Econometric Theory, **42**, 2026, 996–1033.
doi:10.1017/S026646662510011X

DETECTING CHANGES IN GARCH(1,1) PROCESSES WITHOUT ASSUMING STATIONARITY

LAJOS HORVÁTH
University of Utah

SHIXUAN WANG
University of Reading

This article develops a new test to detect changes in generalized autoregressive conditionally heteroscedastic (GARCH(1,1)) processes without imposing a stationary assumption. Specifically, the procedure tests the null hypothesis of a GARCH process with constant parameters, either in (strictly) stationary or explosive regimes, against the alternative hypothesis of parameter changes. We derive the limiting distribution of the test statistics and establish their asymptotic consistency. Monte Carlo simulations show that the proposed test has good size control and high power. We demonstrate a prototype application on a small group of stocks and report a further extensive application to more than ten thousand U.S. stocks.

1. INTRODUCTION

Modeling volatility in financial time series plays an important role in risk management. A key milestone in volatility modeling is the autoregressive conditionally heteroscedastic (ARCH) process introduced by Engle (1982). Following this seminal work, many extensions of the ARCH process have been developed. The most widely used is the generalized ARCH (GARCH) process proposed by Bollerslev (1986). In the last few decades, the GARCH process has received substantial theoretical attention. More importantly, it has become the "workhorse" for volatility modeling in the financial industry (Lee and Hansen, 1994). We refer to Francq and Zakoian (2019) for a comprehensive review of the GARCH model.

In practical applications, financial returns are commonly assumed to be stationary. This assumption has led to a rich body of literature on stationary GARCH, overshadowing the comparatively limited research on nonstationary (explosive)

We would like to thank the editor Peter C. B. Phillips, the co-editor Robert Taylor, and three anonymous referees for their constructive suggestions, which have lead to a much improved paper. We also thank the helpful comments received from the conference participants of the 1st conference on "Econometrics with Data Science" at Reading 2024. All remaining errors are ours. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We report that there are no competing interests to declare. Address correspondence to Shixuan Wang, Department of Economics, University of Reading, Reading, UK; e-mail: shixuan.wang@reading.ac.uk

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

GARCH. We conduct an empirical investigation using the nonstationarity test of Francq and Zakoïan (2012) for the universe of U.S. stocks between 2011 and 2022. At the 1% significance level, the test reveals that the null hypothesis of nonstationarity (explosivity) cannot be rejected for 631 stocks, which constitute 5.94% of the total stocks tested. Further details can be found in Section 4. This empirical evidence shows that nonstationary (explosive) GARCH processes are more prevalent than is typically assumed.

The pioneering work in the limited literature on the explosive and nonstationary behavior of the GARCH(1,1) process is by Jensen and Rahbek (2004). By imposing the constraint that the intercept parameter in the GARCH(1,1) model is fixed, they establish consistency and asymptotic normality for a constrained version of the quasi-maximum likelihood estimator (QMLE) without assuming stationarity. In groundbreaking work of Francq and Zakoïan (2012), they study the asymptotic properties of unconstrained QMLE for GARCH(1,1) models without a strict stationarity assumption. They discover that, except for the intercept, the QMLE remains consistent and asymptotically normal in both stationary and nonstationary (explosive) cases. They further develop econometric tools for testing strict stationarity and nonstationarity in GARCH models. In a follow-up study, Francq and Zakoïan (2013) establish the asymptotic normality of the unconstrained QMLE of the parameters, including the power parameter but excluding the intercept, for asymmetric power-transformed GARCH(1,1) models when strict stationarity does not hold.

The stability of parameter values in an ARCH or GARCH process is crucial for estimating the model parameters. If a researcher fails to account for potential changes in its parameter values, the inference of model parameters will not be valid, and the conditional volatility implied by the estimated model could be misleading, resulting in under- or over-estimation of the underlying risk. Kokoszka and Leipus (1999, 2000) make early contributions to change-point analysis for ARCH processes. They introduce test procedures for ARCH models with at least a finite second moment. Expanding on this, Andreou and Ghysels (2002) develop tests for multiple breaks that can be applied to ARCH processes.

An early contribution to the detection of parameter changes in GARCH models is attributed to Chu (1995), who employs the maximized Lagrange multiplier (LM) test. Using the approximate likelihood scores, Berkes, Horváth, and Kokoszka (2004) introduce a test for parameter constancy in GARCH(p,q) models that does not require observations to have finite variance. From a different perspective, Galeano and Tsay (2010) propose an algorithm to detect multiple change points in individual parameters of a GARCH model. Chen and Hong (2016) propose a test that detects smooth changes in GARCH parameters by testing whether the global GARCH estimates deviate significantly from the local parameter estimates. Different from the retrospective detection framework, Berkes et al. (2004) develop a sequential monitoring scheme for the on-line detection of changes in GARCH(p,q) models in real time. Their procedure is based on quasi-likelihood scores and does not use model residuals.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Almost all the existing literature on detecting changes in GARCH models assumes that the parameters of GARCH models are in a (strictly) stationary regime under the null hypothesis. In contrast, we develop a test to detect changes in GARCH(1,1) processes without assuming stationarity. Namely, we test the null hypothesis of a GARCH process with constant parameters that are in either (strictly) stationary or explosive regimes against the alternative of the existence of parameter changes. Our test complements the existing econometric toolset for detecting changes in GARCH(1,1) processes that could be either (strictly) stationary or explosive under the null. From a practical perspective, our test procedure is the same for both stationary and explosive GARCH(1,1) processes that simplifies its application. We derive the limiting distribution of the test statistics and establish the asymptotic consistency of the test. Monte Carlo simulations show that the proposed test has good size control and high power. In line with Francq and Zakoïan (2012, 2013), a limitation of our study is that our test is applicable to GARCH(1,1) parameters except for the intercept.

Our study is closely related to that of Richter, Wang, and Wu (2023), but the aim and setup differ. They construct a test to identify whether there is a period during which the parameters of a GARCH model have changed. In their test framework, the null hypothesis assumes a GARCH model with constant parameters, whereas the alternative hypothesis has a period during which the parameters of the GARCH model have changed. Notably, their null hypothesis is in the regime where the GARCH process is strictly stationary. In contrast, our test allows the parameters of GARCH(1,1) processes to be in either stationary or explosive regimes under the null hypothesis. We formally prove the limiting distribution of our test statistics and establish the asymptotic consistency of the test for both stationary and explosive GARCH(1,1) processes. As a result, our study makes theoretical contributions that complement the existing research in this area.

Our study is also linked to the broader literature on change-point detection for nonparametric and (semi)parametric measures of volatility and market risks. Andreou and Ghysels (2006) study data-driven processes related to volatility and develop cumulative sum (CUSUM) change-point tests to monitor the stability of dynamic variance processes. Xu (2013) proposes modified CUSUM and LM tests for detecting structural changes in volatility, using a robust long-run variance estimator to achieve higher power through faster divergence under the alternative. Pape, Wied, and Galeano (2016) introduce a model-independent multivariate monitoring procedure for detecting changes in the vector of componentwise unconditional variances. Lazar, Wang, and Xue (2023) propose a test that applies the CUSUM procedure to the Wilcoxon statistic to detect change points in the value-at-risk and expected shortfall estimated by (semi)parametric models. Hoga and Demetrescu (2023) develop monitoring procedures for forecasts of value-at-risk and expected shortfall that are generated by nonparametric and (semi)parametric models.

The remainder of the article is organized as follows: In Section 2, we present the assumptions and formulate the main results of the asymptotic properties. Section 3

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

is dedicated to a simulation study examining the finite sample performance of the proposed test. In Section 4, we present an empirical demonstration of our testing procedure by applying it first to a small group of U.S. stocks and then to a dataset of more than ten thousand U.S. stocks. Section 5 concludes, and Section 6 is devoted to the proofs of all theoretical results.

2. ASSUMPTIONS AND MAIN RESULTS

The time-dependent GARCH (1,1) sequence is defined by the recursion

$$y_i = \sigma_i \epsilon_i \quad \text{and} \quad \sigma_i^2 = \omega_i + \alpha_i y_{i-1}^2 + \beta_i \sigma_{i-1}^2, \quad 1 \le i < \infty,$$

where σ_0^2, y_0^2 are initial values, and α_i, β_i, and ω_i are positive parameters. We observe $y_1, y_2, \ldots, y_N$. Under the null hypothesis, we have

$$H_0 : (\alpha_1, \beta_1, \omega_1) = (\alpha_2, \beta_2, \omega_2) = \cdots = (\alpha_N, \beta_N, \omega_N).$$

If the null hypothesis holds, the model has the form

$$y_i = \sigma_i \epsilon_i \quad \text{and} \quad \sigma_i^2 = \omega_0 + \alpha_0 y_{i-1}^2 + \beta_0 \sigma_{i-1}^2, \quad 1 \le i < \infty. \tag{2.1}$$

We use the notation $\boldsymbol{\theta}_0 = (\alpha_0, \beta_0, \omega_0)^\top$. It is common to assume that the following assumptions.

Assumption 2.1. $\{\epsilon_i, -\infty < i < \infty\}$ are independent and identically distributed random variables, ϵ_0^2 is nondegenerate $E\epsilon_0 = 0$, $E\epsilon_0^2 = 1$, $0 < \mathrm{var}(\epsilon_0^2)$, and $E|\epsilon_0|^\kappa < \infty$ with some $\kappa > 4$

and

Assumption 2.2. $\alpha_0 > 0$, $\beta_0 > 0$, and $\omega_0 > 0$.

Our method employs a quasi-likelihood function based on the normal density; the corresponding quasi-likelihood estimator is given by $\hat{\boldsymbol{\theta}}_N = (\hat{\alpha}_N, \hat{\beta}_N, \hat{\omega}_N)^\top$. If a different density is used in the definition of the quasi-likelihood, as in Berkes and Horváth (2004), our results remain true.

Because the conditional variances σ_{i-1}^2 are not observed, we use the recursion

$$\bar{\sigma}_i^2(\boldsymbol{\theta}) = \omega + \alpha y_{i-1}^2 + \beta \bar{\sigma}_{i-1}^2(\boldsymbol{\theta}), \quad i \ge 1, \tag{2.2}$$

where $\boldsymbol{\theta} = (\alpha, \beta, \omega)^\top$. The corresponding log quasi-likelihood function (ignoring a negative scaling factor) based on $\{y_i, 1 \le i \le k\}$ is

$$\bar{\mathscr{L}}_k(\boldsymbol{\theta}) = \sum_{i=1}^{k} \bar{\ell}_i(\boldsymbol{\theta}), \quad \text{and} \quad \bar{\ell}_i(\boldsymbol{\theta}) = \log \bar{\sigma}_i^2(\boldsymbol{\theta}) + \frac{y_i^2}{\bar{\sigma}_i^2(\boldsymbol{\theta})}.$$

To find $\hat{\boldsymbol{\theta}}_N$, we maximize $-\bar{\mathscr{L}}_k(\boldsymbol{\theta})$ on the set

$$\boldsymbol{\Theta} = \left\{\boldsymbol{\theta} : 1/\bar{\alpha} \le \alpha < \bar{\alpha}, 1/\bar{\beta} \le \beta \le \bar{\beta}, 1/\bar{\omega} \le \omega \le \bar{\omega}\right\},$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

where $\bar{\alpha}, \bar{\beta}$, and $\bar{\omega}$ are arbitrarily large. Thus, we can assume the following assumption.

Assumption 2.3. $\boldsymbol{\theta}_0$ is in the interior of $\boldsymbol{\Theta}$.

Since $y_1, y_2, \ldots, y_N$ are observed, the QMLE satisfies

$$\inf\left\{\bar{\mathscr{L}}_N(\boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta}\right\} = \bar{\mathscr{L}}_N(\hat{\boldsymbol{\theta}}_N),$$

where

$$\bar{\mathscr{L}}_N(\boldsymbol{\theta}) = \sum_{i=1}^{N} \bar{\ell}_i(\boldsymbol{\theta}), \quad \text{and} \quad \bar{\ell}_i(\boldsymbol{\theta}) = \log \bar{\sigma}_i^2(\boldsymbol{\theta}) + \frac{y_i^2}{\bar{\sigma}_i^2(\boldsymbol{\theta})}. \tag{2.3}$$

Nelson (1990) shows that y_i converges to a stationary sequence if and only if $E\log(\alpha_0\epsilon_0^2 + \beta_0) < 0$ (the stationary case). Lumsdaine (1996), Berkes, Horváth, and Kokoszka (2003), Francq and Zakoian (2004), and Straumann and Mikosch (2006) establish the asymptotic consistency and normality of $\hat{\boldsymbol{\theta}}_N$ under slightly weaker restrictions than Assumption 2.1. Nelson (1990) shows that if $E\log(\alpha_0\epsilon_0^2 + \beta_0) \geq 0$, then $|y_i| \to \infty$ in probability, as $i \to \infty$. Therefore, in this case, no stationary solution exists. Jensen and Rahbek (2004) prove that ω_0 cannot be estimated consistently in the explosive case ($E\log(\alpha_0\epsilon_0^2 + \beta_0) > 0$). Using the same arguments as Jensen and Rahbek (2004), our Theorem 6.1 shows that ω_0 cannot be estimated in the "boundary" case ($E\log(\alpha_0\epsilon_0^2 + \beta_0) = 0$) as well. Thus, without assuming stationarity we can only test the null hypothesis

$$H_0^* : (\alpha_1, \beta_1) = (\alpha_2, \beta_2) = \cdots = (\alpha_N, \beta_N).$$

The expected values of the derivatives of the log likelihood function with respect to α and β are zero under the null hypothesis, so we use the sequence of random functions

$$\bar{r}_k(\boldsymbol{\theta}) = \sum_{i=1}^{k} \left(\frac{\partial \bar{\ell}_i(\boldsymbol{\theta})}{\partial \alpha}, \frac{\partial \bar{\ell}_i(\boldsymbol{\theta})}{\partial \beta} \right)^\top$$

in the definition of our test statistics. We normalize the test statistics by

$$\hat{\mathbf{D}}_N = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha}, \frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} \right)^\top \left(\frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha}, \frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} \right).$$

If H_0 and Assumptions 2.1–2.3 hold, then

$$\|\hat{\mathbf{D}}_N - \mathbf{D}\| = o_P((\log\log N)^{-1/2}),$$

where $\mathbf{D}$ is a positive definite matrix. The form of $\mathbf{D}$ depends on whether $E\log(\alpha_0\epsilon_0^2 + \beta_0)$ is negative or nonnegative. Explicit formulas for $\mathbf{D}$ are given by Berkes et al. (2003) and Francq and Zakoian (2004) for the stationary case ($E\log(\alpha_0\epsilon_0^2 + \beta_0) < 0$) and by Jensen and Rahbek (2004) for the explosive case. We note that Jensen and Rahbek (2004) only consider the case when $E\log(\alpha_0\epsilon_0^2 +$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

$\beta_0) > 0$. As a result of Theorem 5.1, their formula also holds when $E\log(\alpha_0\epsilon_0^2 + \beta_0) = 0$. We use the same estimator, and no prior knowledge of stationarity is required.

Our proposed tests are based on the weighted functionals of the stochastic process

$$Z_N(t) = \frac{1}{N^{1/2}} \left\{ \bar{r}_{\lfloor (N+1)t \rfloor}(\hat{\boldsymbol{\theta}}_N) \hat{\mathbf{D}}_N^{-1} \left(\bar{r}_{\lfloor (N+1)t \rfloor}(\hat{\boldsymbol{\theta}}_N) \right)^\top \right\}^{1/2}, \qquad 0 < t < 1,$$

where $\lfloor \cdot \rfloor$ denotes the integer part of a number.

The definition of $Z_N(t)$ implies that $Z_N(t) = 0$ if $0 < t < 1/(N+1)$ and $N/(N+1) < t < 1$; that is, the behavior of $Z_N(t)$ is the same if t is close to 0 and 1. The estimation of the parameters of a GARCH(1,1) sequence requires large sample sizes. We only use the estimator $\hat{\boldsymbol{\theta}}_N$ that is based on the full sample. The partial derivatives in the definition of $\bar{r}_k(\boldsymbol{\theta})$ can be computed quickly using the recursion in (2.2). Another popular method to find changes in the parameters of a time series is based on splitting the data at time k. The estimators are computed from both the first k and the last $N-k$ observations. The corresponding statistic maximizes the difference between the estimators as a function of k. In the case of GARCH(1,1), this would require $2(N-3)$ estimates, and many of these estimators would be based on small subsamples. However, it is well known that estimations of GARCH parameters require large sample sizes (e.g., Huang, Wang, and Yao, 2008). Using the linearization of the quasi-likelihood estimators (cf. our Lemma 6.3 and Jensen and Rahbek, 2004), one can show that this method and our proposed method are asymptotically the same.

We use $w(t)$, a weight function that can be zero at 0 or 1, to get higher power if the change occurs at the end or at the beginning of the observation period. We assume that the weight function $w(t)$ satisfies the following assumption.

Assumption 2.4. (i) $\inf_{\delta \le t \le 1-\delta} w(t) > 0$ for all $0 < \delta < 1/2$, and (ii) $w(t)$ is non-decreasing in a neighborhood of 0 and non-increasing in a neighborhood of 1.

We employ the integral test of Chibisov (1964) and O'Reilly (1974), which is based on the functional

$$I(w,c) = \int_0^1 \frac{1}{s(1-s)} \exp\left(-c \frac{w^2(s)}{s(1-s)} \right) ds.$$

We refer to Csörgő and Horváth (1993) for the properties of $I(w,c)$ and its connection to the upper and lower classes of the Wiener process and the Brownian bridge.

THEOREM 2.1. *We assume that H_0 and Assumptions 2.1–2.4 are satisfied. (i) If $I(w,c) < \infty$ with some $c > 0$, then*

$$\sup_{0<t<1} \frac{Z_N(t)}{w(t)} \xrightarrow{\mathcal{D}} \sup_{0<t<1} \frac{1}{w(t)} \left(B_1^2(t) + B_2^2(t) \right)^{1/2},$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

where $\{B_1(t), 0 \le t \le 1\}$ *and* $\{B_2(t), 0 \le t \le 1\}$ *are independent Brownian bridges.*
(ii) In addition,

$$\lim_{N\to\infty} P\left\{ a(\log N) \max_{1\le k<N} \left\{ \frac{N}{k(N-k)} \bar{r}_k(\hat{\boldsymbol{\theta}}_N) \hat{\mathbf{D}}_N^{-1} \left(\bar{r}_k(\hat{\boldsymbol{\theta}}_N)\right)^\top \right\}^{1/2} \le x + b(\log N) \right\}$$
$$= \exp(-2e^{-x})$$

for all x, *where* $a(x) = (2\log x)^{1/2}$ *and* $b(x) = 2\log x + \log\log x$.

We note that the limit in Theorem 2.1(*i*) is finite with probability one if and only if $I(w,c) < \infty$ with some $c > 0$ (cf. Csörgő and Horváth, 1993). The normalization in Theorem 2.1(*ii*) is proportional to the standard deviation of the coordinates of $\bar{r}_k(\boldsymbol{\theta})$, and it is usually referred to as a maximally selected standardized statistic.

Now, we discuss briefly the consistency of tests based on Theorem 2.1. The parameter $\bar{\boldsymbol{\theta}}_0 = (\alpha_0, \beta_0)^\top$ changes to $\boldsymbol{\theta}_* = (\alpha_*, \beta_*)^\top$ at k^*, and we assume the following assumptions.

Assumption 2.5. $k^* = \lfloor \lambda N \rfloor$ with some $0 < \lambda < 1$,

and

Assumption 2.6. $\alpha_* > 0, \beta_* > 0$, $(\omega_0, \alpha_*, \beta_*)^\top$ is in the interior of $\boldsymbol{\Theta}$, and $\bar{\boldsymbol{\theta}}_0 \neq \boldsymbol{\theta}_*$.

The sequence $y_0, y_i = \sigma_i \epsilon_i, \sigma_i^2 = \omega_0 + \alpha_0 y_{i-1}^2 + \beta_0 \sigma_{i-1}^2, 1 \le i \le k^*$ changes to the new regime $y_i = \sigma_i \epsilon_i, \sigma_i^2 = \omega_0 + \alpha_* y_{i-1}^2 + \beta_* \sigma_{i-1}^2, k^* + 1 \le i \le N$ starting from the "initial value" y_{k^*}.

We use $\|\cdot\|$ for the Euclidean norm of vectors and matrices throughout this article.

THEOREM 2.2. *If Assumptions 2.1–2.3, 2.5, and 2.6 hold, then*

$$\liminf_{N\to\infty} P\left\{ \max\left(\left\| \frac{1}{N} \bar{r}_{k^*}(\hat{\boldsymbol{\theta}}_N) \right\|, \left\| \frac{1}{N} \left(\bar{r}_N(\hat{\boldsymbol{\theta}}_N) - \bar{r}_{k^*}(\hat{\boldsymbol{\theta}}_N) \right) \right\| \right) > \delta \right\} = 1$$

with some $\delta > 0$.

The estimator $\hat{\mathbf{D}}_N$ might not converge in probability, but it converges along subsequences. However, it still holds that

$$\liminf_{N\to\infty} \left\{ \max\left(\frac{1}{N} \left(\bar{r}_{k^*}^\top(\hat{\boldsymbol{\theta}}_N) \hat{\mathbf{D}}_N^{-1} \bar{r}_{k^*}(\hat{\boldsymbol{\theta}}_N) \right)^{1/2}, \right.\right.$$
$$\left.\left. \frac{1}{N} \left(\left(\bar{r}_N(\hat{\boldsymbol{\theta}}_N) - \bar{r}_{k^*}(\hat{\boldsymbol{\theta}}_N) \right)^\top \hat{\mathbf{D}}_N^{-1} \left(\bar{r}_N(\hat{\boldsymbol{\theta}}_N) - \bar{r}_{k^*}(\hat{\boldsymbol{\theta}}_N) \right) \right)^{1/2} \right) > \delta \right\} = 1$$

with some $\delta > 0$. Hence, Theorem 2.2 implies that under the alternative $\sup_{0<t<1} Z_N(t)/w(t) \to \infty$ in probability, the rate of convergence is $N^{1/2}$, and

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

the convergence to ∞ is not affected by the "inconsistency" of $\hat{\mathbf{D}}_N$ under the alternative.

Throughout the article, we assume that ω remains constant and focus on parameter changes in (α, β). Here, we elaborate on the possible effect of a change in ω. Under the "boundary" and explosive GARCH cases, our change-point test would not be affected by a change in ω. Such a change in ω could only cause complications if a stationary GARCH process ($E\log(\alpha_0\epsilon_0^2 + \beta_0) < 0$) changes into a different stationary process ($E\log(\alpha_*\epsilon_0^2 + \beta_*) < 0$). In this scenario, it is possible to reject H_0^* even if a change in ω causes $\sup_{0<t<1} Z_N(t)/w(t)$ to exceed the critical value. If this possibility is suspected, we recommend dividing the data into two subsamples at $\hat{k}_N$, as given by (4.1), and then testing the stationarity on both subsamples. If the two subsamples are indeed stationary, then the parameters can be estimated from each subsample. By comparing the estimates for ω, one can decide whether the change is due to ω or to (α, β).

We also note that Assumption 2.3 excludes parameter values on the boundary, such as $\alpha_0 = 0$ or $\beta_0 = 0$, as is standard in the literature on GARCH models. This is because the QMLE $\hat{\boldsymbol{\theta}}_N$ is not asymptotically normal when $\alpha_0 = 0$ or $\beta_0 = 0$. For these special cases, the asymptotic results are provided by Francq and Zakoian (2007) and Iglesias and Linton (2007). Therefore, one typically needs to have prior knowledge of whether a parameter of the GARCH(1,1) model is zero.

In this article, we only consider GARCH(1,1) processes. Our results can be easily extended to more general GARCH(p, q) processes in the stationary case, where the condition for stationarity is that the Lyapunov exponent is negative. Under this condition, (6.5) holds (cf. Berkes et al., 2003; Francq and Zakoian, 2004). When the Lyapunov exponent is strictly positive, Chan and Ng (2009) show that $\sigma_i^2 \to \infty$ a.s., which allows them to establish the asymptotic normality of the QMLE. However, to obtain the analogues of Theorems 2.1 and 2.2, one needs the rate of convergence of σ_i^2 to ∞. It is necessary for the weighted approximation in (5.16) to hold for GARCH(p, q) processes. Our method can be extended to the more general case if one can establish even a polynomial rate for $\sigma^2 \to \infty$, that is, a weaker result than (6.7) where the Lyapunov exponent is 0 and (6.10) where the Laypunov exponent is positive.

3. MONTE CARLO SIMULATIONS

In this section, we perform a simulation study to evaluate the finite sample performance of the proposed tests under the null and alternative hypotheses. The theory developed in Section 2 leads to two classes of tests, one based on Theorem 2.1(i) and the other based on Theorem 2.1(ii). Our simulation shows that the former has better size control and higher power. For the sake of brevity, we present simulation results for the test based on Theorem 2.1(i) and showcase its performance across all three cases of GARCH(1,1): the stationary case ($E\log(\alpha_0\epsilon_0^2 + \beta_0) < 0$), the "boundary" case ($E\log(\alpha_0\epsilon_0^2 + \beta_0) = 0$), and the explosive case ($E\log(\alpha_0\epsilon_0^2 + \beta_0) > 0$).

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Table 1. Critical values.

κ	10%	5%	1%
0.00	1.452	1.581	1.848
0.15	1.825	1.981	2.305
0.25	2.137	2.314	2.679
0.35	2.529	2.730	3.138

3.1. Empirical Size under H_0

Under the null hypothesis, our data generating process (DGP) is a GARCH(1,1) model:

$$y_i = \sigma_i \epsilon_i \quad \text{and} \quad \sigma_i^2 = \omega_0 + \alpha_0 y_{i-1}^2 + \beta_0 \sigma_{i-1}^2, \quad 1 \le i < N.$$

For the distribution of the innovation term, we have two settings: (*i*) ϵ_i follows a standard normal distribution $\mathcal{N}(0,1)$, and (*ii*) ϵ_i follows the skewed t distribution of Hansen (1994) with degrees of freedom ν and skewness λ. We use $\nu = 10$ and $\lambda = -0.15$, which are taken from Patton (2013) and are close to the typical values implied by empirical financial returns.

We consider all three cases of GARCH(1,1). In the stationary case, the parameter values are set to $(\omega_0, \alpha_0, \beta_0)^\top = (0.014, 0.084, 0.905)^\top$. These values are taken from Francq and Zakoian (2019) (Table 7.4, S&P 500) and reflect real data of stationary financial returns. In the explosive case, the parameter values are set to $(\omega_0, \alpha_0, \beta_0)^\top = (0.014, 0.084, 1.000)^\top$, which violates the strict stationarity condition of GARCH(1,1). In the "boundary" case, the parameter values depend on the distribution of the innovation term: they are set to $(\omega_0, \alpha_0, \beta_0)^\top = (0.014, 0.084, 0.9219)^\top$ for the standard normal distribution, and $(\omega_0, \alpha_0, \beta_0)^\top = (0.014, 0.084, 0.9238)^\top$ for the skewed t distribution.

To implement the test based on Theorem 2.1(*ii*), we require a weight function that satisfies Assumption 2.4. We investigate the impact of different shapes of the weight function by specifying it as $w(t) = (t(t-1))^\kappa$, where $\kappa = 0, 0.15, 0.25$, and 0.35. From Theorem 2.1(*i*), we calculate the test statistic

$$T_N = \sup_{0<t<1} \frac{Z_N(t)}{w(t)}, \tag{3.1}$$

using the steps detailed in the Appendix.

To obtain the critical values, we simulate two independent Brownian bridges $B_1(t)$ and $B_2(t)$ on a grid of 100,000 equally spaced points in $[0,1]$ and calculate the quantity $\sup_{0<t<1} \frac{1}{w(t)} \left(B_1^2(t) + B_2^2(t)\right)^{1/2}$. This is repeated 100,000 times, and we obtain the empirical percentile values of this quantity at the 90%, 95%, and 99%, which are corresponding to the critical values at the 10%, 5%, and 1% levels. The resulting critical values are shown in Table 1.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

TABLE 2. Empirical size.

	$\epsilon_i \sim \mathcal{N}(0,1)$			$\epsilon_i \sim$ skewed t		
	$N = 1,000$	$N = 2,000$	$N = 5,000$	$N = 1,000$	$N = 2,000$	$N = 5,000$
Stationary case						
$\kappa = 0.00$	3.4%	3.4%	3.7%	3.0%	2.8%	3.9%
$\kappa = 0.15$	4.2%	3.8%	3.9%	3.7%	3.3%	4.0%
$\kappa = 0.25$	5.0%	4.7%	4.2%	5.1%	4.5%	4.9%
$\kappa = 0.35$	7.1%	6.8%	6.1%	8.1%	7.7%	7.4%
"Boundary" case						
$\kappa = 0.00$	3.4%	3.7%	4.5%	3.6%	3.8%	4.1%
$\kappa = 0.15$	4.1%	4.2%	4.7%	4.2%	4.2%	4.4%
$\kappa = 0.25$	5.0%	4.9%	5.3%	5.6%	5.1%	5.3%
$\kappa = 0.35$	7.5%	7.3%	6.8%	8.7%	7.8%	7.7%
Explosive case						
$\kappa = 0.00$	4.1%	4.5%	4.9%	4.1%	4.2%	4.6%
$\kappa = 0.15$	5.1%	5.2%	5.3%	4.9%	4.8%	4.9%
$\kappa = 0.25$	6.3%	6.4%	6.1%	6.2%	5.9%	5.7%
$\kappa = 0.35$	8.8%	8.9%	8.4%	9.6%	8.7%	8.4%

We now perform the Monte Carlo simulation 10,000 times and record the computed test statistics. By comparing these simulated statistics with the critical values in Table 1, we obtain the empirical size under the null hypothesis. Table 2 reports the empirical sizes at the 5% significance level with a range of values of κ for the three GARCH(1,1) cases. The differences between the three cases are marginal, with slightly higher sizes observed in the explosive case. Additionally, the empirical sizes under the normal and skewed t innovations are similar. Notably, the empirical size depends on the choice of κ. Recall that κ controls the shape of the weight function, which can be chosen arbitrarily between 0 and 0.5. Our simulation reveals that a smaller κ is more conservative in rejection, while a larger κ leads to a higher rejection percentage. Generally, our proposed test has reasonably good size control with $\kappa = 0.15$ or 0.25, whereas it is under-sized for $\kappa = 0$ and over-sized for $\kappa = 0.35$. Overall, we recommend $\kappa = 0.15$ because its empirical size is typically close to or slightly below the theoretical levels, and there is no compromise in the empirical power with this choice of κ, as shown in Section 3.2.

3.2. Empirical Power under H_A

We now turn to the analysis of the empirical power. The DGP under H_A is similar to that under H_0 but with a change in one model parameter at time k^*. We set

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

the time of the change at $k^* = 0.5N$, corresponding to the middle of the sample period.[1]

There are various possible scenarios under the alternative hypothesis; we consider the following representative scenarios. The parameter values of GARCH(1,1) before k^* are set to be the same as those in the DGP of the three GARCH cases under the null. Since ω cannot be consistently estimated in the explosive regime of GARCH(1,1), we analyze a change in either α or β in the simulation under H_A. We consider both positive and negative changes in α or β. Specifically, after the change at k^*, the parameter values are as follows:

- $H_{A,1}$: a negative change in β, $\beta_* = \beta_0 - 0.05$,
- $H_{A,2}$: a positive change in β, $\beta_* = \beta_0 + 0.05$,
- $H_{A,3}$: a negative change in α, $\alpha_* = \alpha_0 - 0.05$,
- $H_{A,4}$: a positive change in α, $\alpha_* = \alpha_0 + 0.05$.

Table 3 reports the empirical power for changes in the stationary, "boundary", and explosive GARCH(1,1) cases. Several noticeable observations can be made. First, the empirical power increases with the sample size N under all scenarios, indicating the consistency of the proposed test. Second, our test has higher power when detecting changes in the explosive case than in the stationary and "boundary" cases. This can be attributed to the fact by that σ_i^2 and y_i^2 converge to infinity at a much faster rate (see Theorem 6.1), and the terms that are not in the limit are exponentially small. Thus, changes in the explosive case can be detected more easily. Third, the power is slightly lower under the skewed t innovation than under the normal innovation. This could be because the QMLE becomes the MLE for the normal innovation, and the derivatives are calculated exactly. Lastly, the choice of κ has a nontrivial but not substantial impact on the power. Choosing $\kappa = 0.15$ can provide satisfactory and balanced performance across a variety of scenarios.

4. EMPIRICAL APPLICATIONS

In this section, we first present a prototype demonstration by applying our testing procedure to the same five individual stocks analyzed by Francq and Zakoïan (2012). After that, we employ our test on a much larger scale by applying it to 10,620 stocks in the CRSP database. Specifically, we use all stocks with at least two years of price records between 2011 and 2022 in the NYSE, AMEX, and NASDAQ stock markets. We use the test statistic T_N with $\kappa = 0.15$ because it has the best finite sample performance, reflected by its better size control and higher power in our simulation study.

[1] We have also performed the simulation with the time of the change at $k^* = 0.25N$ and $0.75N$, corresponding to early and late changes in the sample period. In both cases, we can observe a modest reduction in the empirical power. Thus, the test is generally more sensitive to a change in the middle of the sample than an early or late change, which is a typical finding in the change point literature.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

TABLE 3. Empirical power.

	$H_{A,1}$			$H_{A,2}$			$H_{A,3}$			$H_{A,4}$		
	$N=1,000$	$N=2,000$	$N=5,000$	$N=1,000$	$N=2,000$	$N=5,000$	$N=1,000$	$N=2,000$	$N=5,000$	$N=1,000$	$N=2,000$	$N=5,000$
Panel A: Stationary case												
$\epsilon_i \sim \mathcal{N}(0,1)$												
$\kappa = 0.00$	84.2%	99.7%	100.0%	99.8%	100.0%	100.0%	92.2%	100.0%	100.0%	79.2%	99.9%	100.0%
$\kappa = 0.15$	81.8%	99.5%	100.0%	99.6%	100.0%	100.0%	90.7%	100.0%	100.0%	76.8%	99.9%	100.0%
$\kappa = 0.25$	78.8%	99.4%	100.0%	99.5%	100.0%	100.0%	88.5%	100.0%	100.0%	74.2%	99.8%	100.0%
$\kappa = 0.35$	73.1%	98.9%	100.0%	99.1%	100.0%	100.0%	84.6%	99.9%	100.0%	70.0%	99.7%	100.0%
$\epsilon_i \sim$ skewed t												
$\kappa = 0.00$	56.2%	91.9%	100.0%	91.7%	100.0%	100.0%	65.0%	96.9%	100.0%	44.5%	91.7%	100.0%
$\kappa = 0.15$	53.3%	90.5%	100.0%	90.3%	100.0%	100.0%	62.5%	96.2%	100.0%	42.2%	90.5%	100.0%
$\kappa = 0.25$	50.3%	88.5%	100.0%	88.6%	100.0%	100.0%	59.8%	95.4%	99.9%	40.1%	89.2%	100.0%
$\kappa = 0.35$	45.7%	85.0%	100.0%	85.0%	99.9%	100.0%	55.2%	94.0%	99.9%	37.9%	86.7%	100.0%
Panel B: "Boundary" case												
$\epsilon_i \sim \mathcal{N}(0,1)$												
$\kappa = 0.00$	96.2%	100.0%	100.0%	100.0%	100.0%	100.0%	98.7%	100.0%	100.0%	94.5%	100.0%	100.0%
$\kappa = 0.15$	95.5%	99.9%	100.0%	100.0%	100.0%	100.0%	98.4%	100.0%	100.0%	93.4%	100.0%	100.0%
$\kappa = 0.25$	94.5%	99.9%	100.0%	99.9%	100.0%	100.0%	97.9%	100.0%	100.0%	92.2%	100.0%	100.0%
$\kappa = 0.35$	92.3%	99.8%	100.0%	99.9%	100.0%	100.0%	96.8%	100.0%	100.0%	90.0%	100.0%	100.0%

(Continued)

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

TABLE 3. Continued

	$H_{A,1}$			$H_{A,2}$			$H_{A,3}$			$H_{A,4}$		
	$N=1,000$	$N=2,000$	$N=5,000$	$N=1,000$	$N=2,000$	$N=5,000$	$N=1,000$	$N=2,000$	$N=5,000$	$N=1,000$	$N=2,000$	$N=5,000$
$\epsilon_i \sim$ skewed t												
$\kappa=0.00$	80.2%	97.9%	100.0%	97.6%	100.0%	100.0%	86.0%	99.4%	100.0%	73.6%	99.4%	100.0%
$\kappa=0.15$	77.7%	97.5%	100.0%	97.2%	100.0%	100.0%	83.8%	99.3%	100.0%	71.3%	99.2%	100.0%
$\kappa=0.25$	75.4%	97.1%	100.0%	96.6%	100.0%	100.0%	82.0%	99.0%	100.0%	69.1%	99.0%	100.0%
$\kappa=0.35$	71.5%	96.0%	100.0%	95.3%	100.0%	100.0%	78.0%	98.7%	100.0%	65.8%	98.7%	100.0%
Panel C: Explosive case												
$\epsilon_i \sim \mathcal{N}(0,1)$												
$\kappa=0.00$	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	96.2%	100.0%	100.0%
$\kappa=0.15$	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	95.5%	100.0%	100.0%
$\kappa=0.25$	99.9%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	94.7%	100.0%	100.0%
$\kappa=0.35$	99.9%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	93.2%	100.0%	100.0%
$\epsilon_i \sim$ skewed t												
$\kappa=0.00$	97.9%	100.0%	100.0%	97.9%	100.0%	100.0%	98.8%	100.0%	100.0%	78.7%	99.3%	100.0%
$\kappa=0.15$	97.5%	100.0%	100.0%	97.5%	100.0%	100.0%	98.3%	100.0%	100.0%	77.0%	99.2%	100.0%
$\kappa=0.25$	96.9%	100.0%	100.0%	97.0%	100.0%	100.0%	97.8%	100.0%	100.0%	75.0%	99.0%	100.0%
$\kappa=0.35$	95.9%	100.0%	100.0%	95.9%	100.0%	100.0%	96.9%	100.0%	100.0%	71.7%	98.7%	100.0%

TABLE 4. Test results of change point.

	ICGN	MCBF	KV-A	BTC	CCME
Start date	1 Jun. 2007	29 Aug. 2007	3 Apr. 2006	2 Jul. 2007	1 Apr. 2009
End date	4 Feb. 2011	7 Feb. 2011	7 Feb. 2011	7 Feb. 2011	7 Feb. 2011
Sample size	928	868	1,221	908	469
$\hat{\alpha}$	0.581	0.023	0.143	0.508	0.413
$\hat{\beta}$	0.696	0.979	0.927	0.765	0.750
FZ(2012) test stat.	−2.297	0.024	1.120	0.491	0.457
FZ(2012) test p-value	0.011	0.510	0.869	0.688	0.676
Our test stat.	1.461	2.186	0.933	0.947	1.197
Rejection at 5%	No	Yes	No	No	No
Time of change	–	9 Feb. 2009	–	–	–

4.1. Prototype Application

In the prototype application, we focus on the daily returns of the five individual stocks analyzed by Francq and Zakoïan (2012), with the same sample period they selected. The five stocks are Icagen (Nasdaq:ICGN), Monarch Community Bancorp (Nasdaq:MCBF), KV Pharmaceutical (NYSE:KV-A), Community Bankers Trust (AMEX:BTC), and China MediaExpress (Nasdaq:CCME). The sample period is generally between 2007 and 2011, with KV-A from 2006 and CCME from 2009. These five stocks provide a useful demonstration because the nonstationary (explosive) assumption cannot be rejected for four of them at any reasonable significance level; see Table 8 in Francq and Zakoïan (2012). Our change-point test can provide further insights into the stability of the two parameters (α, β) of GARCH(1,1) models for those five stocks.

Table 4 presents the results of the parameter estimation, the nonstationarity test of Francq and Zakoïan (2012), and our change-point test for the five individual stock returns. We reproduce the parameter estimation and the nonstationary test results reported in Francq and Zakoïan (2012). Regarding our change-point test, we reveal a change in the two parameters of the GARCH(1,1) model for the stock MCBF, which indicates the rejection of H_0^* at the 5% significance level.

We further use a change-point estimator

$$\hat{k}_N = \lfloor N\hat{t}^* \rfloor, \quad \text{where } \hat{t}^* = \arg\max_{0<t<1} \frac{Z_N(t)}{w(t)}, \tag{4.1}$$

to find the time of the change. It shows that the change date is on 9 February 2009 for MCBF. By splitting the samples, we reestimate the two parameters before and after the change. The estimates are $\hat{\alpha} = 0.118, \hat{\beta} = 0.886$ before the change and $\hat{\alpha} = 0.013, \hat{\beta} = 0.981$ after the change. This result suggests that, in this particular case, the ARCH effect is weaker after the change, while the GARCH effect becomes stronger.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

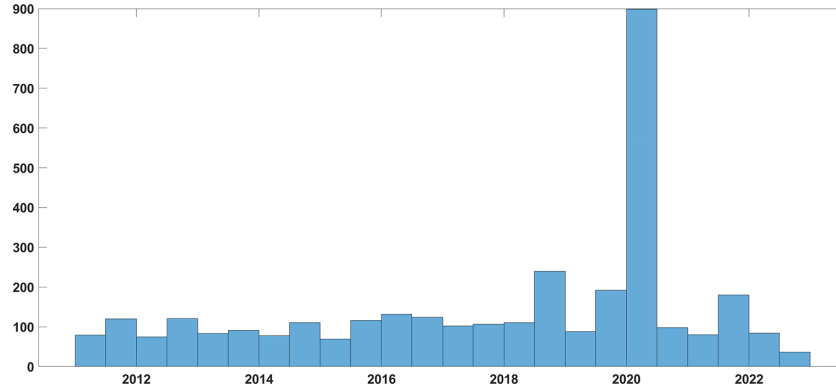


FIGURE 1. Histogram of times of changes in (α, β) for 3,416 U.S. stocks.

4.2. Extensive Application

We repeat our prototype analysis on the universe of U.S. stocks in the CRSP database. We select all stocks with at least 2 years of records between 4 January 2011 and 30 December 2022 in the NYSE, AMEX, and NASDAQ stock markets, resulting in a total of 10,620 stocks. We choose this period to focus on the most recent decade. We note that not all stocks have observations covering the full period, because some stocks had an initial public offering (IPO) after 4 January 2011, or were delisted before 30 December 2022. The nonstationarity test of Francq and Zakoïan (2012) suggests that there are 631 stocks for which the null hypothesis of nonstationarity (explosivity) cannot be rejected at the 1% significance level.

We apply our change-point test to the daily returns[2] of 10,620 stocks, regardless of whether they are in a stationary or explosive regime. We find evidence of change points in (α, β) for 3,416 stocks at the 5% level. Using the change-point estimator defined in (4.1), we further find the time of the change for stocks that are rejected by our test. Figure 1 shows a histogram of the times of detected changes. As can be seen, the largest number of changes occurred in first half of 2020 when the COVID-19 pandemic begun to affect the United States economy. Thus, our test is able to detect change points that align with a well-known market event, providing validation using a large dataset.

5. CONCLUSIONS

This article provides a novel test procedure to detect changes in GARCH(1,1) processes without imposing a stationary assumption. Our methodology complements

[2]We use *daily returns without dividend* from the CRSP database to avoid the complication caused by stock splits when calculating log returns from raw prices. We also performed the analysis using *daily returns with dividend*, and the results showed only slight differences.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

the existing literature by allowing for change points to be detected in GARCH(1,1) models not only in the stationary case but also in the explosive and "boundary" cases. A key advantage of the proposed test is that it does not require practitioners to have prior knowledge on whether their empirical observations are stationary or explosive. The test procedure is applicable in all cases, offering a useful addition to the toolbox of financial econometricians. With other relevant theories developed in the future, it would be interesting to extend the test procedure to more general GARCH(p,q) processes.

6. PROOFS OF TECHNICAL RESULTS

6.1. The Properties of σ_i^2 as $i \to \infty$

The properties of σ_i^2 depend on the sign of $E\log(\alpha_i\epsilon_0^2+\beta_0)$. We consider the following cases:

$$E\log(\alpha_i\epsilon_0^2+\beta_0) < 0, \tag{6.1}$$

$$E\log(\alpha_i\epsilon_0^2+\beta_0) = 0, \tag{6.2}$$

and

$$E\log(\alpha_i\epsilon_0^2+\beta_0) > 0. \tag{6.3}$$

The case (6.1) is known as the stationary case, (6.2) as the "boundary" case, and (6.3) as the explosive case.

In the case (6.1), there is a unique stationary non anticipative sequence satisfying

$$x_i = h_i\epsilon_i, \qquad h_i^2 = \omega_0+\alpha_0 x_{i-1}^2+\beta_0 h_{i-1}^2, \qquad -\infty < i < \infty. \tag{6.4}$$

The next theorem shows that σ_i^2 is close to h_i^2, as $i \to \infty$, if (6.1) holds. We also show that in the "boundary" and explosive cases $\sigma_i^2 \to \infty$ in probability but with different rates. Thus, σ_i^2 is explosive under (6.2) as well as (6.3).

THEOREM 6.1. *We assume that Assumptions 2.1 and 2.2 are satisfied.*
(i) *If (6.1) holds, then there exist* $\bar\kappa > 0$ *and* $0 < \bar\rho < 1$ *such that*

$$E|\sigma_i^2 - h_i^2|^{\bar\kappa} = O\left(\bar\rho^i\right). \tag{6.5}$$

(ii) *If (6.2) holds and*

$$E|\log(\alpha_0\epsilon_0^2+\beta_0)|^{\bar\nu} < \infty \quad \textit{with some} \quad \bar\nu > 2, \tag{6.6}$$

then for all $1/\bar\nu < \xi < 1/2$

$$P\left\{\sigma_i^2 < \exp\left(c_1 i^{\xi}\right)\right\} \le c_2\left(i^{\xi-1/2}+i^{1-\xi\bar\nu}\right) \tag{6.7}$$

with some c_1 *and* c_2.
(iii) *If (6.3) holds, then*

$$\exp(-S(i-1))\sigma_i^2 \to R \quad a.s.,$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

where

$$R = \omega_0 \sum_{k=0}^{\infty} \exp(-S(k)) + \sigma_0^2(\alpha_0\epsilon_0^2 + \beta_0) \geq \omega_0, \qquad P\{R = 0\} = 0,$$

and

$$S(j) = \sum_{k=1}^{j} \log(\alpha_0\epsilon_k^2 + \beta_0), \qquad S(0) = 0. \tag{6.8}$$

For any $m < \bar{m} = E(\log(\alpha_0\epsilon_0^2 + \beta_0)$, *we have*

$$\exp(-mi)\sigma_i^2 \xrightarrow{P} \infty. \tag{6.9}$$

If (6.6) also holds, then for all $1/\bar{\nu} < \xi < 1/2$

$$P\left\{\sigma_i^2 < \exp(\bar{m}i - c_1 i^{\xi})\right\} \leq c_2\left(i^{1-\xi\bar{\nu}} + \exp(-c_3 i)\right) \tag{6.10}$$

with some c_1, c_2, *and* c_3.

Proof. We begin by proving Theorem 6.1(*i*). The recursion in (2.1) yields

$$\sigma_i^2 = \omega_0 + (\alpha_0\epsilon_i^2 + \beta_0)\sigma_{i-1}^2,$$

which can be solved explicitly

$$\sigma_i^2 = \omega_0 \sum_{\ell=1}^{i} \prod_{j=1}^{\ell-1} (\alpha_0\epsilon_{i-j}^2 + \beta_0) + \sigma_0^2 \prod_{\ell=0}^{i-1} (\alpha_0\epsilon_\ell^2 + \beta_0). \tag{6.11}$$

The explicit solution of the equation (6.4) is given by

$$h_i^2 = \omega_0 \sum_{\ell=1}^{\infty} \prod_{k=1}^{\ell-1} (\alpha_0\epsilon_{i-k}^2 + \beta_0)$$

($\prod_{\emptyset} = 1$). If (6.1) holds, then there exists a $\bar{\kappa} > 0$ such that

$$\bar{\rho} = E(\alpha_0\epsilon_0^2 + \beta_0)^{\bar{\kappa}} < 1 \tag{6.12}$$

(cf. Berkes et al., 2003; Francq and Zakoian, 2004). We can assume that $0 < \bar{\kappa} < 1$, so by Jensen's inequality

$$E\left(\sum_{\ell=i+1}^{\infty} \prod_{j=1}^{\ell-1} (\alpha_0\epsilon_{i-j}^2 + \beta_0)\right)^{\bar{\kappa}} \leq \sum_{\ell=i+1}^{\infty} E\left(\prod_{j=1}^{\ell-1} (\alpha_0\epsilon_{i-j}^2 + \beta_0)\right)^{\bar{\kappa}} \leq \sum_{\ell=i+1}^{\infty} \bar{\rho}^{\ell-1},$$

and

$$E\left(\prod_{\ell=0}^{i-1} (\alpha_0\epsilon_\ell^2 + \beta_0)\right)^{\bar{\kappa}} \leq \frac{\bar{\rho}^{i-1}}{1-\bar{\rho}}.$$

Hence, (6.5) in Theorem 6.1(*i*) is proven.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

We now proceed with the proof of Theorem 6.1(*ii*). It follows from (6.11) that

$$\begin{aligned}\sigma_i^2 &\geq \omega_0 \sum_{\ell=1}^{i} \prod_{j=1}^{\ell-1} (\alpha_0 \epsilon_{i-j}^2 + \beta_0) \qquad \textbf{(6.13)}\\ &\geq \omega_0 \max_{1\leq \ell \leq i} \prod_{j=1}^{\ell-1} (\alpha_0 \epsilon_{i-j}^2 + \beta_0)\\ &= \omega_0 \exp\left(\max_{1\leq \ell \leq i} \sum_{j=1}^{\ell-1} \log(\alpha_0 \epsilon_{i-j}^2 + \beta_0) \right).\end{aligned}$$

Using Assumption 2.1, we have

$$P\left\{ \max_{1\leq \ell \leq i} \sum_{j=1}^{\ell-1} \log(\alpha_0 \epsilon_{i-j}^2 + \beta_0) < c_1 i^{\xi} \right\} = P\left\{ \max_{1\leq \ell \leq i-1} S(\ell) < c_1 i^{\xi} \right\},$$

where $S(\ell)$ is defined in (6.8). Let $m_1^2 = E(\log(\alpha_0 \epsilon_0^2 + \beta_0)^2)$. Using the Komlós–Major–Tusnády approximation (cf. Theorem 2.6.7 in Csörgo and Révész, 1981), there is a Wiener process $\{W(x), x \geq 0\}$ such that

$$P\left\{ \max_{0\leq x\leq i} |S(\lfloor x \rfloor) - m_1 W(x)| > x^{\xi} \right\} \leq c_3 i^{1-\xi\bar{\nu}}.$$

By the scale transformation of $W(x)$ and Theorem 1.5.1 in Csörgő and Révész (1981), we have

$$P\left\{ \sup_{0\leq x\leq i} W(x) \leq c_4 i^{\xi} \right\} = P\left\{ |\mathcal{N}| \leq c_4 i^{\xi - 1/2} \right\} \leq c_5 i^{\xi - 1/2},$$

where $\mathcal{N}$ is a standard normal random variable, completing the proof of (6.6) and (6.7) in Theorem 6.1(*ii*).

Lastly, we turn to the proof of Theorem 6.1(*iii*). The formula in (6.11) can be written as

$$\sigma_i^2 = \omega_0 \sum_{\ell=1}^{i} \exp(S(i-1) - S(i-\ell)) + \sigma_0^2 (\alpha_0 \epsilon_0^2 + \beta_0) \exp(S(i-1)), \qquad \textbf{(6.14)}$$

where $S(j)$ is defined in (6.8). Thus, we obtain

$$\sigma_i^2 = \exp(S(i-1)) \left(\omega_0 \bar{R}_i + \sigma_0^2 (\alpha_0 \epsilon_0^2 + \beta_0) \right),$$

where

$$\bar{R}_i = \sum_{k=0}^{i-1} \exp(-S(k)).$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

According to the strong law of large numbers

$$\lim_{k\to\infty}\frac{1}{k}S(k)=E\log(\alpha_0\epsilon_0^2+\beta_0)>0 \qquad \text{a.s.}$$

and therefore

$$\lim_{k\to\infty}\bar{R}_k=\bar{R}=\sum_{\ell=0}^{\infty}\exp(-S(\ell)), \qquad \text{a.s.}$$

and $P\{\bar{R}<\infty\}=1$. We again use (6.13). By the Komlós–Major–Tusnády approximation (cf. Theorem 2.6.7 in Csörgő and Révész, 1981), we can define a Wiener process $\{W(x),x\geq 0\}$ such that

$$P\left\{\max_{0\leq x\leq i}|S(\lfloor x\rfloor)-(m_1W(x)+\bar{m}x)|>i^{\xi}\right\}\leq c_4 i^{1-\xi\bar{\nu}}.$$

By the scale transformation of the Wiener process, we have

$$\sup_{0\leq x\leq i}(m_1W(x)+\bar{m}x)\stackrel{\mathcal{D}}{=}\frac{m_1^2}{\bar{m}}\sup_{0\leq u\leq i\bar{m}^2/m_1^2}(W(u)+u).$$

Applying again the Komlós–Major–Tusnády approximation, we can define a Poisson process $\{\mathscr{P}(x),x\geq 0\}$ such that

$$P\left\{\sup_{1\leq x\leq i}|(W(x)+x)-\mathscr{P}(x)|>i^{\xi}\right\}\leq c_5 i^{1-\xi\bar{\nu}}. \tag{6.15}$$

It can be seen that

$$\sup_{0\leq x\leq i}\mathscr{P}(x)=\mathscr{P}(i).$$

Thus, we can use (6.15) to show that

$$\begin{aligned}P\left\{\max_{1\leq\ell\leq i}S(\ell)<c_6x^{\xi}\right\}&\leq P\left\{W(i)+i\leq 2c_6x^{\xi}\right\}+c_7i^{1-\xi\bar{\nu}}\\&\leq P\left\{\mathcal{N}<-i+2c_6x^{\xi}\right\}\\&\leq 2c_7i^{1-\xi\bar{\nu}}+c_8\exp(-i^2/3),\end{aligned}$$

where $\mathcal{N}$ is a standard normal random variable, completing the proof of (6.8) and (6.9) in Theorem 6.1(*iii*). □

Remark 6.1. It follows from Assumption 2.1 and (6.5) that

$$E|y_i^2-x_i^2|^{\bar{\kappa}}=O\left(\bar{\rho}^i\right).$$

If $F(u)$ denotes the distribution function of ϵ_0^2 and

$$|F(u)-F(v)|\leq c_1|u-v|^{c_2} \qquad \text{in a neighborhood of 0,}$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

then

$$P\left\{\epsilon_0^2 \le i^{-c_3/c_2}\right\} \le c_1 i^{-c_3},$$

and therefore y_i^2 increases to ∞ at an exponential rate.

6.2. Proofs of Theorems 2.1 and 2.1

Proof of Theorem 2.1. We show in Sections 6.3 and 6.4 that we can define Gaussian processes $\{\mathbf{W}_N(x), x \ge 0\}$ such that

$$\max_{x\in[1,N-1]} \left(\frac{N}{x(N-x)}\right)^{\nu} \left\| \bar{\mathscr{F}}_x(\hat{\boldsymbol{\theta}}_N) - \boldsymbol{\Gamma}_N(x) \right\| = O_P(1) \tag{6.16}$$

with some $\nu < 1/2$,

$$\boldsymbol{\Gamma}_N(x) = \mathbf{W}_N(x) - \frac{x}{N}\mathbf{W}_N(N), \quad 0 \le x \le N,$$

$E\mathbf{W}_N(x) = \mathbf{0}$, and $E\mathbf{W}_N(x)\mathbf{W}_N^{\top}(y) = \min(x,y)\mathbf{D}$. Checking the covariance functions, one can verify that

$$\left\{N^{-1/2}\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_N(Nt), 0 \le t \le 1\right\} \overset{\mathscr{D}}{=} \{\mathbf{B}(t), 0 \le t \le 1\}, \tag{6.17}$$

$$\mathbf{B}(t) = (B_1(t), B_2(t))^{\top}, \tag{6.18}$$

where $\{B_1(t), 0 \le t \le 1\}$ and $\{B_2(t), 0 \le t \le 1\}$ are independent Brownian bridges. The approximation in (6.16) can be rewritten as

$$N^{1/2-\nu} \sup_{1/(N+1)\le t\le 1-1/(N+1)} \frac{1}{(t(1-t))^{\nu}} \left\| N^{-1/2}\bar{\mathscr{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N) - N^{-1/2}\boldsymbol{\Gamma}_N(Nt) \right\| = O_P(1), \tag{6.19}$$

which implies that

$$\sup_{1/(N+1)\le t\le 1-1/(N+1)} \frac{1}{(t(1-t))^{1/2}} \left\| N^{-1/2}\bar{\mathscr{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N) - N^{-1/2}\boldsymbol{\Gamma}_N(Nt) \right\| = O_P(1). \tag{6.20}$$

We can show that (6.19) and (6.20) yield

$$\sup_{0<t<1} \frac{1}{w(t)} \frac{1}{N^{1/2}} \left\{ \bar{\mathscr{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\mathbf{D}^{-1} \left(\bar{\mathscr{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\right)^{\top} \right\}^{1/2} \overset{\mathscr{D}}{\to} \sup_{0<t<1} \frac{1}{w(t)} \|\mathbf{B}(t)\|. \tag{6.21}$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Using Assumption 2.4(i) and the approximation in (6.19), we get for all $0 < \delta < 1/2$ that

$$\sup_{\delta\le t\le 1-\delta}\frac{1}{w(t)}\frac{1}{N^{1/2}}\left\{\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\mathbf{D}^{-1}\left(\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\right)^{\top}\right\}^{1/2}\stackrel{\mathscr{D}}{\to}\sup_{\delta\le t\le 1-\delta}\frac{1}{w(t)}\|\mathbf{B}(t)\|. \tag{6.22}$$

If $I(w,c)<\infty$ with some $c>0$, then

$$\lim_{t\to 0}\frac{t^{1/2}}{w(t)}=0 \tag{6.23}$$

(cf. Section 4.1 in Csörgő and Horváth, 1993). We have, by (6.20), that

$$\begin{aligned}&\sup_{1/(N+1)<t\le\delta}\frac{1}{w(t)}\left\|N^{-1/2}\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)-N^{-1/2}\boldsymbol{\Gamma}_N(Nt)\right\|\\ &\quad=\sup_{1/(N+1)<t\le\delta}\frac{t^{1/2}}{w(t)}\frac{1}{t^{1/2}}\left\|N^{-1/2}\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)-N^{-1/2}\boldsymbol{\Gamma}_N(Nt)\right\|\\ &\quad=\sup_{1/(N+1)<t\le\delta}\frac{t^{1/2}}{w(t)}O_P(1).\end{aligned}$$

Hence, we conclude that

$$\lim_{\delta\to 0}\limsup_{N\to\infty}P\left\{\sup_{1/(N+1)\le t\le\delta}\frac{1}{w(t)}\left\|N^{-1/2}\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)-N^{-1/2}\boldsymbol{\Gamma}_N(Nt)\right\|>u\right\}=0$$

for all $u>0$ and by symmetry

$$\lim_{\delta\to 0}\limsup_{N\to\infty}P\left\{\sup_{1-\delta\le t\le 1-1/(N+1)}\frac{1}{w(t)}\left\|N^{-1/2}\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)-N^{-1/2}\boldsymbol{\Gamma}_N(Nt)\right\|>u\right\}=0.$$

Thus, we get

$$\sup_{1/(N+1)\le t\le 1-1/(N+1)}\frac{1}{w(t)}\left\|N^{-1/2}\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)-N^{-1/2}\boldsymbol{\Gamma}_N(Nt)\right\|=o_P(1).$$

Using (6.23) and the definition of $\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}$, we obtain

$$\sup_{0<t\le 1/(N+1)}\frac{1}{w(t)}\left\|N^{-1/2}\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\right\|=o_P(1)$$

and

$$\sup_{N/(N+1)\le t<1}\frac{1}{w(t)}\left\|N^{-1/2}\bar{\mathcal{F}}_{\lfloor (N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\right\|=o_P(1).$$

Because of (6.17) and (6.18)

$$\sup_{0<t\le\delta}\frac{1}{w(t)}\left\|N^{-1/2}\mathbf{D}^{-1/2}\boldsymbol{\Gamma}_N(Nt)\right\|\stackrel{\mathscr{D}}{=}\sup_{0<t\le\delta}\frac{1}{w(t)}\|\mathbf{B}(t)\|$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

and

$$\lim_{\delta\to 0}\sup_{0<t\le\delta}\frac{1}{w(t)}\|\mathbf{B}(t)\| = c_1 \qquad a.s.$$

(cf. Section 4.1 in Csörgő and Horváth, 1993). By symmetry,

$$\sup_{1-\delta\le t<1}\frac{1}{w(t)}\left\|N^{-1/2}\mathbf{D}^{-1/2}\mathbf{\Gamma}_N(Nt)\right\| \stackrel{\mathcal{D}}{=} \sup_{1-\delta\le t<1}\frac{1}{w(t)}\|\mathbf{B}(t)\|,$$

and

$$\lim_{\delta\to 0}\sup_{1-\delta\le t<1}\frac{1}{w(t)}\|\mathbf{B}(t)\| = c_2 \qquad a.s.$$

completing the proof of the first part of Theorem 2.1.

According to the law of the iterated logarithm,

$$\frac{1}{(2\log\log N)^{1/2}}\sup_{1/(N+1)\le t\le 1-1/(N+1)}\frac{\|\mathbf{B}(t)\|}{(t(1-t))^{1/2}} \stackrel{P}{\to} 1, \tag{6.24}$$

Csörgő and Horváth (1993) proved that

$$A_{N,1} = \sup_{1/(N+1)\le t\le(\log N)^4/N}\frac{\|\mathbf{B}(t)\|}{(t(1-t))^{1/2}} = O_P((\log\log\log N)^{1/2}), \tag{6.25}$$

$$A_{N,2} = \sup_{1-(\log N)^4/N\le t\le 1-1/(N+1)}\frac{\|\mathbf{B}(t)\|}{(t(1-t))^{1/2}} = O_P((\log\log\log N)^{1/2}). \tag{6.26}$$

Now putting together (6.20) and (6.24)–(6.26), we obtain

$$\frac{1}{(2\log\log N)^{1/2}}\sup_{1/(N+1)\le t\le 1-1/(N+1)}\frac{1}{(t(1-t))^{1/2}}\left\|N^{-1/2}\mathbf{D}^{-1/2}\bar{\mathbf{r}}_{\lfloor(N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\right\| \stackrel{P}{\to} 1, \tag{6.27}$$

$$A_{N,3} = \sup_{1/(N+1)\le t\le(\log N)^4/N}\frac{1}{(t(1-t))^{1/2}}\left\|N^{-1/2}\mathbf{D}^{-1/2}\bar{\mathbf{r}}_{\lfloor(N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\right\| \tag{6.28}$$

$$= O_P((\log\log\log N)^{1/2}),$$

and

$$A_{N,4} = \sup_{1-(\log N)^4/N\le t\le 1-1/(N+1)}\frac{1}{(t(1-t))^{1/2}}\left\|N^{-1/2}\mathbf{D}^{-1/2}\bar{\mathbf{r}}_{\lfloor(N+1)t\rfloor}(\hat{\boldsymbol{\theta}}_N)\right\| \tag{6.29}$$

$$= O_P((\log\log\log N)^{1/2}).$$

It follows from (6.24) to (6.29) that

$$a(\log N)\max\left\{A_{N,1},A_{N,2}\right\} - b(\log N) \stackrel{P}{\to} -\infty$$

and

$$a(\log N)\max\left\{A_{N,3},A_{N,4}\right\} - b(\log N) \stackrel{P}{\to} -\infty$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

resulting in

$$\lim_{N\to\infty} P\Bigg\{ \sup_{1/(N+1)\le t\le 1-1/(N+1)} \frac{1}{(t(1-t))^{1/2}} \left\| N^{-1/2}\mathbf{D}^{-1/2}\bar{\mathcal{r}}_{\lfloor (N+1)t\rfloor}\right\| = \sup_{(\log N)^4/(N+1)\le t\le 1-(\log N)^4/(N+1)} \frac{1}{(t(1-t))^{1/2}} \left\| N^{-1/2}\mathbf{D}^{-1/2}\bar{\mathcal{r}}_{\lfloor (N+1)t\rfloor}\right\| \Bigg\} = 1.$$

Using (6.19), we conclude that

$$\sup_{(\log N)^4/(N+1)\le t\le 1-(\log N)^4/(N+1)} \frac{1}{(t(1-t))^{1/2}} \left\| N^{-1/2}\mathbf{D}^{-1/2}\left(\bar{\mathcal{r}}_{\lfloor (N+1)t\rfloor} - \mathbf{\Gamma}_N(Nt)\right)\right\| = o_P((\log\log N)^{1/2}).$$

Csörgő and Horváth (1993) proved that

$$\lim_{N\to\infty} P\left\{ a(\log N) \sup_{(\log N)^4/(N+1)\le t\le 1-(\log N)^4/(N+1)} \frac{\|\mathbf{B}(t)\|}{(t(1-t))^{1/2}} \le x + b(\log N)\right\} = \exp\left(-2e^{-x}\right),$$

and therefore the limit result in Theorem 2.1 follows directly from (6.17). □

Proof of Theorem 2.2. Since $\mathbf{\Theta}$ is compact and $\bar{\ell}_i(\boldsymbol{\theta})$ is continuous, the set $\tilde{\mathbf{\Theta}}$ of the possible limit points of $\hat{\boldsymbol{\theta}}_N$ is countable. It follows from the proofs in Sections 6.3 and 6.4 that

$$\frac{1}{N}\bar{\mathcal{r}}_{k^*}(\boldsymbol{\theta}) \xrightarrow{P} \mathcal{a}_1(\boldsymbol{\theta})$$

and

$$\frac{1}{N}\left(\bar{\mathcal{r}}_N(\boldsymbol{\theta}) - \bar{\mathcal{r}}_{k^*}(\boldsymbol{\theta})\right) \xrightarrow{P} \mathcal{a}_2(\boldsymbol{\theta}),$$

uniformly in $\boldsymbol{\theta}\in\mathbf{\Theta}$. We note that $\mathcal{a}_1(\boldsymbol{\theta}) = \mathbf{0}$ if and only if the first two coordinates of $\boldsymbol{\theta}$ are (α_0, β_0) and similarly $\mathcal{a}_2(\boldsymbol{\theta}) = \mathbf{0}$ if and only if the first two coordinates of $\boldsymbol{\theta}$ are (α_*, β_*). Hence, for any $\tilde{\boldsymbol{\theta}}\in\tilde{\mathbf{\Theta}}$,

$$\max\left(\left\|\mathcal{a}_1(\tilde{\boldsymbol{\theta}})\right\|, \left\|\mathcal{a}_2(\tilde{\boldsymbol{\theta}})\right\|\right) > 0.$$

We only need to show that

$$\inf\left\{\max\left(\left\|\mathcal{a}_1(\tilde{\boldsymbol{\theta}})\right\|, \left\|\mathcal{a}_2(\tilde{\boldsymbol{\theta}})\right\|\right) : \tilde{\boldsymbol{\theta}}\in\tilde{\mathbf{\Theta}}\right\} > 0. \tag{6.30}$$

If the inf in (6.30) were 0, then we could construct a subsequence $\hat{\boldsymbol{\theta}}_{N_k}$ that would converge in probability to a point where the first two sets of coordinates are (α_0, β_0) and (α_*, β_*), which contradicts Assumption 2.6. □

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

6.3. Proof of (6.16) When $E\log(\alpha_0\epsilon_0^2+\beta_0)<0$

We use several results obtained by Berkes et al. (2003) and Francq and Zakoian (2004). A survey on volatility processes with detailed proofs is given in Francq and Zakoian (2010). First, we define

$$\bar{h}_i^2(\boldsymbol{\theta})=\omega+\alpha x_{i-1}^2+\beta\bar{h}_{i-1}^2(\boldsymbol{\theta}),$$

the stationary version of $\bar{\sigma}_i^2(\boldsymbol{\theta})$. Berkes et al. (2003) proved that

$$\bar{h}_i^2(\boldsymbol{\theta})=\mathfrak{c}_0(\boldsymbol{\theta})+\sum_{\ell=1}^{\infty}\mathfrak{c}_\ell(\boldsymbol{\theta})x_{i-\ell}^2,$$

where the functions $\{\mathfrak{c}_\ell(\boldsymbol{\theta}),0\le\ell<\infty\}$ are analytical, and

$$\sup\{|\mathfrak{c}_\ell(\boldsymbol{\theta})|:\ \boldsymbol{\theta}\in\boldsymbol{\Theta}\}=O(\rho_1^\ell)\qquad\text{with some}\qquad 0<\rho_1<1.$$

Hence, similarly to (6.5)

$$\sup\left\{|\bar{h}_i^2(\boldsymbol{\theta})-\bar{\sigma}_i^2(\boldsymbol{\theta})|:\ \boldsymbol{\theta}\in\boldsymbol{\Theta}\right\}=O(\rho_2^i)\quad\text{a.s.}\quad\text{with some}\qquad 0<\rho_2<1. \tag{6.31}$$

It is also known that

$$\|\hat{\boldsymbol{\theta}}_N-\boldsymbol{\theta}_0\|=O_P(N^{-1/2}) \tag{6.32}$$

(cf. Berkes et al., 2003; Francq and Zakoian, 2004). Furthermore, there exits a $\boldsymbol{\Theta}_0$, a neighborhood of $\boldsymbol{\theta}_0$, such that

$$\sup\left\{\|\nabla\bar{\sigma}_i^2(\boldsymbol{\theta})-\nabla\bar{h}_i^2(\boldsymbol{\theta})\|:\ \boldsymbol{\theta}\in\boldsymbol{\Theta}_0\right\}=O(\rho_3^i)\quad\text{a.s.}\quad\text{with some}\qquad 0<\rho_3<1. \tag{6.33}$$

It follows from the representation that h_i^2 is a Bernoulli shift; that is, there is a function $z:R^\infty\to R$ such that $h_i^2=z(\epsilon_{i-1}^2,\epsilon_{i-2}^2,\dots)$. We show in the following lemma that h_i^2 is decomposable at an exponential rate. Let

$$h_{i,j}^2=\omega_0\sum_{\ell=1}^{j}\prod_{k=1}^{\ell-1}(\alpha_0\epsilon_{i-k}^2+\beta_0)+\omega_0\sum_{\ell=j+1}^{\infty}\prod_{k=1}^{\ell-1}(\alpha_0(\epsilon_{i,i-k,j}^*)^2+\beta_0),\qquad j\ge 1,$$

and

$$h_{i,j}^2=\omega_0\sum_{\ell=1}^{\infty}\prod_{k=1}^{\ell-1}(\alpha_0(\epsilon_{i,i-k,j}^*)^2+\beta_0),\qquad j\le 0,$$

where $\{\epsilon_{i,k,j}^*,-\infty<i,k,j<\infty\}$ are independent copies of ϵ_0. We also define

$$x_{i,j}=\begin{cases}\epsilon_i h_{i-1,j}, & j\ge 1,\\ \epsilon_{i,i,j}^* h_{i-1,j}, & j\le 0.\end{cases}$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

In the next lemma, we establish the decomposable Bernoulli property of

$$\mathbf{g}_i = \left(\frac{1}{h_i^2} \frac{\partial \bar{h}_i(\boldsymbol{\theta}_0)}{\partial \alpha}, \frac{1}{h_i^2} \frac{\partial \bar{h}_i(\boldsymbol{\theta}_0)}{\partial \beta}, \frac{1}{h_i^2} \frac{\partial \bar{h}_i(\boldsymbol{\theta}_0)}{\partial \omega} \right)^\top,$$

and we approximate $\mathbf{g}_i$ by

$$\mathbf{g}_{i,j} = \left(\frac{1}{h_{i,j}^2} \frac{\partial \bar{h}_{i,j}(\boldsymbol{\theta}_0)}{\partial \alpha}, \frac{1}{h_{i,j}^2} \frac{\partial \bar{h}_{i,j}(\boldsymbol{\theta}_0)}{\partial \beta}, \frac{1}{h_{i,j}^2} \frac{\partial \bar{h}_{i,j}(\boldsymbol{\theta}_0)}{\partial \omega} \right)^\top.$$

LEMMA 6.1. *If H_0, Assumptions 2.1–2.3, and (6.1) are satisfied, then*

$$\left(E|h_i^2 - h_{i,j}^2|^{\bar{\kappa}}\right)^{1/\bar{\kappa}} \le c_1 \bar{\rho}^j, \qquad 1 \le j < \infty \tag{6.34}$$

and

$$\left(E|x_i^2 - x_{i,j}^2|^{\bar{\kappa}}\right)^{1/\bar{\kappa}} \le c_2 \bar{\rho}^j, \qquad 1 \le j < \infty, \tag{6.35}$$

where $c_1, c_2 > 0$, and $0 < \bar{\rho} < 1$, $\bar{\kappa} > 0$ are defined in (6.12). For all $\kappa_1 > 1$

$$\left(E\|\mathbf{g}_i - \mathbf{g}_{i,j}\|^{\kappa_1}\right)^{1/\kappa_1} \le c_3 \rho_1^j, \qquad 1 \le j < \infty, \tag{6.36}$$

with some $c_3 = c_3(\kappa_1)$, and $0 < \rho_1 = \rho_1(\kappa_1) < 1$.

Proof. We can assume without loss of generality that $0 < \bar{\kappa} < 1$, and therefore by Minkowski's inequality,

$$E\left(\sum_{\ell=j+1}^{\infty} \prod_{m=1}^{\ell-1} (\alpha_0 \epsilon_{i-m}^2 + \beta_0) \right)^{\bar{\kappa}} \le c_4 \bar{\rho}^j. \tag{6.37}$$

Using the representation of h_i^2 and the definition of $h_{i,j}^2$, we obtain

$$|h_i^2 - h_{i,j}^2| \le \sum_{\ell=j+1}^{\infty} \prod_{m=1}^{\ell-1} (\alpha_0 \epsilon_{i-m}^2 + \beta_0) + \omega_0 \sum_{\ell=j+1}^{\infty} \prod_{k=1}^{\ell-1} (\alpha_0 (\epsilon_{i,i-k,j}^*)^2 + \beta_0), \qquad j \ge 1,$$

and thus (6.34) follows from (6.37). Observing that

$$E|x_i^2 - x_{i,j}^2|^{\bar{\kappa}} = E|\epsilon_0|^{\bar{\kappa}} E|h_{i-1}^2 - h_{i-1,j}^2|^{\bar{\kappa}},$$

(6.35) is also proven.

It is also shown in Berkes et al. (2003) (cf. also Francq and Zakoian, 2010) that for all $\kappa_2 > 0$,

$$E\left(\frac{1}{h_i^2} \frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha} \right)^{\kappa_2} \le c_5(\kappa_2). \tag{6.38}$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Along the lines of the proof of (6.34), one can verify that there exists a $\kappa_3 > 0$ such that

$$E\left|\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha} - \frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right|^{\kappa_3} \le c_5 \rho_3^j, \quad \text{with some } 0 < \rho_3 < 1. \tag{6.39}$$

Let

$$A_{i,j} = \left\{ |h_i^2 - h_{i,j}^2| > \bar{\rho}^{j/(2\bar{\kappa})} \quad \text{or} \quad \left|\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha} - \frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right| > \rho_3^{j/(2\kappa_3)} \right\}.$$

Then, by Markov's inequality,

$$P\{A_{i,j}\} \le c_6 \max(\bar{\rho}^{j/2}, \rho_3^{j/2}). \tag{6.40}$$

Now

$$\begin{aligned}
&\left|\frac{1}{h_i^2}\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha} - \frac{1}{h_{i,j}^2}\frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right| \\
&\quad \le \left(\left|\frac{1}{h_i^2}\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right| + \left|\frac{1}{h_{i,j}^2}\frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right|\right) I\{A_{i,j}\} + \frac{1}{\omega_0}\left|\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha} - \frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right| I\{\bar{A}_{i,j}\} \\
&\qquad + \frac{1}{\omega_0}\left|\frac{1}{h_{i,j}^2}\frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right| \left|\bar{h}_i^2(\boldsymbol{\theta}_0) - \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)\right| I\{\bar{A}_{i,j}\} \\
&\quad \le \left(\left|\frac{1}{h_i^2}\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right| + \left|\frac{1}{h_{i,j}^2}\frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right|\right) I\{A_{i,j}\} + \frac{1}{\omega_0}\rho_3^{j/(2\kappa_3)} \\
&\qquad + \frac{1}{\omega_0}\left|\frac{1}{h_{i,j}^2}\frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right| \bar{\rho}^{j/(2\bar{\kappa})},
\end{aligned}$$

and therefore by Minkowski's inequality, for any $\kappa_4 \ge 2$, we have

$$\begin{aligned}
&\left(\left|\frac{1}{h_i^2}\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha} - \frac{1}{h_{i,j}^2}\frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right|^{\kappa_4}\right)^{1/\kappa_4} \le 2\left(E\left|\frac{1}{h_i^2}\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right|^{\kappa_4} I\{A_{i,j}\}\right)^{1/\kappa_4} \\
&\qquad + \frac{1}{\omega_0}\rho_3^{j/(2\kappa_3)} + \frac{1}{\omega_0}\bar{\rho}^{j/(2\bar{\kappa})}\left(E\left|\frac{1}{h_{i,j}^2}\frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right|^{\kappa_4}\right)^{\kappa_4}.
\end{aligned}$$

Since by the Cauchy–Schwartz inequality

$$E\left\{\left|\frac{1}{h_i^2}\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right|^{\kappa_4} I\{A_{i,j}\}\right\} \le \left\{E\left(\frac{1}{h_i^2}\frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha}\right)^{2\kappa_4}\right\}^{1/2} \left\{P\{A_{i,j}\}\right\}^{1/2},$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

we conclude that for all $\kappa_5 \geq 2$ there is $0 < \rho_5 < 1$ such that

$$\left(E \left| \frac{1}{h_i^2} \frac{\partial \bar{h}_i^2(\boldsymbol{\theta}_0)}{\partial \alpha} - \frac{1}{h_{i,j}^2} \frac{\partial \bar{h}_{i,j}^2(\boldsymbol{\theta}_0)}{\partial \alpha} \right|^{\kappa_5} \right)^{1/\kappa_5} \leq c_7 \rho_5^j. \tag{6.41}$$

Hence, the decomposability of the first coordinate of $\mathbf{g}_i$ with an exponential rate is proven. The same argument can be applied to the other two coordinates. □

Next, we introduce the stationary version of the likelihood function:

$$\bar{f}_i(\boldsymbol{\theta}) = \log \bar{h}_i^2(\boldsymbol{\theta}) + \frac{x_i^2}{\bar{h}_i^2(\boldsymbol{\theta})}, \qquad -\infty < i < \infty,$$

$$\bar{\ell}_k(\boldsymbol{\theta}) = \sum_{i=1}^{k} \bar{f}_i(\boldsymbol{\theta}),$$

and

$$\bar{f}_{i,j}(\boldsymbol{\theta}) = \log \bar{h}_{i,j}^2(\boldsymbol{\theta}) + \frac{x_i^2}{\bar{h}_{i,j}^2(\boldsymbol{\theta})}.$$

LEMMA 6.2. *If H_0, Assumptions 2.1–2.3, and (6.1) are satisfied, then*

$$\left(\left\| \nabla(\bar{f}_i(\boldsymbol{\theta}_0) - \bar{f}_{i,j}(\boldsymbol{\theta}_0) \right\|^{\kappa/2} \right)^{2/\kappa} \leq c\rho^j \tag{6.42}$$

and

$$\left(\left\| \nabla^2(\bar{f}_i(\boldsymbol{\theta}_0) - \bar{f}_{i,j}(\boldsymbol{\theta}_0) \right\|^{\kappa/2} \right)^{2/\kappa} \leq c\rho^j \tag{6.43}$$

with some $0 < \rho < 1$, and $\kappa > 4$ is defined in Assumption 2.1.

Proof. The result in (6.42) is an immediate consequence of Lemma 6.1 since

$$\nabla \bar{f}_i(\boldsymbol{\theta}_0) = (1 - \epsilon_i^2)\mathbf{g}_i.$$

Similar arguments give (6.43). □

Let

$$\mathbf{H} = E\nabla^2 \bar{f}_0(\boldsymbol{\theta}_0).$$

Berkes et al. (2003) and Francq and Zakoian (2004) prove that $\mathbf{H}$ is not singular.

LEMMA 6.3. *If H_0, Assumptions 2.1–2.3, and (6.1) are satisfied, then*

$$N(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) = -\mathbf{H}^{-1} \nabla \bar{\ell}_N(\boldsymbol{\theta}_0) + O_P(1).$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Proof. Berkes et al. (2003) and Francq and Zakoian (2004) show that there exists a $\boldsymbol{\Theta}_0$, a neighborhood of $\boldsymbol{\theta}_0$, such that

$$\sup\left\{\frac{1}{N}\left\|\nabla^3 \bar{\ell}_N(\boldsymbol{\theta})\right\| : \boldsymbol{\theta} \in \boldsymbol{\Theta}_0\right\} = O_P(1) \quad \text{a.s.} \tag{6.44}$$

By (6.31), we have

$$\|\nabla \mathscr{L}_N(\hat{\boldsymbol{\theta}}_N) - \nabla \bar{\ell}_N(\hat{\boldsymbol{\theta}}_N)\| = O_P(1)$$

and since

$$\nabla \bar{\mathscr{L}}_N(\hat{\boldsymbol{\theta}}_N) = \mathbf{0},$$

by a Taylor expansion, we get from (6.32) and (6.44) that

$$-\nabla \bar{\mathscr{L}}_N(\boldsymbol{\theta}_0) = \nabla \bar{\ell}_N(\hat{\boldsymbol{\theta}}_N) - \nabla \bar{\ell}_N(\boldsymbol{\theta}_0) + O_P(1) = \nabla^2 \bar{\ell}_N(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0) + O_P(1).$$

Because of (6.42) and (6.43), we can use Lemma S2.1 of Aue et al. (2014) to conclude that

$$\|\nabla \bar{\ell}_N(\boldsymbol{\theta}_0)\| = O_P(N^{1/2})$$

and

$$\|\nabla^2 \bar{\ell}_N(\boldsymbol{\theta}_0) - N\mathbf{H}\| = O_P(N^{1/2}).$$

Now the proof of Lemma 6.3 is complete. □

LEMMA 6.4. *If H_0, Assumptions 2.1–2.3, and (6.1) are satisfied, then*

$$\max_{1\le k\le N-1}\left(\frac{N}{k(N-k)}\right)^{\zeta}\left\|\nabla \bar{\mathscr{L}}_k(\hat{\boldsymbol{\theta}}_N) - \left(\nabla \bar{\ell}_k(\boldsymbol{\theta}_0) - \frac{k}{N}\nabla \bar{\ell}_N(\boldsymbol{\theta}_0)\right)\right\| = O_P(1)$$

for any $\zeta > 0$.

Proof. It is enough to prove that

$$\max_{1\le k\le N-1}\left(\frac{N}{k(N-k)}\right)^{\zeta}\left\|\nabla \bar{\ell}_k(\hat{\boldsymbol{\theta}}_N) - \left(\nabla \bar{\ell}_k(\boldsymbol{\theta}_0) - \frac{k}{N}\nabla \bar{\ell}_N(\boldsymbol{\theta}_0)\right)\right\| = O_P(1).$$

We follow the proof of Lemma 6.3. For any ζ, we choose $1/2 < \zeta_1 < 1/2 + \zeta$. Using again (6.43) and Lemma S2.1 of Aue et al. (2014), we have

$$\max_{1\le k\le N}\frac{1}{k^{\zeta_1}}\left\|\nabla^2 \bar{\ell}_k(\boldsymbol{\theta}_0) - k\mathbf{H}\right\| = O_P(1). \tag{6.45}$$

Thus, we obtain

$$\max_{1\le k\le N}\frac{1}{k^{\zeta}}\left\|(\nabla^2 \bar{\ell}_k(\boldsymbol{\theta}_0) - k\mathbf{H})(\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0)\right\| = O_P(1)\max_{1\le k\le N} k^{\zeta_1-\zeta}N^{-1/2} = O_P(1),$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

and similarly by Lemma 6.3

$$\max_{1\le k\le N}\frac{1}{k^{\zeta}}\left\|k\mathbf{H}\left[(\hat{\boldsymbol{\theta}}_N-\boldsymbol{\theta}_0)-\left(-\mathbf{H}\frac{1}{N}\nabla\bar{\mathcal{f}}_N(\boldsymbol{\theta}_0)\right)\right]\right\|=O_P(1).$$

Our estimates imply that

$$\max_{1\le k\le N}\frac{1}{k^{\zeta}}\left\|\nabla\bar{\mathcal{f}}_k(\hat{\boldsymbol{\theta}}_N)-\left(\nabla\bar{\mathcal{f}}_k(\boldsymbol{\theta}_0)-\frac{k}{N}\nabla\bar{\mathcal{f}}_N(\boldsymbol{\theta}_0)\right)\right\|=O_P(1).$$

To prove the same estimate when we use $N-k$ instead of k in the weight function, we observe, as in (6.45), that

$$\max_{1\le k\le N}\frac{1}{(N-k)^{\zeta_1}}\left\|\nabla^2\bar{\mathcal{f}}_N(\boldsymbol{\theta}_0)-\nabla^2\bar{\mathcal{f}}_k(\boldsymbol{\theta}_0)-(N-k)\mathbf{H}\right\|=OP(1) \tag{6.46}$$

for all $\zeta_1>1/2$. Using

$$\begin{aligned}&\max_{1\le k\le N_1}\Big\|\nabla(\bar{\mathcal{f}}_N(\hat{\boldsymbol{\theta}}_N)-\bar{\mathcal{f}}_k(\hat{\boldsymbol{\theta}}_N))\\&\quad-\left[\nabla(\bar{\mathcal{f}}_N(\boldsymbol{\theta}_0)-\bar{\mathcal{f}}_k(\boldsymbol{\theta}_0))+\nabla(\bar{\mathcal{f}}_N(\boldsymbol{\theta}_0)-\bar{\mathcal{f}}_k(\boldsymbol{\theta}_0))(\hat{\boldsymbol{\theta}}_N-\boldsymbol{\theta}_0)\right]\Big\|\\&=O_P(1)\end{aligned}$$

and (6.46), one can verify that

$$\max_{1\le k\le N-1}\frac{1}{(N-k)^{\zeta}}\left\|\nabla\bar{\mathcal{f}}_k(\hat{\boldsymbol{\theta}}_N)-\left(\nabla\bar{\mathcal{f}}_k(\boldsymbol{\theta}_0)-\frac{k}{N}\nabla\bar{\mathcal{f}}_N(\boldsymbol{\theta}_0)\right)\right\|=O_P(1).$$

□

The last lemma of this section is an immediate consequence of Lemma 6.2 and the approximation in Aue et al. (2014).

LEMMA 6.5. *If H_0, Assumptions 2.1–2.3, and (6.1) are satisfied, then we can define independent Gaussian processes $\{\mathbf{W}_{N,1}(x),0\le x\le N/2\}$ and $\{\mathbf{W}_{N,2}(x),0\le x\le N/2\}$ such that*

$$\max_{1\le x\le N/2}\frac{1}{x^{\nu}}\left\|\sum_{i=1}^{x}\left(\frac{\partial\bar{f}_i(\boldsymbol{\theta}_0)}{\partial\alpha},\frac{\partial\bar{f}_i(\boldsymbol{\theta}_0)}{\partial\beta}\right)^{\top}-\mathbf{W}_{N,1}(x)\right\|=O_P(1),$$

and

$$\max_{N/2\le x\le N-1}\frac{1}{(N-x)^{\nu}}\left\|\sum_{i=x+1}^{N}\left(\frac{\partial\bar{f}_i(\boldsymbol{\theta}_0)}{\partial\alpha},\frac{\partial\bar{f}_i(\boldsymbol{\theta}_0)}{\partial\beta}\right)^{\top}-\mathbf{W}_{N,2}(N-x)\right\|=O_P(1),$$

with some $\nu<1/2$, $E\mathbf{W}_{N,1}(x)=E\mathbf{W}_{N,2}(x)=\mathbf{0}$, and $E\mathbf{W}_{N,1}(x)\mathbf{W}_{N,1}^{\top}(y)=E\mathbf{W}_{N,2}(x)\mathbf{W}_{N,2}^{\top}(y)=\min(x,y)\mathbf{D}$.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

6.4. Proof of (6.16) When $E\log(\alpha_0\epsilon_0^2+\beta_0)\geq 0$

We note that Jensen and Rahbek (2004) prove their results under assumption (6.3). However, Theorem 6.1 implies that their arguments are also valid when (6.2) holds. Hence, in the section, we assume that

$$E\log(\alpha_0\epsilon_0^2+\beta_0)\geq 0. \tag{6.47}$$

In Section 6.3, we showed that it is enough to consider the processes based on the stationary variables x_i and h_i^2. There is no stationary solution in the present case, so we need to work directly with $\bar\sigma_i^2(\boldsymbol\theta)$ and $\bar\ell_i(\boldsymbol\theta)$. However, following the method of Jensen and Rahbek (2004), we approximate $\nabla\bar\ell_i$ with a stationary sequence. Elementary arguments lead to

$$\frac{\partial\bar\ell_i(\boldsymbol\theta)}{\partial\alpha}=\left[1-\frac{y_i^2}{\bar\sigma_i^2(\boldsymbol\theta)}\right]\bar p_{i,1}(\boldsymbol\theta),\qquad \frac{\partial\bar\ell_i(\boldsymbol\theta)}{\partial\beta}=\left[1-\frac{y_i^2}{\bar\sigma_i^2(\boldsymbol\theta)}\right]\bar p_{i,2}(\boldsymbol\theta),$$

where

$$\bar p_{i,1}(\boldsymbol\theta)=\frac{1}{\bar\sigma_i^2(\boldsymbol\theta)}\frac{\partial\bar\sigma_i^2(\boldsymbol\theta)}{\partial\alpha},\quad \bar p_{i,2}(\boldsymbol\theta)=\frac{1}{\bar\sigma_i^2(\boldsymbol\theta)}\frac{\partial\bar\sigma_i^2(\boldsymbol\theta)}{\partial\beta}.$$

Since we can estimate only α_0 and β_0, we use a special notation for the estimable parameters, $\bar{\boldsymbol\theta}_0=(\alpha_0,\beta_0)^\top$. Let

$$z_{i,1}=\sum_{j=1}^{\infty}\epsilon_{i-j}^2\frac{1}{\beta_0}\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\quad\text{and}$$

LEMMA 6.6. *If H_0, Assumptions 2.1–2.3, and (6.47) are satisfied, then*

$$\sup_{1/\bar\omega\leq\omega\leq\bar\omega}\left|\bar p_{i,1}(\bar{\boldsymbol\theta}_0,\omega)-z_{i,1}\right|=O(\rho^i)\quad a.s. \tag{6.48}$$

and

$$\sup_{1/\bar\omega\leq\omega\leq\bar\omega}\left|\bar p_{i,2}(\bar{\boldsymbol\theta}_0,\omega)-z_{i,2}\right|=O(\rho^i)\quad a.s. \tag{6.49}$$

with some $0<\rho<1$.

Proof. First, we note that

$$\bar\sigma_i^2(\boldsymbol\theta)=\omega\sum_{k=0}^{i-1}\beta^k+\alpha\sum_{\ell=1}^{i}\beta^{\ell-1}y_{i-\ell}^2+\beta^i\sigma_0^2,$$

which gives the representations

$$\bar\sigma_i^2(\boldsymbol\theta_0)=\omega_0\sum_{k=0}^{i-1}\beta_0^k+\alpha_0\sum_{\ell=1}^{i}\beta_0^{\ell-1}y_{i-\ell}^2+\beta_0^i\sigma_0^2$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

and

$$\bar{\sigma}_i^2(\bar{\boldsymbol{\theta}}_0,\omega)=\omega\sum_{k=0}^{i-1}\beta_0^k+\alpha_0\sum_{\ell=1}^{i}\beta_0^{\ell-1}y_{i-\ell}^2+\beta_0^i\sigma_0^2.$$

We get, by Theorem 6.1(ii) and (iii), that there exist $c_1>0, c_2>0$ such that

$$\exp(-c_1 i^{c_2})\max_{1/\bar{\omega}\le\omega\le\bar{\omega}}|\bar{\sigma}_i^2(\bar{\boldsymbol{\theta}}_0,\omega)-\bar{\sigma}_i^2(\boldsymbol{\theta}_0)|=o(1)\ \text{ a.s.} \tag{6.50}$$

and

$$\max_{1/\bar{\omega}\le\omega\le\bar{\omega}}\left|\frac{\bar{\sigma}_i^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\bar{\sigma}_i^2(\boldsymbol{\theta}_0)}-1\right|=O(\rho_1^i) \tag{6.51}$$

with some $0<\rho_1<1$. Using the recursion, we have

$$\bar{p}_{i,1}(\boldsymbol{\theta})=\frac{1}{\bar{\sigma}_i^2(\boldsymbol{\theta})}\sum_{j=1}^{i}\beta^{j-1}y_{i-j}^2=\sum_{j=1}^{i}\beta^{j-1}\frac{y_{i-j}^2}{\bar{\sigma}_{i-j}^2(\boldsymbol{\theta})}\prod_{k=1}^{j}\frac{\bar{\sigma}_{i-k}^2(\boldsymbol{\theta})}{\bar{\sigma}_{i-k+1}^2(\boldsymbol{\theta})}$$

and therefore,

$$\bar{p}_{i,1}(\bar{\boldsymbol{\theta}}_0,\omega)=\sum_{j=1}^{i}\frac{1}{\beta_0}\frac{y_{i-j}^2}{\bar{\sigma}_{i-j}^2(\bar{\boldsymbol{\theta}}_0,\omega)}\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\bar{\sigma}_{i-k+1}^2(\bar{\boldsymbol{\theta}}_0,\omega)}.$$

Since

$$\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\bar{\sigma}_{i-k+1}^2(\bar{\boldsymbol{\theta}}_0,\omega)}=\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\omega+\alpha_0 y_{i-k}^2+\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}<\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}<1,$$

we obtain

$$\begin{aligned}&\sum_{j=1}^{i/2}\beta_0^{j-1}\left|\frac{y_{i-j}^2}{\bar{\sigma}_{i-j}^2(\bar{\boldsymbol{\theta}}_0,\omega)}-\epsilon_{i-j}^2\right|\prod_{k=1}^{j}\frac{\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\bar{\sigma}_{i-k+1}^2(\bar{\boldsymbol{\theta}}_0,\omega)}\\&\quad\le\frac{1}{\beta_0}\sum_{j=1}^{i/2}\epsilon_{i-j}^2\left|\frac{\sigma_{i-j}^2}{\bar{\sigma}_{i-j}^2(\bar{\boldsymbol{\theta}}_0,\omega)}-1\right|=O(\rho_2^i)\quad\text{a.s.}\end{aligned}$$

on account of (6.51) and

$$\max_{1\le i<\infty}i^{-1/\kappa}\epsilon_i^2=O_P(1)\quad\text{a.s.} \tag{6.52}$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

By applying Minkowski's inequality, for any $\kappa_1 > 2$, we have

$$\begin{aligned}\left(E\left(\sum_{j=i/2}^{i}\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\bar{\sigma}_{i-k+1}^2(\bar{\boldsymbol{\theta}}_0,\omega)}\right)^{\kappa_1}\right)^{1/\kappa_1} &\le \left(E\sum_{j=i/2}^{i}\left(\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right)^{\kappa_1}\right)^{1/\kappa_1} \\ &\le \sum_{j=i/2}^{i}\left(E\left(\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right)^{\kappa_1}\right)^{1/\kappa_1} \\ &= \sum_{j=i/2}^{i}\rho_3^j.\end{aligned} \tag{6.53}$$

Hence, the Borel–Cantelli lemma implies that

$$\sum_{j=i/2}^{i}\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\bar{\sigma}_{i-k+1}^2(\bar{\boldsymbol{\theta}}_0,\omega)} = O(\rho_4^i) \quad \text{a.s.}$$

It follows from the recursion and (6.51) that

$$\frac{y_{i-j}^2}{\bar{\sigma}_{i-j}^2(\bar{\boldsymbol{\theta}}_0,\omega)} = \epsilon_{i-j}^2\frac{\sigma_{i-j}^2}{\bar{\sigma}_{i-j}^2(\bar{\boldsymbol{\theta}}_0,\omega)} \le \epsilon_{i-j}\sup_{1\le\ell<\infty}\sup_{1/\bar{\omega}\le\omega\le\bar{\omega}}\frac{\sigma_\ell^2}{\bar{\sigma}_\ell^2(\bar{\boldsymbol{\theta}}_0,\omega)}$$

and

$$\sup_{1\le\ell<\infty}\sup_{1/\bar{\omega}\le\omega\le\bar{\omega}}\frac{\sigma_\ell^2}{\bar{\sigma}_\ell^2(\bar{\boldsymbol{\theta}}_0,\omega)} = O(1) \quad \text{a.s.}$$

Thus, we conclude that

$$\sup_{1/\bar{\omega}\le\omega\le\bar{\omega}}\left|p_{i,1}(\bar{\theta}_0,\omega) - \sum_{j=1}^{i}\frac{1}{\beta_0}\epsilon_{i-j}^2\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\bar{\sigma}_{i-k+1}^2(\bar{\boldsymbol{\theta}}_0,\omega)}\right| = O(\rho_5^i) \quad \text{a.s.} \tag{6.54}$$

Elementary argument leads to

$$\begin{aligned}&\left|\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\omega+(\alpha_0\epsilon_{i-k}^2+\beta_0)\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)} - \prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right| \\ &\quad\le \sum_{k=1}^{j}\left|\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\omega+(\alpha_0\epsilon_{i-k}^2+\beta_0)\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)} - \frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right| \\ &\quad\le \beta_0\omega\sum_{k=1}^{j}\frac{1}{\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}.\end{aligned} \tag{6.55}$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Using (6.50) and (6.52), we get that

$$\sup_{1/\tilde{\omega}\le\omega\le\tilde{\omega}}\sum_{j=1}^{i/2}\epsilon_{i-j}^2\left|\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\omega+(\alpha_0\epsilon_{i-k}^2+\beta_0)\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}-\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right| \tag{6.56}$$

$$\overset{\text{a.s.}}{=} O(1)\sum_{j=1}^{i/2}\epsilon_{i-j}^2\sum_{k=1}^{j}\exp(-S(i-k))$$

$$\overset{\text{a.s.}}{=} O(\rho_6^i),$$

with some $0<\rho_6<1$. Similarly to (6.53)

$$\sum_{j=i/2}^{i}\epsilon_{i-j}^2\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\overset{\text{a.s.}}{=} O(\rho_7^i),$$

and

$$\sup_{1/\tilde{\omega}\le\omega\le\tilde{\omega}}\sum_{j=i/2}^{i}\epsilon_{i-j}^2\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\omega+(\alpha_0\epsilon_{i-k}^2+\beta_0)\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}\overset{\text{a.s.}}{=} O(\rho_7^i),$$

with $0<\rho_7<1$, completing the proof of (6.48). Since

$$\bar{p}_{i,2}(\bar{\boldsymbol{\theta}}_0,\omega)=\sum_{j=1}^{i}\frac{1}{\beta_0}\prod_{k=1}^{j}\frac{\beta_0\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)}{\omega_0+(\alpha_0\epsilon_{i-k}^2+\beta_0)\bar{\sigma}_{i-k}^2(\bar{\boldsymbol{\theta}}_0,\omega)},$$

the proof of (6.49) is nearly the same but simpler than that of (6.49). □

Let

$$z_{i,j,1}=\sum_{\ell=1}^{j}\epsilon_{i-\ell}^2\frac{1}{\beta_0}\prod_{k=1}^{\ell}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}+\sum_{\ell=j+1}^{\infty}\epsilon_{i,i-\ell,j}^2\frac{1}{\beta_0}$$
$$\times\left(\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right)\left(\prod_{r=j+1}^{\ell}\frac{\beta_0}{\alpha_0\epsilon_{i,i-r,j}^2+\beta_0}\right),$$

and

$$z_{i,j,2}=\sum_{\ell=1}^{j}\frac{1}{\beta_0}\prod_{k=1}^{\ell}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}+\sum_{\ell=j+1}^{\infty}\frac{1}{\beta_0}\left(\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right)\left(\prod_{r=j+1}^{\ell}\frac{\beta_0}{\alpha_0\epsilon_{i,i-r,j}^2+\beta_0}\right),$$

where $\{\epsilon_{i,j,k},-\infty<i,j,k<\infty\}$ are independent copies of ϵ_0.

LEMMA 6.7. *If H_0, Assumptions 2.1–2.3, and (6.47) are satisfied, then*

$$\left(E\left[(1-\epsilon_i^2)(z_{i,1}-z_{i,j,1})\right]^\kappa\right)^{1/\kappa}\le c\rho^j \tag{6.57}$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

and

$$\left(E\left[(1-\epsilon_i^2)(z_{i,2}-z_{i,j,2})\right]^\kappa\right)^{1/\kappa} \le c\rho^j, \tag{6.58}$$

where $0<\rho<1$, *and* κ *is defined in Assumption 2.1.*

Proof. By Assumption 2.1 and Minkowski's inequality, we have

$$\left(E\left[(1-\epsilon_i^2)\sum_{\ell=j+1}^{\infty}\epsilon_{i,i-\ell,j}^2\frac{1}{\beta_0}\left(\prod_{k=1}^{j}\frac{\beta_0}{\alpha_0\epsilon_{i-k}^2+\beta_0}\right)\left(\prod_{r=j+1}^{\ell}\frac{\beta_0}{\alpha_0\epsilon_{i,i-r,j}^2+\beta_0}\right)\right]^\kappa\right)^{1/\kappa}$$
$$\le \left(E(1-\epsilon_0^2)^\kappa\right)\frac{1}{\beta_0}\sum_{\ell=j+1}^{\infty}\left(E\left[\frac{\beta_0}{\alpha_0\epsilon^2+\beta_0}\right]^\kappa\right)^{1/\kappa} \le c_1\rho_1^j$$

with some $0<\rho_1<1$, proving (6.57). The same argument gives (6.58). □

Let

$$\bar{\mathbf{J}}_i(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial^2\bar{\ell}_i(\boldsymbol{\theta})}{\partial\alpha^2}, & \frac{\partial^2\bar{\ell}_i(\boldsymbol{\theta})}{\partial\alpha\partial\beta} \\ \frac{\partial^2\bar{\ell}_i(\boldsymbol{\theta})}{\partial\alpha\partial\beta}, & \frac{\partial^2\bar{\ell}_i(\boldsymbol{\theta})}{\partial\beta^2} \end{pmatrix}$$

be the matrix of the second-order partial derivatives of $\bar{\ell}_i(\boldsymbol{\theta})$. Similarly to Lemma 6.6, it can be approximated with a matrix $\mathbf{Q}_i$ of the form $\mathbf{Q}_i = \mathbf{q}(\epsilon_i,\epsilon_{i-1},\dots)$. The sequence $\mathbf{Q}_i$ is a decomposable Bernoulli shift with $\mathbf{Q}_{i,j} = \mathbf{q}(\epsilon_i,\epsilon_{i-1},\dots,\epsilon_{i-1},\epsilon_{i,j-i-1,j},\epsilon_{i,j-i-2,j},\dots)$.

LEMMA 6.8. *If* H_0, *Assumptions 2.1–2.3, and (6.47) are satisfied, then*

$$\sup_{1/\bar{\omega}\le\omega\le\bar{\omega}}\left\|\bar{\mathbf{J}}_i(\bar{\boldsymbol{\theta}}_0,\omega)-\mathbf{Q}_i\right\| = O(\rho^i) \quad a.s. \tag{6.59}$$

and

$$\left(E\left\|\mathbf{Q}_i-\mathbf{Q}_{i,j}\right\|^\kappa\right)^{1/\kappa} \le c\rho^j, \tag{6.60}$$

where $0<\rho<1$, *and* κ *is defined in Assumption 2.1.*

Proof. The matrix $\bar{\mathbf{J}}_i$ can be computed explicitly, so following the proofs of Lemmas 6.6 and 6.7, we obtain (6.59) and (6.60), respectively. □

The rest of the proof is a repetition of the calculations in Section 6.3, since we have showed again that the partial derivatives can be approximated with decomposable Bernoulli shifts with exponential rates. Thus, we get the following analogue of Lemma 6.4.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

LEMMA 6.9. *If H_0, Assumptions 2.1–2.3, and (6.47) are satisfied, then*

$$\max_{1\le k\le N-1}\left(\frac{N}{k(N-k)}\right)^{\zeta}\left\|\bar{r}_k(\hat{\boldsymbol{\theta}}_N)\right.$$
$$\left.-\left(\sum_{i=1}^{k}(1-\epsilon_i^2)(z_{i,1},z_{i,2})^\top-\frac{k}{N}\sum_{i=1}^{N}(1-\epsilon_i^2)(z_{i,1},z_{i,2})^\top\right)\right\|$$
$$=O_P(1)$$

for any $\zeta>0$.

LEMMA 6.10. *If H_0, Assumptions 2.1–2.3, and (6.47) are satisfied, then we can define independent Gaussian processes $\{\mathbf{W}_{N,1}(x),0\le x\le N/2\}$ and $\{\mathbf{W}_{N,2}(x),0\le x\le N/2\}$ such that*

$$\max_{1\le x\le N/2}\frac{1}{x^\nu}\left\|\sum_{i=1}^{\lfloor x\rfloor}(1-\epsilon_i^2)(z_{i,1},z_{i,2})^\top-\mathbf{W}_{N,1}(x)\right\|=O_P(1),$$

and

$$\max_{N/2\le x\le N-1}\frac{1}{(N-x)^\nu}\left\|\sum_{i=\lfloor x\rfloor+1}^{N}(1-\epsilon_i^2)(z_{i,1},z_{i,2})^\top-\mathbf{W}_{N,2}(N-x)\right\|=O_P(1)$$

with some $\nu<1/2$, $E\mathbf{W}_{N,1}(x)=E\mathbf{W}_{N,2}(x)=\mathbf{0}$, and $E\mathbf{W}_{N,1}(x)\mathbf{W}_{N,1}^\top(y)=E\mathbf{W}_{N,2}(x)\mathbf{W}_{N,2}^\top(y)=\min(x,y)E(1-\epsilon_0^2)^2E[(z_{i,1},z_{i,2})^\top(z_{i,1},z_{i,2})^\top]$.

Proof. The approximations are immediate consequences of Lemma 6.7 and Theorem S2.1 of Aue et al. (2014). □

Appendix

To facilitate the application of the test developed in this work, we provide the following step-by-step procedure based on Theorem 2.1(i) to calculate the test statistic T_N.

1. Use the QMLE to estimate the parameters $\boldsymbol{\theta}$ and the corresponding estimator $\hat{\boldsymbol{\theta}}_N=(\hat{\omega},\hat{\alpha},\hat{\beta})^\top$ that satisfies

   $$\inf\left\{\bar{\mathscr{L}}_N(\boldsymbol{\theta}):\ \boldsymbol{\theta}\in\boldsymbol{\Theta}\right\}=\bar{\mathscr{L}}_N(\hat{\boldsymbol{\theta}}_N),$$

   where

   $$\bar{\mathscr{L}}_k(\boldsymbol{\theta})=\sum_{i=1}^{k}\bar{\ell}_i(\boldsymbol{\theta}),\qquad\text{and}\qquad\bar{\ell}_i(\boldsymbol{\theta})=\log\bar{\sigma}_i^2(\boldsymbol{\theta})+\frac{y_i^2}{\bar{\sigma}_i^2(\boldsymbol{\theta})}.$$

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

2. Calculate the derivatives of the log likelihood function

$$\frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha} = \frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)} \frac{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha} = \frac{1}{\bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)} \left(1 - \frac{y_i^2}{\bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}\right) \frac{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha},$$

$$\frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} = \frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)} \frac{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} = \frac{1}{\bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)} \left(1 - \frac{y_i^2}{\bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}\right) \frac{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}{\partial \beta}.$$

To implement the above calculation, the following recursive equations are required:

$$\frac{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha} = y_{i-1}^2 + \beta \frac{\partial \bar{\sigma}_{i-1}^2(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha},$$

$$\frac{\partial \bar{\sigma}_i^2(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} = \bar{\sigma}_{i-1}^2(\hat{\boldsymbol{\theta}}_N) + \beta \frac{\partial \bar{\sigma}_{i-1}^2(\hat{\boldsymbol{\theta}}_N)}{\partial \beta},$$

and their initial values are set to be $\frac{\partial \bar{\sigma}_0^2(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha} = \frac{\partial \bar{\sigma}_0^2(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} = 0$, following Fiorentini, Calzolari, and Panattoni (1996).

3. Calculate the following two quantities:

$$\bar{\boldsymbol{r}}_k(\hat{\boldsymbol{\theta}}_N) = \sum_{i=1}^{k} \left(\frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha}, \frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} \right),$$

$$\hat{\boldsymbol{D}}_N = \frac{1}{N} \sum_{i=i}^{N} \left(\frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha}, \frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} \right)^{\top} \left(\frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \alpha}, \frac{\partial \bar{\ell}_i(\hat{\boldsymbol{\theta}}_N)}{\partial \beta} \right).$$

4. Then, compute the stochastic process

$$Z_N(t) = \frac{1}{N^{1/2}} \left\{ \bar{\boldsymbol{r}}_{\lfloor (N+1)t \rfloor}(\hat{\boldsymbol{\theta}}_N) \hat{\boldsymbol{D}}_N^{-1} \left(\bar{\boldsymbol{r}}_{\lfloor (N+1)t \rfloor}(\hat{\boldsymbol{\theta}}_N) \right)^{\top} \right\}, \qquad 0 < t < 1.$$

5. Obtain the test statistic

$$T_N = \sup_{0<t<1} \frac{Z_N(t)}{w(t)},$$

where $w(t)$ is a weight function.

REFERENCES

Andreou, E., & Ghysels, E. (2002). Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, 17(5), 579–600.

Andreou, E., & Ghysels, E. (2006). Monitoring disruptions in financial markets. *Journal of Econometrics*, 135(1–2), 77–124.

Aue, A., Hörmann, S., Horváth, L., & Hušková, M. (2014). Dependent functional linear models with applications to monitoring structural change. *Statistica Sinica*, 24, 1043–1073.

Berkes, I., Gombay, E., Horváth, L., & Kokoszka, P. (2004). Sequential change-point detection in GARCH(p, q) models. *Econometric Theory*, 20(6), 1140–1167.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Berkes, I., & Horváth, L. (2004). The efficiency of the estimators of the parameters in GARCH processes. *Annals of Statistics*, 32(2), 633–655.
Berkes, I., Horváth, L., & Kokoszka, P. (2003). GARCH processes: Structure and estimation. *Bernoulli*, 9(2), 201–227.
Berkes, I., Horváth, L., & Kokoszka, P. (2004). Testing for parameter constancy in GARCH(p, q) models. *Statistics & Probability Letters*, 70(4), 263–273.
Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
Chan, N. H., & Ng, C. T. (2009). Statistical inference for non-stationary GARCH (p, q) models. *Electronic Journal of Statistics*, 3, 956–992.
Chen, B., & Hong, Y. (2016). Detecting for smooth structural changes in GARCH models. *Econometric Theory*, 32(3), 740–791.
Chibisov, D. M. (1964). Some theorems on the limiting behaviour of an empirical distribution function. *Selected Translations in Mathematical Statistics and Probability*, 9, 147–156.
Chu, C.-S. J. (1995). Detecting parameter shift in GARCH models. *Econometric Reviews*, 14(2), 241–266.
Csörgő, M., & Horváth, L. (1993). *Weighted approximations in probability and statistics*. Wiley.
Csörgo, M., & Révész, P. (1981). *Strong approximations in probability and statistics*. Academic Press and Akadémiai Kiadó.
Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987–1007.
Fiorentini, G., Calzolari, G., & Panattoni, L. (1996). Analytic derivatives and the computation of GARCH estimates. *Journal of Applied Econometrics*, 11(4), 399–417.
Francq, C., & Zakoian, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4), 605–637.
Francq, C., & Zakoian, J.-M. (2007). Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications*, 117(9), 1265–1284.
Francq, C., & Zakoian, J.-M. (2010). *GARCH models: Structure statistical inference and financial applications*. Wiley.
Francq, C., & Zakoïan, J.-M. (2012). Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models. *Econometrica*, 80(2), 821–861.
Francq, C., & Zakoïan, J.-M. (2013). Inference in nonstationary asymmetric GARCH models. *Annals of Statistics*, 41(4), 1970–1998.
Francq, C., & Zakoian, J.-M. (2019). *GARCH models: Structure, statistical inference and financial applications*. John Wiley & Sons.
Galeano, P., & Tsay, R. S. (2010). Shifts in individual parameters of a GARCH model. *Journal of Financial Econometrics*, 8(1), 122–153.
Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35, 705–730.
Hoga, Y., & Demetrescu, M. (2023). Monitoring value-at-risk and expected shortfall forecasts. *Management Science*, 69(5), 2954–2971.
Huang, D., Wang, H., & Yao, Q. (2008). Estimating GARCH models: When to use what? *The Econometrics Journal*, 11(1), 27–38.
Iglesias, E. M., & Linton, O. B. (2007). Higher order asymptotic theory when a parameter is on a boundary with an application to GARCH models. *Econometric Theory*, 23(6), 1136–1161.
Jensen, S. T., & Rahbek, A. (2004). Asymptotic inference for nonstationary GARCH. *Econometric Theory*, 20(6), 1203–1226.
Kokoszka, P., & Leipus, R. (1999). Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal*, 39, 182–195.
Kokoszka, P., & Leipus, R. (2000). Change-point estimation in ARCH models. *Bernoulli*, 6, 513–539.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press

Lazar, E., Wang, S., & Xue, X. (2023). Loss function-based change point detection in risk measures. *European Journal of Operational Research*, 310(1), 415–431.

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

Nelson, D. B. (1990). Stationarity and persistence in the GARCH (1, 1) model. *Econometric Theory*, 6(3), 318–334.

O'Reilly, N. E. (1974). On the weak convergence of empirical processes in sup-norm metrics. *Annals of Probability*, 2, 642–651.

Pape, K., Wied, D., & Galeano, P. (2016). Monitoring multivariate variance changes. *Journal of Empirical Finance*, 39, 54–68.

Patton, A. (2013). Copula methods for forecasting multivariate time series. *Handbook of Economic Forecasting*, 2, 899–960.

Richter, S., Wang, W., & Wu, W. B. (2023). Testing for parameter change epochs in GARCH time series. *The Econometrics Journal*, 26, 467–491.

Straumann, D., & Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *Annals of Statistics*, 34, 2449–2495.

Xu, K.-L. (2013). Powerful tests for structural changes in volatility. *Journal of Econometrics*, 173(1), 126–142.

https://doi.org/10.1017/S026646662510011X Published online by Cambridge University Press