.08)	-61.89 (-4.25)	-5.39 (-0.76)	-4.64 (-0.66)
Sharpe Ratio Coefficients (β_j^θ)				
Constant	0.33 (0.16)	-1.21 (-2.15)	1.57 (0.76)	0.85 (1.62)
Level	2.69 (1.11)	3.57 (2.46)	-1.90 (-0.71)	-1.04 (-0.75)
Slope	-0.04 (-0.01)	1.91 (0.86)	-5.15 (-0.92)	-3.34 (-1.84)
Curvature	52.58 (1.64)	21.55 (1.86)	-9.51 (-1.09)	-11.52 (-1.54)
Return Constant (α_j)	-3.78 (-0.84)		-0.35 (-0.38)	
No. Moment Conditions	14	14	14	14
J-stat p -value (in %)	69.11	55.20	47.74	58.35
Wald test (1) p -value (in %)	1.38	0.00	0.04	0.01
Wald test (2) p -value (in %)	11.06	3.99	10.91	1.14
Goodness-of-fit (1) (in %)	4.12	4.45	10.03	10.47
Goodness-of-fit (2) (in %)	6.18	5.17	5.31	5.28

Figure 1: Bond Market Growth in China 2002-2021

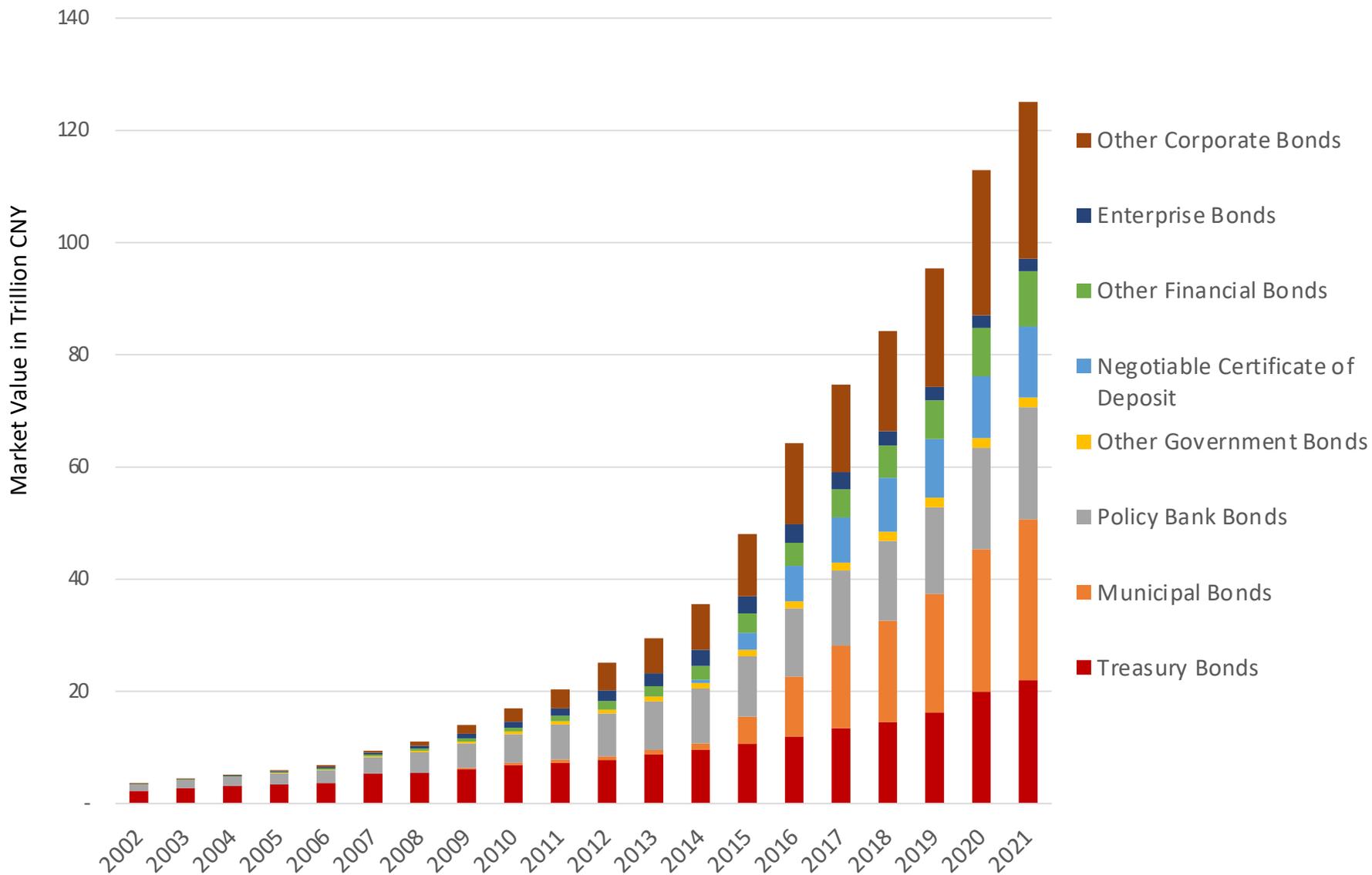
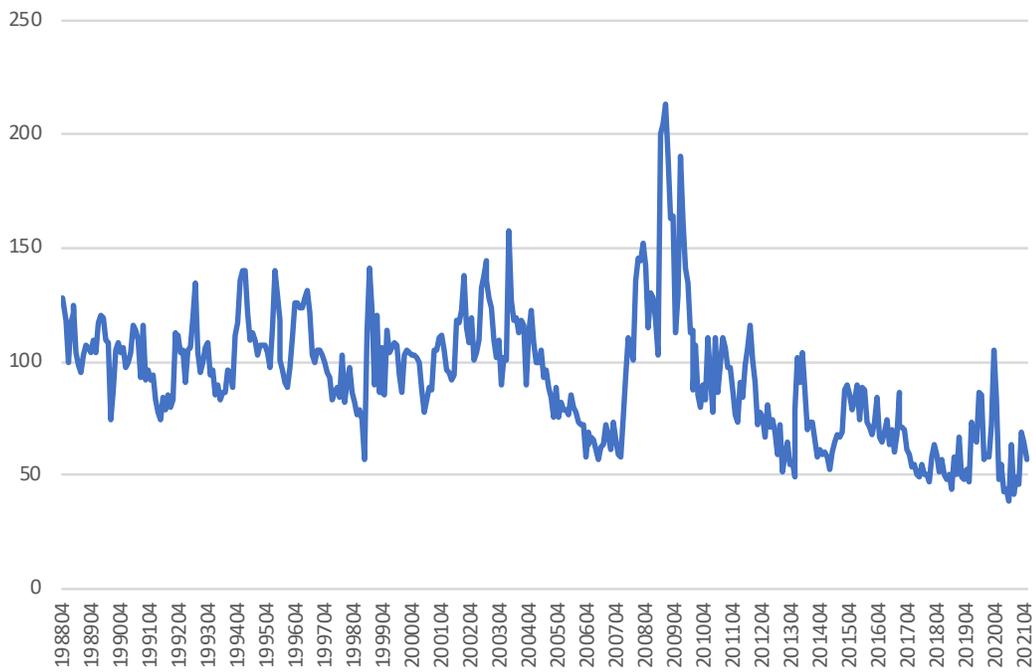
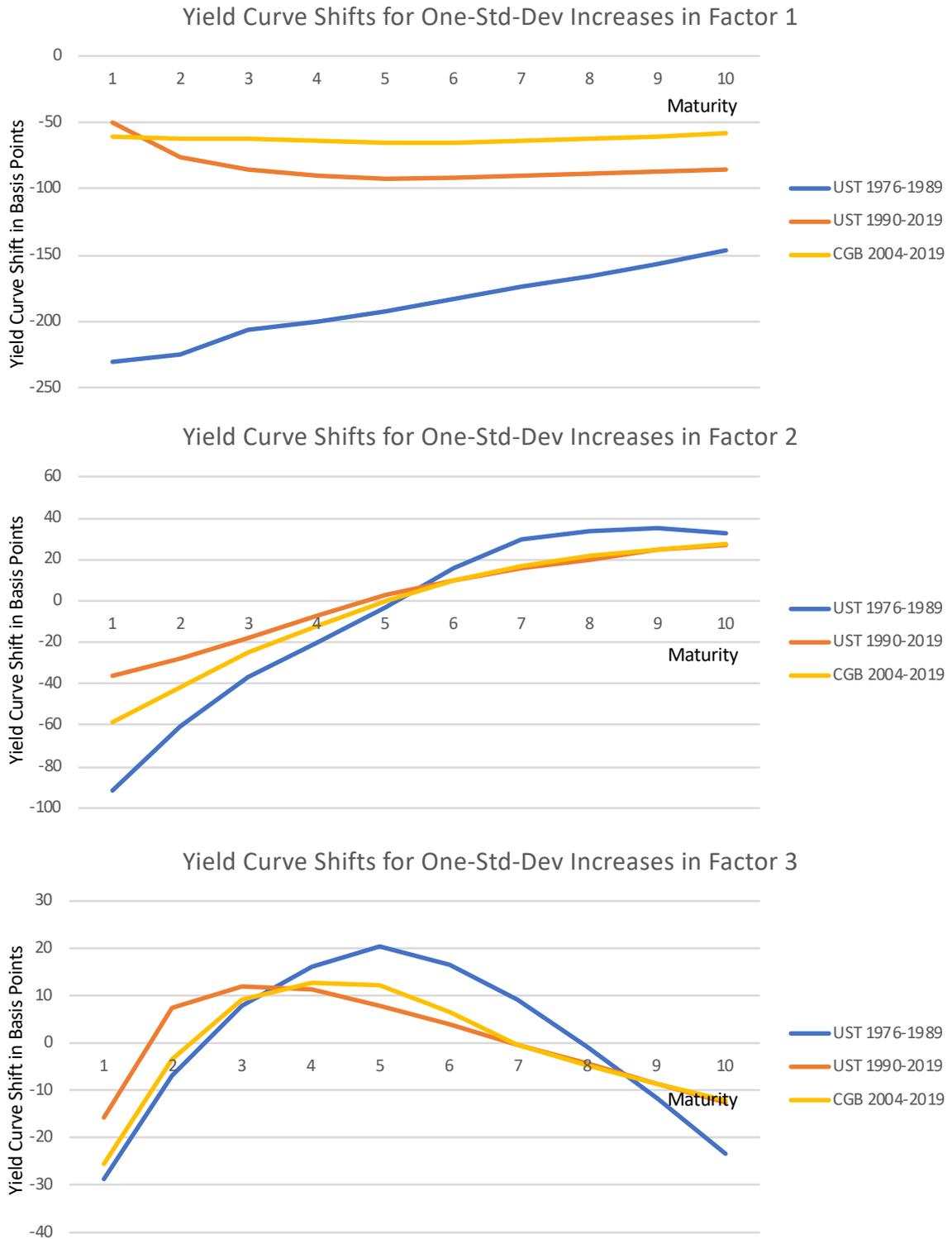


Figure 2: MOVE Index of Implied Yield Volatility from One-Month Treasury Bond Options



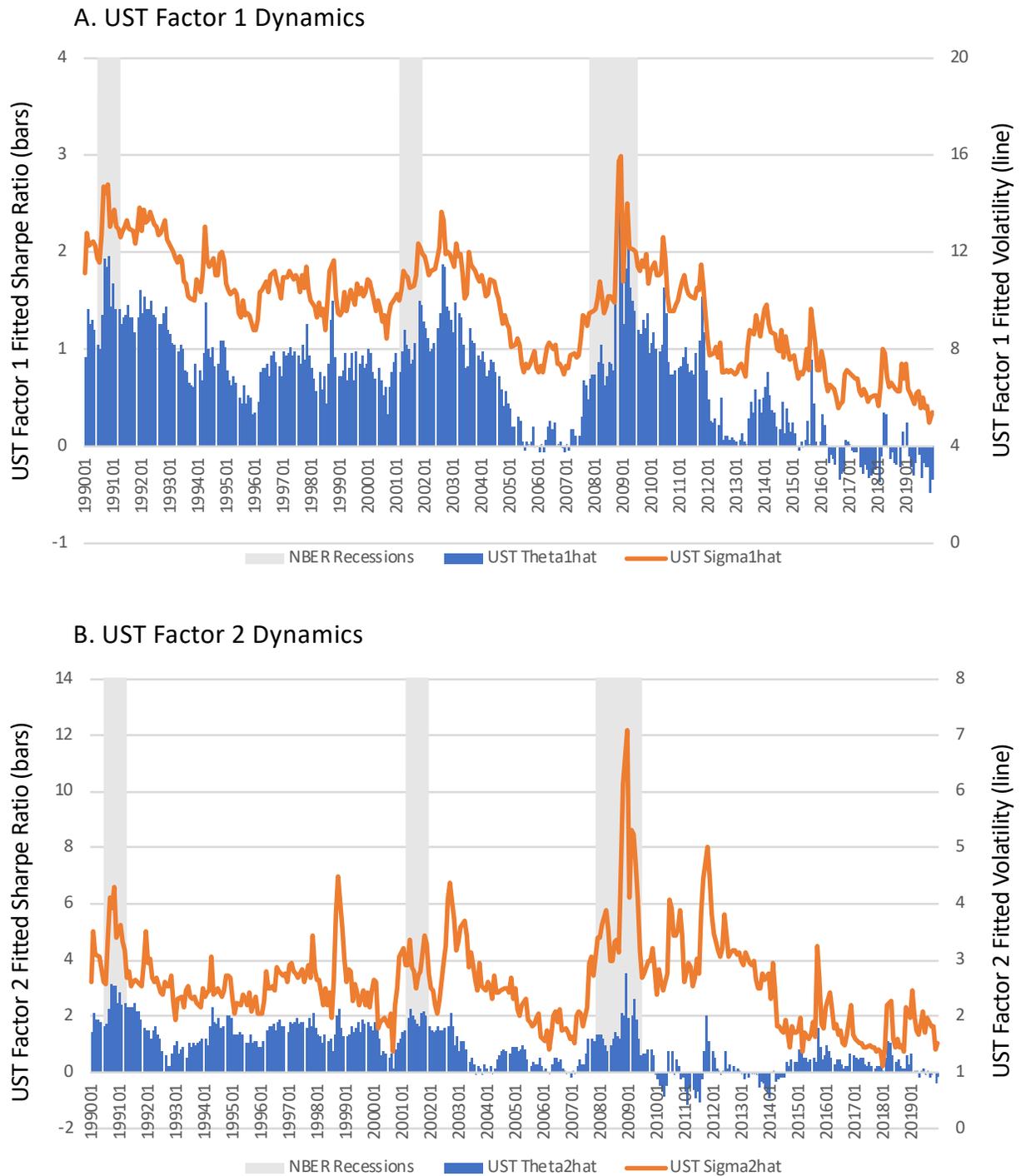
MOVE measures the implied yield volatility of a basket of one-month over-the-counter options on 2-year, 5-year, 10-year, and 30-year Treasuries.

Figure 3: Yield Curve Shifts for One-Std-Dev Increases in Factors



Yield curve shifts associated with one-standard-deviation increases in the returns of the bond market factors described in Table 1.

Figure 4: UST Factor Dynamics



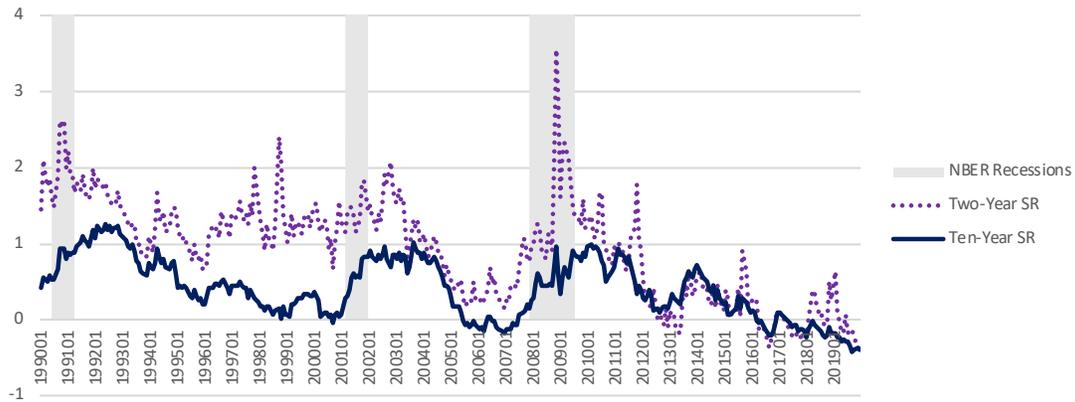
Time series of annualized fitted values of UST Factor 1 and Factor 2 Sharpe ratios and volatilities based on GMM estimates of factor dynamics from Specifications (1c) and (2b) of Table 3, respectively.

Figure 5: UST Bond Dynamics

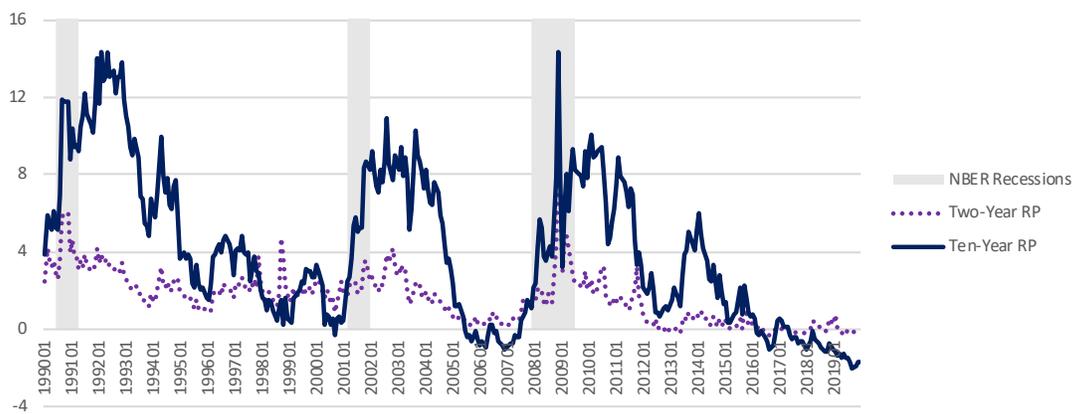
A. UST Bond Volatilities



B. UST Bond Sharpe Ratios

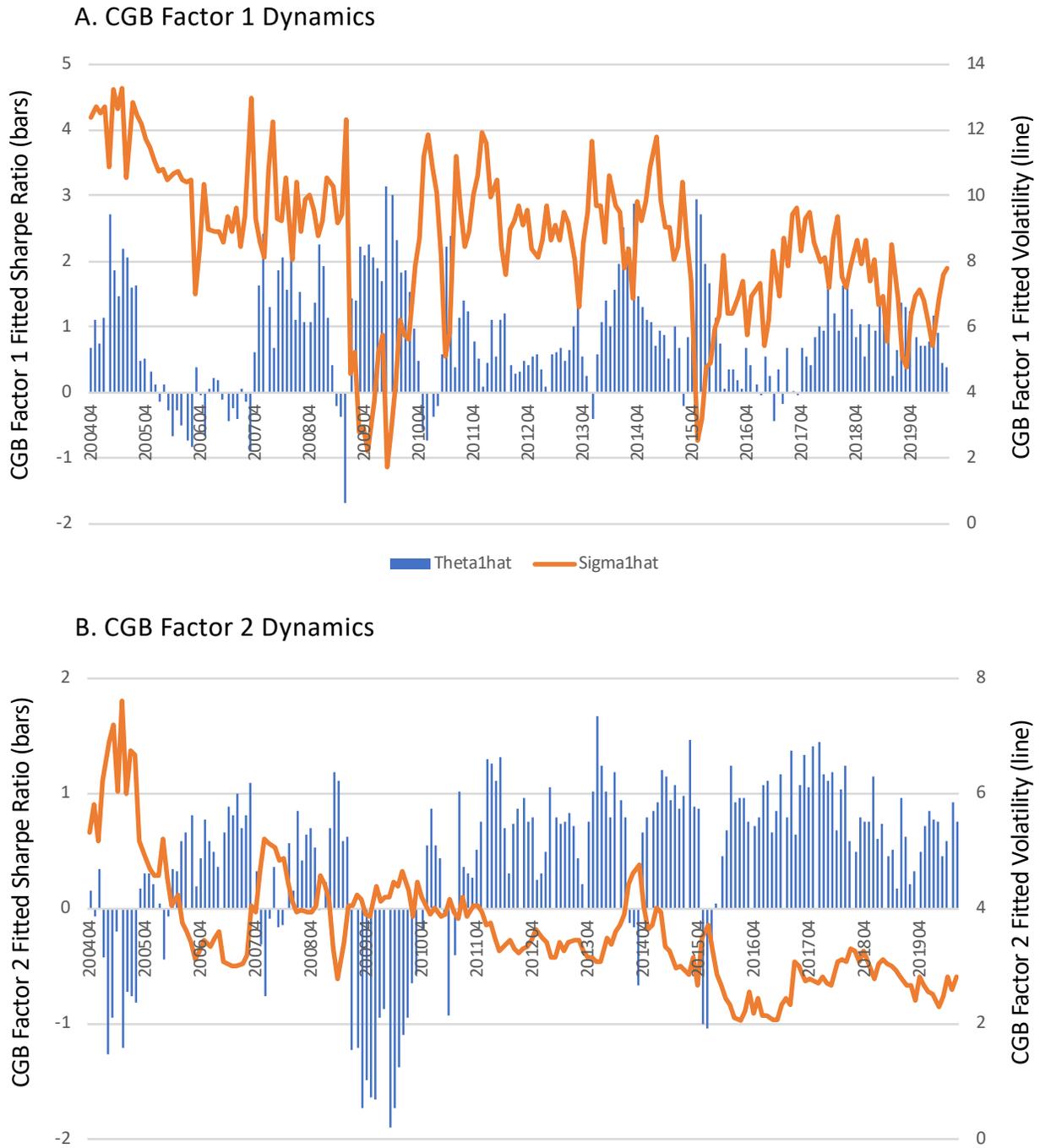


C. UST Bond Risk Premia



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.

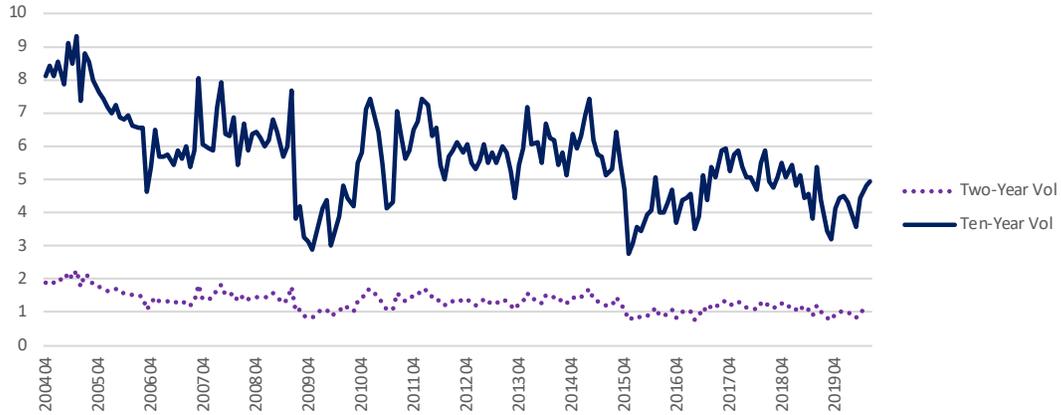
Figure 6: CGB Factor Dynamics



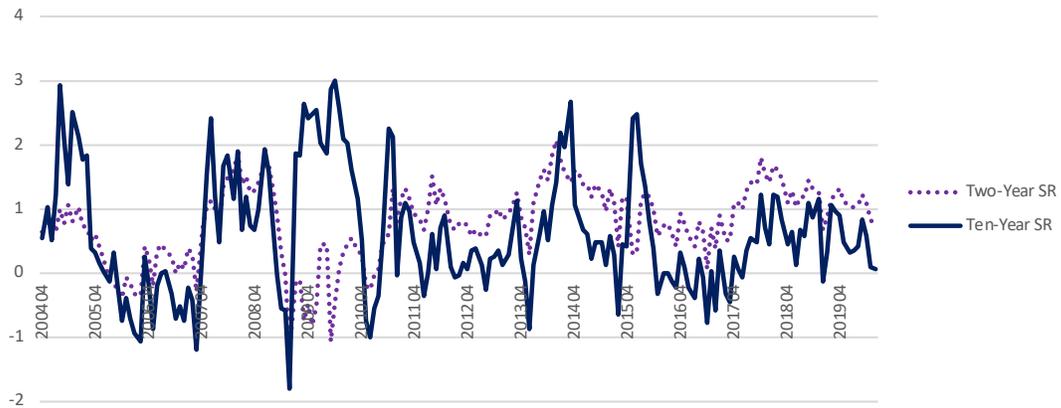
Time series of annualized fitted values of CGB Factor 1 and Factor 2 Sharpe ratios and volatilities based on GMM estimates of factor dynamics from Specifications (1b) and (2b) of Table 5, respectively.

Figure 7: CGB Bond Dynamics

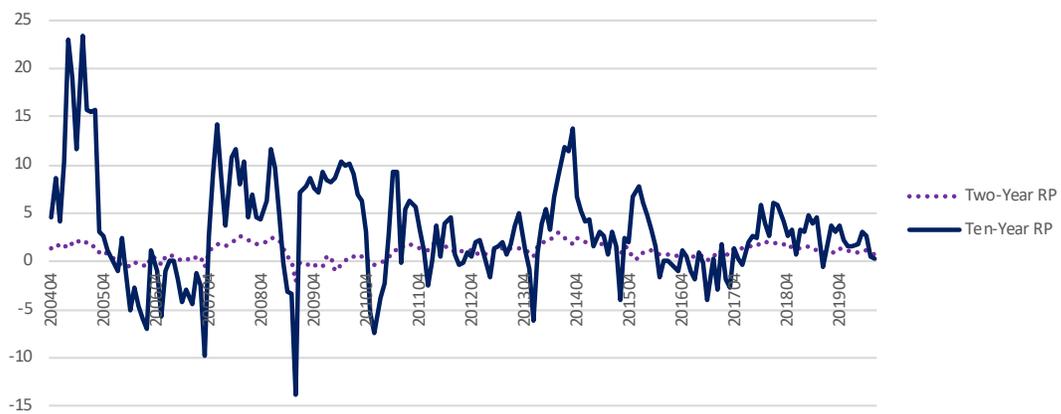
A. CGB Bond Volatilities



B. CGB Bond Sharpe Ratios



C. CGB Bond Risk Premia



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.