Econometric Theory, **22**, 2006, 403–428. Printed in the United States of America.
DOI: 10.1017/S0266466606060208

EMPIRICAL LIKELIHOOD FOR GARCH MODELS

NGAI HANG CHAN
Chinese University of Hong Kong

SHIQING LING
Hong Kong University of Science and Technology

This paper develops an empirical likelihood approach for regular generalized autoregressive conditional heteroskedasticity (GARCH) models and GARCH models with unit roots. For regular GARCH models, it is shown that the log empirical likelihood ratio statistic asymptotically follows a χ^2 distribution. For GARCH models with unit roots, two versions of the empirical likelihood methods, the least squares score and the maximum likelihood score functions, are considered. For both cases, the limiting distributions of the log empirical likelihood ratio statistics are established. These two statistics can be used to test for unit roots under the GARCH framework. Finite-sample performances are assessed through simulations for GARCH models with unit roots.

1. INTRODUCTION

For independent and identically distributed (i.i.d.) random variables $y_1, \dots, y_n$, the distribution F of y_i is usually estimated by the empirical distribution function $F_n = \sum_{i=1}^n p_i I_{\{y_i \le x\}}$ subject to the constraints $\sum_{i=1}^n p_i = 1$ and $p_i \ge 0$. This empirical distribution is maximized at $p_i = 1/n$, $i = 1, \dots, n$. Owen (1988, 1990) defines the empirical likelihood ratio as

$$R(F) = \prod_{i=1}^n np_i,$$

where $p_i = dF(y_i) = P(Y = y_i)$. If the distribution function F is characterized by an unknown parameter λ, then the probabilities p_i should be subject to some restrictions on the parameter λ. For example, when F is characterized by the mean μ, then p_i needs to satisfy the first-order restriction $\sum_{i=1}^n p_i(x_i - \mu) = 0$.

This research was supported in part by Hong Kong Research grants Council Grants CUHK4043/02P and HKUST6273/03H. The authors thank two referees and the Co-Editor, Bruce Hansen, for insightful and helpful comments about the relationship between QMLE and MELE, which led to substantial improvement of the presentation. Computational assistance from Jerry Wong and Chun-Yip Yau is also gratefully acknowledged. Address correspondence to Ngai Hang Chan, Department of Statistics, Chinese University of Hong Kong, Shatin, NT, Hong Kong; e-mail: nhchan@sta.cuhk.edu.hk.

 0266-4666/06 $12.00

More generally, if λ satisfies a k-dimensional vector equation $g(x_i, \lambda) = 0$, then the first-order restriction on p_i is

$$\sum_{i=1}^{n} p_i QQ\breve{Q}QQg(x_i, \lambda) = 0. \tag{1.1}$$

For this restriction, the (profile) empirical likelihood ratio is given by

$$R(\lambda) = \sup\left\{\prod_{i=1}^{n} np_i : p_i \geq 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i g(x_i, \lambda) = 0\right\}.$$

For a given λ, a unique maximum exists if zero lies inside the convex hull of the points $g(x_i, \lambda)$, $i = 1, \ldots, n$. The maximum of $R(\lambda)$ can be found via a simple Lagrange multiplier argument. Let

$$L(\lambda) = \sum_{i=1}^{n} \log p_i + a\left(1 - \sum_{i=1}^{n} p_i\right) - nb' \sum_{i=1}^{n} p_i g(x_i, \lambda),$$

where a and b are Lagrange multipliers. Taking derivatives with respect to p_i and using the restriction (1.1) and $\sum_{i=1}^{n} p_i = 1$, one obtains

$$\sum_{i=1}^{n} \frac{g(x_i, \lambda)}{1 + b' g(x_i, \lambda)} = 0. \tag{1.2}$$

As shown in Qin and Lawless (1994), if the $k \times k$ matrix $\sum_{i=1}^{n} [g(x_i, \lambda) g'(x_i, \lambda)]$ is positive definite, by the inverse function theorem there exists a continuous differentiable function $b(\lambda)$ of λ such that

$$\sum_{i=1}^{n} \frac{g(x_i, \lambda)}{1 + b'(\lambda) g(x_i, \lambda)} = 0. \tag{1.3}$$

The (profile) log empirical likelihood function can be defined as

$$L_E(\lambda) = \sum_{i=1}^{n} \log[1 + b'(\lambda) g(x_i, \lambda)].$$

Its minimizer $\hat{\lambda}_n$ is called the maximum empirical likelihood estimator (MELE). In practice, one is mainly interested in the value of the MELE $\hat{\lambda}_n$ evaluated at the corresponding empirical likelihood ratio statistic defined by

$$W_E(\lambda_0) = -2[L_E(\hat{\lambda}_n) - L_E(\lambda_0)].$$

This statistic can be used to test for the hypothesis $H_0 \colon \lambda = \lambda_0$. Owen (1988) demonstrates that the empirical likelihood approach provides an accurate confidence region for the parameter in finite-sample cases.

When $g(x, \lambda) = x - \mu$ and under some mild conditions, Owen (1988, 1990) proves that $W_E(\mu_0)$ converges in distribution to χ^2_p as $n \to \infty$, where p is the

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

dimension of μ. Owen (1991) and Kolaczyk (1994) extend this methodology to general regression problems including linear, generalized linear, and projection pursuit models, Qin and Lawless (1994) to general independent unbiased estimating functions, and more recently Chuang and Chan (2002) to unstable autoregressive models studied in Chan and Wei (1988).

As pointed out in Owen (1990), the empirical likelihood approach plays a similar role to the bootstrap method. On the other hand, it is also well known that bootstrap methods work less efficiently for serially correlated data than for independent data. It is therefore interesting to explore whether the empirical likelihood approach provides another feasible means to supplement the bootstrap methodology to conduct statistical inference for time series data.

The main goal of this paper is to study the effect of the empirical methodology for unit root models with generalized autoregressive conditional heteroskedasticity (GARCH) errors. It is well known that unit root testing has played an important role in econometrics during the last two decades. Although one can apply the Dickey–Fuller tests for unit root models with GARCH errors (see Pantula, 1989), by accounting for the heterogeneity presented in the GARCH component, more powerful unit root tests based on the quasi-maximum likelihood estimator (QMLE) can be constructed (see Ling and Li, 1998, 2003; Seo, 1999). Simulation studies based on quasi-maximum likelihood estimation can be found in Seo (1999) and Ling, Li, and McAleer (2003). Given these findings, it is interesting to examine unit root tests based on the empirical method (see, e.g., Wright, 1999).

This paper proceeds as follows. Section 2 develops the empirical likelihood for GARCH models and gives its asymptotic properties. Section 3 studies the empirical likelihood for the unit root with GARCH models. The asymptotic properties of the MELEs are obtained, and the unit root test statistics using the empirical likelihood are proposed. Section 4 reports simulation results, and Section 5 concludes. Proofs of the results are given in the Appendix.

Throughout the paper, the following notation will be used: $o(1)$ ($o_p(1)$) denotes a term (a random variable) that converges to zero (in probability); $O(1)$ ($O_p(1)$) denotes a term (a random variable) that is bounded (in probability); $\|\cdot\|$ denotes the Euclidean norm; and $\to_{\mathcal{L}}$ denotes convergence in distribution as the sample size n tends to infinity.

2. EMPIRICAL LIKELIHOOD FOR GARCH MODELS

Consider the GARCH model defined by the equation

$$\varepsilon_t = \eta_t \sqrt{h_t}, \qquad h_t = \omega + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i}, \tag{2.1}$$

where η_t are i.i.d. random variables with mean zero and variance 1. Let $\lambda = (\omega, \alpha_1, \alpha_2, \dots, \alpha_r, \beta_1, \dots, \beta_s)'$ and let the parameter space Θ be a compact sub-

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

set of R^{r+s+1}. Let $\lambda = \lambda_0 \in \Theta$ be the true parameter. Assume that λ_0 is an interior point and for each $\lambda \in \Theta$, it satisfies the following conditions.

Assumption 2.1. $\omega > 0$, $\alpha_i, \beta_i \geq 0$, $\sum_{i=1}^r \alpha_i + \sum_{i=1}^s \beta_i < 1$, and $\sum_{i=1}^r \alpha_i z^i$ and $1 - \sum_{i=1}^s \beta_i z^i$ have no common roots.

Given the observations $\varepsilon_n, \ldots, \varepsilon_1$ and the initial value $\bar{\varepsilon}_0 = \{\varepsilon_t : t \leq 0\}$, the log-likelihood function (modulus of a constant) can be written as

$$L_n(\lambda) = \sum_{t=1}^n l_t(\lambda) \quad \text{and} \quad l_t(\lambda) = -\frac{1}{2}\log h_t(\lambda) - \frac{\varepsilon_t^2}{2h_t(\lambda)}, \tag{2.2}$$

where $h_t(\lambda)$ is a function of ε_t and λ is the vector of parameters defined in (2.1). Because η_t may not be normally distributed, strictly speaking, the function (2.2) is the quasi-likelihood function and its maximizer is the QMLE. The score function and the information matrix are, respectively,

$$\frac{\partial l_t(\lambda)}{\partial \lambda} = \frac{1}{2h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \lambda}\left[\frac{\varepsilon_t^2}{h_t(\lambda)} - 1\right] \quad \text{and}$$

$$\frac{\partial^2 l_t(\lambda)}{\partial \lambda \partial \lambda'} = -\frac{\varepsilon_t^2}{2h_t^3(\lambda)} \frac{\partial h_t(\lambda)}{\partial \lambda} \frac{\partial h_t(\lambda)}{\partial \lambda'} - \frac{1}{2h_t(\lambda)}\left[\frac{\partial^2 h_t(\lambda)}{\partial \lambda \partial \lambda'} - \frac{1}{h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \lambda} \frac{\partial h_t(\lambda)}{\partial \lambda'}\right]\left[1 - \frac{\varepsilon_t^2}{h_t(\lambda)}\right].$$

Let $D_t(\lambda) = \partial l_t(\lambda)/\partial\lambda$ and $P_t(\lambda) = \partial^2 l_t(\lambda)/\partial\lambda\partial\lambda'$. The QMLE is the solution of the score function $\sum_{t=1}^n D_t(\lambda)$. Using this score function $D_t(\lambda)$, the (profile) log empirical likelihood function can be constructed as follows:

$$L_E(\lambda) = \sum_{t=1}^n \log[1 + b'(\lambda) D_t(\lambda)], \tag{2.3}$$

where $b(\lambda)$ is the Lagrange multiplier, that is, a solution to the equation

$$\sum_{t=1}^n \frac{D_t(\lambda)}{1 + b'(\lambda) D_t(\lambda)} = 0.$$

The global minimizer $\hat{\lambda}_n$ of (2.3) is the QMLE and $b(\hat{\lambda}_n) = 0$. Define the empirical likelihood ratio statistic for testing $H_0 : \lambda = \lambda_0$ as

$$W_E(\lambda_0) = -2[L_E(\hat{\lambda}_n) - L_E(\lambda_0)].$$

We have the following result.

THEOREM 2.1. *If Assumption 2.1 holds and* $E\eta_t^4 < \infty$, *then* $W_E(\lambda_0) \to_{\mathcal{L}} \chi^2_{r+s+1}$ *under* H_0.

Remark 2.1. If one is only interested in testing for a subset of parameters, one can construct the empirical likelihood ratio statistic along the lines of Theorem 1 in Kitamura (1997). See Kitamura (1997) and Kitamura, Tripathi, and Ahn (2004) for a proof of consistency of the global minimizer of the log empirical likelihood function.

3. UNIT ROOT WITH GARCH MODELS

Consider the following unit root with GARCH(1,1) model:

$$y_t = y_{t-1} + \varepsilon_t, \tag{3.1}$$

$$\varepsilon_t = \eta_t \sqrt{h_t} \quad \text{and} \quad h_t = \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1},$$

where η_t are i.i.d. random variables with mean zero and variance 1. Let $\delta = (\omega, \alpha, \beta)'$ and let the parameter space Θ be a compact subset of R^3 and $\delta = \delta_0 \in \Theta$ be the true parameter. Assume that δ_0 is an interior point in Θ and for each $\delta \in \Theta$, it satisfies the following conditions.

Assumption 3.1. $\omega, \alpha, \beta > 0$ and $\alpha + \beta < 1$.

The results of this section can be extended to a higher order GARCH(r,s) model by means of the results in Ling and Li (1998). It is also straightforward to extend the results of this section to the nearly unit root case as in Chan and Wei (1987) and Chuang and Chan (2002).

In general, the autoregressive parameter ϕ can be estimated by two methods: least squares estimator (LSE) and QMLE. When the LSE is used, the score function is given by

$$u_t(\phi) = (y_t - \phi y_{t-1}) y_{t-1}.$$

As in Section 1, this equation gives the form of the empirical likelihood function as

$$L_E(\phi) = \sum_{t=1}^{n} \log[1 + b(\phi) u_t(\phi)], \tag{3.2}$$

where $b(\phi)$ is the Lagrange multiplier, which is a solution to the equation

$$\sum_{t=1}^{n} u_t(\phi)/[1 + b(\phi) u_t(\phi)] = 0.$$

The minimizer $\hat{\phi}_n$ of (3.2) is the LSE of $\phi_0 = 1$. The empirical likelihood ratio statistic for testing $H_0\colon \phi_0 = 1$ is

$$W_E(\phi_0) = -2[L_E(\hat{\phi}_n) - L_E(\phi_0)].$$

The following theorem gives the limiting distribution of $W_E(\phi_0)$.

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

THEOREM 3.1. *If Assumption 3.1 holds, then*

$$W_E(\phi_0) \to_{\mathcal{L}} \left(\int_0^1 B^2(\tau)\, d\tau \right)^{-1} \left(\int_0^1 B(\tau)\, dB(\tau) \right)^2,$$

as $n \to \infty$, where $B(\tau)$ is a standard Brownian motion.

Consider now the empirical likelihood ratio statistic based on the quasi-likelihood score. The quasi-likelihood function conditional on the initial value $y_0 = 0$ can be written as

$$L_n(\lambda) = \sum_{t=1}^{n} l_t(\lambda) \quad \text{where } l_t(\lambda) = -\frac{1}{2} \log h_t(\lambda) - \frac{\varepsilon_t^2(\lambda)}{2h_t(\lambda)},$$

where $\lambda = (\phi, \delta)$ and $\varepsilon_t(\lambda) = y_t - \phi y_{t-1}$ is a function of y_t and λ. Then

$$\frac{\partial l_t(\lambda)}{\partial \lambda} = -\frac{\varepsilon_t(\lambda)}{h_t(\lambda)} \frac{\partial \varepsilon_t(\lambda)}{\partial \lambda} + \frac{1}{2h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \lambda} \left(\frac{\varepsilon_t^2(\lambda)}{h_t(\lambda)} - 1 \right),$$

$$\begin{aligned} \frac{\partial^2 l_t(\lambda)}{\partial \lambda \partial \lambda'} = &- \frac{1}{h_t(\lambda)} \frac{\partial \varepsilon_t(\lambda)}{\partial \lambda} \frac{\partial \varepsilon_t(\lambda)}{\partial \lambda'} + \frac{2\varepsilon_t(\lambda)}{h_t^2(\lambda)} \frac{\partial \varepsilon_t(\lambda)}{\partial \lambda} \frac{\partial h_t(\lambda)}{\partial \lambda'} \\ &- \frac{\varepsilon_t^2(\lambda)}{2h_t^3(\lambda)} \frac{\partial h_t(\lambda)}{\partial \lambda} \frac{\partial h_t(\lambda)}{\partial \lambda'} \\ &- \frac{1}{2h_t(\lambda)} \left[\frac{\partial^2 h_t(\lambda)}{\partial \lambda \partial \lambda'} - \frac{1}{h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \lambda} \frac{\partial h_t(\lambda)}{\partial \lambda'} \right] \left[1 - \frac{\varepsilon_t^2(\lambda)}{h_t(\lambda)} \right]. \end{aligned}$$

Let $D_t(\lambda) = \partial l_t(\lambda)/\partial \lambda$ and $P_t(\lambda) = \partial^2 l_t(\lambda)/\partial \lambda \partial \lambda'$. Using this score function as in Section 1, the empirical likelihood function can be constructed as

$$L_E(\lambda) = \sum_{t=1}^{n} \log[1 + b'(\lambda) D_t(\lambda)], \tag{3.3}$$

where $b(\lambda)$ is the Lagrange multiplier that is a solution to the equation

$$\sum_{t=1}^{n} \frac{D_t(\lambda)}{1 + b'(\lambda) D_t(\lambda)} = 0.$$

The minimizer $\hat{\lambda}_n$ of (3.3) is the QMLE of λ_0, the true parameter of λ. The empirical likelihood ratio statistic for testing $H_0\colon \lambda_0 = (1, \delta_0')'$ is

$$W_E(\lambda_0) = -2[L_E(\hat{\lambda}_n) - L_E(\lambda_0)].$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

THEOREM 3.2. *If Assumption 3.1 holds and $E|\eta_t|^{4+\iota} < \infty$ for some $\iota > 0$, then*

$$W_E(\lambda_0) \to_{\mathcal{L}} \left[K \int_0^1 \omega_1^2(\tau)\, d\tau\right]^{-1} \left[\int_0^1 \omega_1(\tau)\, d\omega_2(\tau)\right]^2 + \chi_3^2 \quad \text{under } H_0,$$

as $n \to \infty$, where Ω and κ are defined as in Lemma A.2 with $r = s = 1$ and the two components in the limiting distribution are independent. Herein, $(\omega_1(\tau), \omega_2(\tau))$ is a bivariate Brownian motion with covariance

$$\tau\Sigma = \tau \begin{pmatrix} Eh_t & 1 \\ 1 & K \end{pmatrix},$$

where $F = E(1/h_t) + 2\alpha^2 \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2/h_t^2)$ and $K = E(1/h_t) + \kappa\alpha^2 \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2/h_t^2)$.

If we are interested in testing the unit root model (3.1), we have to find the constrained empirical likelihood $\min_{\phi=1} L_E(\lambda)$. The empirical likelihood test for testing $\phi = 1$ is defined as

$$\widetilde{W}_n = -2\left[L_E(\hat{\lambda}_n) - \min_{\phi=1,\, \delta\in\Theta} L_E(\lambda)\right].$$

Let $L_E(\delta) = L_E(\lambda)|_{\lambda=(1,\delta)}$ and its minimizer be denoted by $\hat{\hat{\delta}}_n$. Here, the number of estimating equations is four, whereas that of unknown parameters is three. Thus, $\hat{\hat{\delta}}_n$ is no longer the same as the QMLE of δ_0 in the finite sample. The following lemma gives the existence of $\hat{\hat{\delta}}_n$.

LEMMA 3.1. *If Assumption 3.1 holds and $E|\eta_t|^{4+\iota} < \infty$ for some $\iota > 0$, then there exists a point $\hat{\hat{\delta}}_n$ in the interior of the ball $V_n = \{\delta : \|\delta - \delta_0\| \le n^{-0.5+\tilde{\iota}}\}$ for some $\tilde{\iota} \in (0, \iota)$ such that $\partial L_E(\hat{\hat{\delta}}_n)/\partial\delta = 0$ with probability approaching 1 and $\lim_{n\to\infty} P\{L_E(\hat{\hat{\delta}}_n) = \min_{\delta\in V_n} L_E(\delta)\} = 1$.*

Because $L_E(\hat{\lambda}_n)$ is always zero, we only need to find $\hat{\hat{\delta}}_n$ numerically. The following corollary gives the limiting distributions of the $\widetilde{W}_n$.

COROLLARY 3.1. *Under the assumptions of Lemma 3.1, it follows that*

$$\widetilde{W}_n^{1/2} \to_{\mathcal{L}} \frac{\rho \int_0^1 B(\tau)\, dB(\tau)}{\left\{\int_0^1 B^2(\tau)\, d\tau\right\}^{1/2}} + \{1 - \rho^2\}^{1/2}\xi,$$

as $n \to \infty$, where $\rho^2 = 1/(\sigma^2 K) \in (0,1)$, $\sigma^2 = Eh_t$, $\xi \sim N(0,1)$, and $B(\tau)$ is a standard Brownian motion that is independent of ξ.

Remark 3.1. $F = K$ when $\kappa = 2$. The limiting distributions in Corollary 3.1 are the same as those given in Seo (1999) and Ling and Li (2003). Some of the critical values can be found in Ling et al. (2003).

4. SIMULATIONS

Finite-sample performances of the unit root tests based on the QMLE and the MELE are examined via Monte Carlo experiments in this section.

The model is given in (3.1) with $\eta_t \sim N(0,1)$ and the parameters (ω, α, β) assume the same values given in Table 1. Furthermore, the parameter ϕ takes the values $\phi = 1.0, 0.99, 0.95$, and 0.9. The sample size is 200, and 1,000 replications are conducted in all cases. Let $\hat{t}^2_{QEn}$ denote the square of the t-test based on the QMLE given in Ling and Li (1998), that is, $\hat{t}^2_{QEn} = \{(1/\sigma\rho)(\sum_{t=2}^n y_{t-1}^2)^{1/2}(\hat{\phi}_{MLE} - 1)\}^2$, which has the same asymptotic distribution as $\widetilde{W}_n$ in Corollary 3.1. The critical values of $\hat{t}^2_{QEn}$ and $\widetilde{W}_n$ are computed by 20,000 replications of the integral of standard Brownian motion appearing on the right-hand side of Corollary 3.1, which is approximated by a discrete random walk model using normal errors. From Table 1, we see that the sizes of both tests are almost identical and are very close to the nominal level 0.05. This is similar to the results reported by Chuang and Chan (2002), who compare the unit root tests based on the empirical likelihood and the LSE methods for the autoregressive (AR) model with i.i.d errors. The powers of $\widetilde{W}_n$ are smaller than

Table 1. Powers of lower tail unit root tests at 5% level for AR(1)-GARCH(1,1) models based on empirical critical values with standard normal errors: $n = 200$ and 1,000 replications

ϕ_0	0.900	0.950	0.990	1.000
$(\omega, \alpha, \beta) = (0.1, 0.4, 0.5)$				
$\hat{t}^2_{QEn}$	1.000	0.938	0.778	0.054
$\widetilde{W}_n$	0.981	0.847	0.705	0.056
$(\omega, \alpha, \beta) = (0.1, 0.3, 0.6)$				
$\hat{t}^2_{QEn}$	0.999	0.910	0.715	0.066
$\widetilde{W}_n$	0.974	0.798	0.625	0.063
$(\omega, \alpha, \beta) = (0.3, 0.2, 0.75)$				
$\hat{t}^2_{QEn}$	0.996	0.892	0.690	0.059
$\widetilde{W}_n$	0.985	0.807	0.586	0.057
$(\omega, \alpha, \beta) = (0.4, 0.5, 0.2)$				
$\hat{t}^2_{QEn}$	0.998	0.932	0.703	0.050
$\widetilde{W}_n$	0.993	0.890	0.640	0.060

Table 2. Powers of lower tail unit root tests at 5% level for AR(1)-GARCH(1,1) models based on empirical critical values with t_5 errors: $n = 200$ and 1,000 replications

ϕ_0	0.900	0.950	0.990	1.000
$(\omega,\alpha,\beta) = (0.1,0.4,0.5)$				
$\hat{t}^2_{QEn}$	0.991	0.879	0.265	0.085
$\widetilde{W}_n$	0.935	0.777	0.275	0.124
$(\omega,\alpha,\beta) = (0.1,0.3,0.6)$				
$\hat{t}^2_{QEn}$	0.984	0.850	0.247	0.068
$\widetilde{W}_n$	0.926	0.739	0.242	0.091
$(\omega,\alpha,\beta) = (0.3,0.2,0.75)$				
$\hat{t}^2_{QEn}$	0.988	0.824	0.236	0.073
$\widetilde{W}_n$	0.946	0.732	0.246	0.123
$(\omega,\alpha,\beta) = (0.4,0.5,0.2)$				
$\hat{t}^2_{QEn}$	0.991	0.852	0.223	0.067
$\widetilde{W}_n$	0.953	0.793	0.276	0.116

those of $\hat{t}^2_{QEn}$, however. One possible explanation is that the calculation of the MELE of δ_0 is not as efficient as the QMLE in finite samples. We need to solve seven highly nonlinear equations to obtain the minimum of $L_E(\delta)$ in $\widetilde{W}_n$ when evaluating the MELE.

In Table 2, we repeat the simulations with the distribution of η_t being replaced by t_5 in Table 1. It is now observed that the sizes of both tests are distorted, with the MELE test $\widetilde{W}_n$ suffering more seriously. As in the normal error case, the powers of $\widetilde{W}_n$ are lower than those of $\hat{t}^2_{QEn}$, except for $\phi = 0.99$. When the sample size is increased to 600 in Table 3, the sizes of both tests become much better, and powers have also been increased. Differences between the powers of two tests become smaller and for $\phi = 0.99$, the power of $\widetilde{W}_n$ becomes lower than $\hat{t}^2_{QEn}$. Because $\phi = 0.99$ is so close to the unit root case, the power of $\widetilde{W}_n$ is affected by its overrejection when the sample size n is 200, but this overrejection becomes less serious when n is increased to 600.

In general, the unit root test based the MELE does not perform better than that based on the QMLE in finite samples. Although the simulation results are somewhat disappointing, if it were not for the asymptotic results established in Theorem 3.2, it would not be possible to conduct the comparison between the empirical likelihood and MLE methods, so we would never know how they perform relative to each other. The knowledge gained by conducting this study is valuable. In the i.i.d. case, Kitamura (2001) showed that the empirical likelihood ratio test is no less powerful than any regular test as the sample size goes to infinity. Our finding only reveals finite-sample situations, which cannot

TABLE 3. Powers of lower tail unit root tests at 5% level for AR(1)-GARCH(1,1) models based on empirical critical values with t_5 errors: $n = 600$ and 1,000 replications

ϕ_0	0.900	0.950	0.990	1.000
$(\omega, \alpha, \beta) = (0.1, 0.4, 0.5)$				
$\hat{t}^2_{QEn}$	1.000	0.999	0.671	0.064
$\widetilde{W}_n$	0.995	0.993	0.607	0.083
$(\omega, \alpha, \beta) = (0.1, 0.3, 0.6)$				
$\hat{t}^2_{QEn}$	1.000	1.000	0.624	0.064
$\widetilde{W}_n$	0.994	0.971	0.561	0.072
$(\omega, \alpha, \beta) = (0.3, 0.2, 0.75)$				
$\hat{t}^2_{QEn}$	0.998	0.999	0.629	0.075
$\widetilde{W}_n$	0.994	0.980	0.522	0.079
$(\omega, \alpha, \beta) = (0.4, 0.5, 0.2)$				
$\hat{t}^2_{QEn}$	1.000	1.000	0.626	0.054
$\widetilde{W}_n$	0.997	0.989	0.600	0.084

be considered evidence contradictory to his findings. It remains, however, a challenging theoretical problem to derive a similar conclusion as in Kitamura (2001) for the MELE in the unit root case.

5. CONCLUDING REMARKS

This paper develops the empirical likelihood approach for GARCH models and GARCH models with unit roots. For GARCH models, it is shown that the log empirical likelihood ratio statistic asymptotically follows a χ^2 distribution. For GARCH models with unit roots, the empirical likelihood method based on the least squares score and the maximum likelihood score functions is investigated. In both cases, the limiting distributions of the log empirical likelihood ratio statistics are established and the unit root tests based on the MELEs are constructed. Numerical simulations are conducted to assess the finite performance.

REFERENCES

Chan, N.H. & C.Z. Wei (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1062.

Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

Chuang, C.S. & N.H. Chan (2002) Empirical likelihood for autoregressive models with applications to unstable time series. *Statistica Sinica* 12, 387–407.
Chung, K.L. (1968) *A Course in Probability Theory.* Academic Press.
Francq, C. & J.M. Zakoïan (2004) Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* 10, 605–637.
Kitamura, Y. (1997) Empirical likelihood methods with weakly dependent processes. *Annals of Statistics* 25, 2084–2102.
Kitamura, Y. (2001) Asymptotic optimality of empirical likelihood for testing moment restrictions. *Econometrica* 69, 1661–1672.
Kitamura, Y., G. Tripathi, & H. Ahn (2004) Empirical likelihood-based inference in conditional moment restriction models. *Econometrica* 72, 1667–1714.
Kolaczyk, E.D. (1994) Empirical likelihood for generalized linear models. *Statistica Sinica* 4, 199–218.
Lee, S.-W. & B.E. Hansen (1994) Asymptotic theory for GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* 10, 29–52.
Ling, S. (2006) Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH /IGARCH models. *Journal of Econometrics*, submitted.
Ling, S. & W.K. Li (1998) Limiting distributions of maximum likelihood estimators for unstable autoregressive moving-average time series with general autoregressive heteroskedastic errors. *Annals of Statistics* 26, 84–125.
Ling, S. & W.K. Li (2003) Asymptotic inference for unit root with GARCH(1,1) errors. *Econometric Theory* 19, 541–564.
Ling, S., W.K. Li, & M. McAleer (2003) Estimation and testing for unit root processes with GARCH (1,1) errors: Theory and Monte Carlo study. *Econometric Reviews* 22, 179–202.
Owen, A. (1988) Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* 75, 237–249.
Owen, A. (1990) Empirical likelihood ratio confidence regions. *Annals of Statistics* 18, 90–120.
Owen, A. (1991) Empirical likelihood for linear models. *Annals of Statistics* 19, 1725–1747.
Qin, J. & J. Lawless (1994) Empirical likelihood and general estimating equations. *Annals of Statistics* 22, 300–325.
Pantula, S.G. (1989) Estimation of autoregressive models with ARCH errors. *Sankhyā B* 50, 119–138.
Phillips, P.C.B. (1989) Partially identified econometric models. *Econometric Theory* 5, 181–240.
Seo, B. (1999) Distribution theory for unit root tests with conditional heteroskedasticity. *Journal of Econometrics* 91, 113–144.
Wright, J. (1999) An empirical likelihood ratio test for a unit root, problem 99.2.1. *Econometric Theory* 15, 257.

APPENDIX: Proofs

For ease of reference, Lemma A.2 of Ling (2006) is repeated as follows. In the lemma, ε_t and $h_t(\lambda)$ are defined as in Section 2. Proof of this statement can be found in Ling (2006).

Lemma A.2 of Ling (2006). *If Assumption 2.1 holds, then there exist a neighborhood Θ_0 of λ_0, a constant $C > 0$, and a constant $\rho \in (0,1)$ such that for any constant $\iota_1 > 0$, it follows that*

$$(a)\quad \sup_{\Theta_0} \left\| \frac{1}{h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \lambda} \right\| \le C\left(1 + \sum_{j=1}^{\infty} \rho^j |\varepsilon_{t-j}|\right)^{\iota_1},$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

$$(b)\quad \sup_{\Theta_0}\left\|\frac{1}{h_t(\lambda)}\frac{\partial^2 h_t(\lambda)}{\partial\lambda\partial\lambda'}\right\| \le C\left(1+\sum_{j=1}^{\infty}\rho^j|\varepsilon_{t-j}|\right)^{\iota_1}.$$

The proof of Theorem 2.1 follows a similar idea as in Owen (1990) and Qin and Lawless (1994). We first present two lemmas.

LEMMA A.1 *If Assumption 2.1 holds and* $E\eta_t^4 < \infty$, *then*

$$(a)\quad E\sup_{\Theta_0}\|D_t(\lambda)\|^2 < \infty,$$

$$(b)\quad E\sup_{\Theta_0}\|P_t(\lambda)\| < \infty,$$

$$(c)\quad E\sup_{\Theta_0}\left|\frac{\partial^3 l_t(\lambda)}{\partial\lambda_i\partial\lambda_j\partial\lambda_k}\right| < \infty,$$

where $i,j,k = 1,\ldots,r+s+1$ *and* Θ_0 *is some neighborhood of* λ_0.

Proof. By Lemma A.2 of Ling (2006), we have

$$\sup_{\Theta_0}\left\|\frac{1}{h_t(\lambda)}\frac{\partial h_t(\lambda)}{\partial\lambda}\right\| \le \xi_{ht-1} \quad \text{and} \quad \sup_{\Theta_0}\left\|\frac{1}{h_t(\lambda)}\frac{\partial^2 h_t(\lambda)}{\partial\lambda\partial\lambda'}\right\| \le \xi_{ht-1},$$

where $\xi_{ht-1} = C(1+\sum_{j=1}^{\infty}\rho^j|\varepsilon_{t-j}|)^{\iota_1}$ for some constants $\rho \in (0,1)$ and $C > 0$ and for any $\iota_1 > 0$. By the two inequalities, it is easy to show that (a) and (b) hold. Using the same method as in Lemma A.2 of Ling (2006), it follows that

$$\sup_{\Theta_0}\left\|\frac{1}{h_t(\lambda)}\frac{\partial^3 h_t(\lambda)}{\partial\lambda_i\partial\lambda_j\partial\lambda_k}\right\| \le \xi_{ht-1},$$

from which part (c) follows. ■

LEMMA A.2. *If Assumption 2.1 holds and* $E\eta_t^4 < \infty$, *then*

$$(a)\quad \max_{1\le t\le n}\|P_t(\lambda)\| = o_p(n),$$

$$(b)\quad \max_{1\le t\le n}\|D_t(\lambda)\| = o_p(n^{0.5}),$$

$$(c)\quad n^{-1}\sum_{t=1}^{n}P_t(\lambda) = \Omega + o_p(1),$$

$$(d)\quad n^{-1}\sum_{t=1}^{n}[D_t(\lambda)D_t'(\lambda)] = \frac{\kappa}{2}\Omega + o_p(1),$$

where $o_p(1)$ *terms hold uniformly in* $\lambda \in V_{0n} = \{\lambda : \|\lambda - \lambda\| \le n^{-0.5}M\}$ *for each constant* $M > 0$, $\Omega = E\{[\partial h_t(\lambda)/\partial\lambda][\partial h_t(\lambda)/\partial\lambda']/2h_t^2(\lambda)]\}|_{\lambda=\lambda_0}$ *and* $\kappa = E\eta_t^4 - 1$.

Proof. By part (b) of the preceding lemma, $E(\sup_{\Theta_0}\|P_t(\lambda)\|^{0.5})^2 < \infty$ for some neighborhood Θ_0 of λ_0. Furthermore, because $\{\sup_{\Theta_0}\|P_t(\lambda)\|\}$ is strictly stationary, it follows that

$$n^{-0.5}\max_{1\le t\le n}\sup_{\Theta_0}\|P_t(\lambda)\|^{0.5} = o_p(1).$$

Here, we use the fact that, if $\{X_n\}$ is an identically distributed sequence with $EX_t^2 < \infty$, then $\max_{1\le t\le n}|X_t| = o_p(n^{-0.5})$; see, for example, Chung (1968, p. 93). Thus,

$$n^{-1}\max_{1\le t\le n}\sup_{\Theta_0}\|P_t(\lambda)\| = \left[n^{-1/2}\max_{1\le t\le n}\sup_{\Theta_0}\|P_t(\lambda)\|^{0.5}\right]^2 = o_p(1).$$

Part (a) follows.

Now consider part (b). By part (a) of the preceding lemma, $E\|D_t(\lambda_0)\|^2 < \infty$. Similar to part (a), we have

$$n^{-0.5}\max_{1\le t\le n}\|D_t(\lambda_0)\| = o_p(1).$$

By the previous two equations,

$$n^{-0.5}\max_{1\le t\le n}\sup_{V_{0n}}\|D_t(\lambda)\| \le n^{-0.5}\max_{1\le t\le n}\|D_t(\lambda_0)\| + n^{-1}\max_{1\le t\le n}\sup_{\Theta_0}\|P_t(\lambda)\| = o_p(1).$$

Part (b) follows.

For part (c), let $\epsilon > 0$ be given.

$$P\left(\frac{1}{n}\sum_{t=1}^{n}\sup_{V_{0n}}\|P_t(\lambda) - P_t(\lambda_0)\| > \epsilon\right) \le \frac{1}{\epsilon}E\sup_{V_{0n}}\|P_t(\lambda) - P_t(\lambda_0)\| \to 0,$$

as $n\to\infty$, by means of part (b) of the preceding lemma and the dominated convergence theorem. By the ergodic theorem,

$$n^{-1}\sum_{t=1}^{n}P_t(\lambda_0) = \Omega + o_p(1).$$

By the previous two equations, (c) holds.

Finally, by Taylor's expansion, and parts (a)–(c) of this lemma, when $\lambda \in V_{0n}$, it follows that

$$\begin{aligned}
&n^{-1}\left\|\sum_{t=1}^{n}[D_t(\lambda)D_t'(\lambda) - D_t(\lambda_0)D_t'(\lambda_0)]\right\| \\
&\quad\le 2n^{-1}\sum_{t=1}^{n}\|n^{-0.5}D_t(\lambda_0)\|\,\|P_t(\lambda^*)\| + n^{-2}\sum_{t=1}^{n}\|P_t(\lambda^*)\|^2 \\
&\quad\le 2\max_{1\le t\le n}\|n^{-0.5}D_t(\lambda_0)\|\frac{1}{n}\sum_{t=1}^{n}\|P_t(\lambda^*)\| \\
&\qquad+ \max_{1\le t\le n}\|n^{-1}P_t(\lambda^*)\|\frac{1}{n}\sum_{t=1}^{n}\|P_t(\lambda^*)\| = o_p(1),
\end{aligned}$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

where the $o_p(1)$ term holds uniformly in V_{0n} and λ^* lies between λ_0 and λ. By the ergodic theorem,

$$n^{-1}\sum_{t=1}^{n} D_t(\lambda_0)D_t'(\lambda_0) = \frac{\kappa}{2}\Omega + o_p(1).$$

By the previous two equations, part (d) follows. ■

Proof of Theorem 2.1. Denote

$$Q_{1n}(\lambda,b) = \frac{1}{n}\sum_{t=1}^{n}\frac{D_t(\lambda)}{1+b'D_t(\lambda)}. \tag{A.1}$$

Let $b = \rho\theta$ with $\|\theta\| = 1$. Observe that

$$\begin{aligned}
0 &= \|Q_{1n}(\lambda,\rho\theta)\| \geq \|\theta' Q_{1n}(\lambda,\rho\theta)\| \\
&= \left|\frac{1}{n}\sum_{t=1}^{n}\theta' D_t(\lambda) - \frac{1}{n}\rho\sum_{t=1}^{n}\frac{\theta' D_t(\lambda)D_t'(\lambda)\theta}{1+\rho\theta' D_t(\lambda)}\right| \\
&\geq -\frac{1}{n}\left|\sum_{t=1}^{n}\theta' D_t(\lambda)\right| + \frac{\rho\theta' S_n(\lambda)\theta}{1+\rho Z_n(\lambda)}, \qquad \text{(A.2)}
\end{aligned}$$

where $S_n(\lambda) = n^{-1}\sum_{t=1}^{n} D_t(\lambda)D_t'(\lambda)$ and $Z_n(\lambda) = \max_{1\leq t\leq n}\|D_t(\lambda)\|$. By the central limit theorem, $n^{-0.5}\sum_{t=1}^{n} D_t(\lambda_0) = O_p(1)$. Furthermore, by Lemma A.2(c), we have

$$\begin{aligned}
n^{-1}\left|\sum_{t=1}^{n}\theta' D_t(\lambda)\right| &\leq n^{-1}\left|\sum_{t=1}^{n}\theta' D_t(\lambda_0)\right| + n^{-1}\left|\sum_{t=1}^{n}\theta' P_t(\lambda^*)\right|\|\lambda-\lambda_0\| \\
&\leq n^{-1}\left|\sum_{t=1}^{n}\theta' D_t(\lambda_0)\right| + n^{-1}\sup_{V_{0n}}\sum_{t=1}^{n}|P_t(\lambda)|\|\lambda-\lambda_0\| \\
&= O_p(n^{-0.5}), \qquad \text{(A.3)}
\end{aligned}$$

uniformly in $\lambda \in V_{0n}$ where V_{0n} is defined in Lemma A.2. By Lemma A.2(d), $\theta' S_n(\lambda)\theta \geq \kappa a/2 + o_p(1)$ uniformly in $\lambda \in V_{0n}$, where a is the smallest eigenvalue of Ω. By virtue of equations (A.2) and (A.3) and this fact, we have

$$\frac{\rho\theta' S_n(\lambda)\theta}{1+\rho Z_n(\lambda)} = O_p(n^{-0.5}),$$

uniformly in $\lambda \in V_{0n}$. Furthermore, by part (b) of our Lemma A.2, we have

$$\rho = \|b\| = O_p(n^{-0.5}),$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

uniformly in $\lambda \in V_{0n}$. Let $\gamma_t = b'D_t(\lambda)$, where b is the solution of equation $Q_{1n}(\lambda, b) = 0$. By part (b) of our Lemma A.2,

$$\max_{1\le t\le n} |\gamma_t| = O_p(n^{-0.5}) \max_{1\le t\le n} \|D_t(\lambda)\| = o_p(1), \tag{A.4}$$

uniformly in $\lambda \in V_{0n}$. Let $Q_{2n}(\lambda, b) = n^{-1}\sum_{t=1}^{n} P_t(\lambda)b/[1 + b'D_t(\lambda)]$. Taking the derivative with respect to λ_i (the ith element of λ), we have

$$\frac{\partial Q_{2n}(\lambda,b)}{\partial \lambda_i} = \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{[\partial P_t(\lambda)/\partial\lambda_i]b}{1+b'D_t(\lambda)} - \frac{P_t(\lambda)bb'[\partial D_t(\lambda)/\partial\lambda_i]}{[1+b'D_t(\lambda)]^2}\right\}.$$

By virtue of the previous equation, Lemma A.1(c), and our Lemma A.2(a) and (c), it follows that

$$\begin{aligned}\left\|\frac{\partial Q_{2n}(\lambda,b)}{\partial \lambda_i}\right\| &\le \frac{\|b\|}{1-\max\limits_{1\le t\le n}|\gamma_t|}\frac{1}{n}\sum_{t=1}^{n}\left\|\frac{\partial P_t(\lambda)}{\partial\lambda_i}\right\| \\ &\quad + \frac{\max\limits_{1\le t\le n}\|P_t(\lambda)\|\,\|b\|^2}{\left(1-\max\limits_{1\le t\le n}|\gamma_t|\right)^2}\frac{1}{n}\sum_{t=1}^{n}\|P_t(\lambda)\| \\ &= o_p(1),\end{aligned} \tag{A.5}$$

where γ_t is defined as in (A.4). Taking the derivative with respect to b,

$$\frac{\partial Q_{2n}(\lambda,b)}{\partial b'} = \frac{1}{n}\sum_{t=1}^{n}\left\{\frac{P_t(\lambda)}{1+b'D_t(\lambda)} - \frac{D_t(\lambda)[P_t(\lambda)b]'}{[1+b'D_t(\lambda)]^2}\right\}.$$

Furthermore, by Lemma A.2(c), we have

$$\begin{aligned}\left\|\frac{\partial Q_{2n}(\lambda,b)}{\partial b'} - \Omega\right\| &\le \left\|\frac{1}{n}\sum_{t=1}^{n}\frac{b'D_t(\lambda)P_t(\lambda)}{1+b'D_t(\lambda)}\right\| + \left\|\frac{1}{n}\sum_{t=1}^{n}[P_t(\lambda)-\Omega]\right\| \\ &\quad + \frac{\max\limits_{1\le t\le n}|\gamma_t|}{\left(1-\max\limits_{1\le t\le n}|\gamma_t|\right)^2}\frac{1}{n}\sum_{t=1}^{n}\|P_t(\lambda)\| \\ &\le \frac{\max\limits_{1\le t\le n}|\gamma_t|}{1-\max\limits_{1\le t\le n}|\gamma_t|}\frac{1}{n}\sum_{t=1}^{n}\|P_t(\lambda)\| + o_p(1) \\ &= o_p(1).\end{aligned} \tag{A.6}$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

Note that $o_p(1)$ in (A.5) and (A.6) holds uniformly in V_{0n}. Similarly, we can show that

$$\partial Q_{1n}(\lambda,b)/\partial\lambda' = \Omega + o_p(1) \quad \text{and} \quad \partial Q_{1n}(\lambda,b)/\partial b' = -\frac{\kappa}{2}\,\Omega + o_p(1),$$

where $o_p(1)$ holds uniformly in V_{0n}. By virtue of this fact and equations (A.5) and (A.6), it follows that

$$\begin{bmatrix} \partial Q_{1n}(\lambda,b)/\partial b' & \partial Q_{1n}(\lambda,b)/\partial\lambda' \\ \partial Q_{2n}(\lambda,b)/\partial b' & \partial Q_{2n}(\lambda,b)/\partial\lambda' \end{bmatrix} = \begin{pmatrix} -\kappa\Omega/2 & \Omega \\ \Omega & 0 \end{pmatrix} + o_p(1), \tag{A.7}$$

where $o_p(1)$ holds uniformly in V_{0n}.

By Lemma A.2(d), it follows that

$$\begin{aligned} 0 = Q_{1n}(\lambda,b) &= \frac{1}{n}\sum_{t=1}^{n} D_t(\lambda) - S_n(\lambda)b + \frac{1}{n}\sum_{t=1}^{n}\frac{D_t(\lambda)\gamma_t^2}{1+\gamma_t} \\ &\le \frac{1}{n}\sum_{t=1}^{n} D_t(\lambda) - S_n(\lambda)b + \frac{\|b\| \max_{1\le t\le n}|\gamma_t|}{1-\max_{1\le t\le n}|\gamma_t|}\,\frac{1}{n}\sum_{t=1}^{n}\|D_t(\lambda)\|^2 \\ &= \frac{1}{n}\sum_{t=1}^{n} D_t(\lambda) - S_n(\lambda)b + o_p(n^{-0.5}), \end{aligned}$$

uniformly in $\lambda \in V_{0n}$. Thus,

$$b_0 = [nS_n(\lambda_0)]^{-1}\sum_{t=1}^{n} D_t(\lambda_0) + o_p(n^{-0.5}).$$

Because $\hat\lambda_n$ is the QMLE, by Theorem 2.2 in Francq and Zakoïan (2004) (see also Lee and Hansen, 1994), we have that

$$\sqrt{n}(\hat\lambda_n - \lambda_0) = -\Omega^{-1}\sqrt{n}Q_{1n}(\lambda_0,0) + o_p(1) \to_{\mathcal{L}} N(0,\kappa\Omega^{-1}/2). \tag{A.8}$$

Using Taylor's expansion at $(\hat\lambda_n, \hat b_n)$, (A.7), and the preceding two equations, we have

$$\begin{aligned} W_E(\lambda_0) &= -2[L_E(\hat\lambda_n) - L_E(\lambda_0)] \\ &= n[(\hat b_n - b_0)', (\hat\lambda_n - \lambda_0)']\left\{\begin{bmatrix} -\kappa\Omega/2 & \Omega \\ \Omega & 0 \end{bmatrix} + o_p(1)\right\}[(\hat b_n - b_0)', (\hat\lambda_n - \lambda_0)']' \\ &= nQ_{1n}'(\lambda_0,0)\left(\frac{2}{\kappa}\,\Omega\right)^{-1} Q_{1n}(\lambda_0,0) + o_p(1) \\ &\to_{\mathcal{L}} \chi^2_{r+s+1}, \end{aligned} \tag{A.9}$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

as $n \to \infty$, where the last step follows in view of (A.8). This completes the proof of Theorem 2.1. ■

The proof of Theorem 3.1 is similar to Theorem 3.2. Thus, we only present the proof of Theorem 3.2. We first need two preliminary lemmas.

LEMMA A.3. *If Assumption 3.1 holds and* $E\eta_t^4 < \infty$, *then it follows that*

$$(a)\ \ -n^{-1/2}\sum_{t=1}^{n}[N_n D_t(\lambda_0)] \to_{\mathcal{L}} \left[\int_0^1 w_1(\tau)\,dw_2(\tau), N\left(0, \frac{\kappa}{2}\Omega\right)\right],$$

$$(b)\ \ n^{-1}\sum_{t=1}^{n} N_n P_t(\lambda_0) N_n \to_{\mathcal{L}} diag\left\{F\int_0^1 w_1^2(\tau)\,d\tau, \Omega\right\},$$

$$(c)\ \ n^{-1}\sum_{t=1}^{n}[N_n D_t(\lambda_0) D_t'(\lambda_0) N_n] \to_{\mathcal{L}} diag\left\{K\int_0^1 w_1^2(\tau)\,d\tau, \frac{\kappa}{2}\Omega\right\},$$

where $N_n = diag\{n^{-0.5}, I_3\}$ *and F and K are defined in Theorem 3.2.*

Proof. Parts (a) and (b) follow from Lemmas 4.7 and 4.8 in Ling and Li (2002) and Theorem 2.2 in Ling et al. (2003). The proof of part (c) is similar to that of Lemma 4.7 in Ling and Li (2003) and Theorem 2.2 in Ling et al. (2003), and hence the details are omitted. ■

LEMMA A.4. *If Assumption 3.1 holds and* $E|\eta_t|^{4+\iota} < \infty$ *for some* $\iota > 0$, *then it follows that*

$$(a)\ \ \max_{1\le t\le n}\|N_n P_t(\lambda) N_n\| = o_p(n^{1-2\tilde{\iota}}),$$

$$(b)\ \ \max_{1\le t\le n}\|N_n D_t(\lambda)\| = o_p(n^{0.5-\tilde{\iota}}),$$

$$(c)\ \ n^{-1}\sum_{t=1}^{n} N_n P_t(\lambda) N_n = n^{-1}\sum_{t=1}^{n} N_n P_t(\lambda_0) N_n + o_p(1),$$

$$(d)\ \ n^{-1}\sum_{t=1}^{n}[N_n D_t(\lambda) D_t'(\lambda) N_n] = n^{-1}\sum_{t=1}^{n}[N_n D_t(\lambda_0) D_t'(\lambda_0) N_n] + o_p(1),$$

where $o_p(1)$ *holds uniformly in* $\lambda \in V_{1n} = \{\lambda : n|\phi - 1| \le M \text{ and } \|\delta - \delta\| \le n^{-0.5+\tilde{\iota}}\}$ *for each constant* $M > 0$ *and some constant* $\tilde{\iota} \in (0,0.5)$.

Proof. By Lemma 4.2 of Ling and Li (2003) and the continuous mapping theorem, it follows that $\max_{1\le t\le n}|y_t| = O_p(n^{1/2})$. Thus,

$$\varepsilon_t(\lambda) = \varepsilon_t - [n(\phi - 1)](n^{-0.5} y_{t-1}) n^{-0.5} = \varepsilon_t + O_p(n^{-0.5}), \quad \textbf{(A.10)}$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

where $O_p(n^{-0.5})$ holds uniformly in V_{1n} and in $t = 1,\dots,n$. Using a similar method as for Lemma A.6(ii) in Ling (2006), we can show that

$$h_t(\lambda) = h_t + o_p(1)h_t, \tag{A.11}$$

where $o_p(1)$ holds uniformly in V_{1n} and in $t = 1,\dots,n$, and h_t is defined as in (2.1) with $r = s = 1$ and $\lambda = \delta_0$. Because $E|\eta_t|^{4+\iota} < \infty$, there is a small $\tilde{\iota} \in [0,\iota)$ such that $E|\eta_t|^{(2+2\iota_0)/(0.5-2\tilde{\iota})} = E|\eta_t|^{(4+4\iota_0)/(1-4\tilde{\iota})} < \infty$ and $Eh_t^{\iota_0/(0.5-2\tilde{\iota})} < \infty$ as ι_0 is zero or small enough. Hence,

$$\begin{aligned} n^{-0.5+2\tilde{\iota}} \max_{1\le t\le n} (\eta_t^{2+2\iota_0} h_t^{\iota_0}) &= \left[n^{-1/2} \max_{1\le t\le n} (\eta_t^{(1+\iota_0)/(0.5-2\tilde{\iota})} h_t^{\iota_0/2(0.5-2\tilde{\iota})}) \right]^{2(0.5-2\tilde{\iota})} \\ &= o_p(1). \end{aligned} \tag{A.12}$$

Because $|\varepsilon_{t-k}(\lambda)| h_t^{-1/2}(\lambda) = O(\beta^{-k/2})$ as $k \ge 1$, and $\max_{1\le t\le n}|y_t| = O_p(n^{1/2})$, we have

$$\begin{aligned} \max_{1\le t\le n} |h_t(\lambda)^{-1/2} \partial h_t(\lambda)/\partial\phi| &= O(1) \max_{1\le t\le n} \left| h_t(\lambda)^{-1/2} \left[\sum_{i=1}^{t-1} \beta^{i-1} y_{t-i-1}\, \varepsilon_{t-i}(\lambda) \right] \right| \\ &\le O(1) \max_{1\le t\le n} |y_t| = O_p(n^{1/2}), \end{aligned} \tag{A.13}$$

$$\max_{1\le t\le n} \left| \frac{\partial^2 h_t(\lambda)}{\partial\phi^2} \right| = O(1) \max_{1\le t\le n} \left| \sum_{i=1}^{t-1} \beta^{i-1} y_{t-i-1}^2 \right| = O_p(n), \tag{A.14}$$

where $O_p(\cdot)$ holds uniformly in $\lambda \in V_{1n}$.

$$\begin{aligned} \frac{\partial^2 l_t(\lambda)}{\partial\phi^2} &= -\frac{y_{t-1}^2}{h_t(\lambda)} - \frac{\varepsilon_t^2(\lambda)}{2h_t^3(\lambda)} \left[\frac{\partial h_t(\lambda)}{\partial\phi} \right]^2 \\ &\quad + \left[\frac{\varepsilon_t^2(\lambda)}{h_t(\lambda)} - 1 \right] \frac{\partial}{\partial\phi} \left[\frac{1}{2h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial\phi} \right] - 2\, \frac{\varepsilon_t(\lambda) y_{t-1}}{h_t^2(\lambda)} \frac{\partial h_t(\lambda)}{\partial\phi}. \end{aligned}$$

Note that $h_t(\lambda) \ge \omega_0/2$ as $\lambda \in V_{1n}$ and n is large enough. By equations (A.10)–(A.14),

$$\begin{aligned} \max_{1\le t\le n} \sup_{V_{1n}} \left| \frac{\partial^2 l_t(\lambda)}{\partial\phi^2} \right| &\le O_p(n) + O_p(n) \max_{1\le t\le n} \sup_{V_{1n}} \frac{|\varepsilon_t(\lambda)|^2}{h_t(\lambda)} \\ &= O_p(n) + O_p(n) \max_{1\le t\le n} \sup_{V_{1n}} \frac{|\varepsilon_t|^2 + o_p(1)}{h_t + o_p(1)h_t} \\ &= O_p(n) + O_p(n) \max_{1\le t\le n} \eta_t^2 = o_p(n^{2-2\tilde{\iota}}), \end{aligned} \tag{A.15}$$

as $\tilde{\iota}$ is small enough. Using the identity $x/(a+x) \le x^{\iota_0/2}$ for any $\iota_0 > 0$ as $a, x > 0$, we have

$$\frac{1}{h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \alpha} = \frac{\sum_{i=1}^{t-1} \beta^i \varepsilon_{t-i}^2(\lambda)}{\alpha \left[c + \sum_{i=1}^{t-1} \beta^i \varepsilon_{t-i}^2(\lambda) \right]} = O_p(1) \sum_{i=1}^{t-1} \rho^i |\varepsilon_{t-i}(\lambda)|^{\iota_0},$$

for any $\iota_0 > 0$ and a constant $0 < \rho < 1$, where $c = \omega(\sum_{i=0}^{t-1} \beta^i)/\alpha$. Similarly, it can be easily seen that

$$\frac{1}{h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \beta} = \frac{\sum_{i=1}^{t-1} i\beta^i \varepsilon_{t-i}^2(\lambda)}{c + \sum_{i=1}^{t-1} \beta^i \varepsilon_{t-i}^2(\lambda)} = O_p(1) \sum_{i=1}^{t-1} \rho^i |\varepsilon_{t-i}(\lambda)|^{\iota_0},$$

$$\frac{\partial^2 l_t(\lambda)}{\partial \delta \partial \delta'} = -\frac{\varepsilon_t^2(\lambda)}{2h_t^3(\lambda)} \frac{\partial h_t(\lambda)}{\partial \delta} \frac{\partial h_t(\lambda)}{\partial \delta'} - \frac{1}{2h_t(\lambda)} \left[\frac{\partial^2 h_t(\lambda)}{\partial \delta \partial \delta'} - \frac{1}{h_t(\lambda)} \frac{\partial h_t(\lambda)}{\partial \delta} \frac{\partial h_t(\lambda)}{\partial \delta'} \right] \left[1 - \frac{\varepsilon_t^2(\lambda)}{h_t(\lambda)} \right].$$

In view of equations (A.10)–(A.12) and the previous three equations, it follows that

$$\begin{aligned} \max_{1\le t\le n} \sup_{V_{1n}} \left\| \frac{\partial^2 l_t(\lambda)}{\partial \delta \partial \delta'} \right\| &\le O_p(1) \max_{1\le t\le n} \sup_{V_{1n}} \frac{|\varepsilon_t(\lambda)|^{2+2\iota_0}}{h_t(\lambda)} \\ &\le O_p(1) \max_{1\le t\le n} \sup_{V_{1n}} \frac{|\varepsilon_t|^{2+2\iota_0} + o_p(1)}{h_t + o_p(1) h_t} \\ &\le o_p(1) + O_p(1) \max_{1\le t\le n} |\eta_t|^{2+2\iota_0} h_t^{\iota_0} = o_p(n^{1-2\tilde{\iota}}), \end{aligned} \quad \textbf{(A.16)}$$

for some small $\tilde{\iota}$. Similarly, it follows that

$$\max_{1\le t\le n} \sup_{V_{1n}} \left\| \frac{\partial^2 l_t(\lambda)}{\partial \delta \partial \phi} \right\| = o_p(n^{3/2-2\tilde{\iota}}).$$

Using this equation with (A.15) and (A.16), part (a) follows.

For part (b), because $\varepsilon_t^2(\lambda_0)/h_t(\lambda_0) = \eta_t^2$ and $\max_{1\le t\le n}|y_t|/\sqrt{n} = O_p(1)$, by (A.12), it follows that

$$\max_{1\le t\le n} \left| \frac{\partial l_t(\lambda_0)}{\partial \phi} \right| \le O(1) \max_{1\le t\le n} |y_t| \left(\max_{1\le t\le n} \eta_t^2 + 1 \right) = o_p(n^{1-\tilde{\iota}}). \quad \textbf{(A.17)}$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

Furthermore, it is straightforward to show that

$$\max_{1\le t\le n}\left\|\frac{\partial l_t(\lambda_0)}{\partial\delta}\right\| \le O(1)\max_{1\le t\le n}|\varepsilon_t|^{\iota_0}\max_{1\le t\le n}(\eta_t^2+1) = o_p(n^{0.5-\tilde{\iota}}).$$

By means of (A.17) and this fact, we have

$$\max_{1\le t\le n}\|N_n D_t(\lambda_0)\| = o_p(n^{0.5-\tilde{\iota}}),$$

for a sufficiently small $\tilde{\iota} > 0$. Using Taylor's expansion and part (a) of this lemma, part (b) is established.

For part (c), using (A.10)–(A.12) and the same arguments used in Theorem C of Ling and Li (2003), it can be shown that

$$\sup_{V_{1n}} n^{-1}\left\|\sum_{t=1}^n N_n[P_t(\lambda)-P_t(\lambda_0)]N_n\right\| = o_p(1).$$

Part (c) follows. Finally, using parts (a)–(c) of this lemma and Lemma A.3(b), we can show that (d) holds. This completes the proof. ■

Proof of Theorem 3.2. Using the idea in Chuang and Chan (2002), let $\tilde{b} = bN_n^{-1}$, $\tilde{D}_t(\lambda) = N_n D_t(\lambda)$, and $\tilde{P}_t(\lambda) = N_n P_t(\lambda) N_n$. Furthermore, let $\tilde{b} = \rho\theta$ with $\|\theta\| = 1$. Denote

$$Q_{1n}(\lambda,\tilde{b}) = \frac{1}{n}\sum_{t=1}^n \frac{\tilde{D}_t(\lambda)}{1+\tilde{b}'\tilde{D}_t(\lambda)}. \qquad \textbf{(A.18)}$$

Similar to proving equation (A.2), it can be shown that

$$0 = \|Q_{1n}(\lambda,\rho\theta)\| \ge -\frac{1}{n}\left|\sum_{t=1}^n \theta'\tilde{D}_t(\lambda)\right| + \frac{\rho\theta'\tilde{S}_n(\lambda)\theta}{1+\rho\tilde{Z}_n(\lambda)}, \qquad \textbf{(A.19)}$$

where $\tilde{S}_n(\lambda) = n^{-1}\sum_{t=1}^n \tilde{D}_t(\lambda)\tilde{D}_t'(\lambda)$ and $\tilde{Z}_n(\lambda) = \max_{1\le t\le n}\|\tilde{D}_t(\lambda)\|$. By Lemmas A.3(a) and (b) and Lemma A.4(c), we have

$$\begin{aligned} n^{-1}\left|\sum_{t=1}^n \theta'\tilde{D}_t(\lambda)\right| &\le n^{-1}\left|\sum_{t=1}^n \theta'\tilde{D}_t(\lambda_0)\right| + n^{-1}\sup_{V_{1n}}\sum_{t=1}^n \|\tilde{P}_t(\lambda)\|\,\|N_n^{-1}(\lambda-\lambda_0)\| \\ &= O_p(n^{-0.5}), \end{aligned} \qquad \textbf{(A.20)}$$

uniformly in $\lambda \in V_{1n}$ defined in Lemma A.4. By Lemmas A.3(c) and A.4(d), $\theta'\tilde{S}_n(\lambda)\theta \ge a + o_p(1)$ for a positive random variable a, as $\lambda \in V_{1n}$. Thus, using equations (A.19) and (A.20), we have

$$\frac{\rho\theta'\tilde{S}_n(\lambda)\theta}{1+\rho\tilde{Z}_n(\lambda)} = O_p(n^{-0.5}).$$

Furthermore, by Lemma A.4(b), it is seen that

$$\rho = \|\tilde{b}\| = O_p(n^{-0.5}),$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

uniformly in $\lambda \in V_{1n}$. Let $\tilde{\gamma}_t = \tilde{b}'\tilde{D}_t(\lambda)$, where $\tilde{b}$ is the solution of $Q_{1n}(\lambda, \tilde{b}) = 0$. Then

$$\max_{1\le t\le n} |\tilde{\gamma}_t| = n^{-0.5} \max_{1\le t\le n} \|\tilde{D}_t(\lambda)\| = o_p(1),$$

uniformly in $\lambda \in V_{1n}$.

Let $Q_{2n}(\lambda, \tilde{b}) = n^{-1}\sum_{t=1}^{n} \tilde{P}_t(\lambda)\tilde{b}/[1 + \tilde{b}'\tilde{D}_t(\lambda)]$. Following the same procedure as in the proof of Theorem 2.1,

$$\frac{\partial Q_{2n}(\lambda, \tilde{b})}{\partial \lambda_i} = o_p(1) \quad \text{and} \quad \frac{\partial Q_{2n}(\lambda, \tilde{b})}{\partial \tilde{b}'} = \frac{1}{n}\sum_{t=1}^{n} \tilde{P}_t(\lambda_0) + o_p(1)$$

and

$$\frac{\partial Q_{1n}(\lambda, \tilde{b})}{\partial \lambda'} = \frac{1}{n}\sum_{t=1}^{n} \tilde{P}_t(\lambda_0) + o_p(1) \quad \text{and} \quad \frac{\partial Q_{1n}(\lambda, \tilde{b})}{\partial \tilde{b}'} = -\tilde{S}_n(\lambda_0) + o_p(1),$$

uniformly in $\lambda \in V_{1n}$.

In view of this fact, Lemma A.3(c), and Lemma A.4(d), it follows that

$$0 = Q_{1n}(\lambda, \tilde{b}) = \frac{1}{n}\sum_{t=1}^{n} \tilde{D}_t(\lambda) - \tilde{S}_n(\lambda)\tilde{b} + o_p(n^{-0.5}),$$

uniformly in $\lambda \in V_{1n}$. By Lemma A.3(a) and (b) and Lemma A.4(c), we have

$$\tilde{b}_0 = [n\tilde{S}_n(\lambda_0)]^{-1}\sum_{t=1}^{n} \tilde{D}_t(\lambda_0) + o_p(n^{-0.5}).$$

Because $\hat{\lambda}_n$ is the QMLE of λ_0, by Theorem 3.2 in Ling, Li, and McAleer (2003), we have

$$\sqrt{n}(\hat{\lambda}_n - \lambda_0) = -\left[\frac{1}{n}\sum_{t=1}^{n} \tilde{P}_t(\lambda_0)\right]^{-1} Q_{1n}(\lambda_0,0) + o_p(1).$$

Using a Taylor's expansion at $(\hat{\lambda}_n, \hat{b}_n)$, the previous two equations, and Lemma 6.3 with the same method as for Theorem 2.1, we can show that

$$\begin{aligned} W_E(\hat{\lambda}_n) &= -2[L_E(\hat{\lambda}_n) - L_E(\lambda_0)] \\ &= nQ'_{1n}(\lambda_0,0)\tilde{S}_n^{-1}(\lambda_0)Q_{1n}(\lambda_0,0) + o_p(1) \\ &\to_{\mathcal{L}} \left[K\int_0^1 \omega_1^2(\tau)\,d\tau\right]^{-1}\left[\int_0^1 \omega_1(\tau)\,d\omega_2(\tau)\right]^2 + \chi_3^2, \end{aligned} \quad \textbf{(A.21)}$$

as $n \to \infty$, where the last two steps follow from Lemma A.3. This completes the proof. ■

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

Proof of Lemma 3.1. Using the idea in Chuang and Chan (2002), let $\tilde{b} = bN_n^{-1}$ and $\tilde{D}_t(\delta) = N_n D_t(\lambda)|_{\lambda=(1,\delta)}$. Furthermore, let $\tilde{b} = \rho\theta$ with $\|\theta\| = 1$. Denote

$$Q_{1n}(\delta,\tilde{b}) = \frac{1}{n}\sum_{t=1}^{n}\frac{\tilde{D}_t(\delta)}{1+\tilde{b}'\tilde{D}_t(\delta)}. \tag{A.22}$$

Similar to proving equation (A.2), it can be shown that

$$0 = \|Q_{1n}(\delta,\rho\theta)\| \geq -\frac{1}{n}\left|\sum_{t=1}^{n}\theta'\tilde{D}_t(\delta)\right| + \frac{\rho\theta'\tilde{S}_n(\delta)\theta}{1+\rho\tilde{Z}_n(\delta)}, \tag{A.23}$$

where $\tilde{S}_n(\delta) = n^{-1}\sum_{t=1}^{n}\tilde{D}_t(\delta)\tilde{D}_t'(\delta)$ and $\tilde{Z}_n(\delta) = \max_{1\leq t\leq n}\|\tilde{D}_t(\delta)\|$. Let $\tilde{D}_t(\delta) = [\tilde{D}_{1t}(\delta), \tilde{D}_{2t}'(\delta)]'$, where $\tilde{D}_{1t}(\delta)$ is the first element of $\tilde{D}_t(\delta)$. Denote $\tilde{P}_t(\delta) = [\partial\tilde{D}_{1t}(\delta)/\partial\delta, \partial\tilde{D}_{2t}'(\delta)/\partial\delta]'$. By Lemma A.3(a) and (b) and Lemma A.4(c), we have

$$\begin{aligned} n^{-1}\left|\sum_{t=1}^{n}\theta'\tilde{D}_t(\delta)\right| &\leq n^{-1}\left|\sum_{t=1}^{n}\theta'\tilde{D}_t(\delta_0)\right| + n^{-1}\sup_{V_{1n}}\sum_{t=1}^{n}\left\|\frac{1}{\sqrt{n}}\tilde{P}_t(\delta)\right\|\|\sqrt{n}(\delta-\delta_0)\| \\ &= O_p(n^{-0.5+\tilde{\iota}}), \end{aligned} \tag{A.24}$$

uniformly in $\lambda \in V_{1n}$. By Lemmas A.3(c) and A.4(d), $\theta'\tilde{S}_n(\delta)\theta \geq a + o_p(1)$ for a positive random variable a, uniformly in $\lambda \in V_{1n}$. Thus, using equations (A.23) and (A.24), we have

$$\frac{\rho\theta'\tilde{S}_n(\delta)\theta}{1+\rho\tilde{Z}_n(\delta)} = O_p(n^{-0.5+\tilde{\iota}}).$$

Furthermore, by Lemma A.4(b), it is seen that

$$\rho = \|\tilde{b}\| = O_p(n^{-0.5+\tilde{\iota}}),$$

uniformly in $\delta \in V_{1n}$. Let $\tilde{\gamma}_t = \tilde{b}'\tilde{D}_t(\delta)$, where $\tilde{b}$ is the solution of $Q_{1n}(\delta,\tilde{b}) = 0$. Then

$$\max_{1\leq t\leq n}|\tilde{\gamma}_t| = n^{-0.5+\tilde{\iota}}\max_{1\leq t\leq n}\|\tilde{D}_t(\delta)\| = o_p(1),$$

uniformly in $\delta \in V_{1n}$.

$$\begin{aligned} 0 = Q_{1n}(\delta,\tilde{b}) &= \frac{1}{n}\sum_{t=1}^{n}\tilde{D}_t(\delta) - \tilde{S}_n(\delta)\tilde{b} + \frac{1}{n}\sum_{t=1}^{n}\frac{\tilde{D}_t(\delta)\tilde{\gamma}_t^2}{1+\tilde{\gamma}_t} \\ &\leq \frac{1}{n}\sum_{t=1}^{n}\tilde{D}_t(\delta) - \tilde{S}_n(\delta)\tilde{b} + \frac{\|\tilde{b}\|\max_{1\leq t\leq n}|\tilde{\gamma}_t|}{1-\max_{1\leq t\leq n}|\tilde{\gamma}_t|}\frac{1}{n}\sum_{t=1}^{n}\|\tilde{D}_t(\delta)\|^2 \\ &= \frac{1}{n}\sum_{t=1}^{n}\tilde{D}_t(\delta) - \tilde{S}_n(\delta)\tilde{b} + o_p(n^{-0.5+\tilde{\iota}}), \end{aligned}$$

uniformly in $\delta \in V_{1n}$. By Lemma A.3(a), $\sum_{t=1}^{n} \tilde{D}_t(\delta_0) = O_p(n^{0.5})$. Furthermore, by Lemma A.4(c),

$$\begin{aligned}\tilde{b} &= [n\tilde{S}_n(\delta)]^{-1} \sum_{t=1}^{n} \tilde{D}_t(\delta) + o_p(n^{-0.5+\tilde{\iota}}) \\ &= [n\tilde{S}_n(\delta_0)]^{-1} \sum_{t=1}^{n} \tilde{D}_t(\delta_0) + [n\tilde{S}_n(\delta_0)]^{-1} \sum_{t=1}^{n} \tilde{P}_t(\delta^*)(\delta - \delta_0) + o_p(n^{-0.5+\tilde{\iota}}) \\ &= o_p(n^{-0.5+\tilde{\iota}}),\end{aligned} \quad \textbf{(A.25)}$$

uniformly in $\delta \in V_{1n}$. By Taylor's expansion,

$$L_E(\delta) = \sum_{t=1}^{n} \log(1+\tilde{\gamma}_t) = \sum_{t=1}^{n} \tilde{\gamma}_t - \frac{1}{2}\sum_{t=1}^{n} \tilde{\gamma}_t^2 + \sum_{t=1}^{n} \tilde{v}_t, \quad \textbf{(A.26)}$$

where, for some finite $B > 0$,

$$P(|\tilde{v}_t| \le B|\tilde{\gamma}_t|^3, 1 \le t \le n) \to 1, \quad \textbf{(A.27)}$$

as $n \to \infty$. Thus, by Lemmas A.3(c) and A.4(d), it follows that

$$|\sum_{t=1}^{n} \tilde{v}_t| \le Bn^{-1+2\tilde{\iota}} \max_{1\le t\le n} |\tilde{\gamma}_t| \sum_{t=1}^{n} \|\tilde{D}_t(\delta)\|^2 = o_p(n^{2\tilde{\iota}}), \quad \textbf{(A.28)}$$

uniformly in $\delta \in V_{1n}$. Denote $\delta = \delta_0 + un^{-0.5+\tilde{\iota}}$ for $\delta \in \{\delta : \|\delta - \delta_0\| = n^{-0.5+\tilde{\iota}}\}$, where $\|u\| = 1$. By means of equations (A.25)–(A.28) and Lemmas A.3(b) and (c) and A.4(c) and (d), we have

$$\begin{aligned}L_E(\delta) &= \sum_{t=1}^{n} \log(1+\tilde{\gamma}_t) \\ &= \frac{n}{2}\left[\frac{1}{n}\sum_{t=1}^{n} \tilde{D}_t(\delta)\right]' \tilde{S}_n^{-1}(\delta) \left[\frac{1}{n}\sum_{t=1}^{n} \tilde{D}_t(\delta)\right] + o_p(n^{2\tilde{\iota}}) \\ &= \frac{n}{2}\left[\frac{1}{n}\sum_{t=1}^{n} \tilde{D}_t(\delta_0) + \frac{1}{n}\sum_{t=1}^{n} \tilde{P}_t(\delta_0)un^{-0.5+\tilde{\iota}} + o_p(n^{-0.5+\tilde{\iota}})\right]' \tilde{S}_n^{-1}(\delta_0) \\ &\quad \times \left[\frac{1}{n}\sum_{t=1}^{n} \tilde{D}_t(\delta_0) + \frac{1}{n}\sum_{t=1}^{n} \tilde{P}_t(\delta_0)un^{-0.5+\tilde{\iota}} + o_p(n^{-0.5+\tilde{\iota}})\right] + o_p(n^{2\tilde{\iota}}) \\ &= \frac{n}{2}[\Omega un^{-0.5+\tilde{\iota}} + o_p(n^{-0.5+\tilde{\iota}})]' \left(\frac{\kappa}{2}\Omega\right)^{-1} [\Omega un^{-0.5+\tilde{\iota}} + o_p(n^{-0.5+\tilde{\iota}})] + o_p(n^{2\tilde{\iota}}) \\ &\ge (c-\epsilon)n^{2\tilde{\iota}},\end{aligned} \quad \textbf{(A.29)}$$

uniformly in u, which happens with probability at least $1 - \epsilon$ for any given $\epsilon > 0$, where $c = \delta_{\min}/\kappa$ and $\delta_{\min}$ is the smallest eigenvalue of Ω. Let $\tilde{b}_0$ be the solution of $Q_{1n}(\delta_0, \tilde{b}) = 0$. By (A.23) and Lemma A.4(b) and (d), it is not difficult to see that

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

$$\tilde{b}_0 = [n\tilde{S}_n(\delta_0)]^{-1} \sum_{t=1}^{n} \tilde{D}_t(\delta_0) + o_p(n^{-0.5}). \quad \textbf{(A.30)}$$

Similar to (A.29), we can show that

$$L_E(\delta_0) = \frac{n}{2}\left[\frac{1}{n}\sum_{t=1}^{n} \tilde{D}_t(\delta_0)\right]' \tilde{S}_n^{-1}(\delta_0)\left[\frac{1}{n}\sum_{t=1}^{n} \tilde{D}_t(\delta_0)\right]' + o_p(1) = O_p(\log\log n). \quad \textbf{(A.31)}$$

Because $L_E(\delta)$ is continuous in δ, by (A.29) and (A.31), $L_E(\delta)$ achieves its minimum value in the interior of V_n so that the minimizer $\hat{\delta}_n$ satisfies $\partial L_E(\delta)/\partial\delta = 0$. Finally, because $\lim_{n\to\infty} P(L_E(\hat{\delta}_n) = \min_{\delta\in V_n} L_E(\delta)) = 1$, the proof is complete. ■

Proof of Corollary 3.1. Let $Q_{2n}(\delta,\tilde{b}) = n^{-1}\sum_{t=1}^{n} \tilde{P}_t(\delta)\tilde{b}/[1 + \tilde{b}'\tilde{D}_t(\delta)]$ and $\hat{b}_n = \tilde{b}(\hat{\delta}_n)$, where $\tilde{b}$ is defined as in (A.22). Then $\partial L_E(\hat{\delta}_n)/\partial\delta = 0$ if and only if

$$Q_{1n}(\hat{\delta}_n,\hat{b}_n) = 0 \quad \text{and} \quad Q_{2n}(\hat{\delta}_n,\hat{b}_n) = 0,$$

where $Q_{1n}(\delta,\tilde{b})$ is defined as in (A.22) and $\tilde{P}_t(\delta)$ is defined as in (A.24) .

Following the same method as for Theorem 2.1 and using Lemmas A.3 and A.4, we can show that

$$\frac{\partial Q_{1n}(\delta,\tilde{b})}{\partial\delta'} = \left[O, \frac{1}{n}\sum_{t=1}^{n}\frac{\partial\tilde{D}'_{2t}(\delta_0)}{\partial\delta}\right]' + o_p(1) = [O,\Omega]' + o_p(1),$$

$$\frac{\partial Q_{1n}(\delta,\tilde{b})}{\partial\tilde{b}'} = -\tilde{S}_n(\delta_0) + o_p(1) = -diag\left\{\tilde{S}_{\phi n}(\delta_0), \frac{\kappa}{2}\Omega\right\} + o_p(1),$$

$$\frac{\partial Q_{2n}(\delta,\tilde{b})}{\partial\delta_i} = o_p(1),$$

$$\frac{\partial Q_{2n}(\delta,\tilde{b})}{\partial\tilde{b}'} = \left[O, \frac{1}{n}\sum_{t=1}^{n}\frac{\partial\tilde{D}'_{2t}(\delta_0)}{\partial\delta}\right] + o_p(1) = [O,\Omega] + o_p(1),$$

uniformly in $\delta \in V_n$, where $O = (0,0,0)'$ and $\tilde{S}_{\phi n}(\delta_0)$ is the (1,1)th element of $\tilde{S}_n(\delta_0)$ defined following (A.23). Using Taylor's expansion, we have

$$\begin{aligned} 0 &= Q_{1n}(\hat{\delta}_n,\hat{b}_n) \\ &= Q_{1n}(\delta_0,0) + [O,\Omega]'(\hat{\delta}_n - \delta_0) + \left[-diag\left\{\tilde{S}_{\phi n}(\delta_0), \frac{\kappa}{2}\Omega\right\} + o_p(1)\right](\hat{b}_n - 0), \end{aligned} \quad \textbf{(A.32)}$$

$$0 = Q_{2n}(\hat{\delta}_n,\hat{b}_n) = Q_{2n}(\delta_0,0) + o_p(1)(\hat{\delta}_n - \delta_0) + [O,\Omega](\hat{b}_n - 0). \quad \textbf{(A.33)}$$

https://doi.org/10.1017/S0266466606060208 Published online by Cambridge University Press

Let $\hat{b}_n = (\hat{b}_{1n}, \hat{b}'_{2n})'$ and $Q_{1n}(\delta_0,0) = (Q_{11n}, Q'_{12n})'$, where $\hat{b}_{1n}$ and Q_{11n} are the first element of $\hat{b}_n$ and $Q_{1n}(\delta_0,0)$, respectively. By (A.32) and (A.33), it follows that

$$\hat{b}_{1n} = \tilde{S}_{\phi n}^{-1}(\delta_0) Q_{11n} + o_p\left(\frac{1}{\sqrt{n}}\right),$$

$$\hat{b}_{2n} = o_p\left(\frac{1}{\sqrt{n}}\right),$$

$$\hat{\delta}_n - \delta_0 = \Omega^{-1} Q_{12n} + o_p\left(\frac{1}{\sqrt{n}}\right).$$

Furthermore, by (A.30) and Lemma A.3, we have

$$\hat{b}_n - \tilde{b}_0 = [0, -(\Omega^{-1} Q_{12n})']' + o_p(n^{-0.5}).$$

Using a Taylor's expansion at $(\hat{\delta}_n, \hat{b}_n)$ and the previous equations, we can show that

$$\begin{aligned} -2[L_E(\hat{\delta}_n) - L_E(\delta_0)] &= n[(\hat{b}_n - \tilde{b}_0)', (\hat{\delta}_n - \delta_0)'] \\ &\quad \times \left\{ \begin{bmatrix} -\tilde{S}_{\phi n}(\delta_0) & O & O \\ O & -\kappa\Omega/2 & \Omega \\ O & \Omega & O \end{bmatrix} + o_p(1) \right\} \\ &\quad \times [(\hat{b}_n - \tilde{b}_0)', (\hat{\delta}_n - \delta_0)']' \\ &= n Q'_{12n} \left(\frac{\kappa}{2}\Omega\right)^{-1} Q_{12n} + o_p(1). \end{aligned} \qquad \textbf{(A.34)}$$

By the expansion in (A.21) and (A.34), it follows that

$$\begin{aligned} W_E(\hat{\delta}_n) &= -2[L_E(\hat{\lambda}_n) - L_E(\delta_n)] \\ &= -2[L_E(\hat{\lambda}_n) - L_E(\lambda_0)] - 2[L_E(\hat{\delta}_n) - L_E(\lambda_0)] \\ &= n Q'_{11n}(\lambda_0,0) \tilde{S}_{1\phi n}^{-1}(\lambda_0) Q_{11n}(\lambda_0,0) + o_p(1) \\ &\to_{\mathcal{L}} \left[K \int_0^1 \omega_1^2(\tau)\, d\tau \right]^{-1} \left[\int_0^1 \omega_1(\tau)\, d\omega_2(\tau) \right]^2 \quad \text{as } n \to \infty. \end{aligned} \qquad \textbf{(A.35)}$$

Let $B_1(\tau) = \omega_1(\tau)/\sigma$ and $B_2(\tau) = -\sigma^{-1}(\sigma^2 K - 1)^{-1/2}\omega_1(\tau) + \sigma(\sigma^2 K - 1)^{-1/2}$ $\omega_2(\tau)$, where $\sigma^2 = Eh_t$. Then $B_1(\tau)$ and $B_2(\tau)$ are two independent standard Brownian motions. Denote the limiting distribution in (A.35) by ζ^2. As shown in Ling and Li (1998), we obtain that

$$\zeta = \frac{\int_0^1 B_1(\tau)\, dB_1(\tau)}{\sigma^2 F \int_0^1 B_1^2(\tau)\, d\tau} + (\sigma^2 K - 1)^{1/2} \frac{\int_0^1 B_1(\tau)\, dB_2(\tau)}{\sigma^2 F \int_0^1 B_1^2(\tau)\, d\tau}.$$

The second term of this equation can be simplified as $[(\sigma^2 K - 1)^{1/2}/(F\sigma^2)]\ (\int_0^1 B_1^2(\tau)\, d\tau)^{-1/2}\xi$, where ξ is a standard normal random variable that is independent of $\int_0^1 B_1^2(\tau)\, d\tau$ (see Phillips, 1989). Then

$$\zeta = \frac{1}{c}\left[\frac{\rho \int_0^1 B_1(\tau)\, dB_1(\tau)}{\int_0^1 B_1^2(\tau)\, d\tau} + (1-\rho^2)^{1/2} \left(\int_0^1 B_1^2(\tau)\, d\tau \right)^{-1/2} \xi \right],$$

where $c = \sigma F / \sqrt{K}$. Furthermore, by (A.35), we have

$$\widetilde{W}_n = \to_{\mathcal{L}} \frac{F}{K} \left(F \int_0^1 \omega_1^2(\tau)\, d\tau \right) \zeta^2 = \left(\int_0^1 B_1^2(\tau)\, d\tau \right) (c\zeta)^2.$$

This completes the proof of Corollary 3.1. ■