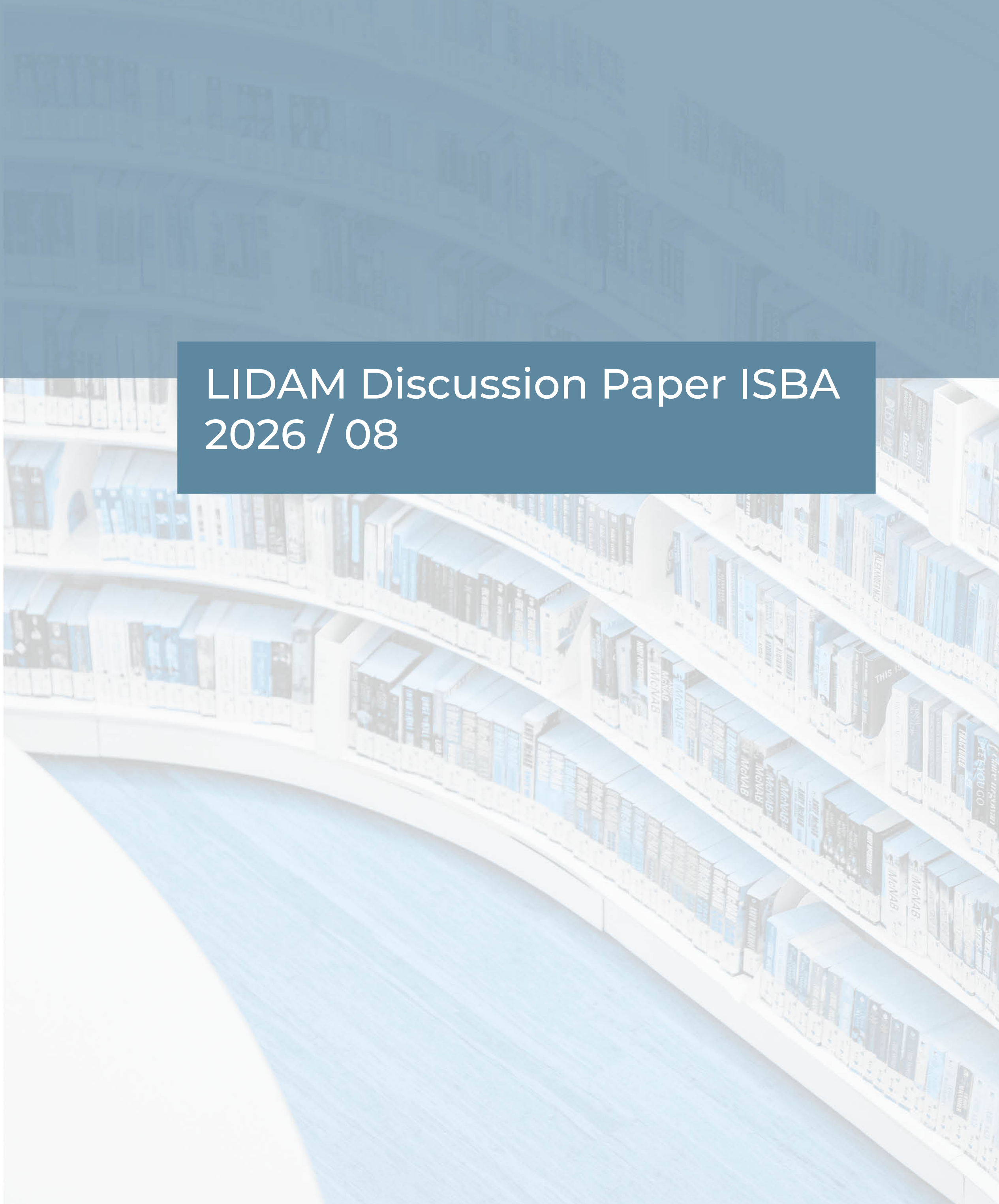
A ZERO INTERCEPT VEC MODEL

Christian M. Hafner, Arie Preminger



LIDAM Discussion Paper ISBA
2026 / 08

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication

A Zero Intercept Vec model

Christian M. Hafner* Arie Preminger†

March 28, 2026

Abstract

This paper introduces a multivariate volatility model that is characterized by nonstationarity irrespective of the parameters. The model is motivated by the multivariate GARCH model in VEC form, setting the intercept term to zero. We first discuss the conditions required for a positive definite conditional variance matrix. For the special case of a diagonal parameter matrix, we derive the conditions for stability of trajectories, meaning that the processes do not diverge to infinity or to zero almost surely. We then develop the asymptotic theory for maximum likelihood estimation, and propose a test of the null hypothesis of a zero Lyapunov exponent, i.e. stability. In a simulation study we demonstrate the good performance of the estimator and the test in finite samples.

Keywords: Volatility, non-stationarity, multivariate GARCH, asymptotic theory, maximum likelihood

*Corresponding author, UCLouvain, LIDAM, Louvain-la-Neuve, Belgium, e-mail `christian.hafner@uclouvain.be`

†Academic College of Ramat Gan, Ramat Gan, Israel. Email : ariepr25@iac.ac.il

1 Introduction

This paper focuses on an extension of a general multivariate GARCH model, known as the VEC model, originally proposed by Bollerslev, Engle, and Wooldridge (1988). The VEC model is a flexible framework that includes several important specifications, such as the BEKK model and various factor GARCH models that are frequently used in practice, see e.g. Bauwens et al. (2006), Silvennoinen and Teräsvirta (2009) and Bauwens et al. (2012).

In this paper, we propose a VEC model with zero intercept (ZI-VEC) and discuss its properties. For the univariate case such models were considered by Hafner and Preminger (2015), Li , Zhang and Zhu (2018) and Li and Zhu (2020). Suppressing the intercept term is not innocuous as it renders the stochastic process non-stationary, irrespective of the parameter values. It turns out, however, that a stable version can be obtained which does not diverge to zero or infinity almost surely, characterized by a zero Lyapunov exponent.

We obtain a closed form expression for the quasi-maximum likelihood estimator (QMLE), for which we establish the strong consistency and asymptotic normality under conditions that assume the existence of the fourth order moment of the innovation term. It should be noted that for the multivariate VEC GARCH, existing work requires the sixth order moment of the observed process which can be quite restrictive due to fat tails, see e.g. Hafner and Preminger (2009). Our results are obtained without a stability assumption of the underlying process. Furthermore, we develop a test for the stability hypothesis of a zero Lyapunov exponent. A simulation study then documents the good performance of both the QMLE and the stability test in finite samples.

The remainder of this paper is organized as follows: Section 2 outlines the theoretical framework of the ZI-VEC model. Section 3 presents simulation studies to assess the performance of the QMLE and the stability test, and Section 4 concludes. Proofs are delegated to an appendix.

The following notation will be used throughout. The norm of a matrix $A = (a_{ij})$ is defined as the Frobenius norm, $\|A\| = \sqrt{\sum_{i,j} a_{ij}^2}$. The Kronecker product is denoted by $\otimes$, the symbol $\to_d$ and $\to_{as}$ denote convergence in distribution and almost surely. respectively.

2 The model

Consider the vector ARCH (1) model for the n-dimensional vector time series $\{y_t\}, t = 0, 1, 2, \ldots, T$,

$$y_t = H_t^{1/2}\varepsilon_t \tag{1}$$

where ε_t is a sequence of i.i.d. random vectors with mean zero and variance-covariance matrix the identity matrix I_n. The conditional covariance matrix is given by

$$h_t := \mathrm{vech}(H_t) = \widetilde{A}\mathrm{vech}(y_{t-1}y_{t-1}^{\intercal}),\ \ \widetilde{A} \in \mathbb{R}^{N\times N} \tag{2}$$

where $N = n(n+1)/2$. This model is motivated by the general multivariate GARCH model proposed by Bollerslev et al. (1988), usually called the VEC model while in our model we have set the intercept to zero (ZI-VEC).

2.1 Positivity conditions

We need to ensure that H_t is positive definite with probability one (wp1). Conditions can be obtained using the equivalent representation:

$$H_t = (I_N \otimes y_{t-1}^{\intercal})B(I_N \otimes y_{t-1})$$

for some $n^2 \times n^2$ parameter matrix B. The link between $\widetilde{A}$ and B is the following. Each row of $\widetilde{A}$ has N components and corresponds to a particular element (i,j) of H_t. Let us call this vector $\widetilde{A}_{ij}$. The model for $H_{ij,t}$ calculates the inner product of this vector with $\mathrm{vech}(y_{t-1}y_{t-1}^{\intercal})$, which has typical elements $y_{t-1,k}y_{t-1,l}$. The coefficients in the inner product $\widetilde{A}_{ij}\mathrm{vech}(y_{t-1}y_{t-1}^{\intercal})$ that correspond to the term $y_{t-1,k}y_{t-1,l}$ will be called $\widetilde{A}_{ij,kl}$. Now, define the $(n \times n)$ matrix

$$A_{ij}^* = \begin{pmatrix} \widetilde{A}_{ij,11} & \widetilde{A}_{ij,12}/2 & \ldots & \widetilde{A}_{ij,1n}/2 \\ \widetilde{A}_{ij,12}/2 & \widetilde{A}_{ij,22} & \ldots & \widetilde{A}_{ij,2n}/2 \\ \vdots & & \ddots & \\ \widetilde{A}_{ij,1n}/2 & \widetilde{A}_{ij,2n}/2 & \ldots & \widetilde{A}_{ij,nn} \end{pmatrix}$$

Then, the equivalence between the two representations is given by the link

$$B = \begin{pmatrix} A_{11}^* & A_{12}^* & \ldots & A_{1n}^* \\ A_{12}^* & A_{22}^* & \ldots & A_{2n}^* \\ \vdots & & \ddots & \\ A_{1n}^* & A_{2n}^* & \ldots & A_{nn}^* \end{pmatrix}$$

For an example with $n = 2$, see Gouriéroux (1997). From the Kronecker representation, it follows that H_t is positive definite iff B is positive definite. Hence, for the vec model, H_t is positive definite iff the matrix B of the corresponding (and equivalent) Kronecker representation is positive definite.

2.2 Diagonality restriction

In the following, we will assume that $\widetilde{A} = \text{diag}(a_1, a_2, ..., a_N)$ is a diagonal matrix. Under this assumption, the ZI-VEC model can be rewritten as

$$H_t = D_t A D_t \tag{3}$$

where $D_t = \text{diag}(y_{1,t-1}, y_{2,t-1}, ..., y_{n,t-1})$ and $A = \text{vech}^{-1}(a_1, a_2, \ldots, a_N)$ is the $n \times n$ matrix obtained by inverting the vech operator. Now, H_t is positive definite wp1 iff A is positive definite.

In model (1), we can define the square root of H_t as $H_t^{1/2} = D_t A^{1/2}$, where $A^{1/2}$ is any decomposition of A satisfying $A^{1/2}(A^{1/2})^\intercal = A$. For example, we may obtain $A^{1/2}$ from the Cholesky decomposition. Then, from (1) and (3) we have

$$y_t = D_t A^{1/2} \varepsilon_t = D_t \eta_t \tag{4}$$

where $\eta_t := A^{1/2}\varepsilon_t$ and $\mathbb{V}\text{ar}(y_t \mid y_{t-1}) = H_t$. Thus, letting $y_t = (y_{1,t}, \ldots, y_{n,t})^\intercal$ and $\eta_t = (\eta_{1,t}, \ldots, \eta_{n,t})^\intercal$, the model can be written elementwise as

$$y_{i,t} = y_{i,t-1}\eta_{i,t}, \qquad i = 1, \ldots, n \tag{5}$$

Assume that $P(\varepsilon_i = 0) = 0$ and that the process starts at a finite vector y_0 with $y_{i,0} \neq 0$ almost surely for each i. Define the Lyapunov exponent $\gamma_i = \mathbb{E}[\log(\eta_{i,t}^2)]$. The following theorem gives the sample path properties of $y_{i,t}^2$.

Theorem 1 *If $0 < \mathbb{V}ar(\log(\eta_{i,t}^2))$ for each $i = 1, ..., n$, then almost surely as $t \to \infty$*

a) If $\gamma_i > 0$, then $y_{i,t}^2 \to \infty$

b) If $\gamma_i < 0$, then $y_{i,t}^2 \to 0$

c) If $\gamma_i = 0$, then $\liminf y_{i,t}^2 = 0, \ \limsup y_{i,t}^2 = \infty$

Let $\gamma_0 = (\gamma_1, \gamma_2, ..., \gamma_n)$. Theorem 1 implies that when $\gamma_0 = 0$, we can say that the process in (1) is "stable" (but non-stationary) in the sense that the observed process has a non-degenerated distribution. In contrast, if $\gamma_0 \neq 0$, we can say that the process is "unstable", in the sense that some components of $y_{i,t}$ will almost surely converge to zero or diverge to infinity. It is interesting to note that if $\eta_t \sim \mathrm{N}(0, A_0)$ the stability condition does not depend on the off diagonal elements of A_0, thus the process is stable if for all i, $A_{ii0} = \frac{1}{2}\exp(-\Psi(\frac{1}{2})) \approx 3.56$, where $\Psi(\cdot)$ is the Euler psi function.

2.3 Conditional correlation

In the diagonal model, the conditional correlation between $y_{i,t}$ and $y_{j,t}$, denoted by $\rho_{ij,t}$ is given by:

$$\rho_{ij,t} = \frac{H_{ij,t}}{\sqrt{H_{ii,t}H_{jj,t}}} = \frac{A_{ij}y_{i,t-1}y_{j,t-1}}{\sqrt{A_{ii}y_{i,t-1}^2 A_{jj}y_{j,t-1}^2}} = \frac{A_{ij}}{\sqrt{A_{ii}A_{jj}}} \cdot \mathrm{sign}(y_{i,t-1}y_{j,t-1}) \tag{6}$$

so it is a constant up to the sign of the cross-product of y_{t-1}. Thus, from equation (6), we observe that the conditional correlation is not purely constant; it varies with the sign of the cross-product of lagged returns. This indicates that even under the diagonality restriction on $\bar{A}$, the model permits a time-varying correlation structure through the interaction of lagged values. Consequently, the ZI-VEC specification is capable of capturing certain dynamic interdependencies beyond a simple diagonal ARCH structure. For an intuitive explanation, note that from (3) if $n = 2$ we have

$$H_t = \begin{pmatrix} a_{11}y_{1,t-1}^2 & a_{12}y_{1,t-1}y_{2,t-1} \\ a_{12}y_{1,t-1}y_{2,t-1} & a_{22}y_{2,t-1}^2 \end{pmatrix}$$

and if $a_{12} \neq 0$ we allow for cross-correlation.

2.4 Estimation

The negative log likelihood function derived from the Gaussian distribution is given by

$$L_T = T^{-1}\sum_{t=2}^{T} \ell_t(\theta), \quad \ell_t(\theta) = \frac{1}{2}\left(\log|H_t(\theta)| + y_t^\intercal H_t^{-1}(\theta)y_t\right) \tag{7}$$

and $\theta \in \Theta \subset \mathbb{R}^N$ a compact set such that the set of induced matrices, $\mathcal{A} = \{A = \mathrm{vech}^{-1}(\theta), \theta \in \Theta\}$ is a compact subset of the space of real symmetric positive definite matrices. The QMLE $\hat{\theta}_T$ of $\theta_0 = \mathrm{vech}(A_0)$, the true parameter, is defined as the minimizer of (7) or equivalently $\hat{\theta}_T = \arg\min_{\theta\in\Theta} L_T(\theta)$. Note that this is not the true likelihood unless ε_t has a Gaussian distribution. In the case of a diagonality restriction, there is a closed form estimator, as $\eta_t = D_t^{-1} y_t$ can be calculated from the data with $\mathbb{E}(\eta_t) = 0$ and $\mathbb{V}\mathrm{ar}(\eta_t) = A$, so that the QMLE is given by

$$\widehat{A} = \frac{1}{T} \sum_{t=1}^{T} \eta_t \eta_t^\intercal,$$

see the appendix.

The following theorem states the consistency and asymptotic normality of $\hat{\theta}_T$. We make the following assumptions.

Assumption 1: The random vector $\{\varepsilon_t\}$ has a continuous density function in some open neighborhood around the origin, with $\mathbb{E}(\varepsilon_t) = 0$ and $\mathbb{V}\mathrm{ar}(\varepsilon_t) = I_n$.

Assumption 2: $\mathbb{E}\|\varepsilon_t\|^4 < \infty$ and $\theta_0 \in \Theta$ is an interior point.

Theorem 2 *Under Assumptions 1-2,*

(i) $\hat{\theta}_T \xrightarrow{a.s.} \theta_0$

(ii) $\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, J)$, *where* $J = 2D_n^+(A_0 \otimes A_0)(D_n^+)^\intercal + D_n^+ \kappa (D_n^+)^\intercal$

where D_n^+ *is the Moore-Penrose inverse of the duplication matrix* D_n *and* κ *is the* $n^2 \times n^2$ *matrix of fourth cumulants of* η_t*.*

Our results are independent of the value of γ_0. Thus, regardless if the underlying process is stable or not, we will get a strongly consistent and asymptotically normal MLE with the same covariance matrix. Further, note that if the error term, ε_t is Gaussian (or symmetric) then $\kappa = 0$ and the covariance matrix simplifies to V. Explicit expressions for κ can be obtained for specific non-Gaussian distributions. For example, if η_t has an elliptically symmetric distribution, meaning that $\xi_t = A_0^{-1/2} \eta_t$ has a spherical distribution with mean zero and variance the identity matrix, then we can show that

$$\kappa = (c-1)\left[(I_{n^2} + K_{nn})(A_0 \otimes A_0) + \mathrm{vec}(A_0)\mathrm{vec}(A_0)^\intercal\right]$$

where $c := \mathbb{E}[\xi_{i,t}^2 \xi_{j,t}^2], i \neq j$. In the Gaussian case, $c = 1$ and $\kappa = 0$. See, for example, Fang et al. (1989). In the general case, an estimator of κ is given by

$$\widehat{\kappa} = \frac{1}{T} \sum_{t=1}^{T} v_t v_t^\intercal - (I_{n^2} + K_{nn})(\hat{A} \otimes \hat{A}) - \mathrm{vec}(\hat{A})\mathrm{vec}(\hat{A})^\intercal,$$

where $v_t = \mathrm{vec}(\eta_t \eta_t^\intercal)$.

In the proof, the diagonality restriction allows us to express the likelihood function in terms of η_t. This simplification facilitates the derivation of the asymptotic results. Extending the analysis to non-diagonal $\tilde{A}$ matrices is left for further research.

Now, let $d_{i,t} = y_{i,t}^2/y_{i,t-1}^2$, $d_t = (d_{1,t}, \ldots, d_{n,t})^\intercal$, and from (5) note that $\mathbb{E}(\log d_t) = \gamma_0$, where the log is applied elementwise. Hence, consistent estimates of γ_0 and $\Omega = \mathbb{V}\mathrm{ar}(\log(d_t))$ are given by $\bar{\gamma}_T = T^{-1}\sum_{t=1}^{T}\log(d_t)$ and $\Omega_T = T^{-1}\sum_{t=1}^{T}(\log(d_t) - \bar{\gamma}_T)(\log(d_t) - \bar{\gamma}_T)^\intercal$, respectively. The following result gives the asymptotic distribution of $\bar{\gamma}_T$.

Theorem 3 *If Ω is positive definite, then $\sqrt{T}(\bar{\gamma}_T - \gamma_0) \to_d N(0, \Omega)$.*

We can use this result to test if $\gamma_0 = 0$, which is equivalent to testing the null hypothesis that the process is stable.

In particular, we construct the test statistic

$$G = T\, \bar{\gamma}_T^\intercal\, \Omega_T^{-1}\, \bar{\gamma}_T$$

which, under the null hypothesis $H_0 : \gamma_0 = 0$, follows asymptocially a chi-square distribution with n degrees of freedom.

3 Simulation study

We simulate the model using various parameter configurations. In the first set of results, reported in Table 1, we assume a stable bivariate ZI-VEC model with Gaussian innovations, parameterized by three parameters, a_{11}, a_{12} and a_{22}. For stability, we set $a_{11} = a_{22} = 3.5621$, and $a_{12} \in \{0, 1, 2, 3\}$. The processes are simulated with two different sample sizes, $T = 100$ and $T = 500$, and the number of independent Monte Carlo replications is set to 10,000. We report the estimation results using the closed form expression for the QMLE. In all cases, the Monte Carlo means of parameter estimates are close to the true parameters, with standard deviations that decrease roughly with a factor of $5^{-1/2}$ when increasing T from 100 to 500, as expected. While the standard deviations of $\widehat{a_{11}}$ and $\widehat{a_{22}}$ seem constant for the different choices of a_{12}, the standard deviation of $\widehat{a_{12}}$ clearly increases for larger values of a_{12}. Note that, since we used the Cholesky decomposition to obtain $A^{1/2}$, the variance of η_{1t} is equal to a_{11} and does not depend on a_{12} and a_{22}. Therefore, the estimate of a_{11} and its standard deviation for a given sample size are the same across the different parameter settings.

Table 1 also reports the proportion of rejections of $H_0 : \gamma_0 = 0$ using the test statistic G at nominal level 5%. The empirical level, denoted p, approaches the nominal level in

	$T = 100$		$T = 500$		$T = 100$		$T = 500$	
	$a_{12} = 0$				$a_{12} = 2$			
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
$\hat{a_{11}}$	3.5707	0.5000	3.5629	0.2248	3.5707	0.5000	3.5629	0.2248
$\hat{a_{12}}$	0.0023	0.3558	0.0014	0.1587	2.0067	0.4076	2.0016	0.1820
$\hat{a_{22}}$	3.5640	0.5054	3.5639	0.2235	3.5682	0.5042	3.5649	0.2244
p	0.0613		0.0518		0.0651		0.0536	
	$a_{12} = 1$				$a_{12} = 3$			
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
$\hat{a_{11}}$	3.5707	0.5000	3.5629	0.2248	3.5707	0.5000	3.5629	0.2248
$\hat{a_{12}}$	1.0046	0.3697	1.0016	0.1648	3.0085	0.4634	3.0014	0.2077
$\hat{a_{22}}$	3.5658	0.5050	3.5646	0.2243	3.5708	0.5027	3.5645	0.2242
p	0.0654		0.0530		0.0637		0.0541	

Table 1: QMLE estimation results for the bivariate stable ZI-VEC model with the parameters a_{11}, a_{12} and a_{22}. Reported are the means and standard deviations of parameter estimates for $10,000$ independent Monte Carlo replications. p gives the proportion of rejections of $H_0 : \gamma_0 = 0$ using the test statistic G and nominal level 5%.

all cases. The test is slightly oversized for $T = 100$, but the empirical level is close to the nominal level for $T = 500$.

In the second set of results, we assume an unstable ZI-VEC process by setting $a_{11} = 3.5621$, $a_{12} = 1$, and $a_{22} \neq 3.5621$. We consider a positive Lyapunov exponent, $\gamma_0 > 0$, by setting $a_{22} = 3.75$ and $a_{22} = 4$, and a negative Lyapunov exponent, $\gamma_0 < 0$, for $a_{22} = 3.25$ and $a_{22} = 3$. Estimation results are similar to the stable case, with all parameter estimates close to the true values already for $T = 100$. The proportion of rejections of $H_0 : \gamma_0 = 0$ using the G-test now increases as the distance of a_{22} from the stability value 3.5621 increases, and as the sample size increases from $T = 100$ to $T = 500$. For example, for the case $a_{22} = 3$ and $T = 500$, the empirical power is 30.15 %. The test shows good power properties even for moderate deviations from stability, and moderate sample sizes. Further results, not reported for space constraints, suggest that the power converges to one as the sample size increases and/or the parameter moves away from the null hypothesis.

	$T = 100$		$T = 500$		$T = 100$		$T = 500$	
	$a_{22} = 3.75$				$a_{22} = 3.25$			
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
$\hat{a_{11}}$	3.5707	0.5000	3.5629	0.2248	3.5707	0.5000	3.5629	0.2248
$\hat{a_{12}}$	1.0047	0.3787	1.0016	0.1688	1.0045	0.3544	1.0015	0.1580
$\hat{a_{22}}$	3.7538	0.5316	3.7526	0.2362	3.2534	0.4607	3.2523	0.2047
p	0.0755		0.0845		0.0631		0.1064	
	$a_{22} = 4$				$a_{22} = 3$			
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
$\hat{a_{11}}$	3.5707	0.5000	3.5629	0.2248	3.5707	0.5000	3.5629	0.2248
$\hat{a_{12}}$	1.0048	0.3902	1.0017	0.1739	1.0044	0.3416	1.0015	0.1523
$\hat{a_{22}}$	4.0040	0.5671	4.0028	0.2519	3.0032	0.4253	3.0022	0.1890
p	0.1042		0.1875		0.0862		0.3015	

Table 2: QMLE estimation results for the bivariate ZI-VEC model. The model is unstable with $\gamma_0 > 0$ for $a_{22} = 3.75$ and $a_{22} = 4$, and $\gamma_0 < 0$ for $a_{22} = 3.25$ and $a_{22} = 3$. Reported are the means and standard deviations of parameter estimates for $10,000$ independent Monte Carlo replications. p gives the proportion of rejections of $H_0 : \gamma_0 = 0$ using the test statistic G and nominal level 5%.

4 Conclusions

In this paper, we have developed the asymptotic theory of the quasi maximum likelihood estimator of a nonstationary multivariate GARCH model obtained by setting the intercept term to zero. We characterized the conditions for stability of the processes and proposed a test for stability. A simulation study has illustrated the good finite sample performance of the estimator and the stability test. Extensions of the proposed model are possible in various directions, for example, by including autoregressive terms in the volatility equation, considering non-Gaussian distributions for the innovation term, or adding a risk premium term to allow for non-zero mean returns. These are left for future research. Future empirical work will also have to show the usefulness of the model when applied to multivariate financial times chracterized by erratic and nonstationary behavior.

Appendix

Proof of Theorem 1: From (5) we note that $y_{i,t}^2 = y_{i,0}^2 \prod_{k=1}^t \eta_{i,k}^2$, hence

$$\log(y_{i,t}^2) = \log(y_{i,0}^2) + \sum_{k=1}^{t} \log(\eta_{i,k}^2) \tag{8}$$

The result then follows by applying the law of large numbers to the partial sums in (8) and using arguments analogous to those in Theorem 1 of Nelson (1990).

Proof of Theorem 2:
(i) From (4) and Assumption 1 we have $\eta_t = D_t^{-1} y_t \sim \text{i.i.d}(0, A_0)$. Thus, from (3) we get

$$y_t^\intercal H_t^{-1}(\theta) y_t = y_t^\intercal (D_t A D_t)^{-1} y_t = y_t^\intercal (D_t^{-1} A^{-1} D_t^{-1}) y_t = (D_t^{-1} y_t)^\intercal A^{-1} D_t^{-1} y_t = \eta_t^\intercal A^{-1} \eta_t$$

and the determinant simplifies as follows:

$$|H_t(\theta)| = |D_t A D_t| = |D_t| \cdot |A| \cdot |D_t| = \left(\prod_{i=1}^{n} y_{i,t-1} \right)^2 |A|.$$

Substituting this into (7) yields

$$L_T(\theta) = C + T^{-1} \sum_{t=2}^{T} \bar{\ell}_t(\theta), \quad \bar{\ell}_t(\theta) = \frac{1}{2}(\log |A| + \eta_t^\intercal A^{-1} \eta_t) \tag{9}$$

where $C = \frac{1}{T}\sum_{t=1}^{T} \log|\prod_{i=1}^{n} y_{i,t-1}|$. Taking the first derivative of (9) with respect to A yields the QMLE

$$\hat{A} = \frac{1}{T}\sum_{t=1}^{T} \eta_t \eta_t^{\intercal} \tag{10}$$

and Assumption 1 and the strong law of large numbers imply that $\hat{\theta}_T \to_{a.s.} \theta_0$.

(ii) Let $v_t = \text{vec}(\eta_t \eta_t^{\intercal})$. The variance of v_t for an iid process is

$$\mathbb{V}\text{ar}(v_t) = (I_{n^2} + K_{nn})(A_0 \otimes A_0) + \boldsymbol{\kappa} \tag{11}$$

where K_{nn} is the commutation matrix and $\boldsymbol{\kappa}$ is the $n^2 \times n^2$ matrix of fourth cumulants of η_t. In the Gaussian case $\boldsymbol{\kappa} = 0$ and the variance reduces to the first term $(I_{n^2} + K_{nn})(A_0 \otimes A_0)$ see e.g. Theorem 10.2 of Magnus (1988). Further, let $w_t = \text{vech}(\eta_t \eta_t^{\intercal})$ and using the duplication matrix D_n we have that $v_t = D_n w_t$ which implies that

$$J = \mathbb{V}\text{ar}(w_t) = D_n^{+} \mathbb{V}\text{ar}(v_t)(D_n^{+})^{\intercal} \tag{12}$$

where $D_n^{+} = (D_n^{T} D_n)^{-1} D_n^{\intercal}$.

For any $n \times n$ matrix B, we have $(I_{n^2} + K_{nn})(B \otimes B) = (B \otimes B)(I_{n^2} + K_{nn})$ by Theorem 3.11 of Magnus (1988). Therefore,

$$\begin{aligned} J &= D_n^{+}((A_0 \otimes A_0)(I_{n^2} + K_{nn}) + \boldsymbol{\kappa})(D_n^{+})^{\intercal} \\ &= D_n^{+}(A_0 \otimes A_0)(I_{n^2} + K_{nn})(D_n^{+})^{\intercal} + D_n^{+}\boldsymbol{\kappa}(D_n^{+})^{\intercal} \\ &= D_n^{+}(A_0 \otimes A_0)(D_n^{+}(I_{n^2} + K_{nn}))^{\intercal} + D_n^{+}\boldsymbol{\kappa}(D_n^{+})^{\intercal} \end{aligned} \tag{13}$$

Using the property that $I_{n^2} + K_{nn} = 2D_n D_n^{+}$, see e.g. Lütkepohl (1996, p.123), we get

$$J = 2D_n^{+}(A_0 \otimes A_0)(D_n^{+})^{\intercal} + D_n^{+}\boldsymbol{\kappa}(D_n^{+})^{\intercal}. \tag{14}$$

Assumptions 1 and 2 are standard regularity conditions which imply that J is finite and positive definite, and the desired result follows from the multivariate Lindeberg-Levy central limit theorem.

Proof of Theorem 3: The desired result follows by applying the Cramer-Wold device with the Lindeberg-Levy central limit theorem.

References

Bauwens, L., Hafner, C. and Laurent, S. (2012). *Handbook of Volatility Models and Their Applications*, New York, NY: Wiley.

Bauwens, L. Laurent, S. and Rombouts, J.V.K. (2006). Multivariate GARCH models: a survey, *Journal of Applied Econometrics*, 21, 79-109.

Bollerslev T, R. Engle and J. Wooldridge (1988). A capital asset pricing model with time varying covariances. *Journal of Political Economy*, 96, 116-131.

Fang, K.T., Kotz, S., Ng, K.W. (1989). *Symmetric Multivariate and Related Distributions*, Monographs on Statistics and Applied Probability 36. London: Chapman & Hall.

Gouriéroux, C. (1997). *ARCH Models and Financial Applications*, Springer Series in Statistics.

Hafner C. M. and Preminger A. (2009). On asymptotic theory for multivariate GARCH models. *Journal of Multivariate Analysis*, 100, 2044-2054.

Hafner C. M. and Preminger A. (2015). An ARCH model without intercept. *Economics Letters*, 129, 13-17.

Horn, R.A. and Johnson, C.R. (1991). *Topics in Matrix Analysis*, Cambridge University Press.

Li D., Zhang, X. and Zhu, K. (2018). The ZD-GARCH model: A new way to study heteroscedasticity, *Journal of Econometrics*, 202, 1-17.

Li D. and Zhu, K. (2020). Inference for asymmetric exponentially weighted moving average models, *Journal of Time Series Analysis*, 41, 154-162.

Lütkepohl, H. (1996). *Handbook of Matrices*, New York, NY: Wiley.

Magnus, J. (1988). *Linear Structures*, London: Griffith.

Magnus, J. and Neudecker, H. (2019). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, New York, NY: Wiley.

Nelson D.B. (1990). Stationarity and Persistence in the GARCH(1,1) Model, *Econometric Theory*, 6(3), 318-334.

Silvennoinen, A., Teräsvirta, T. (2009). Multivariate GARCH Models. In: Mikosch, T., Kreiss, JP., Davis, R., Andersen, T. (eds), *Handbook of Financial Time Series.* Berlin, Heidelberg: Springer.