Econometrics and Statistics 25 (2023) 23–38

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Feasible Panel GARCH Models: Variance-Targeting Estimation and Empirical Application

Manabu Asai

Faculty of Economics, Soka University, Hachioji, Tokyo 192-8577, Japan

ARTICLE INFO

Article history:
Received 23 June 2021
Revised 17 January 2022
Accepted 18 January 2022
Available online 26 January 2022

Keywords:
Panel data
Multivariate GARCH
Fixed effects
Consistency
Asymptotic normality

ABSTRACT

For the panel generalized autoregressive conditional heteroskedasticity (GARCH) model, the conditions for the stationarity and positive definiteness of conditional covariance processes are examined. A new feasible specification is constructed for the class of panel GARCH models, and a three-step estimation technique is developed based on a variance-targeting (VT) approach. The consistency and asymptotic normality of the VT estimator are shown when the time dimension tends to infinity and the cross-sectional dimension is fixed. The stationarity and asymptotic properties are discussed for both time and cross-sectional dimensions tend to infinity. The results of Monte Carlo experiments indicate that the finite sample property of the VT estimator is satisfactory, implying that increasing the cross-sectional dimension does not affect the speed of convergence, but shrinks the asymptotic covariance matrix. The empirical results of the analysis of the inflation rates of G7 countries and growth rates for the value of trade in four economic regions indicate that the feasible specification provides a competitive alternative to the class of panel GARCH models. The empirical results indicate that the global financial crisis affects the growth rates of trades, while the influence of the COVID-19 pandemic shows that its effect on inflation rates is insignificant.

1. Introduction

Interest in allowing heteroskedastic disturbance in panel data regression is ongoing. See Arellano (1987, 2003); Cermeño and Grier (2006); Hansen (2007); Kiefer (1980); Pakel et al. (2011); Stock and Watson (2008), and Vogelsang (2012) among others. Arellano (1987); Kiefer (1980), and Hansen (2007) consider heteroskedasticity-robust standard errors as extensions of the work of White (1980), while Arellano (2003); Stock and Watson (2008), and Vogelsang (2012) use variants of the covariance estimator of Newey and West (1987) to allow heteroskedasticity and autocorrelations. Instead of an unknown form of the heteroskedasticity assumption, Cermeño and Grier (2006) develop a panel data model allowing the generalized autoregressive conditional heteroskedasticity (GARCH) structure. The empirical results of Bouras et al. (2019); Cermeño and Sanin (2015); Deniz et al. (2020); Lee (2010); Ribeiro et al. (2017), and Valera et al. (2017) demonstrate the usefulness of this class of panel GARCH model.

For their panel GARCH model, Cermeño and Grier (2006) consider maximum likelihood estimation for large T with fixed n, where T and n are the time and cross-sectional dimensions, respectively. However, there are several gaps in the theoretical literature. First, no conditions for the stationarity and positive definiteness of the covariance process are yet available.

E-mail address: m-asai@soka.ac.jp

https://doi.org/10.1016/j.ecosta.2022.01.004

Second, we need to verify the conditions for the consistency and asymptotic normality of the estimators. Third, it is important to relax the assumption of Gaussian disturbance to accommodate quasi-maximum likelihood (QML) estimation. Fourth, it is useful to apply the recent works on variance-targeting (VT) estimation developed by Francq et al. (2011); Kristensen and Linton (2004), and Pedersen and Rahbek (2014) to the panel GARCH model.

The first purpose of this study is to derive the conditions for stationarity and positive definiteness. We show that the specification of Cermeño and Grier (2006) is infeasible and/or impractical. The second purpose is to develop a new feasible panel GARCH model that satisfies the conditions for stationarity and positive definiteness as well as show the asymptotic and finite sample properties of the VT estimators. The third purpose is to report the results of an empirical analysis. Once the stationarity condition is obtained, we can apply the works of Bollerslev and Wooldridge (1992) and Hafner and Preminger (2009) to obtain the asymptotic results for the QML estimator. Hence, we concentrate on VT estimation.

The organization of the remainder of the paper is as follows. Section 2 derives the stationarity condition for the panel GARCH model to develop a new feasible specification and discusses how stationarity under n tends to infinity. Section 3 shows the asymptotic properties of the VT estimator when T tends to infinity and n is fixed. Furthermore, this section investigates the finite sample properties of the VT estimator. Section 4 presents an empirical analysis of the inflation rates of G7 countries and growth rates for the value of trade in four economic regions. Section 5 concludes the paper.

We use the following notation throughout the paper. For the product of two matrices, $\circ$ and $\otimes$ denote the Hadamard and Kronecker products, respectively. The Frobenius norm of the matrix, or vector A, is defined as $||A|| = \sqrt{\mathrm{tr}(A'A)}$. For a positive matrix A, we define the square root, $A^{1/2}$, by the spectral decomposition of A.

2. Feasible Panel GARCH Models

2.1. Panel GARCH Model

We explain the panel GARCH model suggested by Cermeño and Grier (2006). By checking the conditions for the positive definiteness of the conditional covariance matrices, we show that their specification is infeasible and/or impractical.

Consider the following dynamic panel data model with fixed effects:

$$y_{it} = \alpha_i + \phi y_{i,t-1} + \boldsymbol{x}_{it}\boldsymbol{\beta} + u_{it}, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \tag{1}$$

where $\boldsymbol{x}_{it}$ is a $1 \times k$ vector of strictly exogenous variables, α_i is an individual effect, ϕ is an AR parameter, and $\boldsymbol{\beta}$ is a $k \times 1$ vector of parameters. Assume the disturbance term u_{it} has zero mean and the following conditional moments:

$$E(u_{it}u_{js}|\Im_{t-1}) = \begin{cases} h_{iit} & \text{for } i = j \text{ and } t = s \\ h_{ijt} & \text{for } i \neq j \text{ and } t = s \\ 0 & \text{for } i = j \text{ and } t \neq s \\ 0 & \text{for } i \neq j \text{ and } t \neq s, \end{cases} \tag{2}$$

where

$$\begin{aligned} h_{iit} &= \kappa_{ii} + \gamma u_{i,t-1}^2 + \varphi h_{ii,t-1}, \\ h_{ijt} &= \kappa_{ij} + \rho u_{i,t-1}u_{j,t-1} + \eta h_{ij,t-1}, \end{aligned} \tag{3}$$

and $\Im_{t-1} = \sigma(y_{1,t-1}, \dots, y_{n,t-1}, y_{1,t-2}, \dots)$ is the sigma field generated by the past information of $(y_{1t}, \dots, y_{nt})$. We assume that $\gamma \neq 0$ and $\rho \neq 0$ to ensure a dynamic covariance structure. The specification on u_{it} does not allow autocorrelation and non-contemporaneous cross-sectional correlation. The former is redundant as the mean equation can be extended to accommodate higher autocorrelation, while the latter is a standard assumption. In addition, the dynamic panel model above allows for a time-varying covariance process. The specification is convenient in a panel data context because the number of parameters is considerably reduced by imposing common dynamics on each of them.

We refer to Eqs. (1)–(3) as the "panel GARCH" model. We can generalize the mean and variance-covariance equations with Pth order autoregression and the (p, q)th order GARCH specification, respectively.

2.2. Feasible Specification

To analyze the time-varying covariance structure, it is convenient to present an alternative form of Eqs. (1)–(3) as

$$\boldsymbol{y}_t = \boldsymbol{\alpha} + Z_t\boldsymbol{\theta} + \boldsymbol{u}_t, \tag{4}$$

$$\boldsymbol{u}_t = H_t^{1/2}\boldsymbol{\varepsilon}_t, \quad E(\boldsymbol{\varepsilon}_t) = \mathbf{0}, \quad E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = I_n, \quad E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s') = O \ (t \neq s), \tag{5}$$

$$H_t = K + C \circ \boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}' + D \circ H_{t-1}, \tag{6}$$

where $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{nt})'$, $\boldsymbol{u}_t = (u_{1t}, \ldots, u_{nt})'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$,

$$Z_t = \begin{pmatrix} \boldsymbol{z}_{1t} \\ \boldsymbol{z}_{2t} \\ \vdots \\ \boldsymbol{z}_{nt} \end{pmatrix}, \quad H_t = \begin{pmatrix} h_{11t} & h_{12t} & \cdots & h_{1nt} \\ h_{21t} & h_{22t} & \cdots & h_{2nt} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1t} & h_{n2t} & \cdots & h_{nnt} \end{pmatrix},$$

with $\boldsymbol{z}_{it} = (y_{i,t-1}\ \boldsymbol{x}_{it})$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)'$, $\boldsymbol{\theta} = (\phi, \boldsymbol{\beta}')'$,

$$K = \begin{pmatrix} \kappa_{11} & \kappa_{12} & \cdots & \kappa_{1n} \\ \kappa_{21} & \kappa_{22} & \cdots & \kappa_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \kappa_{n1} & \kappa_{n2} & \cdots & \kappa_{nn} \end{pmatrix}, \quad C = \rho J_n + (\gamma - \rho) I_n, \quad D = \eta J_n + (\varphi - \eta) I_n, \tag{7}$$

and J_n defined by an $n \times n$ matrix of ones. To guarantee positive definiteness for all H_t, we make the following proposition.

Proposition 1. *Suppose that H_0 and K are positive definite matrices. If the parameters $(\gamma, \varphi, \rho, \eta)$ satisfy the conditions:*

$$\gamma \geq \rho, \quad \gamma + \rho(n-1) \geq 0, \quad \varphi \geq \eta, \quad \varphi + \eta(n-1) \geq 0,$$

then H_t $(t = 1, 2, \ldots)$ is positive definite.

For notational convenience, we define the following mean-subtracted sequence:

$$\tilde{y}_{it} = y_{it} - \bar{y}_i, \quad \tilde{\boldsymbol{x}}_{it} = \boldsymbol{x}_{it} - \bar{\boldsymbol{x}}_i, \quad \tilde{\boldsymbol{z}}_{it} = \boldsymbol{z}_{it} - \bar{\boldsymbol{z}}_i,$$

where $\bar{y}_i = T^{-1}\sum_{t=1}^T y_{it}$, $\bar{\boldsymbol{x}}_i = T^{-1}\sum_{t=1}^T \boldsymbol{x}_{it}$, and $\bar{\boldsymbol{z}}_i = T^{-1}\sum_{t=1}^T \boldsymbol{z}_{it}$. We denote

$$\tilde{\boldsymbol{y}}_t = \begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \\ \vdots \\ \tilde{y}_{nt} \end{pmatrix}, \quad \tilde{X}_t = \begin{pmatrix} \tilde{\boldsymbol{x}}_{1t} \\ \tilde{\boldsymbol{x}}_{2t} \\ \vdots \\ \tilde{\boldsymbol{x}}_{nt} \end{pmatrix}, \quad \tilde{Z}_t = \begin{pmatrix} \tilde{\boldsymbol{z}}_{1t} \\ \tilde{\boldsymbol{z}}_{2t} \\ \vdots \\ \tilde{\boldsymbol{z}}_{nt} \end{pmatrix}. \tag{8}$$

We use the quantities to construct and analyze a conventional estimator for the fixed effect regression models below.

Before we discuss stationarity, we make the following assumptions for $(\boldsymbol{x}_{it}, \varepsilon_{it})$.

Assumption 1.

(a) $(\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{iT})$ and $(\varepsilon_{i1}, \ldots, \varepsilon_{iT})$ are independent and identically distributed (i.i.d.) over $i = 1, \ldots, n$.
(b) $E(\varepsilon_{it}) = 0$ and $E(\varepsilon_{it}\varepsilon_{is}) = 0$ for $t \neq s$.
(c) $\operatorname{plim} \bar{X} = M_x$ and $\operatorname{plim} \bar{Q}_x = Q_x$, where $M_x = E(X_t)$, $Q_x = E(\tilde{X}_t'\tilde{X}_t)$, $\bar{X} = T^{-1}\sum_{t=1}^T X_t$, and $\bar{Q}_x = T^{-1}\sum_{t=1}^T \tilde{X}_t'\tilde{X}_t$. Furthermore, $\bar{Q}_x$ is non-singular.

Remark 2.1. We can relax Assumption 1 using the strict exogeneity condition, as in a standard assumption for regression models. Since our concern is primarily to estimate the conditional covariance model, we use a simpler assumption for $\boldsymbol{x}_{it}$ and ε_{it}.

Assumption 2.

(a) $\boldsymbol{x}_{it}$ is stationary and ergodic and $E[||\boldsymbol{x}_{it}||^4] < \infty$.
(b) $\boldsymbol{\varepsilon}_t$ is i.i.d over $t = 1, \ldots, T$. The distribution of $\boldsymbol{\varepsilon}_t$ is absolutely continuous with respect to the Lebesgue measure on $\Re^m$, and zero is an interior point of the support of the distribution.

Remark 2.2. The fourth moment condition corresponds to the assumption of Theorem 1 of White (1980). Assumption 2(b) is from Theorem 2.4 of Boussama et al. (2011). Although Boussama et al. (2011) derive the condition for BEKK multivariate GARCH models, it is applicable in this context, as explained below.

Regarding the stationarity of $\boldsymbol{u}_t$, we start from the *vec* form of Eq. (6), given by

$$\boldsymbol{h}_t = \boldsymbol{\kappa} + \underline{C}\operatorname{vec}\left(\boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}'\right) + \underline{D}\boldsymbol{h}_{t-1}, \tag{9}$$

where $\boldsymbol{h}_t = \operatorname{vec}(H_t)$, $\boldsymbol{\kappa} = \operatorname{vec}(K)$, $\underline{C} = \operatorname{diag}(\operatorname{vec}(C))$, and $\underline{D} = \operatorname{diag}(\operatorname{vec}(D))$. We can apply the argument of Theorem 2.4 of Boussama et al. (2011) to the *vech* form, which is straightforwardly obtained by the *vec* form (9), to construct a Markov chain for a vector $(\operatorname{vech}(H_t)', \operatorname{vech}(H_{t-1})', \boldsymbol{u}_t', \boldsymbol{u}_{t-1}')'$. The stationarity condition is that the spectral radius of $\underline{C} + \underline{D}$ is less than 1, and it can be simplified as

$$\gamma + \rho < 1, \quad \varphi + \eta < 1. \tag{10}$$

From the condition in Proposition 1, we can verify $\gamma + \rho \geq 0$ since $\gamma + \rho \geq \max(\rho, -\rho(n-1)) + \rho \geq 0$ if $n \geq 2$. In the same manner, we obtain $\varphi + \eta \geq 0$. As we can derive the lower bounds of $\gamma + \rho$ and $\varphi + \eta$ by Proposition 1, they are not stated in Eq. (10). The argument of Theorem 2.4 of Boussama et al. (2011) implies that there exists a unique stationary ergodic solution to the models (5) and (6) and that $E[||\boldsymbol{u}_t||] < \infty$.

For the stationarity of $\boldsymbol{y}_t$, we make the usual assumption that $|\phi| < 1$. By Assumption 2, $\boldsymbol{x}_{it}$ is stationary process with finite fourth moments. Since $\boldsymbol{x}_{it}$ and $\boldsymbol{u}_t$ are stationary ergodic processes with finite second moments, $\boldsymbol{y}_t$ in Eq. (4) is also a stationary ergodic process with finite second moments, $E[||\boldsymbol{y}_t||^2] < \infty$.

Remark 2.3. Define $\Sigma = E(\boldsymbol{u}_t\boldsymbol{u}_t')$. Then, Σ is the solution of

$$\Sigma = K + C \circ \Sigma + D \circ \Sigma, \tag{11}$$

and we can show that Σ is positive definite under the condition of Proposition 1 by applying Theorem 2.2 of Ding and Engle (2001).

Now, we provide the condition for the positive definiteness of K, which is assumed in Proposition 1.

Proposition 2. *For the stationary panel GARCH models (4)–(6), K is positive definite if $\gamma + \varphi - \rho - \eta \le 0$.*

Remark 2.4. Under the parameter restrictions in Proposition 1, the condition, $\gamma + \varphi - \rho - \eta \le 0$, holds if $\gamma = \rho$ and $\varphi = \eta$. Hence, Propositions 1 and 2 imply that Eq. (6) reduces to

$$H_t = (1 - \gamma - \varphi)\Sigma + \gamma \boldsymbol{u}_{t-1}\boldsymbol{u}_{t-1}' + \varphi H_{t-1}, \tag{12}$$

to guarantee the positive definiteness of H_t $(t \ge 1)$ and that the stationarity condition is $\gamma + \varphi < 1$. Proposition 2 supports the empirical results of Cermeño and Grier (2006), since their estimates of γ and φ are similar to those of ρ and η, respectively.

Remark 2.5. We can derive the condition to ensure the positive definiteness of H_t $(t \ge 1)$ for the panel GARCH(p, q) model with the conditional variance and covariance equations defined by

$$\begin{aligned} h_{iit} &= \kappa_{ii} + \sum_{s=1}^{q} \gamma_s u_{i,t-s}^2 + \sum_{r=1}^{p} \varphi_r h_{ii,t-1}, \\ h_{ijt} &= \kappa_{ij} + \sum_{s=1}^{q} \rho_s u_{i,t-1} u_{j,t-1} + \sum_{r=1}^{p} \eta_r h_{ij,t-1}. \end{aligned} \tag{13}$$

Note that the panel GARCH(p, q) model is a special case of the diagonal GARCH model, as in Eq. (9). Scherrer and Ribarits (2007) show that all covariance matrices of the diagonal GARCH model are positive definite if and only if the model has a diagonal BEKK representation. For the panel GARCH(p, q) model, it is straightforward to show that the model has a diagonal BEKK representation if and only if $\gamma_s = \rho_s$ $(s = 1, \ldots, q)$ and $\varphi_r = \eta_r$ $(r = 1, \ldots, p)$.

Our feasible specification for the panel GARCH model is given by Eqs. (1) and (2) with

$$h_{ijt} = (1 - \gamma - \varphi)\sigma_{ij} + \gamma u_{i,t-1} u_{j,t-1} + \varphi h_{ij,t-1}, \tag{14}$$

for all i and j $(i, j = 1, \ldots, n)$.

2.3. Stationarity with Large n

We investigate strict stationarity with $n \to \infty$, as defined by Zhu et al. (2017).

Definition 1. Zhu et al. (2017). Let $\{\boldsymbol{y}_t \in \Re^n\}$ be an n-dimensional time series with $n \to \infty$. Define $\mathcal{W} = \{\omega \in \Re^\infty : \sum |\omega_i| < \infty\}$, where $\boldsymbol{\omega} = (\omega_i \in \Re^1 : 1 \le i \le \infty)' \in \Re^\infty$. For each $\omega in \mathcal{W}$, let $\boldsymbol{\omega}_n = (\omega_1, \ldots, \omega_n)' \in \Re^n$ be the truncated n-dimensional vector. $\{\boldsymbol{y}_t\}$ is then said to be strictly stationary if it satisfies the following conditions: for any $\boldsymbol{\omega} \in \mathcal{W}$, (1) $y_t^\omega = \lim_{n\to\infty} \boldsymbol{\omega}_n'\boldsymbol{y}_t$ exists in the almost sure sense and (2) y_t^ω is strictly stationary.

For the panel GARCH models (1), (2), and (14), we can show that y_t^ω follows an AR(1)-X-GARCH process, as presented below. We assume that $\boldsymbol{x}_t$ is strictly stationary in the sense of Definition 1. Define $\boldsymbol{x}_t^\omega = \lim_{n\to\infty} \boldsymbol{\omega}_n'\boldsymbol{x}_t$, $u_t^\omega = \lim_{n\to\infty} \boldsymbol{\omega}_n'\boldsymbol{u}_t$, and $h_t^\omega = \lim_{n\to\infty} \boldsymbol{\omega}_n' H_t \boldsymbol{\omega}_n$. Then, we obtain

$$\begin{aligned} y_t^\omega &= \alpha^\omega + \phi y_{t-1}^\omega + \boldsymbol{x}_t^\omega \boldsymbol{\beta} + u_t^\omega \\ u_t^\omega &= \sqrt{h_t^\omega}\varepsilon_t, \quad E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = 1, \\ h_t^\omega &= (1 - \gamma - \varphi)\sigma^\omega + \gamma (u_{t-1}^\omega)^2 + \varphi h_{t-1}^\omega, \end{aligned} \tag{15}$$

where $\alpha^\omega = \sum_{i=1}^\infty \omega_i \alpha_i$ and $\sigma^\omega = \sum_{i=1}^\infty \sum_{j=1}^\infty \omega_i \omega_j \sigma_{ij}$. We can show the existence of y_t^ω with probability one by verifying $E|y_t^\omega| < \infty$. Evidently, y_t^ω is strictly stationary if it exists. Hence, $\boldsymbol{y}_t$ is said to be strictly stationary (in the sense of Definition 1).

Remark 2.6. Nijman and Sentana (1996) discuss the contemporaneous aggregation of the diagonal GARCH model (e.g., Ding and Engle, 2001). By virtue of the common parameters in specification (γ, φ), our panel GARCH model yields the strong GARCH process defined by Drost and Nijman (1993) in which ε_t has an independent and identical distribution.

In the next section, we consider the estimation of the feasible panel GARCH model and examine the asymptotic and finite sample properties.

3. Three-Step VT Estimator

3.1. VT Estimator

Rather than the maximum likelihood estimation suggested by Cermeño and Grier (2006), we develop an estimation method based on the feasible specifications in (1), (2), and (14).

Denote the parameter vectors with their parameter spaces for the fixed effect regression model (4) as $\boldsymbol{\beta} \in \Theta_\beta \subset \Re^{k+1}$ and $\boldsymbol{\alpha} \in \Theta_\alpha \subset \Re^n$. Define the parameter vectors for the multivariate GARCH Eq. (12) based on the VT structure (11) as $\boldsymbol{\sigma} = \text{vec}(\Sigma) \in \Theta_\sigma \subset \Re^{n^2}$ and $\boldsymbol{\lambda} = (\gamma, \varphi)' \in \Theta_\lambda \subset \Re^2$. Denote the set of parameter vectors as $\boldsymbol{\psi} = (\boldsymbol{\theta}', \boldsymbol{\alpha}', \boldsymbol{\sigma}', \boldsymbol{\lambda}')' \in \Theta_\psi$, where $\Theta_\psi = \Theta_\theta \times \Theta_\alpha \times \Theta_\sigma \times \Theta_\lambda$. To emphasize the dependence of the parameters, we denote H_t as $H_t(\boldsymbol{\psi})$ and restate Eqs. (5) and (12) as

$$\boldsymbol{u}_t = H_t^{1/2}(\boldsymbol{\psi})\boldsymbol{\varepsilon}_t, \tag{16}$$

$$H_t(\boldsymbol{\psi}) = (1-\gamma-\varphi)\Sigma + \gamma\left[(\boldsymbol{y}_{t-1} - \boldsymbol{\alpha} - Z_{t-1}\boldsymbol{\theta})(\boldsymbol{y}_{t-1} - \boldsymbol{\alpha} - Z_{t-1}\boldsymbol{\theta})'\right] + \varphi H_{t-1}(\boldsymbol{\psi}). \tag{17}$$

We assume that the initial values, $\boldsymbol{y}_0$, $\boldsymbol{y}_{-1}$, X_0, and H_0, are given in our estimation (Hafner and Preminger, 2009; Pedersen and Rahbek, 2014).

We construct a three-step VT estimator as follows. In the first step, we estimate $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ using the fixed effect estimator. The second step estimates $\boldsymbol{\sigma}$ using the sample covariance matrix based on the residuals of the fixed effect regression. The third step conducts the QML estimation by optimizing the log-likelihood function for $\boldsymbol{\lambda}$ conditional on the other estimates. The details of the procedure are as follows.

Step 1. Estimate $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ using the conventional technique for the fixed effect regression model. The fixed effect estimator is given by

$$\hat{\boldsymbol{\theta}} = \left(\sum_{t=1}^T \tilde{Z}_t'\tilde{Z}_t\right)^{-1} \sum_{t=1}^T \tilde{Z}_t'\tilde{\boldsymbol{y}}_t = \left(\sum_{t=1}^T\sum_{i=1}^n \tilde{\boldsymbol{z}}_{it}'\tilde{\boldsymbol{z}}_{it}\right)^{-1} \sum_{t=1}^T\sum_{i=1}^n \tilde{\boldsymbol{z}}_{it}'\tilde{y}_{it}, \tag{18}$$

$$\hat{\boldsymbol{\alpha}} = \bar{\boldsymbol{y}} - \bar{Z}\hat{\boldsymbol{\theta}}, \tag{19}$$

where $\tilde{\boldsymbol{y}}_t$, $\tilde{X}_t$, and $\tilde{Z}_t$ are stated in Eq. (8), $\bar{\boldsymbol{y}} = T^{-1}\sum_{t=1}^T \boldsymbol{y}_t$, and $\bar{Z} = T^{-1}\sum_{i=1}^T Z_t$.

Step 2. Estimate $\boldsymbol{\sigma}$ using the sample covariance matrix, defined by

$$\hat{\boldsymbol{\sigma}} = \text{vec}\big(\hat{\Sigma}\big) = \text{vec}\left(\frac{1}{T}\sum_{t=1}^T \hat{\boldsymbol{u}}_t\hat{\boldsymbol{u}}_t'\right), \tag{20}$$

where $\hat{\boldsymbol{u}}_t = \boldsymbol{y}_t - \hat{\boldsymbol{\alpha}} - Z_t\hat{\boldsymbol{\theta}}$.

Step 3. Estimate $\boldsymbol{\lambda}$ by maximizing the quasi-log-likelihood function conditional on $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\sigma}})$. The VT estimator is given by

$$\hat{\boldsymbol{\lambda}} = \underset{\boldsymbol{\lambda}\in\Theta_\lambda}{\arg\max}\, L_T(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\sigma}}, \boldsymbol{\lambda}), \tag{21}$$

where

$$L_T(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\sigma}, \boldsymbol{\lambda}) = -\frac{NT}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^T \log\det(H_t(\boldsymbol{\psi})) - \frac{1}{2}\sum_{t=1}^T (\boldsymbol{y}_t - \boldsymbol{\alpha} - Z_t\boldsymbol{\theta})' H_t(\boldsymbol{\psi})^{-1}(\boldsymbol{y}_t - \boldsymbol{\alpha} - Z_t\boldsymbol{\theta}). \tag{22}$$

In the following, we examine the asymptotic and finite sample properties of the VT estimator.

3.2. Asymptotic Properties

We first investigate the asymptotic properties of the three-step VT estimator for the case with fixed n and large T. We make the following classical assumptions to state the result for consistency.

Assumption 3.

(a) The process $\{\boldsymbol{y}_t\}$ is strictly stationary and ergodic.
(b) The true parameters $\boldsymbol{\psi}_0 \in \Theta_\psi$ and Θ_ψ are compact.

(c) For $\boldsymbol{\lambda} \in \Theta_\lambda$, if $\boldsymbol{\lambda} \neq \boldsymbol{\lambda}_0$, then $H_t(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0, \boldsymbol{\lambda}) \neq H_t(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\sigma}_0, \boldsymbol{\lambda}_0)$ almost surely, for all $t \geq 1$.
(d) plim $\bar{Q}_z = Q_z$, where $Q_z = E(\tilde{Z}_t'\tilde{Z}_t)$ and $\bar{Q}_z = T^{-1}\sum_{t=1}^T \tilde{Z}_t'\tilde{Z}_t$, and $\bar{Q}_z$ is non-singular.
(e) $E[||\boldsymbol{\varepsilon}_t||^4] < \infty$.

We now state the following result regarding the consistency of the VT estimator.

Proposition 3. *Under Assumptions 1–3, as $T \to \infty$,*

$$\hat{\boldsymbol{\psi}} \xrightarrow{a.s.} \boldsymbol{\psi}_0. \tag{23}$$

Remark 3.1. Assumption 3(e) is necessary for the consistency of $\hat{\boldsymbol{\theta}}$, as in Theorem 1 of Nicholls and Pagan (1983).

We make the following assumption for the asymptotic normality of the VT estimator.

Assumption 4.

(a) $E[||\boldsymbol{u}_t||^6] < \infty$.
(b) $\boldsymbol{\theta}_0$ is in the interior of Θ.
(c) The distribution of $\boldsymbol{\varepsilon}_t$ is symmetric.

We need to assume finite sixth-order moments to show that the second-order derivatives of the log-likelihood function converge uniformly on the parameter space. In the univariate case, we only require finite fourth-order moments (see Francq et al. (2011)).

Proposition 4. *Under Assumptions 1–4, as $T \to \infty$,*

$$\sqrt{T}\left(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0\right) \xrightarrow{d} N(0, \underline{\Omega}_0), \tag{24}$$

where

$$\underline{\Omega}_0 = \begin{pmatrix} \Omega_0^m & O \\ O & P_0\Omega_0^v P_0' \end{pmatrix}, \quad P_0 = \begin{pmatrix} I_{n^2} & O \\ -(\Xi_0^\lambda)^{-1}\Xi_0^\sigma & -(\Xi_0^\lambda)^{-1} \end{pmatrix},$$

with the non-singular matrix Ξ_0^λ and the matrix Ξ_0^σ defined by (A.13) and the non-singular matrices Ω_0^m and Ω_0^v stated in Lemma 2.

Remark 3.2. The asymptotic covariance matrix is block diagonal under Assumption 4(c), which corresponds to the result for the QML estimator for the ARMA-GARCH process, as discussed in Remark 3.7 of Francq and Zakoïan (2004). Hence, the asymptotic accuracy of the estimators of the GARCH parameters is not affected by the mean equation. By relaxing the assumption, we obtain a non-diagonal asymptotic covariance matrix. Note that the specification of H_t will affect the estimate of Ω^m, as discussed later.

Remark 3.3. Instead of the VT estimator, we can consider QML estimation, which maximizes $L_T(\boldsymbol{\psi})$ directly, as in Cermeño and Grier (2006). Under Assumptions 1–4, we can show the consistency and asymptotic normality of the QML estimator by combining the works of Bollerslev and Wooldridge (1992) and Hafner and Preminger (2009). For asymptotic efficiency, the QML estimator is preferred to the VT estimator, as implied by Corollary 2 of Francq et al. (2011). VT estimation is more useful than QML estimation owing to its lower computational time. While the number of parameters in QML estimation is $k + 3 + n(n+3)/2$, the third-step estimator of the VT estimation method uses just two parameters in the optimization step. Furthermore, Francq et al. (2011) discuss the advantages of the VT approach and show that when the model is misspecified, the VT estimator can be superior to the QML estimator for making long-term predictions or calculating value-at-risk.

We can estimate Ω_0^m, Ω_0^v, Ξ_0^λ, and Ξ_0^σ using the sample outer-products and sample Hessian matrices as follows:

$$\hat{\Omega}^m = \frac{1}{T}\sum_{t=1}^T \hat{\boldsymbol{\omega}}_t^m \hat{\boldsymbol{\omega}}_t^{m\prime}, \quad \hat{\Omega}^v = \frac{1}{T}\sum_{t=1}^T \hat{\boldsymbol{\omega}}_t^v \hat{\boldsymbol{\omega}}_t^{v\prime},$$
$$\hat{\Xi}^\lambda = \frac{1}{T}\sum_{t=1}^T \hat{\Xi}_t^\lambda, \quad \hat{\Xi}^\sigma = \frac{1}{T}\sum_{t=1}^T \hat{\Xi}_t^\sigma, \tag{25}$$

where

$$\hat{\boldsymbol{\omega}}_t^m = \begin{pmatrix} \bar{Q}_z^{-1}\tilde{Z}_t'\hat{\boldsymbol{u}}_t \\ (I_n - \tilde{Z}\bar{Q}_z^{-1}\tilde{Z}_t')\hat{\boldsymbol{u}}_t \end{pmatrix}, \quad \hat{\boldsymbol{\omega}}_t^v = \begin{pmatrix} \mathrm{vec}(\hat{\boldsymbol{u}}_t\hat{\boldsymbol{u}}_t') - \hat{\boldsymbol{\omega}} \\ \left.\frac{\partial l_t(\boldsymbol{\psi})}{\partial \boldsymbol{\lambda}}\right|_{\psi=\hat{\psi}} \end{pmatrix},$$
$$\hat{\Xi}_t^\lambda = \left.\frac{\partial^2 l_t(\boldsymbol{\psi})}{\partial\boldsymbol{\lambda}\partial\boldsymbol{\lambda}'}\right|_{\psi=\hat{\psi}}, \quad \hat{\Xi}_t^\sigma = \left.\frac{\partial^2 l_t(\boldsymbol{\psi})}{\partial\boldsymbol{\lambda}\partial\boldsymbol{\sigma}'}\right|_{\psi=\hat{\psi}},$$

with $\bar{Q}_z$ stated in Assumption 3(d). We can then use the model-based estimator for Ω^m as

Table 1
Finite Sample Properties of the VT Estimator for the Panel GARCH Model ($n = 5$)

Parameter	True	$T = 250$			$T = 500$		
		Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
ϕ	0.6	0.5924	0.0287	0.0297	0.5959	0.0205	0.0209
β	1.0	0.9976	0.0974	0.0974	0.9992	0.0722	0.0722
α_1	0.2	0.2131	0.0971	0.0980	0.2063	0.0683	0.0685
α_2	0.4	0.4189	0.1016	0.1033	0.4117	0.0732	0.0741
α_3	0.6	0.6233	0.1106	0.1130	0.6114	0.0796	0.0804
α_4	0.8	0.8238	0.1254	0.1276	0.8144	0.0864	0.0875
α_5	1.0	1.0299	0.1351	0.1383	1.0169	0.0956	0.0970
σ_{11}	1.0	0.9847	0.2381	0.2385	1.0059	0.1888	0.1889
σ_{21}	-0.4	-0.3913	0.1798	0.1800	-0.4039	0.1404	0.1405
σ_{31}	0.16	0.1630	0.1761	0.1761	0.1608	0.1311	0.1311
σ_{41}	-0.064	-0.0626	0.1694	0.1694	-0.0594	0.1323	0.1323
σ_{51}	0.0256	0.0266	0.1702	0.1701	0.0253	0.1348	0.1348
σ_{22}	1.0	0.9973	0.2468	0.2467	0.9988	0.1906	0.1906
σ_{32}	-0.4	-0.4016	0.1845	0.1844	-0.3973	0.1384	0.1384
σ_{42}	0.16	0.1608	0.1729	0.1729	0.1555	0.1307	0.1307
σ_{52}	-0.064	-0.0692	0.1713	0.1714	-0.0642	0.1317	0.1316
σ_{33}	1.0	1.0113	0.2566	0.2568	0.9945	0.1784	0.1785
σ_{43}	-0.4	-0.4040	0.1875	0.1875	-0.3976	0.1378	0.1377
σ_{53}	0.16	0.1695	0.1747	0.1749	0.1604	0.1309	0.1309
σ_{44}	1.0	0.9948	0.2485	0.2485	1.0026	0.1884	0.1884
σ_{54}	-0.4	-0.4045	0.1792	0.1793	-0.3983	0.1441	0.1441
σ_{55}	1.0	1.0069	0.2436	0.2436	1.0036	0.1878	0.1878
γ	0.04	0.0357	0.0090	0.0099	0.0385	0.0057	0.0059
φ	0.94	0.9042	0.0532	0.0641	0.9300	0.0112	0.0150
$\gamma+\varphi$	0.98	0.9398	0.0522	0.0659	0.9685	0.0088	0.0145

Note: The number of replications is 2000.

$$\tilde{\Omega}^m = \frac{1}{T}\sum_{t=1}^{T} \tilde{\boldsymbol{\omega}}_t^m H_t(\hat{\boldsymbol{\psi}})\tilde{\boldsymbol{\omega}}_t^{m\prime}, \quad \tilde{\boldsymbol{\omega}}_t^m = \begin{pmatrix} \bar{Q}_z^{-1}\tilde{Z}_t' \\ (I_n - \bar{Z}\bar{Q}_z^{-1}\tilde{Z}_t') \end{pmatrix}, \tag{26}$$

with $H_1(\hat{\boldsymbol{\psi}}) = \hat{\Sigma}$ being the initial value. Note that $\hat{\Omega}^m$ is robust to model misspecification, while $\tilde{\Omega}^m$ is expected to be efficient when the model is correctly specified. We compare these in our empirical analysis.

Remark 3.4. When n tends to infinity but T is fixed, it is known that $\hat{\boldsymbol{\theta}}$ is inconsistent, leading to the inconsistency of the remaining estimators.

Remark 3.5. When n tends to infinity, $\boldsymbol{Y}_t$ is strictly stationary in the sense of Definition 1, as discussed in Section 2.3. For a large n, the first-step estimator is biased but consistent as T tends to infinity, as in the standard panel data analysis. Hence, we can apply the same argument for the proof of Propositions 3 and 4 to show the consistency and asymptotic normality of the VT estimator when T increases faster than n. We need further examination when $n/T \to k < \infty$ and $n/T^3 \to 0$ (see Section 6.2 of Arellano (2003) for instance). Even if the first-step estimator is consistent, we need to assume sparsity on Σ to obtain its estimate, which is invertible, as in Bickel and Levina (2008a,b) for instance. As their approaches assume independent and identical distributions, they are inapplicable to our case directly. We await further theoretical research on estimating large covariance matrices.

3.3. Monte Carlo Experiment

We examine the finite sample property of the three-step VT estimator for the feasible panel GARCH model defined by Eqs. (1), (2), and (14). To compare the properties for different dimensions and/or different sample sizes, we use two dimensions ($n = 5, 20$) and two sample sizes ($T = 250, 500$). We consider an exogenous variable that follows the independent uniform distribution, that is, $k = 1$ and $x_{it} \sim U(0, 1)$. We specify the parameters as

$$\alpha_i = i/n, \quad \boldsymbol{\theta} = (0.6, 1)', \quad \Sigma = \{\sigma_{ij}\}, \quad \sigma_{ij} = (-0.4)^{|i-j|}, \quad \boldsymbol{\lambda} = (0.04, 0.94)'.$$

The parameters satisfy the stationarity condition and the unconditional covariance matrix of $\boldsymbol{u}_t$, Σ, is positive definite. The number of replications is 2000.

Table 1 presents the sample mean, standard deviation, and root mean squared error (RMSE) of the VT estimates for the dimension $n = 5$. The sample means are close to the true values and the standard deviations are similar to the RMSEs. The biases are negligible even for $T = 250$. As the sample size increases, the RMSEs have lower values. Figure 1 shows the histogram and QQ plots for the VT estimates of (γ, φ) for $(n, T) = (5, 250)$. The distribution of γ is close to the normal distribution, whereas that of φ is skewed to the left. Figure 2 displays the histogram and QQ plots for $(n, T) = (5, 500)$. The

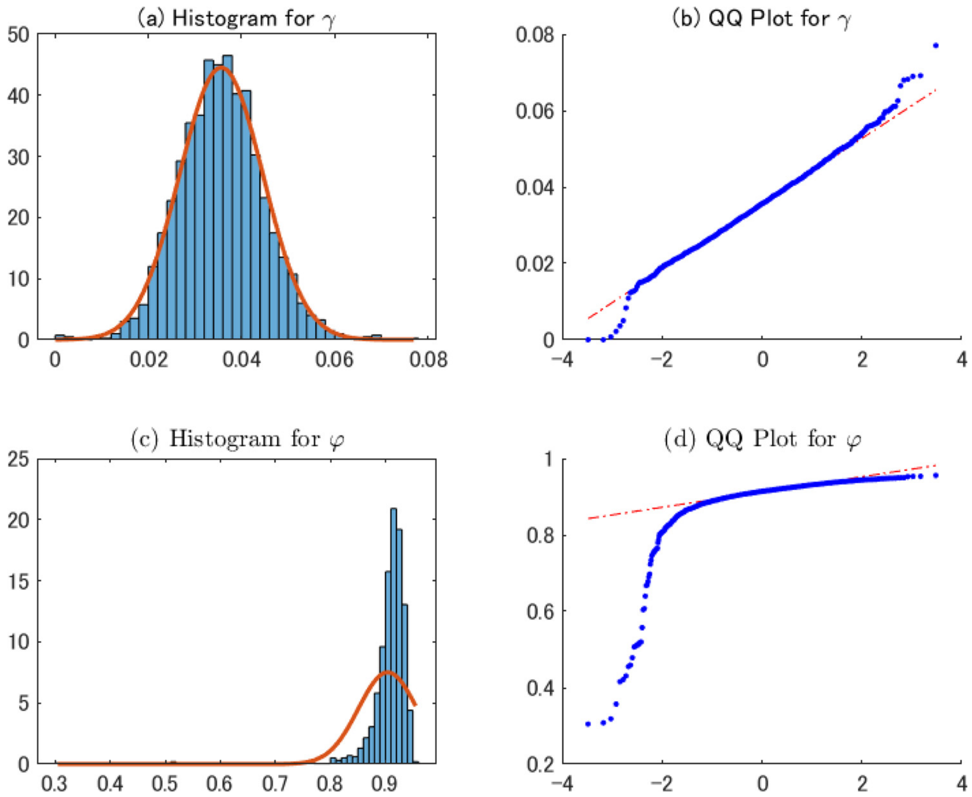

Fig. 1. Histograms and QQ Plots of the VT Estimates for $(n, T) = (5, 250)$. *Note:* The red line in the histograms indicates the normal density with the same mean and variance.

Table 2
Finite Sample Properties of the VT Estimator for the Panel GARCH Model ($n = 20$)

Parameter	True	$T = 250$			$T = 500$		
		Mean	Std. Dev.	RMSE	Mean	Std. Dev.	RMSE
ϕ	0.6	0.5937	0.0178	0.0189	0.5970	0.0133	0.0137
β	1.0	0.9994	0.0495	0.0495	0.9998	0.0358	0.0358
γ	0.04	0.0294	0.0030	0.0110	0.0351	0.0018	0.0052
φ	0.94	0.9123	0.0078	0.0287	0.9335	0.0035	0.0073
$\gamma + \varphi$	0.98	0.9417	0.0070	0.0389	0.9687	0.0028	0.0117

Note: The number of replications is 2000. We omitted the results for $\boldsymbol{\alpha}$ and $\boldsymbol{\sigma}$ to save space.

pattern of skewness is the same as that in Figure 1, but Figure 2 shows that the distributions are closer to normal than those of Figure 1 are. Figure 2 supports the asymptotic normality of the VT estimator, shown by Proposition 4.

Table 2 presents the results for the VT estimates for $n = 20$ with $T = 250$ and 500. We compare the increasing dimension n for $T = 250$. The sample means in Table 2 for $T = 250$ are close to those in Table 1. On the contrary, the RMSEs in Table 2 for $T = 250$ are smaller than those in Table 1. Figure 3 presents the histogram and QQ plots for the VT estimates for $(n, T) = (20, 250)$, which imply that the distributions are close to normal. Hence, increasing the dimension shrinks the asymptotic covariance matrix of $\hat{\boldsymbol{\lambda}}$, while it does not quicken the speed of convergence, as discussed in Remarks 3.4 and 3.5. The result for $n = 20$ with $T = 500$ in Table 2 and Figure 4 also supports Remark 3.5.

4. Empirical Examples

We provide illustrative empirical examples for the estimation of the feasible GARCH model using the inflation rates of G7 countries and growth rates for the value of trade on goods for four economic regions.

We calculate the monthly inflation rates based on the consumer price indices (CPIs), while we obtain the growth rates for the value of trade using the export and import of goods. The four economic regions are advanced economies, emerging and developing Asia, sub-Saharan Africa, and the remaining emerging and developing countries. The data are from International Financial Statistics. The sample period for inflation rates is from January 1978 to February 2021, and we set $(n, T) = (7, 517)$. Since our data include the period in which the COVID-19 pandemic hit, we define a COVID-19 dummy variable that takes

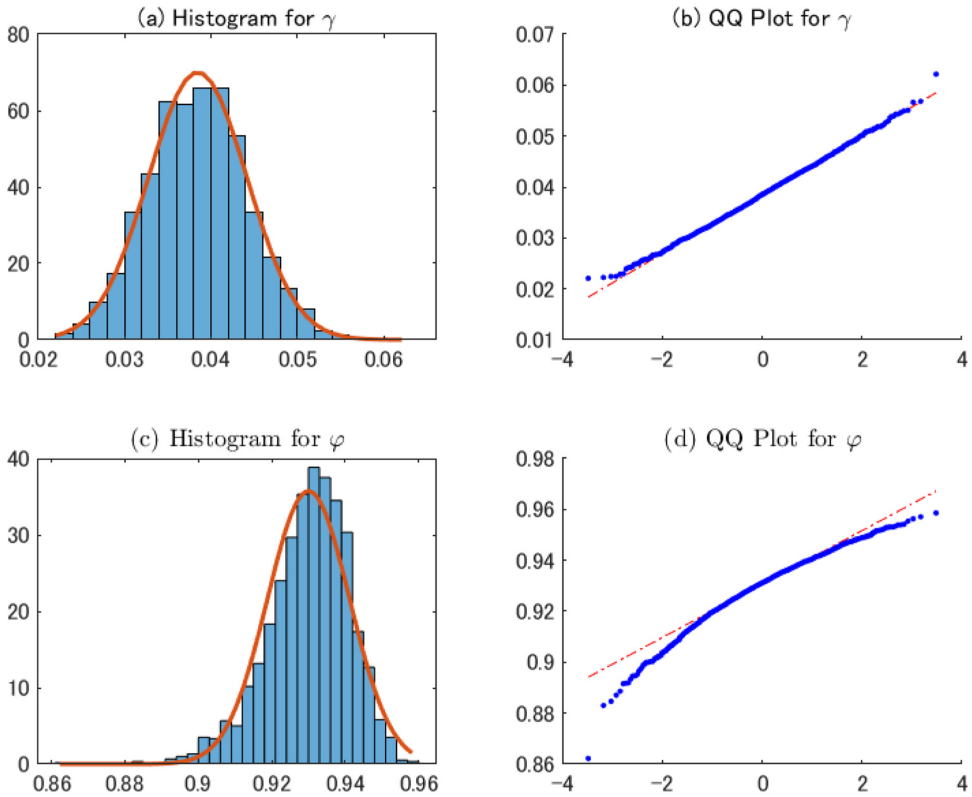

Fig. 2. Histograms and QQ Plots of the VT Estimates for $(n, T) = (5, 500)$. *Note:* The red line in the histograms indicates the normal density with the same mean and variance.

Table 3
First-Step Estimates of the Inflation Rates of G7 Countries

Parameter	Estimate	Robust S.E.	FP S.E.	CCP S.E.
ϕ (AR(1))	0.4086	(0.0299)	(0.0287)	(0.0277)
β (COVID-19)	-0.0859	(0.0527)	(0.0577)	(0.0598)
α_1 (Canada)	0.1595	(0.0183)	(0.0196)	(0.0196)
α_2 (France)	0.1546	(0.0153)	(0.0153)	(0.0154)
α_3 (Germany)	0.1048	(0.0164)	(0.0162)	(0.0163)
α_4 (Italy)	0.2297	(0.0164)	(0.0202)	(0.0201)
α_5 (Japan)	0.0535	(0.0192)	(0.0147)	(0.0148)
α_6 (UK)	0.1870	(0.0186)	(0.0160)	(0.0160)
α_7 (US)	0.1673	(0.0151)	(0.0181)	(0.0181)

Note: Standard errors are in parentheses. "Robust S.E." denotes robust standard errors based on $\hat{\Omega}^m$ stated in Eq. (25), while "FP S.E." and "CCP S.E." denote standard errors based on $\tilde{\Omega}^m$ defined by Eq. (26) with the covariance matrix of the feasible panel GARCH and CCP GARCH models, respectively.

one from February 2020 and zero otherwise. For the growth rates on trade, the sample period is from January 1978 to July 2019, giving $(n, T) = (4, 498)$. We examine the effect of the global financial crisis (GFC) by using a GFC dummy variable that takes one from September to December 2008 and zero otherwise.

We estimate the feasible panel GARCH model in (1), (2), and (14) with the dummy as an exogenous variable in the mean equation. As a benchmark, we consider the constant correlation panel (CCP) GARCH model, as in Cermeño and Sanin (2015) and Ribeiro et al. (2017). The elements conditional covariance matrix of the CCP GARCH model is given by

$$h_{iit} = (1 - \gamma - \varphi)\sigma_{ii} + \gamma u_{i,t-1}^2 + \varphi h_{ii,t-1}, \quad h_{ijt} = r_{ij}\sqrt{h_{iit}h_{jjt}}, \tag{27}$$

where $|r_{ij}| < 1$ for all $i \neq j$ $(i, j = 1, \dots, n)$. Eq. (27) corresponds to (14) in the feasible panel GARCH model. We can obtain the VT estimator for the CCP model, as the estimates of σ_{ii} and r_{ij} are calculated using $\hat{\Sigma}$ in the second step. The number of parameters in the CCP model is the same as that in the feasible GARCH model.

Table 3 shows the first-step estimates for ϕ, β, and α_i $(i = 1, \dots, 7)$ and their standard errors. The robust standard errors based on $\hat{\Omega}^m$ are close to the standard errors based on $\tilde{\Omega}^m$ with the covariance matrix of the feasible panel GARCH and CCP GARCH models. The autoregressive parameter is 0.41, which is positive and significant at the five percent level. The

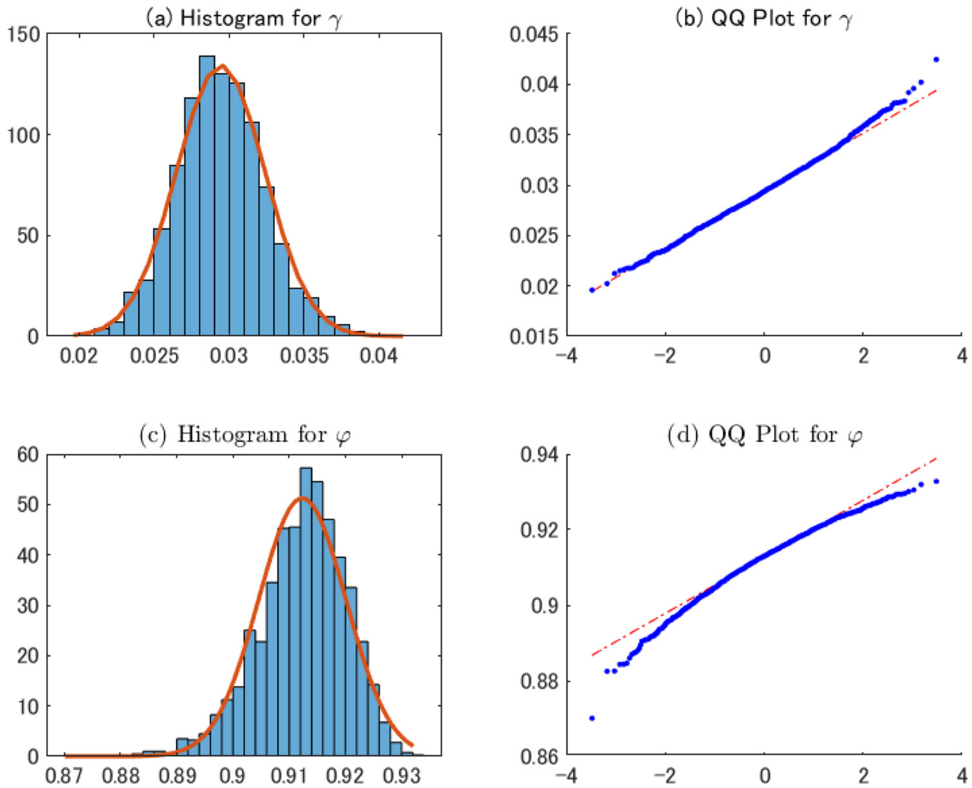

Fig. 3. Histograms and QQ Plots of the VT Estimates for $(n, T) = (20, 250)$. *Note:* The red line in the histograms indicates the normal density with the same mean and variance.

Table 4
Third-Step Estimates of the Inflation Rates of G7 Countries

Parameter	Feasible Panel GARCH		CCP GARCH	
γ	0.0340	(0.0047)	0.0580	(0.0069)
φ	0.9431	(0.0097)	0.9243	(0.0144)
QLike	- 921.10		- 891.05	

Note: Standard errors are in parentheses. "QLike" denotes the quasi-log-likelihood stated in (22).

coefficient for the COVID-19 dummy is negative and insignificant, implying that the effect on inflation rates is negligible. All the estimates of the individual effects are significant. The individual effect for Japan is relatively low, implying low inflation rates over 40 years. Table 4 reports the third-step estimates for the feasible panel GARCH and CCP GARCH models. All the parameters are positive and significant and they satisfy the stationarity condition. The estimates of the feasible panel GARCH model are close to the corresponding values for the CCP GARCH model. The CCP GARCH model has the higher value for the quasi-log-likelihood, implying that the data prefer the CCP GARCH model. It is unnecessary to check information criteria such as the Akaike and Bayesian information criteria, since the number of parameters in the two models is the same. Using the VT estimates as the initial values, we obtain the QML estimates, which are listed in Table 5. There are no major changes in the results as compared to Tables 3 and 4.

Regarding the growth rates for the value of trade in the four regions, we obtain the QML estimates using the VT estimates as the initial values. Table 6 reports the QML estimates. The autoregressive parameter is negative and significant. The coefficient for the GFC dummy is negative and significant, which shows the negative impact of the GFC on the growth rates of trade. All individual effects reject the null hypothesis that the parameter is zero. For the feasible panel GARCH specification, the estimates of γ and φ are positive and significant, and $\gamma + \varphi$ is less than 0.35. The estimate of γ is significant for the CCP GARCH model, whereas φ is insignificant. The values of the quasi-log-likelihood function indicate that the feasible panel GARCH specification is preferable to the CCP GARCH model.

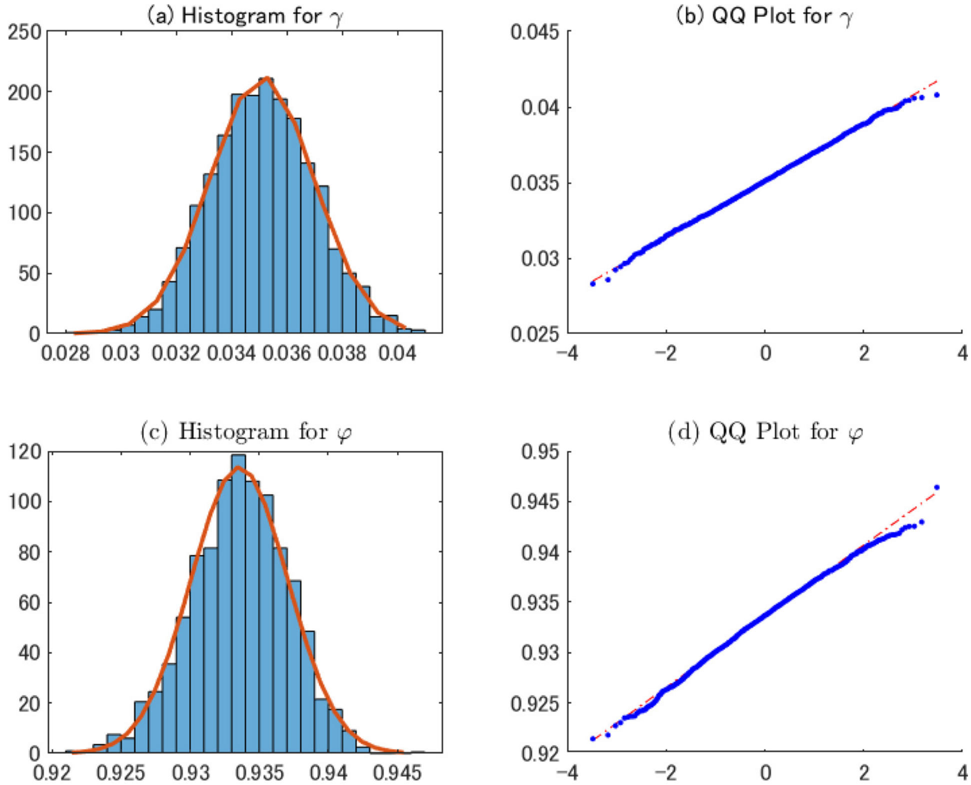

Fig. 4. Histograms and QQ Plots of the VT Estimates for $(n, T) = (20, 500)$. *Note:* The red line in the histograms indicates the normal density with the same mean and variance.

Table 5
QML Estimates of the Inflation Rates of G7 Countries

Parameter	Feasible Panel GARCH		CCP GARCH	
ϕ (AR(1))	0.1373	(0.0213)	0.1355	(0.0191)
β (COVID-19)	-0.0807	(0.0452)	-0.0762	(0.0593)
α_1 (Canada)	0.1585	(0.0190)	0.1686	(0.0196)
α_2 (France)	0.1268	(0.0140)	0.1344	(0.0139)
α_3 (Germany)	0.1125	(0.0164)	0.1196	(0.0166)
α_4 (Italy)	0.1802	(0.0128)	0.1914	(0.0132)
α_5 (Japan)	0.0203	(0.0217)	0.0277	(0.0208)
α_6 (UK)	0.1502	(0.0185)	0.1618	(0.0202)
α_7 (US)	0.1858	(0.0149)	0.1907	(0.0132)
γ	0.0528	(0.0039)	0.0638	(0.0047)
φ	0.9424	(0.0043)	0.9147	(0.0065)
QLike	- 774.98		- 761.61	

Note: Standard errors are shown in parentheses. We omit the estimates for Σ to save space. "QLike" denotes the quasi-log-likelihood stated in (22).

Table 6
QML Estimates of the Growth Rates for the Value of Trade of Four Regions

Parameter	Feasible Panel GARCH		CCP GARCH	
ϕ (AR(1))	-0.3473	(0.0285)	-0.3497	(0.0210)
β (GFC)	-0.1651	(0.0255)	-0.1751	(0.0235)
α_1 (Advanced Economy)	0.0068	(0.0029)	0.0072	(0.0030)
α_2 (Emerging and developing Asia)	0.0156	(0.0042)	0.0159	(0.0041)
α_3 (Sub-Saharan Africa)	0.0111	(0.0028)	0.0106	(0.0029)
α_4 (Others)	0.0080	(0.0030)	0.0089	(0.0032)
γ	0.1279	(0.0239)	0.1512	(0.0204)
φ	0.2140	(0.0990)	0.1052	(0.0704)
QLike	285.69		284.41	

Note: Standard errors are shown in parentheses. We omit the estimates for Σ to save space. "QLike" denotes the quasi-log-likelihood stated in (22).

5. Conclusion

For the panel GARCH model developed by Cermeño and Grier (2006), we derived the conditions for the stationarity and positive definiteness of the covariance matrix. Based on the result, we suggested a feasible panel GARCH model and examined stationarity when n tends to infinity. We considered the three-step VT estimator and demonstrated its consistency and asymptotic normality when T goes to infinity. We investigated the finite sample property of the VT estimator using Monte Carlo experiments. The Monte Carlo results indicate that the finite sample properties of the VT estimator are satisfactory. Furthermore, the results show that the size of n does not affect the speed of convergence, while it does affect the asymptotic covariance matrix. The empirical results for the inflation rates of the G7 countries and growth rates for the value of trade on goods for four economic regions indicate that the feasible panel GARCH specification is a competitive alternative to the CCP GARCH model of Cermeño and Sanin (2015) and Ribeiro et al. (2017). The empirical results indicate that the GFC significantly decreases the growth rates on trades, while the impact of COVID-19 on inflation rates is insignificant for G7.

As in recent empirical work, it is useful to derive the asymptotic theory for the panel GARCH-in-mean and GARCH-X specifications. We may apply the theoretical work of Conrad and Mammen (2016) for the GARCH-in-mean model and extend the work of Han and Kristensen (2014) for GARCH-X models. Furthermore, we may consider time-varying GARCH parameters as in Cho and Korkas (2021) and Silvennoinen and Teräsvirta (2021). The derivation of the asymptotic theory for such extensions is an important direction for future research.

Declaration of Competing Interest

I have no conflict of interest to declare.

Acknowledgements

The author is most grateful to Yoshihisa Baba, Naoki Honda, Erricos Kontoghiorghes, the associate editor, and two anonymous reviewers for helpful comments and suggestions. The author acknowledges the financial support of the Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science (JSPS 19K01594), and Australian Academy of Science.

Appendix A

A1. Derivatives of the Quasi-Log-Likelihood Function

We can restate the quasi-log-likelihood function given by (22) as

$$L_T(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\sigma}, \boldsymbol{\lambda}) = \sum_{t=1}^{T} l_t, \tag{A.1}$$

where l_t is the tth contribution defined by

$$l_t = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log \bar{h}_t - \frac{\bar{u}_t^2}{2\bar{h}_t}. \tag{A.2}$$

The gradient and Hessian of the quasi-log-likelihood function are given by

$$\frac{\partial L_T}{\partial \boldsymbol{\lambda}} = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial l_t}{\partial \boldsymbol{\lambda}}, \quad \frac{\partial^2 L_T}{\partial \boldsymbol{\lambda}\partial \boldsymbol{\lambda}'} = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 l_t}{\partial \boldsymbol{\lambda}\partial \boldsymbol{\lambda}'}.$$

Applying the chain rule and product rule, we obtain

$$\frac{\partial l_t}{\partial \boldsymbol{\lambda}} = \frac{\partial \bar{h}_t}{\partial \boldsymbol{\lambda}}\frac{\partial l_t}{\partial \bar{h}_t}, \quad \frac{\partial^2 l_t}{\partial \boldsymbol{\lambda}\partial \boldsymbol{\lambda}'} = \frac{\partial^2 \bar{h}_t}{\partial \boldsymbol{\lambda}\partial \boldsymbol{\lambda}'}\frac{\partial l_t}{\partial \bar{h}_t} + \frac{\partial \bar{h}_t}{\partial \boldsymbol{\lambda}}\frac{\partial \bar{h}_t}{\partial \boldsymbol{\lambda}'}\frac{\partial^2 l_t}{\partial \bar{h}_t^2}, \tag{A.3}$$

with

$$\frac{\partial l_t}{\partial \bar{h}_t} = \frac{1}{2\bar{h}_t}\left(\frac{\bar{u}_t^2}{\bar{h}_t} - 1\right), \quad \frac{\partial^2 l_t}{\partial \bar{h}_t^2} = \frac{1}{2\bar{h}_t^2}\left(1 - \frac{2\bar{u}_t^2}{\bar{h}_t}\right) \tag{A.4}$$

and

$$\begin{aligned} &\frac{\partial \bar{h}_t}{\partial \gamma} = (\bar{u}_{t-1}^2 - \bar{\sigma}) + \varphi\frac{\partial \bar{h}_{t-1}}{\partial \gamma}, \quad \frac{\partial \bar{h}_t}{\partial \varphi} = (\bar{h}_{t-1} - \bar{\sigma}) + \varphi\frac{\partial \bar{h}_{t-1}}{\partial \varphi}, \\ &\frac{\partial^2 \bar{h}_t}{\partial \gamma^2} = \varphi\frac{\partial^2 \bar{h}_{t-1}}{\partial \gamma^2}, \quad \frac{\partial^2 \bar{h}_t}{\partial \gamma\partial \varphi} = \frac{\partial \bar{h}_{t-1}}{\partial \gamma} + \varphi\frac{\partial^2 \bar{h}_{t-1}}{\partial \gamma\partial \varphi}, \quad \frac{\partial^2 \bar{h}_t}{\partial \varphi^2} = 2\frac{\partial \bar{h}_{t-1}}{\partial \varphi} + \varphi\frac{\partial^2 \bar{h}_{t-1}}{\partial \varphi^2}. \end{aligned} \tag{A.5}$$

To compute $\frac{\partial \boldsymbol{h}_t}{\partial \boldsymbol{\lambda}'}$ and $\frac{\partial^2 \boldsymbol{h}_t}{\partial \lambda_j \partial \lambda_k}$, we set zeros for the pre-sample values as in Fiorentini et al. (1996).

A2. Proofs of the Propositions

Proof of Proposition 1. As discussed in Ding and Engle (2001), specification (6) produces positive definite matrix H_t for all t as long as initial covariance matrix H_0 is positive definite when the matrices of parameters K, C, and D are positive (semi-)definite. Hence, we need to check the latter conditions. For any $n \times 1$ vector $\boldsymbol{x}$, we obtain

$$\begin{aligned}\boldsymbol{x}'C\boldsymbol{x} &= (\gamma-\rho)\sum_{i=1}^n x_i^2 + \rho\left(\sum_{i=1}^n x_i\right)^2 \\ &= (\gamma-\rho)\sum_{i=1}^n (x_i-\bar{x})^2 + (\gamma+\rho(n-1))n\bar{x}^2 \geq 0,\end{aligned}$$

under the condition in Proposition 1, where $\bar{x} = n^{-1}\sum_{i=1}^n x_i$. Hence, C is positive (semi-)definite. Similarly, we can show that D is positive (semi-)definite. Therefore, Proposition 1 holds. □

Proof of Proposition 2. By Eq. (11), we obtain $K = \Sigma \circ (J_n - C - D)$, where J_T is the $n \times n$ matrix of ones. For any vector $\boldsymbol{x}$ and Assumption 1, we obtain

$$\begin{aligned}\boldsymbol{x}'K\boldsymbol{x} &= (1-\rho-\eta)\left(\sum_{i=1}^n x_i\right)^2 - (\gamma+\varphi-\rho-\eta)\sum_{i=1}^n x_i^2 \\ &= n[(1-\gamma-\varphi)+(n-1)(1-\rho-\eta)]\bar{x}^2 \\ &\quad - (\gamma+\varphi-\rho-\eta)\sum_{i=1}^n (x_i-\bar{x})^2 \geq 0,\end{aligned}$$

if $\gamma+\varphi-\rho-\eta \leq 0$. □

Proof of Proposition 3. For the fixed effect estimator of $\boldsymbol{\theta}$, we obtain

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = \left(\frac{1}{T}\sum_{t=1}^T \tilde{Z}_t'\tilde{Z}_t\right)^{-1}\frac{1}{T}\sum_{t=1}^T \tilde{Z}_t'\boldsymbol{u}_t = \left(\frac{1}{T}\sum_{t=1}^T Z_t^{*\prime}Z_t^*\right)^{-1}\frac{1}{T}\sum_{t=1}^T Z_t^{*\prime}\boldsymbol{u}_t^*, \tag{A.6}$$

where Z_t^* and $\boldsymbol{u}_t^*$ are defined by the orthogonal transformation of $\tilde{Z}_t$ and $\boldsymbol{u}_t$ (e.g., Arellano, 2003):

$$\begin{aligned}u_{it}^* &= \sqrt{\frac{T-t}{T-t+1}}\left[u_{it} - \frac{1}{T-t}(u_{i,t+1}+\cdots+u_{iT})\right], \\ \boldsymbol{z}_{it}^* &= \sqrt{\frac{T-t}{T-t+1}}\left[\boldsymbol{z}_{it} - \frac{1}{T-t}(\boldsymbol{z}_{i,t+1}+\cdots+\boldsymbol{z}_{iT})\right],\end{aligned}$$

respectively. Since Z_t and $\boldsymbol{u}_t$ are strictly stationary and ergodic, we can apply the uniform law of large numbers for stationary ergodic processes (see Lemma A.2.2 of White (1994)) for each element of $Z_t'Z_t$ and $Z_t'\boldsymbol{u}_t$ under the existence of second moments. Using the technique of the proof of Theorem 1 in Nicholls and Pagan (1983), we obtain $\hat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}_0$.

An alternative form of $\hat{\boldsymbol{\alpha}}$ is given by

$$\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0 = \frac{1}{T}\sum_{t=1}^T S_t'\boldsymbol{u}_t, \tag{A.7}$$

where

$$S_t = I_n - \tilde{Z}_t\left(\frac{1}{T}\sum_{t=1}^T \tilde{Z}_t'\tilde{Z}_t\right)^{-1}\bar{Z}'.$$

The same technique for $\hat{\boldsymbol{\theta}}$ yields $\hat{\boldsymbol{\alpha}} \xrightarrow{a.s.} \boldsymbol{\alpha}_0$.

For the estimator Σ, we obtain

$$\begin{aligned}\hat{\Sigma} &= \frac{1}{T}\sum_{t=1}^T \left[\boldsymbol{u}_t - (\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}_0) - Z_t(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)\right]\left[\boldsymbol{u}_t - (\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}_0) - Z_t(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0)\right]' \\ &= \frac{1}{T}\sum_{t=1}^T \boldsymbol{u}_t\boldsymbol{u}_t' + o_p(1),\end{aligned} \tag{A.8}$$

using $\hat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\alpha}} \xrightarrow{a.s.} \boldsymbol{\alpha}_0$. Assumption 3(a) and $E||\boldsymbol{u}_t||^2] < \infty$ indicate that according to the ergodic theorem, as $T \to \infty$, $\hat{\boldsymbol{\sigma}} \xrightarrow{a.s.} \boldsymbol{\sigma}_0$.

For the third-step estimator, we can apply the technique from the proof of Theorem 2.1 in Newey and McFadden (1994), as in the proof of Theorem 4.1 of Pedersen and Rahbek (2014), to show $\hat{\boldsymbol{\lambda}} \xrightarrow{a.s.} \boldsymbol{\lambda}_0$. □

Lemma 1. *Under Assumptions 1–4, as $T \to \infty$,*

$$\sqrt{T}\begin{pmatrix} \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \\ \hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0 \\ \hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0 \\ \partial L_T(\boldsymbol{\psi}_0)/\partial \boldsymbol{\lambda} \end{pmatrix} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix} \Upsilon_t^m(\boldsymbol{\psi}_0) & O \\ O & \Upsilon_t^v(\boldsymbol{\psi}_0) \end{pmatrix}\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t - I_n) \end{pmatrix} + o_p(1), \tag{A.9}$$

where $\frac{\partial L_T(\boldsymbol{\psi}_0)}{\partial \lambda} = \frac{\partial L_T(\boldsymbol{\psi})}{\partial \lambda}\Big|_{\psi=\psi_0}$ *and*

$$\Upsilon_t^m(\boldsymbol{\psi}_0) = \begin{pmatrix} Q_z^{-1}(Z_t - M_z)' H_{0t}^{1/2} \\ \left(I_n - M_z Q_z^{-1}(Z_t - M_z)'\right) H_{0t}^{1/2} \end{pmatrix},$$

$$\Upsilon_t^v(\boldsymbol{\psi}_0) = \begin{pmatrix} (1-\gamma_0-\varphi_0)^{-1}(1-\varphi_0)\left(H_{0t}^{1/2} \otimes H_{0t}^{1/2}\right) \\ n^{-2}\left[\sum_{s=0}^{\infty}\varphi_0^s N_{t-1-s}\right]'\left(\iota' H_{0t}^{-1/2} \otimes \iota' H_{0t}^{-1/2}\right) \end{pmatrix},$$

with $N_t = (\bar{u}_t^2 - \bar{\sigma}, \bar{h}_t - \bar{\sigma})'$ *and* $\iota = (1, \ldots, 1)'$.

Proof of Lemma 1. Assumption 3(d) implies $\bar{Z} - M_z = o_p(1)$, and $\hat{Q}_z - Q_z = o_p(1)$. From (A.6), we obtain

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = \bar{Q}_z^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z_t^{*\prime}\boldsymbol{u}_t^* = Q_z^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(Z_t - M_z)'\boldsymbol{u}_t + o_p(1).$$

Similarly, from (A.7), we obtain

$$\sqrt{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[I_n - \tilde{Z}_t\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{Z}_t'\tilde{Z}_t\right)^{-1}\tilde{Z}'\right]'\boldsymbol{u}_t$$

$$= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[I_n - (Z_t - M_z)\bar{Q}_z^{-1}M_z'\right]'\boldsymbol{u}_t + o_p(1),$$

where $M_z = [\boldsymbol{\mu}_y\ M_x]$. Since $\hat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\alpha}} \xrightarrow{a.s.} \boldsymbol{\alpha}_0$ from Proposition 4 and (A.8), following the same argument used in the proof of Lemma B.8 of Pedersen and Rahbek (2014), we obtain

$$\sqrt{T}\begin{pmatrix} \hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0 \\ \partial L_T(\boldsymbol{\psi}_0)/\partial \boldsymbol{\lambda} \end{pmatrix} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Upsilon_t^v(\boldsymbol{\psi}_0)\mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t - I_n) + o_p(1).$$

These results establish (A.9). □

Lemma 2. *Under Assumptions 1–4, as $T \to \infty$,*

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{pmatrix} \Upsilon_t^m(\boldsymbol{\psi}_0) & O \\ O & \Upsilon_t^v(\boldsymbol{\psi}_0) \end{pmatrix}\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t - I_n) \end{pmatrix} \xrightarrow{d} N\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Omega_0^m & O \\ O & \Omega_0^v \end{bmatrix}\right), \tag{A.10}$$

where

$$\Omega_0^m = E\left[\Upsilon_t^m(\boldsymbol{\psi}_0)\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\Upsilon_t^m(\boldsymbol{\psi}_0)'\right]$$

$$\Omega_0^v = E\left[\Upsilon_t^v(\boldsymbol{\psi}_0)\mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t - I_n)\mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t - I_n)'\Upsilon_t^v(\boldsymbol{\psi}_0)'\right].$$

Proof of Lemma 2. From the structure, $(\Upsilon_t^m(\boldsymbol{\psi}_0), \Upsilon_t^v(\boldsymbol{\psi}_0), \boldsymbol{\varepsilon}_t, \mathrm{vec}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t - I_n)$ is an ergodic martingale difference sequence. As in the Appendix of Nicholls and Pagan (1983) and Lemma B.9 of Pedersen and Rahbek (2014), we can demonstrate that the sequence is square-integrable. The regularity conditions of Brown (1971) are satisfied by the ergodic theorem, thereby establishing Lemma 2. □

Proof of Proposition 4. Under Assumption 3(b) and the definition of $\hat{\boldsymbol{\lambda}}$ in (21), we apply the mean value theorem to $\frac{\partial L_T(\psi)}{\partial \lambda}\Big|_{\psi=\hat{\psi}}$ to obtain

$$0 = \frac{\partial L_T(\boldsymbol{\psi}_0)}{\partial \lambda} + \Xi_T^{\theta}(\boldsymbol{\psi}^*)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \Xi_T^{\alpha}(\boldsymbol{\psi}^*)(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0)$$
$$+\Xi_T^{\sigma}(\boldsymbol{\psi}^*)(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0) + \Xi_T^{\lambda}(\boldsymbol{\psi}^*)(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0), \tag{A.11}$$

where

$$\Xi_T^{\theta}(\boldsymbol{\psi}^*) = \frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\theta'}\bigg|_{\psi=\psi^*}, \quad \Xi_T^{\alpha}(\boldsymbol{\psi}^*) = \frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\alpha'}\bigg|_{\psi=\psi^*},$$
$$\Xi_T^{\sigma}(\boldsymbol{\psi}^*) = \frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\sigma'}\bigg|_{\psi=\psi^*}, \quad \Xi_T^{\lambda}(\boldsymbol{\psi}^*) = \frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\lambda'}\bigg|_{\psi=\psi^*},$$

with $\boldsymbol{\psi}^*$ located between $\boldsymbol{\psi}_0$ and $\hat{\boldsymbol{\psi}}$. From the technique in the proof of Theorem 4.2 of Pedersen and Rahbek (2014), we obtain

$$\sqrt{T}(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}_0) = -\Xi_T^{\lambda}(\boldsymbol{\psi}^*)^{-1}\bigg[\Xi_T^{\theta}(\boldsymbol{\psi}^*)\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \Xi_T^{\alpha}(\boldsymbol{\psi}^*)\sqrt{T}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0)$$
$$+\Xi_T^{\sigma}(\boldsymbol{\psi}^*)\sqrt{T}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}_0) + \sqrt{T}\frac{\partial L_T(\boldsymbol{\psi}_0)}{\partial\lambda}\bigg],$$

and

$$\Xi_T^{\lambda}(\boldsymbol{\psi}^*)^{-1} \xrightarrow{p} (\Xi_0^{\lambda})^{-1}, \quad \Xi_T^{\lambda}(\boldsymbol{\psi}^*)^{-1}\Xi_T^{\theta}(\boldsymbol{\psi}^*) \xrightarrow{p} (\Xi_0^{\lambda})^{-1}\Xi_0^{\theta},$$
$$\Xi_T^{\lambda}(\boldsymbol{\psi}^*)^{-1}\Xi_T^{\alpha}(\boldsymbol{\psi}^*) \xrightarrow{p} (\Xi_0^{\lambda})^{-1}\Xi_0^{\alpha}, \quad \Xi_T^{\lambda}(\boldsymbol{\psi}^*)^{-1}\Xi_T^{\sigma}(\boldsymbol{\psi}^*) \xrightarrow{p} (\Xi_0^{\lambda})^{-1}\Xi_0^{\sigma}, \tag{A.12}$$

where

$$\Xi_0^{\theta} = E\left[\frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\theta'}\bigg|_{\psi=\psi_0}\right], \quad \Xi_0^{\alpha} = E\left[\frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\alpha'}\bigg|_{\psi=\psi_0}\right],$$
$$\Xi_0^{\sigma} = E\left[\frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\sigma'}\bigg|_{\psi=\psi_0}\right], \quad \Xi_0^{\lambda} = E\left[\frac{\partial^2 L_T(\psi)}{\partial\lambda\partial\lambda'}\bigg|_{\psi=\psi_0}\right]. \tag{A.13}$$

Since it is straightforward to show that the information matrix for the quasi-log-likelihood function is block diagonal under the assumption of a symmetric distribution, as discussed in Remark 3.7 of Francq and Zakoïan (2004), we obtain $\Xi_0^{\theta} = O$ and $\Xi_0^{\alpha} = O$. The result with Lemma 1 and 2 and the Slutzky theorem produces the asymptotic normality of the estimator. □

References

Arellano, M., 1987. Computing robust standard errors for within-group estimators. Oxford Bulletin of Economics and Statistics 49, 431–434.
Arellano, M., 2003. Panel Data Econometrics. Oxford University Press, Oxford, U.K..
Bickel, P.J., Levina, E., 2008a. Covariance regularization by thresholding. Annals of Statistics 36, 2577–2604.
Bickel, P.J., Levina, E., 2008b. Regularized estimation of large covariance matrices. Annals of Statistics 36, 199–227.
Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. Econometric Reviews 11, 143–172.
Bouras, C., Christou, C., Gupta, R., Suleman, T., 2019. Geopolitical risks, returns, and volatility in emerging stock markets: Evidence from a panel GARCH model. Emerging Markets Finance and Trade 55, 1841–1856.
Boussama, F., Fuchs, F., Stelzer, R., 2011. Stationarity and geometric ergodicity of BEKK multivariate GARCH models. Stochastic Processes and their Applications 121, 2331–2360.
Brown, B., 1971. Martingale central limit theorems. Annals of Mathematical Statistics 42, 59–66.
Cermeño, R., Grier, K.B., 2006. Conditional heteroskedasticity and cross-sectional dependence in panel data: An empirical study of inflation uncertainty in the g7 countries. In: Baltagi, B.H. (Ed.), Panel data econometrics. Elsevier, Amsterdam, pp. 259–277.
Cermeño, R., Sanin, M.E., 2015. Are flexible exchange rate regimes more volatile? panel GARCH evidence for the g7 and latin america. Review of Development Economics 19, 297–308.
Cho, H., Korkas, K.K., 2021. High-dimensional GARCH process segmentation with an application to value-at-risk. In: to appear in econometrics and statistics.
Conrad, C., Mammen, E., 2016. Asymptotics for parametric GARCH-in-mean models. Journal of Econometrics 194, 319–329.
Deniz, P., Stengos, T., Yazgan, M.E., 2020. Revisiting the link between output growth and volatility: Panel GARCH analysis. Empirical Economics doi:10.1007/s00181-020-01878-4.
Ding, Z., Engle, R.F., 2001. Large scale conditional covariance matrix modeling, estimation and testing. Academia Economic Papers 29, 157–184.
Drost, F.C., Nijman, T.E., 1993. Temporal aggregation of GARCH processes. Econometrica 61, 909–927.
Fiorentini, G., Calzolari, G., Panattoni, L., 1996. Analytic derivatives and the computation of GARCH estimates. Journal of Applied Econometrics 11, 399–417.
Francq, C., Horváth, L., Zakoïan, J.M., 2011. Merits and drawbacks of variance targeting in GARCH models. Journal of Financial Econometrics 9, 619–656.
Francq, C., Zakoïan, J.M., 2004. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. Bernoulli 10, 605–637.
Hafner, C.M., Preminger, A., 2009. On asymptotic theory for multivariate GARCH models. Journal of Multivariate Analysis 100, 2044–2054.

Han, H., Kristensen, D., 2014. Asymptotic theory for the QMLE in GARCH-x models with stationary and nonstationary covariates. Journal of Business & Economic Statistics 32, 416–429.
Hansen, C.B., 2007. Asymptotic properties of a robust variance estimator for panel data when t is large. Journal of Econometrics 141, 597–620.
Kiefer, N.M., 1980. Estimation of fixed effect models for time series of cross-sections with arbitrary intertemporal covariance. Journal of Econometrics 14, 195–202.
Kristensen, D., Linton, O., 2004. Consistent standard errors for target variance approach to GARCH estimation. Econometric Theory 20, 990–993.
Lee, J., 2010. The link between output growth and volatility: Evidence from a GARCH model with panel data. Economics Letters 106, 143–145.
Newey, W.K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R.F., McFadden, D. (Eds.), Handbook of econometrics, volume 4. Elsevier, Amsterdam, pp. 2111–2245.
Newey, W.K., West, K.M., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Nicholls, D.F., Pagan, A.R., 1983. Heteroskedasticity in models with lagged dependent variables. Econometrica 51, 1233–1242.
Nijman, T., Sentana, E., 1996. Marginalization and contemporaneous aggregation in multivariate GARCH processes. Journal of Econometrics 71, 71–87.
Pakel, C., Shephard, N., Sheppard, K., 2011. Nuisance parameters, composite likelihoods and a panel of GARCH models. Statistica Sinica 21, 307–329.
Pedersen, R.S., Rahbek, A., 2014. Multivariate variance targeting in the BEKK-GARCH model. Econometrics Journal 17, 24–55.
Ribeiro, P.P., Cermeño, R., Curto, J.D., 2017. Sovereign bond markets and financial volatility dynamics: Panel-GARCH evidence for six euro area countries. Finance Research Letters 21, 107–114.
Scherrer, W., Ribarits, E., 2007. On the parametrization of multivariate GARCH models. Econometric Theory 23, 464–484.
Silvennoinen, A., Teräsvirta, T., 2021. Consistency and asymptotic normality of the maximum likelihood estimators of a multiplicative time-varying smooth transition correlation GARCH model. In: to appear in econometrics and statistics.
Stock, J.H., Watson, M.W., 2008. Heteroskedasticity-robust standard errors for fixed effects panel data regression. Econometrica 76, 155–174.
Valera, H.G.A., Holmes, M.J., Hassan, G., 2017. Stock market uncertainty and interest rate behaviour: A panel GARCH approach. Applied Economics Letters 24, 732–735.
Vogelsang, T.J., 2012. Heteroskedasticity, autocorrelation, and spatial correlation robust inference in linear panel models with fixed-effects. Journal of Econometrics 166, 303–319.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.
White, H., 1994. Estimation, Inference and Specification Analysis. Cambridge University Press, New York.
Zhu, X., Pan, R., Li, G., Liu, Y., Wang, H., 2017. Network vector autoregression. Annals of Statistics 45, 1096–1123.