GARCH Process with Persistent Covariates

Heejoon Han[1]
Department of Economics and Risk Management Institute,
National University of Singapore

Abstract

The paper considers a volatility model that includes a persistent, integrated or nearly integrated, covariate in the standard GARCH(1,1) model. First, we establish the asymptotic distribution theory of the quasi-maximum likelihood estimator (QMLE) for such a model. In general, the QMLE is consistent, and its limit distribution is non-Gaussian and represented as a functional of Brownian motions. However, it becomes Gaussian if the covariate is strictly exogenous or the volatility function is linear in parameter. Second, we provide asymptotic theories, which show that the model successfully explains stylized facts of financial time series, such as long memory property in volatility, IGARCH and leptokurtosis. The autocorrelation of the squared process of the model generates the long memory property following a trend commonly observed in real data. Our theory of misspecification shows that IGARCH could be the result of missing a relevant persistent covariate in the GARCH(1,1) process. Moreover, the kurtosis of the model is randomly bigger than the kurtosis of the GARCH(1,1) model.

This version: December 2007
JEL classification: C22, C50, G12

Keywords; GARCH, persistent covariate, quasi-maximum likelihood estimator, asymptotic distribution theory, long memory property, IGARCH, leptokurtosis

[1]I would like to thank Joon Y. Park and conference participants at the 18th EC^2 meeting on advances in econometric time series analysis (Faro), the third symposium on econometric theory and applications (SETA2007, Hong Kong) and 2007 Korean econometric society meeting (Seoul) for helpful comments and suggestions. I also thank Jeff Fleming for providing the data for firms in the MMI. The research for this paper was supported by the NUS Risk Management Institute. Address correspondence: Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570; phone: (65) 6516-6258; fax: (65) 6775-2646; e-mail: ecshhj@nus.edu.sg

1 Introduction

ARCH type models have been widely used to model the volatility of economic and financial time series since the seminal work by Engle (1982) and the extension made by Bollerslev (1986). However, since most ARCH type models have been univariate, relating the volatility of time series only to the information contained in its own past history, they shed little light on an economic source of volatility.[2]

One of the simplest ways to consider an economic source of volatility is to include economic variables in volatility models. Thus there have been works that consider ARCH type models with macroeconomic variables as covariates. Since the GARCH(1,1) model has been popular and many macroeconomic variables are persistent, these works use the GARCH(1,1) model with a persistent covariate as following;

$$\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + \pi \left| x_{t-1} \right| \text{ (or } \pi x_{t-1}^2\text{)},$$

where (y_t) is a demeaned time series and σ_t^2 is its variance conditional on the information available at time $t-1$. As a covariate (x_t), Glosten et al (1993) and Engle and Patton (2001) used three-month U.S. Treasury bill rates for stock return volatility, and Gray (1996) used the level of interest rates in his generalized regime-switching model of short-term interest rates. Likewise, forward-spot spreads and interest rate differentials between countries were used as covariates respectively by Hodrick (1989) and Hagiwara and Herce (1999) for exchange rate return volatility. Note that these covariates are persistent; interest rates and interest rate differentials between countries may well be modeled as time series with a unit root or a near unit root.

Like most empirical applications of ARCH type models, the quasi-maximum likelihood estimation procedure was used in the literature to estimate the GARCH(1,1) model with a persistent covariate, which has either a unit root or a near unit root. However, there is no existing literature that provides the asymptotic distribution theory of the quasi-maximum likelihood estimator (QMLE) in such a model. The asymptotic distribution theory of the QMLE in the GARCH(1,1) model, stationary GARCH(1,1) as well as IGARCH(1,1), was formally established by Lumsdaine (1996) and Lee and Hansen (1994). Their theories depend on covariance-stationarity or strict stationarity (in case of IGARCH(1,1)) of the GARCH(1,1) process. However, their theories are not valid for the GARCH(1,1) process with a persistent covariate because the process is not stationary due to a nonstationary covariate. Recently, Jensen and Rahbek (2004) established the asymptotic distribution theory of the QMLE in the GARCH(1,1) model with the entire parameter region, including nonstationary behavior. However, their theory also cannot be applied in the GARCH(1,1) model with a persistent covariate.

Since there is no existing literature which investigates the statistical properties of the GARCH(1,1) model with an additional persistent covariate, this paper attempts to fill this

[2] As noted by Engle and Rangel (2005), "After more than 25 years of research on volatility, the central unsolved problem is the relation between the state of the economy and aggregate financial volatility. The number of models that have been developed to predict volatility based on time series information is astronomical, but the models that incorporate economic variables are hard to find."

gap. Recently, Han and Park (2006, 2007) investigated the effect of a persistent covariate in the ARCH type models. However, they examined a simpler model: the ARCH(1) model with a persistent covariate. Considering that a covariate is added predominantly in the GARCH(1,1) model in the literature, we extend their model and investigate the GARCH(1,1) process with a persistent covariate in this paper. For asymptotic theories, we generalize the model and consider the GARCH(1,1) process with a nonlinear function of an integrated or a nearly integrated variable. Therefore, our theories are based on the study of the nonlinear regressions with integrated time series by Park and Phillips (2001).

The contributions of this paper are following. First, we establish the consistency of the QMLE in our model, and we derive its asymptotic distribution theory. In general, the limit distribution of the QMLE is non-Gaussian and represented as a functional of Brownian motions as in other models involving unit root or near unit root processes. However, the limit distribution reduces to be Gaussian in some special cases. In particular, it becomes mixed normal when the covariate is strictly exogenous, being independent of the innovation process of the volatility model. Furthermore, it becomes normal, if the volatility function is specified as being linear in parameter. In these cases, the standard inference is valid and applicable even in the presence of a persistent covariate. Thus, if we simply choose a linear function for a positive covariate, which is the most common way to add a covariate in the GARCH(1,1) model in the literature, then the standard inference is valid.

Second, we investigate if our model can explain stylized facts of financial time series. We consider three commonly observed facts in stock and exchange rate returns, which are long memory property in volatility, IGARCH and leptokurtosis. Our asymptotic theories show that the model successfully explains them due to a persistent covariate. Regarding the long memory property in volatility, it is known that the sample autocorrelation function of squared financial return series (in particular high frequency data) decreases fast at first and remains significantly positive for larger lags. The asymptotic behavior of the sample autocorrelation function of the squared process generated by our model describes this trend and generates the long memory property in volatility; the autocorrelation decreases exponentially at first and finally converges to a positive random limit that is smaller than unity.

Additionally, it is well known that most empirical applications of the GARCH(1,1) model on financial return series suggest the IGARCH(1,1) process if the sample size is sufficiently large. Our theory of misspecification shows that IGARCH could be the result of missing a relevant persistent covariate in the GARCH(1,1) process. Moreover, it is also known that the kurtosis implied by the GARCH(1,1) model with normally distributed innovations tends to be far less than the sample kurtosis observed in many financial return series. The asymptotic limit of the sample kurtosis in our model is randomly bigger than the kurtosis of the GARCH(1,1) model. Hence, our model provides an explanation of the sample kurtosis observed in financial return series without using innovations with fat-tailed distributions.

The rest of the paper is organized as follows. Section 2 introduces the model and assumptions. The asymptotics of the QMLE in our model is presented in Section 3. Section 4 investigates how our model explains the stylized facts of financial time series. Section 5 concludes the paper, and Appendices A and B contain mathematical proofs for the technical results in the paper. Finally a word on notation. We denote by $\mathbb{R}_+$ and $\mathbb{R}_{++}$ the sets of real numbers that are nonnegative and positive, respectively, and $\|\cdot\|$ signifies the usual Euclidean norm for $\mathbb{R}^n$. Standard terminologies and notations in probability and measure theory are used throughout the paper. In particular, notations for various convergence such as $\to_{a.s.}$, $\to_p$ and $\to_d$ will frequently appear. All limits are taken as $n \to \infty$, except where otherwise indicated.

2 The Model and Assumptions

We consider the volatility model specified as

$$y_t = \sigma_t \varepsilon_t,$$

and let $(\mathcal{F}_t)$ be the filtration representing the information available at time t.

Assumption 1 Assume that
(a) (ε_t) is iid (0,1) and adapted to $(\mathcal{F}_t)$
(b) (σ_t) is given by

$$\sigma_t^2 = \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 + f(x_t, \pi) \tag{1}$$

for parameters $\alpha, \beta \in \mathcal{A} \subset \mathbb{R}_+$ and $\pi \in \Pi \subset \mathbb{R}^d$ with $\mathcal{A}$ and Π compact and a nonnegative volatility function $f : \mathbb{R} \times \Pi \to \mathbb{R}_+$, and let

$$x_t = \left(1 - \frac{c}{n}\right) x_{t-1} + v_t$$

assuming that (x_t) is adapted to $(\mathcal{F}_{t-1})$ and $c \geq 0$.

Under Assumption 1, (y_t) becomes the GARCH(1,1) process with a persistent, integrated or nearly integrated, covariate. If $\alpha = \beta = 0$ in (1), then the volatility model is given by $\sigma_t^2 = f(x_t)$, which is referred to as the *Nonstationary Nonlinear Heteroskedasticity* (NNH) model. Park (2002) introduced this model but he considered only the case in which (x_t) has an exact unit root. If $\beta = 0$ in (1), it is the ARCH(1) process with a persistent covariate. Han and Park (2006, 2007) investigated this process; see Han and Park (2006) for the time series properties, empirical applications and forecast evaluation, and refer to Han and Park (2007) for the asymptotic distribution theory of the QMLE and an explanation of IGARCH.

Assumption 2 Assume that
(a) (v_t) is generated by

$$v_t = \varphi(L)\eta_t = \sum_{k=0}^{\infty} \varphi_k \eta_{t-k},$$

where $\varphi_0 = 1$, $\varphi(1) \neq 0$ with $\sum_{k=0}^{\infty} k|\varphi_k| < \infty$, and (η_t) are iid random variables with mean zero and $\mathbb{E}|\eta_t|^p < \infty$ for some $p > 2$,
(b) $\mathbb{E}|\varepsilon_t|^q < \infty$ and $\mathbb{E}\left(\beta + \alpha\varepsilon_t^2\right)^{q/2} < 1$ for some $q > 4$, and
(c) $1/p + 2/q < 1/2$.

Assumption 2 more precisely defines the covariate (x_t) as an integrated or a near-integrated process driven by a general linear process, and introduces moment conditions for the innovation sequences (η_t) and (ε_t). Throughout the paper, we set the long-run variance of (v_t) to be unity because it has only an unimportant scaling effect on our analysis.

The conditions in part (a) are standard. For the conditions in part (b), note that

$$\mathbb{E}\left(\ln\left(\beta+\alpha\varepsilon_t^2\right)\right) \leq \ln\left[\mathbb{E}\left(\beta+\alpha\varepsilon_t^2\right)\right] \leq \frac{2}{q}\ln\left[\mathbb{E}\left(\beta+\alpha\varepsilon_t^2\right)^{q/2}\right]$$

for any $q>2$, by the successive applications of Jensen's inequality. As a result, it follows from part (b) that $\mathbb{E}\left(\ln\left(\beta+\alpha\varepsilon_t^2\right)\right)<0$, which as shown in Nelson (1990) is the necessary and sufficient condition for the GARCH(1,1) process (without covariates) to be strictly stationary and ergodic.[3] Therefore, part (b) implies in particular that the GARCH component of our model is strictly stationary and ergodic. Moreover, note that it also follows from part (b) that

$$\mathbb{E}\left(\beta+\alpha\varepsilon_t^2\right)^2 \leq \left(\mathbb{E}\left(\beta+\alpha\varepsilon_t^2\right)^{q/2}\right)^{4/q} < 1 \tag{2}$$

by Jensen's inequality. This condition is necessary in the investigation of the statistical properties of the GARCH(1,1) process. If $\mathbb{E}\left(\beta+\alpha\varepsilon_t^2\right)^2<1$, the sample autocorrelation of squared process and the sample kurtosis have probability limits in the GARCH(1,1) model. Due to part (c), the conditions in part (b) should hold for large q if p is small in part (a). As $p\to\infty$, we may choose q arbitrarily close to 4.

Now we introduce the conditions for volatility function f. Throughout the paper, we assume that f is asymptotically homogeneous in the sense of Park and Phillips (2001). Roughly, f is given as

$$f(\lambda x,\pi)=\kappa(\lambda,\pi)\bar{f}(x,\pi)+o(\kappa(\lambda,\pi))$$

as $\lambda\to\infty$, over any compact interval for x and uniformly in π over the parameter set Π. We call κ and $\bar{f}$ the asymptotic order and limit homogeneous function of f respectively. The reader is referred to Park and Phillips (2001) for more precise definitions. They show that many functions that are commonly used in nonlinear econometric models are asymptotically homogeneous, including the logarithmic function, logistic function (and other distribution function-like functions) and power function.

The volatility function that is of primary interest is given by

$$f(x,\pi)=\pi_1|x|^{\pi_2} \tag{3}$$

with $\pi=(\pi_1,\pi_2)'\in\Pi\subset\mathbb{R}^2_{++}$. The function f in (3) is often called the constant elasticity of variance (CEV) volatility, since it implies that the elasticity of variance is constant. It is easy to see that the CEV volatility function is asymptotically homogeneous with asymptotic order $\kappa(\lambda,\pi)=\pi_1\lambda^{\pi_2}$ and limit homogeneous function $\bar{f}(x,\pi)=|x|^{\pi_2}$. In what follows, we

[3] Most works pertaining to the asymptotic distribution theory of the QMLE in GARCH models depend on this finding. See Lee and Hansen (1994), Lumsdaine (1996) and Berkes *et al* (2003). Recently, Jensen and Rahbek (2004) established the asymptotic distribution theory of the QMLE in the nonstationary GARCH(1,1) model under $\mathbb{E}\left[\ln\left(\beta+\alpha\varepsilon_t^2\right)\right]\geq 0$.

say that the two asymptotically homogeneous functions are 'equivalent' if they are different only over a compact interval. Clearly, the equivalent asymptotically homogeneous functions have the same asymptotic order and limit homogeneous function.

Assumption 3 The volatility function f satisfies the following conditions:
(a) f is asymptotically homogeneous with asymptotic order κ such that $\inf_{\pi\in\Pi}\kappa(\lambda,\pi)\to\infty$ as $\lambda\to\infty$, and
(b) f has an equivalent asymptotically homogeneous function $\tilde{f}$, say, such that

$$|\tilde{f}(x,\pi)-\tilde{f}(y,\pi)|\leq\nabla f(z,\pi)|x-y|$$

for any x,y in a compact subset of $\mathbb{R}$, where $x\leq z\leq y$ and ∇f is asymptotically homogeneous with asymptotic order $\nabla\kappa$ such that

$$\nabla\kappa(\lambda,\pi)=O(\kappa(\lambda,\pi)/\lambda)$$

for all $\pi\in\Pi$, as $\lambda\to\infty$.

It is not difficult to see that the conditions in Assumption 3 hold for the CEV volatility function introduced in (3). See Han and Park (2007) for details. The introduction of an equivalent asymptotically homogeneous function in part (b) allows us to consider the volatility function, which is not smooth in every point of its domain. For the sake of brevity in exposition, we assume in what follows that the local smoothing has already been done and f itself satisfies the condition in part (b). Moreover, we define $\dot{f}=\partial f/\partial\pi$ and denote the asymptotic order and limit homogeneous function of asymptotically homogeneous $\dot{f}$ by $\dot{\kappa}$ and $\bar{f}^{\cdot}$, respectively. For notational simplicity, we write $\kappa_0(\lambda)=\kappa(\lambda,\pi_0)$, $\dot{\kappa}_0(\lambda)=\dot{\kappa}(\lambda,\pi_0)$, $f_0(x)=f\left(x,\pi_0\right)$, $\bar{f}_0(x)=\bar{f}(x,\pi_0)$ and $\bar{f}_0^{\cdot}(x)=\bar{f}^{\cdot}(x,\pi_0)$.

Now we define

$$z_t=1+\sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\beta+\alpha\varepsilon_{t-i}^2\right) \tag{4}$$

for $t\geq 1$. Note that (z_t) appears in the analysis of the GARCH(1,1) model. In fact, if (y_t) is generated as the GARCH(1,1) process with its conditional variance given by

$$\sigma_t^2=\omega+\alpha y_{t-1}^2+\beta\sigma_{t-1}^2, \tag{5}$$

then we may easily deduce by the recursive substitution that

$$\sigma_t^2=\omega\left(1+\sum_{k=1}^{t-1}\prod_{i=1}^{k}\left(\beta+\alpha\varepsilon_{t-i}^2\right)\right)+\sigma_0^2\prod_{i=1}^{t}\left(\beta+\alpha\varepsilon_{t-i}^2\right)$$

for $t\geq 1$. As Nelson (1990) noted, we can close the system either by defining a probability measure μ_0 for the starting value of σ_0^2 or by assuming that the system extends infinitely far into the past. In the latter case, we have

$$\sigma_t^2=\omega z_t.$$

3 Asymptotic Theory of QMLE

In this section, we derive the asymptotic theory of the QMLE for our model. Let $\theta = (\alpha, \beta, \pi')'$ with the true value $\theta_0 = (\alpha_0, \beta_0, \pi_0')'$ and the parameter set $\Theta = \mathcal{A} \times \Pi$, and let $\sigma_t^2(\theta) = \alpha y_{t-1}^2 + \beta\sigma_{t-1}^2 + f(x_t, \pi)$. Also, we denote by $\ell_t(\theta)$ the conditional log-likelihood for y_t given $\mathcal{F}_{t-1}$ for $t = 1, \ldots, n$. Then the log-likelihood function of the entire sample $(y_1, \ldots, y_n)$ is given by

$$\sum_{t=1}^{n} l_t(\theta) = -\frac{1}{2}\sum_{t=1}^{n}\left(\ln\sigma_t^2 + \frac{y_t^2}{\sigma_t^2}\right)$$

ignoring the unimportant constant term, and the QMLE $\hat{\theta}_n$ is defined as

$$\hat{\theta}_n = \underset{\theta\in\Theta}{\operatorname{argmax}}\sum_{t=1}^{n}\ell_t(\theta).$$

Let $\sigma_t^2 = \sigma_t^2(\theta)$. Then the score vector $s_n(\theta)$ and Hessian matrix $H_n(\theta)$ are given by

$$s_n(\theta) = \sum_{t=1}^{n}\frac{\partial\ell_t(\theta)}{\partial\theta} = \frac{1}{2}\sum_{t=1}^{n}\left(\frac{y_t^2}{\sigma_t^2} - 1\right)\frac{1}{\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta}$$

$$H_n(\theta) = \sum_{t=1}^{n}\frac{\partial^2\ell_t(\theta)}{\partial\theta\partial\theta'} = \frac{1}{2}\sum_{t=1}^{n}\left[\left(1 - \frac{2y_t^2}{\sigma_t^2}\right)\frac{1}{\sigma_t^4}\frac{\partial\sigma_t^2}{\partial\theta}\frac{\partial\sigma_t^2}{\partial\theta'} + \left(\frac{y_t^2}{\sigma_t^2} - 1\right)\frac{1}{\sigma_t^2}\frac{\partial^2\sigma_t^2}{\partial\theta\partial\theta'}\right].$$

The asymptotics of $\hat{\theta}_n$ in our model can be obtained from the first order Taylor expansion of the score vector, i.e.,

$$s_n(\hat{\theta}_n) = s_n(\theta_0) + H_n(\theta_n)(\hat{\theta}_n - \theta_0), \tag{6}$$

where θ_n lies in the line segment connecting $\hat{\theta}_n$ and θ_0. If $\hat{\theta}_n$ is an interior solution, we have $s_n(\hat{\theta}_n) = 0$. Therefore, we may write from (6)

$$\nu_n'(\hat{\theta}_n - \theta_0) = -\left[\nu_n^{-1}H_n(\theta_n)\nu_n^{-1\prime}\right]^{-1}\left[\nu_n^{-1}s_n(\theta_0)\right] \tag{7}$$

for an appropriately defined sequence (ν_n) of normalization matrix.

The following conditions ML1-ML3 are sufficient to derive the asymptotics for $\hat{\theta}_n$ upon appropriately choosing the sequence (ν_n) of normalization matrix.

ML1: $\nu_n^{-1}s_n(\theta_0) \to_d P$ for some P,

ML2: $-\nu_n^{-1}H_n(\theta_0)\nu_n^{-1\prime} \to_d Q$ for some $Q > 0$ a.s., and

ML3: there exists a sequence (μ_n) of invertible normalization matrices such that $\mu_n\nu_n^{-1} \to 0$, and such that

$$\sup_{\theta\in N_n}\left\|\mu_n^{-1}\left(H_n(\theta) - H_n(\theta_0)\right)\mu_n^{-1\prime}\right\| \to_p 0,$$

where $N_n = \{\theta \in \Theta : \|\mu_n'(\theta - \theta_0)\| \leq 1\}$ is a sequence of shrinking neighborhoods of θ_0.

As shown in Wooldridge (1994), we may indeed deduce from ML1-ML3 that

$$\nu_n'(\hat{\theta}_n - \theta_0) = -\left[\nu_n^{-1} H_n(\theta_0) \nu_n^{-1\prime}\right]^{-1} \left[\nu_n^{-1} s_n(\theta_0)\right] + o_p(1) \to_d Q^{-1} P.$$

This is as expected from (7). In particular, ML3 ensures that $s_n(\hat{\theta}_n) = 0$ with probability approaching to one and

$$\nu_n^{-1} \left(H_n(\theta_n) - H_n(\theta_0)\right) \nu_n^{-1\prime} \to_p 0.$$

Now we define $u_t = (u_{1t}, u_{2t}, u_{3t})'$ to be given by

$$u_{1t} = \frac{z_{t-1}\varepsilon_{t-1}^2}{z_t}(\varepsilon_t^2 - 1), \quad u_{2t} = \frac{z_{t-1}}{z_t}(\varepsilon_t^2 - 1) \quad \text{and} \quad u_{3t} = \frac{1}{z_t}(\varepsilon_t^2 - 1),$$

and let $w_t = (u_t', v_t)'$. We let $[c]$ denote the largest integer which does not exceed c. Then we have

Lemma 1 *Under* Assumption 2, *we have*

$$W_n(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t \to_d W(r)$$

for $r \in [0,1]$, *where* W *is a vector Brownian motion with covariance matrix given by the long-run variance of* (w_t).

For the subsequent development of our asymptotics, we decompose $W = (U', V)'$, $U = (U_1, U_2, U_3)'$, conformably with (w_t), and note that the variance of vector Brownian motion U is given by

$$(\kappa^4 - 1)\Omega,$$

where $\kappa^4 = \mathbb{E}\varepsilon_t^4$ and

$$\Omega = \mathbb{E}\begin{pmatrix} (z_{t-1}^2/z_t^2)\varepsilon_{t-1}^4 & (z_{t-1}^2/z_t^2)\varepsilon_{t-1}^2 & (z_{t-1}/z_t^2)\varepsilon_{t-1}^2 \\ (z_{t-1}^2/z_t^2)\varepsilon_{t-1}^2 & z_{t-1}^2/z_t^2 & z_{t-1}/z_t^2 \\ (z_{t-1}/z_t^2)\varepsilon_{t-1}^2 & z_{t-1}/z_t^2 & 1/z_t^2 \end{pmatrix}.$$

Note that (u_t) is a martingale difference sequence. As we show in the proof of Lemma 1, Ω is well defined.

We introduce an Ornstein-Uhlenbeck process V_c, which is defined as

$$V_c(r) = \int_0^r \exp\left(-c(r-s)\right) dV(s)$$

from the limit Brownian motion V appeared in Lemma 1. As is well known, V_c is the solution of the stochastic differential equation

$$dV_c(r) = -cV_c(r)\, dr + dV(r)$$

with the initial condition $V_c(0) = 0$.

The asymptotic distribution theory of the QMLE $\hat{\theta}_n$ in our model is presented in the following theorem.

Theorem 2 *Let* Assumptions 1-3 *and* Regularity Conditions B.1 *in Appendix B hold. Then, all the conditions in* ML1-ML3 *are satisfied and*

$$\nu_n'(\hat{\theta}_n - \theta_0) \to_d \left(\int_0^1 M(r)\Omega M(r)'dr \right)^{-1} \int_0^1 M(r)\, dU(r),$$

where

$$\nu_n = \sqrt{n}\,\mathrm{diag}\left(1, 1, \frac{\dot{\kappa}_0(\sqrt{n})}{\kappa_0(\sqrt{n})}\right)$$

and

$$M(r) = \mathrm{diag}\left(1, 1, \frac{\bar{\dot{f}}_0(V_c(r))}{\bar{f}_0(V_c(r))}\right).$$

Theorem 2 shows that, in general, the QMLE $\hat{\theta}_n$ is consistent and its limit distribution is non-Gaussian. The limit distribution is represented as a functional of Brownian motions as in other models involving unit root or near unit root processes. Therefore, the standard inference is invalid.

However, the distribution reduces to mixed normal or even normal in some cases. The normality and mixed normality in these cases ensure the standard inference to be valid for our model. First, consider the special case, where the limit Brownian motion V is independent of U. This case arises when the innovation (v_t) of the covariate (x_t) is independent of the innovation (ε_t) of the process (y_t), i.e., when (x_t) is strictly exogenous. In this case, it follows immediately from Theorem 2 that

$$\nu_n'(\hat{\theta}_n - \theta_0) \to_d \mathbb{MN}\left(0, (\kappa^4 - 1)\left(\int_0^1 M(r)\Omega M(r)'dr \right)^{-1}\right),$$

where $\mathbb{MN}$ stands for mixed normal distribution.

Second, if the volatility function is *linear in parameter* and given by $f(x, \pi) = \pi g(x)$ for some asymptotically homogeneous function g, then we have $\nu_n = \sqrt{n}\,\mathrm{diag}(1, 1, (1/\pi_0))$ and $M(r)$ becomes an identity matrix for all $r \in [0, 1]$. Consequently, we have

$$\nu_n'(\hat{\theta}_n - \theta_0) \to_d \mathbb{N}\left(0, (\kappa^4 - 1)\Omega^{-1}\right),$$

where $\mathbb{N}$ signifies normal distribution. Since $f(x_t, \pi) = \pi\,|x_t|$ is the most common way to add a covariate in the GARCH(1,1) model in the related literature, Theorem 2 justifies the use of the quasi-maximum likelihood estimation method in the literature.

Note that the convergence rates for $\hat{\alpha}_n$, $\hat{\beta}_n$ and $\hat{\pi}_n$ are generally different. $\hat{\alpha}_n$ and $\hat{\beta}_n$ have the standard rate $\sqrt{n}$, while $\hat{\pi}_n$ has the rate $\sqrt{n}(\dot{\kappa}_0/\kappa_0)(\sqrt{n})$. $\hat{\pi}_n$ has the convergence rate that is dependent upon the asymptotic orders κ and $\dot{\kappa}$ of the volatility function f and its derivative $\dot{f}$ with respect to π.[4] However, $\hat{\pi}_n$ will also have the standard convergence rate if the volatility function is *linear in parameter* as it is shown above.

[4] For the CEV volatility function in (3), the convergence rate for $\hat{\pi}_n$ is

$$\sqrt{n}(\dot{\kappa}_0/\kappa_0)(\sqrt{n}) = \begin{pmatrix} (1/\pi_{10})\sqrt{n} & 0 \\ \sqrt{n}\ln(\sqrt{n}) & \sqrt{n} \end{pmatrix}$$

4 Explanation of Stylized Facts

In this section, we investigate the statistical properties of our model, and show how it explains stylized facts of financial return series, such as long memory property in volatility, IGARCH and leptokurtosis.

4.1 Long Memory Property in Volatility

The sample autocorrelation function of squared financial return series (in particular high frequency data) is known to have a typical trend that decreases fast at first and remains significantly positive for larger lags. Ding *et al.* (1993) found that it is possible to characterize the power transformation of stock returns to be long memory. However, it is known that the GARCH(1,1) model cannot generate this long memory property in volatility. The asymptotic limit of the sample autocorrelation of the squared process generated by the stationary GARCH(1,1) model is

$$R_k^2 = (\alpha + \beta)^{k-1} \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} \tag{8}$$

if $\mathbb{E}\left(\beta + \alpha\varepsilon_t^2\right)^2 < 1$. This shows that the autocorrelation function of the squared GARCH(1,1) process exponentially decreases and quickly converges to zero.

There has been active research to provide an explanation of the long memory property in volatility. See Baillie *et al.* (1996) and Ding and Granger (1996) (fractionality of the order of integration), Engle and Lee (1999) (two components), Diebold and Inoue (2001) (switching regime), Mikosch and Starica (2004) (structural change), Granger and Hyung (2004) (occasional break) and Han and Park (2006) (persistent covariate) for the related literature.

In order to see how our model explains the long memory property in volatility, we investigate the asymptotic behavior of the sample autocorrelation of the squared process generated by the GARCH(1,1) model with a persistent covariate. Define the sample autocorrelation of (y_t^2) by

$$R_{nk}^2 = \frac{\sum_{t=k+1}^{n} \left(y_t^2 - \bar{y}_n^2\right)\left(y_{t-k}^2 - \bar{y}_n^2\right)}{\sum_{t=1}^{n} \left(y_t^2 - \bar{y}_n^2\right)^2},$$

where $\bar{y}_n^2$ denotes the sample mean of (y_t^2). To precisely characterize the asymptotic behavior of R_{nk}^2, we make the following additional assumption.

Assumption 2′ Assume Assumption 2 with $q > 8$.

because we have

$$\dot{\kappa}(\sqrt{n}, \pi) = \begin{pmatrix} \left(\sqrt{n}\right)^{\pi_2} & 0 \\ \pi_1 \left(\sqrt{n}\right)^{\pi_2} \ln\left(\sqrt{n}\right) & \pi_1 \left(\sqrt{n}\right)^{\pi_2} \end{pmatrix} \quad \text{and} \quad \bar{f}^{\cdot}(x, \pi) = \begin{pmatrix} x^{\pi_2} \\ x^{\pi_2} \ln(x) \end{pmatrix}.$$

Proposition 3 *Let Assumptions 1, 2′ and 3 hold and $k \geq 1$. Then,*

$$R_{nk}^2 \to_d R_k^2 = \left[1 - (\alpha+\beta)^{k-1}\frac{\alpha(1-\alpha\beta-\beta^2)}{1-2\alpha\beta-\beta^2}\right] A + (\alpha+\beta)^{k-1}\frac{\alpha(1-\alpha\beta-\beta^2)}{1-2\alpha\beta-\beta^2}$$

where

$$A = \frac{\int_0^1 \bar{f}_0^2(V_c(r))d - \left(\int_0^1 \bar{f}_0(V_c(r))dr\right)^2}{\frac{\left(1-(\alpha+\beta)^2\right)K^4}{1-\left(\alpha^2K^4+2\alpha\beta+\beta^2\right)}\int_0^1 \overline{f}_0^2(V_c(r))dr - \left(\int_0^1 \bar{f}_0(V_c(r))dr\right)^2}.$$

Proposition 3 provides the asymptotic limit of the sample autocorrelation of the squared process generated by our model, which is denoted by R_k^2. Proposition 3 shows that one part of R_k^2 comes from the GARCH(1,1) component, considering (8), and the other part is generated from a persistent covariate in our model.

In order to see how our model explains the commonly observed long memory property in volatility, we consider the behavior of R_k^2 as $k \to \infty$. As $k \to \infty$, R_k^2 will decrease exponentially at first and converge to A. Here, A is random because it includes an Ornstein-Uhlenbeck process V_c, and it is positive and smaller than unity due to the Cauchy-Schwarz inequality[5] and $\frac{\left(1-(\alpha+\beta)^2\right)}{1-\left(\alpha^2K^4+2\alpha\beta+\beta^2\right)} > 1$. This means that the long memory property is generated by a persistent covariate. Moreover, the trend of R_k^2 in our model is quite similar to the typically observed trend of the sample autocorrelation function of squared financial return series, considering that it decreases fast (exponentially) at first and converges to a positive random value smaller than unity. Hence, Proposition 3 shows that our model successfully explains the long memory property in volatility.

4.2 IGARCH

In most empirical applications of the GARCH(1,1) model on financial return series with large sample, the ARCH effects ($\alpha+\beta$ in (5)) are found to be close to unity. Table 1 reports the result of fitting the GARCH(1,1) model to various stock return series, and the ARCH effects are very close to unity in most cases. This is why Engle and Bollerslev (1986) introduced the IGARCH (integrated GARCH) model with $\alpha+\beta=1$ in (5).

<< Insert Table 1 here >>

However, there have been claims that IGARCH could be spurious and could be due to the behavior of estimators under misspecification. One main reason why they think IGARCH could be spurious is that the IGARCH process cannot properly explain the long memory property in volatility. According to the IGARCH(1,1) model, the autocorrelation function of squared return series is unity for all lags, which is not realistic at all. This discordance motivated econometricians to find an explanation of IGARCH.

[5]Since f is not constant according to Assumption 3, we have $\int_0^1 \bar{f}_0^2(V_c(r))d > \left(\int_0^1 \bar{f}_0(V_c(r))dr\right)^2$ by the Cauchy-Schwarz inequality.

Econometricians have proposed several misspecified cases that would generate IGARCH. One example is a case with neglected structural changes. See Diebold (1986) and Lamoureux and Lastrapes (1990). Another example is a case with neglected fractionality of the order of integration. See Baillie *et al.* (1996). These works made use of either simulations or indirect approaches, but Mikosch and Starica (2004) and Hillebrand (2005) showed theoretically that IGARCH could be generated, due to the behavior of estimators, by neglecting structural change.[6] Recently, Han an Park (2007) showed theoretically that IGARCH could be the result of missing a persistent covariate in the ARCH(1) model. Note that the models considered in this context that provide explanations of IGARCH generate the long memory property in volatility.

Following the procedure done by Han and Park (2007), we investigate a misspecified case, which considers the effect of missing a persistent covariate in the GARCH(1,1) model. Define

$$m_t = z_t \varepsilon_t^2,$$

where (z_t) is given in (4). Moreover, we introduce some additional assumption, which are given by

Assumption 4 We let

$$e_{1t}(\beta) = \frac{m_t}{\sum_{k=1}^{\infty} \beta^{k-1} m_{t-k}}, \; e_{2t}(\beta) = \frac{m_{t-1}}{\sum_{k=1}^{\infty} \beta^{k-1} m_{t-k}}, \; e_{3t}(\beta) = \frac{m_t m_{t-1}}{\left(\sum_{k=1}^{\infty} \beta^{k-1} m_{t-k}\right)^2},$$

and assume that $(e_{it}(\beta))$ are strictly stationary and ergodic with $\mathbb{E} e_{it}(\beta)$ finite and continuous for all $\beta \in \mathcal{B} \subset (0,1)$, $i = 1, 2, 3$.

The following proposition establishes the probability limit of the QMLE's $\tilde{\alpha}_n$ and $\tilde{\beta}_n$ respectively for the parameters α and β in the GARCH model (5).[7]

Proposition 4 *Let* Assumptions 1-4 *hold. Then we have*

$$\tilde{\alpha}_n \to_p \alpha_* \quad \textit{and} \quad \tilde{\beta}_n \to_p \beta_*$$

[6] Mikosch and Starica (2004) showed it in the framework of the Whittle estimation, but the theory of Hillebrand (2005) does not depend on estimation method and covers the common framework of the maximum likelihood estimation.

[7] The constant term parameter ω is not identified in our model, precisely for the same reason as in the nonstationary ARCH models studied by Jensen and Rahbek (2004). Therefore, we will not consider its QMLE.

with α_* and β_* defined as the solution of the simultaneous equations

$$\mathbb{E}\left(\frac{m_t}{\alpha_* \sum_{k=1}^{\infty} \beta_*^{k-1} m_{t-k}} - 1\right) = 0$$

$$\mathbb{E}\left[\left(\frac{m_t}{\alpha_* \sum_{k=1}^{\infty} \beta_*^{k-1} m_{t-k}} - 1\right) \frac{m_{t-1}}{\sum_{k=1}^{\infty} \beta_*^{k-1} m_{t-k}}\right] = 0,$$

which we assume to exist.

Proposition 4 shows that the pseudo-true values α_* and β_* of the parameters for the misspecified GARCH(1,1) model in (5) are determined solely by the distribution of (m_t), which is completely specified by the true value α_0 and β_0 of the GARCH parameters and the distribution of (ε_t) in our model. Note that their values depend upon neither the volatility function f nor any characteristics of the covariate (x_t).

It is quite simple to calculate α_* and β_*, at least approximately, once the true value α_0 and β_0 of the GARCH parameters and the distribution of (ε_t) are given. We just consider the data generated from the GARCH(1,1) model with α_0 and β_0 (with an arbitrary constant term, which is unimportant), and fit the data by the GARCH(1,1) model without the constant term, i.e.,

$$\sigma_t^2 = \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{9}$$

using the maximum likelihood estimation method. Further, we denote the QMLE's of the parameters α and β in (9) by $\bar{\alpha}_n$ and $\bar{\beta}_n$, say. Then, we can obtain $\bar{\alpha}_n$ and $\bar{\beta}_n$ that are given by

$$\sum_{t=1}^{n}\left(\frac{m_t}{\bar{\alpha} \sum_{k=1}^{\infty} \bar{\beta}^{k-1} m_{t-k}} - 1\right) = 0$$

$$\sum_{t=1}^{n}\left(\frac{m_t}{\bar{\alpha} \sum_{k=1}^{\infty} \bar{\beta}^{k-1} m_{t-k}} - 1\right) \frac{m_{t-1}}{\sum_{k=1}^{\infty} \bar{\beta}^{k-1} m_{t-k}} = 0.$$

Hence, we may easily see that

$$\bar{\alpha}_n \to_p \alpha_* \quad \text{and} \quad \bar{\beta}_n \to_p \beta_*,$$

where α_* and β_* are defined as in Proposition 4. See the proof of Proposition 4 for details. Since we may assume without loss of generality that $\bar{\alpha}_n$ and $\bar{\beta}_n$ are a.s. bounded, we have

$$\mathbb{E}\bar{\alpha}_n \approx \alpha_* \quad \text{and} \quad \mathbb{E}\bar{\beta}_n \approx \beta_*,$$

for large n. See Han and Park (2007) for more details.

<< Insert Table 2 here >>

Table 2 reports the simulated values of $\mathbb{E}\bar{\alpha}_n$ and $\mathbb{E}\bar{\beta}_n$. We consider both normal and uniform distributions for ε_t when α_0 and β_0 take different values; $0.1 \leq \alpha_0 \leq 0.3$, $0.1 \leq \beta_0$ and $\alpha_0 + \beta_0 \leq 0.7$. The computed values of α_* and β_* are different depending upon both the distribution of (ε_t) and the values of α_0 and β_0. However, we see that the values of $\alpha_* + \beta_*$ are very close to unity in all cases. Hence, we can expect that we would observe the evidence of IGARCH if we omit a relevant covariate in the GARCH(1,1) and fit the GARCH(1,1) model. This is so, regardless of the values of α_0 and β_0 and the distribution of (ε_t). This implies that the empirical findings suggesting the IGARCH(1,1) process could be the result of missing a persistent covariate in the GARCH(1,1) model.

4.3 Leptokurtosis

Many financial time series are known to be leptokurtic. Table 1 shows that the sample kurtosis of stock return series ranges from 5.99 to 27.96. The asymptotic limit of the sample kurtosis of the GARCH(1,1) process is

$$\frac{\left(1-(\alpha+\beta)^2\right)\kappa^4}{1-\left(\alpha^2\kappa^4+2\alpha\beta+\beta^2\right)} \tag{10}$$

if $\mathbb{E}\left(\beta+\alpha\varepsilon_t^2\right)^2 < 1$, which shows that the GARCH(1,1) process is also leptokurtic. The implied kurtosis of (ε_t) in Table 1 is a calculated value of $\kappa^4 \left(= \mathbb{E}\left(\varepsilon_t^4\right)\right)$ from

$$\text{sample kurtosis} = \frac{\left(1-\left(\hat{\alpha}+\hat{\beta}\right)^2\right)\kappa^4}{1-\left(\hat{\alpha}^2\kappa^4+2\hat{\alpha}\hat{\beta}+\hat{\beta}^2\right)}. \tag{11}$$

In Table 1, we divide ten stock return series into two groups depending on the value of the implied kurtosis of (ε_t). For the group 1 in Table 1, the implied kurtosis of (ε_t) is close to three ranging from 2.20 to 3.38. In this case, the GARCH(1,1) model with normally distributed innovations is approximately explaining the observed sample kurtosis.

Meanwhile, the implied kurtosis of (ε_t) is much larger than three for the group 2 in Table 1, which ranges from 4.61 to 8.36. As this shows, it is well known that the kurtosis implied by the GARCH(1,1) model with normally distributed innovations tends to be far less than the sample kurtosis observed for many financial return series. As a typical way to overcome this problem, some econometricians proposed the use of innovation (ε_t) with a fat-tailed distribution while maintaining the GARCH(1,1) model. For example, Bollerslev (1987) advocates the use of innovations with the t-distribution, and Bai *et al.* (2003) considers innovations following a mixture of two normal distributions. We shall refer to these models as the GARCH(1,1) model with fat-tailed innovations.

However, the GARCH(1,1) model with fat-tailed innovations has its limitations. Even if it successfully explains the leptokurtosis of financial time series, it cannot provide explanations on other stylized facts. Most of all, the GARCH(1,1) model with fat-tailed innovations cannot explain the long memory property in volatility as it was shown in the previous section. Regardless of the distribution of (ε_t), the GARCH(1,1) model cannot explain the observed behavior of the autocorrelation function of squared return series.

Now we investigate how a persistent covariate in the GARCH (1,1) model affects the kurtosis of time series. Define the sample kurtosis of (y_t) by

$$K_n^4 = \frac{1}{n}\sum_{t=1}^{n} y_t^4 \bigg/ \left(\frac{1}{n}\sum_{t=1}^{n} y_t^2\right)^2 .$$

We derive the asymptotic limit of the sample kurtosis of the GARCH(1,1) process with a persistent covariate.

Proposition 5 *Let Assumptions 1, 2′* and 3 *hold and* $k \geq 1$. *Then,*

$$K_n^4 \to_d \frac{\left(1-(\alpha+\beta)^2\right)\kappa^4}{1-\left(\alpha^2\kappa^4+2\alpha\beta+\beta^2\right)} \frac{\int_0^1 \bar{f}_0^2(V_c(r))dr}{\left(\int_0^1 \bar{f}_0(V_c(r))dr\right)^2}.$$

Proposition 5 shows that the limit of the sample kurtosis in our model is a product of two parts. One part is the same as (10). The other part is random because it contains an Ornstein-Uhlenbeck process V_c, and its value is larger than unity due to the Cauchy-Schwarz inequality. This means that, asymptotically, our model has a bigger kurtosis than the GARCH(1,1) model. Note that there is no additional assumption on the distribution of (ε_t). Hence, Proposition 5 provides an alternative explanation of the leptokurtosis observed in financial return series without using innovations with fat-tailed distributions.

5 Conclusion

This paper establishes the asymptotic theory of the QMLE in the GARCH(1,1) model with a covariate that is integrated or nearly integrated. For a wide class of volatility functions, including the CEV function, the QMLE in the model is consistent with the convergence rate depending upon the asymptotic order of the volatility function. Their limit distributions are, however, generally non-Gaussian. Only in two special cases do they have limit Gaussian distributions, and the standard testing procedures are valid. If the covariate is strictly exogenous, then the limit distribution of the QMLE becomes mixed normal. Also, if the volatility function is linear in parameter, we have normal limit distribution for the QMLE with the standard convergence rate.

This paper also provides asymptotic theories, which show that the model successfully explains the stylized facts of financial time series, such as long memory property in volatility, IGARCH and leptokurtosis. The autocorrelation of the squared process of the model has a trend that is similar to the trend observed in real data; it decreases fast at first and stays positive for larger lags. Our theory of misspecification shows that IGARCH could be the result of missing a persistent covariate in the GARCH(1,1) process. Moreover, the sample kurtosis of the model is, in asymptotics, randomly bigger than the kurtosis of the GARCH(1,1) model, which provides an explanation of the sample kurtosis observed in real data without using innovations with fat-tailed distributions.

Our model is conventional in the sense that it is a conditional model; it is based on the GARCH(1,1) model. However, our model is unconventional in the sense that it is nonstationary. While most ARCH type models depend on the assumption of stationarity, our model is nonstationary due to the presence of a persistent covariate; the unconditional variance of the model is time-varying. Note that those models that can explain both the long memory property in volatility and IGARCH are nonstationary. For example, like our model, the GARCH(1,1) model with structural change is also nonstationary.

It is interesting to compare our model, a nonstationary conditional variance model, with the work by Starica and Granger (2005), who investigated a nonstationary unconditional variance model of stock return series. They discovered that most of the dynamics of stock return series are concentrated in shifts of the unconditional variance. They argued that their modeling reflects "the belief that fundamental features of the financial markets are continuously and significantly changing." We agree with their idea, and we can deduce that persistent economic variables included in the GARCH(1,1) model account for the change in fundamental features of the financial markets.

Appendix A. Useful Lemmas and Their Proofs

The proofs of the theorems in the paper rely on the results from the following lemmas.

Lemma A1 Under Assumptions 1-3, we have

$$\kappa_0(\sqrt{n})^{-1} \max_{1\leq t\leq n} \left|\sigma_t^2\left(\theta_0\right) - z_t f_0(x_t)\right| = o_p(1)$$

for all large n.

Proof of Lemma A1 Let (τ_n) be a sequence of numbers such that

$$\tau_n = n^r$$

with $0 < r < 1/4 - 1/2p - 1/q$. By the recursive substitution, it can be easily deduced that

$$\sigma_t^2\left(\theta_0\right) = \sum_{k=0}^{\infty} f_0(x_{t-k}) \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)$$

where $\prod_{i=1}^{0}\left(\beta + \alpha\varepsilon_{t-i}^2\right) = 1$. We may write

$$\sigma_t^2\left(\theta_0\right) = z_t f_0(x_t) + e_t$$

with

$$e_t = e_t(A) + e_t(B) + e_t(C),$$

where

$$e_t(A) = \sum_{k=1}^{\tau_n}[f_0(x_{t-k}) - f_0(x_t)] \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)$$
$$e_t(B) = \sum_{k=\tau_n+1}^{\infty} f_0(x_{t-k}) \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)$$
$$e_t(C) = -\sum_{k=\tau_n+1}^{\infty} f_0(x_t) \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right).$$

The stated result follows from the proof of Han and Park (2007, lemma A), where $\prod_{i=1}^{k}\left(\alpha\varepsilon_{t-i}^2\right)$ and $z_t = 1+\sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\alpha\varepsilon_{t-i}^2\right)$ are considered instead of $\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)$ and $z_t = 1+\sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)$. □

Lemma A2 Under Assumptions 1-3, we have

$$\kappa_0(\sqrt{n})^{-1} \max_{1\leq t\leq n} |z_t f_0(x_t) - z_t f_0(x_{t-1})| = o_p(1)$$

for all large n.

Proof of Lemma A2 The stated result follows from the proof of Han and Park (2007, lemma B), where $\prod_{i=1}^{k}\left(\alpha\varepsilon_{t-i}^2\right)$ and $z_t = 1+\sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\alpha\varepsilon_{t-i}^2\right)$ are considered instead of $\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)$ and $z_t = 1+\sum_{k=1}^{\infty}\prod_{i=1}^{k}\left(\beta + \alpha\varepsilon_{t-i}^2\right)$. □

Lemma A3 Under Assumption 1-3, we have

$$(n\kappa_0(\sqrt{n}))^{-1}\sum_{t=1}^{n} f_0(x_t) \quad \to_d \quad \int_0^1 \overline{f}_0(V_c(r))dr$$
$$(n\kappa_0^2(\sqrt{n}))^{-1}\sum_{t=1}^{n} f_0^2(x_t) \quad \to_d \quad \int_0^1 \overline{f}_0^2(V_c(r))dr.$$

Proof of Lemma A3 See Park (2002, lemma A.1) and Park (2003, Theorem 3.8). □

Lemma A4 Let $\widetilde{u}_t$ be a martingale difference sequence with respect to $\mathcal{F}_t$ such that $\mathbb{E}(\widetilde{u}_t^2|\mathcal{F}_{t-1}) = \sigma_{\widetilde{u}}^2$ a.s. for each t and $\sup_t \mathbb{E}(|\widetilde{u}_t|^{2+\epsilon}|\mathcal{F}_{t-1}) < \infty$ a.s. for some $\epsilon > 0$. Under Assumption 1-3, we have

$$(\sqrt{n}\kappa_0(\sqrt{n}))^{-1}\sum_{t=1}^{n} f_0(x_t)\widetilde{u}_t \to_d \int_0^1 \overline{f}_0(V_c(r))d\widetilde{U}(r)$$

where $\widetilde{U}_n(r) = \frac{1}{\sqrt{n}}\sum_{t}^{[nr]} u_t \to_d \widetilde{U}(r)$.

Proof of Lemma A4 See Park (2003, Theorem 3.8). □

Lemma A5 Let $\alpha + \beta < 1$. Under Assumption 1-3, we have

$$\begin{aligned}(n\kappa_0(\sqrt{n}))^{-1}\sum_{t=1}^{n} y_t^2 &\to_d \frac{1}{1-(\alpha+\beta)}\int_0^1 \overline{f}_0(V_c(r))dr \\ (n\kappa_0(\sqrt{n}))^{-1}\sum_{t=1}^{n}\sum_{k=1}^{\infty}\beta^{k-1}y_{t-k}^2 &\to_d \frac{1}{1-\beta}\frac{1}{1-(\alpha+\beta)}\int_0^1 \overline{f}_0(V_c(r))dr.\end{aligned}$$

Proof of Lemma A5 Note that

$$\begin{aligned}(n\kappa_0(\sqrt{n}))^{-1}\sum_{t=1}^{n} y_t^2 &= (n\kappa_0(\sqrt{n}))^{-1}\sum_{t=1}^{n} z_t f_0(x_t)\varepsilon_t^2 + o_p(1) \\ &= (n\kappa_0(\sqrt{n}))^{-1}\sum_{t=1}^{n}\mathbb{E}\left(z_t\varepsilon_t^2\right) f_0(x_t) + o_p(1).\end{aligned}$$

The first line comes from lemma A1, and the second line follows because lemma A4 implies that $\sum_{t=1}^{n}\left[z_t\varepsilon_t^2 - \mathbb{E}\left(z_t\varepsilon_t^2\right)\right] f(x_t) = O_p\left(\sqrt{n}\kappa_0(\sqrt{n})\right)$. Since

$$\mathbb{E}\left(z_t\varepsilon_t^2\right) = \mathbb{E}\left(z_t\right) = \frac{1}{1-(\alpha+\beta)}, \tag{12}$$

the first stated result follows immediately. For the proof of the second asymptotics, we use the result given by Han and Park (2006, lemma 4). □

Appendix B. Proofs of the Main Results

Proof of Lemma 1 The stated result follows from the proof of Han and Park (2007, lemma 1) if we have

$$\mathbb{E}\|w_t\|^2 < \infty.$$

We have

$$\left|\frac{1}{z_t}\right| \leq 1 \text{ a.s.} \tag{13}$$

Moreover, it follows from He and Teräsvirta (1999) that

$$\mathbb{E}z_t^2 < \infty \tag{14}$$

due to (2). Therefore, it follows that

$$\begin{aligned}
\mathbb{E}u_{1t}^2 &= (\kappa^4 - 1)\mathbb{E}\left(\frac{z_{t-1}^2 \varepsilon_{t-1}^4}{z_t^2}\right) \leq (\kappa^4 - 1)\mathbb{E}\left(z_{t-1}^2 \varepsilon_{t-1}^4\right) = \kappa^4(\kappa^4 - 1)\mathbb{E}z_{t-1}^2 < \infty, \\
\mathbb{E}u_{2t}^2 &= (\kappa^4 - 1)\mathbb{E}\left(\frac{z_{t-1}^2}{z_t^2}\right) \leq (\kappa^4 - 1)\mathbb{E}\left(z_{t-1}^2\right) = (\kappa^4 - 1)\mathbb{E}z_{t-1}^2 < \infty
\end{aligned}$$

and

$$\mathbb{E}u_{3t}^2 \leq \kappa^4 - 1 < \infty,$$

due to (13) and (14), and the proof is complete. □

Before we prove Theorem 1, we need additional technical conditions to fully develop the asymptotic theory of the QMLE. Define

$$\dot{f} = \left(\frac{\partial f}{\partial \pi_i}\right), \quad \ddot{f} = \left(\frac{\partial^2 f}{\partial \pi_i \partial \pi_j}\right), \quad \dddot{f} = \left(\frac{\partial^3 f}{\partial \pi_i \partial \pi_j \partial \pi_k}\right)$$

to be all vectors, arranged by the lexicographic ordering of their indices, if they exist. We also define a neighborhood of π_0 by

$$N = \left\{\pi \in \Pi : \left\|(\dot{\kappa}_0/\kappa_0)(\lambda)'(\pi - \pi_0)\right\| \leq \lambda^{-1+\varepsilon}\right\}$$

for $\varepsilon > 0$ given.

Conditions B.1 Assume
(a) $\dot{f}, \ddot{f}$ and $\dddot{f}$ exist, and $\dot{f}$ and $\ddot{f}$ are asymptotically homogeneous,
(b) $\dot{f}/f$ is asymptotically homogeneous with asymptotic order $\dot{\kappa}/\kappa$ and limit homogeneous function $\bar{f}^{\cdot}/\bar{f}$, and
(c) for any $\bar{s} > 0$ given, there exists $\varepsilon > 0$ such that

$$\lambda^{-1}\left\|\kappa_0(\lambda)\left(\dot{\kappa}_0 \otimes \dot{\kappa}_0\right)(\lambda)^{-1}\ddot{\kappa}_0(\lambda)\right\| \to 0 \tag{15}$$

$$\lambda^{-1+\varepsilon}\left\|\kappa_0^2(\lambda)\dot{\kappa}_0(\lambda)^{-1}\sup_{|s|\leq\bar{s}}\left(\sup_{\pi\in N}\left|f(\lambda s,\pi)^{-2}\right|\sup_{\pi\in N}\left|\dot{f}(\lambda s,\pi)\right|\right)\right\| \to 0 \tag{16}$$

$$\begin{aligned}
&\lambda^{-1+\varepsilon}\left\|\kappa_0(\lambda)\dot{\kappa}_0(\lambda)^{-1}\sup_{|s|\leq\bar{s}}\sup_{\pi\in N}\left|\dot{f}(\lambda s,\pi)\right|\right\| \\
&\times\left\|\kappa_0^2(\lambda)\left(\dot{\kappa}_0 \otimes \dot{\kappa}_0\right)(\lambda)^{-1}\sup_{|s|\leq\bar{s}}\left(\sup_{\pi\in N}\left|f(\lambda s,\pi)^{-2}\right|\sup_{\pi\in N}\left|\ddot{f}(\lambda s,\pi)\right|\right)\right\| \to 0
\end{aligned} \tag{17}$$

$$\lambda^{-1+\varepsilon}\left\|\kappa_0^3(\lambda)\left(\dot{\kappa}_0 \otimes \dot{\kappa}_0 \otimes \dot{\kappa}_0\right)(\lambda)^{-1}\sup_{|s|\leq\bar{s}}\left(\sup_{\pi\in N}\left|f(\lambda s,\pi)^{-1}\right|\sup_{\pi\in N}\left|\dddot{f}(\lambda s,\pi)\right|\right)\right\| \to 0 \tag{18}$$

as $\lambda \to \infty$.

It is tedious, but straightforward to show that the CEV volatility function satisfies all the conditions in Conditions B.1.[8]

Proof of Theorem 1 The proof will be done in three steps, for each of ML1-ML3. For notational simplicity, we write $f(x_t) = f(x_t, \pi)$ and $f_0(x_t) = f(x_t, \pi_0)$, and $\sigma_t^2 = \sigma_t^2(\theta)$ and $\sigma_{0t}^2 = \sigma_t^2(\theta_0)$. Also, we let $\kappa_n = \kappa_0(\sqrt{n}, \pi_0), \dot{\kappa}_n = \dot{\kappa}_0(\sqrt{n}, \pi_0)$ and $\ddot{\kappa}_n = \ddot{\kappa}_0(\sqrt{n}, \pi_0)$. We denote by $\ddot{F} = \partial^2 f / \partial \pi \partial \pi'$ the second derivative of f in matrix form..

First Step Since

$$\frac{\partial \sigma_t^2}{\partial \theta} = \left(\frac{\partial \sigma_t^2}{\partial \alpha}, \frac{\partial \sigma_t^2}{\partial \beta}, \frac{\partial \sigma_t^2}{\partial \pi'} \right)' = \left(y_{t-1}^2, \sigma_{t-1}^2, \dot{f}(x_t)' \right)',$$

the score function is given by

$$s_n(\theta) = \frac{1}{2} \sum_{t=1}^n \left(\frac{y_t^2}{\sigma_t^2} - 1 \right) \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} = \frac{1}{2} \sum_{t=1}^n \left(\varepsilon_t^2 - 1 \right) \left(\frac{y_{t-1}^2}{\sigma_t^2}, \frac{\sigma_{t-1}^2}{\sigma_t^2}, \frac{\dot{f}(x_t)'}{\sigma_t^2} \right)'. \tag{19}$$

We have

$$\frac{y_t^2}{\sigma_{0t}^2} = \varepsilon_t^2. \tag{20}$$

Moreover, it follows from Lemmas A1 and A2 that

$$\kappa_n^{-1} \sigma_{0t}^2 = z_t \left(\kappa_n^{-1} f_0(x_t) \right) + o_p(1) = z_t \left(\kappa_n^{-1} f_0(x_{t-1}) \right) + o_p(1)$$

uniformly in $t = 1, \ldots, n$, which in turn yields

$$\frac{y_{t-1}^2}{\sigma_{0t}^2} = \frac{\kappa_n^{-1} \sigma_{0,t-1}^2}{\kappa_n^{-1} \sigma_{0t}^2} \varepsilon_{t-1}^2 = \frac{z_{t-1} \varepsilon_{t-1}^2}{z_t} + o_p(1) \tag{21}$$

$$\left(\frac{\dot{\kappa}_n}{\kappa_n} \right)^{-1} \frac{\dot{f}_0(x_t)}{\sigma_{0t}^2} = \frac{\dot{\kappa}_n^{-1} \dot{f}_0(x_t)}{\kappa_n^{-1} f_0(x_t)} \frac{1}{z_t} + o_p(1) \tag{22}$$

uniformly in $t = 1, \ldots, n$.

Now we let

$$s_n(\theta_0) = (s_{1n}(\theta_0), s_{2n}(\theta_0), s_{3n}'(\theta_0))'.$$

It follows from Lemma 1 and (19) - (22) that

$$\begin{aligned} \frac{1}{\sqrt{n}} s_{1n}(\theta_0) &= \frac{1}{2\sqrt{n}} \sum_{t=1}^n \frac{z_{t-1} \varepsilon_{t-1}^2}{z_t} (\varepsilon_t^2 - 1) + o_p(1) \\ &= \frac{1}{2\sqrt{n}} \sum_{t=1}^n u_{1t} + o_p(1) \to_d \frac{1}{2} U_1(1) \end{aligned} \tag{23}$$

[8] The detailed proof is not given to save the space. The proof is available from the author upon request.

and

$$\frac{1}{\sqrt{n}}\, s_{2n}(\theta_0) = \frac{1}{2\sqrt{n}} \sum_{t=1}^{n} \frac{z_{t-1}}{z_t}(\varepsilon_t^2 - 1) + o_p(1)$$

$$= \frac{1}{2\sqrt{n}} \sum_{t=1}^{n} u_{2t} + o_p(1) \to_d \frac{1}{2} U_2(1). \tag{24}$$

Furthermore, we may also deduce from Lemma 1 and (19) - (22) that

$$\frac{1}{\sqrt{n}} \left(\frac{\dot{\kappa}_n}{\kappa_n}\right)^{-1} s_{3n}(\theta_0) = \frac{1}{2\sqrt{n}} \sum_{t=1}^{n} \frac{\dot{\kappa}_n^{-1} \dot{f}_0(x_t)}{\kappa_n^{-1} f_0(x_t)} \frac{1}{z_t}(\varepsilon_t^2 - 1) + o_p(1)$$

$$= \frac{1}{2\sqrt{n}} \sum_{t=1}^{n} \frac{\dot{\kappa}_n^{-1} \dot{f}_0(x_t)}{\kappa_n^{-1} f_0(x_t)} u_{3t} + o_p(1) \to_d \frac{1}{2} \int_0^1 \frac{\bar{\dot{f}}_0(V_c(r))}{\bar{f}_0(V_c(r))} dU_3(r). \tag{25}$$

Consequently, it follows from (23) and (25) that

$$\nu_n^{-1} s_n(\theta_0) \to_d \frac{1}{2} \int_0^1 M(r)\, dU(r),$$

which establishes ML1.

Second Step Since

$$\frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'} = \begin{pmatrix} y_{t-1}^4 & y_{t-1}^2 \sigma_{t-1}^2 & y_{t-1}^2 \dot{f}' \\ y_{t-1}^2 \sigma_{t-1}^2 & \sigma_{t-1}^4 & \sigma_{t-1}^2 \dot{f}' \\ \dot{f} y_{t-1}^2 & \dot{f} \sigma_{t-1}^2 & \dot{f} \dot{f}' \end{pmatrix} \quad \text{and} \quad \frac{\partial^2 \sigma_t^2}{\partial \theta \partial \theta'} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \ddot{F}(x_t) \end{pmatrix},$$

the Hessian function is given by

$$H_n(\theta) = \frac{1}{2} \sum_{t=1}^{n} \left[\left(1 - \frac{2y_t^2}{\sigma_t^2}\right) \frac{1}{\sigma_t^4} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'} + \left(\frac{y_t^2}{\sigma_t^2} - 1\right) \frac{1}{\sigma_t^2} \frac{\partial^2 \sigma_t^2}{\partial \theta \partial \theta'} \right]$$

$$= \frac{1}{2} \sum_{t=1}^{n} \left[\left(1 - \frac{2y_t^2}{\sigma_t^2}\right) \frac{1}{\sigma_t^4} \begin{pmatrix} y_{t-1}^4 & y_{t-1}^2 \sigma_{t-1}^2 & y_{t-1}^2 \dot{f}' \\ y_{t-1}^2 \sigma_{t-1}^2 & \sigma_{t-1}^4 & \sigma_{t-1}^2 \dot{f}' \\ \dot{f} y_{t-1}^2 & \dot{f} \sigma_{t-1}^2 & \dot{f} \dot{f}' \end{pmatrix} \right.$$

$$\left. + \left(\frac{y_t^2}{\sigma_t^2} - 1\right) \frac{1}{\sigma_t^2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \ddot{F}(x_t) \end{pmatrix} \right]. \tag{26}$$

We let

$$H_n(\theta_0) = \begin{pmatrix} H_{11}^n(\theta_0) & H_{12}^n(\theta_0) & H_{13}^n(\theta_0) \\ H_{21}^n(\theta_0) & H_{22}^n(\theta_0) & H_{23}^n(\theta_0) \\ H_{31}^n(\theta_0) & H_{32}^n(\theta_0) & H_{33}^n(\theta_0) \end{pmatrix}$$

in what follows.

It follows from (20) - (22) and (26) that

$$-\frac{H_{11}^n(\theta_0)}{n} = \frac{1}{2n}\sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right)\frac{\left(\kappa_n^{-1}\sigma_{0,t-1}^2\varepsilon_{t-1}^2\right)^2}{\left(\kappa_n^{-1}\sigma_{0t}^2\right)^2}$$
$$= \frac{1}{2n}\sum_{t=1}^n \left(\frac{z_{t-1}}{z_t}\right)^2 \varepsilon_{t-1}^4 + o_p(1) \to_p \frac{1}{2}\mathbb{E}\left[\left(\frac{z_{t-1}}{z_t}\right)^2 \varepsilon_{t-1}^4\right], \qquad (27)$$

$$-\frac{H_{21}^n(\theta_0)}{n} = \frac{1}{2n}\sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right)\frac{\left(\kappa_n^{-1}\sigma_{0,t-1}^2\varepsilon_{t-1}^2\right)\kappa_n^{-1}\sigma_{0,t-1}^2}{\left(\kappa_n^{-1}\sigma_{0t}^2\right)^2}$$
$$= \frac{1}{2n}\sum_{t=1}^n \left(\frac{z_{t-1}}{z_t}\right)^2 \varepsilon_{t-1}^2 + o_p(1) \to_p \frac{1}{2}\mathbb{E}\left[\left(\frac{z_{t-1}}{z_t}\right)^2 \varepsilon_{t-1}^2\right]$$

and

$$-\frac{H_{22}^n(\theta_0)}{n} = \frac{1}{2n}\sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right)\frac{\left(\kappa_n^{-1}\sigma_{0,t-1}^2\right)^2}{\left(\kappa_n^{-1}\sigma_{0t}^2\right)^2}$$
$$= \frac{1}{2n}\sum_{t=1}^n \left(\frac{z_{t-1}}{z_t}\right)^2 + o_p(1) \to_p \frac{1}{2}\mathbb{E}\left[\left(\frac{z_{t-1}}{z_t}\right)^2\right].$$

Moreover, we have from Lemma 1, (20) - (22) and (26)

$$-\frac{\dot{\kappa}_n^{-1}H_{31}^n(\theta_0)}{n\kappa_n^{-1}} = \frac{1}{2n}\sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right)\frac{\kappa_n^{-1}\sigma_{0,t-1}^2\varepsilon_{t-1}^2}{\kappa_n^{-1}\sigma_{0,t}^2}\frac{\dot{\kappa}_n^{-1}\dot{f}_0(x_t)}{\kappa_n^{-1}\left(z_t f_0(x_t)\right)} + o_p(1)$$
$$= \frac{1}{2n}\sum_{t=1}^n \frac{z_{t-1}}{z_t^2}\varepsilon_{t-1}^2 \frac{\dot{\kappa}_n^{-1}\dot{f}_0(x_t)}{\kappa_n^{-1}f_0(x_t)} + o_p(1) \to_d \frac{1}{2}\mathbb{E}\left[\frac{z_{t-1}}{z_t^2}\varepsilon_{t-1}^2\right]\int_0^1 \frac{\bar{\dot{f}}_0(V_c(r))}{\bar{f}_0(V_c(r))}dr$$

and

$$-\frac{\dot{\kappa}_n^{-1}H_{32}^n(\theta_0)}{n\kappa_n^{-1}} = \frac{1}{2n}\sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right)\frac{\kappa_n^{-1}\sigma_{0,t-1}^2}{\kappa_n^{-1}\sigma_{0,t}^2}\frac{\dot{\kappa}_n^{-1}\dot{f}_0(x_t)}{\kappa_n^{-1}\left(z_t f_0(x_t)\right)} + o_p(1)$$
$$= \frac{1}{2n}\sum_{t=1}^n \frac{z_{t-1}}{z_t^2}\frac{\dot{\kappa}_n^{-1}\dot{f}_0(x_t)}{\kappa_n^{-1}f_0(x_t)} + o_p(1) \to_d \frac{1}{2}\mathbb{E}\left[\frac{z_{t-1}}{z_t^2}\right]\int_0^1 \frac{\bar{\dot{f}}_0(V_c(r))}{\bar{f}_0(V_c(r))}dr.$$

Finally, we consider $H_{33}^n(\theta_0)$, which involves two terms. We write

$$H_{33}^n(\theta_0) = H_{33a}^n(\theta_0) + H_{33b}^n(\theta_0).$$

For the first term, we have from Lemma 1, (20) - (22) and (26)

$$
\begin{aligned}
-\frac{\dot\kappa_n^{-1} H_{33a}^n(\theta_0)\dot\kappa_n^{-1\prime}}{n\kappa_n^{-2}} &= \frac{\kappa_n^2}{2n}\dot\kappa_n^{-1}\left[\sum_{t=1}^n \left(2\varepsilon_t^2-1\right)\frac{\dot f_0(x_t)}{\sigma_{0t}^2}\frac{\dot f_0(x_t)'}{\sigma_{0t}^2}\right]\dot\kappa_n^{-1\prime} \\
&= \frac{1}{2n}\sum_{t=1}^n \frac{1}{z_t^2}\frac{\dot\kappa_n^{-1}\dot f_0(x_t)}{\kappa_n^{-1} f_0(x_t)}\frac{\left(\dot\kappa_n^{-1}\dot f_0(x_t)\right)'}{\kappa_n^{-1} f_0(x_t)} + o_p(1) \\
&\to_d \frac{1}{2}\mathbb{E}\left[\frac{1}{z_t^2}\right]\int_0^1 \frac{\dot{\bar f}_0(V_c(r))\dot{\bar f}_0(V_c(r))'}{\bar f_0^2(V_c(r))}dr. \qquad (28)
\end{aligned}
$$

The second term becomes negligible. In fact, we can deduce that

$$
-\frac{\dot\kappa_n^{-1} H_{33b}^n(\theta_0)\dot\kappa_n^{-1\prime}}{n\kappa_n^{-2}} = \frac{\kappa_n^2}{2n}\dot\kappa_n^{-1}\left[\sum_{t=1}^n \left(1-\varepsilon_t^2\right)\frac{\ddot F_0(x_t)}{\sigma_{0t}^2}\right]\dot\kappa_n^{-1\prime} = o_p(1) \qquad (29)
$$

because

$$
\begin{aligned}
&\operatorname{vec}\left[\frac{\kappa_n^2}{n}\dot\kappa_n^{-1}\sum_{t=1}^n \left(1-\varepsilon_t^2\right)\frac{\ddot F_0(x_t)}{\sigma_{0t}^2}\dot\kappa_n^{-1\prime}\right] \\
&\quad = \left[\frac{\kappa_n}{\sqrt n}\left(\dot\kappa_n^{-1}\otimes\dot\kappa_n^{-1}\right)\ddot\kappa_n\right]\frac{1}{\sqrt n}\sum_{t=1}^n (1-\varepsilon_t^2)\frac{\ddot\kappa_n^{-1}\ddot f_0(x_t)}{\kappa_n^{-1}\sigma_{0t}^2} = o_p(1)
\end{aligned}
$$

due in particular to (15) in Condition B.1(c). Consequently, it follows from (27) - (29) that

$$
-\nu_n^{-1}H_n(\theta_0)\nu_n^{-1\prime} \to_d \frac{1}{2}\int_0^1 M(r)\Omega M(r)' dr,
$$

which establishes ML2.

Third Step To establish ML3, fix δ such that $0<\delta<\varepsilon/6$, and define $\mu_n = \nu_n^{1-\delta}$ so that $\mu_n\nu_n^{-1}\to 0$ as required. Moreover, let $\bar s = \max(s_{\max}, -s_{\min}) + 1$ as in Park and Phillips (2001, Proof of Theorem 5.3). It follows that

$$
\mu_n = \begin{pmatrix} n^{1/2-\delta} & 0 & 0 \\ 0 & n^{1/2-\delta} & 0 \\ 0 & 0 & n^{1/2-\delta}\kappa_n^{-1}\dot\kappa_n \end{pmatrix},
$$

and therefore, we have

$$
\begin{aligned}
\left\| n^{1/2-\delta}\left(\alpha-\alpha_0\right)\right\| &\le 1 \qquad (30) \\
\left\| n^{1/2-\delta}\left(\beta-\beta_0\right)\right\| &\le 1 \qquad (31) \\
\left\| n^{1/2-\delta}(\dot\kappa_n/\kappa_n)'\left(\pi-\pi_0\right)\right\| &\le 1. \qquad (32)
\end{aligned}
$$

for all $\theta \in N_n$. Since

$$\begin{aligned}
&\mu_n^{-1} H_n(\theta) \mu_n^{-1\prime} \\
= & -\begin{pmatrix} n^{-1+2\delta} H_{11}^n(\theta) & n^{-1+2\delta} H_{12}^n(\theta) & n^{-1+2\delta} H_{13}^n(\theta) (\frac{\dot{\kappa}_n}{\kappa_n})^{-1\prime} \\ n^{-1+2\delta} H_{21}^n(\theta) & n^{-1+2\delta} H_{22}^n(\theta) & n^{-1+2\delta} H_{23}^n(\theta) (\frac{\dot{\kappa}_n}{\kappa_n})^{-1\prime} \\ n^{-1+2\delta} (\frac{\dot{\kappa}_n}{\kappa_n})^{-1} H_{31}^n(\theta) & n^{-1+2\delta} (\frac{\dot{\kappa}_n}{\kappa_n})^{-1} H_{32}^n(\theta) & n^{-1+2\delta} (\frac{\dot{\kappa}_n}{\kappa_n})^{-1} H_{33}^n(\theta) (\frac{\dot{\kappa}_n}{\kappa_n})^{-1\prime} \end{pmatrix},
\end{aligned}$$

it suffices to show that

$$\varpi_{1n}^2(\theta) = \left\| \frac{1}{n^{1-2\delta}} \left(H_{11}^n(\theta) - H_{11}^n(\theta_0) \right) \right\| \to_p 0 \tag{33}$$

$$\varpi_{2n}^2(\theta) = \left\| \frac{1}{n^{1-2\delta}} \left(H_{21}^n(\theta) - H_{21}^n(\theta_0) \right) \right\| \to_p 0 \tag{34}$$

$$\varpi_{3n}^2(\theta) = \left\| \frac{1}{n^{1-2\delta}} \left(H_{22}^n(\theta) - H_{22}^n(\theta_0) \right) \right\| \to_p 0 \tag{35}$$

$$\varpi_{4n}^2(\theta) = \left\| \frac{1}{n^{1-2\delta}} (\dot{\kappa}_n/\kappa_n)^{-1} \left(H_{31}^n(\theta) - H_{31}^n(\theta_0) \right) \right\| \to_p 0 \tag{36}$$

$$\varpi_{5n}^2(\theta) = \left\| \frac{1}{n^{1-2\delta}} (\dot{\kappa}_n/\kappa_n)^{-1} \left(H_{32}^n(\theta) - H_{32}^n(\theta_0) \right) \right\| \to_p 0 \tag{37}$$

$$\varpi_{6n}^2(\theta) = \left\| \frac{1}{n^{1-2\delta}} (\dot{\kappa}_n/\kappa_n)^{-1} \left(H_{33}^n(\theta) - H_{33}^n(\theta_0) \right) (\dot{\kappa}_n/\kappa_n)^{-1\prime} \right\| \to_p 0 \tag{38}$$

uniformly for all $\theta \in N_n$.

To derive (33), note that

$$\begin{aligned}
& \sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right) \left(\frac{y_{t-1}^4}{\sigma_t^4} - \frac{y_{t-1}^4}{\sigma_{0t}^4} \right) \\
= & \sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right) y_{t-1}^4 \frac{\left(\sigma_{0t}^2 + \sigma_t^2\right)\left(\sigma_{0t}^2 - \sigma_t^2\right)}{\sigma_t^4 \sigma_{0t}^4} \\
= & \sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right) \frac{y_{t-1}^4}{\sigma_{0t}^2} \frac{\sigma_{0t}^2 - \sigma_t^2}{\sigma_t^4} + \sum_{t=1}^n \left(2\varepsilon_t^2 - 1\right) \frac{y_{t-1}^4}{\sigma_{0t}^4} \frac{\sigma_{0t}^2 - \sigma_t^2}{\sigma_t^2}.
\end{aligned}$$

By the definition of σ_t^2, we have

$$\begin{aligned}
\sigma_{0t}^2 - \sigma_t^2 &= (\alpha_0 - \alpha) y_{t-1}^2 + (\beta_0 - \beta) \sigma_{0t-1}^2 + (f_0(x_t) - f(x_t)) + \beta \left(\sigma_{0t-1}^2 - \sigma_{t-1}^2\right) \\
&= (\alpha_0 - \alpha) \sum_{k=1}^\infty \beta^{k-1} y_{t-k}^2 + (\beta_0 - \beta) \sum_{k=1}^\infty \beta^{k-1} \sigma_{0t-k}^2 \\
&\quad + \sum_{k=1}^\infty \beta^{k-1} \left(f_0(x_{t-k+1}) - f(x_{t-k+1}) \right).
\end{aligned} \tag{39}$$

Using (39), we can divide $\sum_{t=1}^{n}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\frac{\sigma_{0t}^2-\sigma_t^2}{\sigma_t^4}$ into three terms. For the first term, we have

$$\begin{aligned}
&\left\|\frac{1}{n^{1-2\delta}}\sum_{t=1}^{n}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\frac{1}{\sigma_t^4}\left(\alpha_0-\alpha\right)\sum_{k=1}^{\infty}\beta^{k-1}y_{t-k}^2\right\| \\
\leq\ & \frac{1}{n^{1-2\delta}}\sum_{t=1}^{n}\left\|\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\frac{1}{\sigma_t^4}\left(\alpha_0-\alpha\right)\sum_{k=1}^{\infty}\beta^{k-1}y_{t-k}^2\right\| \\
\leq\ & \frac{n^{3\delta}}{\sqrt{n}}\left\|\kappa_n^2\sup_{|s|\leq\bar{s}}\sup_{\theta\in N_n}\left|\frac{1}{f^2\left(\sqrt{n}s,\theta\right)}\right|\right\|\frac{1}{n}\sum_{t=1}^{n}\left|\left(\kappa_n^2\right)^{-1}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\sum_{k=1}^{\infty}\beta^{k-1}y_{t-k}^2\right| \quad (40)
\end{aligned}$$

due to (30) and

$$0<\frac{1}{\sigma_t^2}\leq\frac{1}{f(x_t)}. \tag{41}$$

Note that $\frac{1}{n}\sum_{t=1}^{n}\left|\left(\kappa_n^2\right)^{-1}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\sum_{k=1}^{\infty}\beta^{k-1}y_{t-k}^2\right|=O_p(1)$ due to lemma A5. Similarly, for the second term, we have

$$\begin{aligned}
&\left\|\frac{1}{n^{1-2\delta}}\sum_{t=1}^{n}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\frac{1}{\sigma_t^4}\left(\beta_0-\beta\right)\sum_{k=1}^{\infty}\beta^{k-1}\sigma_{0t-k}^2\right\| \\
\leq\ & \frac{n^{3\delta}}{\sqrt{n}}\left\|\kappa_n^2\sup_{|s|\leq\bar{s}}\sup_{\theta\in N_n}\left|\frac{1}{f^2\left(\sqrt{n}s,\theta\right)}\right|\right\|\frac{1}{n}\sum_{t=1}^{n}\left|\left(\kappa_n^2\right)^{-1}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\sum_{k=1}^{\infty}\beta^{k-1}\sigma_{0t-k}^2\right|. \quad (42)
\end{aligned}$$

due to (31) and (41). Let $\bar{\pi}$ lies in the line segment connecting π and π_0. Then, for the third term, we have

$$\begin{aligned}
&\left\|\frac{1}{n^{1-2\delta}}\sum_{t=1}^{n}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\frac{1}{\sigma_t^4}\sum_{k=1}^{\infty}\beta^{k-1}\left(f_0(x_{t-k+1})-f(x_{t-k+1})\right)\right\| \\
\leq\ & \frac{1}{n^{1-2\delta}}\sum_{t=1}^{n}\left|\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\right|\left\|\frac{1}{\sigma_t^4}\sum_{k=1}^{\infty}\beta^{k-1}\dot{f}(x_{t-k+1},\bar{\pi})'\left(\pi_0-\pi\right)\right\| \\
\leq\ & \left\{\frac{n^{3\delta}}{\sqrt{n}}\left\|\kappa_n^2\dot{\kappa}_n^{-1}\sup_{|s|\leq\bar{s}}\left(\sup_{\theta\in N_n}\left|\frac{1}{f^2\left(\sqrt{n}s,\theta\right)}\right|\sup_{\theta\in N_n}\left|\dot{f}\left(\sqrt{n}s,\theta\right)\right|\right)\right\|\right. \\
&\left.\times\frac{1}{n}\sum_{t=1}^{n}\left|\kappa_n^{-1}\left(2\varepsilon_t^2-1\right)\frac{y_{t-1}^4}{\sigma_{0t}^2}\right|\frac{1}{1-\beta}\right\} \quad (43)
\end{aligned}$$

due to (32) and (41). Similarly, we can obtain

$$
\begin{aligned}
& \left\| \frac{1}{n^{1-2\delta}} \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) \frac{y_{t-1}^4}{\sigma_{0t}^4} \frac{\sigma_{0t}^2 - \sigma_t^2}{\sigma_t^2} \right\| \\
\leq \; & \frac{n^{3\delta}}{\sqrt{n}} \left\| \kappa_n \sup_{|s|\leq \bar{s}} \sup_{\theta \in N_n} \left| \frac{1}{f\left(\sqrt{n}s, \theta\right)} \right| \right\| \frac{1}{n} \sum_{t=1}^{n} \left| \kappa_n^{-1} \left(2\varepsilon_t^2 - 1\right) \frac{y_{t-1}^4}{\sigma_{0t}^4} \sum_{k=1}^{\infty} \beta^{k-1} y_{t-k}^2 \right| \\
& + \frac{n^{3\delta}}{\sqrt{n}} \left\| \kappa_n \sup_{|s|\leq \bar{s}} \sup_{\theta \in N_n} \left| \frac{1}{f\left(\sqrt{n}s, \theta\right)} \right| \right\| \frac{1}{n} \sum_{t=1}^{n} \left| \kappa_n^{-1} \left(2\varepsilon_t^2 - 1\right) \frac{y_{t-1}^4}{\sigma_{0t}^4} \sum_{k=1}^{\infty} \beta^{k-1} \sigma_{0t-k}^2 \right| \\
& + \left\{ \begin{array}{l} \frac{n^{3\delta}}{\sqrt{n}} \left\| \kappa_n \dot{\kappa}_n^{-1} \sup_{|s|\leq \bar{s}} \left(\sup_{\theta \in N_n} \left| \frac{1}{f\left(\sqrt{n}s,\theta\right)} \right| \sup_{\theta \in N_n} \left| \dot{f}\left(\sqrt{n}s, \theta\right) \right| \right) \right\| \\ \times \frac{1}{n} \sum_{t=1}^{n} \left| \left(2\varepsilon_t^2 - 1\right) \frac{y_{t-1}^4}{\sigma_{0t}^4} \right| \frac{1}{1-\beta} \end{array} \right\}. \qquad (44)
\end{aligned}
$$

From (40), (42), (43) and (44), (33) follows immediately due to (16). In the same way done for (33), we can also obtain (34) and (35).

To derive (36), note that

$$
\begin{aligned}
& \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) \left(\frac{\dot{f}(x_t) y_{t-1}^2}{\sigma_t^4} - \frac{\dot{f}_0(x_t) y_{t-1}^2}{\sigma_{0t}^4} \right) \\
= \; & \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) y_{t-1}^2 \left(\frac{\dot{f}(x_t) - \dot{f}_0(x_t)}{\sigma_t^4} + \frac{\dot{f}_0(x_t) \left(\sigma_{0t}^4 - \sigma_t^4\right)}{\sigma_t^4 \sigma_{0t}^4} \right). \qquad (45)
\end{aligned}
$$

For the first term in (45), we have

$$
\begin{aligned}
& \left\| \frac{\kappa_n}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) y_{t-1}^2 \frac{\dot{f}(x_t) - \dot{f}_0(x_t)}{\sigma_t^4} \right\| \\
\leq \; & \frac{1}{n^{3/2-3\delta}} \sum_{t=1}^{n} \left| \kappa_n^{-1} \left(2\varepsilon_t^2 - 1\right) y_{t-1}^2 \right| \left\| \kappa_n^3 \frac{1}{\sigma_t^4} \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \ddot{f}(x_t, \bar{\pi}) \right\| \\
\leq \; & \left\{ \begin{array}{l} \frac{n^{3\delta}}{\sqrt{n}} \left\| \kappa_n^3 \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \sup_{|s|\leq \bar{s}} \left(\sup_{\theta \in N_n} \left| \frac{1}{f^2\left(\sqrt{n}s,\theta\right)} \right| \sup_{\theta \in N_n} \left| \ddot{f}\left(\sqrt{n}s, \theta\right) \right| \right) \right\| \\ \times \frac{1}{n} \sum_{t=1}^{n} \left| \kappa_n^{-1} \left(2\varepsilon_t^2 - 1\right) y_{t-1}^2 \right| \end{array} \right\} \qquad (46)
\end{aligned}
$$

due to (32). For the second term in (45), we have

$$
\begin{aligned}
& \left\| \frac{\kappa_n}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) y_{t-1}^2 \dot{f}_0(x_t) \frac{\left(\sigma_{0t}^4 - \sigma_t^4\right)}{\sigma_t^4 \sigma_{0t}^4} \right\| \\
\leq \; & \frac{1}{n^{1-2\delta}} \sum_{t=1}^{n} \left\| \kappa_n^{-1} \dot{\kappa}_n^{-1} \left(2\varepsilon_t^2 - 1\right) y_{t-1}^2 \dot{f}_0(x_t) \right\| \left| \kappa_n^2 \frac{\left(\sigma_{0t}^4 - \sigma_t^4\right)}{\sigma_t^4 \sigma_{0t}^4} \right| \to_p 0 \qquad (47)
\end{aligned}
$$

by dealing with the last term in (47) as done for (33). From (46) and (47), (36) follows immediately due to (17). In the same way done for (36), we can obtain (37).

To establish (38), at first we will show

$$\left\| \frac{\kappa_n^2}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) \left(\frac{\dot{f}(x_t)\dot{f}(x_t)'}{\sigma_t^4} - \frac{\dot{f}_0(x_t)\dot{f}_0(x_t)'}{\sigma_{0t}^4} \right) \dot{\kappa}_n^{-1\prime} \right\| \to_p 0. \tag{48}$$

Note that

$$\begin{aligned} & \frac{\dot{f}(x_t)\dot{f}(x_t)'}{\sigma_t^4} - \frac{\dot{f}_0(x_t)\dot{f}_0(x_t)'}{\sigma_{0t}^4} \\ = & \frac{\left(\dot{f}(x_t)\dot{f}(x_t)' - \dot{f}_0(x_t)\dot{f}_0(x_t)'\right)\sigma_{0t}^4 - \dot{f}_0(x_t)\dot{f}_0(x_t)'\left(\sigma_t^4 - \sigma_{0t}^4\right)}{\sigma_t^4\sigma_{0t}^4} \\ = & \frac{\dot{f}(x_t)\left[\dot{f}(x_t)' - \dot{f}_0(x_t)'\right]}{\sigma_t^4} + \frac{\left[\dot{f}(x_t) - \dot{f}_0(x_t)\right]\dot{f}_0(x_t)'}{\sigma_t^4} + \frac{\dot{f}_0(x_t)\dot{f}_0(x_t)'\left(\sigma_{0t}^4 - \sigma_t^4\right)}{\sigma_t^4\sigma_{0t}^4}. \end{aligned}$$

For the first term, we have

$$\begin{aligned} & \left\| \frac{\kappa_n^2}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) \frac{\dot{f}(x_t)\left(\dot{f}(x_t)' - \dot{f}_0(x_t)'\right)}{\sigma_t^4} \dot{\kappa}_n^{-1\prime} \right\| \\ \leq & \frac{1}{n^{3/2-3\delta}} \sum_{t=1}^{n} \left|2\varepsilon_t^2 - 1\right| \left\| \dot{\kappa}_n^{-1}\dot{f}(x_t) \right\| \left\| \kappa_n^3 \frac{1}{\sigma_t^4} \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \ddot{f}(x_t, \bar{\pi}) \right\| \\ \leq & \left\{ \frac{n^{3\delta}}{\sqrt{n}} \left\| \dot{\kappa}_n^{-1} \sup_{|s|\leq \bar{s}} \sup_{\theta\in N_n} \left| \dot{f}\left(\sqrt{n}s, \theta\right) \right| \right\| \right. \\ & \left. \times \left\| \kappa_n^3 \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \sup_{|s|\leq \bar{s}} \left(\sup_{\theta\in N_n} \left| \frac{1}{f^2\left(\sqrt{n}s,\theta\right)} \right| \sup_{\theta\in N_n} \left| \ddot{f}\left(\sqrt{n}s, \theta\right) \right| \right) \right\| \frac{1}{n} \sum_{t=1}^{n} \left|2\varepsilon_t^2 - 1\right| \right\}. \end{aligned} \tag{49}$$

For the second term, we have

$$\begin{aligned} & \left\| \frac{\kappa_n^2}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) \frac{\left(\dot{f}(x_t) - \dot{f}_0(x_t)\right)\dot{f}_0(x_t)'}{\sigma_t^4} \dot{\kappa}_n^{-1\prime} \right\| \\ \leq & \frac{1}{n^{3/2-3\delta}} \sum_{t=1}^{n} \left\| \left(2\varepsilon_t^2 - 1\right) \dot{f}_0(x_t)'\dot{\kappa}_n^{-1\prime} \right\| \left\| \kappa_n^3 \frac{1}{\sigma_t^4} \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \ddot{f}(x_t, \bar{\pi}) \right\| \\ \leq & \left\{ \frac{n^{3\delta}}{\sqrt{n}} \left\| \kappa_n^3 \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \sup_{|s|\leq \bar{s}} \left(\sup_{\theta\in N_n} \left| \frac{1}{f^2\left(\sqrt{n}s,\theta\right)} \right| \sup_{\theta\in N_n} \left| \ddot{f}\left(\sqrt{n}s, \theta\right) \right| \right) \right\| \right. \\ & \left. \times \frac{1}{n} \sum_{t=1}^{n} \left\| \left(2\varepsilon_t^2 - 1\right) \dot{f}_0(x_t)'\dot{\kappa}_n^{-1\prime} \right\| \right\}. \end{aligned} \tag{50}$$

For the third term, we have

$$\begin{aligned}
&\left\| \frac{\kappa_n^2}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(2\varepsilon_t^2 - 1\right) \frac{\dot{f}_0(x_t)\dot{f}_0(x_t)' \left(\sigma_{0t}^4 - \sigma_t^4\right)}{\sigma_t^4 \sigma_{0t}^4} \dot{\kappa}_n^{-1\prime} \right\| \\
\leq\ & \frac{1}{n^{1-2\delta}} \sum_{t=1}^{n} \left\| \left(2\varepsilon_t^2 - 1\right) \dot{\kappa}_n^{-1} \dot{f}_0(x_t)\dot{f}_0(x_t)' \dot{\kappa}_n^{-1\prime} \right\| \left| \kappa_n^2 \frac{\sigma_{0t}^4 - \sigma_t^4}{\sigma_t^4 \sigma_{0t}^4} \right| \to_p 0 \qquad (51)
\end{aligned}$$

by dealing with the last term in (51) as done for (33). From (49)~(51), (48) follows immediately due to (17).

As the next step to establish (38), we will show

$$\left\| \frac{\kappa_n^2}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(1 - \varepsilon_t^2\right) \left(\frac{\ddot{F}(x_t)}{\sigma_t^2} - \frac{\ddot{F}_0(x_t)}{\sigma_{0t}^2} \right) \dot{\kappa}_n^{-1\prime} \right\| \to_p 0 \qquad (52)$$

Note that

$$\frac{\ddot{F}(x_t)}{\sigma_t^2} - \frac{\ddot{F}_0(x_t)}{\sigma_{0t}^2} = \frac{\ddot{F}(x_t) - \ddot{F}_0(x_t)}{\sigma_t^2} + \frac{\ddot{F}_0(x_t)\left(\sigma_{0t}^2 - \sigma_t^2\right)}{\sigma_t^2 \sigma_{0t}^2}.$$

For the first term, we have

$$\begin{aligned}
&\left\| \frac{\kappa_n^2}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(1 - \varepsilon_t^2\right) \frac{\ddot{F}(x_t) - \ddot{F}_0(x_t)}{\sigma_t^2} \dot{\kappa}_n^{-1\prime} \right\| \\
\leq\ & \frac{1}{n^{1-2\delta}} \sum_{t=1}^{n} \left|1 - \varepsilon_t^2\right| \left\| \kappa_n^2 \frac{1}{\sigma_t^2} \dot{\kappa}_n^{-1} \left(\ddot{F}(x_t) - \ddot{F}_0(x_t) \right) \dot{\kappa}_n^{-1\prime} \right\| \\
\leq\ & \left\{ \frac{n^{3\delta}}{\sqrt{n}} \left\| \kappa_n^3 \left(\dot{\kappa}_n \otimes \dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \sup_{|s| \leq \bar{s}} \left(\sup_{\theta \in N_n} \left| \frac{1}{f(\sqrt{n}s, \theta)} \right| \sup_{\theta \in N_n} \left| \dddot{f}\left(\sqrt{n}s, \theta\right) \right| \right) \right\| \right. \\
& \left. \times \frac{1}{n} \sum_{t=1}^{n} \left|1 - \varepsilon_t^2\right| \right\}. \qquad (53)
\end{aligned}$$

For the second term, we have

$$
\begin{aligned}
&\left\| \frac{\kappa_n^2}{n^{1-2\delta}} \dot{\kappa}_n^{-1} \sum_{t=1}^{n} \left(1-\varepsilon_t^2\right) \frac{\ddot{F}_0(x_t)\left(\sigma_{0t}^2-\sigma_t^2\right)}{\sigma_t^2 \sigma_{0t}^2} \dot{\kappa}_n^{-1\prime} \right\| \\
\leq\ & \frac{1}{n^{1-2\delta}} \sum_{t=1}^{n} \left| \kappa_n \frac{\left(1-\varepsilon_t^2\right)}{\sigma_{0t}^2} \right| \left\| \kappa_n \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \ddot{f}_0 \right\| \left| \frac{\sigma_{0t}^2-\sigma_t^2}{\sigma_t^2} \right| \\
\leq\ & \left\{ \tfrac{n^{3\delta}}{\sqrt{n}} \left\| \kappa_n^2 \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \sup_{|s|\leq \bar{s}} \left(\left(\sup_{\theta \in N_n} \left| \tfrac{1}{f(\lambda s,\theta)} \right| \right) \left| \ddot{f}\left(\sqrt{n}s,\theta_0\right) \right| \right) \right\| \right. \\
& \times \left[\tfrac{1}{n} \sum_{t=1}^{n} \left| \tfrac{\left(1-\varepsilon_t^2\right)}{\sigma_{0t}^2} \sum_{k=1}^{\infty} \beta^{k-1} y_{t-k}^2 \right| + \tfrac{1}{n} \sum_{t=1}^{n} \left| \tfrac{\left(1-\varepsilon_t^2\right)}{\sigma_{0t}^2} \sum_{k=1}^{\infty} \beta^{k-1} \sigma_{0t-k}^2 \right| \right] \Bigg\} \\
& + \left\{ \tfrac{n^{3\delta}}{\sqrt{n}} \left\| \dot{\kappa}_n^{-1} \sup_{|s|\leq \bar{s}} \sup_{\theta \in N_n} \left| \dot{f}\left(\sqrt{n}s,\theta\right) \right| \right\| \right. \\
& \times \left. \left\| \kappa_n^2 \left(\dot{\kappa}_n \otimes \dot{\kappa}_n\right)^{-1} \sup_{|s|\leq \bar{s}} \left(\left(\sup_{\theta \in N_n} \left| \tfrac{1}{f(\lambda s,\theta)} \right| \right) \left| \ddot{f}\left(\sqrt{n}s,\theta_0\right) \right| \right) \right\| \tfrac{1}{n} \sum_{t=1}^{n} \left| \kappa_n \tfrac{\left(1-\varepsilon_t^2\right)}{\sigma_{0t}^2} \right| \tfrac{1}{1-\beta} \right\}.
\end{aligned} \tag{54}
$$

The third line follows because of (39). From (53) and (54), (52) follows immediately due to (17) and (18). Hence, (48) and (52) yield (38). Therefore, we may easily deduce that $\varpi_{in}^2(\theta) = o_p(1)$ for $i = 1, 2, \cdots, 6$ uniformly in $\theta \in N_n$, which completes the proof. $\square$

Proof of Proposition 3 First, we need to obtain asymptotic limits for the following two sample moments:

$$
\sum_{t=1}^{n} y_t^4 \quad \text{and} \quad \sum_{t=1}^{n} y_t^2 y_{t-k}^2.
$$

The first sample moment is

$$
\begin{aligned}
\frac{\left(\kappa_0(\sqrt{n})\right)^2}{n} \sum_{t=1}^{n} y_t^4 &= \frac{1}{n} \sum_{t=1}^{n} \left(\kappa_0(\sqrt{n})^{-1} z_t f_0(x_t)\right)^2 \varepsilon_t^4 + o_p(1) \\
&= \frac{1}{n} \sum_{t=1}^{n} \mathbb{E}\left(z_t^2 \varepsilon_t^4\right) \left(\kappa_0(\sqrt{n})^{-1} f_0(x_t)\right)^2 + o_p(1) \\
&\rightarrow {}_d \mathbb{E}\left(z_t^2 \varepsilon_t^4\right) \int_0^1 \overline{f}_0^2(V_c(r)) dr
\end{aligned}
$$

due to lemma A1, A3 and A4. Since

$$
z_t = z_{t-1}\left(\beta + \alpha \varepsilon_{t-1}^2\right) + 1,
$$

we have

$$
z_t^2 = z_{t-1}^2 \left(\beta + \alpha \varepsilon_{t-1}^2\right)^2 + 2 z_{t-1} \left(\beta + \alpha \varepsilon_{t-1}^2\right) + 1
$$

and, due to (12) and (14), we can obtain

$$\mathbb{E}\left(z_t^2\right) = \frac{1+(\alpha+\beta)}{1-(\alpha+\beta)}\frac{1}{1-\left(\alpha^2\kappa^4+2\alpha\beta+\beta^2\right)}. \tag{55}$$

Since $\mathbb{E}\left(z_t^2\varepsilon_t^4\right) = \mathbb{E}\left(z_t^2\right)\kappa^4$, we have

$$\frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n}\sum_{t=1}^{n} y_t^4 \to_d \frac{1+(\alpha+\beta)}{1-(\alpha+\beta)}\frac{\kappa^4}{1-\left(\alpha^2\kappa^4+2\alpha\beta+\beta^2\right)}\int_0^1 \overline{f}_0^2(V_c(r))dr. \tag{56}$$

The second moment is

$$\begin{aligned}
&\frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n}\sum_{t=k+1}^{n} y_t^2 y_{t-k}^2 \\
= \; & \frac{1}{n}\sum_{t=k+1}^{n}\left(\kappa_0(\sqrt{n})^{-1}z_t f_0(x_t)\right)\left(\kappa_0(\sqrt{n})^{-1}z_{t-k}f_0(x_{t-k})\right)\varepsilon_t^2\varepsilon_{t-k}^2 + o_p(1) \\
= \; & \frac{1}{n}\sum_{t=k+1}^{n}\left(z_t z_{t-k}\varepsilon_t^2\varepsilon_{t-k}^2\right)\left(\kappa_0(\sqrt{n})^{-1}f_0(x_{t-k})\right)^2 + o_p(1) \\
\to \; & {}_d\mathbb{E}\left(z_t z_{t-k}\varepsilon_t^2\varepsilon_{t-k}^2\right)\int_0^1 \overline{f}_0^2(V_c(r))dr
\end{aligned}$$

due to lemma A1-A4. The third line follows because we have from lemma A2 that

$$\kappa_0(\sqrt{n})^{-1}z_t f_0(x_t) = \kappa_0(\sqrt{n})^{-1}z_t f_0(x_{t-k}) + o_p(1)$$

uniformly in $t = 1, \ldots, n$. Note that this is a different approach from the way by Park (2002) and Han and Park (2006). See the proof of theorem 1 in both papers. Their cases consider more general volatility functions including *integrable* functions, but this paper considers only asymptotically homogeneous functions with Assumption 3. Due to Assumption 3, we have lemma A2 and we can handle $\sum_{t=k+1}^{n} y_t^2 y_{t-k}^2$ in an easier way.

By the recursive substitution, we have

$$z_t = z_{t-k}\prod_{i=1}^{k}\left(\beta+\alpha\varepsilon_{t-i}^2\right) + \sum_{j=1}^{k-1}\prod_{i=1}^{j}\left(\beta+\alpha\varepsilon_{t-i}^2\right) + 1.$$

Therefore,

$$
\begin{aligned}
& \mathbb{E}\left(z_t z_{t-k}\varepsilon_t^2\varepsilon_{t-k}^2\right) = \mathbb{E}\left(z_t z_{t-k}\varepsilon_{t-k}^2\right) \\
= \; & \mathbb{E}\left(z_{t-k}^2\varepsilon_{t-k}^2\prod_{i=1}^{k}\left(\beta+\alpha\varepsilon_{t-i}^2\right) + z_{t-k}\varepsilon_{t-k}^2\sum_{j=1}^{k-1}\prod_{i=1}^{j}\left(\beta+\alpha\varepsilon_{t-i}^2\right) + z_{t-k}\varepsilon_{t-k}^2\right) \\
= \; & \mathbb{E}\left(z_{t-k}^2\right)(\alpha+\beta)^{k-1}\left(\beta+\alpha\kappa^4\right) + \mathbb{E}\left(z_{t-k}\right)\left(\sum_{j=1}^{k-1}(\beta+\alpha)^j + 1\right) \\
= \; & \mathbb{E}\left(z_{t-k}^2\right)(\alpha+\beta)^{k-1}\left(\beta+\alpha\kappa^4\right) + \mathbb{E}\left(z_{t-k}\right)\frac{1-(\alpha+\beta)^k}{1-(\alpha+\beta)}.
\end{aligned}
$$

Using (12) and (55), we have

$$
\begin{aligned}
& \frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n}\sum_{t=k+1}^{n} y_t^2 y_{t-k}^2 \\
\to_d \; & \left[\frac{1-(\alpha+\beta)^k}{[1-(\alpha+\beta)]^2} + \frac{1+(\alpha+\beta)}{1-(\alpha+\beta)}\frac{(\alpha+\beta)^{k-1}\left(\beta+\alpha\kappa^4\right)}{1-\left(\alpha^2\kappa^4+2\alpha\beta+\beta^2\right)}\right]\int_0^1 \bar{f}_0^2(V_c(r))dr. \quad (57)
\end{aligned}
$$

Since

$$
\kappa_0(\sqrt{n})^{-1}\bar{y}_n^2 = \frac{\kappa_0(\sqrt{n})^{-1}}{n}\sum_{t=1}^{n} y_t^2,
$$

it follows from (56), (57) and lemma A5 that

$$
\begin{aligned}
& \frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n}\sum_{t=k+1}^{n}\left(y_t^2-\bar{y}_n^2\right)\left(y_{t-k}^2-\bar{y}_n^2\right) \\
= \; & \frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n}\sum_{t=k+1}^{n} y_t^2 y_{t-k}^2 - \left(\kappa_0(\sqrt{n})^{-1}\bar{y}_n^2\right)^2 + O_p(n^{-1/2}) \\
\to_d \; & \left[\frac{1-(\alpha+\beta)^k}{[1-(\alpha+\beta)]^2} + \frac{1+(\alpha+\beta)}{1-(\alpha+\beta)}\frac{(\alpha+\beta)^{k-1}\left(\beta+\alpha\kappa_\varepsilon^4\right)}{1-\left(\alpha^2\kappa^4+2\alpha\beta+\beta^2\right)}\right]\int_0^1 \bar{f}_0^2(V_c(r))d \\
& -\left(\frac{1}{1-(\alpha+\beta)}\int_0^1 \bar{f}_0(V_c(r))dr\right)^2
\end{aligned}
$$

and

$$\begin{aligned}
&\frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n}\sum_{t=1}^{n}\left(y_t^2-\bar{y}_n^2\right)^2 \\
= \quad & \frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n}\sum_{t=1}^{n}y_t^4-\left(\kappa_0(\sqrt{n})^{-1}\bar{y}_n^2\right)^2+O_p(n^{-1/2}) \\
\to_d \quad & \frac{1+(\alpha+\beta)}{1-(\alpha+\beta)}\frac{\kappa^4}{1-\left(\alpha^2\kappa^4+2\alpha\beta+\beta^2\right)}\int_0^1\bar{f}_0^2(V_c(r))dr \\
& -\left(\frac{1}{1-(\alpha+\beta)}\int_0^1\bar{f}_0(V_c(r))dr\right)^2
\end{aligned}$$

The stated result follows immediately. □

Proof of Proposition 4 Let $\theta=(\alpha,\beta)'$. Since

$$\frac{\partial\sigma_t^2}{\partial\theta}=\left(\frac{\partial\sigma_t^2}{\partial\alpha},\frac{\partial\sigma_t^2}{\partial\beta}\right)'=\left(y_{t-1}^2,\sigma_{t-1}^2\right)',$$

the score function is given by

$$s_n(\theta)=\frac{1}{n}\sum_{t=1}^{n}\left(\frac{y_t^2}{\sigma_t^2(\theta)}-1\right)\left(\frac{y_{t-1}^2}{\sigma_t^2(\theta)},\frac{\sigma_{t-1}^2(\theta)}{\sigma_t^2(\theta)}\right)' \tag{58}$$

where

$$\sigma_t^2(\theta)=\omega+\alpha y_{t-1}^2+\beta\sigma_{t-1}^2=\frac{\omega}{1-\beta}+\alpha\sum_{k=1}^{\infty}\beta^{k-1}y_{t-k}^2.$$

Since it follows from Lemmas A1 and A2 that

$$\kappa_0(\sqrt{n})^{-1}\sigma_t^2(\theta)=\left(\kappa_0(\sqrt{n})^{-1}f_0(x_t)\right)\alpha\sum_{k=1}^{\infty}\beta^{k-1}m_{t-k}+o_p(1) \tag{59}$$

uniformly in $t=1,\ldots,n$ and uniformly in $\theta\in\Theta$, we have

$$s_n(\theta)=\begin{pmatrix}\frac{1}{n}\sum_{t=1}^{n}\left(\frac{m_t}{\alpha\sum_{k=1}^{\infty}\beta^{k-1}m_{t-k}}-1\right)\frac{m_{t-1}}{\alpha\sum_{k=1}^{\infty}\beta^{k-1}m_{t-k}}+o_p(1)\\ \frac{1}{n}\sum_{t=1}^{n}\left(\frac{m_t}{\alpha\sum_{k=1}^{\infty}\beta^{k-1}m_{t-k}}-1\right)\frac{\sum_{k=2}^{\infty}\beta^{k-2}m_{t-k}}{\sum_{k=1}^{\infty}\beta^{k-1}m_{t-k}}+o_p(1)\end{pmatrix}$$

uniformly in $\theta\in\Theta$. Since

$$\frac{\sum_{k=2}^{\infty}\beta^{k-2}m_{t-k}}{\sum_{k=1}^{\infty}\beta^{k-1}m_{t-k}}=\frac{1}{\beta}-\frac{m_{t-1}}{\sum_{k=1}^{\infty}\beta^{k-1}m_{t-k}},$$

the stated result follows from the proof of Han and Park (2007, proposition 3). Note that m_t in their paper is defined by $z_t = 1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(\alpha\varepsilon_{t-i}^2\right)$ instead of $z_t = 1 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \left(\beta + \alpha\varepsilon_{t-i}^2\right)$.
□

Proof of Proposition 5 If we write

$$K_n^4 = \frac{\left(\kappa_0(\sqrt{n})^{-1}\right)^2}{n} \sum_{t=1}^{n} y_t^4 \Bigg/ \left(\frac{\kappa_0(\sqrt{n})^{-1}}{n} \sum_{t=1}^{n} y_t^2 \right)^2$$

the stated result follows directly from the proof of Proposition 4. □

Table 1. QMLE for GARCH(1,1) model and kurtosis of stock return series

Group	Series	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}+\hat{\beta}$	sample kurtosis	implied kurtosis of (ε_t)
Group 1	S&P 500 monthly	0.136	0.836	0.972	14.67	3.28
		(0.042)	(0.039)			
	S&P 500 weekly	0.094	0.898	0.992	6.53	2.20
		(0.036)	(0.038)			
	S&P 500 daily	0.047	0.949	0.996	6.87	3.09
		(0.013)	(0.014)			
	KO	0.037	0.959	0.997	6.43	3.38
		(0.013)	(0.015)			
	XOM	0.052	0.942	0.994	5.99	3.17
		(0.012)	(0.014)			
Group 2	JNJ	0.077	0.898	0.975	9.54	4.94
		(0.020)	(0.024)			
	MCD	0.039	0.953	0.993	7.50	4.61
		(0.014)	(0.016)			
	MMM	0.018	0.978	0.996	6.36	5.28
		(0.032)	(0.047)			
	MO	0.038	0.953	0.990	18.11	8.36
		(0.020)	(0.029)			
	S	0.091	0.877	0.968	27.96	6.79
		(0.034)	(0.042)			

Notes: The table reports the result of fitting the GARCH(1,1) model to various stock return series. The model is of the form

$$\begin{aligned} y_t &= c + \sigma_t \varepsilon_t, \\ \sigma_t^2 &= \omega + \alpha \left(y_{t-1} - c\right)^2 + \beta \sigma_{t-1}^2, \end{aligned}$$

where (y_t) is a stock return series. QMLE standard errors are reported in parentheses. The implied kurtosis of (ε_t) is a calculated value of κ^4 from (11). The sample period of the monthly S&P 500 return series is from January 1919 to June 2003 (1,014 observations). It is 23 October, 1982 to 27 June, 2003 at the weekly frequency (1,079 observations) and 2 November, 1987 to 30 June, 2003 at the daily frequency (3,938 Observations). See Han and Park (2006) for more details. The rest are stock return series of individual firms for the sample period of January 5, 1993 to December 31, 2003 (2,770 observations). These firms are Coca-Cola (KO), Exxon-Mobil (XOM), Johnson & Johnson (JNJ), McDonald's (MCD), 3M (MMM), Philip Morris (MO), and Sears (S).

Table 2. Simulated values of α_* and β_*

		$\varepsilon_t \tilde{} iid\ \mathbb{N}(0,1)$			$\varepsilon_t \tilde{} iid\ \mathbb{U}\left[-\sqrt{3},\sqrt{3}\right]$		
α_0	β_0	α_*	β_*	$\alpha_* + \beta_*$	α_*	β_*	$\alpha_* + \beta_*$
0.1	0.1	0.00	1.00	1.00	0.00	1.00	1.00
	0.2	0.00	1.00	1.00	0.00	1.00	1.00
	0.3	0.00	1.00	1.00	0.00	1.00	1.00
	0.4	0.00	1.00	1.00	0.00	1.00	1.00
	0.5	0.01	1.00	1.00	0.01	0.99	1.00
	0.6	0.01	0.99	1.00	0.01	0.99	1.00
0.2	0.1	0.01	0.99	1.00	0.01	0.99	1.00
	0.2	0.01	0.99	1.00	0.01	0.99	1.00
	0.3	0.03	0.98	1.00	0.02	0.98	1.00
	0.4	0.04	0.96	1.00	0.04	0.96	1.00
	0.5	0.07	0.93	1.01	0.08	0.93	1.00
0.3	0.1	0.05	0.95	1.00	0.05	0.96	1.00
	0.2	0.07	0.94	1.01	0.08	0.92	1.00
	0.3	0.11	0.90	1.01	0.12	0.88	1.01
	0.4	0.16	0.86	1.02	0.18	0.84	1.01

Notes: For the results in Table 2, the GARCH(1,1) processes with the GARCH coefficients α_0 and β_0 are generated and fitted into the GARCH(1,1) model without constant to obtain the MLE $\bar{\alpha}_n$ and $\bar{\beta}_n$ from the samples of size 2,000. The reported values of α_* and β_* are then obtained in each case by taking averages of $\bar{\alpha}_n$ and $\bar{\beta}_n$ over 100 iterations. The values of $\bar{\alpha}_n$ and $\bar{\beta}_n$ show very little variation, and are very close respectively to α_* and β_* in every iteration. The first 1000 observations are discarded from the initial samples of size 3,000 to get rid of the start-up effect. The results are invariant with respect to the value of constant term in the GARCH models.

References

Bai, X., Russell, J.R. and Tiao, G.C. (2003). Kurtosis of GARCH and stochastic volatility models with non-normal innovations, *Journal of Econometrics,* 114, 349-360.

Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics,* 74, 3-30.

Berkes, L., Horvath, L., and Kokoszka, P. (2003): GARCH processes: structure and estimation, *Bernoulli*, 9, 201-207.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics,* 31, 307-327.

Bollerslev, T. (1987), A conditionally heteroskedastic time series for speculative prices and rates of return, *The Review of Economics and Statistics,* 69, 542-547.

Diebold, F.X. (1986). Modelling the persistence of conditional variance: A comment, *Econometric Reviews*, 5, 51-56.

Diebold, F.X. and Inoue, A. (2001), Long memory and regime switching, *Journal of Econometrics,* 105, 131-159.

Ding, Z. and Granger, C.W.J. (1996), Modeling volatility persistence of speculative returns: A new approach, *Journal of Econometrics,* 73, 185-215.

Ding, Z., Granger, C.W.J., and Engle, R.F. (1993), A long memory property of stock market returns and a new model, *Journal of Empirical Finance,* 1, 83-106.

Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. Inflation, *Econometrica*, 50, 987-1008.

Engle, R.F. and Bollerslev, T. (1986). Modelling the persistence of conditional variance, *Econometric Reviews*, 5, 1-50, 81-87.

Engle, R.F. and Lee, G.J. (1999), A long-run and short-run component model of stock return volatility, in Engle R. and H. White ed. *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger*, Chapter 10, 475-497, Oxford University Press.

Engle, R.F. and Patton, A.J. (2001). What good is a volatility model?, *Quantitative Finance*, 1(2), 237-245.

Engle, R.F. and Rangel, J.G. (2005), The spline GARCH model for unconditional volatility and its global macroeconomic causes, mimeograph NYU and UCSD.

Glosten, L.R., Jagannathan, R. and Runkle, D. (1993). On the relation between the expected value and the volatility of nominal excess returns on stocks, *Journal of Finance,* 48, 1779-1801.

Granger, C.W.J. and Hyung, N. (2004), Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns, *Journal of Empirical Finance,* 11, 399-421.

Gray, S.F. (1996), Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics*, 42, 27-62.

Hagiwara, M. and Herce, M.A. (1999). Endogenous exchange rate volatility, trading volume and interest rate differentials in a model of portfolio selection, *Review of International Economics*, 7(2), 202-218.

Han, H. and Park, J.Y. (2006). Time series properties of ARCH processes with persistent covariates, RMI working paper 06/11, Risk Management Institute, National University of Singapore.

Han, H. and Park, J.Y. (2007), ARCH with Persistent Covariate, mimeograph, Department of Economics, National University of Singapore.

He, C. and Teräsvirta, T. (1999). Fourth moment structure of the GARCH(p, q) processes, *Econometric Theory*, 15, 824-846.

Hillebrand, E. (2005). Neglecting parameter changes in GARCH models, *Journal of Econometrics,* 129, 121-138.

Hodrick R.J. (1989), Risk, uncertainty, and exchange rates, *Journal of Monetary Economics*, 23, 433-459.

Jensen, S.T. and Rahbek, A. (2004). Asymptotic inference for nonstationary GARCH, *Econometric Theory*, 20, 1203-1226.

Lamoureux, C.G. and Lastrapes, W.D. (1990). Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics,* 8, 225-234.

Lee, S.W. and Hansen, B.E. (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator, *Econometric Theory*, 10, 29-52.

Lumsdaine, R.L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, *Econometrica*, 64, 575-596.

Mikosch, T. and Starica, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects, *The Review of Economics and Statistics,* 86, 378-390.

Nelson, D.B. (1990). Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory*, 6, 318-334.

Park, J.Y. (2002). Nonstationary nonlinear heteroskedasticity, *Journal of Econometrics*, 110, 383-415.

Park, J.Y. (2003), Weak unit root, mimeograph, Department of Economics, Rice University.

Park, J.Y. and Phillips, P.C.B. (2001). Nonlinear regressions with integrated time series, *Econometrica*, 69, 117-161.

Starica, C. and Granger, C.W.J. (2005), Nonstationarities in stock returns, *The Review of Economics and Statistics,* 87, 503-522.

Wooldridge, J.M. (1994). Estimation and inference for dependent processes, In R.F. Engle and D.L. McFadden, eds., *Handbook of Econometrics*, Vol. 4, 2639-2738, Elsevier: Amsterdam.