Modeling Volatility in Dynamic Term Structure Models*

Hitesh Doshi
University of Houston

Kris Jacobs
University of Houston

Rui Liu
Duquesne University

August 2, 2022

Abstract

We propose no-arbitrage term structure models with volatility factors that follow GARCH processes. The models' tractability is similar to canonical affine term structure models, but they fit yield volatility much better, especially for long-maturity yields. This improvement does not come at the expense of a deterioration in yield fit. Because of the improved volatility fit, the model performs substantially better when hedging Treasury futures options. We conclude that the specification of the volatility factors is critical. Modeling volatility as a function of (lagged) squared innovations to factors improves on models where volatility is a linear function of the factors.

JEL Classification: G12, C58, E43

Keywords: Term Structure; Affine Models; Stochastic Volatility; GARCH; Treasury Futures Options; Hedging.

*We would like to thank an anonymous referee, Liuren Wu, and our discussants Scott Joslin, Anh Le and Guillaume Roussellet for helpful comments. Doshi: hdoshi@bauer.uh.edu; Jacobs (Corresponding Author): kjacobs@bauer.uh.edu; Liu: liur2@duq.edu.

1 Introduction

There is a wealth of evidence in the literature indicating that volatility implied by standard no-arbitrage term structure models corresponds poorly to measures of realized volatility or other model-free estimates. For state-of-the-art affine term structure models (ATSMs) with stochastic volatility, simultaneously matching the properties of the conditional means and variances of yields is indeed the key empirical challenge. These models are characterized by an inherent tension between fitting yields and yield volatility, partly because the mean and variance of yields are driven by the levels of the same state variables.[1]

We propose no-arbitrage term structure models in which the volatility factors follow GARCH processes. Given this specification, the bond prices can be computed analytically. The models' tractability is thus similar to that of canonical affine volatility models, but they capture the time variation in the conditional variances of yields much more accurately. The model is motivated by the well-established literature on ARCH and GARCH models (Engle, 1982; Bollerslev, 1986). We incorporate GARCH dynamics into state-of-the-art term structure models by relating the conditional volatility to the lagged squared *residuals* of the factors driving yields. From a modeling perspective, this is a critical difference with standard ATSMs with stochastic volatility, which model volatility as a linear combination of the *level* of the yield factors.

While this may appear to be a trivial distinction, we show that it greatly matters for the models' ability to capture the stylized facts of the time series and cross-section of conditional volatility. In the empirical analysis, we follow the existing literature and focus on a version of the model with three latent factors, but for simplicity and parsimony we only endow the volatility of the residuals of the first factor with a GARCH dynamic. We estimate the model using monthly Treasury yields from November 1971 to October 2019. Following the existing literature, we

[1] See Dai and Singleton (2000, 2002), Duffee (2002), Longstaff and Schwartz (1992), Collin-Dufresne, Goldstein, and Jones (2009), and Duarte (2004) for evidence on the volatility fit of these models and the trade-off between fitting yields and yield volatility. See Joslin and Le (2020) for a discussion of the role no-arbitrage restrictions play in the tension between fitting first and second moments.

use both realized volatility and EGARCH(1,1) as model-free "true" volatility benchmarks.[2] We find that the no-arbitrage model with GARCH dynamics performs well in capturing the "true" volatility, especially at longer maturities (4-5 and 10 years). The proposed model captures the high volatility periods of the early 1980s well at longer maturities, and fits the low volatility periods well for all bond maturities. The unconditional correlation between model-implied and EGARCH$(1,1)$ volatility is 90% on average across maturities. We also compute the root mean squared errors (RMSEs) based on the difference between model-implied and model-free volatility. The RMSEs on average across all maturities is about 14 basis points. In summary, the proposed model performs well in fitting the conditional second moment of yields, even though it contains only one volatility factor.

We compare the implications of the proposed model with those from standard affine stochastic volatility models. We estimate a three-factor affine model with one factor driving the volatility process, the essentially affine $A_1(3)$ model (see Dai and Singleton, 2000; Duffee, 2002). To facilitate the comparison with our proposed model, we also consider a restricted version of the canonical $A_1(3)$ model, in which the dynamics of the yield curve factors are exactly the same as those in our proposed model. The critical difference with the new model is that in both the canonical and restricted $A_1(3)$ models, the volatility process depends on the level of the factors driving the yield curve. Consistent with the literature, we find that the two benchmark models do not perform as well in capturing the time-variation in the conditional yield volatilities. On average across all maturities, the unconditional correlation with the EGARCH estimates is 67% for the canonical $A_1(3)$ model and 66% for the restricted $A_1(3)$ model.[3] The proposed

[2] Andersen and Benzoni (2010) and Christensen, Lopez, and Rudebusch (2014) use realized volatility as a benchmark to assess the volatility fit of ATSMs. Cieslak and Povala (2016) use realized covariance models to extract stochastic volatility from a term structure model with multivariate volatility components. GARCH and EGARCH benchmarks are used by, among others, Bikbov and Chernov (2011), Collin-Dufresne, Goldstein, and Jones (2009), Dai and Singleton (2003).

[3] The performance of ATSMs in matching yields volatilities is somewhat model- and sample-dependent. We find a positive correlation as in Jacobs and Karoui (2009), because we also use a relative long sample of Treasury yields and include the high inflation period. In Andersen and Benzoni (2010) and Collin-Dufresne, Goldstein, and Jones (2009), the performance of these models is worse.

no-arbitrage model with GARCH volatility therefore substantially outperforms the benchmark models.

The benchmark models cannot capture the high volatility during the monetary experiment in the early 1980s. For short maturities, the estimated volatilities are similar before and after the monetary experiment for both benchmark models. The canonical and restricted $A_1(3)$ models overestimate yield volatility during the low-volatility period between the mid-1980s to 2000, and exhibit very limited time-variation across the maturity spectrum at the start of the sample. The proposed no-arbitrage model with GARCH volatility outperforms the two benchmark models by 33% on average across maturities in terms of yield volatility RMSE. These findings suggest that we can substantially improve the ability of ATSMs to fit yield volatility by endowing the volatility factor with a GARCH process. The state-of-the art stochastic volatility term structure models are not able to accurately model the dynamics of volatility, because by design a linear combination of the yield factors is not able to capture the second moment of yields. Note that the differences are not due to the discrete- versus continuous-time specification of the models. The improved performance of the GARCH models is due to their specification of the variance as a function of (lagged) squared *innovations* rather than as a function of the *levels* of the yield factors.

To demonstrate that the proposed model's improved ability in fitting conditional yield volatility does not come at the expense of a poor fit for the level of yields, we compare the in-sample yield fit of the proposed model and the benchmark models. The yield fit of the proposed model is very similar to that of the two benchmark stochastic volatility models. Moreover, we find that the proposed model captures the empirical patterns in bond risk premia, as characterized by the regression coefficients in the Campbell and Shiller (1991) regressions, very well. We confirm the finding of Dai and Singleton (2000) that benchmark no-arbitrage stochastic volatility models are unable to capture these deviations from the expectations hypothesis in the data. These findings indicate that the proposed model provides an adequate fit to the conditional means as well as the variances of yields.

If the GARCH model better captures yield volatility, it should outperform competing models for securities whose payoffs are highly volatility-dependent. We illustrate this by using the model estimates from Treasury yields to hedge against fluctuations in the prices of Treasury futures options. The first factor determines volatility in both the GARCH model and the competing $A_1(3)$ model, hence we hedge the exposure of the option contract to the first factor using the underlying Treasury yields. We confirm that the the performance of the GARCH model is vastly superior to that of the $A_1(3)$ model. When we compare the models' performance in hedging futures prices, which are not sensitive to volatility and are instead determined by the yield levels, the models' performance is very similar.

We contribute to several strands of literature. As mentioned above, an extensive literature has questioned the ability of ATSMs with stochastic volatility to model conditional volatility.[4] We show that the performance of term structure models in fitting conditional volatility can be improved greatly with a different specification of the volatility factor. This improvement does not come at the expense of the fit of the yield level. Our contribution is related to previous attempts to build ARCH/GARCH volatility into no-arbitrage term structure models. In particular, Heston and Nandi (2003) use a very similar setup. However, they calibrate their model using zero-coupon bond prices for a two-week sample. This limited empirical exercise does not allow them to analyze the modeling of yield volatility or the model's potential to resolve the tension between modeling yield levels and volatilities. We estimate a more general model using a long sample of Treasury yields data. We also explicitly compare the performance of the proposed model to that of state-of-the-art stochastic volatility models. Haubrich, Pennacchi, and Ritchken (2012) use a similar dynamic to investigate the real interest rate and inflation risk premia, but do not investigate the

[4] See for example Collin-Dufresne and Goldstein (2002), and Andersen and Benzoni (2010). See also Heidari and Wu (2003), Fan, Gupta, and Ritchken (2003), Jagannathan, Kaplin, and Sun (2003), and Li and Zhao (2006). See Bikbov and Chernov (2011), and Tang and Xia (2007) for studies using different fixed-income data.

implications of the model for yield volatility.[5]

Our work also complements a rich literature that provides alternative solutions to model interest rate volatility in a no-arbitrage framework. For example, Collin-Dufresne and Goldstein (2002) and Collin-Dufresne, Goldstein, and Jones (2009) propose an unspanned stochastic volatility model by imposing a set of parametric restrictions to break the spanning of conditional volatilities by yields.[6] However, Joslin (2018) shows that the unspanned stochastic volatility restrictions unduly constrain other aspects of model dynamics, and that these restrictions are rejected by the data. Ghysels, Le, Park, and Zhu (2014) impose a component GARCH volatility structure on a no-arbitrage term structure model. Their approach results in unspanned volatility under the physical measure, while our approach is much simpler and falls under the class of spanned volatility models. Finally, the class of models we propose is also related to an existing literature that goes beyond the affine paradigm. Examples include affine-quadratic models (see Ahn, Dittmar, and Gallant, 2002; Ahn, Dittmar, Gao, and Gallant, 2003; Leippold and Wu, 2002), regime-switching models (see for example Dai, Singleton, and Yang, 2007; Bansal and Zhou, 2002; Bansal, Tauchen, and Zhou, 2004; Ang and Bekaert, 2002), and other nonlinear models (see for example Ahn and Gao, 1999; Feldhütter, Heyerdahl-Larsen, and Illeditsch, 2018). We complement these studies by offering a parsimonious yet flexible model for capturing interest rate volatility.

The paper proceeds as follows. Section 2 presents the specification of term structure models with GARCH volatility and the benchmark ATSMs with stochastic volatility. Section 3 discusses the data. Section 4 discusses the parameter estimates for the term structure models, the models' performance in fitting the conditional volatility of yields, the model-implied term structure of unconditional yield volatility, and the trade-off between fitting the conditional means and variances of yields. Section 5 presents robustness results. Section 6 uses the model and the parameter

[5] For other related work, see Koeda and Kato (2015) and Realdon (2018).

[6] Creal and Wu (2015) provide new estimation procedures for ATSMs with spanned or unspanned stochastic volatility. They find that models with spanned volatility fit the cross section of the yield curve better, while those with unspanned volatility fit volatility better.

estimates from Treasury yields to hedge Treasury futures options, and Section 7 concludes.

2 Models

In this section, we first discuss the structure of the proposed ATSMs with GARCH volatility. Subsequently, we briefly discuss the models we use as empirical benchmarks throughout the paper. We use the canonical affine stochastic volatility models as specified in Dai and Singleton (2000), in which the conditional covariance of the state variables is an affine function of the level of the state variables. This class of models is motivated by a rich body of literature showing that the volatility of yield curve is, at least partially, related to the properties of the yield curve. For example, interest rate volatility is usually high when interest rates are high and when the yield curve exhibits higher curvature (see Cox, Ingersoll, and Ross, 1985; Litterman and Scheinkman, 1991; Longstaff and Schwartz, 1992). We also consider a restricted version of the canonical affine stochastic volatility model as an additional benchmark. In this model, the volatility process is an affine function of the state variables, and the dynamics of the state variables are the same as in the model we propose. The only difference between this benchmark and the newly proposed model is the specification of the volatility process.

2.1 ATSMs with GARCH Volatility

We propose a discrete-time term structure model with analytical solutions for bond prices, in which the volatility factors follow a GARCH process. This specification of the volatility process is motivated by the large literature on ARCH and GARCH modeling. There is considerable evidence that these processes provide a good description of interest rate volatility (Koedijk, Nissen, Schotman, and Wolff, 1997; Brenner, Harjes, and Kroner, 1996; Christiansen, 2005). The GARCH literature is formulated in discrete time, which facilitates model implementation.

Our approach retains the tractability of affine models while inheriting the ability of GARCH models to accurately capture time variation in yield volatility. The existing literature has concluded that at least three factors are needed to explain term structure dynamics (see for example Litterman and Scheinkman, 1991; Knez, Litterman, and Scheinkman, 1994). Accordingly, we use ATSMs with three latent state variables, with the following dynamics under the physical measure P and the risk-neutral measure Q:

$$X_{t+1} = K_0^P + K_1^P X_t + \sqrt{\Sigma_{t+1}}\epsilon_{t+1}, \tag{2.1}$$

$$X_{t+1} = K_0^Q + K_1^Q X_t + \sqrt{\Sigma_{t+1}}\epsilon_{t+1}, \tag{2.2}$$

$$r_t = \rho_0 + \rho_1 X_t, \tag{2.3}$$

where X_{t+1}, K_0^P and ϵ_{t+1} are 3×1 vectors, and K_1^P is a 3×3 diagonal matrix. r_t denotes the short rate, ρ_0 is a scalar, ρ_1 is a 1×3 vector, and ϵ_{t+1} is assumed to be distributed $N(0, I)$. The conditional covariance matrix Σ_{t+1} is a 3×3 diagonal matrix with the ith diagonal element $\sigma^2_{i,t+1}$ governed by a GARCH$(1,1)$ dynamic:

$$\sigma^2_{i,t+1} = \omega_i + \beta_i \sigma^2_{i,t} + \alpha_i \epsilon^2_{i,t}, \tag{2.4}$$

where $\epsilon_{i,t}$ is the ith element of vector ϵ_t. ω_i, β_i and α_i are scalars. To ensure that $\sigma^2_{i,t+1}$ is positive, we restrict ω_i, β_i and α_i to be positive numbers. $\sigma^2_{i,t+1}$ is known as of time t, given the history of $X_{i,t}$ and initial variance $\sigma_{i,0}$ as follows

$$\sigma^2_{i,t+1} = \omega_i + \beta_i \sigma^2_{i,t} + \alpha_i \frac{(X_{i,t} - K^P_{0(i)} - K^P_{1(i,i)} X_{i,t-1})^2}{\sigma^2_{i,t}}, \tag{2.5}$$

where $X_{i,t}$ is the ith state variable, $K^P_{0(i)}$ is the ith element of K_0^P, and $K^P_{1(i,i)}$ is the ith diagonal element of K_1^P. Note that this approach filters the time-varying volatility from the observations of the state variables in a very straightforward way. If $\beta_i = \alpha_i = 0$, the volatility for the ith state variable is constant over time.[7]

To link the physical and risk-neutral measures, we specify the pricing kernel to take the form

$$m_{t+1} = \exp(-r_t - \frac{1}{2}\lambda_t^{'}\lambda_t - \lambda_t^{'}\epsilon_{t+1}), \tag{2.6}$$

where λ_t is a 3×1 vector. We use an essentially affine specification for the price of risk (Duffee, 2002; Dai and Singleton, 2002; Cheridito, Filipović, and Kimmel, 2007), which gives:

$$\lambda_t = \left(\sqrt{\Sigma_t}\right)^{-1} (\lambda_0 + \lambda_1 X_t), \tag{2.7}$$

where λ_0 is a 3×1 vector, and λ_1 is a 3×3 diagonal matrix. The P- and Q-parameters in equations (2.1) and (2.2) are therefore related as follows:

$$\begin{aligned} K_0^Q &= K_0^P - \lambda_0, \\ K_1^Q &= K_1^P - \lambda_1. \end{aligned} \tag{2.8}$$

The model-implied price of a zero coupon bond $\widehat{P}_t^n$ with maturity n is given by

$$\widehat{P}_t^n = \exp\left(A_n(\Theta^Q) + B_n^{'}(\Theta^Q)X_t + \sum_i C_{i,n}(\Theta^Q)\sigma_{i,t+1}^2\right), \tag{2.9}$$

where $A_n(\Theta^Q)$, $B_n(\Theta^Q)$ and $C_{i,n}(\Theta^Q)$ are functions of the parameters $\Theta^Q = \{K_0^Q, K_1^Q, \rho_0, \rho_1, \omega_i, \beta_i, \alpha_i\}$ under the Q-dynamics, satisfying the following recursive relations

$$A_n = -\rho_0 + A_{n-1} + B_{n-1}^{'}K_0^Q + \sum_i \left(C_{i,n}\omega_i - \frac{1}{2}\log(1 - 2\alpha_i C_{i,n-1})\right), \tag{2.10}$$

[7] The option pricing literature uses a related dynamic due to Heston and Nandi (2000). That GARCH dynamic contains an additional parameter to capture the so-called leverage effect in index options, which is due to the skewness of the return distribution. We have estimated a more general version of (2.5) with this extra parameter. It does not lead to an improved fit, which intuitively is due to the limited skewness in the yield data.

$$B_n = -\rho_1^{'} + B_{n-1}^{'} K_1^Q, \tag{2.11}$$

$$C_{i,n} = \frac{B_{i,n-1}^2}{2(1 - 2\alpha_i C_{i,n-1})} + \beta_i C_{i,n-1}, \tag{2.12}$$

where A_n and $C_{i,n}$ are scalars, and B_n is a 3×1 vector with $B_{i,n}$ is the ith element. The initial conditions are $A_1 = -\rho_0$, $B_1 = -\rho_1^{'}$ and $C_{i,1} = 0$. The derivation of the recursive relations is provided in Appendix A. For the model with a single time-varying volatility factor, $\beta_i = \alpha_i = 0$ for $i = 2$ and 3. Therefore $C_{i,n} = \frac{1}{2} B_{i,n-1}^2$ for $i = 2$ and 3. For the model with two time-varying volatility factors, $\beta_i = \alpha_i = 0$ and $C_{i,n} = \frac{1}{2} B_{i,n-1}^2$ for $i = 3$.

The model-implied continuously compounded n-maturity yield $\widehat{y}_t^n$ is given by

$$\begin{aligned} \widehat{y}_t^n &= \overline{A}_n + \overline{B}_n^{'} X_t + \sum_i \overline{C}_{i,n} \sigma_{i,t+1}^2 \qquad (2.13) \\ &= \overline{A}_n + \overline{B}_n^{'} X_t + \sum_i \overline{C}_{i,n} \left(\omega_i + \beta_i \sigma_{i,t}^2 + \alpha_i \frac{(X_{i,t} - K_{0(i)}^P - K_{1(i,i)}^P X_{i,t-1})^2}{\sigma_{i,t}^2} \right), \end{aligned}$$

where $\overline{A}_n = -\frac{A_n}{n}$, $\overline{B}_n^{'} = -\frac{B_n^{'}}{n}$, and $\overline{C}_{i,n} = -\frac{C_{i,n}}{n}$.

We extract the conditional volatilities of yields using the filtered time-series of X_t and the estimated model parameters. The model-implied conditional variance of the n-maturity yield is given by

$$\widehat{var}_t(y_{t+1}^n) = \overline{B}_n^{'} var_t(X_{t+1}) \overline{B}_n + \sigma_e^2, \tag{2.14}$$

where σ_e^2 is the variance of the pricing errors. Appendix B provides the computation of the conditional variance based on the Kalman filter algorithm.

2.2 Canonical ATSMs with Stochastic Volatility

The main benchmark model considered in the paper is the widely used canonical affine stochastic volatility model. Using the classification of Dai and Singleton (2000), we denote the class of affine stochastic volatility models as $A_j(3)$, with $j = 1, 2$ or 3 factors driving the conditional variance of the three state variables. The instantaneous spot interest rate r_t is given by

$$r_t = \rho_0 + \rho_1 X_t, \tag{2.15}$$

where ρ_0 is a scalar and ρ_1 is a 1×3 vector. Most of this literature uses continuous-time specifications, and we follow this approach.[8] The state variables X_t follow an affine diffusion under the risk-neutral measure Q

$$dX_t = K^Q_{1\Delta}(K^Q_{0\Delta} - X_t)dt + \sqrt{\Sigma_t} dW^Q_t, \tag{2.16}$$

where $K^Q_{0\Delta}$ is a 3×1 vector, $K^Q_{1\Delta}$ is a 3×3 matrix, W^Q_t is a 3×1 vector of independent standard Brownian motions under the risk-neutral measure Q, and Σ_t is the conditional covariance matrix of X_t. Σ_t is a 3×3 diagonal matrix with the ith diagonal element given by

$$\sigma^2_{i,t} = a_i + b_i^{'} X_t, \tag{2.17}$$

where a_i is a scalar, and b_i is a 3×1 vector. We define $a = [a_1, a_2, a_3]$, which is a 3×1 vector, and $b = [b_1, b_2, b_3]$, which is a 3×3 matrix. In the $A_1(3)$ model, b_i is a vector of zeros for $i = 2$ and $i = 3$, and in the $A_2(3)$ model, b_i is a vector of zeros for $i = 3$. In the canonical $A_j(3)$ model, all three state variables have a time-varying conditional variance, which is an affine function of $j = 1, 2$ or 3 state variables.

The model-implied price of a zero coupon bond $\widehat{P}^n_t$ with maturity n is given by (see Duffie and Kan, 1996)

$$\widehat{P}^n_t = \exp\left(A_n(\Theta^Q) + B_n^{'}(\Theta^Q) X_t\right), \tag{2.18}$$

where $A_n(\Theta^Q)$ and $B_n(\Theta^Q)$ are functions of the parameters under the Q-dynamics, $\Theta^Q = \{K^Q_{0\Delta}, K^Q_{1\Delta}, \rho_0, \rho_1, a, b\}$, through a set of Ricatti ordinary differential equations. The model-

[8] See Le, Singleton, and Dai (2010) for related discrete-time models.

implied continuously compounded n-maturity yield $\widehat{y}_t^n$ is given by

$$\widehat{y}_t^n = \overline{A}_n + \overline{B}_n^{'} X_t, \tag{2.19}$$

where $\overline{A}_n = -\frac{A_n}{n}$, and $\overline{B}_n^{'} = -\frac{B_n'}{n}$.

The pricing kernel π_t is given by

$$\frac{d\pi_t}{\pi_t} = -r_t dt - \Lambda_t' dW_t^P, \tag{2.20}$$

where W_t^P is a 3×1 vector of independent standard Brownian motions under the physical measure P, and the 3×1 vector Λ_t denotes the market price of risk. We adopt the essentially affine specification for the price of risk as in Duffee (2002) and Dai and Singleton (2002).[9]

$$\Lambda_t = \sqrt{\Sigma_t}\lambda_0 + \sqrt{\Sigma_t^-}\lambda_1 X_t, \tag{2.21}$$

where λ_0 is a 3×1 vector and λ_1 is a 3×3 matrix. In $A_j(3)$ models, the diagonal matrix Σ_t^- has zeros in its first j entries and $(a_i + b_i' X_t)^{-1}$ for $i = j + 1,...,3$.

The dynamics of the state variables under the physical measure P can be written in terms of Λ_t and equation (2.16)

$$dX_t = K_{1\Delta}^Q (K_{0\Delta}^Q - X_t)dt + \sqrt{\Sigma_t}\Lambda_t dt + \sqrt{\Sigma_t} dW_t^P. \tag{2.22}$$

The physical dynamic in the essentially affine model is then given by

$$dX_t = K_{1\Delta}^P (K_{0\Delta}^P - X_t)dt + \sqrt{\Sigma_t} dW_t^P, \tag{2.23}$$

[9] Jacobs and Karoui (2009) show that the specification of the price of risk has a minimal impact on modeling conditional volatility.

where

$$K^P_{1\Delta} = K^Q_{1\Delta} - \begin{pmatrix} \lambda_{01} b_1' \\ \lambda_{02} b_2' \\ \lambda_{03} b_3' \end{pmatrix} - I^- \lambda_1,$$

$$K^P_{1\Delta} K^P_{0\Delta} = K^Q_{1\Delta} K^Q_{0\Delta} + \begin{pmatrix} a_1 \lambda_{01} \\ a_2 \lambda_{02} \\ a_3 \lambda_{03} \end{pmatrix}.$$

We denote element i of λ_0 by λ_{0i}. We define I^- as a 3×3 diagonal matrix. The ith diagonal element $I^-_i = 1$ if the ith diagonal element of Σ^-_t is nonzero. If the ith diagonal element of Σ^-_t is zero, we have $I^-_i = 0$

We follow the Dai and Singleton identification scheme to ensure the $\sigma^2_{i,t}$ are strictly positive for all i.[10] The $A_j(3)$ models are different from our proposed model in the parameterization for both the state variables and the volatility process. In the proposed GARCH model, the feedback matrix K^P_1 is a diagonal matrix. In the canonical $A_j(3)$ models, following the admissibility constraints of Dai and Singleton (2000),[11]

$$K^P_{1\Delta} = \begin{bmatrix} K^P_{1\Delta j\times j} & 0_{j\times(3-j)} \\ K^P_{1\Delta(3-j)\times j} & K^P_{1\Delta(3-j)\times(3-j)} \end{bmatrix}, \tag{2.24}$$

where $K^P_{1\Delta j\times j}$ is a $j \times j$ matrix, $K^P_{1\Delta(3-j)\times j}$ is a $(3-j)\times j$ matrix, and $K^P_{1\Delta(3-j)\times(3-j)}$ is a $(3-j)\times(3-j)$ matrix. For the volatility process in equation (2.17),

$$a = \begin{bmatrix} 0_{j\times 1} \\ 1_{(3-j)\times 1} \end{bmatrix}, \tag{2.25}$$

[10] The identification constraints can be applied to either the P- or Q-parameters, see Dai and Singleton (2000), and Singleton (2006).

[11] We refer to Dai and Singleton (2000) equations 15-19 for details on the admissibility restrictions. Joslin and Le (2020) show that for no-arbitrage affine term structure models, these admissibility constraints give rise to tension in joint estimation of the physical and risk-neutral dynamics.

$$b = \begin{bmatrix} I_{j\times j} & b_{j\times(3-j)} \\ 0_{(3-j)\times j} & 0_{(3-j)\times(3-j)} \end{bmatrix}. \tag{2.26}$$

We provide more details on the estimation of these stochastic volatility models in Appendix C.

To facilitate the comparison with our proposed model, we also consider a restricted version of the canonical affine stochastic volatility model as a benchmark model. In the restricted model, we constrain $K_{1\Delta}^P$ and $K_{1\Delta}^Q$ to be diagonal matrices.[12] This model is nested within the canonical specification. The dynamics of the state variables under the restricted model are similar to those in our proposed model. The only difference is the specification of the volatility process. In the restricted $A_j(3)$ model, the conditional variance of the state variables is a linear combination of the levels of $j = 1, 2$ or 3 state variables. The difference with our model is therefore a rather subtle and technical one, and exclusively due to the volatility dynamic.

3 Data

We first discuss the sample of Treasury yields we use to estimate the term structure models. We then discuss the sample of Treasury futures and Treasury futures options data we use for the hedging application.

3.1 Treasury Yield Data

We use monthly data on continuously compounded zero-coupon bond yields with maturities of three and six months, and one, two, three, four, five, and ten years. The three- and six-months yields are obtained from the Federal Reserve Economic Data. The one to five, and ten year yields are from the Gürkaynak, Sack, and Wright (2007, GSW) dataset. The sample period is from November 1971 to October 2019.[13]

[12] To get diagonal $K_{1\Delta}^Q$, restrictions are imposed on the specification of the market price of risk.

[13] The GSW dataset is obtained from the Federal Reserve, available at `http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html`. We use November 1971 as the start date because it is the earliest

Panel A of Table 1 reports the summary statistics of the yields, Panel B reports the realized volatilities, and Panel C reports the EGARCH(1, 1) volatilities. Both measures of "true" volatility are maturity-specific. Note that we do not have high-frequency data available to construct realized volatility over the entire sample period for which yields are available. We therefore follow the approach pioneered by Schwert (1989) in the equity return literature and construct measures of monthly realized volatility using within-month squared changes in daily yields. We present results for one- to five- and ten-year maturities, because three- and six-month daily yields are not available for our sample. The EGARCH(1, 1) is estimated assuming that the conditional mean of changes in monthly yields is generated by an AR(1) process.

On average, the yield curve is upward sloping, and both EGARCH and realized volatilities are hump-shaped as a function of maturity. The yields for all maturities are highly persistent, and exhibit positive skewness and mild excess kurtosis at short maturities. A comparison of the two benchmarks shows that the realized volatilities are much less persistent than the EGARCH(1, 1) volatilities for all maturities. Both volatility measures exhibit excess kurtosis and positive skewness for all maturities.

3.2 Treasury Futures and Futures Options Data

We document the models' performance in hedging derivatives using monthly data on U.S. Treasury futures contracts and options on these futures. The futures and options data are obtained from the Chicago Mercantile Exchange (CME) for the May 1988 to June 2016 sample period. We rely on futures contracts on underlying Treasuries with five- and ten-year maturities. We use options on these underlying futures and require that the option data are available in two consecutive months in order to compute changes in option prices. We also impose the following filters on the option data: i) The option price exceeds 5 cents; ii) Open interest is positive; iii) Maturity is less than 270 days; and iv) Moneyness is between 0.95 and 1.1.

date with uninterrupted availability of 10-year yield data.

These futures and options contracts have some features that are critically important for pricing. Most importantly, the coupon rate for the Treasury underlying these futures contracts is fixed at 8% prior to February 2000 and at 6% after February 2000. We divide the futures data into three groups based on maturity: 1-3 months, 3-6 months, and 6-9 months. Panels A and B of Table 2 present sample summary statistics for the prices of futures contracts on five- and ten-year Treasuries. The futures contracts are quoted in terms of percentage of par. The average percentage prices are not very different across maturities. The averages and medians exceed one hundred percent, which is due to the fact that in our sample period, the fixed coupons exceed the prevailing market yields.

Panels C and D of Table 2 present summary statistics on the implied volatilities of the option contracts in the sample. Implied volatilities are computed using the Black model. Similar to the futures contracts, we divide the sample into three separate maturities for each of the underlying Treasury maturities. The implied volatilities exhibit a smile pattern in the moneyness dimension. As expected, the average implied volatilities are larger for futures options written on ten-year Treasuries compared to futures options on five-year Treasuries. The average at-the-money volatility for the futures options on five-year Treasuries is between 3.27% and 5.71%. The average at-the-money volatility for futures options on ten-year Treasuries is between 5.00% and 6.81%.

4 Term Structure Models: Empirical Results

To retain a low-dimensional structure for the affine term structure model, we focus on a model with one time-varying volatility factor in our empirical analysis. That is, we set $i = 1$ in equation (2.4) in the proposed term structure model with GARCH volatility. Only the first state variable has a time-varying variance. We refer to this model as the no-arbitrage GARCH model. Our main benchmark model is therefore the canonical $A_1(3)$ model with one factor driving the conditional variances of the state variables. Dai and Singleton (2000) conclude that this model offers the best characterization of unconditional yield volatilities and a sufficiently flexible correlation structure among three-factor models with stochastic volatility. Note that this model has a time-varying variance for all three state variables. However, the dynamic of the variance is driven by only one state variable. We also consider a restricted version of the canonical $A_1(3)$ model, where the feedback matrix of the state variables is diagonal, as in the no-arbitrage GARCH model.

In this section, we first discuss how we estimate the term structure models. We then present the parameter estimates for these three models, with a focus on comparing the models' ability to explain the conditional volatility of the yield curve. Subsequently, we examine the unconditional volatility implied by these models. We also document the tension between matching yields and yield volatilities in these models, and we discuss their implications for the expectations hypothesis.

4.1 Estimation Method

The term structure model with GARCH volatility can be expressed using a state-space representation. The observed yield curve $y_t = \widehat{y}_t + e_t$ is the measurement equation, where $\widehat{y}_t$ is the model-implied yield as specified in equation (2.13), and e_t is a vector of measurement errors that is assumed to be $i.i.d.$ normal. We assume that the error variance σ_e^2 is the same across maturities to ensure that all maturities receive similar weight in the likelihood. The state equation is given by equation (2.1). We apply the Kalman filter to the state-space representation of the model. We estimate the parameters $\Theta = \{K_0^P, K_1^P, K_0^Q, K_1^Q, \rho_0, \rho_1, \omega_i, \beta_i, \alpha_i\}$ and filter the state variables X_t using maximum likelihood. The log likelihood of the tth observation is

$$\begin{aligned} \log f_t(\Theta) &= const - \frac{N}{2}\log(\sigma_e^2) - \frac{1}{2}\frac{\|e_t\|^2}{\sigma_e^2} - \frac{1}{2}\log(\det(\Sigma_t)) \\ &\quad -\frac{1}{2}(X_t - K_0^P - K_1^P X_{t-1})'\Sigma_t(X_t - K_0^P - K_1^P X_{t-1}). \end{aligned} \tag{4.1}$$

N denotes the number of available yields in the term structure. In our sample, $N = 8$. $\|e_t\|$ denotes the Euclidean norm of the vector of measurement errors. Appendix B provides more detailed information on the Kalman filter algorithm.[14] The estimation of the benchmark $A_1(3)$ model is discussed in Appendix C.

4.2 Parameter Estimates

Table 3 presents parameter estimates for the no-arbitrage GARCH model (Panel A), the canonical $A_1(3)$ model (Panel B), and the restricted $A_1(3)$ model (Panel C). The no-arbitrage GARCH model has 22 parameters, including the variance σ_e^2 of the measurement errors. The canonical $A_1(3)$ model has 24 parameters, and the restricted $A_1(3)$ model has 14 parameters. In the no-arbitrage GARCH model and the restricted $A_1(3)$ model, the conditional mean under the P-measure is more restricted than in the canonical $A_1(3)$ model. For all models, the estimates satisfy the admissibility conditions under the physical measure P. The state variables in the canonical and restricted $A_1(3)$ models follow a first order VAR process when sampled monthly. To facilitate the comparison with the estimates from the discrete-time no-arbitrage GARCH model, we report the estimated parameters of the discretized VAR process for both benchmark models.

The time series properties of the state variables critically depend on the speed of mean reversion in the feedback matrix K_1^P. For the no-arbitrage GARCH model, the first state variable, which has time-varying volatility, is highly persistent. The second and third state variables are less persistent than the first. The third state variable reverts much more quickly under the Q-measure compared to the P-measure.

Some of the implications of the canonical and restricted $A_1(3)$ models are similar to the no-arbitrage GARCH model. The comparison with the restricted $A_1(3)$ model is more convenient because the structure of the model is similar. Specifically, the K_1^P and K_1^Q matrices are also

[14] See Duffee and Stanton (2012) and Christoffersen, Dorion, Jacobs and Karoui (2014) for estimation using the Kalman filter.

diagonal, which facilitates the interpretation of factor persistence. Similar to the GARCH model, the first state variable is the most persistent in the restricted $A_1(3)$ model and plays the role of level factor, whereas the third variable is strongly mean reverting. However, in contrast to the no-arbitrage GARCH model, the third factor is more mean-reverting under the P-measure in the restricted $A_1(3)$ model.

The structure of volatility in the no-arbitrage GARCH model critically differs from that in the benchmark models. In the canonical and restricted $A_1(3)$ models, the level factor mainly affects the volatility of the second factor, as indicated by the second entry in the b vector. Duffee (2002) and Jacobs and Karoui (2009) find similar results. In our benchmark no-arbitrage GARCH model on the other hand, by design it is the volatility of the first (level) factor that plays a critical role. The estimated volatility dynamic is very persistent ($\beta = 0.9064$). We investigate no-arbitrage GARCH models with additional volatility factors in Section 5.2.

4.3 Conditional Yield Volatility

In this section, we examine the properties of the model-implied conditional yield volatilities. First, consider simple unconditional correlations between the variance factors from the models and the two measures of the "true" variance. Recall that the maturity-specific realized variance is constructed using within-month squared changes in daily yields. The EGARCH model is estimated assuming that the conditional mean of changes in yields follows an AR(1) process, and is also maturity-specific. The variance factor from the no-arbitrage GARCH model has the highest unconditional correlation with the realized variance (65%) as well as the EGARCH estimates (84%). The correlations with the realized variance for the two benchmark models are approximately 46%. The correlations with the EGARCH(1, 1) variance for the two benchmark models are approximately 54%.

Next we present the model-implied conditional volatilities together with the realized volatilities in Figure 1 and with the EGARCH(1, 1) volatilities in Figure 2. Appendices B and C discuss the derivation of the implied one-month conditional volatilities from the no-arbitrage GARCH model and the $A_1(3)$ model respectively. We observe a close correspondence between the realized volatility and the EGARCH volatility, but the realized volatilities are more noisy than the EGARCH(1,1) volatilities for all maturities.[15] This is not surprising, because the realized volatilities are ex-post realizations as opposed to conditional expectations.

The estimated conditional volatilities from the canonical and restricted $A_1(3)$ models are very similar and much less variable than the realized and EGARCH volatilities. The estimated volatility at the short end is similar before and after the monetary experiment in the early 1980s for both benchmark models. In addition, both benchmark models overestimate yield volatility when volatility is low, from the mid-1980s to 2000, and they do not exhibit sufficient variation in yield volatility at the beginning of the sample. At the very short end (3-month and 6-month) and very long end (10-year) of the yield curve, the estimated conditional volatilities from the restricted $A_1(3)$ model exhibit excess movement in the last decade of the sample.

The estimated conditional volatilities from the no-arbitrage GARCH model are more variable than the estimates from the two benchmark models. They comove closely with the realized and EGARCH volatilities at longer maturities (4-5 and 10 years), but do not perform as well for short maturities. For longer maturities, the no-arbitrage GARCH model fits the high volatility periods of the early 1980s well. The canonical and restricted $A_1(3)$ models, on the other hand, cannot capture the high volatility periods for any maturity. The no-arbitrage GARCH model also does a better job in fitting the low volatility periods in our sample (from the mid-1980s to 2000) than the two benchmark models for all maturities. Moreover, it is able to capture the time variation in yield volatility at the beginning of our sample for all maturities. In summary, the no-arbitrage GARCH model appears to capture the time-variation in the second moment of yields quite well, especially at longer maturities.

To further assess the quality of the estimated conditional volatilities, Panel A of Table 4

[15] We do not report realized volatility results for the three- and six-month maturities due to data availability.

reports the unconditional correlations between model-implied and realized volatilities, and Panel B reports the unconditional correlations between model-implied and EGARCH(1, 1) volatilities. A first observation is that the correlations with both realized and EGARCH volatilities are positive at all maturities for all models. The estimated conditional volatilities from the no-arbitrage GARCH model have the highest correlation with both the realized volatilities and the EGARCH estimates at all maturities. For example, the unconditional correlation with the EGARCH(1, 1) volatilities is as high as 95% for the 3-year yield, while it is about 71% for the two benchmark models. The estimated conditional volatilities from the two benchmark models have similar correlations with both realized and EGARCH volatilities at all maturities. On average across all maturities, the unconditional correlation with realized volatilities for the no-arbitrage GARCH model is 67%, for the canonical $A_1(3)$ model it is 46%, and for the restricted $A_1(3)$ model it is 45%. The unconditional correlation with EGARCH(1, 1) volatilities for the no-arbitrage GARCH model is 90% on average across all maturities, for the canonical $A_1(3)$ model it is 67%, and for the restricted $A_1(3)$ model it is 66%.[16]

To provide additional insight into the models' ability to fit volatility, we also examine the root mean squared errors (RMSEs) of model-implied conditional volatilities. Panel A of Table 5 reports the RMSEs between model-implied and realized volatilities, and Panel B of Table 5 reports the RMSEs between model-implied and EGARCH volatilities. Both are expressed in basis points. The no-arbitrage GARCH model outperforms both benchmark models in fitting the volatility across all maturities for both "true" volatility measures. The model does a particularly good job at the intermediate and long end of the yield curve, although it only has a single volatility factor. For example, for 3-5 and 10 years, the RMSEs based on EGARCH volatilities are below 10 basis points.

[16] As discussed in Jacobs and Karoui (2009), the performance of the stochastic volatility model in fitting yield volatility is sensitive to the model and sample under consideration. We find a positive correlation as in Jacobs and Karoui (2009), because we also use a relative long sample of Treasury yields and include the high inflation period. Andersen and Benzoni (2010), and Collin-Dufresne, Goldstein, and Jones (2009) do not find a significant positive correlation.

The RMSE improvement of the no-arbitrage GARCH model over the two benchmark models is about 30% on average across maturities when realized volatility is used as the measure of true volatility. When EGARCH$(1,1)$ volatility is used as the measure of true volatility, the RMSE improvement over the two $A_1(3)$ models is about 33% on average across maturities. The improvement of the RMSEs between model-implied and EGARCH volatilities is more significant at longer maturities. For example, for the 5-year yield, the improvement in RMSEs is about 55%. In summary, the no-arbitrage GARCH model performs better in capturing the time variability of conditional volatility despite using a single volatility factor. This finding suggests that the improvements mainly result from the GARCH specification of the volatility process.

4.4 The Term Structure of Unconditional Volatility

We investigate model-implied unconditional volatility. Figure 3 presents the term structure of unconditional yield volatility implied by the no-arbitrage GARCH model and the canonical and restricted $A_1(3)$ models, and compares these with the "model-free" term structures of unconditional realized and EGARCH$(1,1)$ volatility. The unconditional $A_1(3)$ and no-arbitrage GARCH model volatilities are computed as the averages of the conditional volatility path generated by each model.

Figure 3 shows that the no-arbitrage GARCH implied term structure matches the EGARCH$(1,1)$ implied term structure and the term structure of realized volatility much more closely than the benchmark $A_1(3)$ models. The term structures implied by the two benchmark models are very similar. The volatility curve from these models monotonically decreases as a function of maturity and volatility is too high at all maturities.

4.5 The Trade-off Between Fitting Yield Levels and Volatilities

The tension between matching the first and second moments of Treasury yields in ATSMs has been documented in many studies (Dai and Singleton, 2000, 2002; Duffee, 2002; Duarte, 2004;

Joslin and Le, 2020). In particular, Dai and Singleton (2002) note "a tension in matching simultaneously the historical properties of the conditional means and variances of yields". Joslin and Le (2020) study the mechanism underlying this tension, and argue that imposing a spanning condition may prevent a no-arbitrage model from fully capturing the predictability patterns of bond yields in the data.

The no-arbitrage GARCH model belongs to the class of spanned affine models. Since the GARCH volatility is one of the factors that determine bond yields, it is spanned by the yields. The estimation of this model could therefore be subject to the same tension. We now show that the improved volatility fit is not obtained at the expense of poor yield fit. Panel A of Table 6 reports the RMSEs of yields in basis points for the no-arbitrage GARCH model and the two benchmark models. On average across all maturities, the yield fit is similar for the three models. The in-sample RMSE of yields for the no-arbitrage GARCH model with a single volatility factor is about 19 basis points on average across different maturities. This finding suggests that the improvement of the no-arbitrage GARCH model in fitting conditional volatility, as shown in Section 4.3, does not come at the cost of fitting the conditional mean of yields.

For comparison, we also present the fit of yields for Gaussian models with constant variance-covariance matrix. We consider the canonical representation of Joslin, Singleton, and Zhu (2011, henceforth referred to as JSZ), which allows for stable and tractable estimation of the $A_0(3)$ three-factor Gaussian model.[17] Panel B of Table 6 shows the RMSEs for the maximum flexible specification of the JSZ model and also for three restricted models. In the first restricted model, we use a diagonal variance-covariance matrix with constant variance. In the second restricted model, we set the $(1, 2)$ and $(1, 3)$ entries of the feedback matrix (K_1) to zero and we also restrict the variance-covariance matrix to be diagonal and the variance to be constant. This restricted version is comparable to the canonical $A_1(3)$ model. In the third restricted JSZ specification, we restrict the feedback matrix (K_1) and the constant variance-covariance matrix to be diagonal.

[17] We refer to JSZ (2011) for implementation details.

This restricted form is more comparable to the no-arbitrage GARCH model. Recall that under the newly proposed model, the feedback matrix is a diagonal matrix.

The maximum flexible $A_0(3)$ model provides the best in-sample fit of yields. This is not surprising, since the model has the richest specification for the conditional mean of the state variables. All three restricted forms have marginally higher average RMSEs than the maximum flexible model.

Overall the Gaussian models outperform the models with time-varying volatility in Panel A for the purpose of fitting yields. However, the Gaussian models are by design unable to capture the time variation of yield volatility. These findings show that a tension remains between matching yields and yield volatilities in the proposed model. In comparison to the Gaussian models, all models in Panel A provide a worse fit to the cross-section of yield levels to acquire flexibility in fitting conditional variances. This is inevitable, since in these non-Gaussian models, the state variables driving both yields and yield volatilities are the same and are spanned by the cross section of yields. In the canonical $A_1(3)$ model, all three state variables are non-Gaussian. The non-Gaussian state variables must be positive and enter the conditional variance. As discussed in Joslin and Le (2020), this admissibility constraint creates a tension between fitting the yields and yield volatilities. In the no-arbitrage GARCH model, we have a different specification for the factor with time-varying volatility. The variance of this factor follows a GARCH specification, which results in a very simple admissibility constraint. This simpler constraint does not seem to result in a deterioration of the fit for the conditional mean of yields compared to the benchmark $A_1(3)$ models.

4.6 The Expectations Hypothesis

We conclude that the in-sample fit of the no-arbitrage GARCH model is as good as that of the benchmark models. We now examine the ability of the the no-arbitrage GARCH model to capture the time series properties of the yield data by focusing on the predictability patterns observed in the data. Campbell and Shiller (1991) show that under the expectations hypothesis, a regression coefficient of $\varphi_n = 1$ obtains in the following regression

$$y_{t+1}^{n-1} - y_t^n = \psi_n + \varphi_n \left(\frac{y_t^n - y_t^1}{n-1} \right) + e_{t+1}^n, \tag{4.2}$$

where y_t^n is the n-month yield at time t. However, Dai and Singleton (2002) show that actual estimates of φ_n are negative. They then show that the Gaussian model is consistent with these patterns in the data. No-arbitrage stochastic volatility models are not able to match the observed pattern.

We conduct this regression analysis using our sample and the yields implied by the different models we investigate. Figure 4 presents the results based on a twelve-month holding period. We find deviations from the expectations hypothesis, consistent with the existing literature. The estimated coefficients are all negative, and more negative for longer maturities. Consistent with Dai and Singleton (2002), we find that the Gaussian model is able to capture this pattern, but the canonical and restricted $A_1(3)$ models are not. In contrast to the canonical stochastic volatility models, the proposed no-arbitrage GARCH model can match the empirical patterns of bond risk premia as characterized by the regression coefficients.

We conclude that not only do the improvements provided by the no-arbitrage GARCH for the purpose of fitting yield volatilities not come at the expense of fitting the yield level, the model is also able to rationalize the deviations from the expectations hypothesis observed in the data.

5 Robustness

In this section, we first investigate an alternative model of the realized variance as a measure of "true volatility" to evaluate model performance. Next, we discuss the performance of no-arbitrage GARCH models with multiple volatility factors. Finally, we examine the performance of the no-arbitrage GARCH when the factors are the first three principal components.

5.1 An Alternative Realized Variance Measure

As discussed in Section 3.1, we construct the monthly realized variance using within-month squared changes in daily yields. Assuming that M observations are available within a month, the estimate of the monthly variance for the n-maturity yield is thus computed as

$$var(y_{t+1}^n) = \sum_{m=1}^{M} \left(\Delta y_{t+m/M}^n\right)^2. \tag{5.1}$$

As discussed above, the resulting measures are somewhat noisy because they are ex-post realizations as opposed to conditional expectations. We now repeat our analysis using the (ex ante) conditional expectation of the realized variance instead. The literature has shown that an ARMA$(1,1)$ provides a good fit to the logarithm of the realized variance:

$$\log(var(y_{t+1}^n)) = \gamma \log(var(y_t^n)) + \delta\varepsilon_t + \varepsilon_{t+1}, \tag{5.2}$$

where ε_{t+1} is assumed to be distributed $N(0, \sigma_{t,\varepsilon}^2)$. We refer to this model as the realized variance model. The model-implied one-month conditional variance of the n-maturity yield is then given by:

$$\widehat{var}_t(y_{t+1}^n) = (var(y_t^n))^{\gamma} \exp\left(\delta\varepsilon_t + \frac{\sigma_{t,\varepsilon}^2}{2}\right). \tag{5.3}$$

Figure 5 presents the one-month conditional volatilities implied by this realized variance model and the three no-arbitrage models with time-varying volatility. The estimates from the no-arbitrage GARCH model comove closely with those from the realized variance model. On the other hand, both benchmark models mostly overestimate yield volatilities between 1980 and 2000. These findings are consistent with those in Figures 1 and 2. Panel A of Table 7 reports the unconditional correlations between the one-month conditional volatilities implied by the realized variance model and those implied by the three models. Correlations are again positive at all maturities for all models, but the no-arbitrage GARCH model has a much higher unconditional correlation than the canonical and restricted $A_1(3)$ models for all maturities. Panel B of Table 7 presents the RMSEs of model-implied volatilities when the conditional volatility implied by the realized variance model is used as a model-free volatility measure. Consistent with the findings in Table 5, the no-arbitrage GARCH model outperforms both the canonical and restricted $A_1(3)$ models in fitting conditional volatility for all maturities.

Overall, we conclude that our results in Section 4.3 are robust to the measurement of the realized variance. The no-arbitrage GARCH model performs well for the purpose of modeling conditional volatility.

5.2 Models with Multiple Volatility Factors

In this section, we investigate the performance of no-arbitrage GARCH models with multiple volatility factors. Figure 6 presents the results.[18] The performance of no-arbitrage GARCH models with two and three volatility factors is similar to that of the model with a single volatility factor. The implied conditional volatilities of yields are closely related to both the realized volatilities and the EGARCH estimates, especially at longer maturities. This finding suggests that the first volatility factor plays a dominant role in fitting conditional yield volatility. Table 8 presents the parameter estimates for the no-arbitrage GARCH models with two and three volatility factors as well as the log-likelihood. The estimates of the feedback matrix for the state variables are similar to those in Table 3 for the model with one volatility factor. The likelihood ratio tests indicate that the models with two and three volatility factors do not fit the data better than the model with one volatility factor.

In the model with three volatility factors, the second and third volatility factors mean revert more quickly than the first factor. The estimated mean reversion parameters for the three volatility factors are $\beta_1 = 0.9021$, $\beta_2 = 0.7314$, and $\beta_3 = 0.8431$ respectively. The first variance factor has the highest unconditional correlation with both the realized variance (65%) and the

[18] For ease of exposition, we only present results for four different maturities for realized volatility and EGARCH volatility respectively. The results for the remaining maturities are reported in the internet appendix.

EGARCH estimates (83%).

For the second and third variance factors, the correlations with the realized variance are about 43% and 42% respectively. The correlations with the EGARCH variance are about 64% and 61% respectively for the second and third variance factors. However, these factors are very small in magnitude and do not greatly improve fit.

5.3 The No-Arbitrage GARCH Model with Principal Components as Factors

The main advantage of the GARCH framework is that the filtering of the volatility process is relatively straightforward. The question therefore arises if we can further simplify the model by pre-specifying the factors, rather than filtering them. The term structure literature has demonstrated that Gaussian models fit the data well when using the principal components of the yields as factors. Motivated by this finding, we now investigate the performance of the no-arbitrage GARCH model when its factors are the first three principal components, rather than factors estimated using the Kalman filter. Since the state variables are observable in this case, filtering is no longer required in the estimation. Consistent with the main model analyzed above, we use a model in which only the first principal component has a time-varying variance, and we compare the performance of this model with that of the canonical and restricted $A_1(3)$ models. For these two benchmark models, the results are the same as in Tables 4-6. Panels A and B of Table 9 report the models' performance in matching yield volatility when the realized volatility is used to measure the "true" conditional volatility, while Panels C and D show the results when the EGARCH$(1,1)$ volatility is used.

The estimated conditional volatilities from the no-arbitrage GARCH model with three principal components have the highest correlation with both the realized volatilities and the EGARCH estimates at all maturities. On average across all maturities, the unconditional correlation with realized volatilities for the no-arbitrage GARCH model with three principal components is

68%. The unconditional correlation with EGARCH(1, 1) volatilities for the no-arbitrage GARCH model with three principal components is 87% on average across all maturities. These results are very similar to those of the proposed no-arbitrage GARCH model with three latent state variables as shown in Table 4.

However, while the no-arbitrage GARCH model with three principal components outperforms both benchmark models in fitting the volatility across all maturities for both "true" volatility measures, these improvements are smaller than those of the no-arbitrage GARCH model with three latent state variables as shown in Table 5. For example, the RMSE improvement of the no-arbitrage GARCH model with three principal components over the two benchmark models is about 18% on average across maturities when realized volatility is used as the measure of true volatility, while the improvement for the no-arbitrage GARCH model with three latent state variables is about 30% on average across maturities.

6 Hedging Treasury Futures Options

We first provide a brief discussion of the literature on fixed income option pricing. We then propose a hedging exercise that uses Treasury futures options and is designed to exploit superior modeling of the volatility dynamics, and we provide empirical results for the GARCH model and the competing canonical $A_1(3)$ model. For comparison, we first conduct a similar hedging exercise on Treasury futures, which unlike futures options are not sensitive to volatility.

6.1 Fixed Income Derivatives

The prices of options and many other derivatives are very sensitive to volatility. It therefore stands to reason that improved volatility modeling should lead to more accurate option prices and superior performance when hedging against fluctuations in option prices. To remain as close as possible to the models and parameter estimates in Section 4, we focus on Treasury futures options. For both the newly proposed model with GARCH volatility as well as the competing $A_1(3)$ model, futures prices are available in closed form and option prices can be readily computed. It is therefore straightforward to use the parameters in Section 4 that are estimated from Treasuries for this hedging exercise.

The existing literature on the pricing of Treasury futures options is very limited. Most of the literature on fixed income option pricing instead uses caps or swaption data. See for instance Almeida, Graveline, and Joslin (2011), Collin-Dufresne and Goldstein (2002), Han (2007), Li and Zhao (2006), Heidari and Wu (2009), and Trolle and Schwartz (2009) for examples.[19] While the newly proposed model with GARCH volatility can be used for pricing and hedging swaptions, it is not obvious how to use the parameter estimates from Section 4 for this purpose. We therefore perform a hedging exercise using Treasury futures and options on Treasury futures, using either the model with GARCH volatility or the competing $A_1(3)$ model. Our prior is that the performance of the two models may not be very different when hedging futures, because the prices of these contracts are not very sensitive to volatility. However, we expect the model with GARCH volatility to outperform the $A_1(3)$ model for the purpose of hedging Treasury futures options.

6.2 Hedging Treasury Futures

Our empirical implementation has focused on a GARCH model where only the first factor has time-varying volatility. In the $A_1(3)$ model, by definition only the first factor determines volatility. Both models thus attempt to capture volatility with the first factor, and we therefore focus the hedging exercise on this factor. We first consider an application that hedges the exposure of the futures contracts to the first factor using the underlying Treasury yields. Specifically, we hedge

[19] A growing literature studies other aspects of Treasury futures options. Bakshi, Crosby and Gao (2022) study Treasury option returns. Cieslak and Povala (2016) use option-implied volatilities to improve the estimation of term structure models. Beber and Brandt (2006) and Cremers, Fleckenstein, and Gandhi (2021) study the relation between implied volatility of Treasury options and economic data. Choi, Mueller, and Vedolin (2017) document variance risk premiums, while Bauer and Chernov (2021) study option-implied skewness.

the exposure of the futures contract on the five- (ten-) year Treasury using the five- (ten-) year maturity yield.[20] The initial hedged portfolio is thus defined as:

$$H_t(T) = F_{t,t+\tau}(T) - w_t y_t(T),$$

where τ indicates the maturity of the futures contract, T indicates the maturity of the underlying treasury, $F_{t,t+\tau}(T)$ indicates the time t dollar futures price, and $y_t(T)$ indicates the yield with maturity T. Because we examine the models' performance in hedging exposure to the first factor, we set w_t such that the futures contract exposure to the first factor is hedged using the underlying yield; that is, the change in the initial portfolio value for a given change in $X_{1,t}$ should be zero. This weight w_t is obtained by solving the following equation:

$$\frac{\partial H_t(T)}{\partial X_{1,t}} = \frac{\partial F_{t,t+\tau}(T)}{\partial X_{1,t}} - w_t \frac{\partial y_t(T)}{\partial X_{1,t}} = 0,$$

which gives:

$$w_t = \frac{\frac{\partial F_{t,t+\tau}(T)}{\partial X_{1,t}}}{\frac{\partial y_t(T)}{\partial X_{1,t}}}.$$

Each month t, we compute the derivatives $\frac{\partial F_{t,t+\tau}(T)}{\partial X_{1,t}}$ and $\frac{\partial y_t(T)}{\partial X_{1,t}}$ for both models using the parameters reported in Table 3. Given the estimated weight w_t, we then compute the predicted change in the value of the futures contract and compare it with the actual change in the futures prices. Specifically, the change predicted by the model is defined as $w_t(y_{t+1}(T) - y_t(T))$, while the actual change is the one-month change in the futures price $F_{t+1,t+\tau}(T) - F_{t,t+\tau}(T)$. If the model performs well in estimating the weight required to hedge the exposure to the first factor, we expect the predicted and actual changes to be similar.

Figure 7 presents the results for this exercise that hedges futures contracts with underlying Treasury maturities of five or ten years. For both underlying Treasury maturities, we report the

[20] We hedge using the yield for ease of implementation, because yields are linear in the state variables. Note that hedging using yields is equivalent to hedging using the logarithm of bond prices.

hedging performance for the futures contracts in the 1-3, 3-6, and 6-9 month maturity buckets. The figures scatter plot the predicted changes against the actual changes in the futures prices. If a model performs well, the data should be on the 45-degree line, which is included for reference. The first and third rows of Figure 7 depict the performance of the $A_1(3)$ model, while the second and fourth rows depict the performance of the GARCH model.

Figure 7 clearly indicates that the data nicely line up along the 45 degree line for both models, across all maturities. Both models perform better for futures contracts with longer maturities. We conclude that there is no significant difference in the performance of the two models when hedging the exposure of the futures contract to the first factor. This confirms our prior, since both models perform adequately for the purpose of fitting the underlying Treasury yields, and futures prices are largely determined by the level of Treasury yields.

6.3 Hedging Treasury Futures Options

We now present the results of a similar hedging exercise, but this time we attempt to hedge the exposure of Treasury futures options. The volatility factor, and by extension the specification of the volatility dynamic, is much more critical for option prices compared to futures prices. We therefore expect a model that better captures volatility dynamics to outperform in this dimension. We perform a hedging exercise using options written on Treasury futures contracts with underlying 5- or 10-year bond maturities. For all options with maturity exceeding one month, we compute the option hedge ratio as follows:

$$w_t^O = \frac{\frac{\partial O_{t,t+\tau}(T)}{\partial X_{1,t}}}{\frac{\partial y_t(T)}{\partial X_{1,t}}}$$

where $O_{t,t+\tau}(T)$ is the τ-maturity option on an underlying futures contract based on the T-maturity Treasury. The derivative in the numerator is computed using simulation for both models. We simulate 10,000 paths for each state variable at the monthly frequency, using the parameters from Table 3. This gives us the distribution of the futures prices at the option maturities, which in turn allows us to compute the option prices. We numerically shock the state variable by 1% on both sides and compute the option prices and the corresponding derivatives with respect to the state variable. Similar to the futures hedging exercise, we compute the change predicted by the model as $w_t^O(y_{t+1}(T) - y_t(T))$ and compare it with the actual change in the option price $O_{t+1,t+\tau}(T) - O_{t,t+\tau}(T)$ to examine the models' ability to hedge fluctuations in option prices using the underlying Treasuries.

Figure 8 scatter plots the model-predicted changes against the actual changes in the option prices. Options written on futures on five-year Treasuries are on the left and options on futures on ten-year Treasuries are on the right. Figure 8 clearly indicates that the GARCH model substantially outperforms the $A_1(3)$ model for the purpose of hedging option price fluctuations. While the data of course deviate much more from the 45-degree line compared to the futures hedging exercise in Figure 7, they line up much closer to the 45 degree line for the GARCH model. The scatter plot is much more dispersed in the case of the $A_1(3)$ model.

Table 10 provides additional perspective on these differences in model performance. We regress the actual changes in option prices on the changes predicted by the models. If the model provides a good hedge, we expect high R^2s, a zero intercept, and a loading of one on the model predictions. Panel A of Table 10 presents the results for the $A_1(3)$ model and Panel B presents the results for the GARCH model. We report results based on the full sample (all options) as well as results for put and call options separately. For the $A_1(3)$ model, the regression R^2 is close to zero and the loading on the model change is close to zero, suggesting very limited predictive power and very poor hedging performance. For the GARCH model, we obtain R^2s ranging between 75% and 82% for options on futures contracts with five-year Treasuries as the underlying. The regression loadings are between 0.91 and 1.00. The model performs somewhat worse for options on futures contracts with ten-year Treasuries as the underlying, but its performance remains dramatically better than that of the $A_1(3)$ model.

We conclude that the GARCH model performs substantially better in hedging Treasury futures options compared to the $A_1(3)$ model. This improved performance is due to the fact that option prices are very sensitive to volatility, and the GARCH model vastly outperforms the $A_1(3)$ model for the purpose of volatility modeling.

7 Conclusion

In the term structure literature, state-of-the-art models face difficulties in simultaneously fitting the time variation in yield levels and volatilities. We propose parsimonious yet flexible models with closed-form solutions that outperform benchmark models in this dimension. Our results suggest that the performance of ATSMs in matching yield volatility critically depends on the specification of the volatility dynamics. In standard ATSMs with stochastic volatility, the volatility dynamic is a linear combination of the *levels* of the yield curve factors. We instead propose a no-arbitrage term structure model where the volatility factor is written as a function of the (lagged) squared *innovations* to the yields.

The model combines the tractability of ATSMs with improved modeling of yield volatility. We estimate the model using monthly yield data from 1971 to 2019, and find that the model-implied conditional volatility is highly correlated with measures of "model-free" volatility, especially at longer maturities. The correlation between model-implied and model-free yield volatility is between 80% and 95%. The model significantly outperforms benchmark stochastic volatility models for the purpose of fitting yield volatility. It also provides a good fit to the conditional mean of yields, suggesting that the improved volatility fit is not obtained at the expense of yield fit. These findings are robust to various variations in the empirical setup. We illustrate how these improvements in volatility modeling lead to dramatically better performance when hedging Treasury futures options.

It is worth emphasizing that our approach may not be the only one that provides improved modeling of conditional volatility. Indeed, it may be possible to construct better models, and we plan to address this in future work. Our objective in this paper is merely to show that it is possible to write down parsimonious term structure models that allow for improved modeling of yield volatility as compared to state-of-the art models in this literature, without sacrificing the model's ability to fit the level of yields.

Appendix A. Bond Valuation in ATSMs with GARCH Volatility

To derive the recursions in equations (2.10), (2.11) and (2.12), we first note that the price of a one-period bond, $n = 1$, is as follows

$$\begin{aligned} P(t, t+1) &= E_t^Q \left[\exp(-r_t)\right] \\ &= \exp(-\rho_0 - \rho_1 X_t). \end{aligned} \tag{A.1}$$

Suppose that the price of a n-period bond is given by $P_t^n = \exp\left(A_n + B_n^{'} X_t + \sum_i C_{i,n}\sigma_{i,t+1}^2\right)$. Matching coefficients gives $A_1 = -\rho_0$, $B_1 = -\rho_1^{'}$ and $C_{i,1} = 0$. In order to solve for A_n, B_n and $C_{i,n}$ we derive the bond price under the risk neutral probability measure

$$\begin{aligned} P_t^n &= E_t^Q[\exp(-r_t)P_{t+1}^{n-1}] \\ &= E_t^Q\left[\exp\left(-\rho_0 - \rho_1 X_t\right)\exp\left(A_{n-1} + B_{n-1}^{'} X_{t+1} + \sum_i C_{i,n-1}\sigma_{i,t+2}^2\right)\right] \\ &= \exp\left(-\rho_0 - \rho_1 X_t + A_{n-1} + B_{n-1}^{'}(K_0^Q + K_1^Q X_t) + \sum_i C_{i,n-1}\omega_i + \sum_i C_{i,n-1}\beta_i\sigma_{i,t+1}^2\right) \\ &\quad \times E_t^Q\left[\exp\left(\sum_i \left(B_{i,n-1}\sigma_{i,t+1}\epsilon_{i,t+1} + C_{i,n-1}\alpha_i\epsilon_{i,t+1}^2\right)\right)\right]. \end{aligned} \tag{A.2}$$

Completing the square in the portion to which the expectation applies, using the fact that for a standard normal z, $E\left(a(z+b)^2\right) = \exp\left(-\frac{1}{2}\log(1-2a) + \frac{ab^2}{1-2a}\right)$, and matching coefficients results in the recursive relations in equations (2.10), (2.11), and (2.12).

Appendix B. Implementing ATSMs with GARCH Volatility

This appendix summarizes the derivation of the conditional variance for the affine model in a GARCH framework. The contemporaneous forecast of the state vector and its corresponding covariance matrix are denoted by $X_{t|t}$ and $P_{t|t}$. The Kalman filter algorithm proceeds as follows at any time t:

1. Given $X_{t|t}$ and $P_{t|t}$, compute the one-period ahead forecast of the state vector and its corresponding covariance matrix[21]

$$X_{t+1|t} = K_0^P + K_1^P X_{t|t}, \tag{B.1}$$

and

$$P_{t+1|t} = K_1^{P'} P_{t|t} K_1^P + \Sigma_{t+1|t}. \tag{B.2}$$

The volatility factor can be computed based on equation (2.5)

$$\begin{aligned} \sigma^2_{i,t+2|t} &= \omega_i + \beta_i \sigma^2_{i,t+1|t} + \alpha_i \frac{E_t^P(X_{i,t+1} - K^P_{0(i)} - K^P_{1(i,i)} X_{i,t})^2}{\sigma^2_{i,t+1|t}} \\ &= \omega_i + \beta_i \sigma^2_{i,t+1|t} + \alpha_i. \end{aligned} \tag{B.3}$$

[21] We use the first two unconditional moments in the first step of the recursion.

2. Compute the one-period ahead forecast of y_{t+1} and its corresponding covariance matrix

$$
\begin{aligned}
y_{t+1|t} &= \overline{A}+\overline{B}^{'}X_{t+1|t}+\sum_{i}\overline{C}_i\sigma^2_{i,t+2|t} \\
&= \overline{A}+\overline{B}^{'}X_{t+1|t}+\sum_{i}\overline{C}_i\left(\omega_i+\beta_i\sigma^2_{i,t+1|t}+\alpha_i\right),
\end{aligned} \tag{B.4}
$$

where $y_{t+1|t}$ is an $N\times 1$ vector, $\overline{A}$ and $\overline{C}_i$ are $N\times 1$ vectors, and $\overline{B}$ is a $3\times N$ matrix, where N denotes the number of available yields in the term structure. Furthermore,

$$
V_{t+1|t}=\overline{B}^{'}P_{t+1|t}\overline{B}+R, \tag{B.5}
$$

where R is an $N\times N$ diagonal matrix. We assume that the variance of the pricing errors σ^2_e on the diagonal is the same across maturities.

3. Compute the forecast error of $y_{t+1,}$ $e_{t+1|t}=y_{t+1}\ -y_{t+1|t}$.

4. Update the contemporaneous forecast of the state vector and its corresponding covariance matrix

$$
X_{t+1|t+1}=X_{t+1|t}+P_{t+1|t}\overline{B}V^{-1}_{t+1|t}e_{t+1|t}, \tag{B.6}
$$

$$
P_{t+1|t+1}=P_{t+1|t}-P_{t+1|t}\overline{B}V^{-1}_{t+1|t}\overline{B}^{'}P_{t+1|t}. \tag{B.7}
$$

and compute the smoothed volatility factor

$$
\sigma^2_{i,t+2|t+1}=\omega_i+\beta_i\sigma^2_{i,t+1}+\alpha_i\frac{(X_{i,t+1}-K^P_{0(i)}-K^P_{1(i,i)}X_{i,t})^2}{\sigma^2_{i,t+1}}. \tag{B.8}
$$

5. Return to the first step.

The model-implied conditional variance of yields is computed as

$$
\widehat{var}_t(y_{t+1})=diag(V_{t+1|t}). \tag{B.9}
$$

Note that in the empirical investigation, we use the conditional variance of yield differences, which is also equal to equation (B.9).

Appendix C. Estimation of Canonical ATSMs with Stochastic Volatility

This appendix summarizes the method used to estimate canonical affine stochastic volatility models. A three-factor latent model can be expressed using a state-space representation. The observed yields $y_t = \widehat{y}_t + e_t$ constitute the measurement equation, where $\widehat{y}_t$ is the model-implied yield as specified in equation (2.19), and e_t is a vector of measurement errors that is assumed to be *i.i.d.* normal. We assume that the errors have equal variance σ_e^2 across maturities to ensure similar weights in the likelihood. The state equation (2.23) can be discretized as $X_{t+1} = K_0^P + K_1^P X_t + \epsilon_{t+1}^P$, where $\epsilon_{t+1|t}^P$ is assumed to be distributed $N(0, \Sigma_t)$. We estimate the P- and Q-parameters simultaneously by applying the Kalman filter to the state-space representation. We use quasi-maximum likelihood (QML) as implemented by Jacobs and Karoui (2009).

The contemporaneous forecast of the state vector and its corresponding covariance matrix are denoted by $X_{t|t}$ and $P_{t|t}$. The Kalman filter algorithm works as follows at any time t:

1. Given $X_{t|t}$ and $P_{t|t}$, compute the one-period ahead forecast of the state vector and its corresponding covariance matrix[22]

$$X_{t+1|t} = K_0^P + K_1^P X_{t|t}, \tag{C.1}$$

and

$$P_{t+1|t} = K_1^{P'} P_{t|t} K_1^P + \Sigma_{t|t}. \tag{C.2}$$

[22] We use the first two unconditional moments in the first step of the recursion.

where the volatility factor is computed based on equation (2.17)

$$\sigma_{i,t|t}^{2} = a_i + b_i^{'} X_t, \tag{C.3}$$

where $\sigma_{i,t|t}^{2}$ is the ith diagonal element of the 3×3 matrix $\Sigma_{t|t}$.

2. Compute the one-period ahead forecast of y_{t+1} and its corresponding covariance matrix

$$y_{t+1|t} = \overline{A} + \overline{B}^{'} X_{t+1|t}, \tag{C.4}$$

where $y_{t+1|t}$ is a $N \times 1$ vector, $\overline{A}$ is a $N \times 1$ vector, $\overline{B}$ is a $3 \times N$ matrix, and N denotes the number of available yields in the term structure.

$$V_{t+1|t} = \overline{B}^{'} P_{t+1|t} \overline{B} + R, \tag{C.5}$$

where R is a $N \times N$ diagonal matrix. We assume that the variance of the pricing errors σ_e^2 on the diagonal is the same across maturities.

3. Compute the forecast error of $y_{t+1},\ e_{t+1|t} = y_{t+1} - y_{t+1|t}$.

4. Update the contemporaneous forecast of the state vector and its corresponding covariance matrix

$$X_{t+1|t+1} = X_{t+1|t} + P_{t+1|t} \overline{B} V_{t+1|t}^{-1} e_{t+1|t}, \tag{C.6}$$

and

$$P_{t+1|t+1} = P_{t+1|t} - P_{t+1|t} \overline{B} V_{t+1|t}^{-1} \overline{B}^{'} P_{t+1|t}. \tag{C.7}$$

5. Return to the first step.

The quasi log-likelihood for observation $t+1$ is then

$$\log f_t(\Theta) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log(\det(V_{t+1|t})) - \frac{1}{2}e'_{t+1|t}V_{t+1|t}e_{t+1|t}. \tag{C.8}$$

The model-implied conditional variance of yields is computed as

$$\widehat{var}_t(y_{t+1}) = diag(V_{t+1|t}). \tag{C.9}$$

Note that in the empirical investigation, we use the conditional variance of yield differences, which is also equal to equation (C.9).

References

[1] Ahn, D. H., Dittmar, R. F., and Gallant, A. R. (2002). Quadratic term structure models: Theory and evidence. Review of Financial Studies, 15(1), 243-288.

[2] Ahn, D. H., Dittmar, R. F., Gallant, A. R., and Gao, B. (2003). Purebred or hybrid?: Reproducing the volatility in term structure dynamics. Journal of Econometrics, 116(1-2), 147-180.

[3] Ahn, D. H., and Gao, B. (1999). A parametric nonlinear model of term structure dynamics. Review of Financial Studies, 12(4), 721-762.

[4] Andersen, T., and Benzoni, L. (2010). Do bonds span volatility risk in the US Treasury market? A specification test for affine monetary policy regimes and the term structure of interest rates. Journal of Finance, 65(2), 603-653.

[5] Almeida, C., Graveline, J., Joslin, S. (2011). Do options contain information about excess bond returns. Journal of Econometrics, 164, 35–44.

[6] Ang, A., and Bekaert, G. (2002). Regime switches in interest rates. Journal of Business & Economic Statistics, 20(2), 163-182.

[7] Bakshi, G., Crosby, J., and Gao, X. (2022). The puzzle of Treasury option returns. Working Paper, University of Maryland.

[8] Bansal, R., Tauchen, G., and Zhou, H. (2004). Regime-shifts in term structure, expectations hypothesis puzzle, and the real business cycle. Journal of Business and Economic Statistics, 22(4), 396-409.

[9] Bansal, R., and Zhou, H. (2002). Term structure of interest rates with regime shifts. Journal of Finance, 57(5), 1997-2043.

[10] Bauer, M., and Chernov, M. (2021). Interest rate skewness and biased beliefs. Working Paper, UCLA.

[11] Beber, A., and Brandt, M. (2006). The effect of macroeconomic news on beliefs and preferences: Evidence from the options market. Journal of Monetary Economics, 53(8), 1997-2039.

[12] Bikbov, R., and Chernov, M. (2011). Yield curve and volatility: Lessons from the Eurodollar futures and options. Journal of Financial Econometrics, 9(1), 66-105.

[13] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

[14] Brenner, R. J., Harjes, R. H., and Kroner, K. F. (1996). Another look at models of the short-term interest rate. Journal of Financial and Quantitative Analysis, 31(01), 85-107.

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

[16] Cheridito, P., Filipović, D., and Kimmel, R. L. (2007). Market price of risk specifications for affine models: theory and evidence. Journal of Financial Economics, 83(1), 123-170.

[17] Choi, H., Mueller, P., Vedolin, A. (2017). Bond variance risk premiums. Review of Finance, 21(3), 987-1022.

[18] Christensen, J. H., Lopez, J. A., and Rudebusch, G. D. (2014). Can spanned term structure factors drive stochastic yield volatility? Working Paper, Federal Reserve Bank of San Francisco.

[19] Christiansen, C. (2005). Multivariate term structure models with level and heteroskedasticity effects. Journal of Banking and Finance, 29(5), 1037-1057.

[20] Christoffersen, P., Dorion, C., Jacobs, K., and Karoui, L. (2014). Nonlinear Kalman filtering in affine term structure models. Management Science, 60(9), 2248-2268.

[21] Cieslak, A., and Povala, P. (2016). Information in the term structure of yield curve volatility. Journal of Finance, 71(3), 1393-1436.

[22] Collin-Dufresne, P., and Goldstein, R. S. (2002). Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility. Journal of Finance, 57(4), 1685-1730.

[23] Collin-Dufresne, P., Goldstein, R. S., and Jones, C. S. (2009). Can interest rate volatility be extracted from the cross section of bond yields? Journal of Financial Economics, 94(1), 47-66.

[24] Cox, J. C., Ingersoll Jr, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates. Econometrica, 53(2), 385-407.

[25] Creal, D. D., and Wu, J. C. (2015). Estimation of affine term structure models with spanned or unspanned stochastic volatility. Journal of Econometrics, 185(1), 60-81.

[26] Cremers, M., Fleckenstein, M., and Gandhi, P. (2021). Treasury yield implied volatility and real activity. Journal of Financial Economics, 140(2), 412–445.

[27] Dai, Q., and Singleton, K. J. (2000). Specification analysis of affine term structure models. Journal of Finance, 55(5), 1943-1978.

[28] Dai, Q., and Singleton, K. J. (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. Journal of Financial Economics, 63(3), 415-441.

[29] Dai, Q., and Singleton, K. J. (2003). Term structure dynamics in theory and reality. Review of Financial Studies, 16(3), 631-678.

[30] Dai, Q., Singleton, K. J., and Yang, W. (2007). Regime shifts in a dynamic term structure model of US treasury bond yields. Review of Financial Studies, 20(5), 1669-1706.

[31] Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. Journal of Finance, 57(1), 405-443.

[32] Duffee, G. R., and Stanton, R. H. (2012). Estimation of dynamic term structure models. Quarterly Journal of Finance, 2(2), 1-51.

[33] Duffie, D. and Kan, R. 1996. A yield-factor model of interest rates. Mathematical Finance, 6(4), 379-406.

[34] Duarte, J. (2004). Evaluating an alternative risk preference in affine term structure models. Review of Financial Studies, 17(2), 379-404.

[35] Engle, R. (1982). ARCH with estimates of variance of United Kingdom inflation. Econometrica, 50(4), 987-1007.

[36] Fan, R., Gupta, A., and Ritchken, P. (2003). Hedging in the possible presence of unspanned stochastic volatility: Evidence from swaption markets. Journal of Finance, 58(5), 2219-2248.

[37] Feldhütter, P., Heyerdahl-Larsen, C., and Illeditsch, P. (2018). Risk premia and volatilities in a nonlinear term structure model. Review of Finance, 22(1), 337-380.

[38] Ghysels, E., Le, A., Park, S., and Zhu, H. (2014). Risk and return trade-off in the U.S. Treasury market. Working Paper, University of North Carolina.

[39] Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The US Treasury yield curve: 1961 to the present. Journal of Monetary Economics, 54(8), 2291-2304.

[40] Han, B. (2007). Stochastic volatilities and correlations of bond yields. Journal of Finance, 62(3), 1491-1524.

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

[42] Heidari, M., and Wu, L. (2003). Are interest rate derivatives spanned by the term structure of interest rates?. Journal of Fixed Income, 13(1), 75-86.

[43] Heidari, M., and Wu, L. (2009). A joint framework for consistently pricing interest rates and interest rate derivatives. Journal of Financial and Quantitative Analysis, 44(3), 517-550.

[44] Heston, S. L., and Nandi, S. (2000). A closed-form GARCH option valuation model. Review of Financial Studies, 13(3), 585-625.

[45] Heston, S. L., and Nandi, S. (2003). A two-factor term structure model under GARCH volatility. Journal of Fixed Income, 13(1), 87-95.

[46] Jacobs, K., and Karoui, L. (2009). Conditional volatility in affine term-structure models: Evidence from Treasury and swap markets. Journal of Financial Economics, 91(3), 288-318.

[47] Jagannathan, R., Kaplin, A., and Sun, S. (2003). An evaluation of multi-factor CIR models using LIBOR, swap rates, and cap and swaption prices. Journal of Econometrics, 116(1-2), 113-146.

[48] Joslin, S. (2018). Can unspanned stochastic volatility models explain the cross section of bond volatilities? Management Science, 64(4), 1707-1726.

[49] Joslin, S., and Le, A. (2020). Interest rate volatility and no-arbitrage term structure models. Forthcoming, Management Science.

[50] Joslin, S., Singleton, K. J., and Zhu, H. (2011). A new perspective on Gaussian dynamic term structure models. Review of Financial Studies, 24(3), 926-970.

[51] Knez, P. J., Litterman, R., and Scheinkman, J. (1994). Explorations into factors explaining money market returns. Journal of Finance, 49(5), 1861-1882.

[52] Koeda, J., and Kato, R. (2015). The role of uncertainty in the term structure of interest rates: A GARCH-ATSM approach. Applied Economics, 47(34-35), 3710-3722.

[53] Koedijk, K. G., Nissen, F. G., Schotman, P. C., and Wolff, C. C. (1997). The dynamics of short-term interest rate volatility reconsidered. European Finance Review, 1(1), 105-130.

[54] Le, A., Singleton, K. J., and Dai, Q. (2010). Discrete-time affine$\mathbb{Q}$ term structure models with generalized market prices of risk. Review of Financial Studies, 23(5), 2184-2227.

[55] Leippold, M., and Wu, L. (2002). Asset pricing under the quadratic class. Journal of Financial and Quantitative Analysis, 37(2), 271-295.

[56] Li, H., and Zhao, F. (2006). Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. Journal of Finance, 61(1), 341-378.

[57] Litterman, R. B., and Scheinkman, J. (1991). Common factors affecting bond returns. Journal of Fixed Income, 1(1), 54-61.

[58] Longstaff, F. A., and Schwartz, E. S. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. Journal of Finance, 47(4), 1259-1282.

[59] Realdon, M. (2018). Discrete time affine term structure models with squared Gaussian shocks (DTATSM-SGS). Working Paper, Brunel University London. Available at SSRN 3350750.

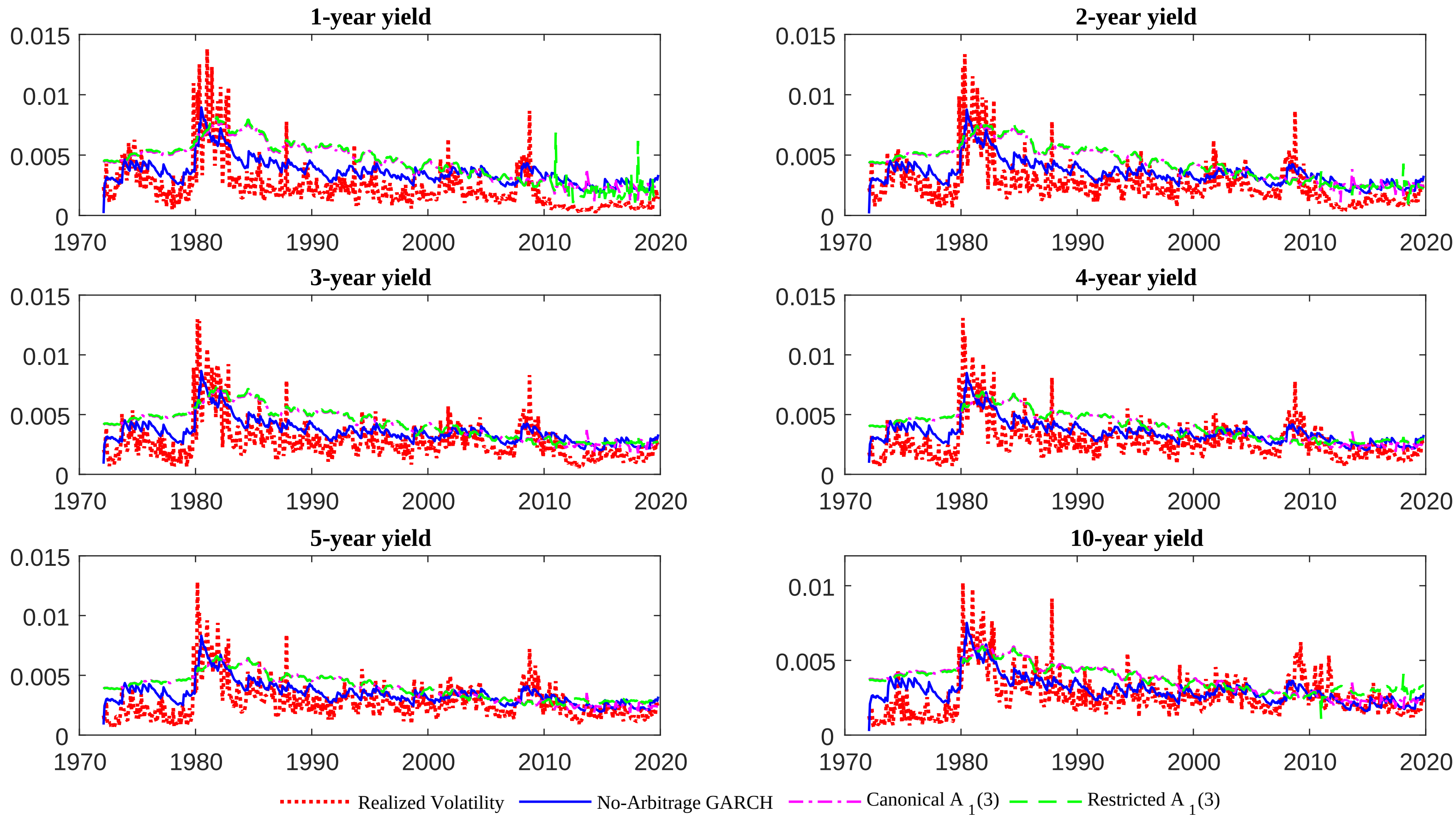
[60] Schwert, G. W. (1989). Why does stock market volatility change over time? Journal of Finance, 44(5), 1115-1153.

[61] Singleton, K. J. (2006). Empirical dynamic asset pricing. Princeton University Press, Princeton, NJ

[62] Tang, H., and Xia, Y. (2007). An international examination of affine term structure models and the expectations hypothesis. Journal of Financial and Quantitative Analysis, 42(1), 41-80.

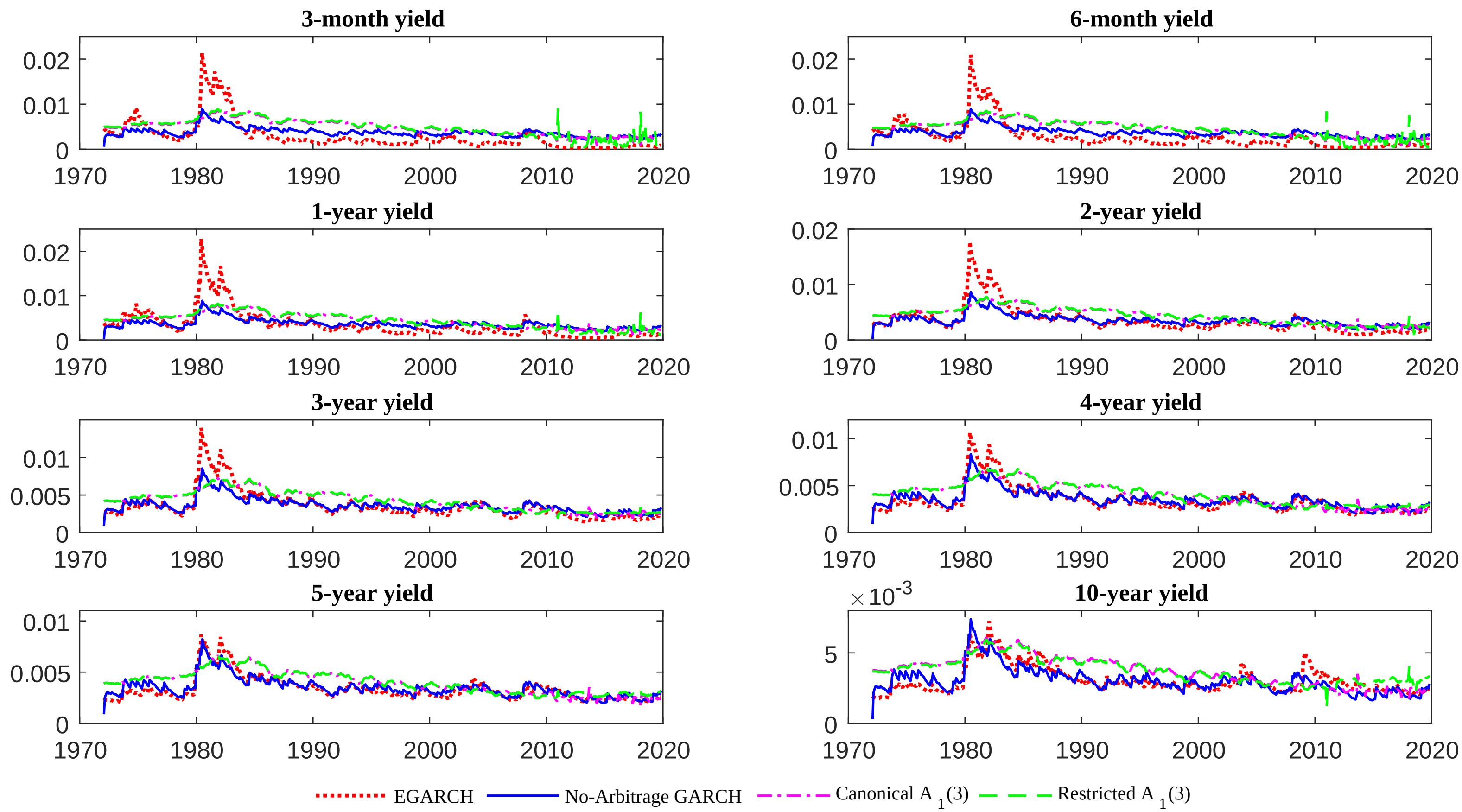
[63] Trolle, A., and Schwartz, E. (2009). A general stochastic volatility model for the pricing of interest rate derivatives. Review of Financial Studies, 22(5), 2007–2057.

Figure 1: Model-Implied Conditional Volatility by Maturity: Comparison with Realized Volatility.



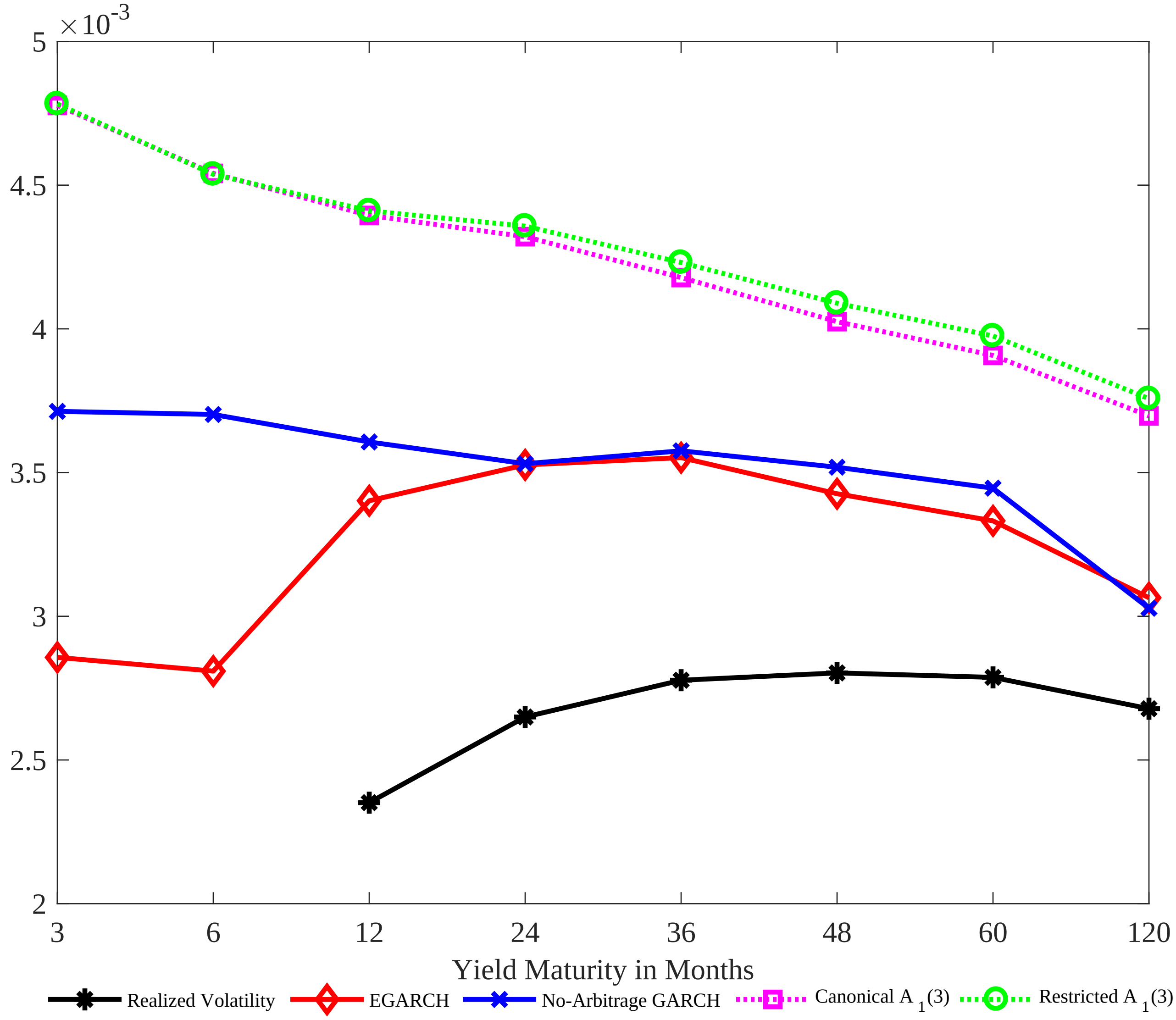
Notes to Figure: We plot the implied one-month conditional yield volatility for different models. For each maturity, the dotted line (red) represents "model-free" volatility, measured by realized volatility. We construct maturity-specific monthly realized variances using within-month squared changes in daily yields. The solid line (blue) represents the conditional volatility from the no-arbitrage GARCH model. The dash-dot line (magenta) represents the conditional volatility from the canonical $A_1(3)$ model. The dashed line (green) represents the conditional volatility from the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix.

Figure 2: Model-Implied Conditional Volatility by Maturity: Comparison with EGARCH Volatility.



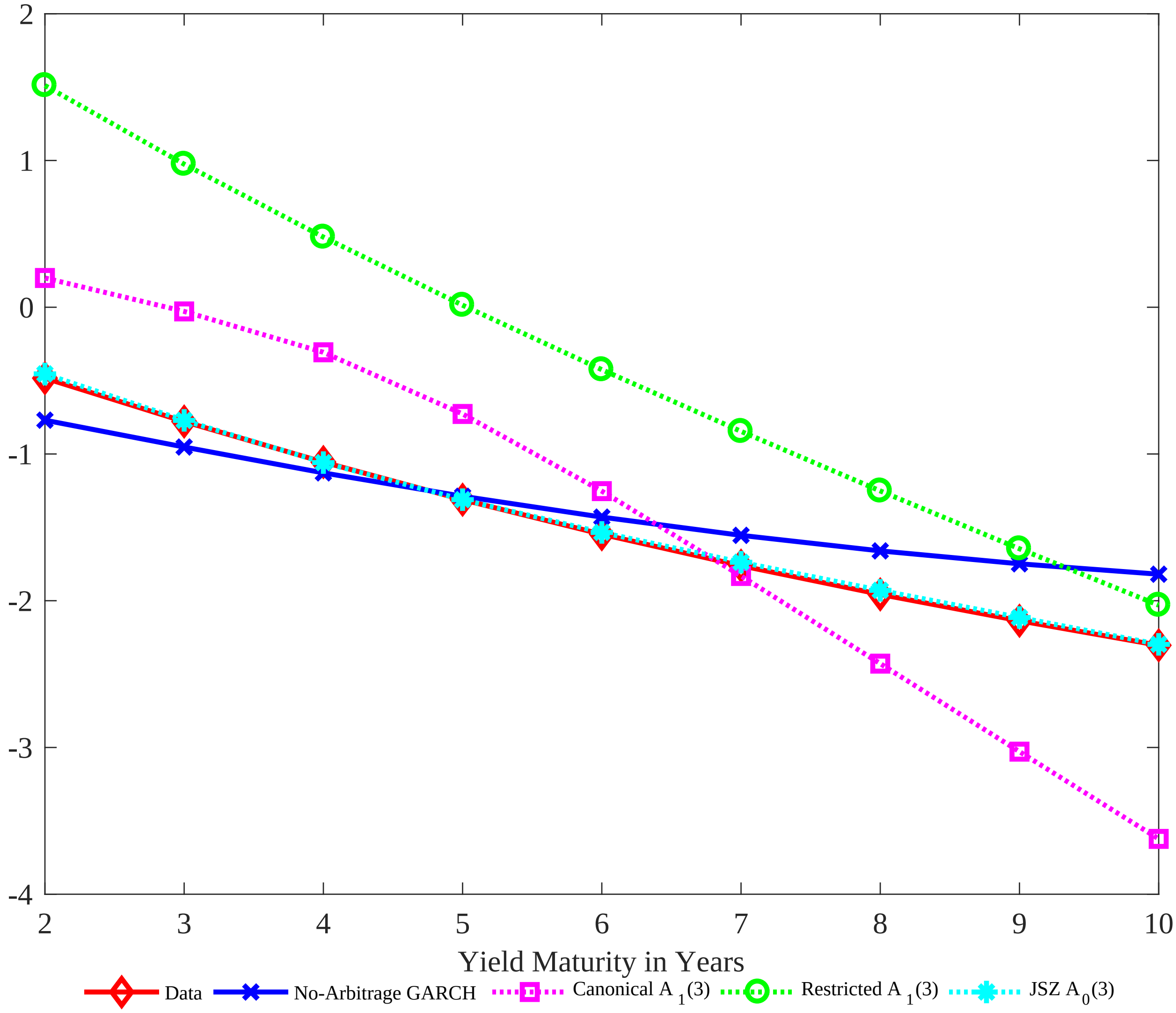
Notes to Figure: We plot the implied one-month conditional yield volatility for different models. For each maturity, the dotted line (red) represents "model-free" volatility, measured by the EGARCH(1, 1) volatility. The maturity-specific EGARCH(1, 1) is estimated assuming that the conditional mean of changes in monthly yields is generated by an AR(1) process. The solid line (blue) represents the conditional volatility from the no-arbitrage GARCH model. The dash-dot line (magenta) represents the conditional volatility from the canonical $A_1(3)$ model. The dashed line (green) represents the conditional volatility from the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix.

Figure 3: The Term Structure of Unconditional Volatility.



Notes to Figure: We plot the term structure of average yield volatility implied by different models together with the term structure of "model-free" realized volatility and EGARCH volatility. The solid line with asterisk (black) represents the term structure of realized volatility, and the diamond line (red) represents the EGARCH$(1,1)$ term structure. The cross line (blue) represents the term structure of volatility from the no-arbitrage GARCH model. The square line (magenta) represents the term structure of volatility from the canonical $A_1(3)$ model. The circle line (green) represents the term structure of volatility from the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. The x-axis is yield maturity in months.

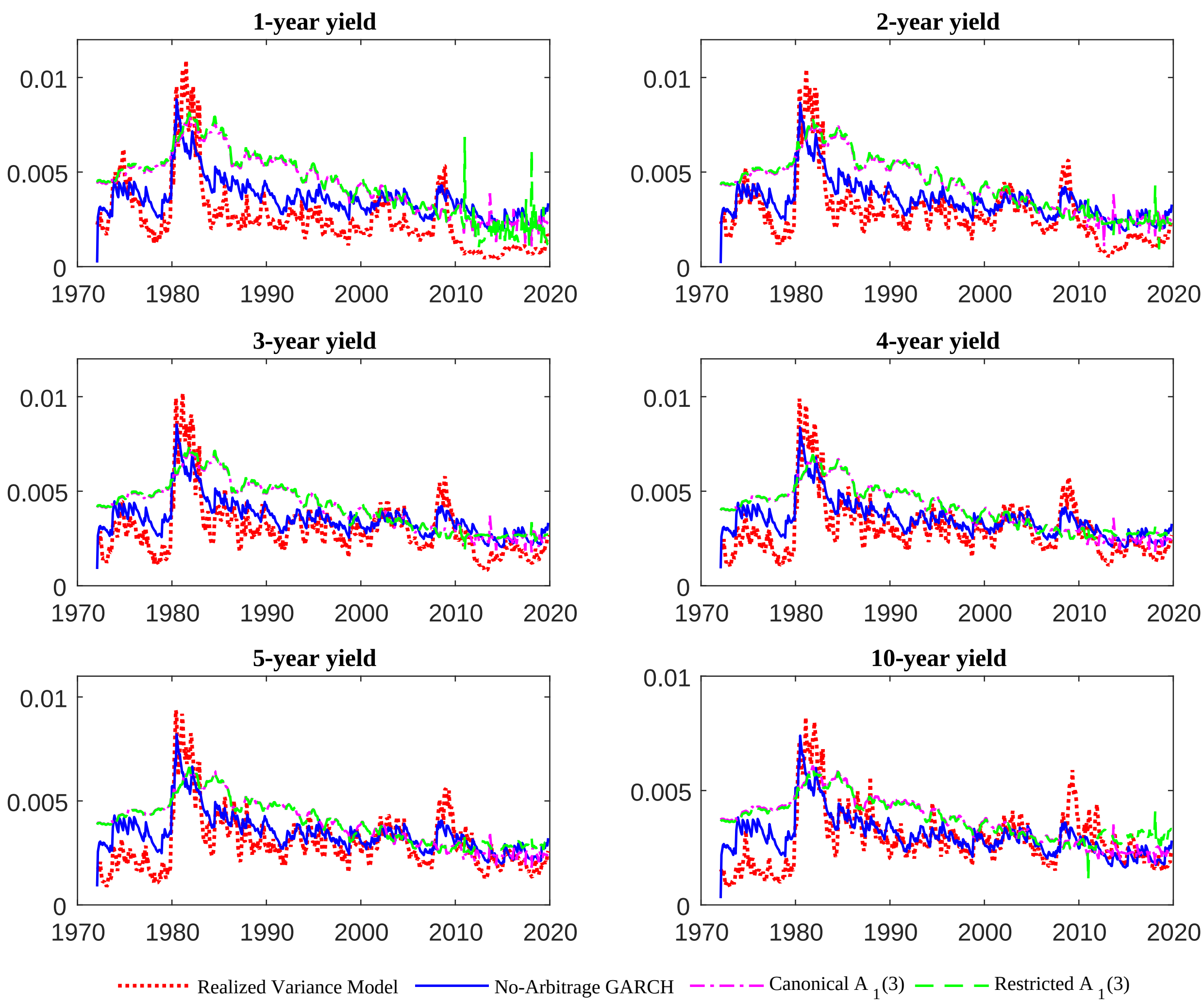
Figure 4: The Campbell-Shiller Expectations Hypothesis Regression.



Notes to Figure: We plot the estimated coefficients from the Campbell-Shiller expectation hypothesis regression using yields implied by different models. The diamond line (red) represents the estimates using the yield data. The cross line (blue) represents the estimates from yields implied by the no-arbitrage GARCH model. The square line (magenta) shows the estimates from yields implied by the canonical $A_1(3)$ model. The circle line (green) is for the estimates from yields implied by the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. The asterisk line (cyan) is for the estimates using yields from the Joslin, Singleton, and Zhu (2011) specification of the Gaussian $A_0(3)$ model. The x-axis is yield maturity in years.

Figure 5: Model-Implied Conditional Volatility:

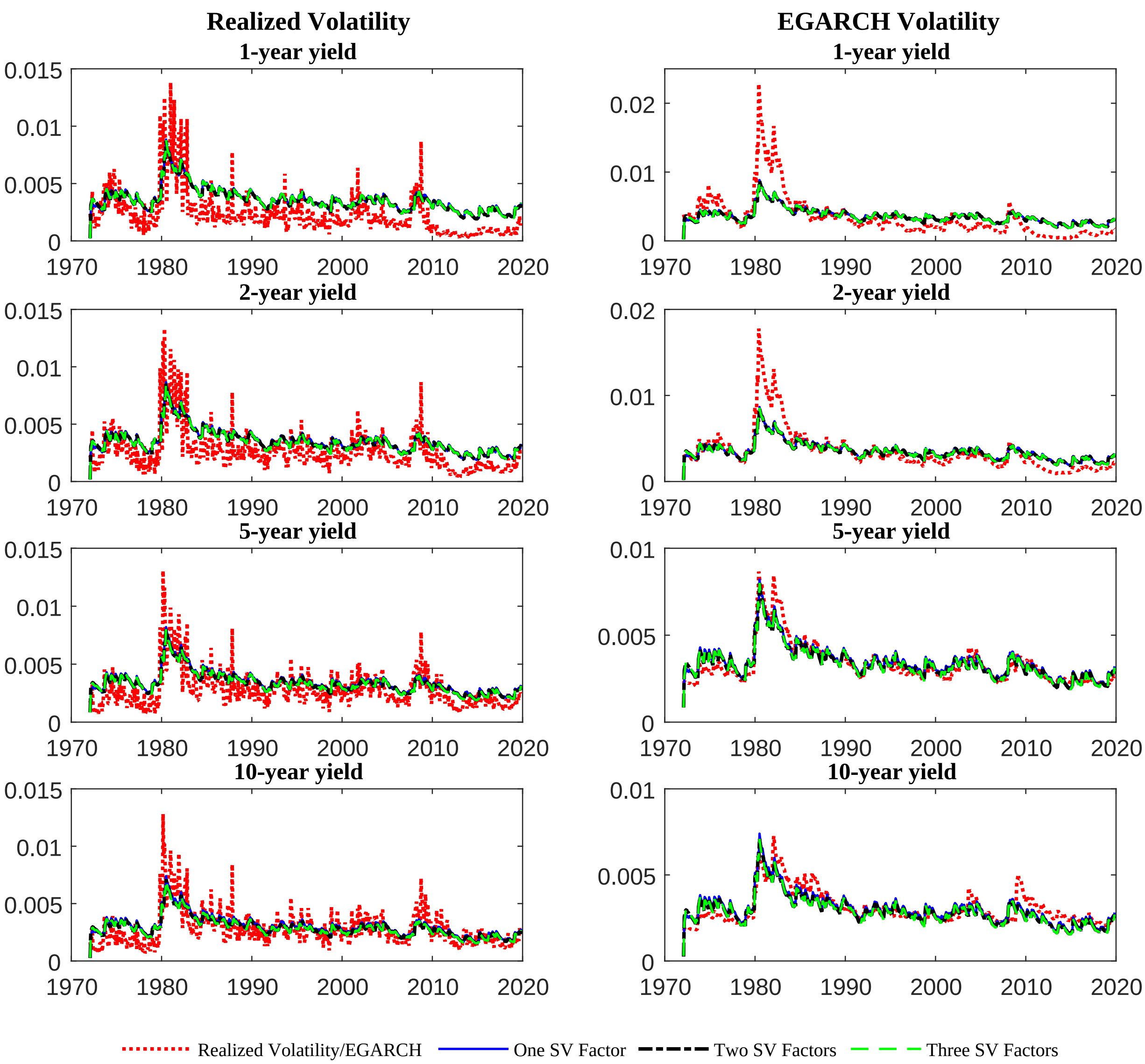
Comparison with the Expected Realized Variance.



Notes to Figure: We plot the one-month conditional volatility of the yield curve implied by different models. For each maturity, the dotted line (red) represents the conditional volatility implied by the realized variance model. We construct a measure of monthly variance using within-month squared changes in daily yields and estimate its conditional expectation assuming that the logarithm of the realized variance follows an ARMA$(1,1)$ process. The solid line (blue) represents the conditional volatility from the no-arbitrage GARCH model. The dash-dot line (magenta) represents the conditional volatility from the canonical $A_1(3)$ model. The dashed line (green) represents the conditional volatility from the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix.

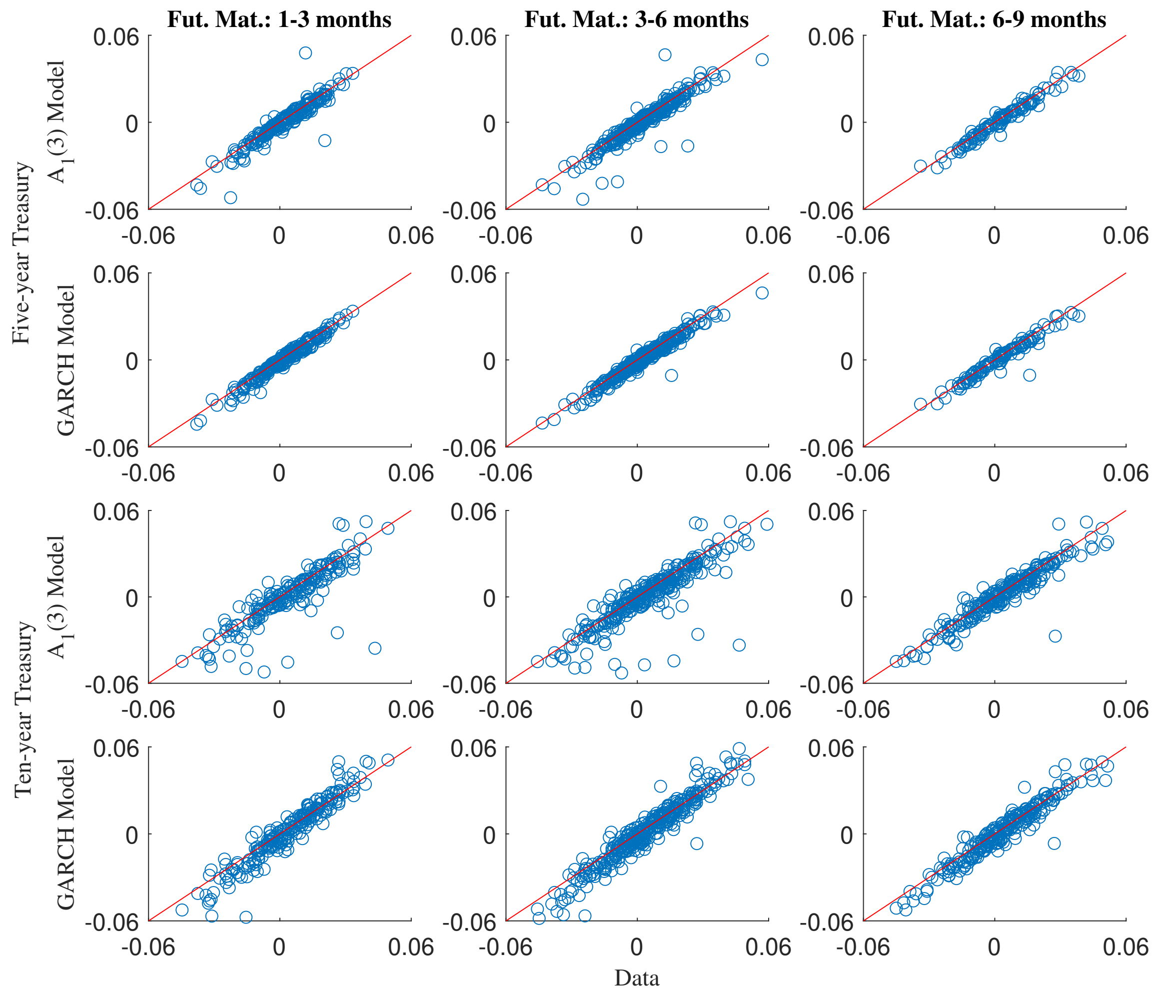
Figure 6: Model-Implied Conditional Volatility in

the No-Arbitrage GARCH Model with Multiple Volatility Factors.



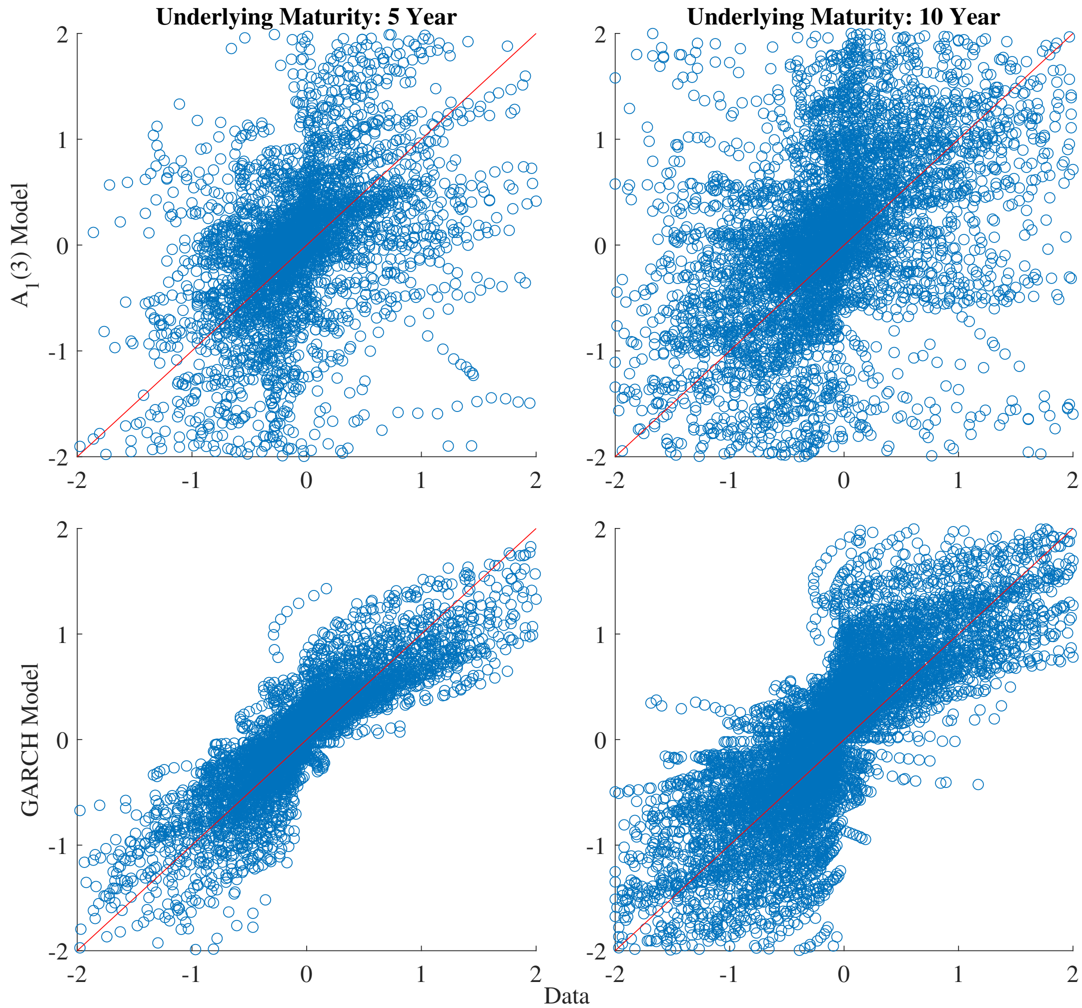
Notes to Figure: We plot the one-month conditional volatility of the yield curve implied by the no-arbitrage GARCH model with different numbers of volatility factors together with the realized volatility (left four panels) and the EGARCH$(1,1)$ volatility (right four panels) respectively. The dotted line (red) in the left panels represents the realized volatility. The dotted line (red) in the right panels represents the EGARCH$(1,1)$ volatility. The solid line (blue) represents the volatility from the no-arbitrage GARCH model with a single volatility factor. The dash-dot line (black) represents the volatility from the model with two volatility factors. The dashed line (green) represents the volatility from the model with three volatility factors.

Figure 7: Hedging Treasury Futures.



Notes to Figure: We scatter plot model-predicted changes in futures prices against actual changes. Model-predicted changes are computed using model-based hedge ratios. We hedge the exposure of each futures contract to the first factor using the underlying Treasury yield. The top two rows present the hedging performance for futures contracts on five-year Treasuries for both models. The bottom two rows present the hedging performance for futures contracts on ten-year Treasuries. The sample period is from 1988 to 2016.

Figure 8: Hedging Treasury Futures Options.



Notes to Figure: We scatter plot model-predicted changes in futures options prices against actual changes. Model-predicted changes are computed using model-based hedge ratios. We hedge the exposure of each futures options contract to the first factor using the underlying Treasury yield. The left two panels present the hedging performance for futures options contracts on five-year Treasuries for both models. The right two panels present the hedging performance for futures options contracts on ten-year Treasuries. The sample period is from 1988 to 2016.

Table 1: Summary Statistics

Panel A: Yields

	Central Moments				Autocorrelation		
	Mean (%)	St.Dev (%)	Skewness	Kurtosis	Lag 1	Lag 12	Lag 30
3 month	4.6508	3.4534	0.6171	3.3040	0.9904	0.8588	0.6391
6 month	4.7758	3.4409	0.5332	3.0506	0.9915	0.8707	0.6632
1 year	5.1349	3.5744	0.4511	2.8107	0.9906	0.8833	0.7019
2 year	5.3581	3.5079	0.3793	2.6595	0.9915	0.8981	0.7500
3 year	5.5477	3.4214	0.3642	2.6022	0.9921	0.9057	0.7775
4 year	5.7130	3.3334	0.3695	2.5853	0.9923	0.9095	0.7938
5 year	5.8587	3.2510	0.3816	2.5896	0.9923	0.9111	0.8032
10 year	6.3640	2.9658	0.4264	2.6891	0.9918	0.9088	0.8110

Panel B: Realized Volatilities

	Central Moments				Autocorrelation		
	Mean (bps)	St.Dev (bps)	Skewness	Kurtosis	Lag 1	Lag 12	Lag 30
1 year	23.4998	19.3417	2.4780	10.8596	0.7292	0.4915	0.1051
2 year	26.4714	17.9137	2.3690	10.6328	0.7045	0.4355	0.0243
3 year	27.7467	16.9367	2.3069	10.4769	0.7047	0.4025	-0.0084
4 year	27.9939	16.0276	2.1956	10.0693	0.7148	0.3943	-0.0107
5 year	27.8329	15.3202	2.0708	9.5533	0.7214	0.4018	0.0029
10 year	26.7380	13.9951	1.6762	7.2509	0.7419	0.4545	0.0609

Panel C: EGARCH Volatilities

	Central Moments				Autocorrelation		
	Mean (bps)	St.Dev (bps)	Skewness	Kurtosis	Lag 1	Lag 12	Lag 30
3 month	28.5701	32.9859	2.7703	11.2903	0.9830	0.7292	0.2958
6 month	28.0909	29.3398	2.7781	12.1228	0.9801	0.6905	0.2998
1 year	34.0188	31.2364	2.6563	11.9225	0.9717	0.6947	0.3562
2 year	35.2671	24.2314	2.6036	11.3548	0.9775	0.6972	0.3708
3 year	35.5186	18.9077	2.4514	10.1627	0.9776	0.6936	0.3635
4 year	34.2632	14.2330	2.2827	8.9541	0.9762	0.6899	0.3539
5 year	33.3196	11.7587	2.0619	7.6196	0.9761	0.7025	0.3603
10 year	30.6445	9.7072	1.5141	4.9637	0.9706	0.6867	0.3639

Notes to Table: We present summary statistics for the data used in estimation. We present the sample mean, standard deviation, skewness, kurtosis, and autocorrelations for each of the yields (Panel A) and yield volatilities (Panels B and C). The yields are continuously compounded monthly zero-coupon bond yields. The monthly realized volatility in Panel B is the square root of monthly realized variance. We construct monthly realized variance using within-month squared changes in daily yields. Panel C reports on the EGARCH$(1,1)$ estimates of changes in monthly yields. The EGARCH$(1,1)$ is estimated assuming that the conditional mean of changes in monthly yields is generated by an AR(1) process. The sample period is from 1971:11 to 2019:10.

Table 2: Summary Statistics: Futures and Options

Panel A: Five-Year Treasury Futures Prices

Maturity	Average	Min	Median	Max
1 - 3 months	110%	94%	109%	125%
3 - 6 months	109%	94%	108%	125%
6 - 9 months	106%	95%	106%	118%

Panel B: Ten-Year Treasury Futures Prices

Maturity	Average	Min	Median	Max
1 - 3 months	112%	92%	111%	135%
3 - 6 months	112%	91%	111%	134%
6 - 9 months	108%	91%	108%	133%

Panel C: Implied Volatility, Options on Five-Year Treasuries

	Moneyness (K/F)					
Maturity	0.95-0.975	0.975-1.00	1.00-1.025	1.025-1.05	1.05-1.10	N
1 - 3 months	4.95%	3.51%	3.27%	4.31%	5.71%	1440
3 - 6 months	4.66%	3.67%	3.56%	4.55%	5.59%	3859
6 - 9 months	4.08%	3.84%	3.78%	3.67%		102
N	408	2542	2170	265	16	

Panel D: Implied Volatility, Options on Ten-Year Treasuries

	Moneyness (K/F)					
Maturity	0.95-0.975	0.975-1.00	1.00-1.025	1.025-1.05	1.05-1.10	N
1 - 3 months	6.26%	5.27%	5.00%	5.45%	6.70%	1856
3 - 6 months	6.22%	5.66%	5.55%	5.88%	6.81%	6225
6 - 9 months	5.97%	5.77%	5.66%	5.67%	6.09%	1070
N	1683	2954	2717	1334	463	

Notes to Table: We present summary statistics for Treasury futures prices and option-implied volatilities from options on Treasury futures. The sample period is from 1988 to 2016. Panels A and B report the average, median, minimum and maximum prices for futures on five- and ten-year Treasuries. Panels C and D report the average implied volatilities from futures options with underlying five- and ten-year Treasuries, for various moneyness and maturity groups. N indicates the number of option contracts in each group.

Table 3: Parameter Estimates

Panel A: The No-Arbitrage GARCH Model

K_0^P		K_1^P			K_0^Q		K_1^Q	
0.0025	0.9979				0.0017	0.9966		
-0.0039		0.9542			-0.0072		0.9466	
0.0007			0.9449		0.0161			0.7212
ρ_0			ρ_1				Log Likelihood	
-0.0001		0.0003	-0.0338	0.0372			19173.32	
$\omega \times 1e2$	α	β		$\sigma_2 \times 1e4$	$\sigma_3 \times 1e4$			
0.0001	0.0794	0.9064		0.0021	0.0153			

Panel B: The Canonical $A_1(3)$ Model

K_0^P		K_1^P			K_0^Q		K_1^Q	
0.0222	0.9958				0.0222	1.0011		
-0.4750	0.0892	0.9697	-0.2707		0.0298	0.1246	0.9658	-0.7689
-0.0850	0.0160	0.0016	0.8095		-0.1022	0.0140	0.0021	0.8867
ρ_0			ρ_1				Log Likelihood	
0.0172		0.0015	0.0015	0.0056			15640.90	
	a				b			
0.0000	1.0000	1.0000		1.0000	10.3716	0.1575		

Panel C: The Restricted $A_1(3)$ Model

K_0^P		K_1^P			K_0^Q		K_1^Q	
0.0226	0.9958				0.0226	1.0010		
0.0000		0.9697			0.0000		0.9657	
0.0000			0.8118		0.0000			0.8890
ρ_0			ρ_1				Log Likelihood	
0.0172		0.0015	0.0015	0.0057			15454.31	
	a				b			
0.0000	1.0000	1.0000		1.0000	10.3615	0.1578		

Notes to Table: We present the estimated parameters and log likelihoods for the no-arbitrage GARCH model (Panel A), the canonical $A_1(3)$ model (Panel B), and the restricted $A_1(3)$ model (Panel C). In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. The estimates are based on monthly zero-coupon bond yields from 1971:11 to 2019:10. The state variables in the canonical and restricted $A_1(3)$ models follow a first order VAR process when sampled monthly.

Table 4: Unconditional Correlations for Conditional Volatility Implied by Different Models

Panel A: Unconditional Correlations with Realized Volatility

	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
No-Arbitrage GARCH	0.69	0.69	0.68	0.67	0.65	0.62	0.67
Canonical $A_1(3)$	0.55	0.50	0.47	0.44	0.41	0.36	0.46
Restricted $A_1(3)$	0.54	0.50	0.46	0.43	0.40	0.35	0.45

Panel B: Unconditional Correlations with EGARCH$(1,1)$ Volatility

	3 month	6 month	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
No-Arbitrage GARCH	0.85	0.86	0.91	0.94	0.95	0.94	0.92	0.80	0.90
Canonical $A_1(3)$	0.62	0.64	0.69	0.71	0.71	0.70	0.69	0.58	0.67
Restricted $A_1(3)$	0.59	0.61	0.67	0.71	0.71	0.70	0.69	0.59	0.66

Notes to Table: Panel A presents the unconditional correlations between the realized volatility and the one-month conditional volatility implied by the no-arbitrage GARCH model, the canonical $A_1(3)$ model, and the restricted $A_1(3)$ model. We construct monthly realized variance using within-month squared changes in daily yields. Panel B presents the unconditional correlations between the EGARCH$(1,1)$ volatility and the one-month conditional volatility implied by the three models. The EGARCH$(1,1)$ is estimated assuming that the conditional mean of changes in monthly yields is generated by an AR(1) process. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. The last column shows the average correlation across all maturities.

Table 5: Conditional Volatility Fit

Panel A: RMSEs of Volatility Based on Realized Volatility

	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
No-Arbitrage GARCH	19.03	15.84	14.89	14.06	13.39	11.62	14.80
Canonical $A_1(3)$	26.51	23.35	21.13	19.59	18.65	17.34	21.09
Restricted $A_1(3)$	27.13	23.79	21.56	19.95	18.92	17.38	21.46
Improvement on Can. $A_1(3)$	28.21%	32.17%	29.56%	28.20%	28.19%	33.01%	29.89%
Improvement on Rest. $A_1(3)$	29.86%	33.43%	30.96%	29.50%	29.24%	33.16%	31.02%

Panel B: RMSEs of Volatility Based on EGARCH(1, 1) Volatility

	3 month	6 month	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
No-Arbitrage GARCH	25.96	22.65	22.09	14.73	9.86	6.10	4.86	5.90	14.02
Canonical $A_1(3)$	32.43	28.60	25.40	18.98	14.68	12.01	10.87	11.09	19.26
Restricted $A_1(3)$	32.89	29.09	25.46	19.07	14.90	12.24	10.98	10.88	19.44
Improvement on Can. $A_1(3)$	19.97%	20.79%	13.04%	22.36%	32.82%	49.21%	55.25%	46.78%	32.53%
Improvement on Rest. $A_1(3)$	21.09%	22.14%	13.23%	22.73%	33.84%	50.16%	55.68%	45.75%	33.08%

Notes to Table: We present the volatility fit for the no-arbitrage GARCH model, the canonical $A_1(3)$ model, and the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. In Panel A, the RMSEs are computed by using realized volatility as a measure of model-free volatility. We construct monthly realized variance using within-month squared changes in daily yields. In Panel B, the RMSEs are computed by using an EGARCH(1, 1) model as a measure of model-free volatility. The EGARCH(1, 1) is estimated assuming that the conditional mean of changes in monthly yields is generated by an AR(1) process. RMSEs are reported in basis points. We also present the percentage RMSE improvement of the no-arbitrage GARCH model over the canonical and restricted $A_1(3)$ models. The last column shows the averages across all maturities.

Table 6: Yield Fit

Panel A: Stochastic Volatility Models

	3 month	6 month	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
No-Arbitrage GARCH	24.22	25.06	4.27	10.44	16.33	18.74	20.10	32.69	18.98
Canonical $A_1(3)$	12.84	17.59	27.12	16.25	9.11	13.36	19.70	36.97	19.12
Restricted $A_1(3)$	18.84	15.09	24.27	15.47	11.53	15.45	21.05	37.92	19.96

Panel B: Gaussian Models

	3 month	6 month	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
JSZ	7.95	10.36	12.88	4.65	5.08	5.67	5.95	7.73	7.53
JSZ Restricted_1	25.04	23.13	5.72	5.36	1.31	3.88	5.02	7.71	9.65
JSZ Restricted_2	7.81	14.18	15.71	7.89	6.20	6.75	7.08	11.81	9.68
JSZ Restricted_3	9.33	11.65	14.75	6.62	5.21	6.37	7.42	8.08	8.68

Notes to Table: We present RMSEs based on the yield fit for the time-varying volatility models (Panel A) and the Gaussian models (Panel B). Panel A reports results for the no-arbitrage GARCH model, the canonical $A_1(3)$ model, and the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. In Panel B, we consider the canonical representation of Joslin, Singleton, and Zhu (2011, referred to as JSZ). We present the results for the maximum flexible specification of the JSZ model and also for three restricted JSZ canonical forms. In the first restricted version, we set the variance-covariance matrix to be a diagonal matrix. In the second restricted version, we also set the $(1, 2)$ and $(1, 3)$ entries of the feedback matrix to be zeros. In the third restricted version, we set the feedback matrix to be a diagonal matrix. The last column in the table shows the average across all maturities. RMSEs are reported in basis points.

Table 7: Conditional Volatility Fit Based on Expected Realized Variance

Panel A: Unconditional Correlations

	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
No-Arbitrage GARCH	0.86	0.87	0.87	0.85	0.83	0.74	0.84
Canonical $A_1(3)$	0.64	0.59	0.54	0.50	0.47	0.40	0.52
Restricted $A_1(3)$	0.63	0.59	0.53	0.49	0.45	0.40	0.51

Panel B: RMSEs of Volatility

	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
No-Arbitrage GARCH	14.73	10.97	10.05	9.51	9.30	8.96	10.59
Canonical $A_1(3)$	23.17	20.01	17.98	16.79	16.15	15.40	18.25
Restricted $A_1(3)$	23.79	20.45	18.40	17.14	16.40	15.33	18.59
Improvement on Can. $A_1(3)$	36.44%	45.18%	44.11%	43.36%	42.39%	41.78%	42.21%
Improvement on Rest. $A_1(3)$	38.10%	46.37%	45.37%	44.51%	43.29%	41.51%	43.19%

Notes to Table: We present the unconditional correlations and the volatility fit for various models when the conditional volatility implied by the realized variance model is used as a measure of model-free yield volatility. We construct the measure of monthly variance using within-month squared changes in daily yields. The logarithm of the realized variance follows an ARMA$(1,1)$ process. Panel A presents the unconditional correlations between the one-month conditional volatility from the realized variance model and the one-month conditional volatility implied by the no-arbitrage GARCH model, the canonical $A_1(3)$ model, and the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. Panel B reports the volatility RMSEs in basis points for the three models. We also present the percentage improvement in RMSE for the no-arbitrage GARCH model over the canonical and restricted $A_1(3)$ models. The last column shows the averages across all maturities.

Table 8: Parameter Estimates for the No-Arbitrage GARCH Model with Multiple Volatility Factors

Panel A: Two Volatility Factors

K_0^P	K_1^P			K_0^Q	K_1^Q		
0.0022	0.9981			0.0011	0.9960		
-0.0048		0.9528		-0.0066		0.9551	
0.0004			0.9428	0.0158			0.7357

ρ_0	ρ_1			Log Likelihood
-0.0001	0.0003	-0.0340	0.0377	19175.59

$\omega \times 1e2$	α	β	$\sigma_3 \times 1e4$
0.0010	0.0720	0.9056	0.0036
0.0000	0.0000	0.8692	

Panel B: Three Volatility Factors

K_0^P	K_1^P			K_0^Q	K_1^Q		
0.0022	0.9985			0.0008	0.9959		
-0.0046		0.9584		-0.0071		0.9544	
0.0007			0.9376	0.0153			0.7313

ρ_0	ρ_1			Log Likelihood
-0.0001	0.0003	-0.0344	0.0349	19176.62

$\omega \times 1e2$	α	β
0.0008	0.0711	0.9021
0.0001	0.0000	0.7314
0.0001	0.0000	0.8431

Panel C: Likelihood Ratio Test

	Three vs One Volatility Factor	Two vs One Volatility Factor
P Value	0.1585	0.1041
Statistics	6.6013	4.5243

Notes to Table: We present the estimated parameters and log likelihoods for the no-arbitrage GARCH model with two volatility factors (Panel A) and the no-arbitrage GARCH model with three volatility factors (Panel B). The estimates are based on monthly zero coupon bond yields from 1971:11 to 2019:10. Panel C presents the p-value and likelihood ratio test statistics for the no-arbitrage GARCH model, testing between the models with three and one volatility factors, and between the models with two and one volatility factors.

Table 9: Results for the No-Arbitrage GARCH Model with Three Principal Components

	3 month	6 month	1 year	2 year	3 year	4 year	5 year	10 year	Avg.
Panel A: Unconditional Correlations with Realized Volatility									
No-Arbitrage GARCH with 3PCs			0.73	0.72	0.71	0.68	0.66	0.59	0.68
Canonical $A_1(3)$			0.55	0.50	0.47	0.44	0.41	0.36	0.46
Restricted $A_1(3)$			0.54	0.50	0.46	0.43	0.40	0.35	0.45
Panel B: RMSEs of Volatility Based on Realized Volatility									
No-Arbitrage GARCH with 3PCs			22.27	18.72	17.06	16.33	15.59	14.46	17.41
Canonical $A_1(3)$			26.51	23.35	21.13	19.59	18.65	17.34	21.09
Restricted $A_1(3)$			27.13	23.79	21.56	19.95	18.92	17.38	21.46
Improvement on Can. $A_1(3)$			15.98%	19.83%	19.27%	16.64%	16.39%	16.61%	17.45%
Improvement on Rest. $A_1(3)$			17.91%	21.32%	20.87%	18.14%	17.61%	16.79%	18.77%
Panel C: Unconditional Correlations with EGARCH(1, 1) Volatility									
No-Arbitrage GARCH with 3PCs	0.84	0.86	0.92	0.93	0.92	0.90	0.87	0.74	0.87
Canonical $A_1(3)$	0.62	0.64	0.69	0.71	0.71	0.70	0.69	0.58	0.67
Restricted $A_1(3)$	0.59	0.61	0.67	0.71	0.71	0.70	0.69	0.59	0.66
Panel D: RMSEs of Volatility Based on EGARCH(1,1) Volatility									
No-Arbitrage GARCH with 3PCs	26.18	23.13	20.37	13.49	9.37	7.88	7.87	8.97	14.66
Canonical $A_1(3)$	32.43	28.60	25.40	18.98	14.68	12.01	10.87	11.09	19.26
Restricted $A_1(3)$	32.89	29.09	25.46	19.07	14.90	12.24	10.98	10.88	19.44
Improvement on Can. $A_1(3)$	19.28%	19.11%	19.79%	28.93%	36.13%	34.42%	27.57%	19.08%	25.54%
Improvement on Rest. $A_1(3)$	20.41%	20.48%	19.96%	29.26%	37.09%	35.65%	28.27%	17.51%	26.08%

Notes to Table: We present the unconditional correlations and the volatility fit for the no-arbitrage GARCH model with three principal components, the canonical $A_1(3)$ model, and the restricted $A_1(3)$ model. In the restricted $A_1(3)$ model, we set the feedback matrix to be a diagonal matrix. Panel A presents the unconditional correlations between the realized volatility and the one-month conditional volatility implied by the three models. We construct monthly realized variance using within-month squared changes in daily yields. Panel C present the unconditional correlations between the EGARCH(1, 1) volatility and the one-month conditional volatility implied by the three models. The EGARCH(1, 1) is estimated assuming that the conditional mean of changes in monthly yields is generated by an AR(1) process. Panels B and D present the volatility fit for the three models. In Panel B, the RMSEs are computed by using realized volatility as a measure of model-free volatility. In Panel D, the RMSEs are computed by using an EGARCH(1, 1) model as a measure of model-free volatility. We also present the percentage RMSE improvement of the no-arbitrage GARCH model with three principal components over the canonical and restricted $A_1(3)$ models. RMSEs are reported in basis points. The last column shows the averages across all maturities.

Table 10: Hedging Treasury Futures Options

Panel A: The Canonical $A_1(3)$ Model

	Full Sample			Put Options			Call Options		
	Intercept	Pred. Change	R^2	Intercept	Pred. Change	R^2	Intercept	Pred. Change	R^2
5 year	-0.03	0.04	0.04	-0.16	0.04	0.04	0.10	0.04	0.04
	(0.01)	(0.003)		(0.01)	(0.004)		(0.01)	(0.004)	
10 year	-0.06	-0.002	0.001	-0.25	-0.003	0.01	0.11	0.002	0.001
	(0.01	(0.001)		(0.02)	(0.001)		(0.02)	(0.001)	

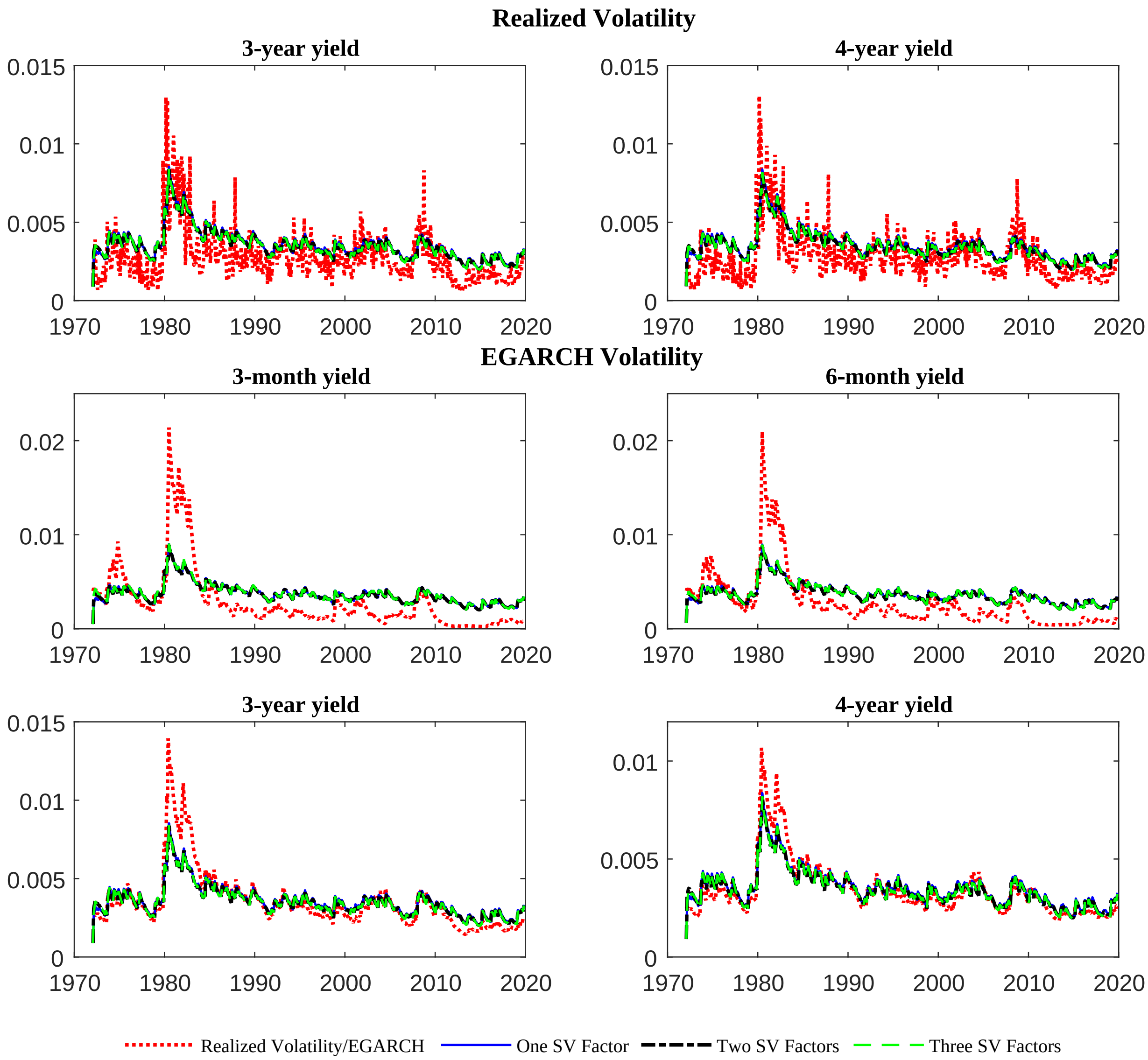
Panel B: The No-Arbitrage GARCH Model

	Full Sample			Put Options			Call Options		
	Intercept	Pred. Change	R^2	Intercept	Pred. Change	R^2	Intercept	Pred. Change	R^2
5 year	-0.07	0.97	0.78	-0.16	0.91	0.75	0.02	1.00	0.82
	(0.004)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)	
10 year	-0.11	0.71	0.70	-0.21	0.91	0.69	-0.001	0.64	0.73
	(0.01)	(0.00)		(0.01)	(0.01)		(0.01)	(0.01)	

Notes to Table: We regress the actual change in option prices over a one-month period on the model-predicted changes. The predicted changes are computed using model-based hedge ratios and the change in yields. We report the intercept and regression slope. Standard errors are in parentheses.

Figure IA1: Model-Implied Conditional Volatility in

the No-Arbitrage GARCH Model with Multiple Volatility Factors.



Notes to Figure: We plot the one-month conditional volatility of the yield curve implied by the no-arbitrage GARCH model with different numbers of volatility factors together with the realized volatility (top two panels) and the EGARCH(1, 1) volatility (bottom four panels) respectively. The dotted line (red) in the top two panels represents the realized volatility. We construct monthly realized variance using within-month squared changes in daily yields. The dotted line (red) in the bottom four panels represents the EGARCH(1, 1) volatility. The EGARCH(1, 1) is estimated assuming that the conditional mean of changes in monthly yields is generated by an AR(1) process. The solid line (blue) represents the volatility from the no-arbitrage GARCH model with a single volatility factor. The dash-dot line (black) represents the volatility from the model with two volatility factors. The dashed line (green) represents the volatility from the model with three volatility factors.