Theoretical Economics Letters, 2016, 6, 647-657
Published Online August 2016 in SciRes. http://www.scirp.org/journal/tel
http://dx.doi.org/10.4236/tel.2016.64070

An Alternative Estimation for Functional Coefficient ARCH-M Model

Xingfa Zhang, Qiang Xiong*

School of Economics and Statistics, Guangzhou University, Guangzhou, China
Email: xingfazhang@hotmail.com, *mathsxq@163.com

Received 4 June 2016; accepted 25 July 2016; published 28 July 2016

Abstract

This article provides an alternative approach to estimate the functional coefficient ARCH-M model given by Zhang, Wong and Li (2016) [1]. The new method has improvement in both computational and theoretical parts. It is found that the computation cost is saved and certain convergence rate for parameter estimation has been obtained.

Keywords

Functional Coefficient, ARCH-M Model, Consistency, Risk Aversion

1. Introduction

ARCH-M model (Engle *et al.* [2]) has been widely studied in last decades due to its various applications. Specially, ARCH-M model gives a way to study the relationship between return and the volatility in finance (for instances, see [3] [4]). Let y_t denote the excess return of a market and h_t denote the corresponding conditional volatility at time *t*. A frequently applied conditional mean in ARCH-M models is $y_t = \delta h_t + \varepsilon_t$ with ε_t being an error term. The above equality gives a straightforward linear relationship between volatility and return: high volatility (risk) causes high return. The volatility coefficient δ can be addressed as relative risk aversion parameter in Das and Sarkar [5] and price of volatility in Chou *et al*. [6]. Many empirical studies have been done based on the above conditional mean. However, some researchers found δ nonconstant and counter-cyclical [7]-[9]. To capture the variation of the volatility coefficient δ, Chou *et al.* [6] studied a time-varying parameter GARCH-M. In their GARCH-M model, the volatility coefficient was assumed to follow a random walk, namely $\delta_t = \delta_{t-1} + v_t$ with v_t being an error term.

Based on Chou *et al.* [6], it makes sense to study the ARCH-M model with a time-varying volatility coefficient. Motivated by the functional coefficient model, Zhang *et al.* [1] consider a class of functional coefficient (G)

*Corresponding author.

How to cite this paper: Zhang, X.F. and Xiong, Q. (2016) An Alternative Estimation for Functional Coefficient ARCH-M Model. *Theoretical Economics Letters*, **6**, 647-657. http://dx.doi.org/10.4236/tel.2016.64070

ARCH-M models. For simplicity, we focus on the functional coefficient ARCH-M model of the form

$$\begin{aligned} &y_t = m(U_t)h_t + \varepsilon_t, \varepsilon_t = e_t\sqrt{h_t}, \\ &e_t \sim i.i.d(0,1), h_t = a_0 + a_1 y_{t-1}^2 + \cdots + a_p y_{t-p}^2. \end{aligned} \tag{1}$$

Here $\{y_t, U_t\}_{t=1}^n$ are observable series and (y_{s-1}, U_s) is independent of $\{e_t\}$ for $t \ge s$. $\theta = (a_0, a_1, \cdots, a_p)^\tau$ is the unknown parameter vector and $m(\cdot)$ is an unknown smooth function. All throughout this article, the superscript τ denotes the transpose of a vector or a matrix. In (1), the volatility coefficient is treated as some unknown smooth function $m(\cdot)$. The conditional variance h_t is assumed to be driven by a new-typed ARCH (p) process: the original ε_{t-i} is replaced by the observable y_{t-i}. Similar to Chou *et al*. [6], the modification for h_t is helpful to estimate the model. In fact, such a setting for the conditional variance in (1) is not new, Ling [10], Ling [11], Zhang *et al*. [12] and Xiong *et al*. [13] have taken advantage of such specifications for the conditional variance. Considering $m(\cdot)$ in (1) as a measure of risk aversion as in Chou *et al*. [6], the improvement of (1) lies in that it gives a way to understand how certain variable impacts the risk aversion.

For model (1), we need to estimate $m(u)$ and $\theta = (a_0, a_1, \cdots, a_p)^\tau$ based on the observable $\{y_t, U_t\}_{t=1}^n$. In Zhang *et al*. [1], the estimation procedures is as follows.

Firstly, given θ, calculating $h_t(\theta)$ based on the second equation of model (1);

Next, getting the estimator $\hat{m}(U_t)$ by functional coefficient regression technique based on the first equation of model (1), by treating $h_t(\theta)$ as observable variable;

Thirdly, calculating residuals $\varepsilon_t(\theta) = y_t - \hat{m}(U_t)h_t(\theta)$ and acquiring $\hat{\theta}$ by minimizing

$$\hat{L}_n(\theta) = \frac{1}{n}\sum_{t=1}^n \left[\log h_t(\theta) + \frac{\hat{\varepsilon}_t(\theta)^2}{h_t(\theta)}\right]\pi(U_t)$$

with respect to θ, where $\pi(\cdot)$ is a known weight function.

It is shown in Zhang *et al*. [1] that the above estimation is consistent. However, there is no concrete convergence rate. Moreover, it can be seen that in the above estimation, $\hat{m}(U_t)$ depends on $h_t(\theta)$ and hence depends on θ. However, there is no simple or explicite expression between them, which will make the calculation a bit time-consuming. In this article, a new simple estimator is given for model (1), which is shown to be consistent and convergence rate is also obtained.

The article is arranged as follows. In Section 2, we explain the idea about estimation approach. Section 3 lists the necessary assumptions to show the convergence results followed in Section 4. We conclude the paper in Section 5. Proofs of lemmas are put in the **Appendix**.

2. Estimation

For model (1), we need to estimate $m(u)$ and $\theta = (a_0, a_1, \cdots, a_p)^\tau$ based on the observable $\{y_t, U_t\}_{t=1}^n$. Denote $f(u)$ to be the probability density function of u_t. Let A be a compact subset of R with nonempty interior and satisfies $\inf_{u\in A} f(u) > 0$. For each $u \in A$, based on (1) we have

$$\begin{aligned} g(u) &:= E(y_t \mid U_t = u) = E\left(m(U_t)h_t + e_t\sqrt{h_t}\right) \\ &= m(u)E(h_t \mid U_t = u) + 0 \\ &= m(u)\left[a_0 + \sum_{k=1}^p a_k E\left(y_{t-k}^2 \mid U_t = u\right)\right] \\ &= m(u)\left[a_0 + \sum_{k=1}^p a_k \sigma_k(u)\right], \end{aligned} \tag{2}$$

where, $\sigma_k(u) := E\left(y_{t-k}^2 \mid U_t = u\right)$. Given $\theta \in \Theta$, define

$$m(\theta, u) = g(u) \Big/ \left[a_0 + \sum_{k=1}^p a_k \sigma_k(u)\right]. \tag{3}$$

Denote $\theta_0 = (a_{00}, a_{10}, \cdots, a_{p0})$ to be the true value for θ. Then, $m(u) = m(\theta_0, u)$ according to (2) and (3). Let $\hat{g}(u)$ and $\hat{\sigma}_k(u)$ be corresponding local linear estimators for $g(u)$ and $\sigma_k(u)$ respectively (Fan and Yao [14]). Then we can define a estimator for $m(\theta,u)$ as

$$\hat{m}(\theta,u) = \hat{g}(u) \Big/ \left[a_0 + \sum_{k=1}^{p} a_k \hat{\sigma}_k(u) \right]. \tag{4}$$

For convenience of notation, we put

$$h_t(\theta) = a_0 + a_1 y_{t-1}^2 + \cdots + a_p y_{t-p}^2, \varepsilon_t(\theta) = y_t - m(\theta, U_t) h_t(\theta), \tag{5}$$

$$h_t = h_t(\theta_0), \varepsilon_t = \varepsilon_t(\theta_0), \hat{\varepsilon}_t(\theta) = y_t - \hat{m}(\theta, U_t) h_t(\theta). \tag{6}$$

Further, define

$$L(\theta) = E\left\{ \left[\log h_t(\theta) + \frac{\varepsilon_t^2(\theta)}{h_t(\theta)} \right] \pi(U_t) \right\}, \tag{7}$$

$$L_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \left\{ \log h_t(\theta) + \frac{\varepsilon_t^2(\theta)}{h_t(\theta)} \right\} \pi(U_t), \tag{8}$$

$$\hat{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \left\{ \log h_t(\theta) + \frac{\hat{\varepsilon}_t^2(\theta)}{h_t(\theta)} \right\} \pi(U_t), \tag{9}$$

where $\pi(\cdot)$ is a nonnegative weight function whose compact support is contained in *A*. Then, in terms of (3) and (9), estimators for θ and $m(u)$ are given as

$$\hat{\theta}_n := \arg\min_{\theta \in \Theta} \hat{L}_n(\theta), \hat{m}(u) := \hat{m}(\hat{\theta}_n, u). \tag{10}$$

In the above estimation procedure, we follow the ideas from Christensen *et al.* [15] and Yang [16]. When $\pi(\cdot) \equiv 1$, $L_n(\theta)$ in (8) becomes the commonly used log-likelihood function in the literature. However the direct minimizer of $L_n(\theta)$ with respect to θ is not practical because the quantity $\varepsilon_t(\theta) = y_t - m(\theta, U_t) h_t(\theta)$ in $L_n(\theta)$ depends on the unknown function $m(\theta, U_t)$. Note that $\hat{L}_n(\theta)$ in (9) can be considered as an approximation to $L_n(\theta)$. Consequently, to obtain a feasible estimator for θ, we switch to minimize $\hat{L}_n(\theta)$. For practical minimization in (10), one can refer the algorithm given by Christensen *et al.* [15].

Remark 1. From (4), it can be seen that there is a simple specification between $\hat{m}(\theta,u)$ and θ. Such a simple explicite expression will greatly improve computational efficiency compared to the method in Zhang *et al.* [1].

3. Assumptions

The following assumptions will be adopted to show some asymptotic results. Throughout this paper, we let M, m denote certain positive constants, which may take different values at different places.

Assumption 1. The kernel function $k(.)$ is a bounded density with a bounded support $[-1,1]$.

Assumption 2. The process $\{U_t\}$ has a continuous pdf $f(u)$ satisfying $\inf_{u\in A} f(u) > 0$, where *A* is a compact subset of *R* with nonempty interior. Further, there are constants *m* and *M* such that $0 < m \le |f'(u)| \le M < \infty$ for $u \in A$.

Assumption 3. The considered parameter space Θ is a bounded metric space. The process $\{(y_t, U_t)\}$ from (1) is strictly stationary and ergodic.

Assumption 4. $0 < m \le \sigma_{i,j}(\theta,u) := E\left(h_t^i(\theta) \|\partial h_t(\theta)/\partial\theta\| \,|\, U_t = u\right) \le M < \infty$ holds uniformly for $\theta \in \Theta$, $u \in A$, where $i = 0,1,2$, $j = 0,1$.

Assumption 5. The function $L(\theta)$ defined in (7) has an unique minimum point at $\theta_0 \in \Theta$.

Assumption 6. $g(u), \sigma_k(u)$ defined in (2) satisfy $0 < m \le g(u), \sigma_k(u) \le M < \infty$ uniformly for $u \in A, k = 1,2,\cdots,p$. The corresponding estimators suffice $\sup_{u\in A} |g(u) - \hat{g}(u)| = O_p\left(b_n^2 + \{nb_n/\log n\}^{-1/2}\right)$,

$\sup_{u\in A}\left|\sigma_k(u)-\hat{\sigma}_k(u)\right| = O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right)$, where b_n is the bandwidth such that $b_n \to 0$ and for some $s>2, \delta>0, \beta>2.5,$

$$n^{1-2/s-2\delta}b_n \to \infty \quad \text{and} \quad n^{(\beta+1.5)\left(s^{-1}+\delta\right)-\beta/2-5/4}b_n^{-\beta/2-5/4} \to 0.$$

Remark 2. Assumptions 1 - 3 are frequently adopted in the literature. Assumptions 4 - 5 have been analogously adopted by Yang [16]. In Assumption 6, the boundness is regular. When the bandwidth b_n suffices the described conditions and the processes $\{y_t,U_t\}_{t=1}^n$ satisfies certain mixing conditions, the uniform convergence holds for local linear regression method (Fan and Yao [14], Theorem 6.5).

4. Asymptotic Results

Theorem 1. *Suppose that Assumptions* 1 - 6 *hold. Then for any* $u\in A$,

$$\hat{\theta}_n-\theta_0 = o_p(1), \hat{m}\left(\hat{\theta}_n,u\right)-m(u) = o_p(1).$$

Theorem 1 shows our estimators are consistent. The following Theorem 2 further gives certain convergence rate.

Theorem 2. *Suppose that Assumptions* 1 - 6 *hold. Then for any* $u\in A$,

$$\hat{\theta}_n-\theta_0 = O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right), \hat{m}\left(\hat{\theta}_n,u\right)-m(u) = O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right).$$

In order to prove Theorem 1 and 2, we need the following lemmas whose proofs can be found in the **Appendix**.

Lemma 1. *For* $m(\theta,u)$ *and* $\hat{m}(\theta,u)$ *given in* (3) *and* (4), *suppose that Assumptions* 1 - 6 *hold. Then for* $l,m=0,1, i,j=0,1,\cdots,p,$

$$\sup_{\theta\in\Theta}\sup_{u\in A}\left|\frac{\partial^{(l+m)}m(\theta,u)}{\partial\theta_i^l\partial\theta_j^m}-\frac{\partial^{(l+m)}\hat{m}(\theta,u)}{\partial\theta_i^l\partial\theta_j^m}\right| = O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right). \tag{11}$$

Lemma 2. *For* $L_n(\theta)$ *and* $\hat{L}_n(\theta)$ *given in* (8) *and* (9), *suppose Assumptions* 1 - 6 *hold. Then for* $k=0,1,2,$

$$\sup_{\theta\in\Theta}\left\|\hat{L}_n^{(k)}(\theta)-L_n^{(k)}(\theta)\right\| = O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right), k=0,1,2. \tag{12}$$

Proof of Theorem 1. From (7)-(8), it is not difficult to get

$$\begin{aligned} & L_n(\theta_1)-L_n(\theta_2) \\ &= (1/n)\sum_{t=1}^n \pi(U_t)\left[1/h_t\left(\theta_{1,t}^*\right)\right]\left[\partial h_t\left(\theta_{1,t}^*\right)/\partial\theta^\tau\right](\theta_1-\theta_2) \\ &\quad +(1/n)\sum_{t=1}^n \pi(U_t)\varepsilon_t^2(\theta_1)\left[1/h_t^2\left(\theta_{2,t}^*\right)\right]\left[\partial h_t\left(\theta_{2,t}^*\right)/\partial\theta^\tau\right](\theta_1-\theta_2) \\ &\quad -(1/n)\sum_{t=1}^n \left[\pi(U_t)/h_t(\theta_2)\right]\left\{\left[\varepsilon_t(\theta_1)+\varepsilon_t(\theta_2)\right]\right. \\ &\quad \times\left.\left[\left(\partial m\left(\theta_{3,t}^*,U_t\right)/\partial\theta^\tau\right)h_t(\theta_1)+m(\theta_2,U_t)\left(\partial h_t\left(\theta_{4,t}^*\right)/\partial\theta^\tau\right)\right]\right\}(\theta_1-\theta_2). \end{aligned} \tag{13}$$

Here, for each $t=1,\cdots,n, i=1,2,3,4, \theta_{i,t}^*$ takes value between θ_1 and θ_2. Similar to (A.18), when $U_t\in A$, it can be shown

$$\varepsilon_t^2(\theta) \le M\left[h_t^2\left(\theta^U\right)+e_t^2h_t\left(\theta^U\right)\right] \tag{14}$$

holds for certain finite *M*. Put

$$B_{3n} = (M/n)\sum_{t=1}^n \pi(U_t)\left[1+h_t^2(\theta_U)+e_t^2h_t(\theta_U)+h_t(\theta_U)+e_t^2\right]. \tag{15}$$

According to (A.18) and (A.19), (13)-(15), for certain *M*, it follows

$$\left|L_n(\theta_1)-L_n(\theta_2)\right|\le B_{3n}\left\|\theta_1-\theta_2\right\|. \tag{16}$$

Note e_t is independent of $\left(h_t\left(\theta^U\right),U_t\right)$ and $E\left[e_t^2\right]=1$. Then similar to (A.22), it can be shown that $E\left[B_{3n}\right]<\infty$, implying $B_{3n}=O_p(1)$. Applying Lemma 1 and Theorem 1 in Andrews [17] to $L_n(\theta)$, then it follows that

$$\sup_{\theta\in\Theta}\left|L_n(\theta)-L(\theta)\right|=o_p(1). \tag{17}$$

(12) and (17) give

$$\sup_{\theta\in\Theta}\left|\hat{L}_n(\theta)-L(\theta)\right|=o_p(1), \tag{18}$$

which implies the consistency of $\hat{\theta}_n$ in (10) by Lemma 14.3 (page 258) and Theorem 2.12 (page 28) in Kosorok [18]. In addition,

$$\begin{aligned}
&\hat{m}\left(\hat{\theta}_n,u\right)-m(u)=\hat{m}\left(\hat{\theta}_n,u\right)-m(\theta_0,u)\\
&=\hat{m}\left(\hat{\theta}_n,u\right)-m\left(\hat{\theta}_n,u\right)+m\left(\hat{\theta}_n,u\right)-m(\theta_0,u)\\
&=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right)+\frac{\partial m\left(\hat{\theta}_n^*,u\right)}{\partial\theta^\tau}\left(\hat{\theta}_n-\theta_0\right)\\
&=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right)+\frac{\partial m(\theta_0,u)}{\partial\theta^\tau}\left(\hat{\theta}_n-\theta_0\right)\left[1+o_p(1)\right]\\
&=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right)+o_p(1)=o_p(1),
\end{aligned}$$

where $\hat{\theta}_n^*$ is between $\hat{\theta}_n$ and θ_0.

Proof of Theorem 2. According to (10) and (12), it follows

$$\frac{\partial\hat{L}_n\left(\hat{\theta}_n\right)}{\partial\theta}=0,\frac{\partial\hat{L}_n(\theta_0)}{\partial\theta}-\frac{\partial\hat{L}_n\left(\hat{\theta}_n\right)}{\partial\theta}=\frac{\partial^2\hat{L}_n\left(\tilde{\theta}_n\right)}{\partial\theta\partial\theta^\tau}\left(\theta_0-\hat{\theta}_n\right),$$

$$\hat{\theta}_n-\theta_0=\left[-\frac{\partial^2\hat{L}_n\left(\tilde{\theta}_n\right)}{\partial\theta\partial\theta^\tau}\right]^{-1}\frac{\partial\hat{L}_n(\theta_0)}{\partial\theta}=\left[-\frac{\partial^2 L_n(\theta_0)}{\partial\theta\partial\theta^\tau}+r_{1n}\right]^{-1}\left[\frac{\partial L_n(\theta_0)}{\partial\theta}+r_{2n}\right], \tag{19}$$

where,

$$r_{1n}=\frac{\partial^2 L_n(\theta_0)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2\hat{L}_n\left(\tilde{\theta}_n\right)}{\partial\theta\partial\theta^\tau},r_{2n}=\frac{\partial L_n(\theta_0)}{\partial\theta}-\frac{\partial\hat{L}_n(\theta_0)}{\partial\theta}. \tag{20}$$

From Theorem 1 and Lemmas 1 - 2,

$$\begin{aligned}
r_{1n}&=\left(\frac{\partial^2 L_n(\theta_0)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2 L_n\left(\tilde{\theta}_n\right)}{\partial\theta\partial\theta^\tau}\right)+\left(\frac{\partial^2 L_n\left(\tilde{\theta}_n\right)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2\hat{L}_n\left(\tilde{\theta}_n\right)}{\partial\theta\partial\theta^\tau}\right)\\
&=o_p(1)+O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right)=o_p(1).
\end{aligned} \tag{21}$$

$$r_{2n}=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right).$$

In the above second equality, the first $o_p(1)$ is from the consistency of $\hat{\theta}_n$. Put

$$\Omega_I=E\left\{\pi(U_t)\left[\frac{1}{h_t^2}\frac{\partial h_t(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta^\tau}+\frac{2}{h_t}\frac{\partial\varepsilon_t(\theta_0)}{\partial\theta}\frac{\partial\varepsilon_t(\theta_0)}{\partial\theta^\tau}\right]\right\},$$

$$\Omega_S=E\left\{\pi(U_t)\left[\frac{Ee_t^4-1}{h_t^2}\frac{\partial h_t(\theta_0)}{\partial\theta}\frac{\partial h_t(\theta_0)}{\partial\theta^\tau}+\frac{4}{h_t}\frac{\partial\varepsilon_t(\theta_0)}{\partial\theta}\frac{\partial\varepsilon_t(\theta_0)}{\partial\theta^\tau}\right]\right\}.$$

From (A.9),

$$\frac{\partial^2 L_n(\theta_0)}{\partial\theta\partial\theta^\tau} \xrightarrow{P} \Omega_I. \tag{22}$$

By the martingale central limit theorem (see, for example, Theorem 35.12 in Billingsley [19]), it is not difficulty to show

$$\sqrt{n}\frac{\partial L_n(\theta_0)}{\partial\theta} \xrightarrow{L} N(0,\Omega_S). \tag{23}$$

According to (19)-(23), it follows that

$$\begin{aligned}\hat{\theta}_n - \theta_0 &= O_p(1)\left[O_p\left(n^{-1/2}\right) + O_p\left(b_n^2 + \{nb_n/\log n\}^{-1/2}\right)\right] \\ &= O_p\left(b_n^2 + \{nb_n/\log n\}^{-1/2}\right).\end{aligned} \tag{24}$$

Moreover,

$$\begin{aligned}&\hat{m}\left(\hat{\theta}_n,u\right) - m(u) \\ &= O_p\left(b_n^2 + \{nb_n/\log n\}^{-1/2}\right) + \frac{\partial m(\theta_0,u)}{\partial\theta^\tau}\left(\hat{\theta}_n - \theta_0\right)\left[1 + o_p(1)\right] \\ &= O_p\left(b_n^2 + \{nb_n/\log n\}^{-1/2}\right).\end{aligned} \tag{25}$$

Conjecture. According to (19)-(25), if one can show $r_{2n} = o_p\left(1/\sqrt{n}\right)$, then we can state the following asymptotic normality:

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{L} N\left(0,\Omega_I^{-1}\Omega_S\Omega_I^{-1}\right),$$

$$\sqrt{n}\left[\hat{m}\left(\hat{\theta}_n,u\right) - m(u) - O_p\left(b_n^2 + \{nb_n/\log n\}^{-1/2}\right)\right] \xrightarrow{L} N\left(0,\Delta M^\tau\Omega_I^{-1}\Omega_S\Omega_I^{-1}\Delta M\right),$$

where $\Delta M = \partial m(\theta_0,u)/\partial\theta$.

5. Conclusions

In this paper, a new approach is proposed to estimate the functional coefficient ARCH-M model. The proposed estimators are more efficient and, under regularity conditions, they are shown to be consistent. Certain convergence rate is also given.

Besides that the proof of conjecture in Section 4 needs further development, it is meaningful to further consider a GARCH type conditional variance in model (1). However, such an improvement is not trivial because the estimation method adopted in this paper can not be applied to the GARCH case. An alternative approach needs further development.

Acknowledgements

We thank the Editor and the referee for their comments. Research of X. Zhang and Q. Xiong is funded by National Natural Science Foundation of China (Grant No. 11401123, 11271095) and the Foundation for Fostering the Scientific and Technical Innovation of Guangzhou University. These supports are greatly appreciated.

References

[1] Zhang, X., Wong, H. and Li, Y. (2016) A Functional Coefficient GARCH-M Model. *Communications in Statistics—Theory and Methods*, **45**, 3807-3821. http://dx.doi.org/10.1080/03610926.2014.906615

[2] Engle, E.F., Lilien, D.M. and Robins, R.P. (1987) Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica*, **55**, 391-407. http://dx.doi.org/10.2307/1913242

[3] Bollerslev, T., Tauchen, G. and Zhou, H. (2009) Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies*, **22**, 4463-4492. http://dx.doi.org/10.1093/rfs/hhp008

[4] Bali, T.G. and Yilmaz, K. (2015) The Intertemporal Relation between Expected Return and Risk on Currency. Handbook of Financial Econometrics and Statistics, Springer, New York, 1105-1141. http://dx.doi.org/10.1007/978-1-4614-7750-1_40

[5] Das, S. and Sarkar, N. (2010) Is the Relative Risk Aversion Parameter Constant over Time? A Multi-Country Study. *Empirical Economics*, **38**, 605-617. http://dx.doi.org/10.1007/s00181-009-0281-y

[6] Chou, R., Engle, R.F. and Kane, A. (1992) Measuring Risk Aversion from Excess Returns on a Stock Index. *Journal of Econometrics*, **52**, 201-224. http://dx.doi.org/10.1016/0304-4076(92)90070-8

[7] Fama, E.F. and French, K.R. (1989) Business Conditions and Expected Returns on Stock and Bonds. *Journal of Financical Economics*, **25**, 23-49. http://dx.doi.org/10.1016/0304-405X(89)90095-0

[8] Backus, D. and Gregory, A. (1993) Theoretical Relations between Risk Premiums and Conditional Variances. *Journal of Business and Economic Statistics*, **11**, 177-185.

[9] Glosten, L.R., Jagannathan, R. and Runkle, D.E. (1993) On the Relationship between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, **48**, 1779-1801. http://dx.doi.org/10.1111/j.1540-6261.1993.tb05128.x

[10] Ling, S. (2004) Estimation and Testing of Stationarity for Double Autoregressive Models. *Journal of Royal Statistical Society B*, **66**, 63-78. http://dx.doi.org/10.1111/j.1467-9868.2004.00432.x

[11] Ling, S. (2007) A Double AR (*p*) Model: Structure and Estimation. *Statistica Sinica*, **17**, 161-175.

[12] Zhang, X., Wong, H., Li, Y. and Ip, W. (2011) A Class of Threshold Autoregressive Conditional Heteroscedastic Models. *Statistics and Its Interface*, **4**, 149-158. http://dx.doi.org/10.4310/SII.2011.v4.n2.a10

[13] Xiong, Q., Li, Y. and Zhang, X. (2014) The Profile Likelihood Estimation for Single-Index ARCH(*p*)-M Model. *Mathematical Problems in Engineering*, 2014, Article ID: 189426. http://dx.doi.org/10.1155/2014/189426

[14] Fan, J. and Yao, Q. (2003) Nonlinear Time Series: Nonparametric and Parametric Methods. Springer, New York. http://dx.doi.org/10.1007/b97702

[15] Christensen, B.J., Dahl, C.M. and Iglesias, E.M. (2012) Semiparametric Inference in a GARCH-in-Mean Model. *Journal of Econometrics*, **167**, 458-472. http://dx.doi.org/10.1016/j.jeconom.2011.09.028

[16] Yang, L. (2006) A Semiparametric GARCH Model for Foreign Exchange Volatility. *Journal of Econometrics*, **130**, 365-384. http://dx.doi.org/10.1016/j.jeconom.2005.03.006

[17] Andrews, D.W.K. (1992) Generic Uniform Convergence. *Econometric Theory*, **8**, 241-257. http://dx.doi.org/10.1017/S0266466600012780

[18] Kosorok, M.R. (2006) Introduction to Empirical Processes and Semiparametric Inference. Springer, New York.

[19] Billingsley, P. (1995) Probability and Measure. 3rd Edition, Wiley, New York.

Appendix

Proof of Lemma 1

Proof. We only show the case of $l = m = 2$. Other situations can be proved by similar argument. Let $\sigma_0(u) \equiv 1$ and $\hat{\sigma}_0(u) \equiv 1$. Then $m(\theta,u)$ can be written as $m(\theta,u) = g(u) \big/ \sum_{k=0}^{p} a_k \sigma_k(u)$, $\hat{m}(\theta,u)$ can be written as $\hat{m}(\theta,u) = \hat{g}(u) \big/ \sum_{k=0}^{p} a_k \hat{\sigma}_k(u)$. Noting, for $i = 0,1,\cdots,p$, $\partial a_k / \partial \theta_i$ equals 1 when $i = k$, and 0 for other cases. Then it is easy to have

$$\begin{aligned}
\frac{\partial m(\theta,u)}{\partial \theta_i} &= \frac{g(u)\sigma_i(u)}{\left[\sum_{k=0}^{p} a_k \sigma_k(u)\right]^2} \\
\frac{\partial^2 m(\theta,u)}{\partial \theta_i \partial \theta_j} &= \frac{-2g(u)\sigma_i(u)\sigma_j(u)}{\left[\sum_{k=0}^{p} a_k \sigma_k(u)\right]^3} \\
\frac{\partial^2 \hat{m}(\theta,u)}{\partial \theta_i \partial \theta_j} &= \frac{-2\hat{g}(u)\hat{\sigma}_i(u)\hat{\sigma}_j(u)}{\left[\sum_{k=0}^{p} a_k \hat{\sigma}_k(u)\right]^3}
\end{aligned} \tag{A.1}$$

Hence,

$$\begin{aligned}
&\left| \frac{\partial^2 m(\theta,u)}{\partial \theta_i \partial \theta_j} - \frac{\partial^2 \hat{m}(\theta,u)}{\partial \theta_i \partial \theta_j} \right| \\
&\le \left| 2\hat{g}(u)\hat{\sigma}_i(u)\hat{\sigma}_j(u) \left(1 \Big/ \left[\sum_{k=0}^{p} a_k \hat{\sigma}_k(u)\right]^3 - 1 \Big/ \left[\sum_{k=0}^{p} a_k \sigma_k(u)\right]^3 \right) \right| \\
&\quad + \left| 1 \Big/ \left[\sum_{k=0}^{p} a_k \sigma_k(u)\right]^3 \times \left[2\hat{g}(u)\hat{\sigma}_i(u)\hat{\sigma}_j(u) - 2g(u)\sigma_i(u)\sigma_j(u) \right] \right|
\end{aligned} \tag{A.2}$$

According to Assumption.6, it is easy to obtain the following equalities:

$$\begin{aligned}
&\sup_{u \in A} \left[2\hat{g}(u)\hat{\sigma}_i(u)\hat{\sigma}_j(u) - 2g(u)\sigma_i(u)\sigma_j(u) \right] = O_p\left(b_n^2 + \{ nb_n / \log n \}^{-1/2} \right), \\
&\sup_{u \in A} \left[\frac{1}{\left[\sum_{k=0}^{p} a_k \hat{\sigma}_k(u)\right]^3} - \frac{1}{\left[\sum_{k=0}^{p} a_k \sigma_k(u)\right]^3} \right] = O_p\left(b_n^2 + \{ nb_n / \log n \}^{-1/2} \right),
\end{aligned} \tag{A.3}$$

Note that $\hat{g}(u)\hat{\sigma}_i(u)\hat{\sigma}_j(u) = O_p(1)$ and $\min\left\{ \sum_{k=0}^{p} a_k \sigma_k(u), \sum_{k=0}^{p} a_k \hat{\sigma}_k(u) \right\} \ge a_0 \ge a_0^U > 0$ implying $1 \Big/ \left[\sum_{k=0}^{p} a_k \sigma_k(u) \right]^3 = O(1)$. Then Equation (11) follows from (A.2)-(A.3).

Proof of Lemma 2

Proof. We only consider the case of $k = 2$, other cases can be obtained with similar and easier arguments. From (5)-(6),

$$\hat{\varepsilon}_t(\theta) = \varepsilon_t(\theta) - h_t(\theta)\left[\hat{m}(\theta, U_t) - m(\theta, U_t) \right]. \tag{A.4}$$

Further,

$$\begin{aligned}
&\hat{\varepsilon}_t^2(\theta)-\varepsilon_t^2(\theta)=h_t^2(\theta)\left[\hat{m}(\theta,U_t)-m(\theta,U_t)\right]^2+2\varepsilon_t(\theta)h_t(\theta)\left[m(\theta,U_t)-\hat{m}(\theta,U_t)\right]\\
&\frac{\partial\hat{\varepsilon}_t(\theta)}{\partial\theta}-\frac{\partial\varepsilon_t(\theta)}{\partial\theta}=h_t(\theta)\left(\frac{\partial m(\theta,u)}{\partial\theta}-\frac{\partial\hat{m}(\theta,u)}{\partial\theta}\right)+\frac{\partial h_t(\theta)}{\partial\theta}\left[m(\theta,U_t)-\hat{m}(\theta,U_t)\right]\\
&\frac{\partial^2\hat{\varepsilon}_t(\theta)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\tau}=\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta^\tau}\left[m(\theta,U_t)-\hat{m}(\theta,U_t)\right]+h_t(\theta)\left(\frac{\partial^2 m(\theta,u)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2\hat{m}(\theta,u)}{\partial\theta\partial\theta^\tau}\right)\\
&\qquad+\left(\frac{\partial m(\theta,u)}{\partial\theta}-\frac{\partial\hat{m}(\theta,u)}{\partial\theta}\right)\frac{\partial h_t(\theta)}{\partial\theta^\tau}+\frac{\partial h_t(\theta)}{\partial\theta}\left(\frac{\partial m(\theta,u)}{\partial\theta^\tau}-\frac{\partial\hat{m}(\theta,u)}{\partial\theta^\tau}\right)
\end{aligned}\tag{A.5}$$

Let

$$l_t(\theta)=\log h_t(\theta)+\frac{\varepsilon_t^2(\theta)}{h_t(\theta)},\hat{l}_t(\theta)=\log h_t(\theta)+\frac{\hat{\varepsilon}_t^2(\theta)}{h_t(\theta)}.\tag{A.6}$$

Then,

$$L_n(\theta)=\sum_{t=1}^n l_t(\theta)\pi(U_t),\hat{L}_n(\theta)=\sum_{t=1}^n\hat{l}_t(\theta)\pi(U_t).\tag{A.7}$$

We can further have

$$\frac{\partial l_t(\theta)}{\partial\theta}=\left(1-\frac{\varepsilon_t^2(\theta)}{h_t(\theta)}\right)\frac{1}{h_t(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}+\frac{2\varepsilon_t(\theta)}{h_t(\theta)}\frac{\partial\varepsilon_t(\theta)}{\partial\theta},\tag{A.8}$$

$$\begin{aligned}
\frac{\partial^2 l_t(\theta)}{\partial\theta\partial\theta^\tau}=&-\frac{1}{h_t^2(\theta)}\left(1-\frac{2\varepsilon_t^2(\theta)}{h_t(\theta)}\right)\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta^\tau}-\frac{2\varepsilon_t(\theta)}{h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial\varepsilon_t(\theta)}{\partial\theta^\tau}\\
&+\frac{2}{h_t(\theta)}\frac{\partial\varepsilon_t(\theta)}{\partial\theta}\frac{\partial\varepsilon_t(\theta)}{\partial\theta^\tau}+\frac{2\varepsilon_t(\theta)}{h_t(\theta)}\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\tau}\\
&-\frac{2\varepsilon_t(\theta)}{h_t^2(\theta)}\frac{\partial\varepsilon_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta^\tau}+\frac{1}{h_t(\theta)}\left(1-\frac{\varepsilon_t^2(\theta)}{h_t(\theta)}\right)\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta^\tau}.
\end{aligned}\tag{A.9}$$

From (A.9), $\partial^2\hat{l}_t(\theta)/\partial\theta\partial\theta^\tau$ can be easily obtained by replacing $\varepsilon_t(\theta)$ with $\hat{\varepsilon}_t(\theta)$. Then

$$\frac{\partial^2\hat{L}_n(\theta)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2 L_n(\theta)}{\partial\theta\partial\theta^\tau}=\sum_{l=1}^6 I_{ln}(\theta).\tag{A.10}$$

Here,

$$I_{1n}(\theta)=\frac{1}{n}\sum_{t=1}^n\pi(U_t)\frac{2}{h_t^3(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\frac{\partial h_t(\theta)}{\partial\theta^\tau}\left[\hat{\varepsilon}_t^2(\theta)-\varepsilon_t^2(\theta)\right]\tag{A.11}$$

$$I_{2n}(\theta)=\frac{1}{n}\sum_{t=1}^n\pi(U_t)\frac{2}{h_t^2(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\left[\hat{\varepsilon}_t(\theta)\frac{\partial\hat{\varepsilon}_t(\theta)}{\partial\theta^\tau}-\varepsilon_t(\theta)\frac{\partial\varepsilon_t(\theta)}{\partial\theta^\tau}\right]\tag{A.12}$$

$$I_{3n}(\theta)=\frac{1}{n}\sum_{t=1}^n\pi(U_t)\frac{2}{h_t(\theta)}\left[\frac{\hat{\varepsilon}_t(\theta)}{\partial\theta}\frac{\partial\hat{\varepsilon}_t(\theta)}{\partial\theta^\tau}-\frac{\varepsilon_t(\theta)}{\partial\theta}\frac{\partial\varepsilon_t(\theta)}{\partial\theta^\tau}\right]\tag{A.13}$$

$$I_{4n}(\theta)=\frac{1}{n}\sum_{t=1}^n\pi(U_t)\frac{2}{h_t(\theta)}\left[\hat{\varepsilon}_t(\theta)\frac{\partial^2\hat{\varepsilon}_t(\theta)}{\partial\theta\partial\theta^\tau}-\varepsilon_t(\theta)\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\tau}\right]\tag{A.14}$$

$$I_{5n}(\theta)=\frac{1}{n}\sum_{t=1}^n\pi(U_t)\frac{2}{h_t^2(\theta)}\left[\hat{\varepsilon}_t(\theta)\frac{\partial\hat{\varepsilon}_t(\theta)}{\partial\theta}-\varepsilon_t(\theta)\frac{\partial\varepsilon_t(\theta)}{\partial\theta}\right]\frac{\partial h_t(\theta)}{\partial\theta^\tau}\tag{A.15}$$

$$I_{6n}(\theta)=\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)\frac{2}{h_t^2(\theta)}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta^\tau}\left[\hat{\varepsilon}_t^2(\theta)-\varepsilon_t^2(\theta)\right]. \tag{A.16}$$

Note $I_{5n}(\theta)=I_{2n}(\theta)^\tau, I_{6n}(\theta)=0$ because of $\partial^2 h_t(\theta)/\partial\theta\partial\theta^\tau=0$. Hence to show (12), it suffices to prove $\sup_{\theta\in\Theta}\|I_{ln}(\theta)\|=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right), l=1,\cdots,4$. To save space, we only give detailed proof of $\sup_{\theta\in\Theta}\|I_{4n}(\theta)\|=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right)$. It is easy to have

$$\begin{aligned}I_{4n}(\theta)=&\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)\frac{2}{h_t(\theta)}\hat{\varepsilon}_t(\theta)\left[\frac{\partial^2\hat{\varepsilon}_t(\theta)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\tau}\right]\\&+\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)\frac{2}{h_t(\theta)}\left[\hat{\varepsilon}_t(\theta)-\varepsilon_t(\theta)\right]\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\tau}.\end{aligned}$$

In terms of (A.4)-(A.5), $I_{4n}(\theta)$ can be written as

$$\begin{aligned}I_{4n}(\theta)=&\frac{1}{n}\sum_{t=1}^{n}\left\{\pi(U_t)\frac{2}{h_t(\theta)}\left[\varepsilon_t(\theta)+h_t(\theta)\left(m(\theta,U_t)-\hat{m}(\theta,U_t)\right)\right]\right.\\&\times\left[h_t(\theta)\left(\frac{\partial^2 m(\theta,u)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2\hat{m}(\theta,u)}{\partial\theta\partial\theta^\tau}\right)+\left(\frac{\partial m(\theta,u)}{\partial\theta}-\frac{\partial\hat{m}(\theta,u)}{\partial\theta}\right)\frac{\partial h_t(\theta)}{\partial\theta^\tau}\right.\\&\left.\left.+\frac{\partial h_t(\theta)}{\partial\theta}\left(\frac{\partial m(\theta,u)}{\partial\theta^\tau}-\frac{\partial\hat{m}(\theta,u)}{\partial\theta^\tau}\right)\right]\right\}\\&+\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)\frac{2}{h_t(\theta)}\left[h_t(\theta)\left(m(\theta,U_t)-\hat{m}(\theta,U_t)\right)\right]\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\tau}.\end{aligned} \tag{A.17}$$

Without loss of generality, there exists a $\theta^U\in\Theta$ such that $h_t(\theta^U)\ge h_t(\theta),\theta\in\Theta$. According to (5), Assumptions 2 and 5, when $U_t\in A$,

$$\begin{aligned}|\varepsilon_t(\theta)|&=|y_t-m(\theta,U_t)h_t(\theta)|\\&=\left|m(\theta_0,U_t)h_t(\theta_0)+e_t\sqrt{h_t(\theta_0)}-m(\theta,U_t)h_t(\theta)\right|\\&\le M\left(h_t(\theta^U)+e_t^2\right).\end{aligned} \tag{A.18}$$

The last inequality comes from the fact $m(\theta,u)$ is uniformly bounded for $\theta\in\Theta,u\in A$. Similarly, when $U_t\in A$ we can show

$$\left\|\frac{1}{h_t(\theta)}\frac{\partial h_t(\theta)}{\partial\theta}\right\|=O(1),\left\|\frac{1}{h_t(\theta)}\frac{\partial^2\varepsilon_t(\theta)}{\partial\theta\partial\theta^\tau}\right\|=O(1). \tag{A.19}$$

From Lemma 1, it follows that

$$\begin{aligned}T_0(n)&:=\sup_{\theta\in\Theta}\sup_{u\in A}\|m(\theta,u)-\hat{m}(\theta,u)\|=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right),\\T_1(n)&:=\sup_{\theta\in\Theta}\sup_{u\in A}\left\|\frac{\partial m(\theta,u)}{\partial\theta}-\frac{\partial\hat{m}(\theta,u)}{\partial\theta}\right\|=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right),\\T_2(n)&:=\sup_{\theta\in\Theta}\sup_{u\in A}\left\|\frac{\partial^2 m(\theta,u)}{\partial\theta\partial\theta^\tau}-\frac{\partial^2\hat{m}(\theta,u)}{\partial\theta\partial\theta^\tau}\right\|=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right).\end{aligned} \tag{A.20}$$

(A.17)-(A.20) gives

$$\begin{aligned}\sup_{\theta\in\Theta}\left\|I_{4n}(\theta)\right\| \le{}& M\left[T_2(n)+2T_1(n)\right]\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)\left[h_t(\theta^U)+e_t^2\right]\\ &+MT_0(n)\left[T_2(n)+2T_1(n)\right]\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)h_t(\theta^U)\\ &+MT_0(n)\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)h_t(\theta^U).\end{aligned} \tag{A.21}$$

Note that e_t is independent of U_t and $E\left[\pi(U_t)h_t(\theta^U)\right]=E\left[\pi(U_t)E\left(h_t(\theta^U)\,|\,U_t\right)\right]$. Based on Assumption 3, we have

$$\begin{aligned}&\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)h_t(\theta^U)\xrightarrow{p}E\left[\pi(U_t)\sigma_{10}(\theta^U,U_t)\right]<\infty,\\ &\frac{1}{n}\sum_{t=1}^{n}\pi(U_t)e_t^2\xrightarrow{p}E\left[\pi(U_t)\right]<\infty.\end{aligned} \tag{A.22}$$

(A.20)-(A.22) implies $\sup_{\theta\in\Theta}\left\|I_{4n}(\theta)\right\|=O_p\left(b_n^2+\{nb_n/\log n\}^{-1/2}\right)$.