HONG KONG INSTITUTE FOR MONETARY RESEARCH

TERM-STRUCTURE MODELLING AT THE ZERO LOWER BOUND: IMPLICATIONS FOR ESTIMATING THE TERM PREMIUM

Tsz-Kin Chung, Cho-Hoi Hui and Ka-Fai Li

HKIMR Working Paper No.21/2015

October 2015

Hong Kong Institute for Monetary Research
香港金融研究中心
(a company incorporated with limited liability)

Term-Structure Modelling at the Zero Lower Bound: Implications for Estimating the Term Premium

Tsz-Kin Chung*

Tokyo Metropolitan University

and

Cho-Hoi Hui**

Hong Kong Monetary Authority

and

Ka-Fai Li**

Hong Kong Monetary Authority

October 2015

Abstract

Although the affine Gaussian term-structure model has been a workhorse model in term-structure modelling, it remains doubtful whether it is an appropriate model in a low interest rate environment because of its inability to preclude negative interest rates. This paper uses an alternative quadratic Gaussian-term structure model which is well known to be as tractable as the affine model and yet is suitable for interest rates close to zero. Compared with the quadratic model under the zero lower bound, we illustrate how the estimated term premium can be biased upward under the affine model. In contrast to the affine model, our numerical study shows that the quadratic model renders the estimated term premium less likely to be affected by the persistence of the data near the zero lower bound.

Keywords: Term Premium, Zero Lower Bound, Quadratic Gaussian Term-Structure Model, Persistence

JEL Classifications: C32, E43, E44, E52, G12

---

* Graduate School of Social Sciences, Tokyo Metropolitan University

** Research Department, Hong Kong Monetary Authority

Email addresses: btkchung@gmail.com (Tsz-Kin Chung), chhui@hkma.gov.hk (Cho-Hoi Hui), kfli@hkma.gov.hk (Ka-Fai Li)

1. Introduction

Shortly after the onset of the 2008 global financial crisis, one type of unconventional monetary policy that the US Federal Open Market Committee employed was quantitative easing (QE).[1] As argued by Bernanke (2013), QE provides stimulus to the economy by lowering the risk premium component of the long-term interest rate, which is commonly referred as the term premium. The term premium reflects the additional compensation that an investor requires for investing in a long-term bond as compared to a strategy of rolling over and reinvesting the short-term bonds.

To obtain a timeliness measure of the term premium, economists typically prefer model-based estimates over infrequently sampled survey-based measures.[2] One commonly used model-based measure is the term premium estimated from an affine Gaussian term-structure model (AGTSM).[3] For example, the commonly used term premium estimates by Kim and Wright (2005) and Adrian et al. (2013) are generated from an AGTSM. Due to the current low interest rate environment, an AGTSM's inability to preclude negative interest rates makes it ineffective and biased. Longstaff (1989), Leippold and Wu (2002) and Ahn et al. (2002) have proposed a quadratic Gaussian term-structure model (QGTSM), and Kim and Singleton (2012) have tested this model in Japan. In this paper, we compare the efficacy of an AGTSM and QGTSM to see which one can produce realistic term premium estimates.

We show that the difficulties of an AGTSM in handling a zero lower bound of interest rates manifest into an upward biased term premium both theoretically and empirically. Specifically, when interest rates are near zero, the expectation of future interest rates is downward biased in an AGTSM because of a non-trivial probability that interest rates could be negative. In the decomposition of long-term interest rates, a downward bias in the expected future interest rates is equivalent to an upward bias in the term premium. Since risk premia are countercyclical in nature - high during recessions and low during expansions - an upward biased term premium may overstate the severity of a recession. On the contrary, the quadratic functional form assumed in QGTSM can avoid negative interest rates and the downward bias in expected future interest rates.[4] We illustrate that the expectation of future interest rates under an AGTSM is always lower than under a QGTSM when current interest rates are persistently close to zero.

Term-structure models, such as the AGTSM or QGTSM, can be re-cast as state-space models. In a

---

1 Other unconventional monetary policies include forward guidance about the expected path of future policy rates and operation twist. For detailed descriptions on how they could be effective when conventional interest rate cuts are not feasible, see Bernanke et al. (2004).

2 Standard surveys such as the Blue Chip Survey of forecaster asks respondents for their long-term forecasts of the short-term interest rates only twice per year. Moreover, as Swanson (2007) argues, rounding errors may be substantial since each respondent only reports very rough estimates of future short-term rates.

3 See Piazzesi (2010) for a survey of AGTSM.

4 In a QGTSM, the bond yields are quadratic function of the state variables, which is in contrast to the linear relationship assumed in an AGTSM.

state-space formulation, there are two equations that completely describe the dynamics of the term-structure model. First, there is a measurement equation that links bond yields and state variables through the bond pricing formula. Second, there is a state equation that specifies the dynamics of the state variables as a vector autoregression (VAR). Recently, Bauer et al. (2012, 2014) further argue that term premium estimates are sensitive to estimation method and may be biased when interest rates are persistent. It is well known that the standard maximum likelihood estimation of VAR is prone to severe small sample biases when the data generating process (DGP) is close to a unit-root process. Bauer et al. (2012, 2014) illustrate that the bias incurred in the estimation of a VAR can undermine the reliability of term premium estimates. Specifically, the mean reversion parameter of the term-structure model would be spuriously upwardly biased. This implies that the forecast of future interest rates may revert too fast to their unconditional mean, leading to too stable risk neutral forward rates and affecting term premium estimates.

The aim of this paper is to compare the term premium estimates generated by a three-factor AGTSM and QGTSM for the US Treasury bond market from 1990 to mid-2014. To facilitate a fair comparison in terms of specification and estimation methods, we assume the two models have the same driving factors and estimate the models using Bayesian Markov Chain Monte Carlo (MCMC) method with the same assumptions on prior distribution and hyper parameters. We employ the conventional level, slope and curvature yield-curve factors, which are known to provide a parsimonious characterisation of US Treasury bond yields. We follow previous studies and introduce measurement errors to the level, slope and curvature proxies.[5] We find that both the AGTSM and QGTSM provide a good fit of bond yields and generate comparable expectations of future short-term interest rates and term premium when interest rates are sufficiently above zero. In contrast, we find that the term premium estimated from AGTSM is biased upward under the zero lower bound.

As the true level of interest rate persistence in the DGP is difficult to infer from a finite data sample, Bayesian methods provide a natural experiment to compare term premium estimates under different assumed levels of interest rate persistence. Different levels of persistence can be obtained by adjusting prior assumptions in the coefficient matrix of the VAR.[6] We find that risk neutral forward rates and term premium in an QGTSM are less likely to be affected by the persistence of interest rates, in contrast to those in an AGTSM. We conjecture that the stickiness feature of interest rates in a QGTSM can counteract the possible bias in the mean reversion parameter when interest rates are persistent.

Our paper is related to the recent studies of term-structure modelling when interest rates are near zero. Anderson and Meldrum (2014) compare the performance of various two-factor AGTSM and QGTSM models for the US Treasury bond market. However, we differ from Anderson and Meldrum (2014) and focus more on both the theoretical and empirical differences of the term premium under a

[5] Hence, the three factors are considered as latent factors in the estimation.

[6] This is because in Bayesian statistics, the posterior estimates of the model parameters are given by their prior distribution and the likelihood of the data.

zero lower bound. It is noteworthy that the shadow rate term–structure model proposed originally by Black (1995) is also a popular candidate for yield curve modelling when interest rates are near zero. Krippner (2013), Bauer and Rudebusch (2013), Ichiue and Ueno (2013), Christensen and Rudebusch (2014) and Wu and Xia (2014) construct shadow rate models and assume that short-term interest rates have an option-like feature. Specifically, it is assumed that actual short-term rates are the maximum of zero and shadow rates, which can be negative, zero, or positive.[7] However, it is well known that estimated shadow rates are sensitive to model specification. Using the Japanese government bond market from 2000 to 2006 as an example, Ueno et al. (2006) estimate a one-factor model and find that the shadow rate can be as low as -15 percent. On the contrary, Kim and Singleton (2012) estimate a two-factor model and find that the shadow rate is only about -1 percent. Similarly, Bauer and Rudebusch (2013) also find significant disparities in estimated shadow rates for the US Treasury bond market under different model specifications. It is also well known that higher order shadow rate models do not have a closed-form bond pricing equation as in the AGTSM or QGTSM. We show how the QGTSM's short rate function can be used to approximate the option-like feature as in shadow rate models. Hence, a QGTSM can retain the spiritual foundation of shadow rate models yet remain analytically tractable.

The paper is organised as follows. Section 2 presents the theoretical framework of the three-factor models that we use. This section also provides a formal definition of the term premium and an illustration that the term premium could be higher for a AGTSM under a zero lower bound. Section 3 presents the data and estimation methods. In Section 4, we compare both short and long-horizon term premia under different models to illustrate that a proper treatment of the zero lower bound is important for evaluating the effectiveness and economic significance of QE. The final section concludes. Technical details, including the derivation of bond pricing formulas and the MCMC algorithm, are covered in the appendices.

2. The Dynamical Term-Structure Model

2.1 General Setup

This paper adopts a discrete time term structure model augmented with Gaussian state factors. The key ingredient of a term-structure model is the linkage between instantaneous short-term interest rates (hereafter short-rate) r_t and an *M*-dimensional state vector $X_t = (x_t^1, \dots, x_t^M)$ as:

$$r_t = \rho(X_t)$$

where ρ is the short-rate function that will be specified later. The state vector X_t is assumed to follow a VAR(1) process:

[7] It is noteworthy that, in the original analysis in Black (1995), the interest rate is assumed to have an option-like feature, but the dynamics of the shadow rate is left unspecified.

$$X_{t+1} = \mu^Q + \Phi^Q X_t + \Sigma\varepsilon_{t+1}$$

with $\varepsilon_t \sim N(0, I_{M\times M})$, $\mu^Q$is a $M \times 1$ vector and Φ^Q is a $M \times M$ matrix. The notation *Q* denotes the risk-neutral probability measure. Following previous studies, we specify the market price of risk as $\lambda_t = \lambda_0 + \lambda_1 X_t$ with λ_0 is a $M \times 1$ vector and λ_1 is a $M \times M$ matrix. Hence, real-world dynamics of the state vector are given by:

$$X_{t+1} = \mu^P + \Phi^P X_t + \Sigma\varepsilon_{t+1}$$

with $\mu^Q = \mu^P - \Sigma\lambda_0$ and $\Phi^Q = \Phi^P - \Sigma\lambda_1$, where *P* denotes the real-world measure. The corresponding pricing kernel is:

$$\xi_{t+1} = \exp\left(-r_t + \tfrac{1}{2}\lambda_t^T\lambda_t - \lambda_t^T\varepsilon_{t+1}\right)\xi_t$$

It can be shown that in the absence of arbitrage opportunities, the time-*t* price of a *n*-period zero-coupon bond can be formulated as:

$$P_t^n = \mathbb{E}_t^P\left[\frac{\xi_{t+1}}{\xi_t}P_t^{n-1}\right] = \mathbb{E}_t^Q\left[\exp\left(-\sum_{i=0}^{n-1} r_{t+i}\right)\right] \quad (1)$$

The *n*-period bond yield is then:

$$y_t^n = -\frac{1}{n}\log P_t^n$$

Given an appropriate specification of the short-rate function $r_t = \rho(X_t)$, it is possible to derive the bond pricing formula in terms of a recursive relationship for both the AGTSM and QGTSM.

2.2 AGTSM

In general, the short-rate function under AGTSM is:

$$r_t = \delta_0 + \delta_1^T X_t \quad (2)$$

where δ_0 is a scalar and δ_1 is a $M \times 1$ vector. Hence, the short-rate is a linear function of the state variables. From Duffie and Kan (1996), we can solve the bond pricing formula in Eq. (1) as:[8]

[8] For details of the derivation, see Appendix 1.

$$P_t^n = \exp(A_n + B_n^T X_t) \quad (3)$$

where A_n is a scalar and B_n is a $M \times 1$ vector. The model-implied bond yield is a linear function to the state variable X_t as:

$$y_t^n = -\frac{1}{n}\log P_t^n = a_n + b_n^T X_t$$

by taking $a_n = -A_n/n$ and $b_n = -B_n/n$ as the factor loadings.

2.3 QGTSM

For the general QGTSM, the short-rate function is:

$$r_t = \alpha_0 + \beta_0^T X_t + X_t^T \Psi_0 X_t \quad (4)$$

where α_0 is a scalar and β_0 is a $M \times 1$ vector and Ψ_0 is a $M \times M$ matrix. This can be seen as a refinement of the linear short-rate function in AGTSM by including the quadratic terms of X_t. The *n*-period zero coupon bond price is:

$$P_t^n = \exp(A_n + B_n^T X_t + X_n^T C_n X_t) \quad (5)$$

where A_n is a scalar and B_n is a $M \times 1$ vector and C_n is a $M \times M$ matrix. The model-implied bond yield can be expressed as:

$$y_t^n = -\frac{1}{n}\log P_t^n = a_n + b_n^T X_t + X_n^T c_n X_t$$

by taking $a_n = -A_n/n$ and $b_n = -B_n/n$ and $c_n = -C_n/n$ as the factor loadings.

2.4 Factors and Short-Rate Function

In this study, we employ the three dimensional state variables $X_t = (x_{1t}, x_{2t}, x_{3t})$ which are proxies of the level, slope and curvature of the yield curve constructed as:

1. Level $(x_{1t}) = y_t^1$, i.e., the three-month Treasury yield;

2. Slope $(x_{2t}) = y_t^{40} - y_t^4$, i.e., the one- to ten-year term spread;

3. Curvature $(x_{3t}) = y_t^{40} - 2y_t^{20} + y_t^4$, i.e., the one- to five- to ten-year butterfly spread).[9]

Diebold et al. (2006), Bikbov and Chernov (2010) and Hamilton and Wu (2012) employ similar proxies in their term-structure models.[10] Furthermore, we assume the short-rate function has dependence on the level factor only, i.e., $r_t = \rho(x_{1t})$. This setup is similar to that used in previous studies such as Bernanke et al. (2004) and Ang et al. (2011).

For the AGTSM, we set $r_t^{AG} = x_{1t}$ such that the short-rate in the model exactly matches 3-month US Treasury yields in the data.[11] For the QGTSM, we consider two specifications that avoid negative interest rates. The first specification sets $r_t^{QG1} = \eta x_{1t}^2$ where η is a parameter used to ensure both the AGTSM and QGTSM produce identical short-rate at a specific level. The 45-degree line and the quadratic curve in Figure 1 represent the short-rate function for the AGTSM and QGTSM respectively. For instance, if we want both models to generate the same short-rate when $x_{1t} = x_u = 2\%$, we can take $\eta = 50$ such that $\eta x_u^2 = 0.02$. We will refer this specification as QGTSM1.

The second specification is set according to the option-like feature inspired by the Black model. The short-rate function in the Black model is $r_t^{Black} = \max(0, x_{1t})$, which is shown graphically in Figure 2 as the green line. It is worth noting that for $x_{1t} > 0$, the short-rate functions of the Black model and the AGTSM (the red line) will coincide. The short-rate function under the second specification for the QGTSM is $r_t^{QG2} = \varphi(x - m)^2$ for $\varphi, m > 0$. φ and m are chosen such that r_t^{Black} and r_t^{QG2} will generate identical short-rates at certain values of the state variable. By construction, r_t^{QG2} is scaled to approximate r_t^{Black}, as shown graphically in Figure 2. For instance, we pick two levels $x_L = -1\%$ and $x_H = 1\%$ respectively. In Figure 2, the conditions of $m = -0.01$ and $\varphi = 25$ ensure that r_t^{QG2} is tangential to r_t^{Black} at the chosen values of -1% and 1%. We will refer the second specification as QGTSM2.

2.5 Term Premium

A common proxy to measure the ex-ante risk premium in bond markets is the term premium. For any maturity pair *m* and *n* with *n* > *m*, the term premium (*TP*) is defined as:

$$TP_t^{m,n} = f_t^{m,n} - \sum_{i=0}^{n-m+1} \mathbb{E}_t^P[r_{t+m+i}] \tag{6}$$

9 The superscript on the bond yield denotes its maturity in quarters.

10 We do not model the factors as the principal components of the yield curve since it is not straightforward to extend the canonical representation of the AGTSM developed by Joslin et al. (2011) to the QGTSM.

11 It is worth noting that this simple short-rate function does not imply bond yields are a function of x_{1t} only.

where $f_t^{m,n}$ are forward interest rates for a *(n-m)*-period bond to be commenced at *m*-periods ahead. Using the terminology in Joslin et al. (2014), $TP_t^{m,n}$ is referred as "in-n-years-for-m-years" term premium. $f_t^{m,n}$ can be computed readily by the bond pricing equation given by $f_t^{m,n} = (n - m)^{-1}(\log P_t^m - \log P_t^n) = (n-m)^{-1}(ny_t^n - mP_t^m)$. Intuitively, the term premium defined in Eq. (6) represents the additional return to holding a *n*-period bond instead of holding a *m*-period bond and then rolling over the proceeds at an uncertain short-rate *(n-m)*-period later.[12] The expectation term $\mathbb{E}_t^P[r_{t+m+i}]$ can be computed by iterating the VAR(1) equation under the P-measure.

For any term-structure model with a reasonable good fit of the bond yield, the discrepancy in forward interest rates should be negligible (i.e., the term $f_t^{m,n}$ in Eq. (6)). Meanwhile, for the expected future short-rate (i.e., $\mathbb{E}_t^P[r_{t+m+i}]$), the expectation is taken at the current time *t*. Hence, there is a high probability that the AGTSM will predict a negative future short-rate when current interest rates are near zero. This in turn leads to overestimation of the term premium.

We can illustrate the potential bias in the AGTSM by calculating the expectation of future interest rates theoretically. Let us rewrite the first equation of the VAR(1) process under Q as:

$$x_{t+1}^1 = \tilde{\mu}_1 + \phi_{11}x_t^1 + \sigma_1\varepsilon_{1,t+1}$$

in which we absorb the interaction terms as $\tilde{\mu}_1 = \mu_1 + \phi_{12}x_t^2 + \phi_{13}x_t^3$. Hence, conditional on the realization of the second and third factors x_t^2 and x_t^3, the level factor x_t^1 can be approximated as a univariate autoregressive model. In a continuous-time limit, if we set $\alpha = \tilde{\mu}_1/\Delta t$, $\beta = (1-\phi_{11})/\Delta t$ and $\sigma = \sigma_1/\sqrt{\Delta t}$ with Δt is a quarter, the state vector $x_t = x_t^1$ follows a mean-reverting process:

$$dx_t = (\alpha - \beta x_t)dt + \sigma dW_t$$

where β is the mean reversion speed to the unconditional mean level α/β and σ is the volatility of the Brownian motion term dW_t. It is well-known that:

$$x_t = x_0 e^{-\beta t} + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right) + \int_0^t \sigma e^{-\beta(t-s)}\, dW_s, \qquad t \geq 0$$

Hence, x_t is a Gaussian random variable distributed as $X \sim N(\bar{\mu}, \bar{\sigma}^2)$ with:

$$\bar{\mu} = \mathbb{E}[x_t] = x_0 e^{-\beta t} + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right) \text{ and } \bar{\sigma}^2 = \frac{\sigma}{2\beta}\left(1 - e^{-2\beta t}\right)$$

To model the persistently low interest rate environment since late 2008, we set $x_0 \to 0$ and $\beta \to 0$

12 Strictly speaking, the term premium defined in Eq. (6) is referred as the forward term premium in previous studies. It can be shown that the term premium can be equivalently defined as the difference between the estimated forward interest rates in the P measure with its counterpart in the Q measure.

(i.e., $\phi_{11} \to 1$). As such, the stochastic process will generate a sequence of x_t which will remain sticky near zero.

For the AGTSM, we have $r_t^{AG} = \delta x_t$ with $\delta > 0$ is a constant.[13] The probability density function is:

$$f(r_t^{AG}) = \phi\left(\frac{r_t^{AG}/\delta - \bar{\mu}}{\bar{\sigma}}\right) = \frac{1}{\sqrt{2\pi\bar{\sigma}^2}} exp\left(-\frac{(r_t^{AG}/\delta - \bar{\mu})^2}{2\bar{\sigma}^2}\right)$$

which does not preclude negative interest rates and the probability of negative interest rates depends on the level α/β and the mean-reverting parameter β. As such, the expected interest rate is:

$$\mathbb{E}[r_t^{AG}] = \delta\mathbb{E}[x_t] = \delta x_0 e^{-\beta t} + \delta\frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right)$$

It is easy to show that the expected interest rate will converge to zero under our assumed condition $x_0 \to 0$ and $\beta \to 0$.

For the QGTSM, we have $r_t^{QG} = \mathrm{a}(x_t + b)^2$ with $\mathrm{a} > 0$ and b are constant. The probability density function is:

$$f\left(r_t^{QG}\right) = \frac{1}{\sqrt{2\pi\bar{\sigma}^2 r_t^{QG}/a}} exp\left(-\frac{r_t^{QG}/a + (\bar{\mu} + b)^2}{2\bar{\sigma}^2}\right)\cosh\left(\frac{\bar{\mu} + b}{\bar{\sigma}^2}\sqrt{r_t^{QG}/a}\right)$$

which is a non-central chi-square distribution (Kim and Singleton, 2012). The model naturally precludes negative interest rates and the distribution is positively skewed with the shape dependent on the ratio $(\bar{\mu} + b)/\bar{\sigma}$. The expected interest rate can be computed as:

$$\mathbb{E}\left[r_t^{QG}\right] = \mathrm{a}(\mathbb{E}[x_t^2] + 2b\mathbb{E}[x_t] + b^2) = \mathrm{a}(\bar{\sigma}^2 + \bar{\mu}^2 - 2b\bar{\mu} + b^2)$$

because $r_t^{QG} = \mathrm{a}(x_t^2 + 2bx_t + b^2)$ and $\mathbb{E}[x_t^2] = \bar{\sigma}^2 + \bar{\mu}^2$. Rearranging terms, we have:

$$\mathbb{E}\left[r_t^{QG}\right] = \mathrm{a}\,\bar{\sigma}^2 + \mathrm{a}(\bar{\mu} + b)^2 > 0$$

which suggests that the expected interest rate is always positive. Hence, if $x_0 \to 0$ and $\beta \to 0$, the expected interest rate converges to a positive level $\mathbb{E}\left[r_t^{QG}\right] = \mathrm{a}\,\bar{\sigma}^2 > 0$. As a result, the expected interest rate under a QGTSM should be higher than under an AGTSM when the interest rate is persistently close to zero. Then, according to Eq. (6), a lower expected interest rate would manifest into a higher term premium when other factors are being held as constant.

13 This is also called the Vasicek model in the literature, after the celebrated work by Vasicek (1977).

3. Data and Estimation Methods

The quarterly (end-of-quarter) dataset of zero-coupon US Treasury yields from January 1990 to June 2014, with yield maturities of one-, two-, four-, eight-, and up to forty-quarter (total 12 maturities) are obtained from Gurkaynak et al. (2007).[14] We re-cast both AGTSM and QGTSM as a non-linear state-space model as follows.

3.1 The Measurement Equation

The measurement equation describes the evolution of the observed bond yields as:

$$y_t^n = a_n + b_n^T X_t + X_n^T c_n X_t + \omega_{n,t} \tag{7}$$

with $n = 1,2, \ldots\ldots N$ and $\omega_{n,t}$ are the yield-measurement errors which are i.i.d. normals.[15] Moreover, we assume that the state variables are observed with state-measurement error $\omega_{X,t}$ as:

$$\hat{X}_t = X_t + \omega_{X,t} \tag{8}$$

where $\hat{X}_t$ is the observed state variables and $\omega_{X,t}$ are i.i.d. normals.

3.2 The State Equation

The state equation is given by the evolution of the latent state vector X_t under the real world measure *P* as:

$$X_{t+1} = \mu^P + \Phi^P X_t + \Sigma\varepsilon_{t+1} \tag{9}$$

Eqs. (7)-(9) together form a non-linear state space model with 15 observables (12 observed bond yields and 3 proxies of the state variables) and 3 latent factors. It is noteworthy that although the short-rate functions Eqs. (2) and (4) are not imposed directly in the estimation, they affect the model-implied bond yield and its likelihood function through the bond pricing equations in Eqs. (3) and (5) for the AGTSM and QGTSM respectively.

We estimate Eqs. (7)- (9) using the Bayesian MCMC method. The Bayes theorem states that the posterior probability is proportional to the product of prior probability and the likelihood given by the data. Hence, the Bayesian method can perturb the persistence in the VAR system to facilitate a model

[14] The authors continue to provide updates of the data; the latest data is available at http://www.federalreserve.gov/econresdata/researchdata/feds200628_1.html.

[15] Obviously for the AGTSM, the term c_n is set to zero.

comparison.[16] To ensure the convergence of the Markov chain, we choose the number of iterations of the Gibbs sampler to be 20,000 and discard the first 10,000 burn-in samples. We conduct statistical inference based on the sample of these remaining draws. The estimation details are provided in Appendix 2.

4. Empirical Results

The estimation coefficients of the VAR system in both P and Q measure are presented in Appendix 3. Here, we focus on the model implications for the pricing errors of bond yields and term premium.

4.1 Fitting of Bond Yields

Table 1 shows the fitting of bond yields and the three latent factors across different models. Pricing errors are measured by the absolute difference between model-implied yields and actual yields. All three models generate a comparable in-sample fit, with the largest pricing error amounting to only around 40 basis points.

Nonetheless, an in-sample forecasting exercise can reveal why the AGTSM is not suitable for yield curve modelling when interest rates are near zero. Given the model parameters and an initial condition, we can simulate the future path of interest rates through iterating the state-space model recursively. For each model, we forecast the last eight observations (i.e., two years) of our data. For the AGTSM, Figure 3 shows that the median forecast for short term bond yields with a maturity up to 1-year are negative, thus severely violating the zero lower bound. Although the median forecast for longer-tenor bond yields are non-negative and largely follow actual bond yields, the confidence interval marked by the fan chart clearly shows a substantial probability of breaching the zero lower bound.

Figures 4 and 5 show that the probability of negative bond yields is always zero for both specifications of the QGTSM, even when the short-term interest rates are extremely close to zero. Indeed, the prediction density shown by the fan chart in each panel is positively skewed because bond yields are bounded below at zero in the QGTSM.

4.2 Implication for Term Premium Estimates

Figure 6 shows the "in-ten-years-for-five-years" term premium ($TP_t^{5,10}$) estimated from the three models. $TP_t^{5,10}$ is commonly used in past studies of risk premia in the bond market (e.g., Wright (2011)). The term premium estimates exhibit a countercyclical pattern, rising notably during

16 Specifically, we use the Minnesota prior in the estimation of the VAR. The Minnesota prior incorporates the prior belief that the variables included in the VAR follow a random walk process, conditional on several hyper parameters. One hyper parameter is commonly referred as the overall tightness parameter in the literature, which determines whether the variables in the VAR would behave more like the random walk process assumed in the prior.

recessions. The three different $TP_t^{5,10}$ move in tandem during most of the sample period. However, it is noted that there are two episodes where the term premium estimated from the AGTSM is tangibly higher than that estimated from the QGTSM1 and QGTSM2. The first episode occurs around 2002 and lasts for about two years. The second episode starts shortly after the global financial crisis in 2008 and has persisted since then. The disparity is due to the downward bias of expected future interest rates in the AGTSM as we argued in Section 3.[17] Indeed, when the FOMC decided to raise the policy rate in late 2004, the ending of a low interest rate environment helped to narrow the disparity because the probability of negative interest rates in AGTSM dwindled significantly.

The shorter-horizon counterpart also depicts a similar pattern to the longer-horizon term premium. Figure 7 shows the "in-two-years-for-one-year" term premium ($TP_t^{1,2}$) estimated from the three models. When compared with Figure 6, the disparity is still pronounced from 2008 onwards, but the disparity in the longer-horizon term premium during 2002-2004 is not present. The reason for this is probably related to the level of interest rates and the forecasting horizon to compute the expectation term in Eq. (6). Specifically, when the federal funds target rate was still one percentage point above the zero lower bound in 2002-2004, the probability of negative interest rates in one-year's time is lower than the corresponding probability in five-year's time since the uncertainty is increasing with the forecasting horizon. On the contrary, since the financial crisis, the AGTSM is more likely to generate negative interest rates as the federal funds target rate has been in the range of 0-25 basis points.

4.3 Persistence of VAR and its Effect on Estimated Term Premium

Bauer et al. (2014) argue that the small sample bias in the maximum likelihood estimation of a AGTSM can make the estimated VAR system less persistent than the true DGP. Equivalently, this means that model implied interest rates revert to their long-term mean faster than usual and the resulting risk-neutral rates are too stable. Due to the bias incurred in standard estimation techniques, Bauer et al. (2014) further challenges the findings by Wright (2011) that term premia in advanced economies have been declining since the early 1990s. In fact, the authors propose several biased-corrected estimation techniques and find that the biased-corrected term premia display a more plausible countercyclical pattern. We re-visit the issue of persistence in both the AGTSM and QGTSM. In our Bayesian estimation, increasing the persistence of the VAR system can be easily achieved by tuning the prior distribution appropriately.

We start with the AGTSM first. We estimate two specifications of an AGTSM such that one version is more persistent than the other. In the more persistent model, the maximum absolute eigenvalue of the estimated VAR system (i.e. $\max(\mathrm{eig}(\Phi^P))$) is 0.996, while the corresponding figure is 0.961 in the less

[17] It is noteworthy that Ichiue and Ueno (2013) find that the term premium estimated in a two-factor AGTSM is downwardly biased when compared with a two-factor Black model, which is contrary to our finding that a AGTSM would generate upward biased term premium. We think that this discrepancy could be due to the specification of the short-rate function. In our AGTSM, the short-rate depends on the level factor only, while Ichiue and Ueno (2013) assume the short-rate is a function of both level and slope factors. Although the level of short-term interest rates remains sticky under the zero lower bound, long-term interest rates still exhibit considerable fluctuations, which would affect the slope factor.

persistent model. Figure 8 decomposes the five-year forwards rate with five years maturity into its risk-neutral component (upper panel) and the associated “in-ten-years-for-five-years” term premium (lower panel). The more stable forward rates in the less persistent model confirm the findings and intuition offered by Bauer et al. (2014) and show a more volatile term premium. Conversely, when the risk neutral forward rate is assumed to be more volatile, achieved by reducing the tendency for it to revert to its unconditional mean, the term premium is more stable as well. Figure 10 illustrates the dilemma of using an AGTSM to generate term premium. On the one hand, as suggested by Bauer et al. (2014), correcting the downward biased in the persistence of the VAR system in an AGTSM can lead to a more realistic pattern of term premium. On the other hand, a more persistent VAR system in an AGTSM may generate negative risk-neutral forward rates and increase the risk of model misspecification under a zero lower bound.

As the model fit and performance in both specifications of the QGTSM are similar, we focus on the first specification, QGTSM1, for the analysis in this section. Figure 9 plots the risk neutral rates and term premium associated with different persistent levels of the VAR system for QGTSM1. Compared with an AGTSM, it appears that the risk neutral rates estimated in a QGTSM are less likely to be affected by the persistence in the VAR system. We conjecture that the stickiness of the QGTSM helps to generate this difference. With a much lower sensitivity to the bias of persistence of the VAR system and its ability to generate non-negative interest rates, a QGTSM offers a more robust framework to analyse term premia.

5. Conclusion

Although the AGTSM has been the workhorse model in term-structure modelling, its inability to preclude negative interest rates undermines its usefulness in a low interest rate environment. Specifically, we show that the term premium estimated from a AGTSM is biased upwards and may lead to an inaccurate assessment of the effectiveness of QE.

A QGTSM shares the analytical tractability of a AGTSM but guarantees non-negative interest rates, and therefore produces plausible term premium estimates with a countercyclical pattern. Compared with an AGTSM, we find that the risk neutral forward rates and term premium estimated in a QGTSM are less likely to be affected by downward bias in the persistence of the VAR system. The model and the estimation methods can be readily extended to macro-finance models with unspanned risk factors, which we leave for future research.

References

Adrian, Tobias, Richard K. Crump and Emanuel Moench (2013), “Pricing the Term Structure with Linear Regressions,” *Journal of Financial Economics*, 110(1): 110-38.

Ang, Andrew, Jean Boivin, Sen Dong and Rudy Loo-Kung (2011), “Monetary Policy Shifts and the Term Structure,” *Review of Economic Studies*, 78(2): 429-57.

Ahn, Dong-hyun., Robert F. Dittmar, A. Ronald Gallant and Henry A. Latane (2002), “Quadratic Term Structure Models: Theory and Evidence,” *Review of Financial Studies*, 15(1): 243-88.

Andreasen, Martin M. and Andrew Meldrum (2014), "Dynamic Term Structure Models: The Best Way to Enforce the Zero Lower Bound," CREATES Working Paper 2014-47.

Bauer, Michael D. and Glenn D. Rudebusch (2013), “Monetary Policy Expectations at the Zero Lower Bound,” Federal Reserve Bank of San Francisco Working Paper Series, 2013-18.

Bauer, Michael D., Glenn D. Rudebusch and Jing Cynthia Wu (2012), "Correcting Estimation Bias in Dynamic Term Structure Models," *Journal of Business & Economic Statistics*, 30(3): 454-67.

Bauer, Michael D., Glenn D. Rudebusch and Jing Cynthia Wu (2014), "Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset: Comment," *American Economic Review*, 104(1): 323-37.

Bernanke, Ben S. (2013), “Communication and Monetary Policy,” FED speech, November 19, 2013

Bernanke, Ben S., Vincent R. Reinhart and Brian P. Sack (2004), “Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment,” *Brookings Papers on Economic Activity*, 2: 1-100.

Bikbov, Ruslan and Mikhail Chernov (2010), "No-Arbitrage Macroeconomic Determinants of the Yield Curve," *Journal of Econometrics*, 159(1): 166–82.

Black, Fischer (1995), “Interest Rates as Options,” *Journal of Finance*, 50(5): 1371-6.

Christensen, Jens H. E. and Glenn D. Rudebusch (2014), “Estimating Shadow-Rate Term Structure Models with Near-Zero Yields,” *Journal of Financial Econometrics*, April 2014.

Del Negro, Marco and Frank Schorfheide (2011), “Bayesian Macroeconometrics,” in John Geweke, Gary Koop and Herman van Dijk, eds., *Oxford Handbook of Bayesian Econometrics*: 293–389, Oxford University Press.

Diebold, Francis X., Glenn D. Rudebusch and S. Boragan Aruoba (2006), "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," *Journal of Econometrics*, 131(1-2): 309–38.

Duffie, Darrell and Rui Kan (1996), “A Yield-Factor Model of Interest Rates,” *Mathematical Finance*, 6(4): 379-406.

Gurkaynak, Refet S., Brian Sack and Jonathan H. Wright (2007), "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54(8): 2291-304.

Hamilton, James D. and Jing Cynthia Wu (2012), "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment," *Journal of Money, Credit and Banking*, 44(1): 3-46.

Ichiue, Hibiki and Yoichi Ueno (2013), “Estimating Term Premia at the Zero Bound: An Analysis of Japanese, US, and UK Yields,” Bank of Japan Working Paper Series 13-E-8.

Joslin, Scott, Marcel Priebsch and Kenneth J. Singleton (2014), "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," *The Journal of Finance*, 69(3): 1197–233.

Joslin, Scott, Kenneth J. Singleton and Haoxiang Zhu (2011), "A New Perspective on Gaussian Dynamic Term Structure Models," *Review Financial Studies*, 24(3): 926-70.

Kim, Don and Jonathan H. Wright (2005), “An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates,” FEDS Working Paper, 2005-33.

Kim, Don and Kenneth J. Singleton (2012), “Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields,” *Journal of Econometrics*, 170(1): 32-49.

Krippner, Leo (2013), "A Tractable Framework for Zero Lower Bound Gaussian Term Structure Models," Reserve Bank of New Zealand Discussion Paper Series DP2013/02, Reserve Bank of New Zealand.

Leippold, Markus and Liuren Wu (2002), “Asset Pricing under the Quadratic Class,” *Journal of Financial and Quantitative Analysis*, 37(2): 271-95.

Longstaff, Francis A. (1989), “A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates,” *Journal of Financial Economics*, 23(2): 195-224.

Piazzesi, Monika (2010), “Affine Term Structure Models,” in Y. Ait-Sahalia and L.P. Hansen, eds., *Handbook of Financial Econometrics*: 691-766, North Holland, Elsevier.

Rogers, John H., Chiara Scotti and Jonathan H. Wright (2014), "Evaluating Asset-Market Effects of Unconventional Monetary Policy: A Multi-Country Review," *Economic Policy*, 29(80): 3-50.

Swanson, Eric T. (2007), “What we Do and Don't know about the Term Premium?” FRBSF Economic Letter, 2007-21.

Ueno, Yoichi, Naohiko Baba and Yuji Sakurai (2006), “The Use of the Black Model of Interest Rates as Options for Monitoring the JGB Market Expectations,” Bank of Japan Working Paper Series 06-E-15.

Wright, Jonathan H. (2011), “Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset,” *American Economic Review*, 101(4): 1514-34.

Wu, Jing Cynthia and Fan Dora Xia (2014), “Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound,” NBER Working Paper No.20117.

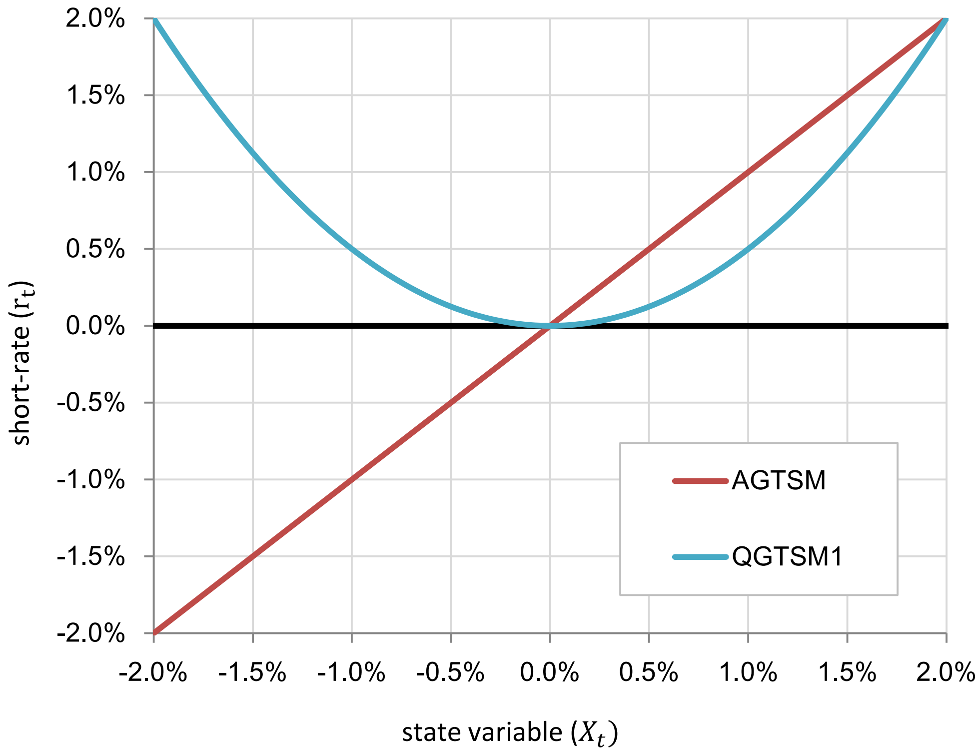
Vasicek, Oldrich (1977), "An Equilibrium Characterisation of the Term Structure," *Journal of Financial Economics*, 5(2): 177–88.

Table 1. Pricing Errors (in Basis Points)

	AGTSM	QGTSM1	QGTSM2
3-month	7.86	19.47	21.72
6-month	17.76	12.31	37.14
1-year	19.75	1.98	20.12
2-year	0.00	5.02	1.89
3-year	0.96	2.93	3.01
4-year	0.43	3.71	1.55
5-year	0.64	2.85	1.94
6-year	0.70	2.72	1.39
7-year	0.48	2.49	0.75
8-year	0.18	2.11	0.18
9-year	0.11	1.60	0.23
10-year	0.31	0.97	0.47
level	0.42	0.22	0.51
slope	0.41	0.62	0.38
curvature	0.28	1.56	0.06

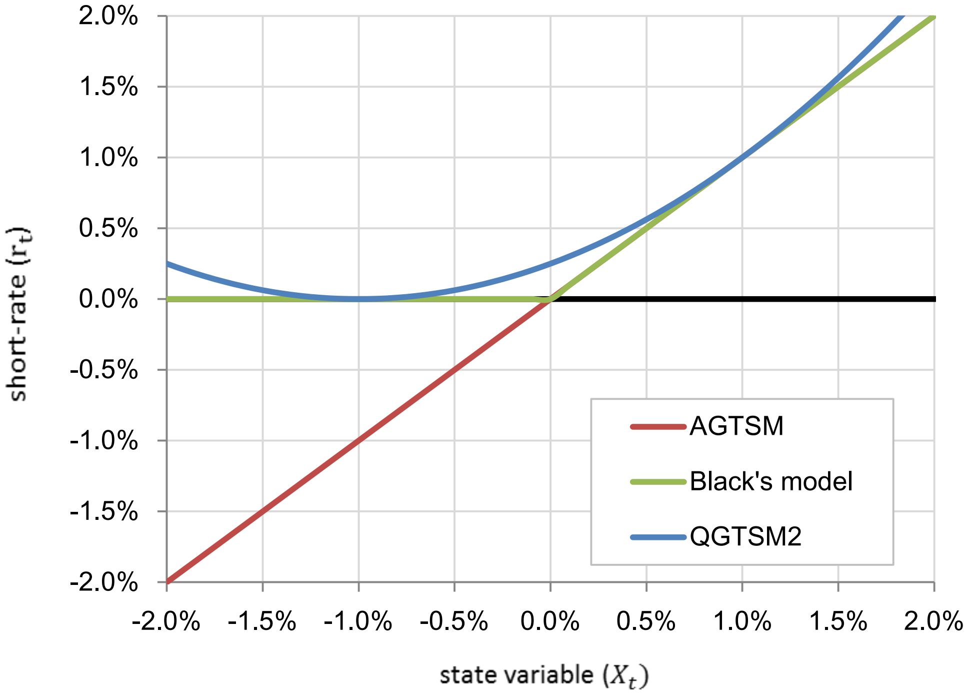
Note: Pricing errors are defined as absolute differences between the actual yields and the model implied yields.

Figure 1. Graphical Illustration of the Short-Rate Functions for AGTSM and QGTSM1



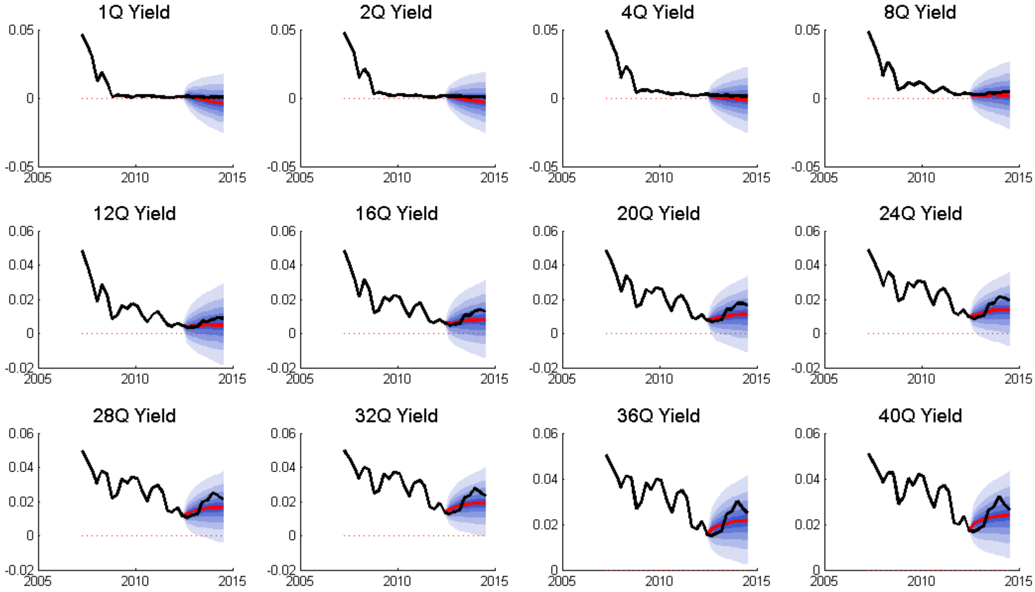
This figure plots the short-rate functions (i.e., $r_t = \rho(x_{1t})$) for the affine model (AGTSM) and the first specification of the quadratic model (QGTSM1). The short-rate functions for the affine and quadratic models are $r_t^{AG} = x_{1t}$ (red line) and $r_t^{QG1} = \eta x_{1t}^2$ (blue line) respectively. We set the parameter η such that both models generate identical short-rates when $x_{1t} = 2\%$.

Figure 2. Graphical Illustration of the Short-Rate Functions for AGTSM and QGTSM2



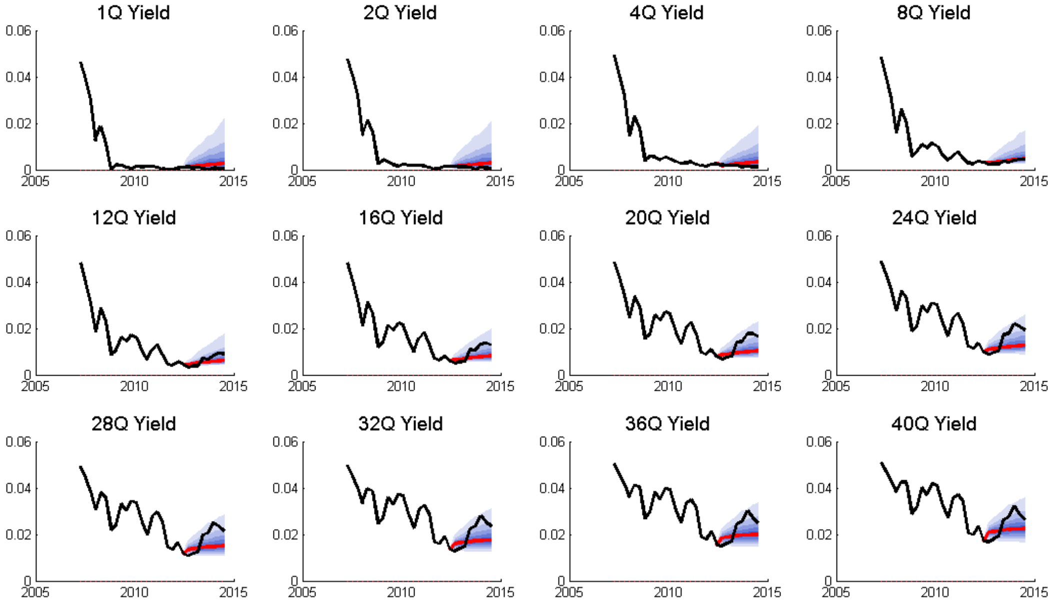
This figure plots the short-rate functions (i.e., $r_t = \rho(x_{1t})$) for the affine model (AGTSM), the Black's model and the second specification of the quadratic model (QGTSM2). The short-rate functions under different models are $r_t^{AG} = x_{1t}$ (red line) $r_t^{Black} = \max(0, x_{1t})$ (green line) and $r_t^{QG2} = \varphi(x_{1t} - m)^2$ respectively. When $x_{1t} > 0$, affine model and black model will generate the same short-rate. We set the parameter φ and m such that the Black model and QGTSM2 generate identical short-rates when $x_{1t} = -1\%$ and 1%.

Figure 3. In-Sample Forecasting for AGTSM



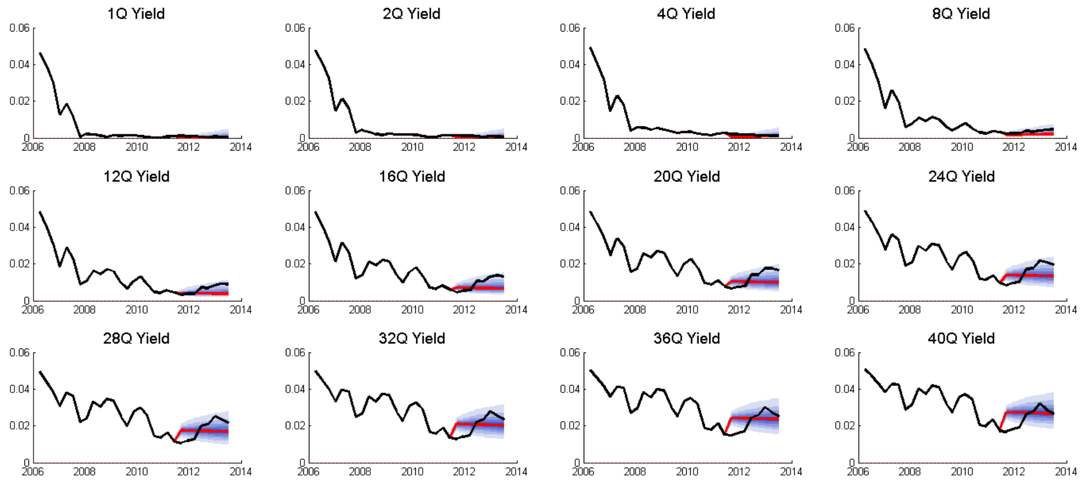
In-sample forecast for the affine model (AGTSM). The forecast horizon is for 8 quarters, which starts from 2012Q2 and ends at 2014Q2. The red line denotes the median forecast (computed as the median of the MCMC draws). The fan chart denotes the confidence interval.

Figure 4. In-Sample Forecasting for QGTSM1



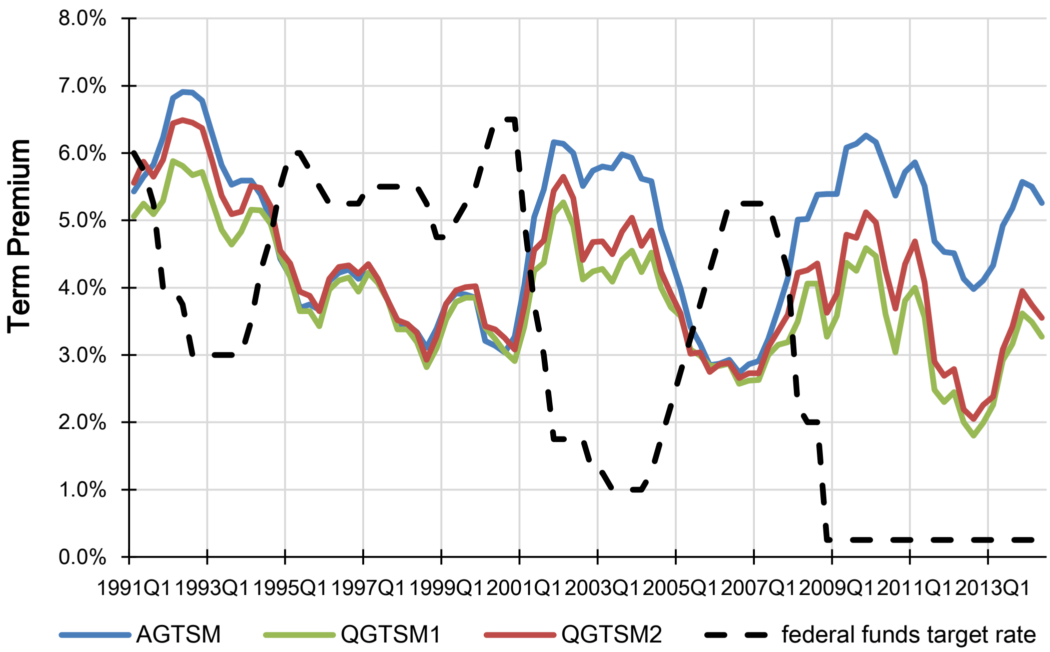
In-sample forecast for the first specification of quadratic model (QGTSM1). The forecast horizon is for 8 quarters, which starts from 2012Q2 and ends at 2014Q2. The red line denotes the median forecast (computed as the median of the MCMC draws). The fan chart denotes the confidence interval.

Figure 5. In-Sample Forecasting for QGTSM2



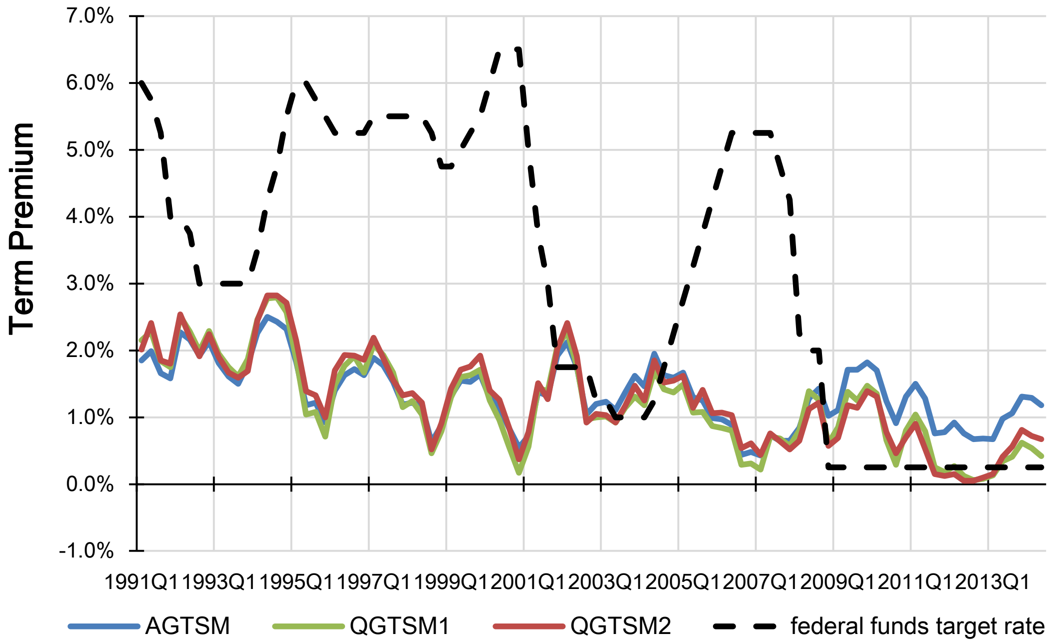
In-sample forecast for the first specification of quadratic model (QGTSM2). The forecast horizon is for 8 quarters, which starts from 2012Q2 and ends at 2014Q2. The red line denotes the median forecast (computed as the median of the MCMC draws). The fan chart denotes the confidence interval.

Figure 6. In-Ten-Years-for-Five-Years Term Premium $\left(TP_t^{5,10}\right)$ of Different Models



This figure plots the estimated "in-ten-years-for-five-years" term premium of different models and the federal funds target rate.

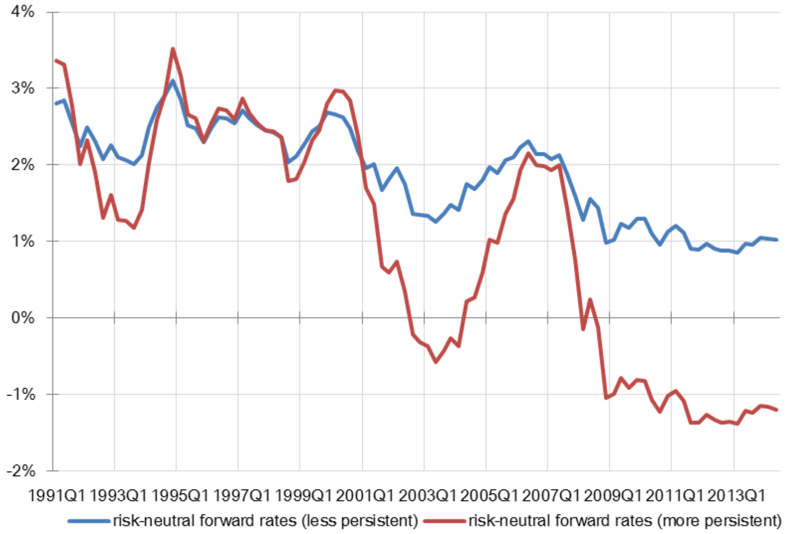
Figure 7. In-Two-Years-for-One-Year Term Premium ($TP_t^{1,2}$) of Different Models



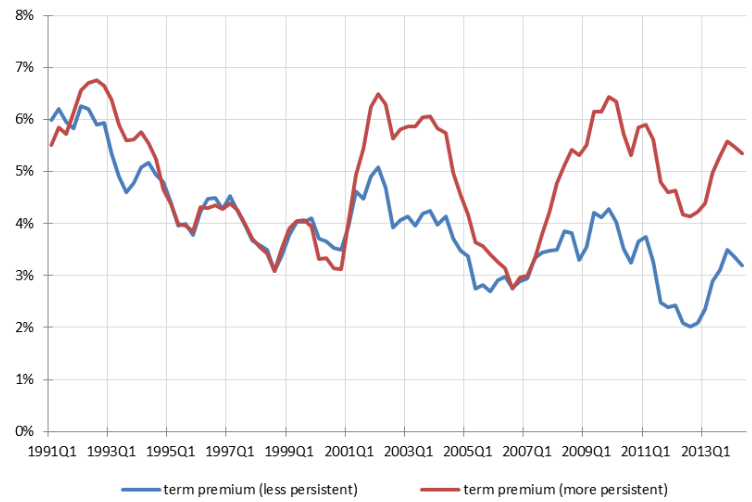
This figure plots the estimated “in-two-years-for-one-years” term premium of different models and the federal funds target rate.

Figure 8. Risk-Neutral Forward Rates and Term Premium of AGTSM under Different Persistence of the VAR System

a. Risk-neutral forward rates



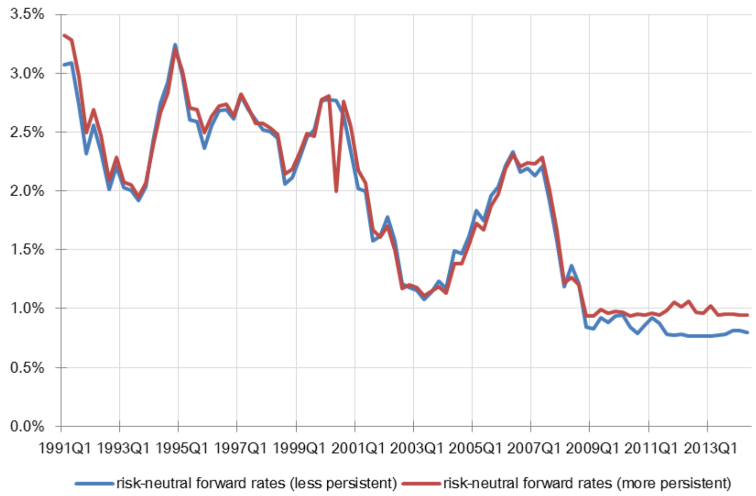
b. Term premia



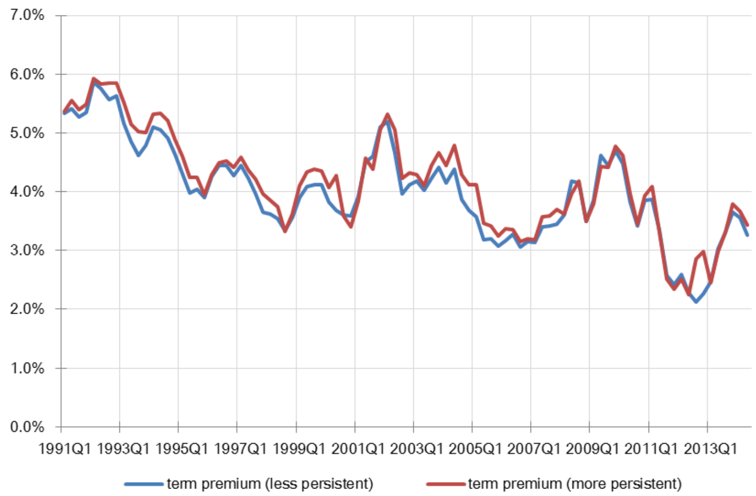
This figure plots the estimated risk-nentral forward rates (upper panel) and the “In-ten-years-for-five-years” term premia (lower panel) for AGTSM under two persistence levels of interest rates. The maximum absolute eigenvalues of the VAR system under the ”more persistent” and “less persistent” scenarios are set to 0.996 and 0.961 respectively.

Figure 9. Risk-Neutral Forward Rates and Term Premium of QGTSM under Different Persistence of the VAR System

a. Risk-neutral forward rates



b. Term premia



This figure plots the estimated risk-nentral forward rates (upper panel) and the “In-ten-years-for-five-years” term premia (lower panel) for QGTSM under two persistence levels of interest rates. The maximum absolute eigenvalues of the VAR system under the ”more persistent” and “less persistent” scenarios are set to 0.996 and 0.961 respectively.

Appendix 1. Bond Pricing Formulas

For notational convenience, we will take $\mu^Q = \mu$ and $\Phi^{\mathrm{Q}} = \Phi$ as the risk-neutral parameters and all expectations are under the risk neutral measure Q.

A.1.1 AGTSM

The n-period zero coupon bond price can be formulated as

$$\mathrm{P_t^n} = \mathbb{E}_t[e^{-r_t} P_{t+1}^{n-1}] \ = \mathbb{E}_t[\exp(-r_t + A_{n-1} + B_{n-1}^T X_{t+1})]$$

where $r_t = \delta_0 + \delta_1^T X_t$ and X_t follows the VAR dynamics $\mathrm{X_{t+1}} = \mu + \Phi X_t + \Sigma\varepsilon_{t+1}$ with $\varepsilon_t \sim N\ (\mathbf{0}, \boldsymbol{I})$. We can substitute the expression of X_{t+1} such that

$$\mathrm{P_t^n} = \exp(-r_t + A_{n-1} + \mu + B_{n-1}^T \Phi X_t)\ \mathbb{E}_t[\exp(B_{n-1}^T \Sigma\varepsilon_{t+1})].$$

Then, we can make use of the moment generating function of $\varepsilon \sim N\ (\mathbf{0}, \boldsymbol{I})$ to compute the expectation as

$$\mathbb{E}_t[\exp(B_{n-1}^T \Sigma\varepsilon)] = \exp[\frac{1}{2} B_{n-1} \Sigma\Sigma^{\mathrm{T}} B_{n-1}^T]$$

by collecting separately the constant terms and linear terms in X_t, we obtain the recursive relationship for AGTSM such that

$$\begin{aligned} A_n &= -\delta_0 + A_{n-1} + B_{n-1}^T \mu^Q + \frac{1}{2} B_{n-1} \Sigma\Sigma^T B_{n-1}^T, \\ B_n^T &= -\delta_1^T + \Phi^Q B_{n-1}^T, \end{aligned}$$

for *n*=1,2,…,*N* with $A_1 = -\delta_0$ and $B_1^T = -\delta_1^T$.

A.1.2 QGTSM

The n-period zero coupon bond price can be formulated as

$$\begin{aligned} \mathrm{P_t^n} &= \mathbb{E}_t[e^{-r_t} P_{t+1}^{n-1}] \\ &= \mathbb{E}[\exp(-r_t + A_{n-1} + B_{n-1}^T X_{t+1} + X_{t+1}^T C_{n-1} X_{t+1})], \end{aligned}$$

where $r_t = \alpha_0 + \beta_0^T X_t + X_t^T \Psi_0 X_t$ and X_t follows the VAR dynamics $\mathrm{X_{t+1}} = \mu + \Phi X_t + \Sigma\varepsilon_{t+1}$ with $\varepsilon_t \sim N\ (\mathbf{0}, \boldsymbol{I})$. Similarly, we substitute the expression of $\mathrm{X_{t+1}}$ such that

$$(\mu + \Phi X_t + \Sigma\varepsilon_{t+1})^T C_{n-1}(\mu + \Phi X_t + \Sigma\varepsilon_{t+1}) = 2(\mu + \Phi \mathrm{X_t})^T C_{n-1}\Sigma\varepsilon_{t+1},$$

and hence

$$\begin{aligned} P_{\mathrm{t}}^{\mathrm{n}} &= \exp(-r_t + A_{n-1} + B_{n-1}^T\mu + \Phi X_t + (\mu + \Phi X_t)^T C_{n-1}(\mu + \Phi X_t)) \\ &\times \mathbb{E}_t[\exp(\Gamma_0^T \varepsilon_{t+1} + \varepsilon_{t+1}^T \Gamma_1 \varepsilon_{t+1})] \end{aligned}$$

Where

$$\Gamma_0^T = B_{n-1}^T \Sigma + 2(\mu + \Phi \mathrm{X_t})^T C_{n-1}\Sigma, \quad \Gamma_1 = \Sigma^{\mathrm{T}} C_{n-1}\Sigma.$$

In this case, we can make use of the (exponential) quadratic-form expectation for $\varepsilon \sim N\ (\mathbf{0}, \boldsymbol{I})$ as

$$\mathbb{E}_t[\exp\left(\Gamma_0^T \varepsilon + \varepsilon^{\mathrm{T}}\Gamma_1\varepsilon\right)] = \exp\left[-\frac{1}{2}\det\left(\mathbf{I} - 2\Gamma_1\right) + \frac{1}{2}\Gamma_0\left(\mathbf{I} - 2\Gamma_1\right)^{-1}\Gamma_0\right].$$

Hence, by collecting separately the constant terms, linear terms in X_t and quadratic terms in X_t, we have the recursive relationship

$$\begin{aligned} A_n = {} & -\delta_0 + A_{n-1} + B_{n-1}^T \mu^Q + (\mu^Q)^T C_{n-1}\mu^Q - \frac{1}{2}\det(\mathrm{I} - 2\Sigma^T C_{n-1}\Sigma) \\ & + \frac{1}{2}(\Sigma^T B_{n-1}^T + 2\Sigma^T C_{n-1}\mu^Q)^T (\mathrm{I} - 2\Sigma^T C_{n-1}\Sigma)^{-1}(\Sigma^T + 2\Sigma^T C_{n-1}\mu^Q), \end{aligned}$$

$$B_n^T = -\beta_0 + B_{n-1}^T \Phi^Q + 2\mu^Q C_{n-1}\Phi^Q + 2(\Sigma^T B_{n-1} + 2\Sigma^T C_{n-1}\mu^Q)^T (\mathrm{I} - 2\Sigma^T C_{n-1}\Sigma)^{-1}\Sigma^T C_{n-1}\Phi^Q,$$

$$C_n = -\Psi_0 + (\Phi^Q)^T C_{n-1}\Phi^Q + 2(\Sigma^T C_{n-1}\Phi^Q)^T (\mathrm{I} - 2\Sigma^T C_{n-1}\Sigma)^{-1}(\Sigma^T C_{n-1}\Phi^Q),$$

for *n*=1,2,…,*N* with $A_1 = -\alpha_0$ and $B_1^T = -\beta_0$ and $C_1 = -\Psi_0$.

Appendix 2. Bayesian Estimation

The Bayesian estimation follows Ang et al. (2011) closely.

A.2.1 State Space Formulation

In this section, we discuss the estimation methodology in more detail. First, it is useful to express more explicitly the state space model in Section 3.4 as follows:

Measurement equation. Factor loadings a_n, b_n, and c_n are derived from the recursive relationship described in Section 4. The measurement equations for the observable bond yields and macro variables are related to the latent factors $X_t = (x_t^1, x_t^2, x_t^3)$ by:

$$\widehat{X_t} = X_t + \omega_{X,t},$$

and

$$\mathrm{y}_{\mathrm{t}}^n = a_n + b_n^T X_t + X_t^T c_n X_t + \omega_{n,t}.$$

where $\widehat{X_t} = (\hat{x}_t^1, \hat{x}_t^2, \hat{x}_t^3)$ is the observed state vector with measurement errors ω_{it} and $\mathrm{X_t}$ is the unobservable state vector. Here, *M* and *N* denote the numbers of factors and yields respectively. The third term of the RHS represents the quadratic multiplication where $0_{M\times M}$ and c_n, are *M* x *M* matrices. When we set the matrix $c_n = 0_{M\times M}$, the QGTSM reduces to the AGTSM and we have a linear state-space model.

State equation. The state equation with the parameters μ^{P} and Φ^{P} is given by:

$$\mathrm{X_{t+1}} = \mu^P + \Phi^{\mathrm{P}} X_t + \Sigma\varepsilon_{t+1}$$

which can be expressed as:

$$\begin{bmatrix} x_{t+1}^1 \\ x_{t+1}^2 \\ x_{t+1}^3 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} x_t^1 \\ x_t^2 \\ x_t^3 \end{bmatrix} + \Sigma \begin{bmatrix} \varepsilon_{t+1}^1 \\ \varepsilon_{t+1}^2 \\ \varepsilon_{t+1}^3 \end{bmatrix}.$$

The equation is a standard VAR(1) system.

A.2.2 MCMC Algorithm

As can be seen from the measurement equation, the state space model is non-linear which makes standard filtering techniques, such as the Forward Filtering Backwards Sampling, inapplicable. Instead, we adopt a single-move sampler for the unobservable state variables $\mathrm{X_t}$.

Step 1: Drawing the latent factor $\mathrm{X_t}$.

We adopt the single-move sampler and generate latent factors using a random walk Metropolis Hasting (MH) acceptance and rejection rule. The conditional posterior density:

$$\mathrm{P}\left(\mathrm{X_t}\middle|X_{t-1},\tilde{Y},\Theta\right) \propto P(X_t|X_{t-1})P\left(\widetilde{Y_t}\middle|\mathrm{X_t},\Theta\right)\mathrm{P}(\mathrm{X_{t+1}}|X_t),$$

Where

$$\mathrm{P}(\mathrm{X_t}|X_{t-1},\Theta) \propto \exp\left(-\frac{1}{2}(X_t-\mu^P-\Phi^{\mathrm{P}}\mathrm{X_{t-1}})^T(\Sigma\Sigma^{\mathrm{T}})^{-1}(X_t-\mu^P-\Phi^{\mathrm{P}}\mathrm{X_{t-1}})\right),$$

and

$$P\left(\widetilde{Y_t}\middle|\mathrm{X_t},\Theta\right) \propto \left(-\frac{1}{2}\sum_n\left[\frac{\left(\widetilde{y_t^n}-(a_n+b_n^TX_t+X_t^Tc_nX_t)\right)^2}{\sigma_n^2}\right]\right),$$

where $\widetilde{Y_t}$ is observable variables including yields and state variables and Θ is parameters. The standard deviation of the random walk MH step is taken to be 0.0005 (i.e., 5 bps).

Step 2: Drawing $(\mu^P,\Phi^{\mathrm{P}})$ and $\boldsymbol{\Sigma\Sigma'}$ under the physical measure P.

This step is implemented as a standard Bayesian VAR. In particular, we use the Minnesota prior, which assumes the mean of the VAR coefficients follow a random walk process. The conditional posterior density for $(\mu^P,\Phi^{\mathrm{P}})$ is:

$$\mathrm{P}\left(\mu^P,\Phi^{\mathrm{P}}\middle|\Theta\_,X,\tilde{Y}\right) \propto P(X|\mu^P,\Phi^{\mathrm{P}},\Sigma)P(\mu^P,\Phi^{\mathrm{P}})$$

where $P(X|\mu^P,\Phi^{\mathrm{P}},\Sigma)$ is the likelihood function and $P(\mu^P,\Phi^{\mathrm{P}})$ is the prior. For the variance $\boldsymbol{\Sigma\Sigma'}$, we have the inverse Wishart distribution as a prior and sample from the proposal density:

$$q(\Sigma\Sigma') = P(X|\mu,\Phi,\Sigma)P(\Sigma\Sigma')$$

where $P(X|\mu,\Phi,\Sigma)$ and $P(\Sigma\Sigma')$ are the likelihood function and prior, respectively. See Del Negro and Schorfheide (2011) for details on Bayesian VAR.

Step 3: Drawing (μ^Q, Φ^Q) under the risk-neutral measure Q.

We use the random walk MH algorithm, and sample μ^{Q} and Φ^Q from a proposal draw using the random walk process $\mathrm{x}^{\mathrm{m}} = x^{m-1} + \varepsilon^m$, where m is iteration and $\varepsilon^m \sim N\ (0, \sigma^2)$. A proposal draw is then accepted with the probability:

$$\alpha = \min\left\{\frac{P\left(\tilde{Y}\middle|(\mu^{\mathrm{Q}}, \Phi^Q)^{m+1}, \Theta\_, \mathrm{X}\right)}{P\left(\tilde{Y}\middle|(\mu^{\mathrm{Q}}, \Phi^Q)^{m}, \Theta\_, \mathrm{X}\right)}, 1\right\}$$

where $P\left(\tilde{Y}\middle|(\mu^{\mathrm{Q}}, \Phi^Q)^{m+1}, \Theta\_, \mathrm{X}\right)$ is the likelihood function or the posterior density as we assume a flat prior as in Ang et al. (2011). The standard deviation of the random walk MH step is taken to be 0.1% of the magnitude of the initial parameters.

Step 4: Drawing the variance of measurement error (σ_u).

We take the inverted Gamma distribution as prior with $IG(0, 10^{-4})$ in order to sample σ_u.

It is important to note that we do not estimate explicitly the loading coefficients for the short-rate functions in both the AGTSM and QGTSM. This allows us to avoid an identification problem (as our state variables are observed with errors) and also gives a more efficient estimation on the model parameters. We follow Ang et al. (2011) to preset the initial loading coefficients such that the moments of the bond yields and state variables are internally consistent.

Appendix 3. Estimated Parameters of the Term-Structure Models

The posterior estimates of the model parameters are presented in the table below. The reported values for the parameters μ and $(\Sigma\Sigma^T)_{ij}$ are multiplied by 10,000.

	AQTSM		QGTSM1		QGTSM2	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
			VAR(1)-system under P-measure			
Φ_{11}	0.9812	0.0165	0.9893	0.0194	0.9795	0.0218
Φ_{12}	0.0049	0.019	0.0036	0.0217	-0.001	0.0273
Φ_{13}	0.0483	0.0381	0.0319	0.0501	0.0542	0.0579
Φ_{21}	-0.0296	0.0295	0.0041	0.044	-0.0232	0.0262
Φ_{22}	0.9812	0.0338	0.9863	0.0511	0.9799	0.0325
Φ_{23}	0.1473	0.0662	0.17	0.1161	0.144	0.0708
Φ_{31}	0.0138	0.0239	-0.0038	0.025	0.0092	0.0209
Φ_{32}	-0.0022	0.0277	-0.0039	0.0292	-0.001	0.026
Φ_{33}	0.8828	0.0529	0.887	0.064	0.8762	0.056
μ_1	-1.0171	2.1848	-0.6267	2.7251	-2.2448	3.2316
μ_2	-2.5093	3.9721	-2.3836	6.2505	-3.7195	3.9006
μ_3	1.0774	3.179	0.9371	3.5622	1.6464	3.0996
			VAR(1)-system under Q-measure			
Φ_{11}	0.8336	0.0033	0.8351	0.0072	0.8616	0.0074
Φ_{12}	0.2936	0.01	0.3246	0.019	0.3068	0.0077
Φ_{13}	0.6613	0.017	0.6819	0.0245	0.6561	0.0223
Φ_{21}	-0.0949	0.0018	-0.0885	0.0039	-0.0842	0.0021
Φ_{22}	1.153	0.0148	1.1024	0.0474	1.1769	0.0116
Φ_{23}	0.4286	0.0107	0.4375	0.0309	0.4918	0.0078
Φ_{31}	0.1679	0.0048	0.1398	0.0036	0.1387	0.0061
Φ_{32}	-0.3785	0.0204	-0.3392	0.0088	-0.4133	0.0082
Φ_{33}	0.2748	0.006	0.277	0.0064	0.2832	0.0071
μ_1	25.6538	1.1095	31.5955	0.7228	23.1007	0.6987
μ_2	12.2652	0.2382	12.9702	0.6465	11.1926	0.5583
μ_3	-26.0571	0.6543	-25.7683	0.6568	-25.5042	0.2993
			Variance Matrix			
$(\Sigma\Sigma^T)_{11}$	0.0146	0.0025	0.0228	0.0064	0.0327	0.0064
$(\Sigma\Sigma^T)_{12}$	0.0137	0.0039	0.0222	0.0073	0.0258	0.0059
$(\Sigma\Sigma^T)_{13}$	-0.0037	0.0029	-0.007	0.0038	-0.002	0.0046
$(\Sigma\Sigma^T)_{21}$	0.0137	0.0039	0.0222	0.0073	0.0258	0.0059
$(\Sigma\Sigma^T)_{22}$	0.0478	0.0115	0.1223	0.0268	0.0469	0.0093
$(\Sigma\Sigma^T)_{23}$	-0.0278	0.0064	-0.0605	0.0152	-0.0218	0.0068
$(\Sigma\Sigma^T)_{31}$	-0.0037	0.0029	-0.007	0.0038	-0.002	0.0046
$(\Sigma\Sigma^T)_{32}$	-0.0278	0.0064	-0.0605	0.0152	-0.0218	0.0068
$(\Sigma\Sigma^T)_{33}$	0.0316	0.0059	0.039	0.0093	0.0295	0.0072

Notes:
1. The posterior mean is computed by averaging the MCMC draws Std. Dev. is computed as the sample standard deviation of the MCMC draws.