Multivariate Testing for Fractional Integration under Conditional Heteroskedasticity

Marina Balboa\textsuperscript{a}, Paulo M.M. Rodrigues\textsuperscript{b}, Antonio Rubia\textsuperscript{a} and A.M. Robert Taylor\textsuperscript{c}
\textsuperscript{a} University of Alicante
\textsuperscript{b} Banco de Portugal and Nova SBE, Universidade Nova de Lisboa
\textsuperscript{c} University of Essex

\textbf{Work in Progress.}
This version: February 28, 2018

\section*{Abstract}

This paper introduces a new testing approach to detect fractional integration in a multivariate time series context. We build under fairly general assumptions. These allow stationary short-term errors to be driven by shocks or innovations that can exhibit time-varying conditional heteroskedasticity on a martingale difference sequence basis. Furthermore, we do not require the order of fractional integration to lie in a particular region, thereby allowing for both stationary and nonstationary dynamics, nor impose any particular assumption on the distribution of the data. To the best of our knowledge, none of the existing methods in the extant literature achieve this degree of flexibility. Monte Carlo analysis shows that the test ensures correct size performance to detect the order of integration in the presence of GARCH-type dependencies and heavy-tailed conditional distributions. We implement the new procedure to jointly infer the order of fractional integration in trading volume and volatility measures in a sample of major stocks from the U.S. market. The evidence from this analysis suggest that volatility tends to be more persistent than trading volume.

\textbf{Keywords:} Long memory, realized volatility, trading volume.
\textbf{JEL classification:} C20, C22.
1 Introduction

Long-range dependence, also known as long memory or fractional integration, is an important stylized feature of many economic and financial time series. It manifests itself through long-lasting statistically significant empirical autocorrelations and impulse-response coefficients that distinctively decay to zero at an hyperbolic rate. Fractionally integrated models generalize the standard ARIMA setting to accommodate this form of persistence and have received considerable attention in the extant literature; see e.g. Baillie (1996) and Robinson (2003) for surveys on the applications of these models in economics. To formally detect fractional integration, a number of alternative testing procedures adopting an univariate perspective, both in the frequency and time domains, have been proposed; see, for instance, Agiakloglou and Newbold (1994), Robinson (1994), Tanaka (1999), Breitung and Hassler (2002), Nielsen (2005), Demetrescu et al. (2008) and Hassler et al. (2009, 2016).

In this paper, we introduce a new testing approach to address multivariate fractional integration within the class of fractionally integrated vector autoregressive (FIVAR) models. This method can also be used to construct confidence sets that include the true values of the long-memory coefficients with a given asymptotic coverage level. Multivariate testing is naturally intended to address hypotheses that involve the degree of persistence of different variables. For instance, there has been great interest in theoretical and empirical finance to understand the link between trading volume and return volatility. Several papers have analyzed if the long-run dynamics of these variables share a common order of fractional integration, a premise that can theoretically be founded in a long-run interpretation of the so-called mixture of distributions hypothesis (MDH) introduced by Clark (1973). However, the evidence supporting the existence of such commonalities is mixed; see, for instance, Bollerslev and Jubinski (1999), Lobato and Velasco (2000), Luu and Martens (2003) and Fleming and Kirby (2011). Similarly, Andersen et al. (2003) analyzed the long-run dynamics of the realized volatility of several exchange rates reporting comparable estimates of the order of fractional integration. The authors addressed the hypothesis of common fractional integration using a multivariate extension of the GPH estimator developed by Robinson (1995), finding supportive evidence. While testing for a common order of integration is a possibility in our context, our testing approach is more
general and can be used to address joint and individual hypotheses involving het
erogenous orders of integration. Multivariate testing can enhance efficiency over uni
variate testing because it explicitly acknowledges the existence of endogenous cross-depen
dencies in the variables analyzed. Exploiting such patterns empirically may reduce the variabil
ity in the estimation errors and, hence, improve the efficiency in estimation and testing.

More specifically, we propose a parametric multivariate Lagrange multiplier (LM) test
statistic in the time domain which extends the multivariate regression-based setting in
Breitung and Hassler (2002) under fairly general assumptions. Under certain conditions,
our test is asymptotically equivalent to the multivariate LM test in Nielsen (2004,2005)
and, consequently, it can be shown to be asymptotically efficient under Gaussian inno
vations, although we remark that we do not impose such assumption in our test. Our
approach can also be related to the class of frequency-domain methods for detecting mul
tivariate fractional integration in the FIVAR and, more generally, the FIVARMA setting,
including, among others, Robinson (1995), Lobato and Robinson (1998), Lobato (1999),

A key contribution of our paper to this literature is to build on less restrictive theo
retical assumptions. Like in all these previous studies, the short-term errors in the data
generating process (DGP) of our setting are allowed to exhibit stationary dependencies
with contemporaneously cross-correlated innovations, but we additionally permit such
innovations to exhibit time-varying conditional heteroskedasticity on a martingale differ
cence sequence (MDS) basis. This allows for both serial and cross-sectional dependence
in conditional second-order moments under suitable restrictions, which is particularly rel
vant for financial data. Furthermore, we do not require the order of integration to lie
in a particular region, thereby allowing for both stationary and nonstationary dynamics,
nor impose any particular assumption on the distribution of the data. To the best of our
knowledge, none of the existing methods in the extant literature achieve this degree of
flexibility. Our test can readily be implemented in a regression-based context building
on feasible generalized least squares (FGLS) estimation. We analytically show that the
asymptotic null distribution of the test statistic is a standard Chi-square distribution,
independently of the value of the long-memory parameters, and that the test exhibits
non-trivial power against sequences of local alternatives. An in-depth Monte Carlo analy
sis shows that the multivariate test has approximately correct asymptotic size in finite
samples and good power performance given different forms of empirically relevant features of the data, such as short-run dependence and time-varying volatility with both Gaussian and non-Gaussian innovations.

Using this new approach, we jointly infer the order of fractional integration of trading volumes and different measures of return volatility for the stocks that compose the Dow Jones Industrial Average Index (DJ) over the period 02/01/2003 through 31/12/2014. This analysis also allows us to address the existence of long-run commonalities, thereby providing new evidence on the general suitability of the MDH. We stress that since our test builds on general assumptions which do not require a particular distribution and permit time-varying second-order moments, the results may be more robust than those reported in the previous literature and which are based on estimation techniques within the FIVAR setting that either require Gaussianity and/or neglect multivariate conditional heteroskedasticity. Together with daily log-volume, we consider the log-transformations of three alternative measures of volatility with increasing degree of efficiency, namely, absolute-valued returns, the range-based estimator of Garman and Klass (1980), and realized variance constructed from 5-minute returns. The main goal of this analysis is to appraise the influence of measurement errors on the conclusions. The main picture that emerges from this analysis suggests that common fractional integration cannot be generally rejected when volatility is proxied by absolute-valued returns, but it does when volatility is proxied by more accurate estimates such as the range or realized variance. This result is consistent with previous literature and helps to understand the disparity of empirical results whenever different proxies are used. According to our study, volatility tends to exhibit a higher order of fractional integration than trading volume in the sample analyzed, with long-term behavior possibly driven by non-stationary dynamics.

The remainder of the paper is organized as follows. Section 2 introduces the DGP and states the main assumptions that characterize our theoretical analysis. It also presents the new test statistics for multivariate fractional integration and formally states its asymptotic distribution under the null hypothesis and under a sequence of local alternatives. Section 3 discusses the results of an in-depth Monte Carlo analysis to evaluate the performance of the test in finite samples and compares them to other multivariate tests available in the literature. Section 4 analyzes the empirical relation between trading volume and volatility given a sample of individual DJ stocks. Finally, Section 5 summarizes and
concludes. The mathematical proofs of the main statements in Section 2 are collected in a technical appendix.

In what follows, ‘⇒’ and ‘P’ denote weak convergence and convergence in probability, respectively, as the sample size is allowed to diverge. \( I(\cdot) \) is an indicator function that takes a value of one if the condition in parenthesis is fulfilled and zero otherwise. The operator \( \otimes \) denotes the Kronecker product of two matrices. The terms \( \mathbf{I}_n \) and \( \mathbf{0}_{n \times m} \) denote an \( n \)-dimensional identity matrix and an \( n \times m \) zero matrix, respectively. Throughout the paper, vectors and matrices are represented in bold letters.

## 2 Multivariate fractional integration

We start our analysis by stating the DGP and the set of sufficient assumptions that underlie the theoretical analysis in our paper. This section also introduces notation and certain variables that play a major role in our framework.

### 2.1 The data generating process

Consider an observable \( k \)-dimensional time series vector \( \{y_t\}_{t=1}^T \) generated as:

\[
\Delta^{d+\theta} (L) y_t = \varepsilon_t I(t \geq 1)
\]

where \( \Delta^{d+\theta} (L) \) is a \( k \times k \) diagonal matrix polynomial in the lag operator \( L \) with characteristic element given by \( (1 - L)^{d_i + \theta_i} \), where the real-valued \( d_i + \theta_i \) coefficient is usually referred to as the long-memory or fractional integration parameter, \( d+\theta := (d_1 + \theta_1, \ldots, d_k + \theta_k)' \), and the \( k \)-dimensional vector \( \{\varepsilon_t\} \) is a weakly-dependent noise process with spectral density that is bounded and bounded away from zero at the origin. For instance, in the bivariate case, model (1) is given by

\[
\begin{bmatrix}
(1 - L)^{d_1 + \theta_1} & 0 \\
0 & (1 - L)^{d_2 + \theta_2}
\end{bmatrix} y_t = \varepsilon_t I(t \geq 1)
\]

with \( y_t = (y_{1t}, y_{2t})' \) and \( \varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \). Assumption 1 below lays out the properties of the short-term component \( \{\varepsilon_t\} \) in (1) and completes the characterization of the DGP used in our analysis.
Assumption 1. \{\varepsilon_t\} in (1) is generated as $\Pi(L)\varepsilon_t = e_t$, with $\Pi(L) := I_p - \sum_{j=1}^{p} \Pi_j L^j$, and $\Pi_j$ being $k \times k$ parameter matrices such that $\Pi(L)$ has all its roots outside the unit root circle and $\{\varepsilon_t\}$ satisfies the conditions:

(A1) $E(e_t) = 0$ and $E(e_t e_t') := \Sigma$, with $\Sigma > 0$.

(A2) $\sup_t E(||e_t||^{4+\eta}) < \infty$ for some $\eta > 0$.

(A3) $\{e_t, \mathcal{F}_t\}_{t=-\infty}^{\infty}$ is a strictly stationary and ergodic vector martingale difference sequence, with $\mathcal{F}_t$ denoting the $\sigma$-field of $\{e_s : s \leq t\}$.

(A4) $\sum_{i=1}^{\infty} \sum_{j=1,i\neq j}^{\infty} E|e_{ht} e_{st} e_{rt-i} e_{ut-j}| < \infty$, for any $1 \leq h, s, r, u \leq k$, with $e_{lt}$ denoting the $l$-th element in $e_t$.

(A4*) $\sum_{s=1}^{k} \sum_{i=1}^{\infty} \sum_{j=1,i\neq j}^{\infty} \xi_{sij} < \infty$, $\xi_{sij} := ||E(e_{st} e_{t-i} e_{t-j}')||$ and $e_{st}$ denoting the $s$-th element in $e_t$.

Some comments follow. Assumption 1 allows the short-run component of $\{y_t\}$ to be driven by a stationary VAR($p$) process. Accordingly, (1) is a FIVAR time-series model in which each component $\{y_{it}\}$, $i = 1, ..., k$, is a (Type-II) ARFIMA($p, d_i + \theta_i, 0$) process. An interesting feature of this theoretical setting is that the fractional parameter $d_i + \theta_i$ is not restricted to lie in the interval $(-0.5, 0.5)$, a necessary condition to ensure stationarity and invertibility. (A1) and (A2) define usual moment conditions. (A3) allows innovations to exhibit time-varying conditional variances, with the absolute summability condition in (A4) limiting the amount of temporal and cross-sectional dependence in the second-order moments. Essentially, (A4) is equivalent to requiring absolutely summable 4th-order joint cumulants. This condition is weaker than requiring $\{e_t, \mathcal{F}_t\}$ to be conditionally homoskedastic, and considerably weaker than requiring independence, both of which imply (A4). Finally, (A1) and (A3) imply that $E(e_{it} e_{jt+h}) = 0$ whenever $h \neq 0$, but allows $E(e_{it} e_{jt}) \neq 0$ when $i \neq j$ because $\Sigma$ is not restricted to be diagonal.

\textsuperscript{1}The DGP can be analogously be rewritten as $\Pi(L) \Delta^{d_i + \theta_i} (L) y_t = e_t$. Under the stationary restriction, and for a large enough value of $p$, the FIVAR representation can be seen as an approximation of the more general class of FIVAR moving average (FIVARMA) models. Given its generality and methodological convenienic, this setting has been widely considered both in theoretical (e.g., Nielsen 2004, 2005) and empirical papers (e.g., Andersen et al. 2003).
2.2 The test regression framework

Given the observable time series vector \( \{y_t\} \) generated as in (1) and an arbitrary real-valued vector \( g := (g_1, ..., g_k)' \), define the \( k \)-dimensional stochastic processes

\[
\varepsilon_{t,g} := (1 - L)^k y_t = \sum_{j=0}^{t-1} \Lambda_j(g) y_{t-j},
\]

and

\[
z_{t-1,g}^* := \sum_{j=1}^{t-1} j^{-1} \varepsilon_{t-j,g}, \quad t = 2, ..., T
\]

with \( \{\Lambda_j(g)\}_{j=0}^{t-1} \) denoting a sequence of \( k \times k \) diagonal matrices with \( ii \)-th element given by

\[
\lambda_0(g_i) := 1, \quad \text{and} \quad \lambda_j(g_i) := \frac{j - 1 - g_i}{j} \lambda_{j-1}(g_i), \quad j \geq 1,
\]

corresponding to the truncated series of polynomial coefficients in the binomial expansion \((1 - L)^g := \sum_{j=0}^{\infty} \lambda_j(g) L^j\). These variables are straightforward generalizations of the univariate processes in Breitung and Hassler (2002) to the multivariate context, with the characteristic harmonic weighting in (4) arising from the derivative of a (Gaussian) score function.

Let \( \Phi = \{\phi_{ii}\} \) denote a \( k \times k \) diagonal matrix of fixed coefficients. Given Assumption 1, testing the null hypothesis that \( d \) is the order of integration of \( \{y_t\} \), or \( H_0 : \theta = 0 \), is equivalent to testing \( H_0 : \Phi = 0 \) in the multivariate linear regression model

\[
\varepsilon_{t,d} = \Phi z_{t-1,d}^* + \sum_{j=1}^{p} \Pi_j \varepsilon_{t-j,d} + v_t, \quad t = p + 1, ..., T
\]

because this characterization holds exactly with \( \phi_{ii} = 0 \) and \( v_t = e_t \) under \( H_0 : \theta = 0 \); see also Breitung and Hassler (2002), Demetrescu et al. (2008), and Hassler et al. (2009). Since the error covariance matrix \( \Sigma := E(e_t e_t') \) is not necessarily diagonal, (5) defines a seemingly unrelated regression equation system. Under Assumption 1 and \( H_0 : \theta = 0 \), the main parameters can be consistently estimated using FGLS, building on a preliminary consistent estimate of \( \Sigma \) which can be obtained from least squares (LS) residuals estimated equation-by-equation.

In the next section, we shall discuss the asymptotic distribution of the FGLS-based test for \( H_0 : \Phi = 0 \). To this end, it is convenient to write (5) using compact matrix notation. Noting that \( \varepsilon_{t,d} = (\varepsilon_{1t,d_1}, ..., \varepsilon_{kt,d_k})' \) and \( z_{t-1,d}^* := (z_{1,t-1,d_1}^*, ..., z_{k,t-1,d_k}^*)' \), we can
generally write for any of the \( i \)-th variables that,

\[
\mathbf{Y}_{i;t,d_i} = \mathbf{X}_{i,t-1,d}^* \beta_i + \mathbf{u}_{i,t}, \quad 1 \leq i \leq k
\]  

(6)

where \( \mathbf{Y}_{i;t,d_i} := (\varepsilon_{ip+1,d_i}, \ldots, \varepsilon_{iT,d_i})' \) is a \((T - p) \times 1\) vector, \( \beta_i := (\phi_{ii}, \pi_{i1}, \ldots, \pi_{ip})' \) is a \( k' \)-dimensional parameter vector, with \( k' := pk + 1 \), and \( \pi_{ij} \) denotes the \( i \)-th row of \( \Pi_j \), \( j = 1, ..., p \). \( \mathbf{u}_{i,t} := (v_{ip+1}, ..., v_{iT})' \) is a \((T - p) \times 1\) vector of residuals, and \( \mathbf{X}_{i,t-1,d}^* \) is the \((T - p) \times k' \) matrix of observations of the right-hand side variables \( \mathbf{x}_{i,t-1,d}^* := (z_{i,t-1,d}; \varepsilon_{i-1,d}; \ldots, \varepsilon_{i-p,d})' \). With the exception of the first regressor, all other right-hand side variables that characterize the \( i \)-th equation (6) are the same, since these always correspond to lagged values of \( \varepsilon_{i,t,d} \). Then, given \( T' := k (T - p) \), we can write the system of equations compactly as \( \mathbf{Y}_{t,d} = \mathbf{X}_{t-1,d}^* \mathbf{\beta} + \mathbf{u}_t \), with these terms defined implicitly as:

\[
\begin{bmatrix}
\mathbf{Y}_{1,t,d_1} \\
\mathbf{Y}_{2,t,d_2} \\
\vdots \\
\mathbf{Y}_{k,t,d_2}
\end{bmatrix}_{T' \times 1} =
\begin{bmatrix}
\mathbf{X}_{1,t-1,d}^* & 0_{(T-p) \times k'} & \cdots & 0_{(T-p) \times k'} \\
0_{(T-p) \times k'} & \mathbf{X}_{2,t-1,d}^* & \cdots & 0_{(T-p) \times k'} \\
\vdots & \vdots & \ddots & \vdots \\
0_{(T-p) \times k'} & 0_{(T-p) \times k'} & \cdots & \mathbf{X}_{k,t-1,d}^*
\end{bmatrix}_{T' \times kk'} \mathbf{\beta} +
\begin{bmatrix}
\mathbf{u}_{1t} \\
\mathbf{u}_{2t} \\
\vdots \\
\mathbf{u}_{kt}
\end{bmatrix}_{T' \times 1}
\]

2.3 The multivariate fractional integration test

Under Assumption 1 and \( \mathcal{H}_0 : \mathbf{\theta} = \mathbf{0} \), it follows that \( E(\mathbf{u}_t \mathbf{u}_t') = \mathbf{\Sigma} \otimes \mathbf{I}_{k'} \), recalling that \( k' := pk + 1 \) is the number of parameters to be estimated in each equation. Since \( \mathbf{\Sigma} \) may not be diagonal, equation-by-equation LS estimation renders consistent estimates of \( \mathbf{\beta} \), but is no longer efficient. Hence, alternatively, we consider the FGLS estimator of \( \mathbf{\beta} \), defined as:

\[
\widehat{\mathbf{\beta}} := \left( \mathbf{X}_{t-1,d}^* \hat{\mathbf{\Sigma}}^{-1} \otimes I_{T-p} \mathbf{X}_{t-1,d}^* \right)^{-1} \left( \mathbf{X}_{t-1,d}^* \hat{\mathbf{\Sigma}}^{-1} \otimes I_{T-p} \mathbf{Y}_{t,d} \right)
\]

(7)

where \( \hat{\mathbf{\Sigma}} = \{\hat{\sigma}_{ij}\} \) is estimated as \( \hat{\sigma}_{ij} = T^{-1} \hat{\mathbf{u}}_{it} \hat{\mathbf{u}}_{jt} \), with \( \hat{\mathbf{u}}_{st} := \mathbf{Y}_{st,d,s} - \mathbf{X}_{st-1,d,s} \hat{\mathbf{\beta}}_s \), \( s = 1, ..., k \), and \( \hat{\mathbf{\beta}}_s \) denotes the equation-by-equation LS estimate of (6). Theorem 1 below characterizes the asymptotic null distribution of \( \widehat{\mathbf{\beta}} \) under Assumption 1.
Theorem 1. Let \( \hat{\beta} \) be a vector of FGLS estimates as characterized in (7). Under Assumption 1 and \( H_0 : \theta = 0 \),

\[
\sqrt{T} \left( \hat{\beta} - \beta_0 \right) \Rightarrow \mathcal{N}(0, \Omega_{\beta})
\]

where \( \beta_0 = (\beta_{01}, \ldots, \beta_{0k})' \), \( \beta_0 := (0, \pi_{s1}, \ldots, \pi_{sp})' \), \( s = 1, \ldots, k \), and \( \Omega_{\beta} := A_{\beta}^{-1} B_{\beta} A_{\beta}^{-1} \), with \( A_{\beta} := \text{plim} E \left( \frac{1}{T} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p} X_{t-1,d} \right) \), \( B_{\beta} := \text{plim} E \left( \frac{1}{T} w_{t-1,d}^* w_{t-1,d}^* \right) \), and \( w_{t-1,d}^* := X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p} u_t \).

If Assumption 1 holds with i.i.d. observations, then \( \Omega_{\beta} = A_{\beta}^{-1} \). More, generally, however, the asymptotic covariance matrix is given by \( \Omega_{\beta} = A_{\beta}^{-1} B_{\beta} A_{\beta}^{-1} \). This term can be estimated consistently using the Eicker-Huber-White approach building on the preliminary estimate of \( \Sigma \) and on the FGLS residuals \( \hat{u}_t = Y_t - X_{t-1}^* \hat{\beta} \). This strategy ensures robustness against unknown conditional heteroskedasticity under Assumption 1. In particular, define \( A_T := Z_{t-1}^* Z_{t-1}^* / T \), with \( Z_{t-1}^* := \hat{\Sigma}^{-1/2} X_{t-1,d} \otimes I_{T-p} \), and \( B_T := w_{T,t-1}^* W_{T,t-1}^* / T \), with \( w_{T,t-1}^* := X_{t-1,d}^* \hat{\Sigma}^{-1/2} \otimes I_{T-p} \hat{u}_t \). Hence, it can be shown that \( \Omega_T := A_T^{-1} B_T A_T^{-1} \) is a consistent estimate for \( \Omega_{\beta} \); see Appendix for details. Given this result, it is straightforward to construct a test statistic for the joint hypothesis \( H_0 : \theta = 0 \) building on the LM principle. This is formally stated in the following theorem.

Theorem 2. Let \( R = \{r_{ij}\} \) be a \( k \times k' \) indicator matrix taking a value equal to one when \( j = (i-1)k' + 1, i = 1, \ldots, k \), and zero otherwise, and let \( \Omega_T := A_T^{-1} B_T A_T^{-1} \) be the sample-based estimate of the asymptotic covariance matrix of \( \beta \). Thus, considering that Assumption 1 holds, it follows under \( H_0 : \theta = 0 \), that,

\[
LM_{d}^{FGLS} := T \left[ R \hat{\Omega} \right] \left[ R \Omega_T \ R' \right]^{-1} \left[ R \hat{\Omega} \right] \Rightarrow \chi^2_{(k)}
\]

whereas under a sequence of local alternatives \( H_1 : \theta = c / \sqrt{T} \), for \( ||c|| > 0 \),

\[
LM_{d}^{FGLS} \Rightarrow \chi^2_{(k, ||c||^2)}
\]

where \( \chi^2_{(k)} \) and \( \chi^2_{(k, ||c||^2)} \) denote a standard Chi-square distribution with \( k \) degrees of freedom, and a non-central Chi-square distribution with \( k \) degrees of freedom and non-centrality parameter \( ||c||^2 \), respectively.

Remark 1: The test procedure can be generalized to account for non-zero means following the approach in Robinson (1994). For instance, consider \( y_t = \mu + \Delta (L)^{-d} \theta \varepsilon_t I(t \geq 1) \),
where $\mu := (\mu_1, \ldots, \mu_k)'$ is a fixed (unknown) vector. Under $H_0: \theta = 0$, $(1 - L)^d y_{st} = \mu_s (1 - L)^d + \varepsilon_t I(t \geq 1), 1 \leq s \leq k$, and hence the $s$-th element of $\mu$ can be consistently estimated under Assumption 1 from the LS regression of $(1 - L)^d y_{st} \equiv \sum_{j=0}^{t-1} \lambda_j (d_s) y_{st-j}$ on $h_{t,ds} := \sum_{j=0}^{t-1} \lambda_j (d_s), t = 2, \ldots, T$, with $\{\lambda_j (d_s)\}$ as defined in (4). The residuals of this regression correspond to $\{\varepsilon_{t,ds}\}$ and, therefore, it suffices to redefine the elements in (2) as $\varepsilon_{t,ds} := (1 - L)^d y_{st} - \hat{\mu}_s h_{t,ds}$. The results can be extended along the same lines to account for deterministic time trends and seasonalities. \hfill \Box

**Remark 2:** Theorem 2 provides a theoretical basis for the construction of confidence sets that include the values of the fractional parameter vector with $100 (1 - \lambda)$% asymptotic coverage. This can be achieved by inverting the non-rejection region of the test statistic; see Hassler et al. (2016). More specifically, let $LM_g$ denote the value of the LM statistic when testing $H_0: \theta = 0$ for an arbitrary $g \in \mathbb{R}^k$, and let $\Psi$ be an arbitrary compact set in $\mathbb{R}^k$. Define $D_\lambda = \left\{ g \in \Psi : \Pr \left[ \chi^2_{(k)} > LM_g \right] \leq 1 - \lambda \right\}$ with $\lambda \in (0,1)$, i.e., the subset of $\Psi$ for which the null hypothesis cannot be rejected at the $\lambda$ significance level. From Theorem 2, it follows that if $\Psi$ is large enough as to contain the true values of the long-memory parameter vector, then the probability of the true order of integration lying within $D_\lambda$ is at least $(1 - \lambda)$. Thus, a confidence ellipsoid can be constructed through a grid-search process. While $\Psi$ can be taken arbitrarily large, the approach is computationally feasible because the fractional parameters typically lie in the $(0,1)$ interval. We shall implement this technique later in the empirical section. \hfill \Box

**Remark 3:** Consider that Assumption 1 holds with i.i.d. innovations $e_t$ as in Nielsen (2005). Then, we can write,

\[
LM^{GLS}_d = \left( X_{t-1,d}' \tilde{\Sigma}^{-1} \otimes I_{T-p} Y_{t,d} \right)' \left( \frac{1}{T} X_{t-1,d}' \tilde{\Sigma}^{-1} \otimes I_{T-p} X_{t-1,d} \right)^{-1} \left( X_{t-1,d}' \tilde{\Sigma}^{-1} \otimes I_{T-p} Y_{t,d} \right)
\]

Furthermore, assume that $p = 0$, such that $X_{t-1,d} := \text{diag}\{X_{t-1,1:d}, \ldots, X_{t-1,k-d}^*\}$ with $X_{t-1,i:d} := Z_{t-1,i,d} a_i, Z_{t-1,i,d} := \{z_{t-1,i:d}, \ldots, z_{t-1,k-d}^*\}$, and $a_i$ denoting the $i$-th unit $k$-dimensional vector. Noting that $z_{t-1,i,g} := \sum_{j=1}^{t-1} e_{t-j,i} = - \ln (1 - L) e_{t-j,i}$ under the null hypothesis, the term $X_{t-1,d}' \tilde{\Sigma}^{-1} \otimes I_{T-p} Y_{t,d}$ corresponds to the Gaussian score vector $S_T := J_{k} \text{vec} \left( \tilde{\Sigma}^{-1} S_{10} \right)$ in Nielsen (2005, eq.11), where $S_{10} := \sum_{t=2}^{T} e_{t-1} e_t'$ with $e_{t-1} := \sum_{j=1}^{t-1} e_{t-j}$, and $J_{k} := (\text{vec}(A_{11}), \ldots, \text{vec}(A_{kk}))$ with $A_{ii} := a_i a_i'$. Since $\frac{1}{T} X_{t-1,d}' \tilde{\Sigma}^{-1} \otimes I_{T-p} X_{t-1,d} \overset{p}{\rightarrow} A_{\beta}$, and $\text{plim} E (S_T S_T') = A_{\beta}$ under the restrictions considered, $LM^{GLS}_d$ is asymptotically equivalent to the Gaussian LM test in Nielsen (2005) and, therefore, asymptotically
efficient when \( e_t \) is i.i.d. normal distributed; see also Nielsen (2004). This result carries over when \( p > 0 \), noting that the two tests differ crucially on how short-run dependencies are dealt with at this point. While \( LM_d^{FGLS} \) uses \( p \)th-order augmentation in the auxiliary regression, Nielsen’s (2005) test uses a two-staged procedure and relies on the residuals of a VAR(\( p \)) model. Augmentation and prewhitening are asymptotically equivalent strategies, but it is widely agreed in the literature of stochastic-trend detection that augmentation typically leads to more efficient results in finite samples. We shall address this question in the Monte Carlo section.

\[ \Box \]

3 Monte Carlo analysis

3.1 Experiment design

In this section, we analyze the finite-sample properties of the FGLS multivariate tests for fractional integration using Monte Carlo simulation. We focus on the bivariate case \( y_t = (y_{1t}, y_{2t})' \), considering the DGP

\[
\begin{bmatrix}
(1 - L)^{1+\theta_1} & 0 \\
0 & (1 - L)^{1+\theta_2}
\end{bmatrix}
y_t = \epsilon_t I(t \geq 1), \quad t = 1, \ldots, T,
\]

(8)

such that \( \Pi(L)\epsilon_t = e_t \) with

\[
\Pi(L) = \begin{bmatrix}
1 - \pi_1 L & 0 \\
0 & 1 - \pi_2 L
\end{bmatrix}
\]

(9)

and \( (\pi_1, \pi_2) \in \{(0, 0), (0.4, 0.4)\} \). The case \( \pi_1 = \pi_2 = 0 \) corresponds to white noise, and \( \pi_1 = \pi_2 = 0.4 \) leads to stationary VAR(1) errors. Since the particular values of the long-memory coefficients play no role in our context, we set \( d_1 = d_2 = 1 \) with no loss of generality, corresponding to the unit-root case under the null hypothesis. We set finite sample lengths \( T \in \{500, 1000\} \).

The innovations \( \{\epsilon_t\} \) exhibit time-varying conditional second-order moments according to:

\[
e_t = \begin{bmatrix}
\sigma_{1t} & 0 \\
0 & \sigma_{2t}
\end{bmatrix} \eta_t; \quad E(\eta_t) = 0, \quad E(\eta_t \eta_t') = \Omega_\rho = \begin{bmatrix}
1 & \rho \\
\rho & 1
\end{bmatrix}
\]

where \( \eta_t := (\eta_{1t}, \eta_{2t})' \) is an i.i.d. vector drawn from either a multivariate Gaussian or a multivariate Student-\( t \) distribution with 5 degrees of freedom. The latter is motivated
by the empirical observation that the conditional distributions of high-frequency financial
time series typically exhibit heavy tails. The covariance matrix $\Omega_\rho$ depends on the con-
temporaneous correlation coefficient $\rho$, for which we set $\rho \in \{0, 0.2, 0.4, 0.6, 0.8\}$. The
conditional variances $\{\sigma^2_t\}$ are driven by a stationary GARCH(1,1) process characterized by:

$$
\sigma^2_t = (1 - \alpha - \beta) + \alpha \epsilon^2_{t-1} + \beta \sigma^2_{t-1}, \; i = 1, 2
$$

with $\alpha, \beta \geq 0$ and $\alpha + \beta < 1$ such that $E(\epsilon^2_t) = 1$. For parsimony, we consider the same
GARCH dynamics for the two series, focusing on GARCH parameter configurations that
characterize different degrees of (empirically relevant) persistence in conditional variance
as measured by $\alpha + \beta$, namely, $(\alpha, \beta) \in \{(0, 0), (0.1, 0.5), (0.1, 0.7), (0.1, 0.8), (0.1, 0.85)\}$. The case $\alpha = \beta = 0$ corresponds to conditional homoskedasticity, which coupled with $\rho = 0$ leads to independent series with i.i.d. observations. Non-zero values of these
parameters introduce serial and/or cross-sectional dependencies in the short-run dynamics
of the process.

For ease of presentation and discussion, we fix $\theta_2 = 0$ in all simulations and consider
the sequence $\theta_1 \in \{-0.3, -0.25, ..., 0, ..., 0.25, 0.3\}$. Consequently, while the true order of
integration of $\{y_{2t}\}$ is always one, the true order of integration of $\{y_{1t}\}$ is $1 + \theta_1$. Hence,
the analysis of the rejection frequencies under the case $\theta_1 = 0$ determines the empirical
size of the multivariate test, while the cases $\theta_1 \neq 0$ characterize its finite-sample power
behavior. For each of the parameter configurations $(\alpha, \beta, \rho, \pi_1, \pi_2, \theta_1)$, the two sample
lengths, and the two conditional distributions, we compute $LM_d^{\text{FGLS}}$ and determine the
average frequencies of rejection at the usual 5% nominal size level given 5,000 replications.

The performance of the test is benchmarked with two alternative (but related) multivari-
tate tests, namely, the Gaussian-based multivariate LM test in Nielsen (2004, 2005),
denoted $LM_d^{\text{MLE}}$, and the multivariate trace test of Breitung and Hassler (2002), denoted
as $BH_d$. The $LM_d^{\text{MLE}}$ test statistic can be computed as

$$
LM_d^{\text{MLE}} = vec\left(\tilde{\Sigma}^{-1}S_{10}'\right)'J_2\left(S_{11} \odot \tilde{\Sigma}^{-1} + (\tilde{\Sigma}^{-1}S_{20}) \odot I_2\right)^{-1}J_2 vec\left(\tilde{\Sigma}^{-1}S_{10}'\right)
$$

(10)

with $k = 2$ in the bivariate setting, $S_{10} := \sum_{t=2}^T e^*_{t-1}e^*_t$ with $e^*_{t-1} := \sum_{j=1}^{t-1} j^{-1}e_{t-j}$, $S_{11} := \sum_{t=2}^T e^*_{t-1}e^*_{t-1}$, $S_{20} := \sum_{t=1}^T e^*_{t-2}e^*_t$ with $e^*_{t-2} := \sum_{j=1}^{t-2} j^{-1} e^*_{t-j}$, and $J_2 := (vec(A_{11}), vec(A_{22}))$, where $A_{ii} := aia'_i$ with $a_i$ denoting the $i$-th unit $k$-vector. The operator $\odot$ denotes the
Hadamard product. Similarly, the $BH_d$ test statistic can be computed as,

$$BH_d = tr \left( \hat{\Sigma}^{-1} S_{10}^{-1} S_{11}^{-1} S_{10} \right) \Rightarrow \chi^2_{(k^2)}$$ (11)

noting that the asymptotic null distribution of the test is a Chi-square distribution with $k^2 = 4$ degrees of freedom in the bivariate context.

Whereas $BH_d$ is specifically intended to address the null hypothesis of common order of integration (which holds under the null hypothesis in our experiments), $LM_d^{MLE}$ can be used in a more general context and, therefore, it seems the most natural candidate to benchmark the performance of our test. As discussed in Remark 3, $LM_d^{MLE}$ and $LM_d^{FGLS}$ are asymptotically equivalent under suitable restrictions, but the performance of these tests may differ in finite samples, particularly, in the presence of short-run dependencies. Furthermore, since $LM_d^{FGLS}$ builds on more general theoretical assumptions than $LM_d^{MLE}$ and $BH_d$, the different conditions in our experimental study may lead to relevant differences, not necessarily related to finite-sample considerations.

### 3.2 Experiment results

#### 3.2.1 Empirical sizes under no augmentation/prewhitening.

We first report the empirical size of the multivariate tests in the absence of short-term dynamics in mean ($\pi_1 = \pi_2 = 0$), i.e., when there is no need for augmentation/prewhitening. This gives a sense of the performance of the tests without interferences and, hence, is the natural starting point in our analysis. Thus, Table 1 reports the rejection frequencies of $LM_d^{FGLS}$, $LM_d^{MLE}$ and $BH_d$ when $\theta_1 = \theta_2 = 0$ for different values of the GARCH parameters ($\alpha, \beta$) and the correlation coefficient $\rho$, given the two sample lengths $T$ and the conditional distributions considered. According to the average rejection frequencies, the multivariate FGLS test exhibits finite-sample size close to the 5% nominal level in all cases, with mild size distortions that tend to vanish as the sample size increases.

In relative terms, the size performance of $LM_d^{FGLS}$ is similar to or better than that of the alternative procedures analyzed. More specifically, when innovations are i.i.d. ($\alpha = \beta = 0$), all the multivariate tests display correct finite-sample size, independently of the underlying conditional distribution. This evidence is expected from the theoretical considerations in the previous section and agrees with the experimental results commented
in Nielsen (2005). On the other hand, when innovations exhibit conditional heteroskedasticity, $LM_{d}^{MLE}$ and $BH_{d}$ suffer noticeable size distortions. In contrast to $LM_{d}^{FGLS}$, whose empirical size is close to the 5% nominal level in all cases, the alternative multivariate tests become oversized, particularly, when volatility is strongly persistent and innovations are drawn from a heavy-tailed distribution. Such distortions depend on the degree of endogenous correlation and tend to increase with $|\rho|$. More importantly, the size departures do not seem to be attenuated when the number of observations in the sample increases. This evidence suggests directly that the asymptotic limit distribution of those test statistics under conditional heteroskedasticity may no longer correspond to that discussed under the i.i.d. setting. For instance, for $T = 500$ and $\rho = 0.8$, the empirical sizes of $LM_{d}^{MLE}$ and $BH_{d}$ with GARCH errors driven by $(\alpha, \beta) = (0.10, 0.85)$ and Student-t innovations are 36% and 38.8%, respectively. In contrast, the size of $LM_{d}^{FGLS}$ is only slightly oversized, 7.1%. When $T = 1000$, the empirical sizes of $LM_{d}^{MLE}$ and $BH_{d}$ increase significantly up to 48.3% and 52.8%, respectively, whereas the size of $LM_{d}^{FGLS}$ slightly reduces to 6.1%, approaching the required asymptotic nominal level.

Two main conclusion can be drawn from this experiment. First, Monte Carlo analysis confirms the empirical suitability of the theoretical results discussed in the previous section and the reliability of the asymptotic critical values taken from the Chi-squared distribution for finite samples. Second, it is widely agreed that time-varying conditional variance and excess kurtosis are stylized features of high-frequency financial variables, which furthermore often exhibit long-range dependence. When addressing hypotheses involving the order of fractional integration on data with such characteristics, the multivariate test proposed in this paper can provide reliable inference. In contrast, the multivariate tests that require conditional homoskedasticity may lead to biased conclusions resulting from oversized inference. The magnitude of such departures is data-dependent and is critically affected by the extent of non-Gaussian features of the conditional distribution, the degree of persistence in conditional variance, and the cross-correlation of innovations.

[ Insert Table 1 around here ]
3.2.2 Frequencies of rejection under augmentation/prewhitening.

Secondly, we characterize the size and power profiles of the multivariate tests under augmentation/prewhitening. We allow the errors in the DGP to display stationary VAR dynamics with either \( \pi_1 = \pi_2 = 0 \) (unnecessary augmentation) or \( \pi_1 = \pi_2 = 0.4 \) (correct augmentation). In both cases, the auxiliary regression that gives rise to \( LM_{d,FGLS} \) is augmented with one lag of \( \varepsilon_{t,d} \) and, analogously, we compute \( LM_{d,MLE} \) and \( BH_d \) on the residuals of a VAR(1) model on \( \varepsilon_{t,d} \). Note that augmentation/prewhitening is redundant when \( \pi_1 = \pi_2 = 0 \). While adding unnecessary lags does not hurt the empirical size asymptotically, it reduces power and may lead to size departures in finite samples. Monte Carlo simulation allows us to analyze the performance of the multivariate tests in this context.

[ Insert Table 2 around here ]

Table 2 reports the rejection frequencies of the tests in the Gaussian homoskedastic case (\( \alpha = \beta = 0 \)) given the different values of \( \theta_1 \) and the two sample lengths considered. Since the results from the Student-\( t \) distribution are completely similar in this context, we do not report them to save space, but we remark that complete results are available upon request. Similarly, for ease of exposition, we only present the results for the values of the correlation coefficient \( \rho = 0 \) and \( \rho = 0.8 \), omitting intermediate cases. Several issues are worth noting from this experiment.

First, the average frequency of rejection for \( \theta_1 = 0 \) shows that augmentation is a valid instrument to handle short-run dependence even in finite samples. The empirical size of the FGLS-based test is always close to the nominal level, even under unnecessary augmentation. The same statement is true for the other two tests, but we note that prewhitening leads to slightly larger size distortions. According to the results in Table 2, augmentation tends to generate more stable size around the nominal level than prewhitening, which generally agrees with the results in the literature of stochastic-trend detection.

Secondly, our test exhibits increasing power on the size of \( |\theta_1| \) and the sample length \( T \), which fully supports the theoretical statements in Theorem 2, as well as on the magnitude of the correlation parameter \( \rho \), which perfectly illustrates the benefits from multivariate modelling when variables are cross-correlated. In relative terms, our test presents comparable power to the other two methods, although we notice large relative gains in many
cases. Overall, the Monte Carlo analysis suggests that for a fixed $T$ the power function of $LM_{dFGLS}$ tends to dominate the power functions of $LM_{dMLE}$ and $BH_d$. The gains are small or negligible when $\theta_1 < 0$, but can be relatively large in the region $\theta_1 > 0$ (i.e., when the process is more persistent than posited under the null). For instance, for $\pi_1 = \pi_2 = 0.4$, $\rho = 0$, $T = 500$, and $\theta_1 = 0.3$ the power of $\{LM_{dFGLS}, LM_{dMLE}, BH_d\}$ is $\{57.4\%, 32.3\%, 26.3\\% \}$. Accordingly, the relative gains in efficiency of $LM_{dFGLS}$ with respect to $LM_{dMLE}$ and $BH_d$ are about $78\%$ and $118\%$, respectively. This improved performance stems because augmentation is a more efficient way of handling short-run dependencies than prewhitening. Finally, it is worth noticing that the power function of $LM_{dFGLS}$ exhibits data-dependent asymmetries as a function of $\rho$ and the sign of $\theta_1$. For instance, given $\pi_1 = \pi_2 = 0.4$, $\rho = 0$, $T = 500$, the power of $LM_{dFGLS}$ to detect $\theta_1 = 0.3$ ($\theta_1 = -0.3$) is $57.4\%$ ($87.5\%$), such that overdifferentiation is more easily rejected. Similar patterns arise in the power functions of the other two tests. Breitung and Hassler (2002) also report data-dependent asymmetries in power in the univariate context.

[ Insert Tables 3 and 4 around here ]

Finally, we address the influence of GARCH-type dependencies in the size/power profile of the test. Given $T = 500$ and for both Gaussian and Student-$t$ innovations, Table 3 reports the rejection frequencies of the multivariate tests when $\pi_1 = \pi_2 = 0.4$, $\rho \in \{0, 0.8\}$, and $(\alpha, \beta) \in \{(0.10, 0.80), (0.10, 0.85)\}$. Table 4 reports the analogous results when the sample length is $T = 1000$. When $\theta_1 = 0$, the empirical size of the tests under GARCH innovations are similar to the values reported under no augmentation/prewhitening in Table 1 above. In particular, whereas the empirical size of $LM_{dFGLS}$ is close to the $5\%$ nominal level in all cases (size departures are not greater than $1.6\%$ for $T = 500$ and $0.8\%$ for $T = 1000$), the wrong assumption of conditional homoskedasticity leads $LM_{dMLE}$ and $BH_d$ to overreject. For example, for $T = 500$, $\rho = 0.8$, and $(\alpha, \beta) = (0.10, 0.85)$, $LM_{dMLE}$ and $BH_d$ exhibit empirical sizes that reach, respectively, $9.2\%$ and $8.7\%$ under Gaussian innovations, and $28.2\%$ and $32.4\%$ under Student-$t$ innovations; see Table 3. As discussed earlier, increasing the sample length does not seem to help to reduce these distortions. For $T = 1000$, empirical size levels under the same parameter configuration reach $10.4\%$ and $9.9\%$ under Gaussian innovations, and $40.4\%$ and $46.2\%$ under Student-$t$ innovations; see Table 4.
For non-zero values of $\theta_1$, we observe similar patterns in the power function of $LM_{d}^{FGLS}$ as those reported in Table 2 in the homoskedastic case, noting on the other hand that persistent GARCH-type dependencies erode the ability of $LM_{d}^{FGLS}$ to reject the null hypothesis with respect to the i.i.d. case, particularly, when the distribution is heavy-tailed. For instance, for $\pi_1 = \pi_2 = 0.4, \rho = 0, T = 500$, the power of $LM_{d}^{FGLS}$ to detect $\theta_1 = 0.3$ ($\theta_1 = -0.3$) in the i.i.d. case is 57.4% (87.5%); see Table 2. Under GARCH dependence with $(\alpha, \beta) = (0.10, 0.85)$ the respective probabilities are 52.7% (75.7%) in the Gaussian case, and 37.3% (49.3%) in the Student-t case; see Table 3. Fortunately, the power performance of the test largely increases with the availability of data. When $T = 1000$, the rejection frequencies under this parameter configuration increase up to 98.7% (99.9%) in the Gaussian case, and up to 71.6% (81.1%) in the Student-t case; see Table 4. In this regard, it should be mentioned that time-varying patterns in conditional variance are a distinctive feature of high-frequency financial variables, for which available samples usually span thousands of observations.

The main conclusion from this experiment is that the multivariate test proposed in this paper exhibits increasing power to reject the null hypothesis against departures from the null when the sample length and/or the magnitude of $||\theta||$ increases, which confirms the theoretical statements in the previous section. The test exhibits good power performance given empirically relevant features of the data. In a conditionally homoskedastic environment, our test has finite-sample power that either is comparable to that of prewhitening-based methods or can improve it substantially. Under conditional heteroskedasticity, the tests that assume conditional homoskedasticity may experience important size distortions, whereas our test may still ensure approximately correct size and acceptable power performance. All these characteristics make the test proposed in this paper a better alternative to address hypotheses involving the order of fractional integration of (multivariate) variables exhibiting the stylized features of financial data.

4 Long-run dynamics in volume and volatility

Understanding the links that intertwine return volatility, liquidity and trading activity has taken a predominant position in the finance literature. In this section, we implement the multivariate testing approach developed in the previous section to jointly infer the
order of fractional integration of trading volume and volatility on a sample of major stocks traded in the U.S. market. Given the resultant estimates, we can furthermore address the hypothesis that these variables exhibit the same order of fractional integration within the FIVAR setting. This premise might be backed up by a long-run view of the MDH, which generally posits that volume and volatility are jointly subordinate to a latent information-arrival process that measures the rate at which information arrives to the market. Bollerslev and Jubinski (1999) argue that, if the latent directing process possesses long-memory characteristics, then trading volume and volatility should exhibit fractional integration dynamics of the same order.

This hypothesis has been tested empirically within the FIVAR(MA) setting in different papers. Apart from the sheer differences in the sample composition and periods analyzed, the related studies differ mainly in the testing approach implemented and in the variable used to proxy for volatility. Bollerslev and Jubinski (1999) and Lobato and Velasco (2000) use semiparametric multivariate periodogram-based estimators in the frequency domain, proxying volatility with absolute-valued returns. The main conclusion of these studies is that for most of the stocks analyzed the hypothesis that trading volume and volatility share the same order of fractional integration cannot be rejected. On the other hand, Fleming and Kirby (2011) argue that periodogram-based estimators have a low rate of convergence, which raises concerns about efficiency in estimation. As an alternative, these authors implement a parametric Gaussian quasi maximum likelihood (QML) approach as in Nielsen (2004) to estimate a bivariate FIVAR model, allowing for short-run dependencies, but assuming conditional homoskedasticity. Additionally, these authors focus their main analysis on realized variance to improve accuracy over absolute-valued returns. The results in this analysis show a rejection frequency of the hypothesis of common fractional integration of 20%, which leads the authors to question the general suitability of the long-run view of the MDH.

Our testing procedure can be used to shed further light on this issue for two main reasons. First, as shown in Theorem 1, the FGLS-based test achieves the usual $\sqrt{T}$ rate of convergence in parametric testing, which ensures greater finite-sample power performance over periodogram-based estimators; see, for instance, Tanaka (1999). This consideration overcomes the potential concerns about efficiency raised by Fleming and Kirby (2011). In addition, our testing approach is formally valid under conditional heteroskedasticity pro-
vided absolute-summability conditions. As shown in the Monte Carlo section, whereas our
test can control size in the presence of conditionally time-varying second-order moments
and heavy-tailed innovations, alternative multivariate procedures that impose the i.i.d.
condition may result in oversized inference. Evidently, the QML approach in Nielsen
(2004) and Fleming and Kirby (2011) fall into this pitfall, so given the characteristic
stylized features of financial data our procedure may lead to more robust results.

4.1 Data

Our analysis focuses on the individual stocks that composed the DJ index over the period
02/01/2003 through 31/12/2014. In contrast to trading volume, volatility cannot be
observed directly. The literature has suggested different proxies which do not involve
parameter estimation. We implement three methods in increasing degree of accuracy to
estimate daily volatility in this spirit. The simplest approach uses absolute-valued returns
computed from close-to-close daily prices. Unfortunately, this measure is known to be
very inefficient and subject to large estimation errors. More accurate estimates can be
constructed building on intra-day information. Thus, following Garman and Klass (1980),
we proxy daily variability as $u_t^2/2 - (2 \ln 2 - 1) c_t^2$, where $u_t$ and $c_t$ are the differences in
the natural logarithms of the high and low, and of the closing and opening prices, respectively.
Range-based estimators produce more efficient estimates than absolute-valued returns
computed from close-to-close prices (Parkinson, 1980) and, as discussed in Anderson and
Bollerslev (1998), can be as efficient a measure of volatility as realized volatility computed
on the basis of three to four hour returns. Efficiency can be improved further by using the
whole path of intra-day variation, which leads to realized variance and related measures;
see, among others, Andersen et al. (2003) and references therein. Consequently, the last
estimator we consider is realized variance computed from aggregating 5-minute squared
returns over the trading session. Daily volumes and high, low, opening and closing prices
are obtained from CRSP. High-frequency prices necessary to compute realized variance
are obtained from the NYSE Trade and Quote (TAQ) database.

As customary in the related literature, we implement the log-transforms in both vol-
ume and volatility variables prior to the analysis of long memory in order to reduce the
effects associate with extreme observations; see, for instance, Andersen et al. (2003).
The log-transform is known to induce Gaussian-like behavior in realized variance, but we
noticed that all the variables in our sample seem to exhibit significant degrees of skewness and/or excess kurtosis. Table 5 details the stocks involved in our analysis and reports standard descriptive statistics.

[ Insert Table 5 around here ]

4.2 Main results

We adopt the following approach in the joint estimation of the order of fractional integration $d$ of log-trading volume and log-volatility. First, in order to account for non-zero deterministic effects in the level of these series, we implement the demeaning process suggested by Robinson (1994); see Remark 1. While the vast majority of papers in empirical finance and financial econometrics do not consider deterministic trends as a stylized feature of volatility, trading volume is widely accepted to exhibit trending paths which conform with increasing growth in the number of traders and trading activity; see Fleming and Kirby (2011) and reference therein for a discussion. Consequently, for the log-volatility measures, the main analysis is carried out by simply including a constant to capture a non-zero drift in the series, as in Hassler et al. (2016). In the case of log-volume, we set a quadratic time trend polynomial $\mu_{it} = \mu_{0i} + \mu_{1i}(t/T) + \mu_{2i}(t/T)^2$ to capture non-linear patterns as in, among others, Gallant et al. (1992), Luu and Martens (2003) and Fleming and Kirby (2011). For any potential value of the long-memory parameter vector, the parameters that characterize these deterministic functions are estimated endogenously through LS, as described in Remark 1, and then the multivariate test is computed on the resultant residuals.

Secondly, the FGLS auxiliary regression system is augmented with $p$ lags of the depen-

\footnote{Campbell et al. (2001) document a positive trend in idiosyncratic volatility of individual firms during the 1962–1997 period, giving rise to the so-called idiosyncratic volatility puzzle. However, Brandt et al. (2010) argue that idiosyncratic volatility has fallen substantially over recent years, reversing any potential time trend pattern. In the view of these authors, the overall behavior of idiosyncratic volatility is more likely to reflect episodic phenomena which may be associated with market drivers, such as the activity of retail investors, than a true time trend. From an econometric perspective, it should be noted that the episodic occurrence of apparent time trends is compatible with the dynamics of fractional integration, since long memory causes variables to exhibit local trends and cycles; see Sibbertsen et al. (2017).}
dent variable to account for short-run dependence. We adopt the same criterion advocated by Demetrescu et al. (2008) in the univariate context and determine the lag order according to Schwert’s rule, namely, \( p = [4(T/100)^{1/4}] \), where \([\cdot]\) denotes the integer part of the argument.\(^3\) Finally, following the estimation strategy discussed in Remark 2, we construct 95% confidence-set estimators for \( \mathbf{d} \) by inverting the non-rejection region of the multivariate test in a grid search over the support \( \Psi = [-0.5, 1.5] \times [-0.5, 1.5] \). Point estimates of the long-memory parameter vector can be generated by taking the values that minimize the absolute value of the test statistic, i.e., the values which provide maximum evidence for the null. This strategy has been intensively applied in the univariate context; see, for instance, Gil-Alaña and Robinson (1997). We shall denote the point estimates of the long-memory parameter for log-trading volatility and log-volatility obtained in this way as \( \hat{d}_{\text{min}}(vlm) \) and \( \hat{d}_{\text{min}}(\sigma) \), respectively.

For any of the stocks analyzed, and given the three alternative measures of volatility, Table 6 reports the main outcomes from this analysis, namely, point estimates \( \hat{d}_{\text{min}}(vlm) \) and \( \hat{d}_{\text{min}}(\sigma) \), the bounds of the 95% confidence ellipsoids, and the set of values for which the hypothesis that log-volume and log-volatility have the same order of fractional integration cannot be rejected at the 5% nominal size level. If this region is non-empty, the resultant set is a 95% confidence-interval estimate of the common order of fractional integration. In addition, given \( \hat{d}_{\text{min}}(vlm) \) and \( \hat{d}_{\text{min}}(\sigma) \), we can compute the residuals from the multivariate FGLS regression and estimate the contemporaneous correlation between the innovations to log-volume and a given volatility measure, \( \hat{\rho}_x \). Large values of this coefficient will give support to the usefulness of joint multivariate estimation over univariate analysis. For each stock in the sample, the last column of Table 6 reports these estimates, showing sizeable positive correlations around 40%.

![Insert Table 6 around here](image)

We first discuss the results from the analysis of the joint dynamics of log-volume and

\(^3\)In the univariate context, Demetrescu et al. (2008) argue that sample-dependent rules are more effective in controlling the empirical size of the regression-based test for fractional integration than standard information criteria. Schwert’s rule ensures a relatively large lag order selection, which provides a safe rule to guarantee that the short-run component of log-volume and log-realized volatility is captured in the auxiliary regression. Andersen et al. (2003) adopt a similar strategy in their analysis and use a relatively large number of lags in the estimation of the FIVAR model.
log absolute returns. Consistent with previous literature, the multivariate test generally rejects the hypotheses that the order of integration of the series is I(0) (weakly dependent process) or I(1) (unit root). The only exceptions are INTC (Intel) and MSFT (Microsoft), for which the multivariate test cannot reject that log-volume is I(0) at the 5% level. The cross-section average of $\hat{d}_{\text{min}}(vlm)$ and $\hat{d}_{\text{min}}(\sigma)$ is 0.41 and 0.39, respectively, matching the “characteristic” value 0.40 typically found in literature; see, Andersen et al. (2003). While for many stocks the point estimates of $d$ are below the stationary threshold, we note that most of the stocks the respective confidence sets include values on both the stationary and nonstationary regions, preventing us from drawing clear conclusions on the stationarity of the underlying series. Owing to the strong similitudes of the long-memory estimates in the bivariate system, the hypothesis that trading volume and volatility are driven by a FIVAR model with the same long-memory parameter can be rejected at the 5% level in only six cases. This represents a rejection frequency of 20.69% in our sample. This value is considerably greater than the 8% reported by Bollerslev and Jubinski (1999) and completely agrees with the 20% rejection rate reported by Lobato and Velasco (2000).

The results on log-volume and log absolute returns in Fleming and Kirby (2011) show a rejection rate of 100% because QML yields systematically larger parameter estimates for trading volume.

For illustrative purposes, the upper plot in Figure 1 shows the confidence ellipsoids at the 95%, 99% and 99% levels for the long-memory coefficients of the log-volume and log-volatility of the IBM stock, a representative example of the main evidence in the sample. Note that the three sets only include non-integer values, so the I(0) and I(1) hypotheses would be rejected at any of the usual confidence levels for this stock. Additionally, the

---

4These comparisons should be taken with caution because the samples involved include different stocks and different time periods. Nevertheless, it is tempting to conclude in favor of the superior efficiency of our procedure since we reach similar on-average conclusions to those in Lobato and Velasco (2000) on the basis of a sample period which roughly represents 3/8 of the sample period analyzed therein.

5Fleming and Kirby (2011) attribute this anomalous result to some form of bias and conjecture that departures from normality in log absolute returns may cause a pervasive effect on QML estimation; see Fleming and Kirby (2011, pp. 1721-1722). The simulations in the Monte Carlo section in our paper shed light on this issue owing to the similitudes between QML-based inference in the context of Nielsen (2004) and the Gaussian-based LM test in Nielsen (2005). Accordingly, it is the combination of persistent time-varying volatility and non-Gaussian features that can indeed introduce sizeable biases in conducting inference when conditional heteroskedasticity is neglected.
three sets include values along the 45° line (red dashed line), i.e., pairs of the long-memory coefficient for which \( H_0 : \theta = 0 \) cannot be rejected at the related nominal size level and which have the additional property of having the same value. The range of such values for a given confidence level represents a confidence-interval estimation of the common long-memory parameter in volume and volatility. Consequently, the hypothesis that the two series have the same order of fractional integration cannot be rejected at any of the usual confidence levels for this asset.

We now discuss the results associated to log-range and log-realized variance. The overall evidence that emerges is remarkably similar in both cases. Multivariate estimation provides even stronger support for fractional integration in this context, with the I(0) and I(1) hypotheses being rejected at the 5% level in all cases. For log-trading volume, the joint estimation of the order of integration leads to mild changes with respect to the results discussed previously. The average values across stocks of \( \hat{d}_{\min}(vlm) \) given the log-range and log-realized variance estimators are 0.45 and 0.46, respectively, which is fairly close to the 0.41 value reported before. In contrast, the estimates of the order of integration of volatility based on log-range or log-realized variance increase considerable. In both cases, the average value of \( \hat{d}_{\min}(\sigma) \) is 0.58 and the overall evidence is strongly suggestive that volatility displayed nonstationarity dynamics because the lower bound of the confidence set is not smaller than the 0.5 threshold in many cases. The evidence of nonstationary fractionally integrated dynamics in realized variance over this period (which included the financial crisis) is consistent with the results reported by Hassler et al. (2016). Consequently, and since the persistence of log-realized variance tends to be greater than that of log-volume, the hypothesis that both variables have a common degree of fractional integration in a FIVAR model is rejected at the 5% level in a significantly larger proportion of cases, namely, 51.07% (log-range) and 62.07% (log-realized variance).

The bottom part of Figure 1 shows the 99%, 95%, and 90% confidence ellipsoids of \( d \) for IBM based on log-volume and log-realized variance. Like in the case of absolute-valued returns, the I(0) and I(1) are strongly rejected at any of the usual confidence levels. In addition, the most noticeable feature is that all these sets are now above the 45° degrees line, suggesting that log-realized variance is significantly more persistent that volume at any of the usual confidence levels.

[Insert Figure 1 around here]
The econometrics literature has documented that measurement errors in absolute returns cause biases in (univariate) long-memory estimation. Essentially, log absolute returns are subject to noisy additive measurement errors with large variability, which make the underlying process to look as if it was less persistent than it actually is, leading to downward-biased estimates of the true order of integration; see, for instance, Bollerslev and Wright (2000), Haldrup and Nielsen (2007), and Dalla (2015). This feature provides a straightforward explanation to understand the systematic differences in long-memory estimation given the different volatility measures in Table 6. According to these results, there is a positive relationship between higher accuracy (efficiency) in volatility estimation and the rejection rates of the hypothesis of a common order of integration, suggesting that downward biases in the estimation of the long-memory parameter on absolute returns lead spuriously toward the evidence of a common order of integration. Using more accurate measures reduces biases in estimation and leads to new evidence suggesting that volatility is more persistent than volume. In summary, because the evidence from more efficient volatility measures must necessarily be deemed as more reliable, we conclude that the hypothesis that trading volume and volatility have the same order of integration is not supported empirically for most of the stocks analyzed in our sample.

4.3 Robustness checks

We checked the robustness of the main conclusion from the main analysis to the order of augmentation and to the inclusion of a time trend in the volatility measures. As in Fleming and Kirby (2011), we considered a linear time trend in volatility and a low-order VAR($p$). Table 7 reports the main results from this analysis, focusing directly on log-volume and log-realized variance, with $p = 2$ and with and without a trend in volatility. We also report the related results when $p$ is chosen according to Schwert’s rule and volatility includes a determinist trend, thereby completing the analysis reported in the previous section). While results are not totally insensitive to these considerations, they exhibit a considerable degree of resilience and the main qualitative picture that emerges is completely similar to that discussed previously.

[Insert Table 7 around here]

Additionally, in order to benchmark the empirical results with a different method and
provide additional evidence, for each stock in the sample we implemented the test in Hualde (2003) to address the null hypothesis of equality of long-memory coefficients in a bivariate system formed by log-volume and log-realized variance. Hualde’s test has the outstanding advantage of being valid both in the FIVARMA setting and in a fractionally cointegrated system but, unfortunately, it has an unusually low rate of divergence under the alternative, which implies little power in finite samples. In particular, for two series with order of integration $d_1$ and $d_2$, the test statistic is asymptotically normal distributed under the null $H_0: d_1 = d_2$, and diverges at rate $T^{d_1-d_2}$ when $d_1 \neq d_2$. Based on univariate parametric estimations of the long-memory coefficients obtained from inverting the test statistic in Demetrescu et al. (2008), the results from the implementation of this test (not reported here, but available under request) showed significant rejections at some of the usual confidence levels in 8 cases, which indicates an overall rejection rate of 27.59% in our sample.\footnote{Namely, CSCO, DIS, INTC, JNJ, MCD, TRV, VZ, WMT and XOM.} All the cases in which Hualde’s test was able to reject the null correspond to stocks for which our test failed to find a non-empty confidence interval for the common order of integration, thereby rejecting the null of common order of fractional integration as well. Consequently, for the stocks in this subsample, the agreement in conclusions from two independent statistical methods provides an extra layer of evidence suggesting that rejections are truly caused by differences in the order of integration. The overall evidence in this paper, therefore, supports the main conclusions in Fleming and Kirby (2011), who cast doubts on the general suitability of the long-run view of the MDH. The empirical dynamics of volume and volatility prove in many cases to be too complex as to generally accept that they can be explained by a bivariate FIVAR model with a single long-memory parameter.

5 Concluding remarks

The main contribution of this paper to the extant literature is to propose a new testing approach to infer fractional integration in a FIVAR multivariate context under fairly general conditions. These conditions permit conditional heteroskedasticity, do not require the order of integration to lie in a certain region (thereby allowing for both stationary and nonstationary dynamics) and do not impose a particular distribution, such as normality.
To the best of our knowledge, none of the existing methods in the extant literature achieves this degree of flexibility. The testing approach is a LM test that can readily be implemented from FGLS estimation in a regression-based context. We have analytically shown that the null asymptotic distribution is a Chi-square distribution, and that the test exhibits non-trivial power to reject a sequence of local alternatives.

Our approach can be used to test hypotheses involving the orders of fractional integration of multiple series in the FIVAR setting, or as an estimation procedure to generate confidence sets that include the unknown values with a given confidence level. Given the flexibility of the underlying assumptions, this testing approach is well suited to be implemented on data exhibiting the stylized features of financial data, and which typically include time-varying second-order moments, non-Gaussian features and long-range persistence. Monte Carlo analysis shows that the test ensures approximately correct size in finite samples exhibiting such characteristics, whereas alternative approaches that neglect conditional heteroskedasticity may lead to severe oversized inference. Furthermore, the test present comparable or improved power in comparison to alternatives that rely on prewhitening, since our test handles short-term dependencies more efficiently through augmentation.

Given the suitability of the test to be implemented on financial data, we have used it to jointly infer the order of fractional integration of trading volume and volatility in a sample of major stocks traded in the U.S. market. Volatility is proxied with three different measures in increasing degree of accuracy, namely, absolute returns, a range-based estimator, and realized variance computed over 5-minute returns. The evidence from the analysis on realized variance and range-based estimates leads to similar conclusions, namely, that for many stocks in the sample volatility is more persistent than trading volume. On the other hand, the analysis based on log absolute returns shows a different picture suggesting that volume and volatility have the same order of fractional integration. Because long-memory estimation in absolute returns is known to be downward-biased, this evidence is likely to be spuriously caused by measurement errors in the data.
References


A Technical Appendix

A.1 Preliminary Results

Before presenting the proofs of the main results put forward in the main text it will be

convenient to state first the following Lemmas:

Lemma A1. Let \( x_{st-1,d}^{**} := (x_{st-1,d,s}^{**}, \epsilon_{t-1,d}^{**}, \ldots, \epsilon_{t-p,d}^{**})^T \), with \( x_{st-1,d,s}^{**} := \sum_{l=1}^{\infty} l^{-1} \epsilon_{st-l,d,s} \), and \( \epsilon_{st,d,s} \) denoting the \( s \)-th element of \( \epsilon_{t,d} \), \( 1 \leq s \leq k \). Let \( \nu_{ij} \) be the \( ij \)-th element of \( \Sigma^{-1} \). Under Assumption 1 and \( H_0 : \theta = 0 \), \( \nu_{ij} (X_{it-1,d}^{**}X_{jt-1,d}^{**}) / T \xrightarrow{a.s.} \Omega_{\lambda_{ij}} \), with \( \Omega_{\lambda_{ij}} := \nu_{ij}E(x_{st-1,d}^{**}x_{st-1,d}^{**}) \) bounded, and bounded away from zero if \( \nu_{ij} \neq 0 \).

Proof of Lemma A1. Denote \( \lambda_{t-1,d} := (\epsilon_{t-1,d}^{**}, \ldots, \epsilon_{t-p,d}^{**})^T \), such that for any \( 1 \leq i, j \leq k \) we have

\[
T^{-1} X_{it-1,d}^{**}X_{jt-1,d}^{**} = \begin{bmatrix}
T^{-1} \sum_{t=p}^{T} z_{it-1,d}^{**}z_{jt-1,d}^{**} & T^{-1} \sum_{t=p}^{T} z_{it-1,d}^{**}\lambda_{t-1,d}^{**} \\
T^{-1} \sum_{t=p}^{T} \lambda_{t-1,d}^{**}z_{it-1,d}^{**} & T^{-1} \sum_{t=p}^{T} \lambda_{t-1,d}^{**}\lambda_{t-1,d}^{**}
\end{bmatrix}.
\]

Under Assumption 1 and \( H_0 : \theta = 0 \), \( \{X_{st-1,d}^{**}\} \) is a measurable function of a strictly stationary and ergodic process and, therefore, it is a strictly stationary and ergodic process as well, and so is \( \{X_{st-1,d}^{**}X_{st-1,d}^{**}\} \). The required result then follows from the Ergodic Theorem (ET) because \( x_{st-1,d}^{**} \) is (uniformly) \( L_2 \)-bounded, so the elements in \( \{x_{st-1,d}^{**}X_{st-1,d}^{**}\} \) have finite absolute expected values. To see this, first note that for any \( 1 \leq i \leq k \), there exists some finite \( K > 0 \) such that \( E(z_{it-1,d}^{**}2) = \sum_{l=1}^{\infty} \omega_{il} E(\epsilon_{t-l}^{2}) < K \), since \( \omega_{il} = O(1/l) \) and \( \{\epsilon_{t}\} \) is uniformly \( L_2 \)-bounded under (A2). From this result, it follows from the Cauchy-Schwarz inequality that \( E \left( |x_{st-1,d}^{**}z_{st-1,d}^{**}| \right) \leq \sqrt{E \left( x_{st-1,d}^{**2} \right) \sqrt{E \left( z_{st-1,d}^{**2} \right)}} < K \), \( 1 \leq i, j \leq k \). Similarly, since \( \lambda_{t-1,d} \) is uniformly \( L_2 \)-bounded under (A2), there exists some finite \( C > 0 \) for which \( E \|\lambda_{t-1,d}^{**}x_{st-1,d}^{**}\| \leq \sqrt{E \|\lambda_{t-1,d}^{**}\|^2 \sqrt{E \left( x_{st-1,d}^{**2} \right)}} < C \) and \( E \|\lambda_{t-1,d}^{**}x_{t-1,d}^{**}\| \leq E \|\lambda_{t-1,d}^{**}\|^2 < C \). Consequently, the ET ensures \( \nu_{ij} (X_{it-1,d}^{**}X_{jt-1,d}^{**}) / T \xrightarrow{a.s.} \nu_{ij}E(x_{it-1,d}^{**}x_{jt-1,d}^{**}) \). Finally, from stationarity, \( x_{st-1,d}^{**} = \sum_{l=1}^{\infty} \Gamma_{il} \Sigma_{jl}^{d} e_{j-l} \) with \( ||\Gamma_{il}|| = O(1/l) \), so \( \Omega_{\lambda_{ij}} = \nu_{ij} \sum_{l=1}^{\infty} \Gamma_{il} \Sigma_{jl}^{d} \) < \( \infty \). Clearly, the condition \( \Sigma > 0 \) rules out the degenerate case \( E(x_{it-1,d}^{**}x_{jt-1,d}^{**}) = 0 \), from which the required results follow. Furthermore, for \( i = j \), \( \Omega_{\lambda_{ij}} = \nu_{ii} \sum_{l=1}^{\infty} \Gamma_{il} \Sigma_{jl}^{d} \), so \( \Omega_{\lambda_{ij}} \) is positive definite (Davidson, 1999, Corollary 14.2.10, p.216).
Lemma A2. Under Assumption 1 and $H_0 : \theta = 0$, for $1 \leq i, j \leq k$, it follows that,

(i) $T^{-1} \| X_{it-1,d}^* X_{jt-1,d}^* - X_{it-1,d}^* X_{jt-1,d}^* \| = O_p \left( T^{-1/2} \right)$;

(ii) $T^{-1/2} \| (X_{it-1,d}^* - X_{it-1,d}^*) \mathbf{u}_{jt} \| = O_p \left( \frac{\log T}{\sqrt{T}} \right)$.

Proof of Lemma A2. For (i), we can write $T^{-1} \left( X_{it-1,d}^* X_{jt-1,d}^* - X_{it-1,d}^* X_{jt-1,d}^* \right)$ as

$$
\begin{bmatrix}
T^{-1} \sum_{t=p+1}^{T} \left( z_{it-1,d}^* z_{jt-1,dj}^* - z_{it-1,dj}^* z_{jt-1,d}^* \right) \\
T^{-1} \sum_{t=p+1}^{T} \lambda_{t-1,d} \left( z_{it-1,dj}^* - z_{it-1,dj}^* \right)
\end{bmatrix}
= O_p \left( T^{-1/2} \right)
$$

From the Cauchy-Schwarz inequality that,

$$
E \left[ T^{-1} \sum_{t=p+1}^{T} \lambda_{t-1,d} \left( z_{it-1,dj}^* - z_{it-1,dj}^* \right) \right] \leq T^{-1} \sum_{t=p+1}^{T} \sqrt{E \left[ \| \lambda_{t-1,d} \|^2 \right]} \sqrt{E \left[ \| z_{it-1,dj}^* - z_{it-1,dj}^* \|^2 \right]}
$$

$$
= O \left( T^{-1} \sum_{t=p+1}^{T} \frac{1}{\sqrt{t}} \right) = O \left( T^{-1/2} \right).
$$

Hence, $T^{-1} \sum_{t=p+1}^{T} \lambda_{t-1,d} \left( z_{it-1,dj}^* - z_{it-1,dj}^* \right) = O_p \left( T^{-1/2} \right)$ by Markov’s Theorem. Next, write $z_{st-1,d}^* = z_{st-1,d}^* + b_{s,t-1}$, with $b_{s,t-1} := \sum_{t=s}^{\infty} \omega_{st} e_{st-1}$. Since $\omega_{st} = O \left( 1/t \right)$ and $b_{s,t-1} = O_p \left( 1/\sqrt{T} \right)$, it follows that,

$$
z_{it-1,dj}^* = \left( z_{it-1,dj}^* + b_{i,t-1} \right) \left( z_{jt-1,dj}^* + b_{j,t-1} \right) = z_{it-1,dj}^* z_{jt-1,dj}^* + r_{ij,t-1}
$$

with $r_{ij,t-1} = O_p \left( 1/\sqrt{T} \right)$ defined implicitly. Therefore,

$$
E \left[ T^{-1} \sum_{t=p+1}^{T} \left( z_{it-1,dj}^* z_{jt-1,dj}^* - z_{it-1,dj}^* z_{jt-1,dj}^* \right) \right] \leq T^{-1} \sum_{t=p+1}^{T} E \left[ r_{ij,t-1} \right] = O \left( T^{-1/2} \right) = o(1)
$$

and the required results holds from Markov’s Theorem. For part (b), note that the column vector $(X_{it-1,d}^* - X_{it-1,d}^*) \mathbf{u}_{jt}$ has a first entry $T^{-1/2} \sum_{t=p+1}^{T} b_{i,t-1} e_{jt}$, and all the remaining values are equal to zero. Owing to the MDS property of $\{ e_t \}$ and stationarity in (A3), as well as the moment conditions in (A2) and (A4) it follows that,

$$
E \left[ T^{-1/2} \sum_{t=p+1}^{T} b_{i,t-1} e_{jt} \right]^2 = T^{-1} \sum_{t=p+1}^{T} E \left( b_{i,t-1}^2 e_{jt}^2 \right)
$$

$$
= T^{-1} \sum_{t=p+1}^{T} \sum_{l_1=t}^{\infty} \sum_{l_2=t}^{\infty} \omega_{it_1} \omega_{it_2} E \left( e_{jt}^2 e_{it_1} e_{it_2} \right) = O \left( \frac{\log T}{T} \right)
$$

31
since
\[
\sum_{l_1=t}^{\infty} \sum_{l_2=t}^{\infty} \omega_{l_1} \omega_{l_2} E \left( e_{jt}^2 e_{it-l_1} e_{it-l_2} \right) = \sum_{l_1=t}^{\infty} \omega_{l_1}^2 E \left( e_{jt}^2 \right) + \sum_{l_1=t}^{\infty} \sum_{l_2 \neq l_1}^{\infty} \omega_{l_1} \omega_{l_2} E \left( e_{jt}^2 e_{it-l_1} e_{it-l_2} \right)
\]
\[
= \mathcal{O} \left( \sum_{l_1=t}^{\infty} \frac{1}{l_1^2} \right) + o \left( \sum_{l_1=t}^{\infty} \frac{1}{l_1^2} \right) = \mathcal{O} \left( 1/t \right)
\]
given that \( E \left( e_{jt}^2 e_{it-l_1} \right) \leq \left( E \left( e_{jt}^4 \right) \right)^{1/2} < K \) for all \( t \), and (A4) implies that \( E \left( e_{jt}^2 e_{it-l_1} e_{it-l_2} \right) \leq E \left( |e_{jt}^2 e_{it-l_1} e_{it-l_2}| \right) = o \left( \frac{1}{l_1^2} \right) \) for any \( l_1, l_2 > 0, l_1 \neq l_2 \) under absolute summability; see Lemmas B.1i) and B.5 in Hassler et al. (2009). The required result holds from Markov’s Theorem. ■

**Lemma A3.** Let \( \hat{\Sigma} = \{ \hat{\sigma}_{ij} \} \) denote the LS-based estimator of \( \Sigma = \{ \sigma_{ij} \} \), namely, \( \hat{\sigma}_{ij} = T^{-1} \hat{u}_i \hat{u}_j' \), \( \hat{u}_i := Y_{s_t,d_s} - X_{s_t-1,d} \widetilde{\beta}_s \), with \( \widetilde{\beta}_s \) denoting the LS estimator of \( \beta_s \) in the corresponding equation. Then, under Assumption 1 and \( H_0 : \theta = 0 \),

1) \( \hat{\Sigma} \rightarrow \Sigma \);
2) \( T^{-1/2} X_{t-1,d}^* \left( \hat{\Sigma}^{-1} - \Sigma^{-1} \right) \otimes I_{T-p} X_{t-1,d}^* = \mathcal{O}_p \left( T^{-1/2} \right) \); and
3) \( T^{-1/2} X_{t-1,d}^* \left( \hat{\Sigma}^{-1} - \Sigma^{-1} \right) \otimes I_{T-p} u_t \rightarrow \mathcal{L} 0.

**Proof of Lemma A3.** Part i) follows from consistancy of the LS estimator \( \widetilde{\beta}_s \) in the equation-by-equation estimation under Assumption 1 and \( H_0 : \theta = 0 \), which can be proven as in Demetrescu et al. (2007). Part ii) follows from \( \sqrt{T} \)-consistency in i) because \( X_{t-1,d}^* \) is uniformly \( L_2 \)-bounded. Finally, for part iii), note that

\[
T^{-1/2} X_{t-1,d}^* \left( \hat{\Sigma}^{-1} - \Sigma^{-1} \right) \otimes I_{T-p} u_t = \begin{bmatrix}
\sum_{s=1}^{k} T^{-1/2} \left( \hat{\nu}_{1s} - \nu_{1s} \right) \left( \frac{1}{T} X_{t-1,d}^* u_{st} \right) \\
\sum_{s=1}^{k} T^{-1/2} \left( \hat{\nu}_{2s} - \nu_{2s} \right) \left( \frac{1}{T} X_{t-1,d}^* u_{st} \right) \\
\vdots \\
\sum_{s=1}^{k} T^{-1/2} \left( \hat{\nu}_{ks} - \nu_{ks} \right) \left( \frac{1}{T} X_{t-1,d}^* u_{st} \right)
\end{bmatrix}_{km \times 1}
\]
and the required result follows noting that \( T^{-1/2} \left( \hat{\nu}_{ij} - \nu_{ij} \right) = \mathcal{O}_p (1) \) for all \( 1 \leq i, j \leq k \) from i) above, whereas \( X_{t-1,d}^* u_{jt}/T \stackrel{a.s.}{\to} 0 \) from the ET, since \( \{ X_{t-1,d}^* u_{jt} \} \) is a strictly stationary and ergodic vector MDS, and \( E \| X_{t-1,d}^* u_{jt} \| \leq \sqrt{E \| X_{t-1,d}^* u_{jt} \|^2 E \| u_{jt} \|^2} < \infty \) under Assumption 1 and \( H_0 : \theta = 0 \). ■

**Lemma A4.** Define \( D_{rsij} := e_{rt} e_{st} X_{t-1,d}^* X_{j-1,d}^* \) for all \( 1 \leq r, s, i, j \leq k \). Under Assumption 1 and \( H_0 : \theta = 0 \), \( E \| D_{rsij} \| < \infty \).
**Proof of Lemma A4.** The proof follows from the Cauchy-Schwarz inequality given that $e_{it}x_{jt-1,d}^{**}$ is (uniformly) $L_2$-bounded for any $1 \leq i, j \leq k$. To see this, note that

$$E \| e_{it}x_{jt-1,d}^{**} \|^2 = E \| e_{it} \lambda_{t-1,d} \|^2 + E \left( e_{it}^2 r_{jt-1,d}^{2**} \right) < K$$

because for some finite $C > 0$,

$$E \| e_{it} \lambda_{t-1,d} \| \leq \sqrt{E \left( e_{it}^2 \right) E \| \lambda_{t-1,d} \|^2} < C$$

and

$$E \left( e_{it}^2 r_{jt-1,d}^{2**} \right) = \sum_{l_1=1}^{\infty} \sum_{l_2=1}^{\infty} \omega_{l_1} \omega_{l_2} E \left( e_{it}^2 r_{jt-1,l_1} r_{jt-1,l_2} \right)$$

$$= \sum_{l_1=1}^{\infty} \omega_{l_1}^2 E \left( e_{it}^2 r_{jt-1,l_1} r_{jt-1,l_1} \right) + \sum_{l_1=1}^{\infty} \sum_{l_2 \neq l_1}^{\infty} \omega_{l_1} \omega_{l_2} E \left( e_{it}^2 r_{jt-1,l_1} r_{jt-1,l_2} \right)$$

where $E \left( e_{it}^2 r_{jt-1,l_1} r_{jt-1,l_1} \right) \leq E \left( e_{it}^4 \right)^{1/4} \times E \left( e_{jt}^4 \right)^{3/4} < K$ and, as in Lemma A2,

$$\sum_{l_1=1}^{\infty} \sum_{l_2 \neq l_1}^{\infty} \omega_{l_1} \omega_{l_2} E \left( e_{it}^2 r_{jt-1,l_1} r_{jt-1,l_2} \right) = O \left( \sum_{l_1=1}^{\infty} \frac{1}{l_1^2} \right) + o \left( \sum_{l_1=1}^{\infty} \sum_{l_2 \neq l_1}^{\infty} \frac{1}{l_1 l_2^2} \right) = O \left( 1 \right).$$

Consequently, $E \| e_{it} e_{it} x_{it-1,d}^{**} x_{jt-1,d}^{**} \| \leq \sqrt{E \| e_{it} x_{it-1,d}^{**} \|^2} \sqrt{E \| e_{it} x_{jt-1,d}^{**} \|^2} < \infty$. \(\blacksquare\)

### A.2 Proofs of main results

**Proof of Theorem 1.** Under Assumption 1 and $H_0 : \theta = 0$, the FGLS estimator of $\beta$ can be written as

$$\hat{\beta} = \beta_0 + \left( X_{t-1,d}^{**} \Sigma^{-1} \otimes I_{T-p} X_{t-1,d}^{**} \right)^{-1} \left( X_{t-1,d}^{**} \Sigma^{-1} \otimes I_{T-p} u_t \right).$$

Using consecutively Lemmas A2 and A3, we have that,

$$\sqrt{T} \left( \hat{\beta} - \beta_0 \right) = \left( \frac{1}{T} X_{t-1,d}^{**} \Sigma^{-1} \otimes I_{T-p} X_{t-1,d}^{**} \right)^{-1} \left( \frac{1}{\sqrt{T}} X_{t-1,d}^{**} \Sigma^{-1} \otimes I_{T-p} u_t \right) + o_p \left( 1 \right).$$

Recall that $\nu_{ij}$ denotes the $(i, j)$-th element of $\Sigma^{-1}$, and define $A_{ij}^T := \frac{1}{T} X_{t-1,d}^{**} \Sigma^{-1} \otimes I_{T-p} X_{t-1,d}^{**}$, noting that $A_{ij}^{**}$ can be represented as a partitioned matrix with $ij$-block

33
From Lemma A4, \( \nu_{ij}X_{i,t-1,d}^*X_{j,t-1,d}^*/T \). From Lemma A2, \( \nu_{ij}X_{i,t-1,d}^*X_{j,t-1,d}^*/T \xrightarrow{a.s.} \Omega_{Aij} \) for all \( 1 \leq i, j \leq k \). Consequently, \( A_T^{**} \xrightarrow{a.s.} A_\beta \), where \( A_\beta \) is a partitioned matrix with \( ij \)-th submatrix given by \( \Omega_{Aij} \). Since the columns of \( A_\beta \) cannot be written as linear combinations of the other elements, \( \det(A_\beta) > 0 \), and consequently

\[
\left( \frac{1}{T} X_{i,t-1,d}^* \Sigma^{-1} \otimes I_{T-p} X_{j,t-1,d}^* \right)^{-1} \xrightarrow{a.s.} A_\beta^{-1}
\]

by Slutsky’s Theorem. We now discuss the asymptotic behavior of the the second term in (1). To this end, define the column vector \( \mathbf{w}_{t-1}^* := X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p} \mathbf{u}_t \), noting that

\[
\mathbf{w}_{t-1}^* = \left( \sum_{s=1}^{k} \nu_{1s} X_{1,t-1,d} \mathbf{u}_{st}, \sum_{s=1}^{k} \nu_{2s} X_{2,t-1,d} \mathbf{u}_{st}, \ldots, \sum_{s=1}^{k} \nu_{ks} X_{k,t-1,d} \mathbf{u}_{st} \right)'
\]

with \( \mathbf{u}_{st} = (e_{s1}, \ldots, e_{sT})' \) under \( H_0 : \theta = 0 \). Given that \( E(\mathbf{w}_{t-1}^* | \mathcal{F}_{t-1}) = 0 \) and \( \mathbf{w}_{t-1}^* \) is a measurable function of \( \{e_i\} \), \( \{\mathbf{w}_{t-1}^*, \mathcal{F}_t\} \) is a strictly stationary and ergodic vector MDS. The covariance matrix of \( \mathbf{w}_{t-1}^* \) is \( \mathbf{B}_\beta := E(\mathbf{w}_{t-1}^* \mathbf{w}_{t-1}^{**,T}) \), which can be represented as a partitioned matrix with \( ij \)-th block \( \Omega_{Bij} \) given by

\[
\Omega_{Bij} := \sum_{r=1}^{k} \sum_{s=1}^{k} \nu_{ir} \nu_{is} E(e_{rt}e_{st} X_{i,t-1,d}^* X_{j,t-1,d}^*)
\]

From Lemma A4, \( E(\|e_{rt}e_{st}X_{i,t-1,d}^* X_{j,t-1,d}^*\|) < \infty \), and consequently \( \Omega_{Bij} < \infty \) for all \( 1 \leq i, j \leq k \), so \( \mathbf{B}_\beta < \infty \). Furthermore, the condition \( \Sigma > 0 \) trivially rules out the degenerate case \( \|\Omega_{B\beta}\| = 0 \), so \( \mathbf{B}_\beta \) is bounded away from zero. Consequently, the Central Limit Theorem (CLT) for MDS (Davidson, 1994, Theorem 24.3) and the Cramér-Wold device ensure that

\[
\sqrt{T} \left( \mathbf{w}_{t-1} - \mathbf{w}_{t-1}^{**} \right) \Rightarrow \mathcal{N}(0, \mathbf{B}_\beta)
\]

and so we conclude that,

\[
\sqrt{T} \left( \mathbf{\hat{w}} - \mathbf{\beta}_0 \right) \Rightarrow \mathcal{N}(0, \mathbf{A}^{-1}_\beta \mathbf{B}_\beta \mathbf{A}_\beta^{-1})
\]

as required. ■

**Proof of Theorem 2.** We first prove the convergence under the null hypothesis, which follows directly from Theorem 1 and the consistency of \( \hat{\Omega}_\beta := A_T^{**} B_T^{**} A_T^{**} \), with \( A_T^{**} := \frac{1}{T} X_{i,t-1,d}^* \Sigma^{-1} \otimes I_{T-p} X_{j,t-1,d}^* \) and \( B_T^{**} := \frac{1}{T} \mathbf{w}_{t-1}^* \mathbf{w}_{t-1}^{**,T} \). As discussed previously, \( A_T^{**} \xrightarrow{a.s.} A_\beta \), so Lemma A2 and the AEL allow us
to conclude that \( A_T \xrightarrow{p} A_\beta \). From \( \sqrt{T} \)-consistency in Theorem 1, \( \hat{u}_t := u_t + o_p(T^{-1/2}) \) and, therefore, \( w_T^{*} := X_{t-1,d}^* \hat{\Sigma}^{-1} \otimes I_{T-p}u_t + o_p(T^{-1/2}) \). Then, applying consecutively Lemmas A2 and A3, we can write \( B_T^{*} = B_T^{**} + o_p(1) \), where \( B_T^{**} := w_T^{**}w_T^{**}T \) can be represented as a partitioned matrix with \( ij \)-th block

\[
\mathbf{B}_{ij}^{**} := \sum_{r=1}^k \sum_{s=1}^k \nu_{irs} \left[ \frac{1}{T} X_{it-1,d}^* u_{rt} u_{st} X_{jt-1,d}^* \right]
\]

such that

\[
E \left( \mathbf{B}_{ij}^{**} \right) = \sum_{r=1}^k \sum_{s=1}^k \nu_{irs} \left( T^{-1} \sum_{t=p+1}^T E \left( e_{rt} \epsilon_{st} X_{it-1,d}^* X_{jt-1,d}^* \right) \right) = \Omega_{B_{ij}}
\]

from stationarity and the MDS property of \( \{ e_t \} \). Since \( \{ w_T^{**}, w_T^{**} \} \) is strictly stationary, ergodic, and \( L_2 \)-bounded by Lemma A6, the ET ensures that \( B_T^{**} \xrightarrow{a.s.} B_\beta \), so the AEL implies \( B_T^{**} \xrightarrow{P} B_\beta \). By Slutsky’s Theorem, \( \sqrt{T} \mathbf{R} \hat{\beta} \Rightarrow \mathcal{N} (0, \mathbf{R} \Omega_\beta \mathbf{R}') \), and since \( \mathbf{R} \Omega_\beta \mathbf{R}' \) is symmetric and nonnegative, there exists an upper triangular matrix \( \mathbf{L} \) such that \( \mathbf{R} \Omega_\beta \mathbf{R}' = \mathbf{L}' \mathbf{L} \). Consequently, \( \sqrt{T} \mathbf{L}^{-1} \mathbf{R} \hat{\beta} \Rightarrow \mathcal{N} (0, \mathbf{I}) \), and, hence,

\[
LM_d = T \left[ \mathbf{R} \hat{\beta} \right]' \left( \mathbf{R} \Omega_\beta \mathbf{R}' \right)^{-1} \left[ \mathbf{R} \hat{\beta} \right] + o_p(T^{-1/2})
\]

\[
= T \left[ \mathbf{L}'^{-1} \mathbf{R} \hat{\beta} \right]' \left[ \mathbf{L}'^{-1} \mathbf{R} \hat{\beta} \right] \Rightarrow \chi^2 (k)
\]

We now show the asymptotic convergence under the alternative hypothesis. Let \( c := (c_1, ..., c_k)' \), with \( ||c|| > 0 \). Under Assumption 1 and \( H_1 : \theta = c / \sqrt{T} \), we have \( Y_{t,d} = X_{t-1,d} \beta_0 + u_{\theta t} \) where \( u_{\theta t} := (u'_{\theta t1}, ..., u'_{\theta tk}) \), \( u_{\theta st} := u_{st} + \frac{1}{T} \sum_{t=1}^{T} X_{t-1,d}^* \theta \psi_{cs} \), and

\[
\psi_{cs} := (c_s, -\pi_{s1} \otimes c', ..., -\pi_{sp} \otimes c')'
\]

for \( 1 \leq s \leq k \); see Tanaka (1990). In this context, the FGLS estimator is given by

\[
\hat{\beta} = \beta_0 + \left( X_{t-1,d}^* \hat{\Sigma}^{-1} \otimes I_{T-p}X_{t-1,d}^* \right)^{-1} \left( \frac{1}{T} X_{t-1,d}^* \hat{\Sigma}^{-1} \otimes I_{T-p}u_{\theta t} \right) + o_p(1). \]

Lemma A3i) applies under the alternative hypothesis because, although the LS estimator is no longer consistent, it still follows that \( \hat{u}_{st} = u_{st} + O_p(T^{-1/2}) \) and, hence, \( \hat{\nu}_{ij} = \nu_{ij} + O(T^{-1/2}) \). Consequently,

\[
\sqrt{T} \left( \hat{\beta} - \beta_0 \right) = \left( \frac{1}{T} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p}X_{t-1,d}^* \right)^{-1} \left( \frac{1}{\sqrt{T}} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p}u_{\theta t} \right) + o_p(1). \]

Define \( \psi := (\psi'_{e1}, ..., \psi'_{ek})' \). Then,

\[
\frac{1}{\sqrt{T}} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p}u_{\theta t} = \frac{1}{T} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p}X_{t-1,d}^* \theta \psi_c + \frac{1}{\sqrt{T}} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p}u_{t}
\]
where we can show that,
\[
\frac{1}{T} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p} X_{t-1,d+\theta}^* \psi_c \xrightarrow{p} A_\beta \psi_c
\]
from the ET and the EL because
\[
T^{-1} \| (X_{t-1,d}^{**} - X_{t-1,d+\theta}^{**}) \Sigma^{-1} \otimes I_{T-p} (X_{t-1,d+\theta}^{**} - X_{t-1,d+\theta}^{*}) \| = O_p \left( \frac{\log T}{\sqrt{T}} \right) = o_p(1).
\]
Similarly,
\[
\frac{1}{\sqrt{T}} X_{t-1,d}^{**} \Sigma^{-1} \otimes I_{T-p} u_{st} \Rightarrow \mathcal{N}(0, B_\beta).
\]
from the CLT for MDS, given that
\[
T^{-1/2} \| (X_{t-1,d}^{**} - X_{t-1,d+\theta}^{**}) \Sigma^{-1} \otimes I_{T-p} u_{st} \| = O_p \left( \frac{\log T}{\sqrt{T}} \right)
\]
and, finally,
\[
\frac{1}{T} X_{t-1,d}^* \Sigma^{-1} \otimes I_{T-p} X_{t-1,d}^* \xrightarrow{p} A_\beta.
\]
Consequently, under Assumption 1 and $H_1: \theta = c/\sqrt{T}$,
\[
\sqrt{T} (\hat{\beta} - \beta_0) \Rightarrow \mathcal{N}(\psi_c, \Omega_\beta).
\]
Since $A_T^* \xrightarrow{p} \Omega_A$ and the previous results ensures $\hat{u}_t = u_t + O_p(T^{-1/2}), B_T^* \xrightarrow{p} \Omega_B$.
Finally, since $\sqrt{T} \hat{R} \beta \Rightarrow \mathcal{N}(c, \Omega_\beta)$ and, hence, $\sqrt{T} \left[ L^{-1} \hat{R} \beta \right] \Rightarrow \mathcal{N}(c, I)$, the result $\Omega_T \xrightarrow{p} \Omega_\beta$ implies
\[
LM_d = T \left[ L^{-1} \hat{R} \beta \right]' \left[ L^{-1} \hat{R} \beta \right] + O_p(T^{-1/2})
\]
\[
\Rightarrow \chi^2_{[k, ||c||^2)}
\]
as required. $\blacksquare$
B Figures and tables

Figure 1. Confidence sets at the 90%, 95%, and 99% confidence levels of the long-memory parameters of log-trading volume and log-volatility of IBM, with the latter proxied by absolute-valued returns (top) and realized variance (bottom). The sets are obtained from the empirical level curves of the $L_{d}^{FGLS}$ test statistic evaluated at different values given the sample observations, with level curves corresponding to the 90%, 95%, and 99% percentiles of $\chi^2(2)$, namely, 4.61, 5.99, and 9.21, respectively. The central point denotes the coordinates given by $\hat{d}_{\min}(vlm)$ and $\hat{d}_{\min}(\sigma)$. Finally, the red dashed line represents the 45° line.
Table 1. Rejection frequencies (empirical sizes) for $\theta_1 = \theta_2 = 0$ at the 5% significance level of the joint multivariate tests introduced in this paper, $LM^{\text{FGLS}}_d$, as well as the LM tests in Nielsen (2005), $LM^{\text{MLE}}_d$, and the trace-type test in Breitung and Hassler (2002), $BH_d$, given different values for the contemporaneous correlation parameter $\rho$, the GARCH parameters ($\alpha, \beta$), and sample lengths $T$. Innovations are drawn from either a multivariate normal distribution or a multivariate Student-$t$ distribution with 5 degrees of freedom.

<table>
<thead>
<tr>
<th>$\theta_1 = \theta_2 = 0$</th>
<th>Gaussian Innovations</th>
<th>Student-$t$ Innovations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$LM^{\text{FGLS}}_d$</td>
<td>$LM^{\text{MLE}}_d$</td>
</tr>
<tr>
<td>$\alpha = 0, \beta = 0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.0$</td>
<td>0.051</td>
<td>0.045</td>
</tr>
<tr>
<td>$0.2$</td>
<td>0.050</td>
<td>0.047</td>
</tr>
<tr>
<td>$0.4$</td>
<td>0.053</td>
<td>0.050</td>
</tr>
<tr>
<td>$0.6$</td>
<td>0.055</td>
<td>0.051</td>
</tr>
<tr>
<td>$0.8$</td>
<td>0.055</td>
<td>0.050</td>
</tr>
<tr>
<td>$\alpha = 0.1, \beta = 0.5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.0$</td>
<td>0.054</td>
<td>0.071</td>
</tr>
<tr>
<td>$0.2$</td>
<td>0.053</td>
<td>0.072</td>
</tr>
<tr>
<td>$0.4$</td>
<td>0.054</td>
<td>0.072</td>
</tr>
<tr>
<td>$0.6$</td>
<td>0.055</td>
<td>0.074</td>
</tr>
<tr>
<td>$0.8$</td>
<td>0.056</td>
<td>0.075</td>
</tr>
<tr>
<td>$\alpha = 0.1, \beta = 0.7$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.0$</td>
<td>0.053</td>
<td>0.080</td>
</tr>
<tr>
<td>$0.2$</td>
<td>0.054</td>
<td>0.080</td>
</tr>
<tr>
<td>$0.4$</td>
<td>0.054</td>
<td>0.080</td>
</tr>
<tr>
<td>$0.6$</td>
<td>0.055</td>
<td>0.080</td>
</tr>
<tr>
<td>$0.8$</td>
<td>0.056</td>
<td>0.080</td>
</tr>
<tr>
<td>$\alpha = 0.1, \beta = 0.8$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.0$</td>
<td>0.052</td>
<td>0.088</td>
</tr>
<tr>
<td>$0.2$</td>
<td>0.052</td>
<td>0.088</td>
</tr>
<tr>
<td>$0.4$</td>
<td>0.052</td>
<td>0.089</td>
</tr>
<tr>
<td>$0.6$</td>
<td>0.056</td>
<td>0.091</td>
</tr>
<tr>
<td>$0.8$</td>
<td>0.055</td>
<td>0.093</td>
</tr>
<tr>
<td>$\alpha = 0.1, \beta = 0.85$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0$</td>
<td>0.053</td>
<td>0.103</td>
</tr>
<tr>
<td>$0.2$</td>
<td>0.050</td>
<td>0.104</td>
</tr>
<tr>
<td>$0.4$</td>
<td>0.054</td>
<td>0.103</td>
</tr>
<tr>
<td>$0.6$</td>
<td>0.056</td>
<td>0.107</td>
</tr>
<tr>
<td>$0.8$</td>
<td>0.058</td>
<td>0.114</td>
</tr>
</tbody>
</table>

Note: The table shows rejection frequencies for different values of $\rho$, $\alpha$, and $\beta$, and sample lengths $T=500$ and $T=1000$. The columns for $LM^{\text{FGLS}}_d$, $LM^{\text{MLE}}_d$, and $BH_d$ represent the rejection frequencies for each test statistic.
Table 2. Rejection frequencies at the 5% nominal size level for $\theta_2 = 0$ and the sequence of values $\theta_1 = 0$ (empirical size) and $|\theta_1| > 0$ (empirical power) of the joint multivariate tests introduced in this paper, $LM_d^FGLS$, as well as the joint LM test introduced by Nielsen (2005), $LM_d^{MLE}$, and the joint test in Breitung and Hassler (2002), $BH_d$, given the correlation coefficient $\rho$, and the sample length $T$. Short-term errors in the DGP obey VAR(1) dynamics with on-diagonal coefficients $\pi_1$ and $\pi_2$ and off-diagonal coefficients $\pi_{12} = \pi_{21} = 0$. The $LM_d^FGLS$ test statistic is computed from an augmented auxiliary regression with one lag of the dependent variables. $LM_d^{MLE}$ and $BH_d$ are computed from VAR(1) residuals. Innovations are drawn from a multivariate Gaussian distribution.

<table>
<thead>
<tr>
<th>$\theta_1$</th>
<th>$\rho$</th>
<th>$LM_d^FGLS$</th>
<th>$LM_d^{MLE}$</th>
<th>$BH_d$</th>
<th>$LM_d^FGLS$</th>
<th>$LM_d^{MLE}$</th>
<th>$BH_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.30</td>
<td>0.0</td>
<td>1.000</td>
<td>0.999</td>
<td>0.994</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>-0.25</td>
<td>0.0</td>
<td>0.987</td>
<td>0.982</td>
<td>0.946</td>
<td>1.000</td>
<td>1.000</td>
<td>0.999</td>
</tr>
<tr>
<td>-0.20</td>
<td>0.0</td>
<td>0.899</td>
<td>0.879</td>
<td>0.781</td>
<td>0.999</td>
<td>0.999</td>
<td>0.990</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.0</td>
<td>0.644</td>
<td>0.626</td>
<td>0.493</td>
<td>0.941</td>
<td>0.934</td>
<td>0.868</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.0</td>
<td>0.338</td>
<td>0.321</td>
<td>0.239</td>
<td>0.609</td>
<td>0.601</td>
<td>0.478</td>
</tr>
<tr>
<td>-0.05</td>
<td>0.0</td>
<td>0.118</td>
<td>0.112</td>
<td>0.091</td>
<td>0.183</td>
<td>0.179</td>
<td>0.134</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.057</td>
<td>0.051</td>
<td>0.054</td>
<td>0.052</td>
<td>0.050</td>
<td>0.049</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0</td>
<td>0.121</td>
<td>0.112</td>
<td>0.091</td>
<td>0.179</td>
<td>0.174</td>
<td>0.142</td>
</tr>
<tr>
<td>0.10</td>
<td>0.0</td>
<td>0.321</td>
<td>0.311</td>
<td>0.243</td>
<td>0.584</td>
<td>0.569</td>
<td>0.472</td>
</tr>
<tr>
<td>0.15</td>
<td>0.0</td>
<td>0.600</td>
<td>0.581</td>
<td>0.498</td>
<td>0.888</td>
<td>0.876</td>
<td>0.814</td>
</tr>
<tr>
<td>0.20</td>
<td>0.0</td>
<td>0.812</td>
<td>0.788</td>
<td>0.701</td>
<td>0.983</td>
<td>0.978</td>
<td>0.957</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0</td>
<td>0.926</td>
<td>0.895</td>
<td>0.831</td>
<td>0.998</td>
<td>0.996</td>
<td>0.988</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0</td>
<td>0.969</td>
<td>0.942</td>
<td>0.902</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\pi_1 = \pi_2 = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LM_d^FGLS$</td>
</tr>
<tr>
<td>0.875</td>
</tr>
<tr>
<td>0.720</td>
</tr>
<tr>
<td>0.496</td>
</tr>
<tr>
<td>0.289</td>
</tr>
<tr>
<td>0.154</td>
</tr>
<tr>
<td>0.074</td>
</tr>
<tr>
<td>0.053</td>
</tr>
<tr>
<td>0.080</td>
</tr>
<tr>
<td>0.145</td>
</tr>
<tr>
<td>0.272</td>
</tr>
<tr>
<td>0.392</td>
</tr>
<tr>
<td>0.496</td>
</tr>
<tr>
<td>0.574</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\pi_1 = \pi_2 = 0.4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LM_d^FGLS$</td>
</tr>
<tr>
<td>0.998</td>
</tr>
<tr>
<td>0.958</td>
</tr>
<tr>
<td>0.908</td>
</tr>
<tr>
<td>0.678</td>
</tr>
<tr>
<td>0.350</td>
</tr>
<tr>
<td>0.124</td>
</tr>
<tr>
<td>0.057</td>
</tr>
<tr>
<td>0.117</td>
</tr>
<tr>
<td>0.324</td>
</tr>
<tr>
<td>0.567</td>
</tr>
<tr>
<td>0.743</td>
</tr>
<tr>
<td>0.845</td>
</tr>
<tr>
<td>0.898</td>
</tr>
</tbody>
</table>
Table 3. Rejection frequencies at the 5% nominal size level for $\theta_2 = 0$ and the sequence of values $\theta_1 = 0$ (empirical size) and $|\theta_1| > 0$ (empirical power) of $LM_{d}^{FGLS}$, $LM_{d}^{MLE}$, and $BH_d$, given the correlation coefficient $\rho$. Short-term errors are driven from a VAR(1) model with $\pi_1 = \pi_2 = 0.4$ and GARCH innovations with parameters $\alpha$ and $\beta$. The $LM_{d}^{FGLS}$ test statistic is computed from an augmented auxiliary regression with one lag of the dependent variables. $LM_{d}^{MLE}$ and $BH_d$ are computed from VAR(1) residuals. The sample length is $T=500$.

<table>
<thead>
<tr>
<th>$\theta_1$</th>
<th>$\rho$</th>
<th>GARCH: $\alpha = 0.1, \beta = 0.8$</th>
<th>GARCH: $\alpha = 0.1, \beta = 0.85$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Gaussian</td>
<td>Student-t</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$LM_{d}^{FGLS}$</td>
<td>$LM_{d}^{MLE}$</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.0</td>
<td>0.795</td>
<td>0.859</td>
</tr>
<tr>
<td>-0.25</td>
<td>0.0</td>
<td>0.625</td>
<td>0.718</td>
</tr>
<tr>
<td>-0.20</td>
<td>0.0</td>
<td>0.419</td>
<td>0.512</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.0</td>
<td>0.250</td>
<td>0.328</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.0</td>
<td>0.142</td>
<td>0.182</td>
</tr>
<tr>
<td>-0.05</td>
<td>0.0</td>
<td>0.076</td>
<td>0.102</td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
<td>0.051</td>
<td>0.072</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0</td>
<td>0.074</td>
<td>0.102</td>
</tr>
<tr>
<td>0.10</td>
<td>0.0</td>
<td>0.136</td>
<td>0.164</td>
</tr>
<tr>
<td>0.15</td>
<td>0.0</td>
<td>0.251</td>
<td>0.266</td>
</tr>
<tr>
<td>0.20</td>
<td>0.0</td>
<td>0.358</td>
<td>0.334</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0</td>
<td>0.461</td>
<td>0.377</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0</td>
<td>0.543</td>
<td>0.346</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.8</td>
<td>0.992</td>
<td>0.998</td>
</tr>
<tr>
<td>-0.25</td>
<td>0.8</td>
<td>0.954</td>
<td>0.981</td>
</tr>
<tr>
<td>-0.20</td>
<td>0.8</td>
<td>0.836</td>
<td>0.901</td>
</tr>
<tr>
<td>-0.15</td>
<td>0.8</td>
<td>0.588</td>
<td>0.603</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.8</td>
<td>0.300</td>
<td>0.378</td>
</tr>
<tr>
<td>-0.05</td>
<td>0.8</td>
<td>0.121</td>
<td>0.159</td>
</tr>
<tr>
<td>0</td>
<td>0.8</td>
<td>0.054</td>
<td>0.077</td>
</tr>
<tr>
<td>0.05</td>
<td>0.8</td>
<td>0.103</td>
<td>0.152</td>
</tr>
<tr>
<td>0.10</td>
<td>0.8</td>
<td>0.296</td>
<td>0.351</td>
</tr>
<tr>
<td>0.15</td>
<td>0.8</td>
<td>0.519</td>
<td>0.550</td>
</tr>
<tr>
<td>0.20</td>
<td>0.8</td>
<td>0.696</td>
<td>0.689</td>
</tr>
<tr>
<td>0.25</td>
<td>0.8</td>
<td>0.809</td>
<td>0.763</td>
</tr>
<tr>
<td>0.30</td>
<td>0.8</td>
<td>0.870</td>
<td>0.735</td>
</tr>
</tbody>
</table>
Table 4. Rejection frequencies at the 5% nominal size level for $\theta_2 = 0$ and the sequence of values $\theta_1 = 0$ (empirical size) and $|\theta_1| > 0$ (empirical power) of $LM_{d}^{FGLS}$, $LM_{d}^{MLE}$, and $BH_d$, given the correlation coefficient $\rho$. Short-term errors are driven from a VAR(1) model with $\pi_1 = \pi_2 = 0.4$ and GARCH innovations with parameters $\alpha$ and $\beta$. The $LM_{d}^{FGLS}$ test statistic is computed from an augmented auxiliary regression with one lag of the dependent variables. $LM_{d}^{MLE}$ and $BH_d$ are computed from VAR(1) residuals. The sample length is $T=1000$.

| $T=1000$ | \begin{tabular}{c c c c c c c c c c} \hline \ \hline $\theta_1$ & $\rho$ & $LM_{d}^{FGLS}$ & $LM_{d}^{MLE}$ & $BH_d$ & $LM_{d}^{FGLS}$ & $LM_{d}^{MLE}$ & $BH_d$ & $LM_{d}^{FGLS}$ & $LM_{d}^{MLE}$ & $BH_d$ \\ \hline -0.30 & 0.0 & 0.983 & 0.993 & 0.983 & 0.753 & 0.961 & 0.949 & 0.967 & 0.991 & 0.979 \\ -0.25 & 0.0 & 0.916 & 0.959 & 0.920 & 0.632 & 0.915 & 0.884 & 0.891 & 0.956 & 0.914 \\ -0.20 & 0.0 & 0.744 & 0.841 & 0.744 & 0.457 & 0.811 & 0.754 & 0.702 & 0.836 & 0.742 \\ -0.15 & 0.0 & 0.471 & 0.597 & 0.481 & 0.287 & 0.645 & 0.592 & 0.437 & 0.603 & 0.489 \\ -0.10 & 0.0 & 0.237 & 0.316 & 0.248 & 0.155 & 0.457 & 0.415 & 0.220 & 0.332 & 0.269 \\ -0.05 & 0.0 & 0.093 & 0.132 & 0.107 & 0.080 & 0.302 & 0.286 & 0.093 & 0.150 & 0.124 \\ 0.0 & 0.0 & 0.051 & 0.072 & 0.066 & 0.058 & 0.243 & 0.237 & 0.053 & 0.089 & 0.083 \\ 0.05 & 0.0 & 0.091 & 0.122 & 0.098 & 0.077 & 0.262 & 0.247 & 0.084 & 0.139 & 0.112 \\ 0.10 & 0.0 & 0.246 & 0.284 & 0.220 & 0.176 & 0.385 & 0.353 & 0.226 & 0.293 & 0.228 \\ 0.15 & 0.0 & 0.437 & 0.461 & 0.376 & 0.288 & 0.508 & 0.470 & 0.414 & 0.464 & 0.384 \\ 0.20 & 0.0 & 0.635 & 0.603 & 0.523 & 0.434 & 0.607 & 0.557 & 0.607 & 0.607 & 0.531 \\ 0.25 & 0.0 & 0.756 & 0.653 & 0.576 & 0.588 & 0.670 & 0.625 & 0.729 & 0.652 & 0.578 \\ 0.30 & 0.0 & 0.830 & 0.620 & 0.559 & 0.673 & 0.664 & 0.623 & 0.811 & 0.621 & 0.555 \\ \hline \end{tabular} | \begin{tabular}{c c c c c c c c c c} \hline \ \hline $\theta_1$ & $\rho$ & $LM_{d}^{FGLS}$ & $LM_{d}^{MLE}$ & $BH_d$ & $LM_{d}^{FGLS}$ & $LM_{d}^{MLE}$ & $BH_d$ & $LM_{d}^{FGLS}$ & $LM_{d}^{MLE}$ & $BH_d$ \\ \hline -0.30 & 0.8 & 1.000 & 1.000 & 1.000 & 0.929 & 0.998 & 0.999 & 1.000 & 1.000 & 1.000 \\ -0.25 & 0.8 & 1.000 & 1.000 & 1.000 & 0.927 & 0.999 & 0.999 & 0.997 & 1.000 & 1.000 \\ -0.20 & 0.8 & 0.988 & 0.998 & 0.993 & 0.752 & 0.965 & 0.965 & 0.977 & 0.995 & 0.992 \\ -0.15 & 0.8 & 0.893 & 0.951 & 0.906 & 0.555 & 0.876 & 0.860 & 0.846 & 0.940 & 0.898 \\ -0.10 & 0.8 & 0.535 & 0.658 & 0.554 & 0.295 & 0.670 & 0.635 & 0.491 & 0.655 & 0.560 \\ -0.05 & 0.8 & 0.164 & 0.228 & 0.180 & 0.112 & 0.386 & 0.391 & 0.155 & 0.254 & 0.207 \\ 0.0 & 0.8 & 0.058 & 0.083 & 0.080 & 0.057 & 0.271 & 0.296 & 0.058 & 0.104 & 0.099 \\ 0.05 & 0.8 & 0.175 & 0.228 & 0.180 & 0.113 & 0.349 & 0.352 & 0.162 & 0.246 & 0.196 \\ 0.10 & 0.8 & 0.510 & 0.578 & 0.477 & 0.307 & 0.591 & 0.569 & 0.471 & 0.574 & 0.485 \\ 0.15 & 0.8 & 0.822 & 0.848 & 0.780 & 0.542 & 0.771 & 0.760 & 0.781 & 0.839 & 0.768 \\ 0.20 & 0.8 & 0.940 & 0.940 & 0.908 & 0.714 & 0.873 & 0.860 & 0.916 & 0.928 & 0.897 \\ 0.25 & 0.8 & 0.977 & 0.969 & 0.949 & 0.819 & 0.911 & 0.903 & 0.964 & 0.961 & 0.943 \\ 0.30 & 0.8 & 0.993 & 0.958 & 0.943 & 0.885 & 0.900 & 0.908 & 0.987 & 0.946 & 0.930 \\ \hline \end{tabular} |
Table 5. Descriptive statistics (mean, standard deviation, minimum, maximum, skewness and kurtosis of the log-volume, log absolute returns, log range estimator and log-realized variance.

<table>
<thead>
<tr>
<th>Ticker</th>
<th>Company</th>
<th>Log Trading Volume</th>
<th>Log Absolute Returns</th>
<th>Log Range</th>
<th>Log Realized Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPL</td>
<td>16.70</td>
<td>0.74</td>
<td>20.36</td>
<td>14.11</td>
<td>14.14</td>
</tr>
<tr>
<td>AXP</td>
<td>15.65</td>
<td>0.62</td>
<td>18.32</td>
<td>13.70</td>
<td>0.63</td>
</tr>
<tr>
<td>BA</td>
<td>15.32</td>
<td>0.45</td>
<td>17.61</td>
<td>13.13</td>
<td>0.34</td>
</tr>
<tr>
<td>CXT</td>
<td>15.17</td>
<td>0.64</td>
<td>18.03</td>
<td>12.87</td>
<td>-0.14</td>
</tr>
<tr>
<td>CSKO</td>
<td>17.69</td>
<td>0.41</td>
<td>20.15</td>
<td>17.58</td>
<td>0.25</td>
</tr>
<tr>
<td>CVX</td>
<td>15.79</td>
<td>0.30</td>
<td>17.69</td>
<td>13.52</td>
<td>0.35</td>
</tr>
<tr>
<td>DRS</td>
<td>16.01</td>
<td>0.41</td>
<td>18.56</td>
<td>14.20</td>
<td>0.44</td>
</tr>
<tr>
<td>GE</td>
<td>17.65</td>
<td>0.60</td>
<td>20.44</td>
<td>15.54</td>
<td>0.56</td>
</tr>
<tr>
<td>GS</td>
<td>15.51</td>
<td>0.30</td>
<td>18.56</td>
<td>13.46</td>
<td>0.73</td>
</tr>
<tr>
<td>HB</td>
<td>16.16</td>
<td>0.52</td>
<td>18.36</td>
<td>14.41</td>
<td>0.20</td>
</tr>
<tr>
<td>IBM</td>
<td>15.55</td>
<td>0.44</td>
<td>17.24</td>
<td>14.16</td>
<td>0.29</td>
</tr>
<tr>
<td>INTC</td>
<td>17.79</td>
<td>0.43</td>
<td>19.55</td>
<td>16.26</td>
<td>0.21</td>
</tr>
<tr>
<td>JNJ</td>
<td>16.06</td>
<td>0.42</td>
<td>18.40</td>
<td>14.06</td>
<td>0.33</td>
</tr>
<tr>
<td>JPM</td>
<td>16.81</td>
<td>0.75</td>
<td>19.20</td>
<td>14.11</td>
<td>0.30</td>
</tr>
<tr>
<td>KO</td>
<td>15.97</td>
<td>0.43</td>
<td>18.41</td>
<td>13.89</td>
<td>0.16</td>
</tr>
<tr>
<td>MCD</td>
<td>15.63</td>
<td>0.44</td>
<td>18.28</td>
<td>13.06</td>
<td>0.47</td>
</tr>
<tr>
<td>MMM</td>
<td>14.97</td>
<td>0.43</td>
<td>17.01</td>
<td>13.39</td>
<td>0.52</td>
</tr>
<tr>
<td>MKR</td>
<td>16.25</td>
<td>0.51</td>
<td>18.79</td>
<td>14.45</td>
<td>0.38</td>
</tr>
<tr>
<td>MSFT</td>
<td>17.83</td>
<td>0.42</td>
<td>20.30</td>
<td>16.22</td>
<td>0.11</td>
</tr>
<tr>
<td>NKE</td>
<td>14.67</td>
<td>0.57</td>
<td>18.69</td>
<td>12.71</td>
<td>0.34</td>
</tr>
<tr>
<td>PFE</td>
<td>17.31</td>
<td>0.30</td>
<td>19.49</td>
<td>15.18</td>
<td>0.26</td>
</tr>
<tr>
<td>PG</td>
<td>15.83</td>
<td>0.35</td>
<td>18.09</td>
<td>13.15</td>
<td>0.26</td>
</tr>
<tr>
<td>TRV</td>
<td>14.97</td>
<td>0.53</td>
<td>17.14</td>
<td>13.15</td>
<td>0.12</td>
</tr>
<tr>
<td>UNH</td>
<td>15.56</td>
<td>0.63</td>
<td>17.76</td>
<td>13.38</td>
<td>0.01</td>
</tr>
<tr>
<td>UXJ</td>
<td>15.10</td>
<td>0.45</td>
<td>18.12</td>
<td>12.92</td>
<td>0.03</td>
</tr>
<tr>
<td>VZ</td>
<td>16.31</td>
<td>0.30</td>
<td>20.24</td>
<td>14.17</td>
<td>0.37</td>
</tr>
<tr>
<td>V</td>
<td>14.39</td>
<td>1.47</td>
<td>18.25</td>
<td>10.52</td>
<td>0.45</td>
</tr>
<tr>
<td>WMT</td>
<td>16.23</td>
<td>0.52</td>
<td>18.39</td>
<td>14.56</td>
<td>0.23</td>
</tr>
<tr>
<td>XOM</td>
<td>16.69</td>
<td>0.47</td>
<td>18.59</td>
<td>14.21</td>
<td>0.34</td>
</tr>
</tbody>
</table>
Table 6: Main results from the implementation of the FGLS multivariate test statistic on log-volume and log-absolute returns. By columns, $\hat{d}_m$ and $\hat{d}_m$ denote the values that minimize the test statistic; 95% CI denotes the bounds of the 95% confidence ellipsoid for the long-memory parameter. When non-empty, Common $d$ reports the 95% confidence interval of the common order of integration. If empty, no common value is detected. The column $\hat{b}_e$ reports the sample cross-correlation coefficient of the residuals from the FGLS auxiliary regression given the values $\hat{d}_m$ ($v_{lm}$) and $\hat{d}_m$ ($v$).

<table>
<thead>
<tr>
<th>Stock</th>
<th>$\hat{d}_m$</th>
<th>$\hat{d}_m$</th>
<th>95% CI</th>
<th>Common $d$</th>
<th>$\hat{b}_e$</th>
<th>$\hat{d}_m$</th>
<th>$\hat{d}_m$</th>
<th>95% CI</th>
<th>Common $d$</th>
<th>$\hat{b}_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPL</td>
<td>0.38</td>
<td>0.38</td>
<td>[0.32, 0.44]</td>
<td>0.35</td>
<td>[0.28, 0.42]</td>
<td>0.52</td>
<td>0.52</td>
<td>[0.45, 0.59]</td>
<td>0.50</td>
<td>[0.43, 0.57]</td>
</tr>
<tr>
<td>AMH</td>
<td>0.32</td>
<td>0.32</td>
<td>[0.25, 0.39]</td>
<td>0.30</td>
<td>[0.23, 0.36]</td>
<td>0.48</td>
<td>0.48</td>
<td>[0.41, 0.55]</td>
<td>0.46</td>
<td>[0.40, 0.52]</td>
</tr>
<tr>
<td>BRK-A</td>
<td>0.52</td>
<td>0.52</td>
<td>[0.44, 0.60]</td>
<td>0.50</td>
<td>[0.43, 0.57]</td>
<td>0.67</td>
<td>0.67</td>
<td>[0.60, 0.74]</td>
<td>0.65</td>
<td>[0.59, 0.72]</td>
</tr>
<tr>
<td>C</td>
<td>0.62</td>
<td>0.62</td>
<td>[0.55, 0.69]</td>
<td>0.60</td>
<td>[0.53, 0.67]</td>
<td>0.75</td>
<td>0.75</td>
<td>[0.68, 0.82]</td>
<td>0.73</td>
<td>[0.66, 0.79]</td>
</tr>
<tr>
<td>GS</td>
<td>0.47</td>
<td>0.47</td>
<td>[0.40, 0.54]</td>
<td>0.45</td>
<td>[0.38, 0.52]</td>
<td>0.63</td>
<td>0.63</td>
<td>[0.56, 0.70]</td>
<td>0.61</td>
<td>[0.54, 0.68]</td>
</tr>
<tr>
<td>HON</td>
<td>0.32</td>
<td>0.32</td>
<td>[0.25, 0.39]</td>
<td>0.30</td>
<td>[0.23, 0.36]</td>
<td>0.48</td>
<td>0.48</td>
<td>[0.41, 0.55]</td>
<td>0.46</td>
<td>[0.40, 0.52]</td>
</tr>
<tr>
<td>IBM</td>
<td>0.52</td>
<td>0.52</td>
<td>[0.44, 0.60]</td>
<td>0.50</td>
<td>[0.43, 0.57]</td>
<td>0.67</td>
<td>0.67</td>
<td>[0.60, 0.74]</td>
<td>0.65</td>
<td>[0.59, 0.72]</td>
</tr>
<tr>
<td>INTC</td>
<td>0.18</td>
<td>0.18</td>
<td>[0.10, 0.26]</td>
<td>0.16</td>
<td>[0.08, 0.23]</td>
<td>0.35</td>
<td>0.35</td>
<td>[0.28, 0.43]</td>
<td>0.33</td>
<td>[0.26, 0.40]</td>
</tr>
<tr>
<td>INJ</td>
<td>0.14</td>
<td>0.14</td>
<td>[0.06, 0.22]</td>
<td>0.12</td>
<td>[0.04, 0.19]</td>
<td>0.28</td>
<td>0.28</td>
<td>[0.21, 0.35]</td>
<td>0.26</td>
<td>[0.19, 0.33]</td>
</tr>
<tr>
<td>JPM</td>
<td>0.36</td>
<td>0.36</td>
<td>[0.28, 0.44]</td>
<td>0.34</td>
<td>[0.26, 0.41]</td>
<td>0.53</td>
<td>0.53</td>
<td>[0.45, 0.61]</td>
<td>0.51</td>
<td>[0.44, 0.58]</td>
</tr>
<tr>
<td>MCD</td>
<td>0.18</td>
<td>0.18</td>
<td>[0.10, 0.26]</td>
<td>0.16</td>
<td>[0.08, 0.23]</td>
<td>0.35</td>
<td>0.35</td>
<td>[0.28, 0.43]</td>
<td>0.33</td>
<td>[0.26, 0.40]</td>
</tr>
<tr>
<td>MMM</td>
<td>0.30</td>
<td>0.30</td>
<td>[0.22, 0.38]</td>
<td>0.28</td>
<td>[0.20, 0.34]</td>
<td>0.49</td>
<td>0.49</td>
<td>[0.41, 0.56]</td>
<td>0.47</td>
<td>[0.40, 0.53]</td>
</tr>
<tr>
<td>MRK</td>
<td>0.35</td>
<td>0.35</td>
<td>[0.28, 0.42]</td>
<td>0.33</td>
<td>[0.25, 0.39]</td>
<td>0.53</td>
<td>0.53</td>
<td>[0.45, 0.61]</td>
<td>0.51</td>
<td>[0.44, 0.58]</td>
</tr>
<tr>
<td>NKE</td>
<td>0.10</td>
<td>0.10</td>
<td>[0.02, 0.18]</td>
<td>0.08</td>
<td>[0.00, 0.15]</td>
<td>0.25</td>
<td>0.25</td>
<td>[0.17, 0.33]</td>
<td>0.23</td>
<td>[0.15, 0.30]</td>
</tr>
<tr>
<td>PFE</td>
<td>0.32</td>
<td>0.32</td>
<td>[0.25, 0.39]</td>
<td>0.30</td>
<td>[0.22, 0.36]</td>
<td>0.48</td>
<td>0.48</td>
<td>[0.41, 0.55]</td>
<td>0.46</td>
<td>[0.39, 0.52]</td>
</tr>
<tr>
<td>PC</td>
<td>0.38</td>
<td>0.38</td>
<td>[0.30, 0.45]</td>
<td>0.37</td>
<td>[0.29, 0.43]</td>
<td>0.53</td>
<td>0.53</td>
<td>[0.46, 0.60]</td>
<td>0.51</td>
<td>[0.44, 0.58]</td>
</tr>
<tr>
<td>TEF</td>
<td>0.40</td>
<td>0.40</td>
<td>[0.32, 0.47]</td>
<td>0.38</td>
<td>[0.30, 0.44]</td>
<td>0.55</td>
<td>0.55</td>
<td>[0.48, 0.62]</td>
<td>0.53</td>
<td>[0.46, 0.59]</td>
</tr>
<tr>
<td>TXN</td>
<td>0.43</td>
<td>0.43</td>
<td>[0.35, 0.49]</td>
<td>0.41</td>
<td>[0.33, 0.46]</td>
<td>0.58</td>
<td>0.58</td>
<td>[0.51, 0.65]</td>
<td>0.56</td>
<td>[0.49, 0.63]</td>
</tr>
<tr>
<td>VZ</td>
<td>0.73</td>
<td>0.73</td>
<td>[0.65, 0.80]</td>
<td>0.71</td>
<td>[0.63, 0.78]</td>
<td>0.82</td>
<td>0.82</td>
<td>[0.75, 0.89]</td>
<td>0.80</td>
<td>[0.73, 0.88]</td>
</tr>
<tr>
<td>WMT</td>
<td>0.40</td>
<td>0.40</td>
<td>[0.32, 0.47]</td>
<td>0.38</td>
<td>[0.30, 0.44]</td>
<td>0.55</td>
<td>0.55</td>
<td>[0.48, 0.62]</td>
<td>0.53</td>
<td>[0.46, 0.59]</td>
</tr>
<tr>
<td>XOM</td>
<td>0.38</td>
<td>0.38</td>
<td>[0.30, 0.45]</td>
<td>0.36</td>
<td>[0.28, 0.42]</td>
<td>0.51</td>
<td>0.51</td>
<td>[0.44, 0.58]</td>
<td>0.49</td>
<td>[0.42, 0.56]</td>
</tr>
<tr>
<td>Average</td>
<td>0.41</td>
<td>0.41</td>
<td>[0.33, 0.49]</td>
<td>0.39</td>
<td>[0.31, 0.44]</td>
<td>0.50</td>
<td>0.50</td>
<td>[0.43, 0.57]</td>
<td>0.48</td>
<td>[0.41, 0.54]</td>
</tr>
</tbody>
</table>

Rejection rate: 20.69%
Table 7. Robustness checks in the joint analysis on log-volume and log-realized volatility against the choice of p and the inclusion of a time trend in volatility. Auxiliary regressions are augmented with \( p \) lags, with either \( p=2 \) or \( p \) determined according to Schwert’s rule.

<table>
<thead>
<tr>
<th>Stock</th>
<th>( p = 2 ), no linear trend</th>
<th>( p = 2 ), linear trend</th>
<th>Schwert’s rule, linear trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>95% CI VLM</td>
<td>95% CI RV</td>
<td>Common ( d )</td>
</tr>
<tr>
<td>AAPL</td>
<td>0.49, 0.57</td>
<td>0.45, 0.53</td>
<td>0.49, 0.53</td>
</tr>
<tr>
<td>AXP</td>
<td>0.38, 0.47</td>
<td>0.44, 0.52</td>
<td>0.45, 0.47</td>
</tr>
<tr>
<td>BA</td>
<td>0.27, 0.40</td>
<td>0.40, 0.49</td>
<td>- 0.52</td>
</tr>
<tr>
<td>CAT</td>
<td>0.34, 0.46</td>
<td>0.41, 0.48</td>
<td>0.42, 0.46</td>
</tr>
<tr>
<td>CSCO</td>
<td>0.28, 0.40</td>
<td>0.38, 0.46</td>
<td>- 0.53</td>
</tr>
<tr>
<td>CVX</td>
<td>0.38, 0.48</td>
<td>0.43, 0.51</td>
<td>0.43, 0.48</td>
</tr>
<tr>
<td>DIS</td>
<td>0.28, 0.39</td>
<td>0.40, 0.49</td>
<td>- 0.42</td>
</tr>
<tr>
<td>GE</td>
<td>0.38, 0.47</td>
<td>0.43, 0.52</td>
<td>0.45, 0.46</td>
</tr>
<tr>
<td>GS</td>
<td>0.41, 0.51</td>
<td>0.43, 0.50</td>
<td>0.43, 0.50</td>
</tr>
<tr>
<td>HD</td>
<td>0.36, 0.47</td>
<td>0.42, 0.49</td>
<td>0.43, 0.47</td>
</tr>
<tr>
<td>IBM</td>
<td>0.31, 0.41</td>
<td>0.39, 0.47</td>
<td>- 0.50</td>
</tr>
<tr>
<td>INTC</td>
<td>0.23, 0.38</td>
<td>0.38, 0.46</td>
<td>- 0.50</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.31, 0.42</td>
<td>0.41, 0.50</td>
<td>- 0.44</td>
</tr>
<tr>
<td>JPM</td>
<td>0.41, 0.49</td>
<td>0.44, 0.51</td>
<td>0.44, 0.49</td>
</tr>
<tr>
<td>KO</td>
<td>0.32, 0.44</td>
<td>0.40, 0.48</td>
<td>0.41, 0.44</td>
</tr>
<tr>
<td>MCD</td>
<td>0.28, 0.41</td>
<td>0.38, 0.46</td>
<td>- 0.45</td>
</tr>
<tr>
<td>MMM</td>
<td>0.26, 0.40</td>
<td>0.37, 0.45</td>
<td>- 0.46</td>
</tr>
<tr>
<td>MRK</td>
<td>0.31, 0.44</td>
<td>0.38, 0.46</td>
<td>0.38, 0.44</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.22, 0.39</td>
<td>0.39, 0.47</td>
<td>- 0.50</td>
</tr>
<tr>
<td>NKE</td>
<td>0.28, 0.39</td>
<td>0.38, 0.46</td>
<td>- 0.50</td>
</tr>
<tr>
<td>PFE</td>
<td>0.30, 0.43</td>
<td>0.39, 0.48</td>
<td>0.42, 0.43</td>
</tr>
<tr>
<td>PG</td>
<td>0.28, 0.42</td>
<td>0.38, 0.47</td>
<td>0.40, 0.41</td>
</tr>
<tr>
<td>TRV</td>
<td>0.35, 0.49</td>
<td>0.45, 0.56</td>
<td>- 0.43</td>
</tr>
<tr>
<td>UNH</td>
<td>0.28, 0.42</td>
<td>0.40, 0.48</td>
<td>- 0.52</td>
</tr>
<tr>
<td>UTX</td>
<td>0.31, 0.42</td>
<td>0.41, 0.49</td>
<td>- 0.44</td>
</tr>
<tr>
<td>V</td>
<td>0.51, 0.66</td>
<td>0.43, 0.51</td>
<td>- 0.50</td>
</tr>
<tr>
<td>VZ</td>
<td>0.32, 0.46</td>
<td>0.44, 0.52</td>
<td>- 0.41</td>
</tr>
<tr>
<td>WMT</td>
<td>0.28, 0.40</td>
<td>0.36, 0.44</td>
<td>- 0.51</td>
</tr>
<tr>
<td>XOM</td>
<td>0.33, 0.45</td>
<td>0.38, 0.48</td>
<td>0.39, 0.45</td>
</tr>
</tbody>
</table>

Average: 0.48 | 44.83% | 0.48 | 55.17%