IMES DISCUSSION PAPER SERIES

Design and Estimation of a Quadratic Term Structure Model with a Mixture of Normal Distributions

Kentaro Kikuchi

Discussion Paper No. 2012-E-8

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

2-1-1 NIHONBASHI-HONGOKUCHO
CHUO-KU, TOKYO 103-8660
JAPAN

IMES Discussion Paper Series 2012-E-8
June 2012

Design and Estimation of a Quadratic Term Structure Model with a Mixture of Normal Distributions

Kentaro Kikuchi*

Abstract

To keep yields non-negative in a quadratic Gaussian term structure model (QGTM), the short rate is represented by the quadratic form of the Gaussian state variables. The QGTM is among the most attractive candidate tools for analyzing yield curves for countries with low interest rates. However, the model is unlikely to capture the fat-tailed feature of changes in yields observed in actual bond markets. This study extends the QGTM by introducing state variables whose future distributions follow a mixture of normal distributions. This extension allows our model to accommodate vast changes in non-negative yields. As an illustrative empirical study, we applied our model to Japanese government bond (JGB) yields using the unscented Kalman filter. We then used the parameters obtained to investigate market views on past JGB interest rates by simulating future interest rate probability distributions under the physical measure and by decomposing interest rates into expected short rates and term premia.

Keywords: affine term structure model; quadratic Gaussian term structure model; mixture of normal distributions; unscented Kalman filter; maximum likelihood method

JEL classification: C13, E43, G12

*Deputy Director and Economist, Institute for Monetary and Economic Studies, Bank of Japan (E-mail: kentarou.kikuchi@boj.or.jp)

The author would like to thank Don H. Kim (Yonsei University), participants of the October 2011 JAFEE Workshop and the December 2011 Nakanoshima Workshop on Mathematical Finance, and the staff of the Bank of Japan for their useful comments. Views expressed in this paper are those of the author and do not necessarily reflect the official views of the Bank of Japan.

1. Introduction

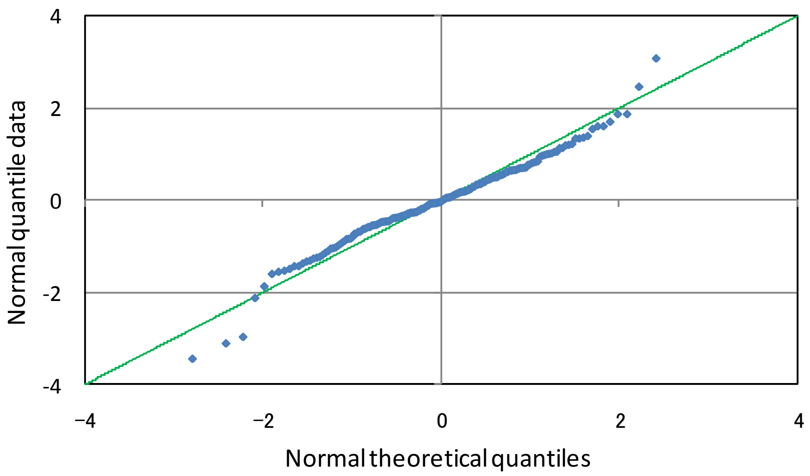
In empirical studies of the term structure of interest rates, the affine Gaussian term structure model (AGTM), one of the affine term structure models originally introduced by Duffie and Kan [1997], is the model most commonly applied, due to the ease of estimation it offers. In the AGTM, the short rate is represented as the affine function of state variables following Gaussian processes. Together with the above setting, the no-arbitrage condition leads to yields that are also represented as the affine function of these Gaussian state variables. Model parameters and state variables are readily estimated, since this model has a Gaussian affine feature. The maximum likelihood method with the Kalman filter is often applied to estimate the AGTM. This method is analytically precise and does not require approximations. This is one likely reason why many empirical studies of yield curves are based on the AGTM.

The AGTM has certain drawbacks. The best-known is that it cannot ensure the non-negativity of theoretical yields. This suggests the AGTM would be hard to use to accurately analyze yield curves for countries with low interest rates. We can resort to an affine term structure model with a square root diffusion process to circumvent this problem. However, as pointed out by Dai and Singleton [2000], the square root affine model imposes non-negative correlations between state variables to generate well-defined bond prices. This constraint is restrictive and makes the model less flexible for application to actual data.

Ahn, Dittmar, and Gallant [2002] and Leippold and Wu [2003] developed the quadratic Gaussian term structure model (QGTM) to address the drawbacks of the AGTM. In the QGTM, the short rate is represented by the quadratic form of the state variables that follow Gaussian processes. The authors of the preceding two papers derived an analytical pricing formula for bonds under the no-arbitrage condition based on this setting. Several empirical studies based on the QGTM have appeared recently (Kim and Singleton [2011] and Nyholm and Vidova-Koleva [2010]). If we attempt to estimate the QGTM, the Kalman filter is invalid, since theoretical yields are the nonlinear functions of state variables in the model. Instead, nonlinear filtering methods such as the extended Kalman filter and the unscented Kalman filter developed by Julier and Uhlmann [1997] are used to estimate the model parameters and state variables of the model.

The QGTM results in non-negative yields, an advantage over the AGTM. However, the QGTM is unlikely to adequately capture the fat-tailed feature of changes in yields. This feature is often observed in bond and interest rate markets. In fact, monthly changes in Japanese government bond (JGB) yields exhibit the fat-tailed feature shown as a normal Q-Q plot in Figure 1.

Figure 1: Normal Q-Q plot of monthly changes in 10-year JGB yields



Note: In this plot, monthly changes in 10-year JGB yields are standardized to mean zero and unit variance for the period from January 1996 to December 2010. The straight line with a 45 degree slope in the above figure indicates the data follows a normal distribution.

This study extends the QGTM to explicitly incorporate this fat-tailed feature into the term structure model. In our model, the short rate is represented by the quadratic form of the state variables whose future conditional distributions follow a mixture of normal distributions. We call this model the quadratic mixture of Gaussian term structure model (QMGTM). Under the no-arbitrage condition, we derive the bond pricing formula based on the QMGTM using a log-linear approximation. The model generates non-negative yields and exhibits the fat-tailed feature of changes in yields.

Our model provides more accurate yield estimates from actual data. Moreover, because it lets us capture probability distributions for future interest rates under the physical probability measure, our model lets us analyze market views on interest rates. We applied this model to investigate market views on JGB yields, simulating future interest rate probability distributions under the physical measure and decomposing interest rates into expected future short rates and term premia.

The remainder of this paper is organized as follows. Section 2 describes our model, the QMGTM, as well as the AGTM and the QGTM. Section 3 describes the estimation methodology. Section 4 presents a performance comparison of the AGTM, QGTM, and QMGTM and the results of an empirical study for JGB yields. Section 5 presents our conclusions.

2. Term Structure Models

This section describes the structures of the AGTM, QGTM, and our model, the QMGTM. It also describes detailed settings for parameters in the discrete time framework for each model.

2.1 Affine Gaussian Term Structure Model

In the AGTM, the short rate is represented as the affine function of state variables following Gaussian processes. To extract market views from observed yields, it is convenient to represent state variable processes under both the risk neutral measure **Q** and the physical measure **P** as follows:

$$X_{t+1} = a^{\mathbf{Q}} + \Phi^{\mathbf{Q}} X_t + \Sigma \varepsilon_{t+1}^{\mathbf{Q}},\ \varepsilon_{t+1}^{\mathbf{Q}} \sim i.i.d.N(0,I), \tag{1}$$

$$X_{t+1} = a^{\mathbf{P}} + \Phi^{\mathbf{P}} X_t + \Sigma \varepsilon_{t+1}^{\mathbf{P}},\ \varepsilon_{t+1}^{\mathbf{P}} \sim i.i.d.N(0,I), \tag{2}$$

where X_t denotes an $N \times 1$ vector of state variables at time t and $N(0,I)$ is the normal distribution with zero mean and covariance matrix I, the $N \times N$ unit matrix. We denote shock at time t under **Q** by $\varepsilon_t^{\mathbf{Q}}$ and shock at time t under **P** by $\varepsilon_t^{\mathbf{P}}$. The shock $\varepsilon_t^{\mathbf{Q}}$ and the shock $\varepsilon_t^{\mathbf{P}}$ have the following relationship mediated by market price of risk Λ_t:

$$\varepsilon_t^{\mathbf{Q}} = \varepsilon_t^{\mathbf{P}} + \Lambda_t . \tag{3}$$

Throughout this paper, as in Duffee [2002], we assume that Λ_t is a linear function of a vector of state variables X_t:

$$\Lambda_t = \lambda_0 + \Lambda_1 X_t . \tag{4}$$

This is the "essentially affine" assumption.

From equations (1)–(4), we find that

$$a^{\mathbf{P}} = a^{\mathbf{Q}} + \Sigma\lambda_0,\ \Phi^{\mathbf{P}} = \Phi^{\mathbf{Q}} + \Sigma\Lambda_1 .$$

In the AGTM, the short rate r_t is represented as the affine function of X_t as shown below:

$$r_t = \delta_0 + \delta_1' X_t , \tag{5}$$

where δ_0 is a scalar and δ_1 an $N \times 1$ vector. In addition, δ_1' represents the transpose of δ_1. Hereafter, we assume that A' denotes the transpose of a vector or a matrix *A*.

Under the no-arbitrage condition, in addition to the above setting, the following relation must hold between a zero coupon bond price P_t^n with *n* period maturity at time *t* and a zero coupon bond price P_{t+1}^{n-1} with *n* − 1 period maturity at time *t* + 1:

$$P_t^n = E_t^{\mathbf{Q}}[e^{-r_t} P_{t+1}^{n-1}], \tag{6}$$

where $E_t^{\mathbf{Q}}[\cdot]$ denotes the conditional expectation operator under measure **Q** conditioned on information obtained at time *t*.

The boundary condition of equation (6) is $P_t^0 = 1$. Taking into consideration this boundary condition, we can derive the following analytical bond pricing formula (7) from equation (1), (5), and (6).

$$\begin{aligned} &P_t^n = \exp(a_n + b_n' X_t), \\ &a_n = -\delta_0 + a_{n-1} + b_{n-1} a^{\mathbf{Q}} + \frac{1}{2} b'_{n-1} \Sigma\Sigma' b_{n-1},\ a_0 = 0, \\ &b_n = -\delta_1' + b'_{n-1} \Phi^{\mathbf{Q}},\ b_0 = \mathbf{0}. \end{aligned} \tag{7}$$

Therefore, the theoretical yield y_t^n with *n* period maturity is written as the affine function of X_t in the AGTM:

$$y_t^n = -\frac{1}{n} a_n - \frac{1}{n} b_n' X_t . \tag{8}$$

As implied by the form of the yields in equation (8), yields under the AGTM are likely to be negative.

When we attempt an estimate based on actual observed yields, we interpret the observed yield $\tilde{y}_t^n$ as the sum of the theoretical yield y_t^n and Gaussian noise η_t^n with zero mean. That is,

$$\tilde{y}_t^n = y_t^n + \eta_t^n . \tag{9}$$

We will proceed to specify detailed parameter settings for the AGTM required for estimates of model parameters.

The number of factors for all models analyzed in this paper is assumed to be three. Using the invariant transformation discussed by Dai and Singleton [2000], we can set several parameters for the AGTM without loss of generality as follows:

$$a^{\mathbf{P}} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Phi^{\mathbf{P}} = \begin{pmatrix} \Phi_{11}^{\mathbf{P}} & 0 & 0 \\ \Phi_{21}^{\mathbf{P}} & \Phi_{22}^{\mathbf{P}} & 0 \\ \Phi_{31}^{\mathbf{P}} & \Phi_{32}^{\mathbf{P}} & \Phi_{33}^{\mathbf{P}} \end{pmatrix}, \Sigma = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{pmatrix}, \tag{10}$$
$$0 < \Phi_{11}^{\mathbf{P}} < 1,\ 0 < \Phi_{22}^{\mathbf{P}} < 1,\ 0 < \Phi_{33}^{\mathbf{P}} < 1.$$

Furthermore, the parameters for the market price of risk, λ_0 and Λ_1 in equation (4) are assumed to be a three-dimensional full vector or a full matrix:

$$\lambda_0 = \begin{pmatrix} \lambda_{0,1} \\ \lambda_{0,2} \\ \lambda_{0,3} \end{pmatrix}, \Lambda_1 = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{pmatrix}. \tag{11}$$

2.2 Quadratic Gaussian Term Structure Model

As pointed out above, yields obtained from the AGTM can take negative values. To address this drawback, Ahn, Dittmar, and Gallant [2002] and Leippold and Wu [2003] developed the QGTM.

We assume that the stochastic process of the vector of state variables X_t is of the same form as equations (1) and (2). We also assume that the market price of risk of X_t is the same as equation (4). This is the essentially affine market price of risk. In this model, the short rate r_t is represented by the following quadratic form of X_t:

$$r_t = X_t' A X_t . \tag{12}$$

To ensure the non-negativity of the short rate, the $N \times N$ matrix A in equation (12) is assumed to be positive semi-definite.

Under the no-arbitrage condition, we can derive the bond pricing formula from equations (1), (6), and (12), as follows:

$$\begin{aligned}
P_t^n &= \exp(X_t' A_n X_t + b_n' X_t + c_n), \\
A_n &= -A + \Phi^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{n-1}^{-1}A_{n-1}), A_0 = -A, \\
b_n &= \Phi^{\mathbf{Q}}{}' b_{n-1} + 2\Phi^{\mathbf{Q}}{}' A_{n-1}\Psi_{n-1}^{-1} b_{n-1} + 2\Phi^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{n-1}^{-1}A_{n-1}) a^{\mathbf{Q}}, b_0 = \mathbf{0}, \\
c_n &= c_{n-1} + b_{n-1}' a^{\mathbf{Q}} - 0.5\log|\Psi_{n-1}| - 0.5\log|\Sigma\Sigma'| + 2a^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{n-1}^{-1}A_{n-1}) a^{\mathbf{Q}} \\
&\quad + 0.5 b_{n-1}'\Psi_{n-1}^{-1} b_{n-1} + 2 b_{n-1}'\Psi_{n-1}^{-1} A_{n-1} a^{\mathbf{Q}}, \ c_0 = 0,
\end{aligned} \tag{13}$$

where $\Psi_n = (\Sigma\Sigma')^{-1} - 2A_n$.

Theoretical yield y_t^n with n period maturity under the QGTM is written as the quadratic form of X_t, using A_n, b_n, and c_n, defined in equation (13) as follows:

$$y_t^n = -\frac{1}{n} X_t' A_n X_t - \frac{1}{n} b_n' X_t - \frac{1}{n} c_n. \tag{14}$$

As implied by the form of the yields in equation (14), the yields must be non-negative.

We will specify detailed parameter settings for the QGTM required by the estimate. Using an invariant transformation for the QGTM discussed by Leippold and Wu [2003], we can set several parameters without loss of generality as follows:

$$\begin{aligned}
&a^{\mathbf{P}} = \begin{pmatrix} a_1^{\mathbf{P}} \\ a_2^{\mathbf{P}} \\ a_3^{\mathbf{P}} \end{pmatrix}, \Phi^{\mathbf{P}} = \begin{pmatrix} \Phi_{11}^{\mathbf{P}} & 0 & 0 \\ \Phi_{21}^{\mathbf{P}} & \Phi_{22}^{\mathbf{P}} & 0 \\ \Phi_{31}^{\mathbf{P}} & \Phi_{32}^{\mathbf{P}} & \Phi_{33}^{\mathbf{P}} \end{pmatrix}, \Sigma = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{pmatrix}, \\
&0 < \Phi_{11}^{\mathbf{P}} < 1, \ 0 < \Phi_{22}^{\mathbf{P}} < 1, \ 0 < \Phi_{33}^{\mathbf{P}} < 1.
\end{aligned} \tag{15}$$

Based on the singular value decomposition of the positive semi-definite matrix, we represent A in equation (12) as the product of certain matrices:

$$\begin{aligned}
&A = \begin{pmatrix} 1 & 0 & 0 \\ s_{21} & 1 & 0 \\ s_{31} & s_{32} & 1 \end{pmatrix} \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix} \begin{pmatrix} 1 & s_{21} & s_{31} \\ 0 & 1 & s_{32} \\ 0 & 0 & 1 \end{pmatrix} \equiv SDS^T, \\
&0 \le d_1 < 1, \ 0 \le d_2 < 1, \ 0 \le d_3 < 1,
\end{aligned} \tag{16}$$

where S^T denotes the transpose of matrix S.

For the parameters of market price of risk in the QGTM, we assume that the setting is the same as equation (11).

2.3 Our Model: Quadratic Mixture of Gaussian Term Structure Model

In this subsection, we propose a new term structure model, an extension of the QGTM. While the QGTM ensures non-negative yields, the model may not capture extremely the large changes in yields infrequently observed in bond markets. Thus, we model the stochastic process of a vector of state variables not as a Gaussian process, but as a process whose marginal densities follow a mixture of normal distributions. More specifically, the process of a vector of state variables X_t is defined using a new stochastic process s_t under the physical measure $\mathbf{P}$, as follows:

$$\begin{aligned} &X_{t+1} = (a^{\mathbf{P}} + \mu_{s_{t+1}}) + \Phi^{\mathbf{P}} X_t + \Sigma_{s_{t+1}} \varepsilon^{\mathbf{P}}_{t+1},\ \varepsilon^{\mathbf{P}}_{t+1} \sim i.i.d.N(0, I), \\ &\mathbf{P}(s_{t+1} = 1 \mid I_t) = p,\ \mathbf{P}(s_{t+1} = 2 \mid I_t) = 1 - p,\ s_t \text{ are } i.i.d.,\ s_{t+1} \perp \varepsilon^{\mathbf{P}}_{t+1}, \\ &p\mu_1 + (1-p)\mu_2 = 0. \end{aligned} \tag{17}$$

We denote the information set obtained at time t by I_t. Equation (17) implies that one-period conditional distributions of future state variables at time t under $\mathbf{P}$ follow a mixture of normal distributions. This model lets us incorporate the fat-tailed feature of yields into a term structure model. Lemke [2006] assumes the same setting for X_t's dynamics in equation (17). However, the short rate there is assumed to be an affine function; in contrast, we set r_t as the quadratic function of X_t.

We assume that the market price of risk of our model also follows the essentially affine form of equation (4), as with the AGTM and the QGTM. Equations (3), (4), and (17) give the dynamics of a vector of state variables under the risk neutral measure $\mathbf{Q}$, as follows:

$$\begin{aligned} X_{t+1} &= (a^{\mathbf{P}} + \mu_{s_{t+1}} - \lambda_0 \Sigma_{s_{t+1}}) + (\Phi^{\mathbf{P}} - \Lambda_1 \Sigma_{s_{t+1}}) X_t + \Sigma_{s_{t+1}} \varepsilon^{\mathbf{Q}}_{t+1},\ \varepsilon^{\mathbf{Q}}_{t+1} \sim i.i.d.N(0, I), \\ &\equiv a^{\mathbf{Q}}_{s_{t+1}} + \Phi^{\mathbf{Q}}_{s_{t+1}} X_t + \Sigma_{s_{t+1}} \varepsilon^{\mathbf{Q}}_{t+1}. \end{aligned} \tag{18}$$

In addition, we assume that the conditional probability for s_{t+1} under $\mathbf{Q}$ at time t follows equation (19):

$$\begin{aligned} &\mathbf{Q}(s_{t+1} = 1 \mid I_t) = \mathbf{P}(s_{t+1} = 1 \mid I_t) = p, \\ &\mathbf{Q}(s_{t+1} = 2 \mid I_t) = \mathbf{P}(s_{t+1} = 2 \mid I_t) = 1 - p. \end{aligned} \tag{19}$$

We also assume that the short rate has the same form as equation (12) in the QGTM. This assumption allows the short rate to remain non-negative.

Under the certain assumptions indicated above and the no-arbitrage condition, we seek to derive the zero coupon bond pricing formula. However, we must rely on an approximation due to the difficulty of deriving the closed-form solution within this framework. Thus, we use the log-linear approximation used by Bansal and Zhou [2002], who studied the affine term structure model in regime switching.

First, we assume that a zero coupon bond price can be described $P_t^n = \exp(X_t' A_n X_t + b_n' X_t + c_n)$. Under this assumption and the no-arbitrage condition (equation (6)), we obtain the following equation,

$$\begin{aligned} \exp(X_t' A_n X_t + b_n' X_t + c_n) &= E_t^{\mathbf{Q}}[\exp(-r_t)\exp(X_{t+1}' A_{n-1} X_{t+1} + b_{n-1}' X_{t+1} + c_{n-1})] \\ &= \exp(-X_t' A_n X_t) E_t^{\mathbf{Q}}[\exp(X_{t+1}' A_{n-1} X_{t+1} + b_{n-1}' X_{t+1} + c_{n-1})]. \end{aligned} \tag{20}$$

Next, when we try to derive the recursive relationship of coefficients A_n, b_n, and c_n from equation (20), we substitute equation (18) into the right hand side of equation (20). Here, we use a log-linear approximation, $\exp(x) \approx 1 + x$, to solve equation (20). Appendix 1 gives the recursive solutions of coefficients A_n, b_n, and c_n based on this approximation. As the maturity of yield is longer, this degrades the precision of the approximation.[1] Nevertheless, since yields estimated from our model demonstrate a good fit with observation yields with maturities of six months and two, five, 10, and 20 years, as indicated in Section 4, we see that it suffices for our model to extract information included in the observed yields. For this reason, our use of the approximation formula is valid for analyses of short rate and 10-year yields over the next six months and term premia of two and 10-year yields in Section 4.

Hereafter, we refer to our model as the quadratic mixture of Gaussian term structure model (QMGTM).

We will now proceed to specify the details of the QMGTM to estimate model parameters and state variables X_t.

First, the parameters in equation (17) are set using the following invariant

[1] We compared yields based on our model approximation formula to those given by Monte Carlo simulations using state variables and parameters estimated from JGB zero coupon yields indicated in Appendix 2. Error ratios averaged during the sample period were 0.40% for six-month yields, 0.50% for two-year yields, 0.74% for five-year yields, 2.31% for 10-year yields, and 5.52% for 20-year yields.

transformation proposed by Leippold and Wu [2002]:

$$a^{\mathbf{P}} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \Phi^{\mathbf{P}} = \begin{pmatrix} \varphi_{11} & 0 & 0 \\ \varphi_{21} & \varphi_{22} & 0 \\ \varphi_{31} & \varphi_{32} & \varphi_{33} \end{pmatrix}, \mu_1 = \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \end{pmatrix},$$
$$\Sigma_1 = \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{pmatrix}, \ \Sigma_2 = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}. \tag{21}$$

The invariant transformation enables $a^{\mathbf{P}}$, $\Phi^{\mathbf{P}}$, and Σ_1 in equation (18) to turn into the form of equation (21). Note that Σ_2 is generally represented as a triangular matrix.

Next, we assume that the "mixture" probability p of the state variable transition in equation (17) is a fixed value close to 1, not a free parameter. This is because dealing with p as a free parameter destabilizes model parameter estimates.[2] Specifically, we set $p = 0.95$ as a fixed value close to 1. This allows us to capture the fat-tailed feature of changes in yields with a low probability of occurrence.[3] Based on this, in Section 4, we compare QMGTM performance to AGTM and QGTM performance.

Finally, for the parameters of market price of risk in the QMGTM, we assume that the setting is the same as for equation (11).

Section 4 applies these models to actual data and compares the performance of our model to the two other models. We will also consider estimates of parameters and state variables of the models from the actual data.

3. Estimation Methodology

As is well-known, when the term structure model is described as a state space model with an affine Gaussian structure, the maximum likelihood estimation with a Kalman filter is useful for simultaneous estimation of latent state variables and model parameters. However, once the state space model is non-Gaussian or nonlinear, we

---

[2] In the QMGTM setting, the case in which $\mu_1 = 0, \Sigma_1 = \Sigma_2, 0 < \forall p < 1$ and the other case in which $p = 0$ or $p = 1$ both represent the QGTM. These two cases are indistinguishable in the estimation. If the QGTM captures actual data accurately, we confront the problem of distinguishing between model parameters in the QMGTM.

[3] We also estimated the model parameters of the QMGTM when $p = 0.9$, aside from the case in which $p = 0.95$. The results with $p = 0.9$ and $p = 0.95$ were near-identical.

cannot use the maximum likelihood estimation with a Kalman filter. We must rely instead on a nonlinear filtering method. A nonlinear filtering method is required for estimates of the QGTM and QMGTM, since these two models are not affine. In this study, we use the unscented Kalman filter developed by Julier and Uhlmann [1997] for estimations of the QGTM and QMGTM.

The three models described in the previous section are all regarded as state space models. State transition equations for the AGTM and the QGTM are written in the following general form:

$$X_{t+1} = f(X_t) + \Sigma \tilde{\varepsilon}^{\mathbf{P}}_{t+1}, \quad \tilde{\varepsilon}^{\mathbf{P}}_{t+1} \sim i.i.d.N(0, I), \tag{22}$$

where $f(X)$ denotes some affine function of X.

The state transition equation for the QMGTM is nearly identical to equation (22). However, it depends on the random state variable s_{t+1}, which takes the value of 1 with probability p or 2 with probability $1-p$. The state dynamics of the QMGTM is represented as follows:

$$\begin{aligned} &X_{t+1} = f_{s_{t+1}}(X_t) + \Sigma_{s_{t+1}} \tilde{\varepsilon}^{\mathbf{P}}_{t+1}, \quad \tilde{\varepsilon}^{\mathbf{P}}_{t+1} \sim i.i.d.N(0, I), \\ &\mathbf{P}(s_{t+1} = 1 \mid I_t) = p, \; \mathbf{P}(s_{t+1} = 2 \mid I_t) = 1 - p, \; s_t \; i.i.d., \\ &s_t \perp \tilde{\varepsilon}^{\mathbf{P}}_t \,. \end{aligned} \tag{23}$$

The observation equations for the AGTM, QGTM, and QMGTM are written in the following common form:

$$\tilde{Y}_t = G(X_t) + \eta_t, \quad \eta_t \sim i.i.d.N(0, \sigma^2 I), \tag{24}$$

where $\tilde{Y}_t$ is a vector with elements of some yields observed in bond markets or computed from bond prices; $G(X)$ is a vector with elements of theoretical yields corresponding to observational yields; and η_t is an observational error vector that follows a multivariate normal distribution. Here, we note that the form of the above function $G(X)$ varies depending on the choice of models. For the AGTM, $G(X)$ is the affine function of X. For the QGTM and the QMGTM, $G(X)$ is the nonlinear function of X.

We will discuss the unscented Kalman filter algorithm, taking the QMGTM as an example.

The first step of the algorithm is an initialization:

$$\hat{X}_0 := E^{\mathbf{P}}[X_0],\ P_0 := E^{\mathbf{P}}[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T]. \tag{25}$$

In this initialization step, we compute the unconditional expectation $\hat{X}_0 = E^{\mathbf{P}}[X_0]$ and variance P_0 of state variable X_t.

Next, we calculate sigma points. We denote the state dimension by N. For $k = 1, 2, \ldots, T$ where T is the number of sample periods, we calculate $2N+1$ points called sigma points:

$$\chi_{k-1} = [\hat{X}_{k-1}, \hat{X}_{k-1} + \gamma\sqrt{P_{k-1}}, \hat{X}_{k-1} - \gamma\sqrt{P_{k-1}}], \tag{26}$$

where χ_{k-1} is an $N \times (2N+1)$ matrix and $\sqrt{P_{k-1}}$ represents the square root matrix of P_{k-1}. For example, the square root matrix is computed using Cholesky decomposition. We define $\hat{X}_{k-1} + \gamma\sqrt{P_{k-1}}$ of equation (26) as follows:

$$\hat{X}_{k-1} + \gamma\sqrt{P_{k-1}} := (\hat{X}_{k-1}, \hat{X}_{k-1}, \ldots, \hat{X}_{k-1}) + \gamma\sqrt{P_{k-1}},$$

where $\gamma = \sqrt{\alpha^2(N+\kappa)}$ is called the scaling parameter. We define $\hat{X}_{k-1} - \gamma\sqrt{P_{k-1}}$, similarly. For the estimation presented in this paper, we assume that $\alpha = 1$ and $\kappa = 2$.

The next step is the time update of state variables. In the time update, we transform sigma points in equation (26) through equation (23).

$$\begin{aligned} \chi^{*,1}_{k|k-1} &= [f_1(\hat{X}_{k-1}), f_1(\hat{X}_{k-1} + \gamma\sqrt{P_{k-1}}), f_1(\hat{X}_{k-1} - \gamma\sqrt{P_{k-1}})], \\ \chi^{*,2}_{k|k-1} &= [f_2(\hat{X}_{k-1}), f_2(\hat{X}_{k-1} + \gamma\sqrt{P_{k-1}}), f_2(\hat{X}_{k-1} - \gamma\sqrt{P_{k-1}})]. \end{aligned} \tag{27}$$

Using elements of the matrices in equation (27), we calculate the "mean" and "covariance" of the state variables as below,

$$\begin{aligned} \hat{X}_k^- &= p\sum_{i=0}^{2N} W_i^{(m)} \chi^{*,1}_{i,k|k-1} + (1-p)\sum_{i=0}^{2N} W_i^{(m)} \chi^{*,2}_{i,k|k-1}, \\ P_k^- &= p\sum_{i=0}^{2N} W_i^{(c)} (\chi^{*,1}_{i,k|k-1} - \hat{X}_k^-)(\chi^{*,1}_{i,k|k-1} - \hat{X}_k^-)^T \\ &\quad + (1-p)\sum_{i=0}^{2N} W_i^{(c)} (\chi^{*,2}_{i,k|k-1} - \hat{X}_k^-)(\chi^{*,2}_{i,k|k-1} - \hat{X}_k^-)^T \\ &\quad + p\Sigma_1\Sigma_1^T + (1-p)\Sigma_2\Sigma_2^T, \end{aligned} \tag{28}$$

where weights are assigned as follows:

$$W_0^{(m)} = \frac{\alpha^2(N+\kappa)-N}{\alpha^2(N+\kappa)}, W_0^{(c)} = W_0^{(m)} + (1-\alpha^2+\beta),$$
$$W_i^{(m)} = W_i^{(c)} = \frac{1}{2\alpha^2(N+\kappa)} \quad i = 1,\ldots,2N, \tag{29}$$

where β above is a nonnegative weighting parameter introduced to adjust the weighting of the first sigma point for the covariance. This parameter can be used to incorporate knowledge of higher order distribution moments. In this paper, we set $\beta = 0$. We define sigma points of the predicted state variables as follows:

$$\chi_{k|k-1} = [\hat{X}_k^-, \hat{X}_k^- + \gamma\sqrt{P_k^-}, \hat{X}_k^- - \gamma\sqrt{P_k^-}]. \tag{30}$$

Next, we transform those sigma points $\chi_{k|k-1}$ to predicted sigma points of the yields.

$$Y_{k|k-1} = [G(\hat{X}_k^-), G(\hat{X}_k^- + \gamma\sqrt{P_k^-}), G(\hat{X}_k^- - \gamma\sqrt{P_k^-})],$$
$$\hat{Y}_k^- = \sum_{i=0}^{2N} W_i^{(m)} Y_{i,k|k-1}. \tag{31}$$

Finally, we obtain the following measurement update equations:

$$P_{Y_kY_k} = \sum_{i=0}^{2N} W_i^{(c)} (Y_{i,k|k-1} - \hat{Y}_k^-)(Y_{i,k|k-1} - \hat{Y}_k^-)^T + \sigma^2 I, \tag{32}$$

$$P_{X_kY_k} = \sum_{i=0}^{2N} W_i^{(c)} (\chi_{i,k|k-1} - \hat{X}_k^-)(Y_{i,k|k-1} - \hat{Y}_k^-)^T, \tag{33}$$

$$K_k = P_{X_kY_k} P_{Y_kY_k}^{-1}, \tag{34}$$

$$X_k = \hat{X}_k^- + K_k(Y_k - \hat{Y}_k^-), \tag{35}$$

$$P_k = P_k^- - K_k P_{Y_kY_k} K_k^T. \tag{36}$$

The unscented Kalman filter approximates the distribution of state variables relatively accurately, since it uses more information through the sigma points, as indicated above. Some previous studies have validated the performance of the unscented Kalman filter in empirical studies (e.g., Christoffersen *et al*. [2008]).

Through the unscented Kalman filter, we can compute the conditional expectation and covariance matrix of the yield vector at time *t*, $Y_t := G(X_t)$ conditioned on an observation at $t-1$. Using the conditional expectation and covariance matrix of Y_t, we then apply the quasi-maximum likelihood method to parameter estimations. The log likelihood function $L(\Theta)$ is computed as follows:

$$L(\Theta) = -\frac{1}{2}\sum_{k=0}^{T-1}(\log|P_{Y_kY_k}| + (Y_{k+1} - \widetilde{Y}_{k+1})^T (P_{Y_kY_k})^{-1}(Y_{k+1} - \widetilde{Y}_{k+1})), \tag{37}$$

where T denotes the number of observations and vector $\widetilde{Y}_{k+1}$ consists of observed yields at time $k+1$. We choose the model parameters to maximize the above log likelihood of the data series.

4. Estimation Results

In this section, we apply the AGTM, QGTM, and the QMGTM to Japanese government bond (JGB) yield data and discuss the results of the estimation. We first compare how different models fit the data, then simulate the probability distributions of future interest rates under the physical measure using estimated latent variables and model parameters. In particular, we examine the expectation as well as percentile points of the future interest rate for the past JGB interest rate, which in turn are calculated from the simulated distributions. Finally, we investigate the development of the JGB yields by decomposing them into expected future short rates and term premia for the sample period.

4.1 Data and Model Fit

We use monthly data from January 1996 through December 2010 for JGB zero coupon yields. This data is computed from the Broker's Broker JGB prices by the method presented in McCulloch [1975]. The maturities included are six months and two, five, 10, and 20 years.

Parameter and state variable estimations for the AGTM, QGTM, and the QMGTM are implemented based on the unscented Kalman filter with the quasi-maximum likelihood method specified in the previous section. Although the maximum likelihood method estimate with the Kalman filter with for the AGTM can be theoretically possible, we also apply the unscented Kalman filter to the AGTM to compare the three models. Tables A, B, and C in Appendix 2 display the estimated parameters for each model. Figure A in Appendix 2 gives the time series of the estimated state variables for each model.

We compare the fit to the data for the three models. Table 1 presents comparative statistics of fit for the three models. We compare models not just through likelihood, but two information criteria. These information criteria are determined based on the way in which the log-likelihood is penalized by the number of model parameters. Table 1 indicates that the QMGTM by any criterion gives the best performance of the three models:

Table 1: Comparative statistics for fit

	AGTM	QGTM	QMGTM
$\log(L)$	8219.85	8246.37	8296.35
AIC	-16393.7	-16436.7	-16518.7
BIC	-16320.3	-16347.3	-16400.6

Note: *L* is likelihood. AIC is Akaike information criterion. BIC is Bayesian information criterion.

Next, we see the results of model fitting for each maturity. Table 2 presents statistics on prediction errors between observed yields and predicted yields for each maturity. Comparing model fits for the three models, we find that the mean squared error (MSE) of the QMGTM for each maturity is the lowest of the three models. This is attributable to a low standard deviation of the prediction error of the QMGTM for each maturity compared to those of the two other models. The maximum value and the absolute value of the minimum value of the QMGTM prediction errors are also significantly lower than for the two other models.

Table 2: Prediction error statistics

Six-month

	AGTM	QGTM	QMGTM
Mean	0.0757	0.00104	-0.00658
Stdev	7.94	7.88	7.83
Max	25.9	27.4	23.2
Min	-33.1	-44.4	-37.9
MSE	0.00628	0.00624	0.00614

Two-year

	AGTM	QGTM	QMGTM
Mean	-0.447	-1.110	-0.809
Stdev	11.7	11.3	11.2
Max	48.9	43.8	39.5
Min	-41.1	-51.3	-40.4
MSE	0.0137	0.0130	0.0124

Five-year

	AGTM	QGTM	QMGTM
Mean	0.628	-1.645	-0.813
Stdev	15.5	15.3	15.2
Max	82.0	72.7	65.0
Min	-35.5	-47.9	-35.3
MSE	0.0239	0.0238	0.0230

10-year

	AGTM	QGTM	QMGTM
Mean	0.146	-0.756	-0.958
Stdev	15.6	15.9	15.2
Max	88.0	79.1	75.6
Min	-56.9	-64.5	-58.1
MSE	0.0241	0.0254	0.0230

20–year

	AGTM	QGTM	QMGTM
Mean	0.315	-0.675	1.13
Stdev	16.6	17.1	15.5
Max	89.6	76.3	78.8
Min	-64.2	-75.1	-67.7
MSE	0.0274	0.0291	0.0240

Note: Stdev is the standard deviation of the prediction errors. MSE is the mean squared error. Units excluding MSE are basis points. The unit for MSE is square of percent.

4.2 Probability Distributions for Future Interest Rates under the Physical Measure

This subsection discusses simulations of probability distributions for future interest rates under the physical measure based on estimated state variables and model parameters. The QGTM and QMGTM are more likely to capture future interest rates accurately than the AGTM because they ensure the non-negativity of yields. Moreover, the AGTM is clearly undesirable for this analysis since distributions for the future interest rate always follow the normal distribution in the AGTM, and its standard deviations also remain constant for the whole period represented by the data sample. Hence, we exclude the AGTM from our analysis and address only the QGTM and the QMGTM to simulate probability distributions of future interest rates.

Simulations of probability distributions for interest rates over the next six months under the physical measure are implemented based on estimates of model parameters and filtered state variables $E_t^{\mathbf{P}}[X_t]$. The simulation process is described as follows.

First, we simulate the future shocks of X_t, $\varepsilon_{t+1}^{\mathbf{P}}, \varepsilon_{t+2}^{\mathbf{P}}, \ldots, \varepsilon_{t+6}^{\mathbf{P}}$ at time *t*. Here, the number of simulations is set to 30,000. The next step for the QGTM is to compute X_{t+6} from the simulated shocks $\varepsilon_{t+1}^{\mathbf{P}}, \varepsilon_{t+2}^{\mathbf{P}}, \ldots, \varepsilon_{t+6}^{\mathbf{P}}$ based on equation (2). In contrast, for the QMGTM, the next step is to compute X_{t+6} from the simulated shocks $\varepsilon_{t+1}^{\mathbf{P}}, \varepsilon_{t+2}^{\mathbf{P}}, \ldots, \varepsilon_{t+6}^{\mathbf{P}}$ and the simulated variables $s_{t+1}, s_{t+2}, \ldots, s_{t+6}$ using equation (17). Once we obtain the simulated latent variable X_{t+6}, we can compute simulated yields over the next six months based on the zero coupon bond pricing formula (equation (13) for the QGTM and equation (20) for the QMGTM).

Figure 2: Distributions for 10-year yields over the subsequent six months as of December 2010

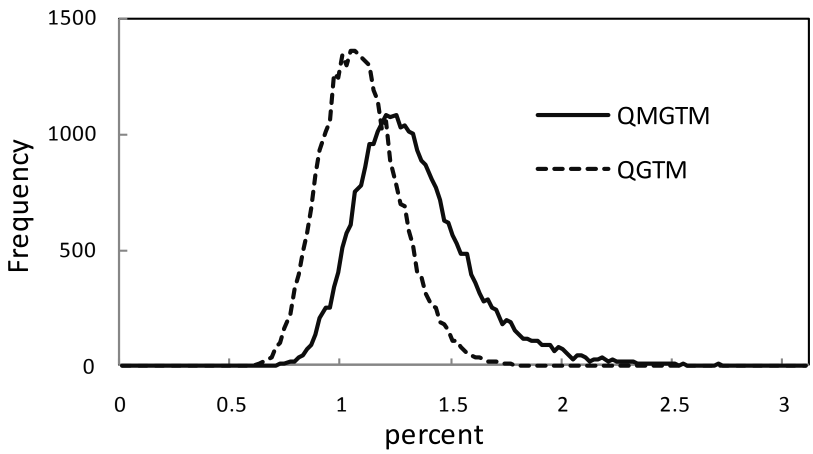
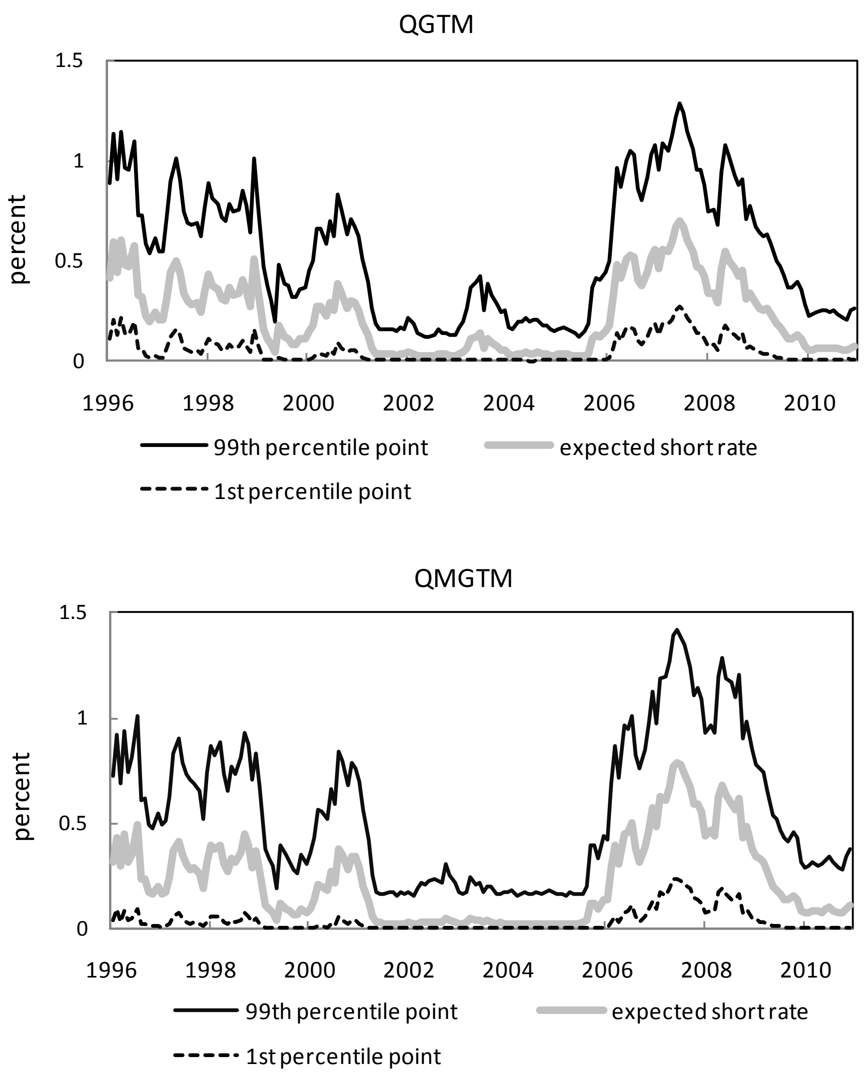


Figure 2 displays probability distributions under $\mathbf{P}$ for 10-year yields over the next six months under the QGTM and the QMGTM evaluated on December 2010 as histograms. Computing the kurtosis for each distribution, we find that the QGTM's distribution kurtosis is equal to 3.21, while that of QMGTM is equal to 5.09. This implies that the probability distribution under the QMGTM has fatter tails than that of the QGTM. This is consistent with our prediction that the QMGTM captures a fat-tailed feature of changes in yields.

Figure 3 displays the time series not just of expectations, but 1st and 99th percentile points of probability distributions under the physical measure $\mathbf{P}$ for one-month yields over the next six months, based on two models. We focus on the rise in the expected one-month yields in summer 2003 under the QGTM. During this period, prospects for short-term interest rates in the near future did not almost change. This increase in expectation in the QGTM appears strange in light of contemporary market conditions. In contrast, the expectation obtained based on the QMGTM at that time is near-zero. This appears better than the QGTM result. Looking at the QMGTM 99th percentile points from the latter half of 2001 to the first half of 2005, which accounts for most of the period during the time of quantitative easing, we find that it remained nearly steady. This means the potential risks assumed by market participants for the short term interest rate over the next six months did not change during this period.

Figure 3: Expectations and percentile points of distributions for one-month yields over the subsequent six months



4.3 Expected Short Rate and Term Premium

This subsection discusses the modified version of the expectations hypothesis, which incorporates the time-varying term premium. Under this hypothesis, the medium to long-term interest rates are decomposed into the average of expected future short rates under the physical measure $\mathbf{P}$ and the term premium. We decompose two- and 10-year yields into the expected short rates and the term premia under this hypothesis for the QMGTM and the QGTM and investigate the development of JGB yields and the decomposed elements.

Figure 4: Two-year yield decomposition

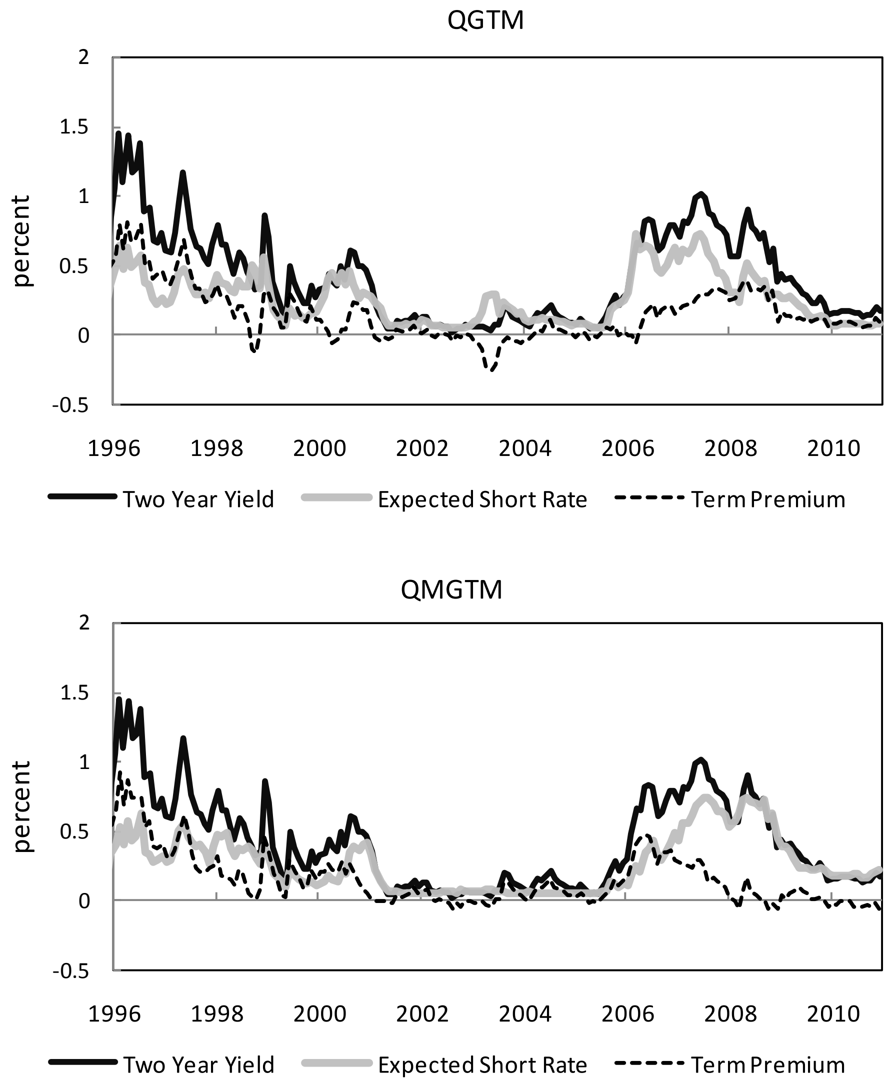


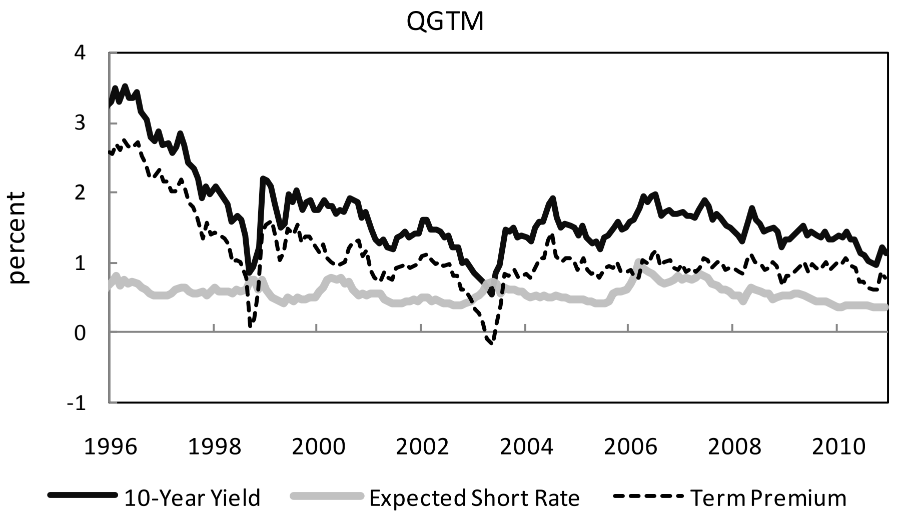
Figure 4 shows the decompositions of two-year yields. One feature of Figure 4 is that the development of the expected future short rates and the term premia differs significantly between the QGTM and the QMGTM from the latter half of 2005 up to the latter half of 2007. Both the expected future short rate and the term premium of the QMGTM rose from July 2005 to May 2006. In contrast, the expected future short rate of the QGTM rose during this period, while the term premium of the QGTM remained near-zero.

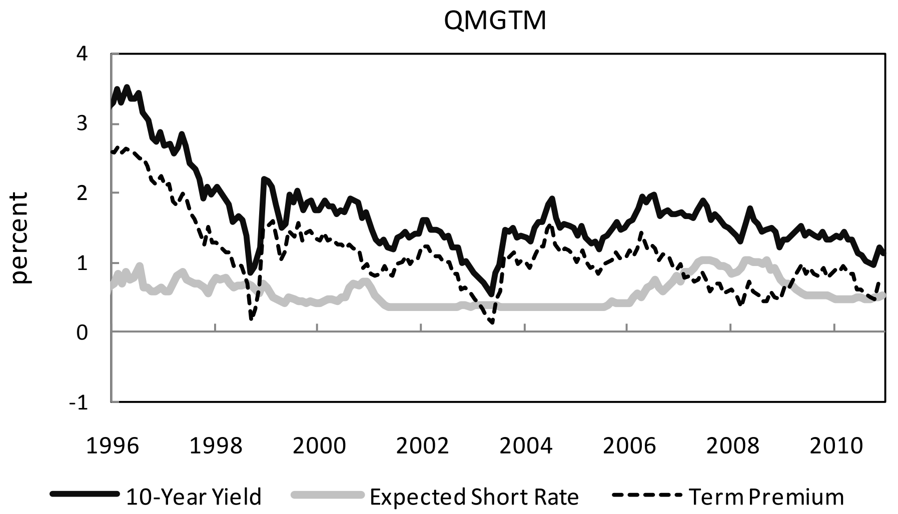
From June 2006 until July 2007, the expected future short rate of the QMGTM rose while that of the QGTM moved within a narrow range. In particular, the expected future short rate for the QGTM in July 2007 was nearly at the same level as the estimate in March 2006, when the quantitative easing policy ended. This appears strange, since the Bank of Japan raised the policy rate twice during this period; many market participants believed the Bank of Japan would raise the policy rate at a pace paralleling economic recovery over the next one or two years. Decomposition using the QMGTM would therefore appear more valid than the QGTM. Regarding the term premium, we can interpret the development of the term premium from July 2005 to July 2007 as follows: First, the increase in the QMGTM term premium from July 2005 to May 2006 could be attributable to the increase in uncertainties concerning the development of the short term interest rate after the end of quantitative easing. The decrease in the term premium from June 2006 to July 2007 may imply that uncertainties about the development of the short term interest rate in fact declined after the end of the quantitative easing and zero interest rate policies.

Figure 5 shows decompositions of 10-year yields under the modified version of expectations hypothesis for the QGTM and the QMGTM. In summer 2003, medium- to long-term interest rates rose sharply as a result of one-side selling by the banks which breached the internal risk limits in terms of value at risk (VaR) due to the increase in the JGB volatility. This episode is called “the VaR shock”. Figure 5 shows that 10-year yields changes before and after the VaR shock was caused by the development of term premium under both two models. In those days, economic conditions showed little change; therefore, the above mentioned development of term premium appears consistent with the reality.

The developments in the first half of 2009 (Figure 5) point to a key difference between the QGTM and the QMGTM. In the QGTM, the expected future short rate and the term premium stayed nearly constant. In contrast, in the QMGTM, the expected future short rate rose, while the term premium declined. According to Bank of Japan [2009], the sharp deterioration in economic conditions applied downward pressure, while the supply and demand conditions of JGBs reflecting the growing fiscal deficit applied an upward pressure on JGB long term interest rates. The decrease in the expected future short rate in the QMGTM can be explained by the sharp deterioration in economic conditions. The increase in the term premium in the QMGTM may reflect the imbalance between supply and demand conditions of JGBs.

Figure 5: 10-year yield decomposition





5. Conclusion

Our paper proposes a new term structure model that allows interest rates to remain non-negative and captures the fat-tailed feature of changes in interest rates. In our model, the short rate is formulated as the quadratic form of the state variables whose future probability distributions follow a mixture of normal distributions under both risk neutral and physical measures in a discrete time setting. In this framework, we derived the approximated pricing formula for zero coupon bond prices by a log-linear approximation. We then estimated state variables and model parameters using the unscented Kalman filter joint with the maximum likelihood method.

Our estimation results showed the QMGTM has the greatest likelihood. It also provided the best AIC and BIC of all three models. The QMGTM also accounted for the smallest mean squared error for each observed maturity. The estimates of the latent state variables and model parameters allow us to compute decompositions of JGB interest rates into expected future short rates and term premia and to simulate probability distributions of future interest rates under the physical measure.

Our model appears to explain actual trends better than other models. For example, the decomposition of a two-year yield based on our model showed a continually increasing expected short rate from June 2006 to July 2007. The expected short rate under the QGTM did not change during this time. During this period, the Bank of Japan raised the policy rate twice. Our model's expected short rate throughout this period appears at least to reflect actual market views. Several other examples discussed in this paper indicate that our model appears to capture actual market conditions more accurately than the other models.

With regard to the empirical aspects of past JGB interest rates, estimation results suggest the following observations. First, both the 99th percentile point and the expectation of the short rate over a six-month period remained steady during most of the quantitative easing policy period from 2001 to 2006. This implies the policy duration was highly effective in stabilizing market expectations regarding the future short rate. Second, much of the development of the JGB long term interest rate in summer 2003 is explained by changes in its term premium, not by changes in its expected short rate. Third, the term premium with a two year yield rose from June 2005 to May 2006 and declined from June 2006 through summer 2007. The rise may reflect uncertainties regarding future policy rate developments after the end of the period of quantitative easing. The decline in the term premium from June 2006 may be attributable to receding uncertainties regarding the future policy rate against the backdrop of the actual end of quantitative easing and zero interest rate policies. Finally, in the first half of 2009, the expected short rate of a 10-year yield declined, while its term premium increased. This may be because the sharp deterioration in economic conditions applied downward pressure on the expected short rate while the deteriorating supply and demand conditions of JGB applied an upward pressure on the term premium.

Although our study focused on the JGB yield, our model can be applied to yields in other countries. In particular, our model should be highly effective for examining yields in countries with low interest rates. The United States and European countries have maintained an expansionary monetary policy since the emergence of the financial crisis in 2007, and interest rates in these countries remained low. One possible direction for future research is to explore the effects of unconventional monetary policies in recent years, based on our model, by simulating probability distributions of the future interest rate under the physical measure or by decomposing the interest rate into the expected short rate and term premium.

References

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, **15**(1), 2002, pp.243–288.

Bank of Japan, "Financial Markets Report –Developments in the first half of 2009," BOJ Reports and Research Papers, 2009.

Bansal, Ravi and Hao Zhou, "Term Structure of Interest Rates with Regime Shifts," *Journal of Finance*, **57**(5), 2002, pp.1997–2043.

Christoffersen, Peter F., Kris Jacobs, Lofti Karoui, and Karim Mimouni, "Nonlinear Filtering in Affine Term Structure Models: Evidence from the Term Structure of Swap Rates," working paper, 2008.

Dai, Qiang and Kenneth J. Singleton, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, **55**(5), 2000, pp.1943–1978.

Duffee, Gregory, "Term Premia and Interest Rates Forecasts in Affine Models," *Journal of Finance*, **57**(1), 2002, pp.405–443.

Duffie, Darell and Rui Kan, "A Yield Factor Model of Interest Rates," *Mathematical Finance*, **6**(4), 1996, pp.379–406.

Julier, J. Simon and Jeffrey K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, 1997.

Kim, Don H. and Kenneth J. Singleton, "Term Structure Models with Zero Bound: an Empirical Investigation of Japanese Yields," working paper, 2011.

Leippold, Markus and Liuren Wu, "Design and Estimation of Quadratic Term Structure Models," *European Finance Review*, **7**(1), 2003, pp.47–73.

Lemke, Wolfgang, "Term Structure Modeling and Estimation in a State Space Framework," *Lecture Notes in Economics and Mathematical Systems*, 565, Springer, 2006.

McCulloch, J. Huston, "The Tax-adjusted Yield Curve," *Journal of Finance*, **30**(3), 1975, pp.811–830.

Nyholm, Ken and Rositsa Vidova-Koleva, "Nelson –Siegel Affine and Quadratic Yield Curve Specifications Which One is Better at Forecasting?," European Central Bank Working Paper, 1205, 2010.

Appendix 1

We derive the approximated pricing formula for the zero coupon bonds based on the QMGTM.

The model is represented from equations (12), (18), and (19), as follows:

$$r_t = X_t' A X_t \tag{A-1}$$

$$X_{t+1} = a^{\mathbf{Q}}_{s_{t+1}} + \Phi^{\mathbf{Q}}_{s_{t+1}} X_t + \Sigma_{s_{t+1}} \varepsilon^{\mathbf{Q}}_{t+1},\ \varepsilon^{\mathbf{Q}}_{t+1} \sim i.i.d.N(0, I), \tag{A-2}$$

$$\begin{aligned} \mathbf{Q}(s_{t+1} = 1 \mid I_t) &= p, \\ \mathbf{Q}(s_{t+1} = 2 \mid I_t) &= 1 - p. \end{aligned} \tag{A-3}$$

Under the no-arbitrage condition, the following formula must hold:

$$P_t^n = E_t^{\mathbf{Q}}[e^{-r_t} P_{t+1}^{n-1}]. \tag{A-4}$$

From equation (A-1)-(A-4), we want to derive the closed pricing formula for P_t^n; however, it is not possible to do that. Thus, we derive the approximated pricing formula for P_t^n.

The price for P_t^n is assumed to have the form $P_t^n = \exp(X_t' A_n X_t + b_n' X_t + c_n)$. Substituting this into equation (A-4), we obtain the following equation:

$$\begin{aligned}
&\exp(X_t' A_n X_t + b_n' X_t + c_n) = E_t^{\mathbf{Q}}[\exp(-X_t' A X_t + X_{t+1}' A_{n-1} X_{t+1} + b_{n-1}' X_{t+1} + c_{n-1})] \\
&= e^{-X_t' A X_t} \left(\begin{aligned} & p E_t^{\mathbf{Q}} \left[\exp \left(\begin{aligned} &(a_1^{\mathbf{Q}} + \Phi_1^{\mathbf{Q}} X_t + \Sigma_1 \varepsilon^{\mathbf{Q}}_{t+1})' A_{n-1} (a_1^{\mathbf{Q}} + \Phi_1^{\mathbf{Q}} X_t + \Sigma_1 \varepsilon^{\mathbf{Q}}_{t+1})' \\ &+ b_{n-1}'(a_1^{\mathbf{Q}} + \Phi_1^{\mathbf{Q}} X_t + \Sigma_1 \varepsilon^{\mathbf{Q}}_{t+1}) + c_{n-1} \end{aligned} \right) \right] \\
& + (1-p) E_t^{\mathbf{Q}} \left[\exp \left(\begin{aligned} &(a_2^{\mathbf{Q}} + \Phi_2^{\mathbf{Q}} X_t + \Sigma_2 \varepsilon^{\mathbf{Q}}_{t+1})' A_{n-1} (a_2^{\mathbf{Q}} + \Phi_2^{\mathbf{Q}} X_t + \Sigma_2 \varepsilon^{\mathbf{Q}}_{t+1})' \\ &+ b_{n-1}'(a_2^{\mathbf{Q}} + \Phi_2^{\mathbf{Q}} X_t + \Sigma_2 \varepsilon^{\mathbf{Q}}_{t+1}) + c_{n-1} \end{aligned} \right) \right] \end{aligned} \right) \\
&= e^{-X_t' A X_t} \left(\begin{aligned} & p e^{\left(\begin{aligned} &(a_1^{\mathbf{Q}} + \Phi_1^{\mathbf{Q}} X_t)' A_{n-1} (a_1^{\mathbf{Q}} + \Phi_1^{\mathbf{Q}} X_t) \\ &+ b_{n-1}'(a_1^{\mathbf{Q}} + \Phi_1^{\mathbf{Q}} X_t) + c_{n-1} \end{aligned} \right)} E_t^{\mathbf{Q}} \left[\exp \left(\begin{aligned} &\left(\begin{aligned} &2(a_1^{\mathbf{Q}} + \Phi_1^{\mathbf{Q}} X_t)' A_{n-1} \\ &+ b_{n-1}' \end{aligned} \right) \Sigma_1 \varepsilon^{\mathbf{Q}}_{t+1} \\ &+ (\Sigma_1 \varepsilon^{\mathbf{Q}}_{t+1})' A_{n-1} \Sigma_1 \varepsilon^{\mathbf{Q}}_{t+1} \end{aligned} \right) \right] \\
& + (1-p) e^{\left(\begin{aligned} &(a_2^{\mathbf{Q}} + \Phi_2^{\mathbf{Q}} X_t)' A_{n-1} (a_2^{\mathbf{Q}} + \Phi_2^{\mathbf{Q}} X_t) \\ &+ b_{n-1}'(a_2^{\mathbf{Q}} + \Phi_2^{\mathbf{Q}} X_t) + c_{n-1} \end{aligned} \right)} E_t^{\mathbf{Q}} \left[\exp \left(\begin{aligned} &\left(\begin{aligned} &2(a_2^{\mathbf{Q}} + \Phi_2^{\mathbf{Q}} X_t)' A_{n-1} \\ &+ b_{n-1}' \end{aligned} \right) \Sigma_2 \varepsilon^{\mathbf{Q}}_{t+1} \\ &+ (\Sigma_2 \varepsilon^{\mathbf{Q}}_{t+1})' A_{n-1} \Sigma_2 \varepsilon^{\mathbf{Q}}_{t+1} \end{aligned} \right) \right] \end{aligned} \right).
\end{aligned} \tag{A-5}$$

The following formula is helpful in resolving the right-hand side of equation (A-5):

$$E_t^{\mathbf{Q}}[\exp(\varepsilon_{t+1}^{\mathbf{Q}}{}'\Sigma' K\varepsilon_{t+1}^{\mathbf{Q}}\Sigma + d'\Sigma\varepsilon_{t+1}^{\mathbf{Q}})] = \frac{\exp\left\{\frac{1}{2}d'\left[(\Sigma\Sigma')^{-1} - 2K\right]^{-1} d\right\}}{|\Sigma\Sigma'|^{0.5}|(\Sigma\Sigma')^{-1} - 2K|^{0.5}}, \qquad \text{(A-6)}$$

where $\varepsilon_{t+1}^{\mathbf{Q}} \sim N(0, I)$.

Using equation (A-6), equation (A-5) is computed as follows:

$$\exp(X_t{}'A_n X_t + b_n{}'X_t + c_n)$$
$$= e^{-X_t{}'AX_t}\begin{pmatrix} pe^{\begin{pmatrix}(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)'A_{n-1}(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)\\ +b_{n-1}{}'(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)+c_{n-1}\end{pmatrix}} \\ \exp\left(\begin{pmatrix}2(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}'\left[(\Sigma_1\Sigma_1{}')^{-1}-2A_{n-1}\right]^{-1}\begin{pmatrix}2(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}\right) \\ |\Sigma_1\Sigma_1{}'|^{-0.5}|(\Sigma_1\Sigma_1{}')^{-1}-2A_{n-1}|^{-0.5} \\ +(1-p)e^{\begin{pmatrix}(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)'A_{n-1}(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)\\ +b_{n-1}{}'(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)+c_{n-1}\end{pmatrix}} \\ \exp\left(\begin{pmatrix}2(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}'\left[(\Sigma_2\Sigma_2{}')^{-1}-2A_{n-1}\right]^{-1}\begin{pmatrix}2(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}\right) \\ |\Sigma_2\Sigma_2{}'|^{-0.5}|(\Sigma_2\Sigma_2{}')^{-1}-2A_{n-1}|^{-0.5}\end{pmatrix}. \qquad \text{(A-7)}$$

Introducing the log-linear approximation used in Bansal and Zhou [2000] in equation (A-7), we obtain:

$$1 = 1 - X_t{}'(A_n + A)X_t - b_n{}'X_t - c_n + c_{n-1}$$
$$+p\begin{pmatrix}(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)'A_{n-1}(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)+b_{n-1}{}'(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)\\ +\begin{pmatrix}2(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}'\left[(\Sigma_1\Sigma_1{}')^{-1}-2A_{n-1}\right]^{-1}\begin{pmatrix}2(a_1^{\mathbf{Q}}+\Phi_1^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}\\ \log(|\Sigma_1\Sigma_1{}'|^{-0.5}|(\Sigma_1\Sigma_1{}')^{-1}-2A_{n-1}|^{-0.5})\end{pmatrix} \qquad \text{(A-8)}$$
$$+(1-p)\begin{pmatrix}(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)'A_{n-1}(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)+b_{n-1}{}'(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)\\ +\begin{pmatrix}2(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}'\left[(\Sigma_2\Sigma_2{}')^{-1}-2A_{n-1}\right]^{-1}\begin{pmatrix}2(a_2^{\mathbf{Q}}+\Phi_2^{\mathbf{Q}}X_t)'A_{n-1}\\ +b_{n-1}{}'\end{pmatrix}\\ \log(|\Sigma_2\Sigma_2{}'|^{-0.5}|(\Sigma_2\Sigma_2{}')^{-1}-2A_{n-1}|^{-0.5})\end{pmatrix}.$$

Hence, we obtain the following recursive equations for coefficients A_n, b_n, c_n:

$$
\begin{aligned}
P_t^n &= \exp(X_t' A_n X_t + b_n' X_t + c_n),\\
A_n &= -A + p\Phi_1^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{1,n-1}^{-1}A_{n-1}')\Phi_1^{\mathbf{Q}}\\
&\quad + (1-p)\,\Phi_2^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{2,n-1}^{-1}A_{n-1}')\Phi_2^{\mathbf{Q}}, A_0 = -A,\\
b_n &= p(\Phi_1^{\mathbf{Q}}{}' b_{n-1} + 2\Phi_1^{\mathbf{Q}}{}' A_{n-1}\Psi_{1,n-1}^{-1}b_{n-1} + 2\Phi_1^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{1,n-1}^{-1}A_{n-1})a_1^{\mathbf{Q}})\\
&\qquad + (1-p)(\Phi_2^{\mathbf{Q}}{}' b_{n-1} + 2\Phi_2^{\mathbf{Q}}{}' A_{n-1}\Psi_{2,n-1}^{-1}b_{n-1}\\
&\qquad + 2\Phi_2^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{2,n-1}^{-1}A_{n-1})a_2^{\mathbf{Q}}), b_0 = \mathbf{0},\\
c_n &= c_{n-1} + p(b_{n-1}a_1^{\mathbf{Q}}{}' - 0.5\log|\Psi_{1,n-1}| - 0.5\log|\Sigma_1\Sigma_1'| + 2a_1^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{1,n-1}^{-1}A_{n-1})a_1^{\mathbf{Q}}\\
&\qquad + 0.5b_{n-1}'\Psi_{1,n-1}^{-1}b_{n-1} + 2b_{n-1}'\Psi_{1,n-1}^{-1}A_{n-1}a_1^{\mathbf{Q}})\\
&\quad + (1-p)(b_{n-1}a_2^{\mathbf{Q}}{}' - 0.5\log|\Psi_{2,n-1}| - 0.5\log|\Sigma_2\Sigma_2'| + 2a_2^{\mathbf{Q}}{}'(A_{n-1} + 2A_{n-1}\Psi_{2,n-1}^{-1}A_{n-1})a_2^{\mathbf{Q}}\\
&\qquad + 0.5b_{n-1}'\Psi_{2,n-1}^{-1}b_{n-1} + 2b_{n-1}'\Psi_{2,n-1}^{-1}A_{n-1}a_2^{\mathbf{Q}}),\ \ c_0 = 0,\\
\Psi_{i,n} &= (\Sigma_i\Sigma_i')^{-1} - 2A_n.
\end{aligned}
\tag{A-9}
$$

Appendix 2

Here we begin by showing estimation results for model parameters. We then show the estimated state variables for each model.

Tables A, B, and C, respectively, give the estimated parameters for the AGTM, QGTM, and the QMGTM.

Table A: Estimated Parameters of the AGTM

$$\Phi^{\mathbf{P}}: \begin{pmatrix} 0.953(2.981\times10^{-5}) & 0 & 0 \\ -0.0832(3.246\times10^{-7}) & 0.945(5.649\times10^{-5}) & 0 \\ -0.0432(3.598\times10^{-5}) & 0.00123(2.229\times10^{-8}) & 0.934(0.00154) \end{pmatrix}$$

$$\delta_0: \quad -0.000121(4.853\times10^{-5})$$

$$\delta_1': \begin{pmatrix} -0.000244(9.146\times10^{-6}) & 0.000325(5.444\times10^{-7}) & -0.000101(9.107\times10^{-6}) \end{pmatrix}$$

$$\lambda_0: \begin{pmatrix} 0.403(0.0718) \\ 1.607(0.122) \\ 1.184(0.0930) \end{pmatrix}$$

$$\Lambda_1: \begin{pmatrix} -0.545(1.405\times10^{-5}) & 0.0516(0.00203) & 0.250(0.000798) \\ -0.233(0.00718) & -0.780(1.607\times10^{-5}) & -0.510(0.0157) \\ -0.981(3.370\times10^{-5}) & -0.206(2.450\times10^{-6}) & -0.323(0.0959) \end{pmatrix}$$

Note: Standard errors are given in parentheses.

Table B: Estimated Parameters of the QGTM

$$a^{\mathbf{P}}: \begin{pmatrix} 0.0378(3.33\times10^{-5}) \\ 0.0438(0.000105) \\ -0.0156(6.04\times10^{-5}) \end{pmatrix}$$

$$\Phi^{\mathbf{P}}: \begin{pmatrix} 0.919(0.000106) & 0 & 0 \\ -0.0820(6.22\times10^{-5}) & 0.972(7.53\times10^{-5}) & 0 \\ -0.0118(7.66\times10^{-6}) & 0.770(2.26\times10^{-5}) & 0.953(9.82\times10^{-5}) \end{pmatrix}$$

$$S: \begin{pmatrix} 1 & 0 & 0 \\ -0.0145(0.0145) & 1 & 0 \\ 3.89(0.0129) & -2.03(0.00385) & 1 \end{pmatrix}$$

$$D: \begin{pmatrix} 3.93\times10^{-7}(5.48\times10^{-8}) & 0 & 0 \\ 0 & 2.56\times10^{-7}(5.17\times10^{-7}) & 0 \\ 0 & 0 & 5.79\times10^{-5}(8.27\times10^{-10}) \end{pmatrix}$$

λ_0:	$\begin{pmatrix} 0.311(0.000103) \\ 0.701(0.000201) \\ 0.334(0.000435) \end{pmatrix}$
Λ_1:	$\begin{pmatrix} -0.791(0.000644) & 0.00209(0.000149) & -0.0494(4.12\times10^{-5}) \\ -0.998(0.000823) & -0.0932(0.000251) & 0.148(4.35\times10^{-5}) \\ -1.191(0.00149) & 0.228(7.04\times10^{-5}) & -0.386(2.29\times10^{-6}) \end{pmatrix}$

Note: Standard errors are given in parentheses.

Table C: Estimated Parameters of the QMGTM

$a^{\mathbf{P}}$	$\begin{pmatrix} 0.0285(0.00124) \\ 0.0207(0.000266) \\ 0.0395(0.000199) \end{pmatrix}$
$\Phi^{\mathbf{P}}$	$\begin{pmatrix} 0.929(7.33\times10^{-4}) & 0 & 0 \\ -0.106(0.00147) & 0.960(0.00396) & 0 \\ 0.0159(1.18\times10^{-5}) & 0.0646(0.00117) & 0.979(0.00643) \end{pmatrix}$
S	$\begin{pmatrix} 1 & 0 & 0 \\ 1.489\,(0.0167) & 1 & 0 \\ 2.237\,(0.706) & 2.511\,(0.0296) & 1 \end{pmatrix}$
D	$\begin{pmatrix} 1.465\times10^{-6}\ (1.114\times10^{-9}) & 0 & 0 \\ 0 & 2.593\times10^{-6}\ (6.315\times10^{-6}) & 0 \\ 0 & 0 & 4.309\times10^{-6}\ (1.450\times10^{-9}) \end{pmatrix}$
λ_0	$\begin{pmatrix} 0.281(0.0800) \\ -0.0681(0.00312) \\ 0.146(0.000838) \end{pmatrix}$
Λ_1	$\begin{pmatrix} -0.670(0.00301) & 0.0754(0.00675) & -0.00655(0.000109) \\ -0.841(0.00801) & -0.131(0.0238) & 0.134(0.0203) \\ -0.358(0.0312) & 0.140(0.0181) & -0.165(0.0183) \end{pmatrix}$
μ_1	$\begin{pmatrix} -0.0177\ (0.000140) \\ -0.00588(0.000352) \\ 0.0112\ (0.000282) \end{pmatrix}$
Σ_2	$\begin{pmatrix} 0.105(0.0110) & 0 & 0 \\ -0.083(0.00905) & 0.673(0.00755) & 0 \\ -0.067(0.000205) & -0.243(0.00861) & 0.276(0.00640) \end{pmatrix}$

Note: Standard errors are given in parentheses.

Figure A gives the estimated state variables for each model.

Figure A: Estimated State Variables $E_t^{\mathbf{P}}[X_t]$