Bayesian Estimation of Wishart Term Structure Models *

Ming Lin, Linlin Niu and Chen Zhang†

We propose a novel method for estimating the parameters and latent factors of Wishart term structure models (WTSM) under the Bayesian framework. The lack of a closed-form transition density for the Wishart autoregressive (WAR) latent factors makes the maximum likelihood method or the generalised method of moments infeasible for joint inference on model parameters and states. To address this difficulty, we transform the WAR process into the outer product of a first-order vector autoregressive (VAR(1)) process with a Gaussian conditional density. A procedure combining Markov chain Monte Carlo (MCMC) and sequential Monte Carlo (SMC) is then employed for Bayesian inference. We apply the WTSM and the proposed estimation method to analyze the US yield curves over the period from 1961Q2 to 2025Q4. The model fits the data remarkably well and identifies three economically meaningful risk factors governing yield dynamics via time-varying cross-factor correlation. The estimated specification successfully reproduces short-rate dynamics across the zero lower bound episode without resorting to shadow-rate modeling, and generates empirically realistic time-varying correlations across yields of distinct maturities.

Primary G13, G17; secondary C11.
Keywords and phrases: Bayesian Inference, Sequential Monte Carlo, Wishart Term Structure Model.

1. INTRODUCTION

Time-varying volatility and changing correlations are pervasive features of individual asset returns as well as aggregate financial prices such as the Treasury yield curve. Economic theory suggests that aggregate risks should be priced in asset returns; consequently, the time-varying volatility and covolatility of interest rate movements should be reflected in the dynamic pricing of yield curves across maturities. Within the framework of arbitrage-free term structure models—pioneered by [34] and [10], several strands have emerged to incorporate stochastic volatility (SV) under consistent pricing conditions. These models are often divided into spanned and unspanned volatility classes, depending on whether the SV risk factors directly enter the cross-sectional pricing of yields.

Spanned volatility models date back to the single-factor square-root process of the CIR model [10]. In the generalised affine term structure framework [15, 12], SV is introduced via square-root processes alongside Gaussian autoregressive factors. However, ensuring positive definiteness of the stochastic variance-covariance matrix imposes restrictive parameter constraints, creating a trade-off between modelling conditional volatility flexibly and capturing conditional correlation, which often becomes negative between factors [12]. Quadratic term structure models (QTSMs) [2] partly address this deficiency by allowing multivariate heteroscedastic volatility and flexible factor correlations, but they tend to underestimate conditional yield volatility (especially at short to medium maturities) and produce correlation patterns that are much smoother than those observed empirically [2, 1].

Unsatisfactory performance of spanned models in fitting both first and second moments of yields led to the development of unspanned volatility models, where the SV factors do not directly price yields but interact with the pricing factors over time [8, 9]. Although unspanned models offer greater flexibility for the second moment, their overall fit often does not improve upon affine models [11]. Moreover, [9] find that unspanned volatility extracted from such models is negatively correlated with short-rate volatility estimated from reduced-form GARCH models, and these models cannot quantify the effect of SV on cross-sectional yields, a theoretically and practically important issue.

As a continuation of efforts to improve the theoretical properties of spanned SV models under no-arbitrage, [21] propose a matrix-valued affine model based on the Wishart autoregressive (WAR) process. The WAR process guarantees positive definiteness of the stochastic variance-covariance matrix at every point in time, has a well-defined marginal density, and admits closed-form forecasts [19, 20]. This class, known as the Wishart term structure model (WTSM), extends QTSMs and permits an economic interpretation of time-varying variances and covariances across yields of distinct maturities via the stochastic discount factor. [4, 28, 7]. An additional appealing feature of

*We thank seminar participants at Wang Yanan Institute for Studies in Economics (WISE), Xiamen University for useful comments, WISE Information Center for computational service, and support from National Science Foundation of China (72273118, 72033008).
†Corresponding author.

WTSMs is that, through time-varying correlation between risk factors, they can generate both positive and negative excess bond returns—a sign-switching property that is not supported by completely affine term structure models unless complex market prices of risk are introduced [6].

Despite these theoretical advantages, most empirical applications of WTSMs have assumed that the underlying volatility process is known ex ante. For instance, [20] and [5] compute realised volatility matrices from high-frequency data and then apply the method of moments. [7] use an unscented Kalman filter with likelihood approximation and augment the information with realised yield volatilities, but this approach does not guarantee consistency. In a related strand, [23] develop a Riesz state-space model for observable realized covariance matrices, also using Bayesian MCMC. However, their approach relies on high-frequency data and an observable proxy for the covariance matrix, which is not available in many settings of macro-finance studies where the volatility-covolatility process must be treated as latent.

The major obstacle to inference on the latent WAR process is that its conditional distribution follows a non-centered Wishart distribution, whose density has no closed-form expression [20]. This complexity creates a severe computational burden. In this paper, we bridge this gap by developing an innovative Bayesian estimation procedure based on Markov chain Monte Carlo (MCMC). Our method is not only useful for studying WTSMs of yield curves but also applicable to derivative pricing problems that rely on WAR processes [17, 18, 20, 5].

Specifically, we focus on the WAR(1) process with integer degrees of freedom. To circumvent the lack of a closed-form transition density, we exploit an alternative representation that expresses the WAR process as the outer product of a mean-zero vector autoregressive process with Gaussian innovations. This representation yields a Gaussian state transition, enabling efficient MCMC sampling. We adopt a two-step procedure: first, conditional on a given degrees of freedom, we estimate all remaining parameters and the latent variance-covariance process simultaneously using an MCMC algorithm; second, we select the optimal degrees of freedom using the deviance information criterion (DIC) evaluated via sequential Monte Carlo (SMC).

Using this estimation procedure, we apply a 2-by-2 WTSM model with three distinct variance and covariance elements to span the cross-section of US quarterly yields from 1961Q2 to 2025Q4. From the estimated loadings and extracted latent factors, we identify three economically meaningful factors: a long-term risk factor (Lt) primarily affecting long-term yields, a short-term risk factor (St) capturing short-rate dynamics, and a correlation risk factor (Cr) representing the time-varying correlation between Lt and St.

Our empirical results reveal several interesting features consistent with observed yield behaviour:

- Economically interpretable latent factors – We extract long-term, short-term, and correlation risk factors with intuitive maturity loadings consistent with standard yield curve behavior.
- Zero lower bound (ZLB) fit – The WTSM replicates empirical short-rate movements throughout the zero lower bound period with no need for supplementary shadow-rate assumptions.
- Realistic cross-yield correlation – The model produces time-varying cross-maturity yield correlations matching key post-crisis and recession-era empirical patterns.

The remainder of this paper is organized as follows. Section 2 presents the Wishart term structure model and discusses the estimation challenges. Section 3 details the Bayesian inference procedure, including the MCMC algorithm and the SMC-based DIC for selecting the degrees of freedom. Section 4 illustrates the method's performance through simulation studies. Section 5 applies the model to the US yield curve, interprets the extracted factors, and documents the model's ability to capture short rate patterns including the ZLB phenomenon, and time-varying cross-maturity yield correlations. Section 6 concludes.

2. THE WISHART TERM STRUCTURE MODEL

2.1 Model Setup

Following [21], we assume that the stochastic discount factor (SDF) from period t to $t+1$ is

$$D_{t,t+1} = \exp\left\{d + Tr\big(C\Omega_{t+1}\big)\right\}, \tag{1}$$

where $Tr(\cdot)$ denotes the matrix trace, d is a scalar parameter, C is an $n \times n$ symmetric matrix, and Ω_{t+1} is a latent volatility matrix that represents the latent risk factors at time $t+1$. We further assume that the matrix series $\{\Omega_t\}$ follows a first-order Wishart autoregressive process (WAR(1)), defined through its conditional Laplace transform as follows [20]:

$$\begin{aligned}\Psi_t(\Gamma) :=& E\left[\exp\left\{Tr(\Gamma\Omega_{t+1})\right\} \mid \Omega_1, \cdots \Omega_t\right]\\ =& \frac{\exp\left\{Tr\left[M'\Gamma(I_n - 2\Sigma\Gamma)^{-1}M\Omega_t\right]\right\}}{\left[\det(I_n - 2\Sigma\Gamma)\right]^{K/2}},\end{aligned} \tag{2}$$

where I_n denotes the $n \times n$ identity matrix, Γ is an $n \times n$ matrix such that $I_n - 2\Sigma\Gamma$ is positive definite. Here, M is an $n \times n$ autoregressive coefficient matrix, Σ is an $n \times n$ symmetric positive-definite innovation covariance matrix, and $K \geq n$ is the degrees of freedom of the WAR(1) process. The WAR(1) process ensures that Ω_t stays positive definite, which is essential for a well-defined variance-covariance structure of the underlying risk factors.

Let $P_{t,\tau}$ be the price of a zero-coupon bond with payoff 1 and maturity τ at time t. Under the SDF $D_{t,t+1}$ defined in (1) and the no-arbitrage condition, bond prices are exponentially affine in Ω_t. Specifically, we have

$$P_{t,0} = 1 = \exp\left\{A_0 + Tr(B_0\Omega_t)\right\},$$

where A_0 is an $n \times 1$ zero vector and B_0 is an $n \times n$ zero matrix, and for $\tau = 1, 2, \cdots$, we can recursively obtain that

$$\begin{aligned}
&P_{t,\tau}\\
&= E\Big(D_{t,t+1}P_{t+1,\tau-1}\,|\,\Omega_1, \cdots \Omega_t\Big)\\
&= E\Big[\exp\{d + Tr(C\Omega_{t+1})\} \cdot \exp\{A_{\tau-1} + Tr(B_{\tau-1}\Omega_{t+1})\}\,|\,\Omega_1, \cdots \Omega_t\Big]\\
&= E\Big[\exp\Big\{d + A_{\tau-1} + Tr\Big((C + B_{\tau-1})\Omega_{t+1}\Big)\Big\}\,|\,\Omega_1, \cdots \Omega_t\Big]\\
&:= \exp\{A_\tau + Tr(B_\tau\Omega_t)\},
\end{aligned}$$

where

$$A_\tau = d + A_{\tau-1} - \frac{K}{2}\log\left[\det(I_n - 2\Sigma(C + B_{\tau-1}))\right] \tag{3}$$

and

$$B_\tau = M'(C + B_{\tau-1})[I_n - 2\Sigma(C + B_{\tau-1})]^{-1}M \tag{4}$$

according to Equation (2). Given $P_{t,\tau}$ and accounting for pricing errors, the zero-coupon bond yield of maturity τ at time t becomes

$$\begin{aligned}
y_{t,\tau} &= -\frac{1}{\tau}\log P_{t,\tau} + \varepsilon_{t,\tau} &(5)\\
&= -\frac{1}{\tau}\left\{A_\tau + \mathrm{Tr}\big(B_\tau\Omega_t\big)\right\} + \varepsilon_{t,\tau}, &(6)
\end{aligned}$$

where $\varepsilon_{t,\tau} \sim N(0, \sigma^2)$ denotes the pricing error.

2.2 Challenges in Estimating the WTSM

If the matrix factors $\{\Omega_t\}$ were observable, the parameters M, Σ and K of the WAR(1) model can be estimated by the method of moments [20]. However, since $\{\Omega_t\}$ is typically unobservable and we need to rely on the observable zero-coupon bond yields $\{y_{t,\tau}\}$ for inference on the model, the method of moments estimator is infeasible in this circumstance. Consequently, the WTSM can be formulated as a state-space model:

(7)

$$\begin{aligned}
\text{Observation equation:} \quad & y_{t,\tau} = -\frac{1}{\tau}\left\{A_\tau + \mathrm{Tr}[B_\tau\Omega_t]\right\} + \varepsilon_{t,\tau},\\
\text{State equation:} \quad & \{\Omega_t\} \sim \text{WAR}(1) \text{ process.}
\end{aligned}$$

We are interested in estimating the model parameters and the latent factors $\{\Omega_t, t = 1, 2, \cdots\}$.

Common approaches for estimating state-space models include Markov chain Monte Carlo (MCMC) [29], simulated maximum likelihood [16, 31], and nonlinear filtering [7]. MCMC is attractive because it simultaneously samples parameters and latent states from the full posterior distribution, thereby avoiding the bias of approximate filters. However, the lack of a closed-form transition density for the WAR process has precluded the use of MCMC for inference on the WTSM in existing studies.

We circumvent this difficulty by focusing on the case of integer degrees of freedom K and exploiting an equivalent representation that transforms the WAR(1) process into a Gaussian state-space form. This idea is conceptually related to recent work by [23] on matrix-variate state-space models for observable realized covariance measures, but our setting is fundamentally different because $\{\Omega_t\}$ is latent and no high-frequency proxy is available.

2.3 WTSM with Integer Degrees of Freedom

When the degrees of freedom K is an integer, the WTSM in (7) can be represented as[20, 21]:

(8)

$$\begin{aligned}
\text{Observation equation:} \quad & y_{t,\tau} = -\frac{1}{\tau}\left\{A_\tau + Tr[B_\tau\Omega_t]\right\} + \varepsilon_{t,\tau},\\
\text{State equations:} \quad & \Omega_t = Z_tZ_t',\\
& Z_t = MZ_{t-1} + \Xi_t,
\end{aligned}$$

where A_τ and B_τ are given in (3) and (4), respectively, the pricing errors $\varepsilon_{t,\tau} \sim N(0, \sigma^2)$ are independent across time t and maturity τ, $Z_t = (z_{t,1}, \cdots, z_{t,K})$ and $\Xi_t = (\xi_{t,1}, \cdots, \xi_{t,K})$ are $n \times K$ matrices, $\{\xi_{t,1}\}, \cdots, \{\xi_{t,K}\}$ are independent n-dimensional Gaussian white noises with $\xi_{t,k} \sim N(0, \Sigma)$, $k = 1, \cdots, K$, and M, Σ and K are the parameters in the WAR(1) process. In this new representation, $\{\Omega_t\}$ is the outer product of a latent vector autoregressive Gaussian process $\{Z_t\}$. Conditional on Z_{t-1}, the transition density of Z_t has a closed form, which resolves the main computational obstacle to implementing MCMC. We also let $Y_t = (y_{t,\tau_1}, \cdots, y_{t,\tau_N})'$ denote the vector of yields $y_{t,\tau}$ observed at time t for different maturities.

In the above model, notice that multiplying Z_t by an orthogonal matrix leaves $\Omega_t = Z_tZ_t'$ unchanged, so Σ is not identifiable up to a multiplicative orthogonal matrix. Furthermore, the conditional likelihood of $y_{t,\tau}$ depends on Σ through $C\Omega_{t+1}$ and $C\Sigma_{t+1}$, which induces an identification problem for Σ and C. To address the identification issue, we follow [23] and fix the innovation covariance matrix Σ to the identity matrix I_n.

3. BAYESIAN INFERENCE

In this section, for model (8) with observed yield vectors $Y_1, \cdots, Y_T$, we consider the estimation of the integer degrees of freedom K, the remaining unfixed model parameters $\Theta := (M, d, C, \sigma^2)$, and the latent risk factors $\{\Omega_t = Z_tZ_t', t = 1, \cdots, T\}$ under the Bayesian framework.

3.1 Inference on Θ and the Latent States Given K

We first consider Bayesian inference on $\Theta = (M, d, C, \sigma^2)$ and $\Omega_t = Z_tZ_t'$, $t = 1, \cdots, T$, when the integer degrees of freedom K is given. In this subsection, we suppress the dependence on K in the model whenever it will not cause confusion. The joint posterior distribution of $(Z_{1:T}, \Theta)$ in model (8) is

$$
\begin{aligned}
P(Z_{1:T}, \Theta \,|\, Y_{1:T}) \quad &\propto \quad P(Z_{1:T}, Y_{1:T}, M, d, C, \sigma^2) \\
&= \quad P(Z_1, M, d, C, \sigma^2) P(Y_1 \,|\, Z_1, M, d, C, \sigma^2) \\
&\quad \times \prod_{t=2}^{T} P(Z_t \,|\, Z_{t-1}, M) P(Y_t \,|\, Z_t, M, d, C, \sigma^2),
\end{aligned}
$$

where $Z_{1:T} = (Z_1, \cdots, Z_T)$ and $Y_{1:T} = (Y_1, \cdots, Y_T)$, $P(Z_1, M, d, C, \sigma^2)$ is the joint prior distribution of the initial state Z_1 and the parameters (M, d, C, σ^2), $P(Z_t \,|\, Z_{t-1}, M)$ and $P(Y_t \,|\, Z_t, M, d, C, \sigma^2)$ are defined by the state equation and observation equation in (8), respectively.

We use the MCMC method to draw samples $\left\{\left(Z_{1:T}^{(s)}, \Theta^{(s)}\right)\right\}_{s=1}^{S}$ from the posterior distribution $P(Z_{1:T}, \Theta \,|\, Y_{1:T})$. Then each parameter, say θ, can be estimated by

$$\hat{\theta} = \frac{1}{S} \sum_{s=1}^{S} \theta^{(s)},$$

and the latent factor Ω_t, $t = 1, \cdots, T$, can be estimated by

$$\hat{\Omega}_t = \frac{1}{S} \sum_{s=1}^{S} Z_t^{(s)} (Z_t^{(s)})'.$$

A more detailed discussion of the MCMC method can be found in [30] and [25]. In the following, we describe some implementation details of the MCMC procedure.

3.1.1 Prior Distributions

For simplicity, we choose the joint prior distribution

$$P(Z_1, M, d, C, \sigma^2) = P(Z_1)\, P(M)\, P(d)\, P(C)\, P(\sigma^2),$$

that is, Z_1, M, d, C, and σ^2 are independent in the prior distribution.

For ease of representation, we rearrange the n^2 entries in M into a vector denoted as $\overrightarrow{M}$. Similarly, we arrange the $(n^2 + n)/2$ distinct entries in C into a vector $\overrightarrow{C}$. Without extra prior information, we adopt non-informative priors for there parameters. Specifically, we set

$$P(\overrightarrow{M}) \propto 1, \quad P(d) \propto 1, \quad P(\overrightarrow{C}) \propto 1, \tag{9}$$

and

$$P(\sigma^2) \propto 1/\sigma^2. \tag{10}$$

For the prior distribution $P(Z_1)$ of $Z_1 = (z_{1,1}, \cdots, z_{1,K})$, we let $z_{1,1}, \cdots, z_{1,K}$ be independent and identically distributed as $N\left(\mu_{z,prior}, \Sigma_{z,prior}\right)$, where $\mu_{z,prior}$ is a zero vector and $\Sigma_{z,prior}$ is set to the identity matrix.

3.1.2 Metropolis-within-Gibbs Sampling

We draw samples $\left\{\left(Z_{1:T}^{(s)}, \Theta^{(s)}\right)\right\}_{s=1}^{S}$ from the posterior distribution $P(Z_{1:T}, \Theta \,|\, Y_{1:T})$ by iteratively updating each component of Θ and $Z_{1:T}$ using the Metropolis-within-Gibbs sampling method. Specifically, given the initial sample $\left(Z_{1:T}^{(1)}, \Theta^{(1)}\right)$, the algorithmic steps to recursively generate $\left(Z_{1:T}^{(s)}, \Theta^{(s)}\right)$, $s = 2, 3, \cdots$, are described as follows.

1. Update $Z_{1:T}$: For $t = 1, \cdots, T$, generate $Z_t^{(s)}$ as follows.
   (a) Draw $z_{t,k}^{*}$, $k = 1, \cdots, K$, from the trial distribution

   $$Q(Z_t^{*} \,|\, Z_{1:t-1}^{(s)}, Z_{t+1:T}^{(s-1)}, \overrightarrow{M}^{(s-1)}, \Sigma^{(s-1)}) \sim N\left(A_{Z,t,k}^{-1} B_{Z,t,k}, A_{Z,t,k}^{-1}\right), \tag{11}$$

   where

   $$B_{Z,t,k} = \begin{cases} (M^{(s)})'\Sigma^{-1} z_{k,t+1}^{(s-1)} + \Sigma_{z,prior}^{-1}\mu_{z,prior}, & t = 1; \\ (M^{(s)})'\Sigma^{-1} z_{k,t+1}^{(s-1)} + \Sigma^{-1}(M^{(s)} z_{k,t-1}^{(s)}), & 1 < t < T; \\ \Sigma^{-1}(M^{(s)} z_{k,t-1}^{(s)}), & t = T \end{cases} \tag{12}$$

   and

   $$A_{Z,t,k} = \begin{cases} (M^{(s)})'\Sigma^{-1} M^{(s)} + \Sigma_{z,prior}^{-1}, & t = 1; \\ (M^{(s)})'\Sigma^{-1} M^{(s)} + \Sigma^{-1}, & 1 < t < T; \\ \Sigma^{-1}, & t = T. \end{cases} \tag{13}$$

   (b) Let $Z_t^{*} = \left(z_{t,1}^{*}, \cdots, z_{t,K}^{*}\right)$. Accept Z_t^{*} as $Z_t^{(s)}$ with probability

   $$\max\left\{1, \frac{P\left(Y_t \,|\, Z_t^{*}, \overrightarrow{M}^{s-1}, d^{(s-1)}, \overrightarrow{C}^{(s-1)}, (\sigma^2)^{(s-1)}\right)}{P\left(Y_t \,|\, Z_t^{(s)}, \overrightarrow{M}^{(s-1)}, d^{(s-1)}, \overrightarrow{C}^{(s-1)}, (\sigma^2)^{(s-1)}\right)}\right\}, \tag{14}$$

   and let $Z_t^{(s)} = Z_t^{(s-1)}$ if Z_t^{*} is rejected.
2. Draw $\overrightarrow{M}^{*}$ from a random walk proposal

   $$Q(\overrightarrow{M} \,|\, \overrightarrow{M}^{(s-1)}) \sim N(\overrightarrow{M}^{(s-1)}, \sigma_M^2 \cdot I_{n^2}), \tag{15}$$

   where σ_M is a prespecified step size. Accept $\overrightarrow{M}^{*}$ as $\overrightarrow{M}^{(s)}$ with probability

   $$\max\left\{1, \frac{\prod_{t=1}^{T} P\left(Y_t \,|\, Z_t^{(s)}, \overrightarrow{M}^{*}, d^{(s-1)}, \overrightarrow{C}^{(s-1)}, (\sigma^2)^{(s-1)}\right)}{\prod_{t=1}^{T} P\left(Y_t \,|\, Z_t^{(s)}, \overrightarrow{M}^{(s-1)}, d^{(s-1)}, \overrightarrow{C}^{(s-1)}, (\sigma^2)^{(s-1)}\right)} \times \frac{\prod_{t=2}^{T} P(Z_t^{(s)} | Z_{t-1}^{(s)}, \overrightarrow{M}^{*})}{\prod_{t=2}^{T} P(Z_t^{(s)} | Z_{t-1}^{(s)}, \overrightarrow{M}^{(s-1)})}\right\}, \tag{16}$$

   and let $\overrightarrow{M}^{(s)} = \overrightarrow{M}^{(s-1)}$ if $\overrightarrow{M}^{*}$ is rejected.
3. Draw d^{*} from the trial distribution

   $$Q(d \,|\, d^{(s-1)}) \sim N(d^{(s-1)}, \sigma_d^2). \tag{17}$$

Accept d^* as $d^{(s)}$ with probability

$$\max\left\{1, \frac{\prod_{t=1}^{T} P\big(Y_t \,|\, Z_t^{(s)}, \overrightarrow{M}^{(s)}, d^*, \overrightarrow{C}^{(s-1)}, (\sigma^2)^{(s-1)}\big)}{\prod_{t=1}^{T} P\big(Y_t \,|\, Z_t^{(s)}, \overrightarrow{M}^{(s)}, d^{(s-1)}, \overrightarrow{C}^{(s-1)}, (\sigma^2)^{(s-1)}\big)}\right\}, \tag{18}$$

and let $d^{(s)} = d^{(s-1)}$ if d^* is rejected.

4. Draw $\overrightarrow{C}^*$ from the trial distribution

$$Q(\overrightarrow{C} \,|\, \overrightarrow{C}^{(s-1)}) \sim N(\overrightarrow{C}^{(s-1)}, \sigma_C^2 \cdot I_{(n^2+n)/2}) \tag{19}$$

Accept $\overrightarrow{C}^*$ as $\overrightarrow{C}^{(s)}$ with probability

$$\max\left\{1, \frac{\prod_{t=1}^{T} P\big(Y_t \,|\, Z_t^{(s)}, \overrightarrow{M}^{(s)}, d^{(s)}, \overrightarrow{C}^*, (\sigma^2)^{(s-1)}\big)}{\prod_{t=1}^{T} P\big(Y_t \,|\, Z_t^{(s)}, \overrightarrow{M}^{(s)}, d^{(s)}, \overrightarrow{C}^{(s-1)}, (\sigma^2)^{(s-1)}\big)}\right\}, \tag{20}$$

and let $\overrightarrow{C}^{(s)} = \overrightarrow{C}^{(s-1)}$ if $\overrightarrow{C}^*$ is rejected.

5. Draw $(\sigma^2)^{(s)}$ from the conditional distribution

$$P(\sigma^2 \,|\, d^{(s)}, \overrightarrow{C}^{(s)}, M^{(s)}, Z_{1:T}^{(s)}, Y_{1:T}) \tag{21}$$

$$\sim \text{inverse-Gamma}(\alpha,\ \beta), \tag{22}$$

where $\alpha = TN/2$ and

$$\beta = \sum_{t=1}^{T}\sum_{l=1}^{N}\left(y_{t,\tau_l} + \frac{1}{\tau_l}\left(A_{\tau_l}^{(s)} + \mathrm{Tr}\left[C_{\tau_l}^{(s)}\Omega_t^{(s)}\right]\right)\right)^2 /2$$

Alternatively, we can draw $h \equiv \sigma^{-2}$, which follows a corresponding Gamma distribution.

3.2 Selection of the Degrees of Freedom

In empirical studies, it is important to choose appropriate degrees of freedom, K, for model (8), where the dimension of the state variable Z_t depends on K. [22] proposes the reversible jump MCMC method for Bayesian model determination when the dimension of the model space is not fixed. However, how to choose an efficient jump between state spaces of differing dimensionality remains a challenging problem in general. We consider employing the deviance information criterion (DIC) to select K [32, 33], which is defined as

$$\begin{aligned} &DIC(K) \\ &:= E\big[L(\Theta;K) \,|\, Y_{1:T}, K\big] + \big\{ E\big[L(\Theta;K) \,|\, Y_{1:T}, K\big] - L(\overline{\Theta};K)\big\} \\ &= 2E\big[L(\Theta;K) \,|\, Y_{1:T}, K\big] - L(\overline{\Theta};K), \end{aligned} \tag{23}$$

where $L(\Theta;K) = -2\log P\big(Y_{1:T}|\Theta,K\big)$ and $\overline{\Theta} = E\big(\Theta \,|\, Y_{1:T}, K\big)$. We select the value of K that minimizes $DIC(K)$.

Using the MCMC samples $\big\{\big(Z_{1:T}^{(s)}, \Theta^{(s)}\big)\big\}_{s=1}^{S}$ generated from the posterior distribution $P(Z_{1:T}, \Theta \,|\, Y_{1:T}, K)$ as presented in Section 3.1, the DIC value can be estimated as

$$\widehat{DIC}(K) = \frac{2}{S}\sum_{s=1}^{S} L(\Theta^{(s)};K) - L\big(\hat{\Theta};K\big), \tag{24}$$

where $\hat{\Theta} = \frac{1}{S}\sum_{s=1}^{S}\Theta^{(s)}$. However, computing $P\big(Y_{1:T}|\Theta,K\big)$ in $L(\Theta;K)$ involves a a high-dimensional integral of the form

$$\begin{aligned} &P\big(Y_{1:T}|\Theta,K\big) \\ &= \int P\big(Z_{1:T}, Y_{1:T} \,|\, \Theta, K\big)\, dZ_{1:T} \\ &= \int P(Z_1 \,|\, K)P(Y_1 \,|\, Z_1, \Theta, K)\prod_{t=2}^{T} P(Z_t \,|\, Z_{t-1}, \Theta, K)P(Y_t \,|\, Z_t, \Theta, K)\, dZ_{1:T}, \end{aligned} \tag{25}$$

which does not have a closed-form solution. We consider employing the sequential Monte Carlo (SMC) method to evaluate $P\big(Y_{1:T}|\Theta,K\big)$ [24, 3]. The algorithmic steps are provided below.

1. At time $t = 1$, draw samples $Z_1^{(j)}$, $j = 1, \cdots, m$, from the prior distribution $P(Z_1 \,|\, K)$, and compute weights

$$w_1^{(j)} = P(Y_1 \,|\, Z_1^{(j)}, \Theta, K). \tag{26}$$

2. At times $t = 2, \cdots, T$,
   (a) Draw samples $Z_{1:t-1}^{*(j)}$, $j = 1, \cdots, m$, from a discrete distribution constructed using the weighted sample set $\{(Z_{1:t-1}^{(j)}, w_{1:t-1}^{(j)})\}_{j=1}^{m}$, which is

$$\sum_{j=1}^{m}\frac{w_{t-1}^{(j)}}{\sum_{k=1}^{m} w_{t-1}^{(k)}}\,\delta\left(Z_{0:t-1} - Z_{0:t-1}^{(j)}\right), \tag{27}$$

   where $\delta(\cdot)$ is the Dirichlet function.
   (b) Draw $Z_t^{(j)}$, $j = 1, \cdots, m$, from the distribution $P(Z_t \,|\, Z_{t-1}^{*(j)}, \Theta, K)$.
   (c) Compute weights

$$w_t^{(j)} = P(Y_t \,|\, Z_t^{(j)}, \Theta, K).$$

According to the principle of sequential importance sampling, the discrete distribution (27) constitutes an approximation to the distribution $P(Z_{1:t-1} \,|\, Y_{1:t-1}, \Theta, K)$, and the generated samples $Z_t^{(j)}$ $j = 1, \cdots, m$, follow the distribution $P(Z_t \,|\, Y_{1:t-1}, \Theta, K)$ approximately. Therefore, as the SMC sample size $m \to \infty$, we have

$$\begin{aligned} &\widehat{P}(Y_1 \,|\, \Theta, K) \\ &:= \frac{1}{m}\sum_{j=1}^{m} w_1^{(j)} \xrightarrow{a.s.} E\big[w_1^{(j)}\big] \\ &= \int P(Y_1 \,|\, Z_1, \Theta, K)P(Z_1 \,|\, K)\, dZ_1 = P(Y_1 \,|\, \Theta, K) \end{aligned} \tag{28}$$

and

$$\widehat{P}(Y_t \,|\, Y_{1:t-1}, \Theta, K) \tag{29}$$

$$:= \frac{1}{m}\sum_{j=1}^{m} w_t^{(j)} \xrightarrow{a.s.} \int P(Y_t \,|\, Z_t, \Theta, K) P(Z_t \,|\, Y_{1:t-1}, \Theta, K)\, dZ_1$$
$$= P(Y_t \,|\, Y_{1:t-1}, \Theta, K).$$

Consequently, $L(\Theta; K)$ in (24) can be estimated by

(30)

$$\begin{aligned} &\widehat{L}(\Theta; K) \\ &= -2\log \widehat{P}\big(Y_{1:T}|\Theta, K\big) \\ &:= -2\log\left[\widehat{P}(Y_1 \,|\, \Theta, K) + \sum_{t=1}^{T} \log \widehat{P}(Y_t \,|\, Y_{1:t-1}, \Theta, K)\right]. \end{aligned}$$

The reader is referred to [25] and [14] for further details on the SMC method.

4. SIMULATION STUDIES

Since three-factor models are commonly used to capture diverse shapes and explain the majority of variation in yield curves [27, 15, 12, 13], we consider a WAR(1) process with $n = 2$ and $K = 3$ in model (8) for simulation and empirical study, in which $\{\Omega_t\}$ is a 2×2 symmetric matrix series that contains three different dynamic risk factors.

We simulate the data using model (8) with $T = 500$ and maturities $\tau \in \{1, 4, 8, 16, 20\}$. The parameters in the stochastic discount factor are set as $d = -4 \times 10^{-3}$ and

$$C = \begin{pmatrix} -8 & -5 \\ -5 & -30 \end{pmatrix} \times 10^{-5}.$$

The coefficient matrix in the WAR(1) process is given by

$$M = \begin{pmatrix} 0.96 & -0.05 \\ 0.1 & 0.85 \end{pmatrix}, \quad \text{and} \quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For the initial state $Z_1 = (z_{1,1}, \cdots, z_{1,K})$, we draw $z_{1,k}$, $k = 1, \cdots, K$, independently from the normal distribution

$$N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

Figure 1 displays one realization of the simulated WAR(1) process $\{\Omega_{1:T}\}$. The time-varying correlation between the two volatility risks is presented in Figure 2, which is computed as $\rho_t = \frac{\Omega_{t,12}}{\sqrt{\Omega_{t,11}\Omega_{t,22}}}$. It is evident that substantial variation exists not only in the variances but also in the correlations. Figure 3 presents the average upward-sloping yield curve, together with yield curves of distinct shapes generated in the sample.

[Figure 1. Simulated Wishart factors in $\Omega_{1:T}$]

[Figure 2. Time-varying correlation $\rho_t = \frac{\Omega_{t,12}}{\sqrt{\Omega_{t,11}\Omega_{t,22}}}$]

[Figure 3. Average yields and specific shapes generated in the sample]

Given that $K = 3$, we estimate the parameters $\Theta = (M, d, C, \sigma^2)$ and the risk factor series $\Omega_{1;T}$ with the MCMC procedure described in Section 3.1. We generate 40,000 samples via the MCMC procedure, with the first 20,000 samples discarded as the burn-in period. The posterior distributions of the parameters are shown in Figure 4. It can been seen that the posterior distributions concentrate closely around the true parameter values even under non-informative priors. The trace plots in Figure 5 indicate satisfactory convergence of the Markov chain.

[Figure 4. Posterior distribution of M, C, d, and h]

[Figure 5. Trace plots of posterior draws]

Figure 6 plots the the 95% highest posterior density (HPD) intervals of $\Omega_{1:T}$, which closely track the true $\Omega_{1:T}$. Figure 7 displays the implied correlation derived from the simulated data alongside the estimated median correlation series. While the estimated correlation series is smoother than its simulated counterpart, it effectively captures the time-varying patterns of the true correlation series.

[Figure 6. The 95% HPD intervals of $\Omega_{1:T}$ and the true $\Omega_{1:T}$]

[Figure 7. True and estimated time-varying correlation series]

Table 1 summarizes the goodness-of-fit metrics for the yields $\{y_{t,\tau}\}$. The mean absolute error (MAE) and root mean squared error (RMSE) between the estimated posterior mean and the simulated data range from 10 to 15 basis points (bps), which are comparable to those reported in typical empirical studies of three-factor yield curve models.

Next, we use $\widehat{DIC}(K)$ in (24) to select the degrees of freedom, K, for the simulated data. To save computational time, we retain one sample of the parameter Θ out of every 20 consecutive MCMC samples generated above; that is, we keep 1000 samples of $\Theta^{(s)}$. The SMC procedure described in Section 3.2 with $m = 5000$ is used to compute $L(\Theta, K)$ for each $\Theta^{(s)}$. Figure 8 plots $\widehat{DIC}(K)$ for $K = 2, 3, 4, 5$, which shows that the model attains the smallest $\widehat{DIC}(K)$ at $K = 3$, the true degrees of freedom.

[Figure 8. Logarithm of Bayesian posterior probabilities for alternative K]

5. EMPIRICAL STUDY OF U.S. YIELD CURVE WITH THE WTSM

We apply the WTSM model with $n = 2$ to the quarterly U.S. zero-coupon equivalent yield curve data from 1961Q2 to 2025Q4, among which the 1-quarter yield, 1-, 2-, 3-, 4-, and 5-year yields from [26]. The yield data are plotted in Figure 9. The statistics for its mean and standard deviation are reported in Table 2.

[Figure 9. U.S. Treasury yield curve from 1961:Q2 to 2025:Q4]

Using the SMC procedure described in Section 3.2, we calculate $\widehat{DIC}(K)$ for $K = 2, 3, 4, 5$ and report the values in Table 3. It suggests that we use $K = 3$ in the WTSM. In the following, we focus on Bayesian inference of the model for the $K = 3$ case.

5.1 Parameter Estimation and Goodness of Fit

Table 5 reports the posterior mean and 95% HPD interval for each parameter. It shows that M_{11} is close to 1 and is far higher than M_{22}, indicating that the process of $\Omega_{t,11}$ is more persistent than that of $\Omega_{t,22}$.

The results in Table 4 show that the model fits the yield curve very well with an average error of 12 bps measured in MAE and around 16 bps in RMSE, similar to the performance in our simulation study.

The tight in-sample fitting across all maturities confirms the WTSM's strong empirical performance in matching U.S. Treasury yield levels over 1961Q2–2025Q4, laying the foundation for reliable interpretation of latent risk factors and yield dynamics below.

5.2 Dynamic Factors

The dynamic factors of the underlying volatility and covalatility factors are plotted in Figure 10, where the solid lines indicate posterior means and dashed lines are for 95% HPD intervals. All three factors display an overall pattern of peaks around the beginning of 1980s, and the first factor is the most persistent which captures the most prominent common movement of the yield curve.

[Figure 10. Posterior means and 95% HPD intervals of $\Omega_{1:T}$]

To uncover the economic meaning of these three factors in their relationship with yields, we plot the factor loadings of these three factors along maturity in Figure 11:

(1) Long-term (Lt) risk factor, $\Omega_{t,11}$. The first loading on factor $\Omega_{t,11}$ is upward sloping, meaning that the long term yield is affected the most, thus a Long-term (Lt) risk factor.

(2) Short-term (St) risk factor, $\Omega_{t,22}$. The second loading on factor $\Omega_{t,22}$ is downward sloping, indicating that the effect of this factor mostly affects the short term yield and has diminishing effect on long term yields.

(3) Covariance (Cr) risk factor $\Omega_{t,12}$ and $\Omega_{t,21}$. The third loading on the covariance factor $\Omega_{t,12}$ and $\Omega_{t,21}$ is hump shaped with the highest effect on yield around 2-3 quarters (more precisely 7 months).

[Figure 11. Factor loadings of three Wishart factors]

These three economically distinct latent components constitute the core drivers of cross-sectional yield variation in our empirical WTSM setup.

5.3 Zero Lower Bound

A key empirical advantage of our three-factor WTSM manifests in its ability to reproduce observed short-term interest rate movements during the prolonged zero lower bound period, as illustrated in Figure 12. Unlike many conventional affine and quadratic term structure models that require auxiliary shadow-rate constructs to accommodate bounded nominal short rates, our specification built on latent Wishart volatility-covariance factors naturally generates empirically consistent short-rate dynamics across the ZLB episode without any additional shadow-rate assumptions. This finding constitutes a novel empirical merit of the Wishart-based arbitrage-free yield framework.

[Figure 12. Short interest rate fit during the ZLB period]

5.4 Model-Implied Time-Varying Correlation Between Yields

Figure 13 plots the implied time-varying correlations between the 3-month yield and 1- and 5-year yields, respectively. Consistent with stylized U.S. fixed-income facts, cross-maturity yield correlations decline as maturity gaps widen, drop notably during economic recessions, and sink to historically low levels post the 2008 financial crisis amid large-scale quantitative easing programs.

[Figure 13. Model-implied conditional correlations of yields across maturities]

6. CONCLUSION

We have developed a Bayesian estimation framework for the Wishart term structure model (WTSM) with a latent WAR(1) process. By reparameterising the volatility process as the outer product of a Gaussian vector autoregression, we overcome the lack of a closed-form transition density and enable efficient Markov chain Monte Carlo (MCMC) inference on parameters and latent states. The optimal degrees of freedom are selected via the deviance information criterion using sequential Monte Carlo.

Our empirical implementation delivers economically interpretable long-term, short-term and correlation risk factors and generates realistic time-varying cross-maturity yield correlations consistent with historical U.S. Treasury data. Most notably, the model reliably fits short-rate dynamics during the zero lower bound period without imposing shadow-rate restrictions, a key practical advantage relative to conventional affine and quadratic yield curve specifications.

Our procedure delivers accurate parameter estimates and latent factor reconstructions, with in-sample fit errors of about 11–14 basis points. Extensions to higher-dimensional WAR processes, non-integer degrees of freedom, or VAR-WAR state dynamics with observable macroeconomic variables in the VAR process are promising avenues for future research.

7. TABLES AND FIGURES

Table 1. Goodness of fit for the yields

	3-month	1-year	2-year	4-year	5-year
MAE(bp)	12.18	11.42	11.47	10.67	10.88
RMSE(bp)	15.46	14.47	14.25	13.78	13.66

Table 2. Mean and standard deviation of U.S. yields

Maturity	Mean(%)	Standard deviation(%)
3-month	4.47	3.19
1-year	4.81	3.25
2-year	5.00	3.22
3-year	5.15	3.16
4-year	5.30	3.10
5-year	5.40	3.03

Table 3. DIC($\div 10^3$) for alternative K

K	2	3	4	5
DIC	-1.713	-1.506	-1.871	-1.779

Table 4. In-sample fit of the three-factor WTSM with U.S. data

	3-month	1-year	2-year	3-year	4-year	5-year
MAE(bp)	13.99	12.93	11.19	11.81	9.77	14.94
RMSE(bp)	18.20	17.00	14.15	15.66	14.54	20.66

Table 5. Posterior mean and 95% HPD intervals of model parameters

Parameter	$\bar{C}_{11}\times 10^6$	$\bar{C}_{12}\times 10^6$	$\bar{C}_{21}\times 10^6$	$\bar{C}_{22}\times$
Mean	**-1.185**	**-2.385**	**-2.385**	**-1.7**
95% HPD interval	(-3.067 -0.183)	(-4.339 -0.943)	(-4.339 -0.943)	(-1.764 -
Parameter	M_{11}	M_{12}	M_{21}	M_2
Mean	**0.941**	**0.064**	**0.106**	**0.8**
95% HPD interval	(0.934 0.948)	(0.054 0.073)	(0.100 0.113)	(0.860 (
Parameter	$d\times 10^4$	$h/10^6$		
Mean	**5.172**	**5.732**		
95% HPD interval	(5.042 5.287)	(5.308 6.157)		

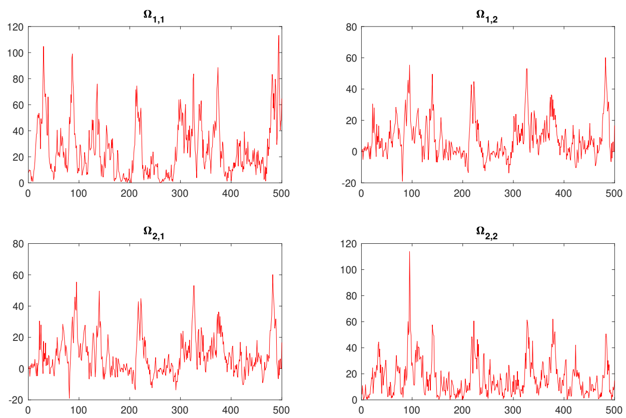

Figure 1. Simulated Wishart factors in $\Omega_{1:T}$

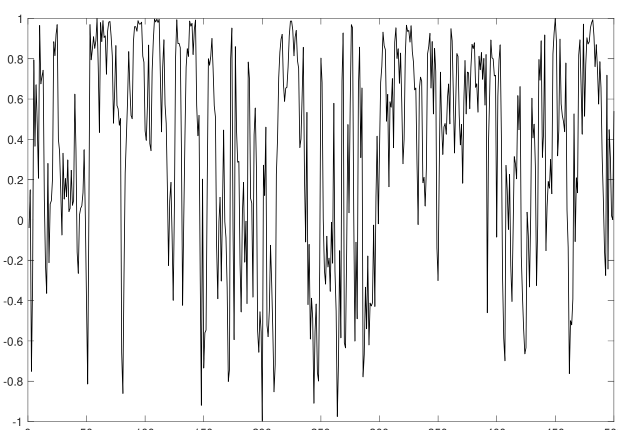

Figure 2. Time-varying correlation $\rho_t = \frac{\Omega_{t,12}}{\sqrt{\Omega_{t,11}\Omega_{t,22}}}$

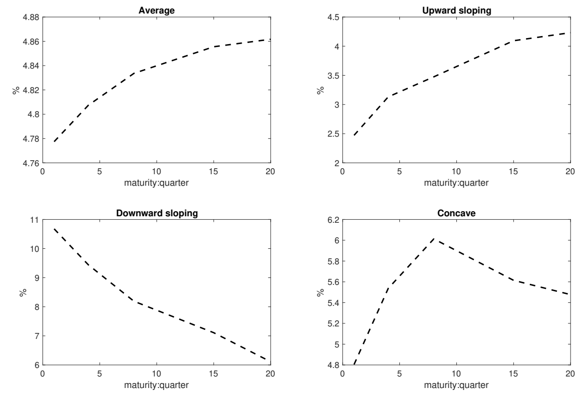

Figure 3. Average yields and specific shapes generated in the sample

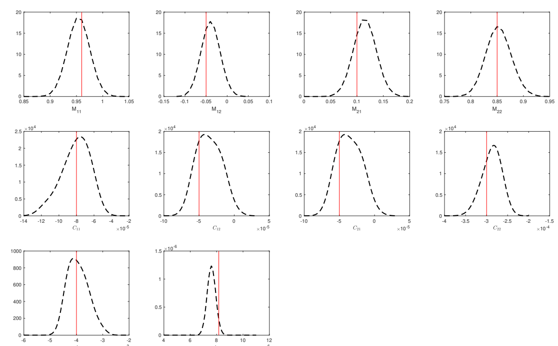

Figure 4. Posterior distribution of M, C, d, and h. (The vertical solid line represents the actual value of parameter.)

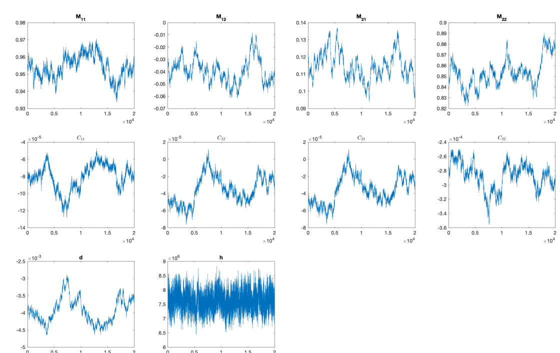

Figure 5. Trace plots of posterior draws

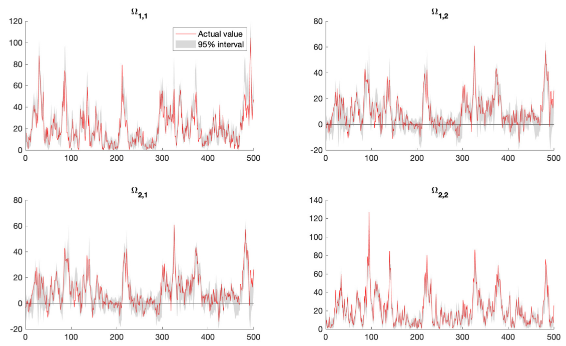

Figure 6. The 95% HPD intervals of $\Omega_{1:T}$ and the true $\Omega_{1:T}$

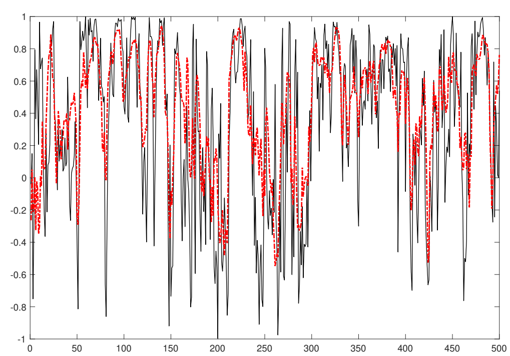

Figure 7. True (solid line) and estimated (dashed line) time-varying correlation series

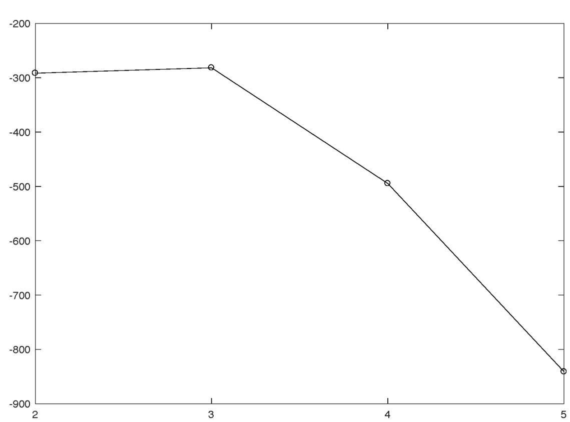

Figure 8. Logarithm of Bayesian posterior probabilities for alternative K

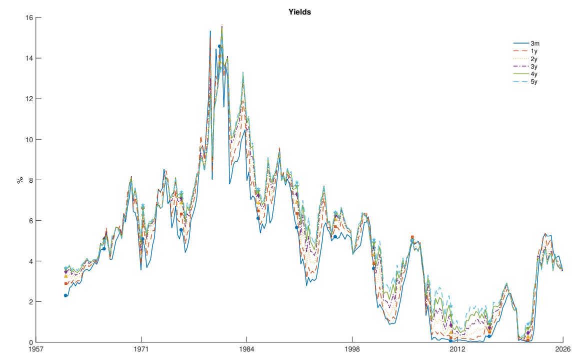

Figure 9. U.S. Treasury yield curve from 1961:Q2 to 2025:Q4

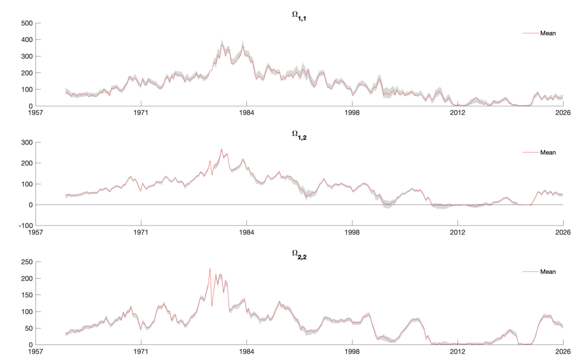

Figure 10. Posterior means and 95% HPD intervals of $\Omega_{1:T}$

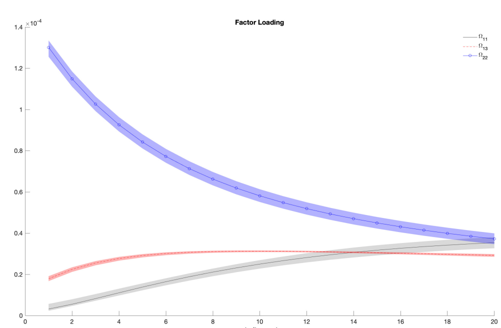

Figure 11. Factor loadings of three Wishart factors

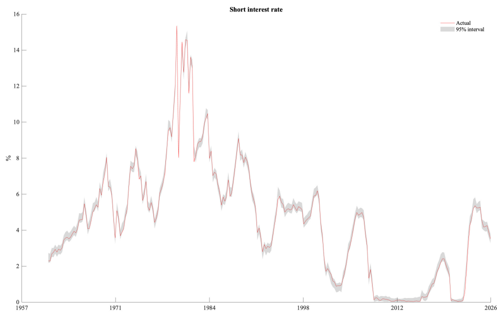

Figure 12. Short interest rate fit during the ZLB period

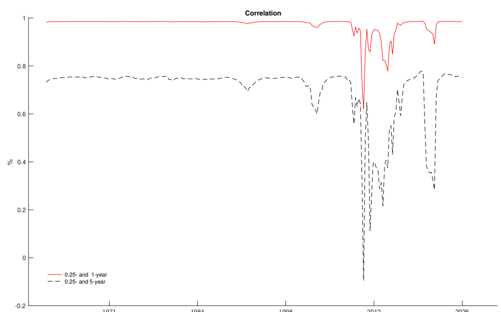

Figure 13. Model-implied conditional correlations of yields across maturities

CODE AND DATA AVAILABILITY

The code is available upon request and the data are publicly available at https://sites.google.com/view/jingcynthiawu/yield-data.

ACKNOWLEDGEMENTS

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Ahn, D.-H., R. F. Dittmar, A. Gallant, and B. Gao (2003). Purebred or Hybrid?: Reproducing the Volatility in Term Structure Dynamics. *Journal of Econometrics 116*(1), 147–180.

[2] Ahn, D.-H., R. F. Dittmar, and A. R. Gallant (2002). Quadratic Term Structure Models: Theory and Evidence. *The Review of Financial Studies 15*(1), 243–288.

[3] Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society, Series B 72*(3), 269–342.

[4] Buraschi, A., A. Cieslak, and F. Trojani (2008). Correlation Risk and the Term Structure of Interest Rates. *Working Paper*.

[5] Buraschi, A., P. Porchia, and F. Trojani (2010). Correlation Risk and Optimal Portfolio Choice. *The Journal of Finance 65*(1), 393–420.

[6] Cheridito, P., D. FilipoviĆ, and R. L. Kimmel (2007). Market Price of Risk Specifications for Affine Models: Theory and Evidence. *Journal of Financial Economics 83*(1), 123 – 170.

[7] Cieslak, A. and P. Povala (2016). Information in the Term Structure of Yield Curve Volatility. *The Journal of Finance 71*(3), 1393–1436.

[8] Collin-Dufresne, P. and R. S. Goldstein (2002). Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility. *The Journal of Finance 57*(4), 1685–1730.

[9] Collin-Dufresne, P., R. S. Goldstein, and C. S. Jones (2009). Can Interest Rate Volatility be Extracted from the Cross Section of Bond Yields? *Journal of Financial Economics 94*(1), 47–66.

[10] Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica 53*(2), 385–407.

[11] Creal, D. D. and J. C. Wu (2015). Estimation of Affine Term Structure Models with Spanned or Unspanned Stochastic Volatility. *Journal of Econometrics 185*(1), 60–81.

[12] Dai, Q. and K. J. Singleton (2000). Specification Analysis of Affine Term Structure Models. *The Journal of Finance 55*(5), 1943–1978.

[13] Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*, 337–364.

[14] Doucet, A. and A. M. Johansen (2009). A tutorial on particle filtering and smoothing: Fifteen years later. In D. Crisan and B. Rozovsky (Eds.), *Handbook of Nonlinear Filtering*. Cambridge: Cambridge University Press.

[15] Duffie, D. and R. Kan (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance 6*(4), 379–406.

[16] Durbin, J. and S. Koopman (1997). Monte carlo maximum likelihood estimation for non-gaussian state space models. *Biometrika 84*, 669–684.

[17] Fonseca, J. D., M. Grasselli, and C. Tebaldi (2007). Option Pricing when Correlations are Stochastic: An Analytical Framework. *Review of Derivatives Research 10*(2), 351–180.

[18] Fonseca, J. D., M. Grasselli, and C. Tebaldi (2008). A Multifactor Volatility Heston Model. *Quantitative Finance 8*(6), 591–604.

[19] Gourieroux, C. (2006). Continuous Time Wishart Process for Stochastic Risk. *Econometric Reviews 25*(2-3), 177–217.

[20] Gourieroux, C., J. Jasiak, and R. Sufana (2009). The Wishart Autoregressive Process of Multivariate Stochastic Volatility. *Journal of Econometrics 150*(2), 167–181.

[21] Gourieroux, C. and R. Sufana (2011). Discrete Time Wishart Term Structure Models. *Journal of Economic Dynamics and Control 35*(6), 815–824.
[22] Green, P. J. (1995). Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika 82*, 711–732.
[23] Gribisch, B. and J. P. Hartkopf (2023). Modeling realized covariance measures with heterogeneous liquidity: A generalized matrix-variate wishart state-space model. *Journal of Econometrics 235*(1), 43–64.
[24] Liu, J. and R. Chen (1998). Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association 93*(443), 1032–1044.
[25] Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. New York: Springer-Verlag.
[26] Liu, Y. and J. C. Wu (2021). Reconstructing the yield curve. *Journal of Financial Economics 142*(3), 1395–1425.
[27] Nelson, C. R. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *Journal of Business 60*, 473–489.
[28] Niu, L. (2010). An Affine Term Structure Model with Auxiliary Stochastic Volatility-Covolatility. *Working Paper*.
[29] Philipov, A. and M. Glickman (2006). Factor multivariate stochastic volatility via wishart processes. *Econometric Reviews 25*, 311–334.
[30] Robert, R. and G. Casella (1999). *Monte Carlo Statistical Methods*. New York: Springer.
[31] Shephard, N. and M. Pitt (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika 84*, 653–667.
[32] Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B 64*(4), 583¨C639.
[33] Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society, Series B 76*(3), 485¨C493.
[34] Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics 5*(2), 177–188.

Ming Lin
School of Economics, Xiamen University, 361005, Xiamen
China
E-mail address: linming50@xmu.edu.cn

Linlin Niu
WISE, Xiamen University, 361005, Xiamen
China
E-mail address: llniu@xmu.edu.cn

Chen Zhang
Lingnan College, Sun Yat-Sen University, 510275, Guangzhou
China
E-mail address: zhangch597@mail.sysu.edu.cn
URL: https://lingnan.sysu.edu.cn/faculty/ZhangChen