Bulletin of the University of Electro-Communications, 15-2, pp.209-214 (2003)

On Selection of Moment Conditions in GMM from Viewpoints of Quasi-Likelihood

Koichi Miyazaki*, Hiroe Tsubaki**

Abstract

In the estimation of parameters of 1-factor spot rate model based on GMM, Chan, Karolyi, Longstaff and Sanders (1992) introduced moment conditions, which work very well in practice. However, they didn't refer how to select the conditions and how sensitive the estimation results are against the selection of the moment conditions. This article suggests the optimal strategy of the above selection in the CKLS model from the viewpoints of Heyde (1997)'s quasi-likelihood and optimal estimating equation theory, which may not be fairly popular in the econometric field as in the biometric field and also clarifies that the derived optimal GMM estimators are asymptotically equivalent to their quasi-maximum likelihood estimates. In numerical examples, the parameter estimation results based on several moment conditions are provided in the case of CIR model and Vasicek model.

Keywords : 1-factor spot rate model, selection of the moment conditions, GMM, quasi-likelihood

I. Introduction

Since the likelihood of most spot rate models or forward rate model models cannot be explicitly represented with the transition density except for simple cases as deterministic models whose diffusion functions are independent of the state variable, Hansen (1982)'s generalized method of moments (GMM) is one of the most widely used estimation methods for identifying such complex models as Chan, Karolyi, Longstaff and Sander (1992)'s one-factor model (CKLS model) for an instantaneous spot rate, X,

$$dX=(\alpha+\beta X)dt+\sigma X^{\gamma}dZ \quad (1),$$

where α, β, γ, σ are parameters of the model, and t and Z represent the time and the standard Brownian motion, respectively. Miyazaki and Tsubaki (1999) also adopted the GMM for estimating 11 unknown parameters in the proposed bi-variate extension of the model (1) in order to describe the interest rate and credit spreads simultaneously.

After converting a continuous-time model as (1) to an approximately corresponding discrete-time model, the traditional first step of the GMM is to introduce appropriate moment conditions or estimating functions, $f_t(\theta)$ with $E[f_t(\theta)]=0$,. here θ is the unknown parameter vectors and the final step of GMM is to minimize the quadratic form of $g_T(\theta)=\Sigma f_t(\theta)/T$ with θ.

In fact Chan et al. (1992) and Miyazaki et al. (1999) used four and eleven moment conditions to attain just identifiability of their models, respectively. Their introduced moment conditions may work well in practice, however they didn't refer how to select the conditions and how sensitive the estimation results are against the selection of the moment conditions.

This paper suggests the optimal strategy of the above selection in the CKLS model from the viewpoints of Heyde(1997)'s quasi-likelihood and optimal estimating equation theory, which may not be fairly popular in the econometric field as in the biometric field. Wedderburn(1974) suggested that usage of the likelihood of a natural exponential family

Received on September 17, 2002
* Department of Systems Engineering
** Graduate School of Business Sciences, University of Tsukuba

could be justified as a quasi-likelihood of more general semi-parametric distributions with the same variance structure of the exponential family. It means the quasi-maximum-likelihood estimation can be regarded as the natural extension of the method of the least square, since it will result in the estimates based on the asymptotically optimal linear estimating equation in Godambe (1960)'s sense. Heyde(1997) further extended the quasi-likelihood approach to state dependent time series models as Brennan and Schwartz (1980).

This paper also clarifies that the derived optimal GMM estimators are asymptotically equivalent to their quasi-maximum likelihood estimates.

The organization of the paper is following. In section II, the GMM and the Heyde(1997)'s quasi-likelihood estimate are briefly summarized in the case of CKLS model at first and the optimal class of moment conditions are shown to be naturally derived through the quasi-likelihood approach by which the estimates are insensitive against the different selection of individual moment conditions. The authors illustrated numerical examples to confirm their conclusions in section III.

II. GMM and Quasi-Likelihood Method in One-Factor Spot Rate Models

The author treated the discretized version of the CKLS model as

$$X_t - X_{t-1} = \alpha + \beta X_{t-1} + h_t \tag{2a}$$

and

$$E\left[h_t\right] = 0,\ E\left[h_t^2\right] = \sigma^2 X_{t-1}^{2\gamma} \tag{2b}$$

In this case Chan et al.(1992) have already introduced following estimating functions for the GMM,

$$f_t(\theta) = \begin{bmatrix} h_t \\ h_t X_{t-1} \\ h_t^2 - \sigma^2 X_{t-1}^{2\gamma} \\ \left(h_t^2 - \sigma^2 X_{t-1}^{2\gamma}\right) X_{t-1} \end{bmatrix} \tag{3}$$

where $\theta = (\alpha,\ \beta,\ \sigma^2,\ \gamma)^{\mathrm{T}}$ is a vector of unknown parameters. Thus the GMM estimators of θ is obtained by minimizing the quadratic form,

$$J_T(\theta) = g_T'(\theta) W_T(\theta) g_T(\theta),$$

where $g_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} f_t(\theta)$. Hansen(1982) proved that the optimal choice of the above positive definite weight matrix, $W_T(\theta)$ shall be $W_T(\theta)=S^{-1}(\theta)$ for minimizing asymptotic standard errors of the estimates, where $S(\theta)=S[f_t(\theta)f_t'(\theta)]$. We will simply call the optimal GMME to the GMM estimator based on the optimal weight matrix.

Now the authors will introduce quasi-likelihood approach for the model (2). Since the drift and diffusion terms in (2) are $\alpha+\beta X$ and σX^{γ}, respectively, the corresponding natural martingales become $\sum_{t=1}^{T} h_t$ and $\sum_{t=1}^{T} k_t$, respectively, where $k_t = h_t^2 - \sigma^2 X_{t-1}^{2\gamma}$. Then, according to Heyde (1997) the optimal composite quasi-likelihood estimates, or the optimal estimating equation estimates will be obtained by solving the following quasi-likelihood equation system with θ,

$$\begin{aligned} &\sum_{t=1}^{T}\left\{\left(d\bar{H}_t\right)'\left(d\langle H\rangle_t\right)^{-1}\left(I-R_tS_t\right)^{-1} - \left(d\bar{K}_t\right)'\left(d\langle K\rangle_t\right)^{-1}\left(I-S_tR_t\right)^{-1}S_t\right\}h_t \\ &+\sum_{t=1}^{T}\left\{\left(d\bar{K}_t\right)'\left(d\langle K\rangle_t\right)^{-1}\left(I-S_tR_t\right)^{-1} - \left(d\bar{H}_t\right)'\left(d\langle H\rangle_t\right)^{-1}\left(I-R_tS_t\right)^{-1}R_t\right\}k_t \end{aligned} \tag{4}$$

where,

$$\left(d\bar{H}_t\right)' = \left(E\left(\dot{h}_t(\theta)\middle|F_{t-1}\right)\right)',\ \left(d\langle H\rangle_t\right)^{-1} = \left(E\left(h_t(\theta)h_t'(\theta)\middle|F_{t-1}\right)\right)^{-1},$$

$$\left(d\bar{K}_t\right)' = \left(E\left(\dot{k}_t(\theta)\middle|F_{t-1}\right)\right)',\ \left(d\langle H\rangle_t\right)^{-1} = \left(E\left(h_t(\theta)h_t'(\theta)\middle|F_{t-1}\right)\right)^{-1},$$

$$R_t = E\left(h_t(\theta)k_t'(\theta)\ \middle|F_{t-1}\right)\left(E\left(k_t(\theta)k_t'(\theta)\middle|F_{t-1}\right)\right)^{-1},$$

$$S_t = E\left(k_t(\theta)h_t'(\theta)\ \middle|F_{t-1}\right)\left(E\left(h_t(\theta)h_t'(\theta)\middle|F_{t-1}\right)\right)^{-1}.$$

Next several important properties will be introduced to show some similarity between the optimum GMME and the optimal composite maximum quasi-likelihood estimator (QMLE).

Lemma 1

In case that $\alpha = \gamma = 0$ and σ is known in the model (2), in other words, X_t is AR(1), the optimum GMME and the QMLE are completely equivalent if the natural martingales h_t and k_t are adopted as the estimating functions for the optimum GMM.

(The proof is shown in Appendix A)

In case of the model (2), where two more moment conditions is required for the just identification of the four parameters, their appropriate selection will be suggested by the following proposition.

Proposition 1

One of the optimum choices of the moment condition for the model (2) is h_t, k_t, h_tX_t and k_tX_t.

(The proof is shown in Appendix B).

Corollary 1

Instead of h_tX_t, and k_tX_t, other optimum choices of the additional moment conditions are

(1) $h_tX_t^2$ and $k_tX_t^2$,

(2) $h_tX_t^3$ and $k_tX_t^3$,

(3) h_tX_t and $k_tX_t^2$,

(4) $h_tX_t^2$ and k_tX_t and so on.

(The proof is shown in Appendix C)

III. Numerical Examples

Now, a numerical result of Monte Carlo simulation should be illustrated to clarify effects of the property shown in the properties in the section II. Chan et al. (1992) fitted the model (2) to monthly spot-rate data from 1964/06 to 1989/12 and got estimates by their GMM as $\hat{\alpha} = 0.0408$, $\hat{\beta} = -0.5921$, $\hat{\sigma} = 1.29244$ and $\hat{\gamma} = 1.4999$. We should reanalyze these data by QMLE, however, we could not utilize the same data so that we generated virtual time-series data sets of the length, $T=100$, 1000 and 10000, numerically following the model (2) with $\alpha = 0.0408$, $\beta = -0.5921$, $\sigma = 1.29244$ and $\hat{\gamma} = 1.4999$, which are the same values as the above estimates.

We show the results of fitting the model (2), the Vasicek (1977) model and the Cox, Ingersoll and Ross (1985) model in the tables 1and 2 respectively, by using the five different sets of the optimum moment conditions suggested in the proposition 1 and the corollary 1. Here the Vasicek and CIR models are sub models of (2) with a parameter constraint. We numerically confirm that there is little discrepancy of estimates among the 5 sets of moment conditions. Strictly speaking, the different moment conditions may be considered to lead a slightly different speed of convergence of the estimates as $T \to \infty$.

Table1

The Estimated Parameters of Vasicek Model

T: Number of Data Simulated Based on the Estimation Result of KCLS

T=100	X, X	X^2, X^2	X^3, X^3	X, X^2	X^2, X	TRUE
Alpha	0.029708	0.029715	0.029861	0.029793	0.029557	0.0408
	−0.01109	−0.01108	−0.01094	−0.01101	−0.01124	
Beta	−0.42667	−0.42681	−0.429	−0.42787	−0.42448	−0.5921
	0.165432	0.165287	0.1631	0.164227	0.167621	
Sigma	0.021018	0.021095	0.021175	0.021095	0.021022	1.29244
	−1.27142	−1.27134	−1.27126	−1.27134	−1.27142	
Gamma	0	0	0	0	0	1.4999
	−1.4999	−1.4999	−1.4999	−1.4999	−1.4999	

T=1000	X, X	X^2, X^2	X^3, X^3	X, X^2	X^2, X	TRUE
Alpha	0.039346	0.039187	0.038966	0.039256	0.039296	0.0408
	−0.00145	−0.00161	−0.00183	−0.00154	−0.0015	
Beta	−0.57037	−0.56802	−0.56477	−0.56902	−0.56966	−0.5921
	0.021732	0.024076	0.027326	0.023082	0.022442	
Sigma	0.021988	0.022112	0.022242	0.022112	0.021988	1.29244
	−1.27045	−1.27033	−1.2702	−1.27033	−1.27045	
Gamma	0	0	0	0	0	1.4999
	−1.4999	−1.4999	−1.4999	−1.4999	−1.4999	

T=10000	X, X	X^2, X^2	X^3, X^3	X, X^2	X^2, X	TRUE
Alpha	0.042735	0.042987	0.043219	0.042725	0.043007	0.0408
	0.001935	0.002187	0.002419	0.001925	0.002207	
Beta	−0.62053	−0.62425	−0.62764	−0.62043	−0.62451	−0.5921
	−0.02843	−0.03215	−0.03554	−0.02833	−0.03241	
Sigma	0.022227	0.022407	0.022586	0.022403	0.022231	1.29244
	−1.27021	−1.27003	−1.26985	−1.27004	−1.27021	
Gamma	0	0	0	0	0	1.4999
	−1.4999	−1.4999	−1.4999	−1.4999	−1.4999	

Table2

The Estimated Parameters of CIR Model

T : Number of Data Simulated Based on the Estimation Result of KCLS

T=100	X, X	X^2, X^2	X^3, X^3	X, X^2	X^2, X	TRUE
Alpha	0.029708	0.029715	0.029861	0.029793	0.029557	0.0408
	−0.01109	−0.01108	−0.01094	−0.01101	−0.01124	
Beta	−0.42667	−0.42681	−0.429	−0.42787	−0.42448	−0.5921
	0.165432	0.165287	0.1631	0.164227	0.167621	
Sigma	0.021018	0.021095	0.021175	0.021095	0.021022	1.29244
	−1.27142	−1.27134	−1.27126	−1.27134	−1.27142	
Gamma	0	0	0	0	0	1.4999
	−1.4999	−1.4999	−1.4999	−1.4999	−1.4999	

T=1000	X, X	X^2, X^2	X^3, X^3	X, X^2	X^2, X	TRUE
Alpha	0.039346	0.039187	0.038966	0.039256	0.039296	0.0408
	−0.00145	−0.00161	−0.00183	−0.00154	−0.0015	
Beta	−0.57037	−0.56802	−0.56477	−0.56902	−0.56966	−0.5921
	0.021732	0.024076	0.027326	0.023082	0.022442	
Sigma	0.021988	0.022112	0.022242	0.022112	0.021988	1.29244
	−1.27045	−1.27033	−1.2702	−1.27033	−1.27045	
Gamma	0	0	0	0	0	1.4999
	−1.4999	−1.4999	−1.4999	−1.4999	−1.4999	

T=10000	X, X	X^2, X^2	X^3, X^3	X, X^2	X^2, X	TRUE
Alpha	0.042735	0.042987	0.043219	0.042725	0.043007	0.0408
	0.001935	0.002187	0.002419	0.001925	0.002207	
Beta	−0.62053	−0.62425	−0.62764	−0.62043	−0.62451	−0.5921
	−0.02843	−0.03215	−0.03554	−0.02833	−0.03241	
Sigma	0.022227	0.022407	0.022586	0.022403	0.022231	1.29244
	−1.27021	−1.27003	−1.26985	−1.27004	−1.27021	
Gamma	0	0	0	0	0	1.4999
	−1.4999	−1.4999	−1.4999	−1.4999	−1.4999	

IV. Conclusion

The quasi-likelihood theory can naturally suggest the reasonable moment conditions for the GMM to identify the spot rate models; in fact the naturally introduced martingales should be used as the moment conditions themselves in case of just identifiable models. In overidentifiable cases there can be different choices of the moment conditions, however performance of the estimates is practically independent of the choice. More theoretical investigations might clarify the higher order difference of efficiencies of these estimates in future.

Appendix A: Proof of Lemma 1

Put GMME $f_i(\theta)$ as $f_i(\beta)=\begin{pmatrix}\varepsilon_i\\ \varepsilon_i^2-\sigma^2\end{pmatrix}=\begin{pmatrix}h_i\\ k_i\end{pmatrix}$.

Minimizing $J_t(\beta)$ with respect to parameter β is equivalent to solving the following equation.

$$D'(\beta)W_T(\beta)g_T(\beta)=0,$$

where $D(\beta)$ is Jacobian of $g_T(\beta)$ with respect to parameter β.

$S(\theta)=E\lfloor f_i(\beta)f_i'(\beta)\rfloor$ can be computed as next.

$$S(\theta)=\begin{pmatrix}E\left[h_i^2\right] & E\left[h_i^3\right]-\sigma^2E\left[h_i\right]\\ E\left[h_i^3\right]-\sigma^2E\left[h_i\right] & E\left[h_i^4\right]-2\sigma^2E\left[h_i^2\right]+\sigma^4\end{pmatrix},$$
$$=\begin{pmatrix}\sigma^2 & \gamma\sigma^3\\ \gamma\sigma^3 & \sigma^4(\kappa+2)\end{pmatrix}$$

where γ are skewness and kurtosis of h_i respectively. Following Hansen(1982), we choose the optimal weighting matrix as

$$W_T(\theta)=\frac{1}{\sigma^4(\kappa+2-\gamma^2)}\begin{pmatrix}\sigma^2(\kappa+2) & -\gamma\sigma\\ -\gamma\sigma & 1\end{pmatrix}.$$

Using

$$\frac{\partial h_i}{\partial\beta}=-X_{i-1}\ ,\quad E\left[\frac{\partial h_i}{\partial\beta}\middle|F_{i-1}\right]=-E\left[X_{i-1}\middle|F_{i-1}\right],\quad \frac{\partial k_i}{\partial\beta}=-2X_{i-1}h_i\ ,$$
$$E\left[-2X_{i-1}h_i\middle|F_{i-1}\right]=-2X_{i-1}E\left[h_i\middle|F_{i-1}\right]=0 \text{ and } D'(\theta)=\left(-E\left[X_{i-1}\middle|F_{i-1}\right]\ \ 0\right),$$
$$D'(\beta)W_T(\beta)g_T(\beta)$$
$$=\left(-E\left[X_{i-1}\middle|F_{i-1}\right]\ \ 0\right)\frac{1}{\sigma^4(\kappa+2-\gamma^2)}\begin{pmatrix}\sigma^2(\kappa+2) & -\gamma\sigma\\ -\gamma\sigma & 1\end{pmatrix}\begin{pmatrix}E\left[h_i\middle|F_{i-1}\right]\\ E\left[k_i\middle|F_{i-1}\right]\end{pmatrix}$$
$$=\left(-E\left[X_{i-1}\middle|F_{i-1}\right]\ \ 0\right)\frac{1}{\sigma^4(\kappa+2-\gamma^2)}\begin{pmatrix}\sigma^2(\kappa+2)E\left[h_i\middle|F_{i-1}\right]-\gamma\sigma E\left[k_i\middle|F_{i-1}\right]\\ -\gamma\sigma E\left[h_i\middle|F_{i-1}\right]+E\left[k_i\middle|F_{i-1}\right]\end{pmatrix}$$
$$=\frac{1}{\sigma^4(\kappa+2-\gamma^2)}\left(-\sigma^2(\kappa+2)E\left[X_{i-1}\middle|F_{i-1}\right]E\left[h_i\middle|F_{i-1}\right]+\gamma\sigma E\left[X_{i-1}\middle|F_{i-1}\right]E\left[k_i\middle|F_{i-1}\right]\right)$$
$$=\frac{1}{\sigma^4(\kappa+2-\gamma^2)}\left(-\sigma^2(\kappa+2)E\left[X_{i-1}h_i\middle|F_{i-1}\right]+\gamma\sigma E\left[X_{i-1}k_i\middle|F_{i-1}\right]\right)$$

Thus, in the case that γ, κ, σ are known, β is estimated based on the following equation.

$$\frac{1}{\sigma^4(\kappa+2-\gamma^2)}\frac{1}{T}\sum_{i=1}^{T}\left\{-\sigma^2(\kappa+2)X_{i-1}h_i+\sigma X_{i-1}\gamma k_i\right\}=0$$

On the other hand, noticing $\dot{\sigma}=d\sigma/d\theta=0$, the optimal estimating equation (4) is reduced to the following form.

$$\frac{1}{\sigma^4(\kappa+2-\gamma^2)}\sum_{i=1}^{T}\left\{-\sigma^2(\kappa+2)X_{i-1}h_i+\sigma X_{i-1}\gamma k_i\right\}=0$$

Therefore, the optimum GMME and the QMLE are completely equivalent.

Appendix B: Proof of Proposition 1

The proof of proposition2 is pretty similar fashion as that of proposition1 with some more computational burden. To make the notation concise, we omit suffix i and the conditional information F_{i-1} temporary.

We choose the estimating function (3) as GMME. In this case $D'(\theta)$, $W_T(\theta)$, $g_T(\theta)$, (θ) are computed as next.

$$D'(\theta)=\begin{pmatrix} -1 & -E[X] & 0 & 0 \\ -E[X] & -E[X^2] & 0 & 0 \\ 0 & 0 & -2\sigma E[X^{2\gamma}] & -2\sigma E[X^{2\gamma+1}] \\ 0 & 0 & -2\sigma^2 E[X^{2\gamma}\ln X] & -2\sigma^2 E[X^{2\gamma+1}\ln X] \end{pmatrix}, \quad \text{(B-1)}$$

$$W_T(\theta)=\frac{1}{AB}\begin{pmatrix} E[X^2]E[k^2] & -E[X]E[k^2] & -E[X^2]E[hk] & E[X]E[hk] \\ -E[X]E[k^2] & E[k^2] & E[X]E[hk] & -E[hk] \\ -E[X^2]E[hk] & E[X]E[hk] & E[X^2]E[h^2] & -E[X]E[h^2] \\ E[X]E[hk] & -E[hk] & -E[X]E[h^2] & E[h^2] \end{pmatrix}, \quad \text{(B-2)}$$

$$g_T(\theta)=\begin{pmatrix} E[h] \\ E[hX] \\ E[k] \\ E[kX] \end{pmatrix}, \quad \text{(B-3)}$$

where $A = E[k^2]E[h^2]-E[hk]^2,\ B = E[X^2]-E[X]^2$.

Using (B-1), (B-2) and (B-3),

$$W_T(\theta)g_T(\theta)=\frac{1}{A}\begin{pmatrix} E[k^2]E[h]-E[hk]E[k] \\ 0 \\ E[h^2]E[k]-E[hk]E[h] \\ 0 \end{pmatrix}, \quad \text{(B-4)}$$

$$D'(\theta)W_T(\theta)g_T(\theta)=\frac{1}{A}\begin{pmatrix} -E[k^2]E[h]+E[hk]E[k] \\ -E[X]\left(E[k^2]E[h]-E[hk]E[k]\right) \\ -2\sigma E[X^{2\gamma}]\left(E[h^2]E[k]-E[hk]E[h]\right) \\ -2\sigma^2 E[X^{2\gamma}\ln X]\left(E[h^2]E[k]-E[hk]E[h]\right) \end{pmatrix}. \quad \text{(B-5)}$$

Using the notation such as

$$E[hk]=E[h_i^3|F_{i-1}]=\tilde{h}_i^3,\ \ E[k^2]=E[h_i^4|F_{i-1}]-(\sigma X_{i-1}^{\gamma})^4=\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4$$

and $E[k^2]=E[h_i^4|F_{i-1}]-(\sigma X_{i-1}^{\gamma})^4=\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4$

$$D'(\theta)W_T(\theta)g_T(\theta)=\sum_{i=1}^{T}\frac{1}{A_i}\begin{pmatrix} -h_i\left(\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4\right)+\tilde{h}_i^3k_i \\ X_{i-1}\left\{-h_i\left(\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4\right)+\tilde{h}_i^3k_i\right\} \\ 2\sigma X_{i-1}^{\gamma}\left\{\tilde{h}_i^3h_i-(\sigma X_{i-1}^{\gamma})^2k_i\right\} \\ 2\sigma X_{i-1}^{\gamma}\ln X_{i-1}\left\{\tilde{h}_i^3h_i-(\sigma X_{i-1}^{\gamma})^2k_i\right\} \end{pmatrix}.$$

On the other hand,
using

$$1-R_iS_i=1-\frac{(\tilde{h}_i^3)}{(\sigma X_{i-1}^{\gamma})^2\left(\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4\right)},\ (d\overline{H}_i)=\left(E(\dot{h}_i(\theta)|F_{i-1})\right)=(-1\ \ -X_{i-1}\ \ 0\ \ 0),$$

$$(d\overline{H}_i)=\left(E(\dot{h}_i(\theta)|F_{i-1})\right)=(-1\ \ -X_{i-1}\ \ 0\ \ 0)$$

optimal estimating equation (4) is reduced to the following form.

$$\sum_{i=1}^{T}\left\{\frac{\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4}{A_i}\begin{pmatrix} -1 \\ -X_{i-1} \\ 0 \\ 0 \end{pmatrix}-\frac{\tilde{h}_i^3}{A_i}\begin{pmatrix} 0 \\ 0 \\ -2\sigma X_{i-1}^{2\gamma} \\ -2\sigma^2 X_{i-1}^{2\gamma}\ln X_{i-1} \end{pmatrix}\right\}h_i$$

$$+\sum_{i=1}^{T}\left\{\frac{(\sigma X_{i-1}^{\gamma})^2}{A_i}\begin{pmatrix} 0 \\ 0 \\ -2\sigma X_{i-1}^{2\gamma} \\ -2\sigma^2 X_{i-1}^{2\gamma}\ln X_{i-1} \end{pmatrix}-\frac{\tilde{h}_i^3}{A_i}\begin{pmatrix} -1 \\ -X_{i-1} \\ 0 \\ 0 \end{pmatrix}\right\}k_i$$

$$=\sum_{i=1}^{T}\frac{1}{A_i}\begin{pmatrix} -h_i\left(\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4\right)+\tilde{h}_i^3k_i \\ X_{i-1}\left\{-h_i\left(\tilde{h}_i^4-(\sigma X_{i-1}^{\gamma})^4\right)+\tilde{h}_i^3k_i\right\} \\ 2\sigma X_{i-1}^{\gamma}\left\{\tilde{h}_i^3h_i-(\sigma X_{i-1}^{\gamma})^2k_i\right\} \\ 2\sigma X_{i-1}^{\gamma}\ln X_{i-1}\left\{\tilde{h}_i^3h_i-(\sigma X_{i-1}^{\gamma})^2k_i\right\} \end{pmatrix}$$

Therefore, GMME based on the choice of the moment condition such as ht, kt, htXt and ktXt matches to the QMLE and it is one of the optimum choices of the moment condition for the model (2).

Appendix C: Proof of Corollary 1

We provide the proof in the case of moment condition (1) i.e. $h_i\boldsymbol{X}_i^2$. In this case, even though $\boldsymbol{X}$, $\boldsymbol{X}^2$ and $\boldsymbol{B}=E[\boldsymbol{X}^2]-E[\boldsymbol{X}]^2$ in (B-1), (B-2) and (B-3) are replaced by $\boldsymbol{X}^2$,$\boldsymbol{X}^4$ and $\boldsymbol{B}=E[\boldsymbol{X}^4]-E[\boldsymbol{X}^2]^2$ respectively, (B-4) remains unchanged. Regarding as $D'(\boldsymbol{\theta})$, $\boldsymbol{X}$ and $\boldsymbol{X}^2$ in $D'(\boldsymbol{\theta})$ are replaced by $\boldsymbol{X}^2$ and $\boldsymbol{X}^4$. However, because the second and the fourth row of (B-4) are zero, $D'(\boldsymbol{\theta})$, $W_T(\boldsymbol{\theta})g_T(\boldsymbol{\theta})$ is not affected and it matches (B-5). In the case of other moment conditions (2)-(4), the proof is given in a same manner.

REFERENCES

Brennan, M. J. and E. S. Schwartz, "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency," *Journal of Finance and Quantitative Analysis*, 17, 1982, pp.75-100.

Chan, K. C., G.A. Karolyi, F. A. Longstaff, and A. B. Sanders, "An Emprical Comparison of Alternative Models of the Short-Term Interest Rate," *The Journal of Finance*, Vol. XLVII, No.3, July,1992, pp.1209-1227.

Cox,J.,J.E.Ingersoll,andS.A.Ross, "A Theory of the Term Structure of the Interest Rates," *Econometrica*, 53,1985, pp.385-407.

Godambe, V.P., "An optimum property of regular maximum-likelihood estimation", Ann. Math. Statist. 31, 1960, pp.1208-1211.

Hansen, L. P., "Large sample properties of Generalized Method of Moments estimators," *Ecomnometrica* 50, 1982, pp.1029-1054.

Heyde,C.C, "*Quasi-Likelihood And Its Application*," Springer-Verlag New York, Inc.,1997

Miyazaki, K. and H. Tsubaki, "Comparison of JGB and Bank Debenture Credit Spread Models," *The Journal of Fixed Income*, Vol.9, No.1,June, 1999, pp.63-70.

Vasicek, O., "An Equilibrium Characterization of the Term Structure," *Journal of Financial Econometrics*, 5,1977, pp.177-188.

Wedderburn, R.W.M., "Quasi-likelihood functions, generalized linear models, and Gauss-Newton method," Biometrica, 61, 1974, pp.439-447.