SCIENTIFIC LETTERS

Econometrics Letters Volume (6), 2, 2019, December

A Short Note on Error Estimation of Multi-Vasicek Model Parameters

Hong Mao[1]

Wei Hao[2]

Abstract

In this article, a novel study is presented to measure the errors of the multi-Vasicek model parameters. The application is illustrated with historical data from two US insurers. The results show that the multi-Vasicek model can substantially reduce the fitting errors of model parameters for small samples.

Keywords: The multi-Vasicek model; Asymptotic variance; Multi-financial indices.
JEL Codes : C51.

[1] Shanghai Second Polytechnic University e-mail: hmaoi@vip.126..com
[2] State Farm Insurance e-mail: wei.hao.pc3n@statefarm.com

ISSN: 2148-6212
eISSN: 2148-6212

1. Introduction

The Vasicek model first presented by Vasicek in 1977 is widely used to model the term structure of short-term interest rates. It is found that its extension to multi-dimensional situation can display the stochastic changes of multiple variables that may include the returns of risky assets invested, different financial indices, the annual growth rate of wage and the values of both assets and liabilities. Albano et al. (2018) analyze the small-sample properties of the ML estimation procedure in the Vasicek and the CIR models. Their results indicate that the estimator bias may be quite large for small samples and the relative estimator bias could even become worse when the true model parameters are close to the non-stationary case. In this paper, we present the error estimation of the multi-Vasicek model parameters. Our results show that in a small-sample situation, the multi-Vasicek model help to reduce the estimation errors of parameters of the model substantially. To our knowledge, this article is first to study the error estimation of the multi-Vasicek model parameters.

Let us first introduce the Vasicek model. Assume that the instantaneous interest rate follows the stochastic differential equation:

$$dr_t = a(b-r)dt + \sigma dW_t \tag{1}$$

where W_t is a Wiener process under the risk neutral framework, b is its long term equilibrium means, $b-r_t$ is the gap between its current mean and long-run equilibrium level, a is a parameter measuring the speed at which the gap diminishes.

In what follows, we construct a multi-Vasicek model. Assume that the i-th stochastic process $r_i(t)$ for i=1, 2, $\cdots$, n. at time t follows the Vasicek (1977) stochastic process:

$$dr_i(t) = a_i(b_i - r_i)dt + \sigma_i dW^i(t),\ i = 1, 2, \cdots, n. \tag{2}$$

where $W^i(t), i = 1,2,...,n$ are mutually dependent standard Wiener process defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$. For simplicity, it is assumed that the underlying filtration, $\mathcal{F}_t$, coincides with the one generated by the Wiener process, that is, $\mathcal{F}_t = \sigma\left(W_2(s) : 0 \leq s \leq t\right)$ and σ denotes the volatility matrix.

σ_i in equation (2) is the instantaneous volatility of the randomness measure of the i-th-stochastic process, b_i is its equilibrium means in the long run and a_i is a parameter measuring the speed at which the gap diminishes. A diversified portfolio consists of n kinds of stochastic processes following the multi-Vasicek model; the fraction of each process is $\alpha_i, i = 1,2,...,n$, or $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)$,

$$\sum_{i=1}^{n} \alpha_i = 1 \tag{3}$$

and

$$r = \sum_{i=1}^{n} \alpha_i r_i . \tag{4}$$

The differential of equation (4) can be further expressed as

$$\sum_{i=1}^{n} \alpha_i dr_i = \sum_{i=1}^{n} \alpha_i \left(a_i (b_i - r_i) dt + \sigma_i dW^i \right) = \sum_{i=1}^{n} \alpha_i a_i (b_i - r_i) dt + \sum_{i=1}^{n} \alpha_i \sigma_i dW^i \tag{5}$$

If the correlation between $W^i(t)$ and $W^j(t)$ is $\rho_{ij}^{(d)}$, then the variance of the portfolio is

$$V_t \left(\sum_{i=1}^{n} \alpha_i r_i \right) = \left(\sum_{i=1}^{n} \alpha_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \alpha_i \alpha_j \sigma_i \sigma_j \rho_{ij}^{(d)} \right) \tag{6}$$

and the standard deviation is

$$\sigma_r = \sqrt{\sum_{k=1}^{n} \alpha_k^2 \sigma_k^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \alpha_i \alpha_j \rho_{ij}^{(d)} \sigma_i \sigma_j} \tag{7}$$

As in Momon (2004), the expected value of multi-variable portfolio is

$$E(r) = \mu_r(t) = E\sum_{i=1}^{n} \alpha_i r_i = \sum_{i=1}^{n} \alpha_i \mu_i(t) = \sum_{i=1}^{n} \alpha_i e^{-a_i t} \left(r_i(0) + b_i \left(e^{a_i t} - 1 \right) \right), \tag{8}$$

where $\mu_i(t)$ is the expected value of the i-th stochastic variable at time t in the real world measure and $r_i(0)$ is the value of the i-th stochastic variable at $t = 0$. Let

$$\begin{aligned} \sigma_r^2(t) &= Var\left(r\right), \\ \sigma_r^2(t) &= Var\left(\sum_{i=1}^{n} \alpha_i r_i \right) = \sum_{k=1}^{n} \alpha_i^2 \sigma_i^2(t) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \alpha_i \alpha_j \sigma_{ij}(t) \end{aligned} \tag{9}$$

where $\sigma_i^2(t)$ is the volatility of the i-th stochastic variable in the real world measure,

$$\sigma_i^2(t) = \sigma_i^2 \left(\frac{1 - e^{-2a_i t}}{2a_i} \right) \tag{10}$$

and $\sigma_{ij}(t)$ is the covariance between the *i*-th stochastic variable and the *j*-th stochastic variable in the real world measure,

$$\sigma_{ij}(t) = \frac{\sigma_i \sigma_j}{a_i + a_j}\left(1 - e^{-(a_i+a_j)t}\right) \tag{11}$$

For the proof of equation (10), please see Appendix A.

Based on Mamon (2004), $r_i(t) \sim N(\mu_i(t), \sigma_i^2(t))$, $r_i(t)$ also satisfies the stochastic differential equation: $dr_i(t) = \mu_i(t)dt + \sigma_i(t)dB_i$, where B_i is the *i*-th dimensional standard Brownian Motion, $i = 1, 2, \cdots, n$ and $r(t)$ satisfies

$$dr(t) = \mu_r(t)dt + \sigma_r(t)dB_r \tag{12}$$

In what follows, we derive the formulas of error estimation of parameters $\mu_r(t)$, and $\sigma_r(t)$.

We use *ML* estimation and the δ-method introduced in Gourieroux and Jasiak (2001) (Section 12.1.2) to estimate the parameters of the Vasicek model: $\hat{a}_i$, $\hat{b}_i$, $\hat{\sigma}_i$, $i = 1, 2, \cdots, n$ and their asymptotic variances, $V_{asy}(\hat{a}_i)$, $V_{asy}(\hat{b}_i)$, $V_{asy}(\hat{\sigma}_i)$, $i = 1, 2, \cdots, n$. in which historical data are used to fit the multi-Vasicek model.

Let $\varepsilon = (\varepsilon_{ij})$ be an $n \times T$ matrix of residuals for an *n*-dimensional stochastic process, where ij is *i*-th stochastic variable at time j, $i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, T$, and T is sample size.

Let

$$\hat{\eta}_i^2 = \frac{1}{T}\sum_{j=1}^{T} \varepsilon_{ij}^2 \tag{13}$$

where

$$\varepsilon_{ij} = y_{ij} - \bar{y}_i - \hat{\rho}_i(y_{ij-1} - \bar{y}_i) \tag{14}$$

$$\bar{y}_i = \frac{1}{T}\sum_{j=1}^{T} y_{ij} \tag{15}$$

y_{ij} is the value of *i*-th stochastic variable at time j, $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, T$,

and

$$\hat{\rho}_i = \frac{1}{T}\sum_{j=1}^{T}\left(y_{ij}-\bar{y}_i\right)\left(y_{ij-1}-\bar{y}_i\right) / \frac{1}{T}\sum_{j=1}^{T}\left(y_{ij}-\bar{y}_i\right)^2 \tag{16}$$

Based on Gourieroux and Jasiak (2001), the *ML* estimators of parameters $\hat{a}_i$, $\hat{b}_i$ and $\hat{\sigma}_i, i = 1, 2, \cdots, n$ can be written as follows:

$$\hat{a}_i = -\ln \hat{\rho}_i \tag{17}$$

$$\hat{b}_i = \frac{1}{T}\sum_{j=1}^{T} y_{ij} \tag{18}$$

$$\hat{\sigma}_i^2 = -\frac{2\log\left(\hat{\rho}_i\right)\hat{\eta}_i^2}{\left(1-\hat{\rho}_i^2\right)} \tag{19}$$

where $\hat{\eta}_i$ and $\hat{\rho}_i$ satisfy equations (13) and (16) respectively.

Their asymptotic variances are derived by the $\delta-$method as follows:

$$V_{asy}(\hat{a}_i) = \left(\frac{\partial a_i}{\partial \rho_i}\right)^2 V_{asy}\left(\hat{\rho}_i\right) = \frac{1}{T}\frac{1-\hat{\rho}_i^2}{\hat{\rho}_i^2} \tag{20}$$

$$V_{asy}(\hat{b}_i) = \frac{\hat{\eta}_i^2}{T\left(1-\hat{\rho}_i^2\right)} \tag{21}$$

$$V_{asy}(\hat{\sigma}_i) = \left(\frac{\partial}{\partial \rho_i}\sqrt{\hat{\sigma}_i^2}\right)^2 = \left(-\frac{1/\hat{\rho}_i\left(1-\hat{\rho}_i^2\right)+2\hat{\rho}_i\ln\hat{\rho}_i}{\left(1-\hat{\rho}_i^2\right)^2\sqrt{-\frac{2\ln\hat{\rho}_i}{1-\hat{\rho}_i^2}}}\right)^2 V_{asy}\left(\hat{\eta}_i\right) \tag{22}$$

where $\hat{\eta}_i$ and $\hat{\rho}_i$ satisfy equations (13) and (16) respectively, and

$$V_{asy}\left(\hat{\eta}_i\right)=\frac{2\hat{\eta}_i^4}{T} \tag{23}$$

Gourieroux and Jasiak (2001, Section 12.1.2) only gives the formulas for estimating the asymptotic variances of parameters a and b. We present the formula of estimating the asymptotic variance of parameter σ in the Vasicek model in equation (22). Equations (20), (21), (22) and (23) indicate that all of the asymptotic variances of parameters a, b and σ are inversely related to the sample size T which means that decreasing sample size increases the estimation errors of all parameters of the Vasicek model.

In what follows, we derive the formulas of the asymptotic variances of parameters of $\hat{\mu}_r(t)$ and $\hat{\sigma}_r(t)$ of the multi-Vasicek model by the $\delta-$method.

Based on equation (8), we have

$$V_{asy}\left(\hat{\mu}_r(t)\right)=\left(\sum_{i=1}^{n}\alpha_i\left(e^{-\hat{a}_i t}\cdot t\left(\hat{b}_i-r_i(0)\right)\left(V_{asy}\left(\hat{a}_i\right)\right)^{\frac{1}{2}}+\left(1-e^{-\hat{a}_i t}\right)\left(V_{asy}\left(\hat{b}_i\right)\right)^{\frac{1}{2}}\right)\right)^2 \tag{24}$$

Based on equations (9) and (11), we have

$$V_{asy}\left(\sqrt{\hat{\sigma}_r^2}\right)=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j\left(\left(\frac{\left(V_{asy}\left(\hat{\sigma}_i\right)^{\frac{1}{2}}\hat{\sigma}_i+V_{asy}\left(\hat{\sigma}_j\right)^{\frac{1}{2}}\hat{\sigma}_i\right)}{\hat{a}_i+\hat{a}_j}-\frac{\hat{\sigma}_i\hat{\sigma}_j\left(V_{asy}\left(\hat{a}_i\right)^{\frac{1}{2}}+V_{asy}\left(\hat{a}_j\right)^{\frac{1}{2}}\right)}{\left(\hat{a}_i+\hat{a}_j\right)^2}\right)\left(1-e^{-\left(\hat{a}_i+\hat{a}_j\right)t}\right)+\frac{\hat{\sigma}_i\hat{\sigma}_j e^{-\left(\hat{a}_i+\hat{a}_j\right)t}}{\hat{a}_i+\hat{a}_j}\left(V_{asy}\left(\hat{a}_i\right)^{\frac{1}{2}}+V_{asy}\left(\hat{a}_j\right)^{\frac{1}{2}}\right)\cdot t\right)\right)^2\times\left(-\frac{1}{2\sqrt{\hat{\sigma}_r^2}}\right)^2 \tag{25}$$

Examples

From Mao and Hao (2019), we select two examples to illustrate the error estimation of parameters $\mu_r(t)$ and $\sigma_r(t)$ of the multi-Vasicek model, $\varepsilon_{\mu_r}(t)$ and $\varepsilon_{\sigma_r}(t), t = 1, 2, \cdots, T$, where

$$\varepsilon_{\mu_r}(t) = \left(V_{asy}\ \overset{\wedge}{\mu}_r(t)\right)^{\frac{1}{2}} \text{ and } \varepsilon_{\sigma_r}(t) = \left(V_{asy}\ \overset{\wedge}{\sigma}_r(t)\right)^{\frac{1}{2}}$$

Assume that $\boldsymbol{x}(t)$ denotes the n-dimensional financial indices as follows:

$$\boldsymbol{x}(t) = \left(x_1(t), x_2(t), ..x_n(t)\right) \tag{26}$$

The *i*-th standardized financial index is expressed as

$$z_i(t) = \frac{x_i(t) - \min(x_i(t))}{\max(x_i(t)) - \min(x_i(t))}, i = 1, 2, \cdots, n, 0 \le t \le T \tag{27}$$

Since the importance of each financial index may vary, we assign different weights $\alpha_i, i = 1, 2, .., n$ to each different standardized financial index that satisfies:

$$\sum_{i=1}^{n} \alpha_i = 1$$

Mao and Hao (2019) shows that the optimal weights distributed to each financial index are equal weights, that is,

$$\alpha_1 = \alpha_2 = \cdots, = \alpha_i = \cdots, = \alpha_n = \frac{1}{n} \tag{28}$$

Table 1 and Table 2 list the parameters and estimation errors of the Vasicek model for each financial index of a US life insurer and a US property and casualty insurer respectively. Table 3 and Table 4 list the parameters of the means and the volatilities of the weighted average of all financial indices, $\overset{\wedge}{\mu}_r(t)$ and $\overset{\wedge}{\sigma}_r(t)$ and their estimation errors of the multi-Vasicek model for these two insurers. Figure 1 displays the change patterns of their estimation errors with time. The results in Table 1 through Table 4 and Figure 1 indicate that the estimation errors of the parameter of $\overset{\wedge}{\mu}_r(t)$ of the multi-Vasicek model are very small even though those of the parameters $\hat{a}_i, i = 1, 2, \cdots, 6(7)$ are rather large (the highest relative errors are over 0.90 and the lowest ones are over 0.60). However, the estimation errors of the parameter of $\overset{\wedge}{\sigma}_r(t)$ of the multi-Vasicek model are rather large, especially for LifeNo.2., but they are still within acceptable range for most situations. As stated by Albano (2019), the impact of a small size sample on the range of estimation errors of parameter $\hat{a}$ is significant. However, our results show that the multi-Vasicek model can partially offset the estimation errors of parameter $\hat{a}_i, i = 1, 2, \cdots, 6(7)$. Finally, Figure 1 shows that the estimation errors of the parameter $\hat{\mu}_r(t)$ of the multi-Vasicek model are extremely small ($\leq 10^{-5}$) while those of $\hat{\sigma}_r(t)$ first increase with time at first but then decrease and finally converge to a constant.

Table 1. Values of the parameters and the estimation errors of the Vasicek model with seven indices for one US life insurer[3] from 2001–2017

Life No.2						
Index	a_i	ε_{a_i}	b_i	ε_{b_i}	σ_i	ε_{σ_i}
x_1	0.8751	0.5289	0.3147	0.0416	0.2267	0.0105
x_2	.01390	0.1373	0.4783	0.0873	0.1899	0.0063
x_3	1.1725	0.7449	0.3911	0.0533	0.3363	0.0249
x_4	0.8502	0.5132	0.1134	0.0745	0.4003	0.0325
x_5	0.4809	0.3084	0.4602	0.0712	0.2878	0.0155
x_6	0.2667	0.2036	0.5031	0.0659	0.1985	0.0071
x_7	0.6844	0.4152	0.4624	0.0488	0.2354	0.0108

Table 2. Values of the parameters and the estimation errors of the Vasicek model with six indices for one US P&C insurer from 2001–2017

P&CNo.1						
Index	a_i	ε_{a_i}	b_i	ε_{b_i}	σ_i	ε_{σ_i}
x_1	0.4504	0.2932	0.3154	0.0595	0.2328	0.0101
x_2	0.2474	0.1941	0.3039	0.0548	0.1590	0.0045
x_3	0.8164	0.4922	0.6113	0.0619	0.3263	0.0214
x_4	0.3384	0.2386	0.6998	0.0620	0.2102	0.0080
x_5	0.3121	0.2258	0.6560	0.0602	0.1960	0.0070
x_6	0.4593	0.2976	0.2830	0.0586	0.2318	0.0100

[3] $\varepsilon_{a_i}, \varepsilon_{b_i}$ and ε_{σ_i} in Tables 1 and 2 express the errors of estimation of parameters of a_i, b_i and σ_i for $i = 1, 2, \cdots, 7$ in the multi-Vasicek model, where $\varepsilon_{a_i} = \left(V_{asy}\left(\hat{a}_i\right)\right)^{\frac{1}{2}}$, $\varepsilon_{a_i} = \left(V_{asy}\left(\hat{b}_i\right)\right)^{\frac{1}{2}}$ and $\varepsilon_{\sigma_i} = \left(V_{asy}\left(\hat{\sigma}_i\right)\right)^{\frac{1}{2}}$.

Table 3. Values of the parameters and the estimation errors of the multi-Vasicek model with seven indices for one US life insurer from 2001–2017

t	1	2	3	4	5	6
$\hat{\mu}_r(t)$	0.3144	0.3443	0.3630	0.3751	0.3827	0.3879
$\varepsilon_{\mu_r}(t)$	0.0000135	0.0000018	0.000016	0.000014	0.000011	0.000008
$\hat{\sigma}_r(t)$	0.1991	0.2270	0.2365	0.2407	0.2427	0.2439
$\varepsilon_{\sigma_r}(t)$	0.2535	0.2373	0.2376	0.2406	0.2432	0.2448
t	7	8	9	10	11	12
$\hat{\mu}_r(t)$	0.3911	0.3930	0.3941	0.4604	0.4606	0.4605
$\varepsilon_{\mu_r}(t)$	0.0000062	0.0000047	0.0000037	0.0000030	0.0000024	0.0000021
$\hat{\sigma}_r(t)$	0.2441	0.2451	0.2454	0.2456	0.2457	0.2458
$\varepsilon_{\sigma_r}(t)$	0.2457	0.2460	0.2461	0.2459	0.2457	0.2454
t	13	14	15	16	17	
$\hat{\mu}_r(t)$	0.4603	0.4599	0.4595	0.4590	0.4585	
$\varepsilon_{\mu_r}(t)$	0.0000018	0.0000016	0.0000015	0.0000014	0.0000014	
$\hat{\sigma}_r(t)$	0.2459	0.2459	0.2460	0.2460	0.2460	
$\varepsilon_{\sigma_r}(t)$	0.2452	0.2449	0.2447	0.2445	0.2443	

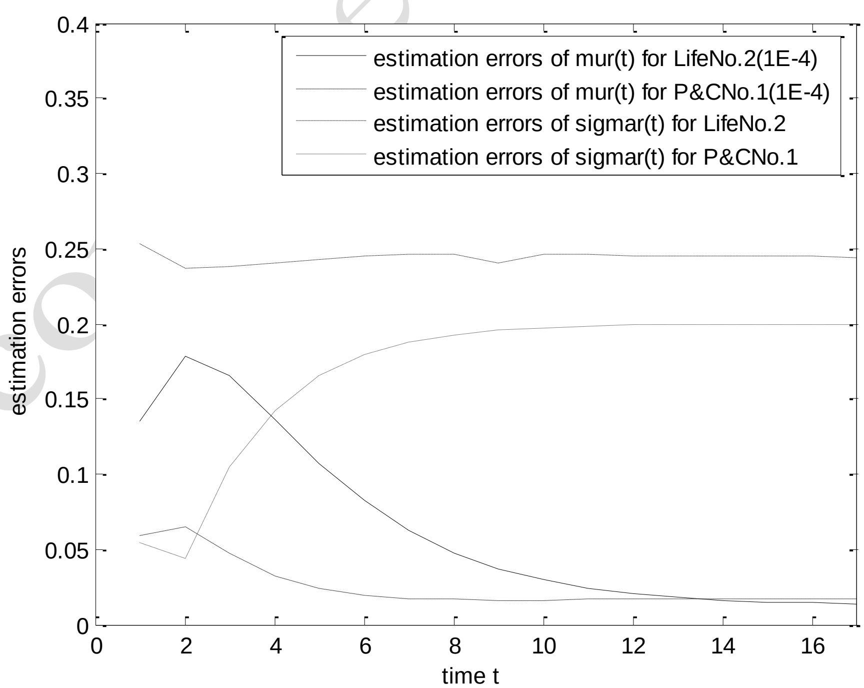

Figure 1. the estimation errors of parameters of the multi-Vasicek model

Table 4. Values of the parameters and the estimation errors of the multi-Vasicek model with six indices for one US P&C insurer from 2001–2017

t	1	2	3	4	5	6
$\hat{\mu}_r(t)$	0.4505	0.4646	0.4712	0.4743	0.4758	0.4765
$\varepsilon_{\mu_r}(t)$	0.0000059	0.0000065	0.0000047	0.0000032	0.0000039	0.0000020
$\hat{\sigma}_r(t)$	0.1822	0.2165	0.2299	0.2358	0.2385	0.2399
$\varepsilon_{\sigma_r}(t)$	0.0542	0.0434	0.1044	0.1421	0.1651	0.1791
t	7	8	9	10	11	12
$\hat{\mu}_r(t)$	0.4769	0.4771	0.4773	0.4774	0.4775	0.4776
$\varepsilon_{\mu_r}(t)$	0.0000018	0.0000017	0.0000016	0.0000016	0.0000016	0.0000017
$\hat{\sigma}_r(t)$	0.2405	0.2409	0.2410	0.2411	0.2412	0.2412
$\varepsilon_{\sigma_r}(t)$	0.1875	0.1926	0.1956	0.1974	0.1984	0.1991
t	13	14	15	16	17	
$\hat{\mu}_r(t)$	0.4777	0.4778	0.4779	0.4779	0.4780	
$\varepsilon_{\mu_r}(t)$	0.0000016	0.0000017	0.0000017	0.0000017	0.0000017	
$\hat{\sigma}_r(t)$	0.2412	0.2412	0.2412	0.2412	0.2412	
$\varepsilon_{\sigma_r}(t)$	0.1994	0.1997	0.1998	0.1999	0.1999	

Conclusions

A novel study is proposed for the error estimation of multi-Vasicek model parameters. The application is illustrated with two examples from both a US life insurer and a US P&C insurer. The results indicate that the multi-Vasicek model can substantially reduce the estimation errors of the model parameter, $\hat{\mu}_r(t)$ for small sample size. The results also show that the estimation errors of the parameter, $\hat{\sigma}_r(t)$ decrease with time at first and then increase, finally, tend to plateau.

References

Albano, G., M. L. Rocca and C. Perna, 2018, Small sample properties of *ML* estimator in Vasicek and CIR models: a simulation experiment. https://doi.org/10.1007/s10203-019-00237-y

Korn, O. and C. Koziol, 2006, Bond portfolio optimization: a risk-return approach, *Journal of Fixed Income*, 15: 48–60.

Mamon, R. S., 2004, Three ways to solve for bond prices in the Vasicek model, *Journal of Applied Mathematics and Decision Science* 8, 1–14.

Mao, H. and W. Hao, 2019, Dynamic monitoring and forecasting of the soundness of U.S. insurers in a cyclical environment, *Asia Pacific Journal of Risk and Insurance*, 13, 1-15.

Vasicek, O. A., 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177–188.

Appendix A:

We calculate the time varying covariance of $\sigma_{ij}(t)$. In a manner similar to Korn and Koziol (2006), and Mamon (2004), we have

Vasicek model: $dr_t = a\left(b-r\right)dt + \sigma dW_t$.

Write $X\left(u\right) = r_u - b$,

$dX\left(t\right) = -aX\left(t\right) + \ \sigma dW_t$,

and $X\left(u\right) = e^{-au}\left(X\left(0\right) + \int_0^u \sigma e^{as} dw_s\right)$

In the following, we calculate the time varying covariance, $\sigma_{ij}(t)$, By Itô isometry, we have:

$$\sigma_{ij}(t) = Cov\left(r_{it}(t), r_{jt}(t)\right) = E(r_{it}(t)r_{jt}(t)) - E\left(r_{it}(t)\right)E\left(r_{jt}(t)\right)$$

$$= E\left(\left(b_i + (r_{i0} - b_i)e^{-a_i t} + \sigma_i e^{-a_i t}\int_0^t e^{a_i s} dw_s\right)\left(b_j + (r_{j0} - b_j)e^{-a_j t} + \sigma_j e^{-a_j t}\int_0^t e^{a_j s} dw_s\right)\right)$$

$$-\left(b_i + (r_{i0} - b_i)e^{-a_i t}\right)\left(b_j + (r_{j0} - b_j)e^{-a_j t}\right)$$

$$= \sigma_i \sigma_j e^{-a_i t - a_j t} E(\int_0^t e^{a_i s} dw_s \int_0^t e^{a_j s} dw_s)$$

$$= \sigma_i \sigma_j e^{-a_i t - a_j t} \int_0^t e^{(a_i + a_j)s} ds$$

$$= \frac{\sigma_i \sigma_j}{a_i + a_j}\left(1 - e^{-(a_i + a_j)t}\right)$$