Essex Finance Centre
Working Paper Series

Working Paper No 65: 01-2021

"Multivariate Fractional Integration Tests allowing for Conditional Heteroskedasticity with an Application to Return Volatility and Trading Volume"

"Marina Balboa, Paulo M.M. Rodrigues,
Antonio Rubia, A.M. Robert Taylor"

Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ
Web site: http://www.essex.ac.uk/ebs/

Multivariate Fractional Integration Tests allowing for Conditional Heteroskedasticity with an Application to Return Volatility and Trading Volume*

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

January 18, 2021

Abstract

We introduce a new joint test for the order of fractional integration of a multivariate fractionally integrated vector autoregressive [FIVAR] time series based on applying the Lagrange multiplier principle to a feasible generalised least squares estimate of the FIVAR model obtained under the null hypothesis. A key feature of the test we propose is that it is constructed using a heteroskedasticity-robust estimate of the variance matrix. As a result, the test has a standard χ^2 limiting null distribution under considerably weaker conditions on the innovations than are permitted in the extant literature. Specifically, we allow the innovations driving the FIVAR model to follow a vector martingale difference sequence allowing for both serial and cross-sectional dependence in the conditional second-order moments. We also do not constrain the order of fractional integration of each element of the series to lie in a particular region, thereby allowing for both stationary and non-stationary dynamics, nor do we assume any particular distribution for the innovations. A Monte Carlo study demonstrates that our proposed tests avoid the large over-sizing problems seen with extant tests when conditional heteroskedasticity is present in the data. We report an empirical case study for a sample of major U.S. stocks investigating the order of fractional integration in trading volume and different measures of volatility in returns, including realized variance. Our results suggest that both return volatility and trading volume are fractionally integrated, but with the former generally found to be more persistent (having a higher fractional exponent) than the latter, when more reliable proxies for volatility such as the range or realized variance are used.

Keywords: Fractional integration, conditional heteroskedasticity, Lagrange multiplier test, return volatility, trading volume.
JEL classifications: C20, C22.

*Correspondence to: Robert Taylor, Essex Business School, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Email: `robert.taylor@essex.ac.uk`

1 Introduction

Long memory models have been used to model time series data in a wide range of fields of application. The class of (multivariate) fractionally integrated autoregressive moving average [ARFIMA] models provides a parsimonious means of simultaneously modelling the patterns of long and short range dependence typically seen in many macroeconomic and financial data sets; see, for example, the surveys in Baillie (1996) and Robinson (2003). In the context of the ARFIMA class of models the long memory parameter, or fractional exponent (vector of exponents in the multivariate case), is the key parameter driving the behaviour of the series. Where this is zero a weakly dependent (short memory) ARMA series obtains. If it is less than one-half the series is weakly stationary, otherwise it is non-stationary, the familiar autoregressive unit root case occurring where the exponent is unity. Consequently, considerable interest has been paid to developing methods of inference on the fractional exponent both as a parameter of interest in its own right and for preliminary data analysis. A leading example is a test of the null hypothesis of weak dependence (against fractional alternatives); here a non-rejection would allow for the use of standard methods for conducting, among other things, causality, structural vector autoregression, or impulse response analyses. More generally, such tests could usefully be employed to indicate what order of differencing of the data is required for such methods to be suitably employed.

In the univariate setting a number of hypothesis tests on the fractional exponent have been proposed; see, among others, Robinson (1994), Tanaka (1999), Breitung and Hassler (2002), Nielsen (2004b), Demetrescu et al. (2008), Hassler et al. (2009, 2016) and Cavaliere et al. (2017). In the context of a vector series one could perform such univariate fractional integration tests separately on each element of the vector. However, the overall size of such a testing procedure would be hard to control. Moreover, multivariate testing can improve efficiency relative to univariate testing because it explicitly acknowledges and exploits the existence of any endogenous cross-dependencies in the vector series which can reduce the variability in the estimation errors and, hence, improve efficiency in estimation and testing. In this paper we develop multivariate fractional integration tests designed to test joint null hypotheses concerning the values of the long memory parameters of the elements of a fractionally integrated vector autoregressive [FIVAR] model. Specifically, we propose a parametric multivariate Lagrange multiplier [LM]-type test in the time-domain which generalises the univariate regression-based LM-type test of Demetrescu et al. (2008) to the multivariate case. The method we propose can also be used to construct confidence sets, at a given asymptotic coverage level, for the true values of the long memory coefficients.

Our testing procedure is implemented in a regression-based context, based on feasible generalised least squares [FGLS] estimation of the multivariate FIVAR model under the null hypothesis, coupled with a heteroskedasticity-robust variance matrix estimate. A key advantage of, and motivation for, this approach is that it allows us to significantly weaken the technical conditions needed on the innovations, relative to existing multivariate fractional tests including, among others, Robinson (1995), Lobato and Robinson (1998), Lobato (1999), Lobato and Velasco (2000), Marinucci and Robinson (2001), Breitung and Hassler (2002), Shimotsu (2007), and (Nielsen, 2004b, 2005, 2011). In particular, we allow the driving innovations in the data generating process [DGP] to follow a vector martingale difference sequence [MDS] which is permitted to exhibit time-varying conditional heteroskedasticity. This therefore allows for both serial and cross-sectional dependence in the conditional second-order moments, which is of particular empirical relevance when modelling financial data and is not, to the best of our knowledge, allowed by any extant multivariate fractional integration test.

Like Nielsen (2004a, 2005), we work within the context of a multivariate FIVAR model. This model allows each series within the vector process to have different fractional exponents irrespective of the parameters of the short-run component of the model. This property is not guaranteed when using the class of vector autoregressive fractionally integrated [VARFI] models where the orders of integration of the elements of the vector series are not constant throughout the parameter space of the model; for further details see Nielsen (2005, pp.381-382). This is important for the empirical case study we consider in this paper with respect to trading volume and return volatility where we aim to explicitly investigate whether the data support the hypothesis that these series admit a common fractional exponent or not. For a further recent empirical application using FIVAR models, investigating the effects of monetary policy on the economy, where it is important to allow the elements of the vector time series to have potentially different fractional exponents, see Lovcha and Perez-Laborda (2018). An implication of the FIVAR model, however, is that fractional cointegration is not possible between the elements of the vector time series. In common with the tests in Nielsen (2005) we do not restrict the fractional exponents to lie within a particular region, thereby allowing for both stationary and non-stationary dynamics. We also do not impose any particular distributional law on the innovations.

Under the conditionally heteroskedastic setting outlined above, our proposed test retains a standard χ^2 limiting null distribution (irrespective of the null values of the long memory parameters being tested) and exhibits non-trivial power against a sequence of local (Pitman drift) alternatives. Moreover, where the errors are independent and identically distributed [i.i.d.] our test statistic is asymptotically equivalent to the multivariate LM statistic discussed in Nielsen (2004a, 2005). As a consequence, the LM-type test we propose is asymptotically (locally) efficient when the errors are i.i.d. Gaussian. Monte Carlo simulation experiments show that our proposed multivariate fractional integration test displays good finite sample size control and power performance in the presence of empirically relevant data features, such as short-run dependence and time-varying GARCH-type conditional variances for both Gaussian and non-Gaussian innovations. In contrast, extant tests are shown to display quite poor finite sample size control in the presence of such features.

Multivariate testing is naturally intended to address joint hypotheses involving the degree of persistence of a set of variables. This has important practical applications. As a leading example, there has been considerable interest in both the theoretical and empirical finance literatures on understanding the link between trading volume and return volatility. A number of papers have analysed if the long-run dynamics of these variables share a common order of fractional integration, with mixed evidence; see, for example, Bollerslev and Jubinski (1999), Lobato and Velasco (2000), Luu and Martens (2003), Fleming and Kirby (2011) and Rossi and de Magistris (2013).

In our empirical analysis we apply our new approach to conduct joint inference on the order of fractional integration of trading volumes and different measures of return volatility focusing on 30 major U.S. stocks from the Dow-Jones Industrial Average Index [DJI]. We also investigate the existence of a common order of fractional integration between volume and these measures of return volatility. Because our tests do not require a particular distribution and, more importantly, allow for time-varying conditional second-order moments, our results are likely to be more robust than those reported in previous studies which are based on estimation techniques which neglect these empirically relevant data features (e.g. Bollerslev and Jubinski (1999), Lobato and Velasco (2000) and Fleming and Kirby (2011)).

Together with daily log-volume, we consider the log-transformations of three alternative measures of return volatility with increasing degrees of efficiency, namely: absolute-valued returns, the range-based estimator of Garman and Klass (1980), and a measure of realised variance constructed from 5-minute returns. An important aspect of this analysis is to investigate the influence that measurement errors have on the conclusions drawn from the data. Our empirical findings suggest that a common fractional exponent cannot in general be rejected when return volatility is proxied by absolute-valued returns, but can be rejected when it is proxied by more accurate estimates such as the range or realised variance. These findings are consistent with the previous literature and help us to understand the disparity between empirical results where different proxies for volatility are used. Our empirical results indicate that return volatility tends to exhibit a larger fractional integration exponent than trading volume, with long-term behaviour possibly driven by non-stationary dynamics. Heterogeneous degrees of fractional integration, such that return volatility tends to be more persistent than trading volume, could originate in certain types of trading strategies associated with imitation and herding in investors and market microstructure environmental conditions; see, e.g., LeBaron and Yamamoto (2008) and Yamamoto (2011).

The remainder of the paper is organised as follows. Section 2 introduces the DGP and the main assumptions underlying our theoretical analysis. In section 3 we detail our new LM-type multivariate fractional integration test and derive its asymptotic distribution under both the null hypothesis and a sequence of local alternatives. Section 4 discusses the results of our finite sample Monte Carlo study. Section 5 analyses the empirical relationship between trading volume and return volatility for stocks from the DJI. Section 6 concludes. An on-line supplementary appendix contains mathematical proofs of the large sample results given in section 3 together with additional material relating to the Monte Carlo analysis in section 4 and to the empirical application in section 5.

In what follows, $\Rightarrow$ and $\xrightarrow{p}$ denote weak convergence and convergence in probability, respectively, as $T \to \infty$. $I(\cdot)$ is an indicator function that equals one if the condition in parenthesis is fulfilled, and equals zero otherwise. The operators $\otimes$ and $\odot$ correspond to the Kronecker and Hadamard products, respectively. The quantities $\boldsymbol{I}_n$ and $\mathbf{0}_{n\times m}$ denote an n-dimensional identity matrix and an $n \times m$ zero matrix, respectively. The notation $\boldsymbol{A} = \{a_{ij}\}$ denotes that the (i, j)th element of the matrix $\boldsymbol{A}$ is given by a_{ij}.

2 A FIVAR Model with Heteroskedasticity

We consider the observable k-dimensional time series vector $\{\boldsymbol{y}_t\}_{t=1}^T$, where $\boldsymbol{y}_t \equiv (y_{1,t}, ..., y_{k,t})'$, is generated according to the DGP:

$$\boldsymbol{\Delta}^{\boldsymbol{d}+\boldsymbol{\theta}}(L)\, \boldsymbol{y}_t = \boldsymbol{\varepsilon}_t I(t \geq 1) \tag{1}$$

where $\boldsymbol{\Delta}^{\boldsymbol{d}+\boldsymbol{\theta}}(L)$ is a $k \times k$ diagonal matrix polynomial in the conventional lag operator, L, with characteristic element given by $(1 - L)^{d_i+\theta_i}$, $i \in \{1, ..., k\}$. The real-valued fractional exponent, $d_i + \theta_i$, is commonly referred to as the long memory or fractional integration parameter, such that $\boldsymbol{d} + \boldsymbol{\theta} \equiv (d_1 + \theta_1, ..., d_k + \theta_k)'$. The k-dimensional vector $\boldsymbol{\varepsilon}_t \equiv (\varepsilon_{1,t}, ..., \varepsilon_{k,t})'$ is a weakly-dependent (short memory or $I(0)$) noise process with bounded spectral density that is bounded away from zero at the origin. Our focus is on developing tests of the null hypothesis that $\boldsymbol{d}$ is the true order of integration of $\{\boldsymbol{y}_t\}$; that is, $\text{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, against the alternative hypothesis that at least one element of $\boldsymbol{\theta}$ is non-zero.

Assumption 1 details the formal properties which we will assume to hold on $\{\boldsymbol{\varepsilon}_t\}$ in (1).

Assumption 1. *$\{\boldsymbol{\varepsilon}_t\}$ in (1) is generated as $\mathbf{\Pi}(L)\boldsymbol{\varepsilon}_t = \boldsymbol{e}_t \equiv (e_{1,t}, ..., e_{k,t})'$, with $\mathbf{\Pi}(L) := \boldsymbol{I}_p - \sum_{j=1}^{p} \mathbf{\Pi}_j L^j$, where $\mathbf{\Pi}_j$ are $k \times k$ parameter matrices such that $\mathbf{\Pi}(L)$ has all of its roots lying outside the unit circle and $\{\boldsymbol{e}_t\}$ satisfies the following conditions:*
(A1) *$E(\boldsymbol{e}_t) = 0$ and $E(\boldsymbol{e}_t\boldsymbol{e}_t') =: \mathbf{\Sigma}$, with $\mathbf{\Sigma}$ positive definite.*
(A2) *$sup_t E(||\boldsymbol{e}_t||^{4+\eta}) < \infty$ for some $\eta > 0$.*
(A3) *$\{\boldsymbol{e}_t, \mathcal{F}_t\}_{t=-\infty}^{\infty}$ is a strictly stationary and ergodic vector MDS, with respect to the natural filtration $\mathcal{F}_t$, the σ-field generated by $\{\boldsymbol{e}_s : s \leq t\}$.*
(A4) *$\sum_{i=1}^{\infty} \sum_{j=1, i \neq j}^{\infty} E|e_{h,t}e_{s,t}e_{r,t-i}e_{u,t-j}| < \infty$, for any $1 \leq h, s, r, u \leq k$.*

Remark 1. Assumption 1 allows the short memory component of $\{\boldsymbol{y}_t\}$ to be driven by a stationary VAR(p) process. Accordingly, (1) is a FIVAR model in which each component $\{y_{i,t}\}$, $i = 1, ..., k$, follows a Type-II ARFIMA$(p, d_i + \theta_i, 0)$ process. The choice of Type-II fractional integration in our setting has the desirable feature that the same definition is valid for an arbitrarily large range of admissible values of the fractional parameters, $d_i + \theta_i$, $i = 1, ..., k$; in particular, these are not restricted to lie in the interval $(-0.5, 0.5)$, a necessary condition for stationarity and invertibility. ◇

Remark 2. (A1) and (A2) are standard moment conditions. Unlike the existing multivariate fractional integration tests discussed in Section 1, (A3) allows the innovations to exhibit time-varying conditional variances. The absolute summability condition (A4) limits the amount of temporal and cross-sectional dependence in the second-order moments, and is equivalent to requiring absolutely summable 4th-order joint cumulants. Our conditions are weaker than requiring $\{\boldsymbol{e}_t, \mathcal{F}_t\}$ to be either conditionally homoskedastic or independent, both of which imply (A4). Finally, (A1) and A(3) imply that $E(e_{i,t}e_{j,t+h}) = 0$ whenever $h \neq 0$, but allows $E(e_{i,t}e_{j,t}) \neq 0$ when $i \neq j$ because $\mathbf{\Sigma}$ is not restricted to be diagonal. ◇

Remark 3. Assumption 1 imposes that the unconditional variance matrix of $\boldsymbol{e}_t$, $\mathbf{\Sigma}$, is constant. However, the pivotal χ^2 limiting null distribution of our proposed FGLS statistic,

$LM_{\boldsymbol{d}}^{FGLS}$, defined below in (8), in Theorem 2 under the conditions of Assumption 1 remains valid in so-called non-stationary volatility cases where $E(\boldsymbol{e}_t\boldsymbol{e}_t') = \boldsymbol{\Sigma}_t = \boldsymbol{\sigma}_t\boldsymbol{\sigma}_t'$, provided the unconditional volatility matrix, $\boldsymbol{\sigma}_t$, satisfies the regularity conditions detailed in, e.g., Assumption 2(a) of Boswijk et al. (2016).[1] In particular, these entail that $\boldsymbol{\sigma}_t := \boldsymbol{\sigma}(t/T)$, for all $t = 1, ..., T$, where $\boldsymbol{\sigma}(\cdot)$ is a non-stochastic element of the space of $k \times k$ matrices of càdlàg functions on $[0, 1]$ equipped with the Skorokhod metric, and is such that $\boldsymbol{\Sigma}(u) := \boldsymbol{\sigma}(u)\boldsymbol{\sigma}(u)'$ is positive definite for all $u \in [0, 1]$. For further discussion, including a number of examples satisfying these conditions, see Boswijk et al. (2016, p.66). ♢

Remark 4. Under Assumption 1, the model in (1) can be re-written as $\boldsymbol{\Pi}(L)\boldsymbol{\Delta}^{\boldsymbol{d}+\boldsymbol{\theta}}(L)\boldsymbol{y}_t = \boldsymbol{e}_t$. Given the stationarity restriction imposed under Assumption 1, for a sufficiently large value of p the FIVAR representation could be viewed as an approximation to the more general class of FIVARMA models, although we treat p as fixed (independent of the sample size) in this paper. We conjecture that it should be possible to extend our analysis to allow p to increase with the sample size but this would considerably complicate the theoretical analysis and is beyond the scope of this paper. ♢

Remark 5. The FIVAR model in (1) under Assumption 1 rules out the possibility of fractional cointegration between the elements of $\{\boldsymbol{y}_t\}$; for further discussion see, among others, Sela and Hurvich (2009) and Nielsen (2005, pp.381-382). The maintained assumption of no fractional cointegration is also made in all of the extant multivariate fractional integration tests cited in the Introduction. However, noting from Remark 10 below that the feasible GLS multivariate fractional integration test we propose in section 3.2 is asymptotically equivalent to the multivariate LM fractional test in Nielsen (2005), then for the same reasons as are discussed in Nielsen (2005, pp.378-379), the LM-type test, $LM_{\boldsymbol{d}}^{FGLS}$, developed in section 3 is also implicitly a test of the null of no fractional cointegration (in the sense defined in Nielsen (2005, p.378)) and will diverge at rate $O_p(T)$ under fractional cointegration.[2] It therefore seems advisable to consider the tests proposed in this paper alongside tests for fractional cointegration. We adopt this approach in the empirical application in section 5 by also considering the procedures developed in Nielsen and Shimotsu (2007). ♢

[1]Numerical experiments investigating the properties of the $LM_{\boldsymbol{d}}^{FGLS}$ test for data with a one-time break in unconditional variance are reported in the supplementary appendix. These results suggest that even quite large variance breaks have very little impact on the finite sample size of the FGLS-based tests.

[2]Numerical experiments investigating rejection rates of the $LM_{\boldsymbol{d}}^{FGLS}$ test and the tests of Nielsen (2005) and Breitung and Hassler (2002) in a fractionally cointegrated model are reported in the supplementary appendix. These show that, as expected, all three tests display empirical rejection frequencies in excess of the nominal level which are larger, other things equal, the larger is T or the strength of cointegration. Of the three tests, our FGLS test tends to reject with slightly lower frequency than the other two tests.

3 A Multivariate LM-type Fractional Integration Test

3.1 Preliminaries

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

$$\boldsymbol{\varepsilon}_{t,\boldsymbol{g}} \equiv (\varepsilon_{1,t,g_1}, ..., \varepsilon_{k,t,g_k})' := (1-L)_+^{\boldsymbol{g}}\, \boldsymbol{y}_t = \sum_{j=0}^{t-1} \boldsymbol{\Lambda}_j (\boldsymbol{g})\, \boldsymbol{y}_{t-j}, \tag{2}$$

where $(1-L)_+^{\boldsymbol{g}} := \sum_{j=0}^{t-1} \boldsymbol{\Lambda}_j (\boldsymbol{g})\, L^j$, and

$$\boldsymbol{z}_{t-1,\boldsymbol{g}}^{*} \equiv \left(z_{1,t-1,g_1}^{*}, ..., z_{k,t-1,g_k}^{*}\right)' := \sum_{j=1}^{t-1} j^{-1} \boldsymbol{\varepsilon}_{t-j,\boldsymbol{g}},\ \ t = 2, ..., T \tag{3}$$

with $\{\boldsymbol{\Lambda}_j (\boldsymbol{g})\}_{j=0}^{t-1}$ denoting a sequence of $k \times k$ diagonal matrices with ith diagonal element

$$\lambda_0 (g_i) := 1,\ \text{and}\ \lambda_j (g_i) := \frac{j-1-g_i}{j} \lambda_{j-1} (g_i)\ ,\ j \geq 1, \tag{4}$$

corresponding to the truncated series of polynomial coefficients in the binomial expansion $(1-L)^{g_s} := \sum_{j=0}^{\infty} \lambda_j (g_s)\, L^j$. These variables are straightforward generalisations of the corresponding univariate processes in Breitung and Hassler (2002) to the multivariate context, with the characteristic harmonic weighting in (3) arising from the derivative of a (Gaussian) score function. Remark 10 below gives further insight into the key role played by these variables in the construction of our proposed LM-type test statistic.

Let $\boldsymbol{\Phi}$ denote a $k \times k$ diagonal matrix with ith diagonal element ϕ_{ii}, $i = 1, ..., k$. Under Assumption 1, testing the null hypothesis that $\boldsymbol{d}$ is the true order of integration of $\{\boldsymbol{y}_t\}$, $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, is equivalent to testing $\mathrm{H}_0 : \boldsymbol{\Phi} = \mathbf{0}$ in the multivariate linear regression model

$$\boldsymbol{\varepsilon}_{t,\boldsymbol{d}} = \boldsymbol{\Phi} \boldsymbol{z}_{t-1,\boldsymbol{d}}^{*} + \sum_{j=1}^{p} \boldsymbol{\Pi}_j \boldsymbol{\varepsilon}_{t-j,\boldsymbol{d}} + \boldsymbol{v}_t, \quad t = p^* + 1, ..., T \tag{5}$$

where $p^* := max(1, p)$. This equivalence holds because, under $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, (5) and (1) are bijective with $\phi_{ii} = 0$ for all $i = 1, ..., k$ and $\boldsymbol{v}_t = \boldsymbol{e}_t$ in (5); see also Breitung and Hassler (2002), Demetrescu et al. (2008), and Hassler et al. (2009).

It will prove convenient to re-write (5) in matrix notation. First, corresponding to the time series of observations for each element of $\boldsymbol{y}_t$, we have the equivalent representation,

$$\boldsymbol{Y}_{i,d_i} = \boldsymbol{X}^*_{i,-1,\boldsymbol{d}}\boldsymbol{\beta}_i + \boldsymbol{u}_i,\ 1 \leq i \leq k \tag{6}$$

where $\boldsymbol{Y}_{i,d_i} := (\varepsilon_{i,p^*+1,d_i}, ..., \varepsilon_{i,T,d_i})'$ is a $(T-p^*) \times 1$ vector, $\boldsymbol{\beta}_i := (\phi_{ii}, \boldsymbol{\pi}_{i1}, ..., \boldsymbol{\pi}_{ip})'$ is a k'-dimensional parameter vector, with $k' := pk+1$, and $\boldsymbol{\pi}_{ij}$ denotes the i-th row of $\boldsymbol{\Pi}_j$, $j = 1, ..., p$, $\boldsymbol{u}_i := (v_{i,p^*+1}, ..., v_{i,T})'$ is a $(T-p^*) \times 1$ vector of innovations, and $\boldsymbol{X}^*_{i,-1,\boldsymbol{d}}$ is the $(T-p^*) \times k'$ matrix of observations of the (lagged) right-hand side variables $\boldsymbol{x}^*_{i,t-1,\boldsymbol{d}} := \left(z^*_{i,t-1,d_i}, \boldsymbol{\varepsilon}'_{t-1,\boldsymbol{d}}, ..., \boldsymbol{\varepsilon}'_{t-p,\boldsymbol{d}}\right)'$. With the exception of the first regressor, all other right-hand side variables that characterise the i-th equation (6) are the same, since these always correspond to lagged values of $\boldsymbol{\varepsilon}_{t,\boldsymbol{d}}$. Then, given $T' := k(T-p^*)$, we can write the system of equations (6) compactly as $\boldsymbol{Y}_{\boldsymbol{d}} = \boldsymbol{X}^*_{-1,\boldsymbol{d}}\ \boldsymbol{\beta} + \boldsymbol{u}$, with these terms defined implicitly as:

$$\begin{bmatrix} \boldsymbol{Y}_{1,d_1} \\ \boldsymbol{Y}_{2,d_2} \\ \vdots \\ \boldsymbol{Y}_{k,d_k} \end{bmatrix}_{T'\times 1} = \begin{bmatrix} \boldsymbol{X}^*_{1,-1,\boldsymbol{d}} & \mathbf{0}_{(T-p^*)\times k'} & \cdots & \mathbf{0}_{(T-p^*)\times k'} \\ \mathbf{0}_{(T-p^*)\times k'} & \boldsymbol{X}^*_{2,-1,\boldsymbol{d}} & \cdots & \mathbf{0}_{(T-p^*)\times k'} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{(T-p^*)\times k'} & \mathbf{0}_{(T-p^*)\times k'} & \cdots & \boldsymbol{X}^*_{k,-1,\boldsymbol{d}} \end{bmatrix}_{T'\times kk'} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_k \end{bmatrix}_{kk'\times 1} + \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_k \end{bmatrix}_{T'\times 1}.$$

3.2 A Heteroskedasticity-Robust LM Test

Under Assumption 1 and $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, it follows that $E\left(\boldsymbol{u}\boldsymbol{u}'\right) = \boldsymbol{\Sigma} \otimes \boldsymbol{I}_{T-p^*}$. Equation (5) defines a seemingly unrelated regression equation [SURE] system. Although equation-by-equation ordinary least squares [OLS] estimation will deliver consistent estimates of $\boldsymbol{\beta}$, these estimates will not be efficient unless $\boldsymbol{\Sigma}$ is diagonal (recalling that the regressors differ across the equations in the system). We will therefore consider a FGLS estimator of $\boldsymbol{\beta}$ based on a preliminary consistent estimate of $\boldsymbol{\Sigma}$ (obtained using OLS residuals estimated on an equation-by-equation basis). The resulting FGLS estimator of $\boldsymbol{\beta}$ is defined as:

$$\widehat{\boldsymbol{\beta}} := \left(\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^*_{-1,\boldsymbol{d}}\right)^{-1}\left(\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{Y}_{\boldsymbol{d}}\right) \tag{7}$$

where $\widetilde{\boldsymbol{\Sigma}} = \{\widetilde{\sigma}_{ij}\}$ is estimated as $\widetilde{\sigma}_{ij} := T^{-1}\widetilde{\boldsymbol{u}}'_i\widetilde{\boldsymbol{u}}_j$, with $\widetilde{\boldsymbol{u}}_s := \boldsymbol{Y}_{s,d_s} - \boldsymbol{X}^*_{s,-1,\boldsymbol{d}}\widetilde{\boldsymbol{\beta}}_s$, and $\widetilde{\boldsymbol{\beta}}_s$ denotes the equation-by-equation OLS estimate of $\boldsymbol{\beta}_s$, $s = 1, ..., k$ in (6).[3]

[3]Some numerical experiments comparing the finite sample properties of the equation-by-equation OLS estimate and the FGLS estimate in (7) of $\boldsymbol{\beta}$ are given in the supplementary appendix. These clearly demonstrate the efficiency gains that can be obtained by FGLS over OLS.

In Theorem 1 we now characterise the asymptotic distribution of the FGLS estimate, $\widehat{\boldsymbol{\beta}}$, under Assumption 1 and $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$.

Theorem 1. *Let $\boldsymbol{y}_t$ be generated according to* (1) *and let $\widehat{\boldsymbol{\beta}}$ be the vector of FGLS estimates defined in* (7). *Under Assumption 1 and $H_0 : \boldsymbol{\theta} = \mathbf{0}$, $\sqrt{T}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) \Rightarrow \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\beta}}\right)$ where $\boldsymbol{\beta}_0 \equiv (\boldsymbol{\beta}_{01}', ..., \boldsymbol{\beta}_{0k}')'$ with $\boldsymbol{\beta}_{0s} := (0, \boldsymbol{\pi}_{s1}, ..., \boldsymbol{\pi}_{sp})'$, $s = 1, ..., k$, and $\boldsymbol{\Omega}_{\boldsymbol{\beta}} := \boldsymbol{A}_{\boldsymbol{\beta}}^{-1} \boldsymbol{B}_{\boldsymbol{\beta}} \boldsymbol{A}_{\boldsymbol{\beta}}^{-1}$, with $\boldsymbol{A}_{\boldsymbol{\beta}} := \underset{T\to\infty}{plim} E\left(\frac{1}{T} \boldsymbol{X}_{-1,\boldsymbol{d}}'^{*} \left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right] \boldsymbol{X}_{-1,\boldsymbol{d}}^{*}\right)$, $\boldsymbol{B}_{\boldsymbol{\beta}} := \underset{T\to\infty}{plim} E\left(\frac{1}{T} \boldsymbol{w}_{-1,\boldsymbol{d}}^{*} \boldsymbol{w}_{-1,\boldsymbol{d}}'^{*}\right)$, and $\boldsymbol{w}_{-1,\boldsymbol{d}}^{*} := \boldsymbol{X}_{-1,\boldsymbol{d}}'^{*} \left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right] \boldsymbol{u}$.*

The dependence of the asymptotic variance of the FGLS estimator on nuisance parameters arising from any weak dependence and/or cross sectional correlation in $\boldsymbol{\varepsilon}_t$ implies that asymptotically pivotal inference on the long memory parameters will need to be based on a heteroskedasticity-robust statistic formed using a consistent estimate of $\boldsymbol{\Omega}_{\boldsymbol{\beta}}$. This can be achieved by using the familiar Eicker-Huber-White approach building on the preliminary OLS estimate $\widetilde{\boldsymbol{\Sigma}}$ and the FGLS residuals $\widehat{\boldsymbol{u}} := \boldsymbol{Y}_{\boldsymbol{d}} - \boldsymbol{X}_{-1,\boldsymbol{d}}^{*} \widehat{\boldsymbol{\beta}}$. In particular, a heteroskedasticity-robust estimate of the variance matrix $\boldsymbol{\Omega}_{\boldsymbol{\beta}}$ is given by $\widehat{\boldsymbol{\Omega}}_{\boldsymbol{\beta}} := \boldsymbol{A}_T^{*-1} \boldsymbol{B}_T^{*} \boldsymbol{A}_T^{*-1}$, where $\boldsymbol{A}_T^{*} := \boldsymbol{X}_{-1,\boldsymbol{d}}'^{*} \left[\widetilde{\boldsymbol{\Sigma}}^{-1/2} \otimes \boldsymbol{I}_{T-p^*}\right] \boldsymbol{X}_{-1,\boldsymbol{d}}^{*}/T$ and $\boldsymbol{B}_T^{*} := \widehat{\boldsymbol{w}}_{-1,\boldsymbol{d}}^{*} \widehat{\boldsymbol{w}}_{-1,\boldsymbol{d}}'^{*}/T$, with $\widehat{\boldsymbol{w}}_{-1,\boldsymbol{d}}^{*} := \boldsymbol{X}_{-1,\boldsymbol{d}}'^{*} \left[\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right] \widehat{\boldsymbol{u}}$. It is shown in the supplementary appendix that $\widehat{\boldsymbol{\Omega}}_{\boldsymbol{\beta}}$ is a consistent estimate of $\boldsymbol{\Omega}_{\boldsymbol{\beta}}$ under the conditions given in Assumption 1.

Based on the heteroskedasticity-robust estimate, $\widehat{\boldsymbol{\Omega}}_{\boldsymbol{\beta}}$, it is then straightforward to construct a test statistic for the joint hypothesis $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$ using the LM testing principle. Specifically, we can form a heteroskedasticity-robust LM-type test which rejects $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$ for large values of the statistic

$$LM_{\boldsymbol{d}}^{FGLS} := T \ \left[\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right]' \left[\boldsymbol{R}\, \widehat{\boldsymbol{\Omega}}_{\boldsymbol{\beta}}\, \boldsymbol{R}'\right]^{-1} \left[\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right] \tag{8}$$

where $\boldsymbol{R} = \{r_{ij}\}$ is a $k \times kk'$ indicator matrix taking a value equal to one when $j = (i-1)k' + 1$, $i = 1, ..., k$, and zero otherwise. In Theorem 2 we next derive the large sample behaviour of $LM_{\boldsymbol{d}}^{FGLS}$ under both the null hypothesis, $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, and under the sequence of local alternatives $\mathrm{H}_{\boldsymbol{c}} : \boldsymbol{\theta} = \boldsymbol{c}/\sqrt{T}$, where $\boldsymbol{c} \equiv (c_1, ..., c_k)'$ is a k-vector of finite constants (Pitman drifts) at least one of which is non-zero.

Theorem 2. *Let $\boldsymbol{y}_t$ be generated according to* (1) *and let Assumption 1 hold. Let $LM_{\boldsymbol{d}}^{FGLS}$ be as defined in* (8)*. Then: (i) under the null hypothesis, $H_0 : \boldsymbol{\theta} = \mathbf{0}$, $LM_{\boldsymbol{d}}^{FGLS} \Rightarrow \chi^2_{(k)}$, and (ii) under the sequence of local alternatives, $H_{\boldsymbol{c}} : \boldsymbol{\theta} = \boldsymbol{c}/\sqrt{T}$, with at least one element of $\boldsymbol{c}$ non-zero, $LM_{\boldsymbol{d}}^{FGLS} \Rightarrow \chi^2_{(k,\xi)}$, where $\chi^2_{(k)}$ and $\chi^2_{(k,\xi)}$ denote a standard χ^2 distribution with* *k degrees of freedom, and a non-central χ^2 distribution with k degrees of freedom and non-centrality parameter $\xi := (\boldsymbol{L}'^{-1}\boldsymbol{c})'(\boldsymbol{L}'^{-1}\boldsymbol{c})$, respectively, with $\boldsymbol{L}$ denoting an upper triangular matrix such that $\boldsymbol{L}'\boldsymbol{L} = \boldsymbol{R\Omega}_{\boldsymbol{\beta}}\boldsymbol{R}'$.*

Remark 6. The result in part (i) of Theorem 2 shows that the limiting null distribution of $LM_{\boldsymbol{d}}^{FGLS}$ is pivotal and that a test of $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$ can be run using standard critical values from the $\chi^2_{(k)}$ distribution, where k is the dimension of $\boldsymbol{y}_t$. Part (ii) of Theorem 2 establishes that the asymptotic distribution of $LM_{\boldsymbol{d}}^{FGLS}$ displays a non-trivial positive offset under the local alternative, $\mathrm{H}_{\boldsymbol{c}} : \boldsymbol{\theta} = \boldsymbol{c}/\sqrt{T}$, *vis-à-vis* the null, H_0, but that its asymptotic local power function will, in general, depend on nuisance parameters arising from any weak dependence or cross sectional correlation present in $\boldsymbol{\varepsilon}_t$. The same is also true of the extant multivariate fractional integration tests discussed in section 1, except that these do not, in general, have pivotal limiting null distributions when conditional heteroskedasticity is present in $\boldsymbol{e}_t$, as a consequence of the fact that they are not based around a heteroskedasticity-robust estimate of the variance matrix $\boldsymbol{\Omega}_{\boldsymbol{\beta}}$. ♢

Remark 7. Theorem 2 provides a theoretical basis for the construction of confidence sets. This can be achieved by inverting the non-rejection region of the test statistic; see Hassler et al. (2009). More specifically, let $LM_{\boldsymbol{g}}$ denote the value of the LM statistic when testing $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$ for an arbitrary $\boldsymbol{g} \in \mathbb{R}^k$, and let Ψ be an arbitrary compact set in $\mathbb{R}^k$. Define $\mathcal{D}_\lambda := \left\{\boldsymbol{g} \in \Psi : \Pr\left[\chi^2_{(k)} > LM_{\boldsymbol{g}}\right] \leq 1-\lambda\right\}$ with $\lambda \in (0,1)$, i.e., the subset of Ψ for which H_0 cannot be rejected at the λ significance level. From Theorem 2, it follows that if Ψ is large enough so as to contain the true values of the long memory parameter vector, then the probability of the true order of integration lying within $\mathcal{D}_\lambda$ is at least $(1-\lambda)$. ♢

Remark 8. Our proposed test procedure can be generalised to account for non-zero means following the approach in Robinson (1994). To that end, consider the extended form of the DGP in (1) given by $\boldsymbol{y_t} = \boldsymbol{\mu} + \boldsymbol{\Delta}(L)^{-\boldsymbol{d}-\boldsymbol{\theta}}\boldsymbol{\varepsilon}_t I(t \geq 1)$, where $\boldsymbol{\mu} \equiv (\mu_1, ..., \mu_k)'$ is a fixed vector. Under $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, $(1-L)_+^{d_i} y_{it} = (1-L)_+^{d_i} \mu_i + \varepsilon_t I(t \geq 1)$, $1 \leq i \leq k$. Following Robinson (1994), we regress the differences $(1-L)_+^{d_i} y_{it} := \sum_{j=0}^{t-1} \lambda_j(d_i)\, y_{it-j}$ on $h_{t,d_i} := \sum_{j=0}^{t-1} \lambda_j(d_i)$, $t = 2, ..., T$, with $\{\lambda_j(d_i)\}$ as defined in (4). Denote the resulting estimates as $\widetilde{\mu}_i$, $i = 1, ..., k$, and the corresponding residuals as $\widetilde{\varepsilon}_{it,d_i} := (1-L)_+^{d_i} y_{it} - \widetilde{\mu}_i h_{t,d_i}$. One then simply redefines the ith element of the vector $\boldsymbol{\varepsilon}_{t,\boldsymbol{d}}$ from (2) to be $\widetilde{\varepsilon}_{it,d_i}$, $i = 1, ..., k$, and then proceeds as before. Let $\tilde{\boldsymbol{\beta}}$ denote the FGLS estimator obtained in this way. Then, following the approach taken in Proposition 4 of Demetrescu et al. (2008), it can be shown that Theorem 1 holds with $\widehat{\boldsymbol{\beta}}$ replaced by $\tilde{\boldsymbol{\beta}}$ since $\|\tilde{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}\| = o_p(T^{-1/2})$ under the restrictions considered and the additional condition that $\boldsymbol{d} > 0$. More generally, the results can be extended to account for, among other things, deterministic polynomial time trends and deterministic seasonal effects; see also Nielsen (2005) and Demetrescu et al. (2008). The large sample results given in this section are not affected by accounting for such deterministics.[4] ♢

Remark 9. In practical applications of the tests, the lag order p will typically be unknown and so could be selected using a standard consistent information criterion such as the Bayes information criterion [BIC]. Demetrescu et al. (2008) argue that these can lead to substantial finite-sample biases in the context of the tests considered here. As an alternative, Demetrescu et al. (2008) advocate the use of a deterministic lag selection rule, such as the popular Schwert (1989) rule which sets $p = \lfloor K(T/100)^{1/4} \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part of its argument and K is a finite positive constant. Provided the true lag order p is finite, as we assume in this paper, then the limiting distribution theory given in this section will remain apposite for tests based on a lag length determined according to such deterministic rules. We will implement Schwert's rule in the empirical application considered in section 5. ♢

Remark 10. It is useful to compare the large sample properties of our proposed test with the Gaussian LM test of Nielsen (2005) in comparable settings. To this end, consider the case where, as required by the conditions imposed in Theorem 3 of Nielsen (2005, p.381), Assumption 1 is restricted such that $p = 0$ and $\boldsymbol{e}_t$ is an i.i.d. innovation sequence. It is straightforward to show that under these additional restrictions

$$LM_{\boldsymbol{d}}^{FGLS} = \left(\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes \boldsymbol{I}_{T-1}\right]\boldsymbol{Y}_{\boldsymbol{d}}\right)'\left(\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes \boldsymbol{I}_{T-1}\right]\boldsymbol{X}^{*}_{-1,\boldsymbol{d}}\right)^{-1}\left(\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes \boldsymbol{I}_{T-1}\right]\boldsymbol{Y}_{\boldsymbol{d}}\right) + o_p(1)$$

where $\boldsymbol{X}^{*}_{-1,\boldsymbol{d}}$:=diag$\left\{\boldsymbol{X}^{*}_{1,-1,\boldsymbol{d}}, ..., \boldsymbol{X}^{*}_{k,-1,\boldsymbol{d}}\right\}$ in which $\boldsymbol{X}^{*}_{i,-1,\boldsymbol{d}} = \boldsymbol{Z}^{*}_{-1,\boldsymbol{d}}\boldsymbol{a}_i$, where $\boldsymbol{Z}^{*}_{-1,\boldsymbol{d}} := (\boldsymbol{z}^{*}_{1,\boldsymbol{d}}, ..., \boldsymbol{z}^{*}_{T-1,\boldsymbol{d}})'$ and $\boldsymbol{a}_i$ denotes the i-th unit k-dimensional vector. Noting, moreover, that $\boldsymbol{z}^{*}_{t-1,\boldsymbol{d}} := \sum_{j=1}^{t-1} j^{-1}\boldsymbol{e}_{t-j,\boldsymbol{d}} = -\ln(1-L)_{+}\,\boldsymbol{e}_{t-j,\boldsymbol{d}}$ under the null hypothesis, the vector $\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes \boldsymbol{I}_{T-1}\right]\boldsymbol{Y}_{\boldsymbol{d}}$ corresponds to the Gaussian score vector $\boldsymbol{S}_T := \boldsymbol{J}_k' vec\left(\widetilde{\boldsymbol{\Sigma}}^{-1}\boldsymbol{S}_{10}'\right)$ given in Equation (11) of Nielsen (2005), where $\boldsymbol{S}_{10} := \sum_{t=2}^{T}\boldsymbol{e}^{*}_{t-1}\boldsymbol{e}_t'$ with $\boldsymbol{e}^{*}_{t-1} := \sum_{j=1}^{t-1} j^{-1}\boldsymbol{e}_{t-j}$, and $\boldsymbol{J}_k := (vec(\boldsymbol{A}_{11}), ..., vec(\boldsymbol{A}_{kk}))$ with $\boldsymbol{A}_{ii} := \boldsymbol{a}_i\boldsymbol{a}_i'$. Because $\frac{1}{T}\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes \boldsymbol{I}_{T-1}\right]\boldsymbol{X}^{*}_{-1,\boldsymbol{d}} \xrightarrow{p} \boldsymbol{A}_{\boldsymbol{\beta}}$, where $\boldsymbol{A}_{\boldsymbol{\beta}} = \frac{\pi^2}{6}\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}^{-1}$ under the additional restrictions outlined above, it can

[4] Numerical experiments investigating the finite sample rejection rates of our tests when a non-zero mean is allowed for are given in section B.3 of the supplementary appendix. These confirm the (exact) invariance of such tests and the lack of invariance of tests which do not allow for a non-zero mean. The loss of finite sample power from allowing for a non-zero mean appears very modest.

be seen that $LM_{\boldsymbol{d}}^{FGLS}$ is asymptotically equivalent to the Gaussian LM test proposed in Nielsen (2005). Consequently, $LM_{\boldsymbol{d}}^{FGLS}$ is asymptotically locally efficient when $\boldsymbol{e}_t$ is a Gaussian i.i.d. sequence; see Nielsen (2004a). Where $p > 0$ the two tests differ crucially on how the short-run autocorrelation is handled. While $LM_{\boldsymbol{d}}^{FGLS}$ uses pth-order augmentation in (5), Nielsen's (2005) test relies on pre-whitening using the residuals from a VAR(p) model in a two-stage procedure. Augmentation and pre-whitening are asymptotically equivalent strategies but will differ in finite samples, as will be explored in the next section. $\diamondsuit$

4 Monte Carlo Simulations

We consider the simulation DGP,

$$\begin{bmatrix} (1-L)^{1+\theta_1} & 0 \\ 0 & (1-L)^{1+\theta_2} \end{bmatrix} \boldsymbol{y}_t = \boldsymbol{\varepsilon}_t I(t \geq 1), \quad t = 1, ..., T, \tag{9}$$

where $\boldsymbol{y}_t \equiv (y_{1t}, y_{2t})'$, $\boldsymbol{\Pi}(L)\boldsymbol{\varepsilon}_t = \boldsymbol{e}_t$ with $\boldsymbol{\Pi}(L) = \text{diag}\{1 - \pi_1 L,\, 1 - \pi_2 L\}$, and $(\pi_1, \pi_2) \in \{(0,0), (0.4, 0.4)\}$; such that the former corresponds to white noise, while the latter yields weakly stationary VAR(1) errors. As the particular values of the long memory coefficients play no role in our context, we set $d_1 = d_2 = 1$. We report results for $T \in \{500, 1000\}$.

The innovations $\{\boldsymbol{e}_t\}$ are generated to exhibit time-varying conditional second-order moments according to the design,

$$\boldsymbol{e}_t = \begin{bmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{bmatrix} \boldsymbol{\eta}_t; \quad E(\boldsymbol{\eta}_t) = 0, \quad E(\boldsymbol{\eta}_t \boldsymbol{\eta}_t') =: \boldsymbol{\Omega}_\rho = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

where $\boldsymbol{\eta}_t := (\eta_{1t}, \eta_{2t})'$ is an i.i.d. vector drawn from either a multivariate Gaussian distribution or a (heavy-tailed) multivariate Student-t distribution with 5 degrees of freedom. The covariance matrix $\boldsymbol{\Omega}_\rho$ depends on the contemporaneous correlation coefficient ρ, $\rho \in \{0,$ $0.2,\ 0.4,\ 0.6,\ 0.8\}$. The conditional variances $\{\sigma_{it}^2\}$ are driven by (normalised) stationary GARCH(1,1) processes $\sigma_{it}^2 = (1 - \alpha - \beta) + \alpha e_{i,t-1}^2 + \beta \sigma_{i,t-1}^2$, $i = 1, 2$ with $\alpha, \beta \geq 0$ and $\alpha + \beta < 1$, such that $E(e_{it}^2) = 1$. We consider $(\alpha, \beta) \in \{(0,0),$ (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.

To simplify our discussion, we fix $\theta_2 = 0$ in all of the reported simulations and vary θ_1 among $\{-0.3, -0.25, ..., 0, ..., 0.25, 0.3\}$. Consequently, while the true order of integration of $\{y_{2t}\}$ is always one, the true order of integration of $\{y_{1t}\}$ is $1 + \theta_1$. The case where $\theta_1 = 0$ allows us to investigate the empirical size properties of $LM_{\boldsymbol{d}}^{FGLS}$, while the cases where

$\theta_1 \neq 0$ allow us to investigate its finite sample power against an alternative where one of the long memory parameters deviates from the null hypothesis. 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 $LM_{\boldsymbol{d}}^{FGLS}$ and determine the empirical rejection frequencies [ERFs] at the 5% nominal (asymptotic) level over $5,000$ replications.

We also benchmark the performance of $LM_{\boldsymbol{d}}^{FGLS}$ against two alternative (but related) tests. The first is the multivariate LM test of Nielsen (2004a, 2005) discussed in Remark 10 above, denoted $LM_{\boldsymbol{d}}^{MLE}$ in what follows, and the second is the multivariate trace test of Breitung and Hassler (2002), which we denote $BH_{\boldsymbol{d}}$. While $LM_{\boldsymbol{d}}^{FGLS}$ corrects for stationary serial correlation in $\boldsymbol{\varepsilon}_t$ via lag augmentation in (5), both $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ use a pre-whitening approach. Both $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ require that $\{\boldsymbol{e}_t\}$ is i.i.d., and so neither allows for the presence of conditional heteroskedasticity in $\{\boldsymbol{e}_t\}$. Under these conditions, $LM_{\boldsymbol{d}}^{MLE}$ has a limiting $\chi^2_{(k)}$ null distribution, while $BH_{\boldsymbol{d}}$ has a limiting $\chi^2_{(k^2)}$ null distribution. Nielsen's $LM_{\boldsymbol{d}}^{MLE}$ is designed to test the same null hypotheses as $LM_{\boldsymbol{d}}^{FGLS}$ and so is the most natural candidate to benchmark our test against. In contrast, the $BH_{\boldsymbol{d}}$ test is for the null hypothesis of a common order of integration between the elements of the vector time series. Our simulation DGP is such that this condition holds under the null hypothesis, but not under the alternative, so a comparison with this test is appropriate.

4.1 ERFs with no Augmentation/Pre-whitening

Table 1 reports the empirical size properties ($\theta_1 = \theta_2 = 0$), for $LM_{\boldsymbol{d}}^{FGLS}$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ where no short-run dynamics are present ($\pi_1 = \pi_2 = 0$), and where, accordingly, no lag augmentation or pre-whitening is needed. This allows us to first investigate the impact of GARCH effects, contemporaneous correlations, and the conditional distribution of the innovations on each test.

The results show that $LM_{\boldsymbol{d}}^{FGLS}$ displays ERFs close to the nominal asymptotic 5% level in almost all cases. Some mild over-sizing is seen for the smaller sample size considered when the innovations are conditionally Student-t distributed with relatively high GARCH persistence, $\alpha = 0.1$ and $\beta \geq 0.80$, and significant levels of endogeneity, $\rho \geq 0.4$. These distortions are largely ameliorated as the sample size increases. Where the innovations are i.i.d. ($\alpha = \beta = 0$), both $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ display good finite sample size control regardless of the conditional distribution or the degree of endogeneity. However, where the innovations exhibit conditional heteroskedasticity a very different pattern emerges for both $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$. These tests display a tendency to strong over-sizing, with these distortions being larger (other things equal): the stronger the degree of persistent of the GARCH process; the larger the degree of endogenous correlation $|\rho|$; and for innovations drawn from a heavy-tailed distribution. Moreover, these size distortions are not ameliorated as the sample size increases. To illustrate, for $T = 500$ and $\rho = 0.8$, the ERFs of $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ with GARCH errors driven by $(\alpha, \beta) = (0.10, 0.85)$ and Student-t innovations are 36% and 38.8%, respectively. In contrast, the $LM_{\boldsymbol{d}}^{FGLS}$ test is only slightly oversized at 7.1%. For $T = 1000$, the corresponding ERFs of $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ increase significantly to 48.3% and 52.8%, respectively, while that of $LM_{\boldsymbol{d}}^{FGLS}$ reduces to 6.1%.

4.2 ERFs with Augmentation/Pre-whitening

We now analyse the finite sample size and power properties of $LM_{\boldsymbol{d}}^{FGLS}$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ in the case where the errors, $\boldsymbol{\varepsilon}_t$, can display first-order stationary VAR dynamics. Accordingly, we set $p = 1$ in (5) in relation to the $LM_{\boldsymbol{d}}^{FGLS}$ test, while analogously we use a VAR(1) for pre-whitening in connection with the $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ tests. For $\boldsymbol{\varepsilon}_t$ we consider: (i) $\pi_1 = \pi_2 = 0$, so that augmentation/pre-whitening is in fact unnecessary, and (ii) $\pi_1 = \pi_2 = 0.4$, so that the correct order of augmentation/pre-whitening is employed.

Table 2 reports ERFs of the three tests in the Gaussian homoskedastic case ($\alpha = \beta = 0$). Results for the Student-t case are not reported as these are almost identical to the results reported in Table 2. Also, to keep the size of the subsequent tables to manageable proportions we will only report results for two values of the correlation coefficient, namely $\rho = 0$ and $\rho = 0.8$. Corresponding results for $\rho \in \{0.2, 0.4, 0.6\}$ can be obtained on request.

Consider first the results for the case where $\theta_1 = 0$ so that the null hypothesis holds. Here we see that the ERFs of the augmented $LM_{\boldsymbol{d}}^{FGLS}$ test lie close to the nominal asymptotic level throughout, even where the lag augmentation is unnecessary. Pre-whitening also appears to be effective for the $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ tests, with the exception of the case where $\rho = 0.8$ where these tests are somewhat oversized for $T = 1000$.

Turning next to the empirical power results for $\theta_1 \neq 0$, we see that $LM_{\boldsymbol{d}}^{FGLS}$ displays good finite sample power properties with power increasing, other things equal, both as $|\theta_1|$ increases and as T increases, as would be expected. Power is also larger, other things equal, for $\rho = 0.8$ than for $\rho = 0$, illustrating the efficiency benefits gained from multivariate modelling when the variables are cross-correlated. In terms of a comparison between the three tests, overall the finite sample power properties of $LM_{\boldsymbol{d}}^{FGLS}$ and $LM_{\boldsymbol{d}}^{MLE}$ seen in

Table 2 are very similar for alternatives where $\theta_1 < 0$, as might be expected given the asymptotic equivalence of these tests when the innovations are i.i.d.; cf. Remark 10. For alternatives where $\theta_1 > 0$ (i.e., when the process is more persistent than posited under the null) $LM_{\boldsymbol{d}}^{FGLS}$ can display somewhat higher power than $LM_{\boldsymbol{d}}^{MLE}$, particularly in the case where the errors are first-order autocorrelated, $\pi_1 = \pi_2 = 0.4$; for example, for $\pi_1 = \pi_2 = 0.4$, $\rho = 0$, $T = 500$, and $\theta_1 = 0.3$ the power of $LM_{\boldsymbol{d}}^{FGLS}$ and $LM_{\boldsymbol{d}}^{MLE}$ are 57.4% and 32.3%, respectively. These differences are likely attributable to the use of lag augmentation rather than pre-whitening in the construction of $LM_{\boldsymbol{d}}^{FGLS}$. Both $LM_{\boldsymbol{d}}^{FGLS}$ and $LM_{\boldsymbol{d}}^{MLE}$ clearly dominate $BH_{\boldsymbol{d}}$ on power; in the previous example the power of $BH_{\boldsymbol{d}}$ is only 26.3%. The power functions of all of the tests are asymmetric in the sign of θ_1, for a given DGP, such that a false null hypothesis which leads to an over-differenced series ($\theta_1 < 0$) is seen to be more easily rejected than an incorrect null which leads to an under-differenced series ($\theta_1 > 0$) where the magnitude of the under/over difference is the same. To illustrate, for $\pi_1 = \pi_2 = 0.4$, $\rho = 0$ and $T = 500$, the power of $LM_{\boldsymbol{d}}^{FGLS}$ to detect $\theta_1 = 0.25$ and $\theta_1 = -0.25$ is 49.6% and 72.0%, respectively. Breitung and Hassler (2002) report a similar asymmetry in the power properties of their univariate tests.

Finally, we turn to the case where the innovations may display GARCH effects and excess kurtosis. Table 3 ($T = 500$) and Table 4 ($T = 1000$) report the ERFs for $LM_{\boldsymbol{d}}^{FGLS}$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ for both Gaussian and Student-t innovations for $\pi_1 = \pi_2 = 0.4$, $\rho \in \{0, 0.8\}$, $(\alpha, \beta) \in \{(0.10, 0.80), (0.10, 0.85)\}$. The results for $\theta_1 = 0$ show that the empirical size properties of the tests in the presence of GARCH are similar to the corresponding results reported previously for the serially uncorrelated case with no augmentation/prewhitening in Table 1. In particular, while the empirical size of $LM_{\boldsymbol{d}}^{FGLS}$ is reasonably close to the nominal asymptotic 5% level throughout (size departures are not greater than 1.6% for $T = 500$ and not greater than 0.8% for $T = 1000$), incorrectly assuming conditional homoskedasticity causes significant over-sizing in both $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ which is not ameliorated by increasing the sample size. To illustrate, for $\rho = 0.8$, and $(\alpha, \beta) = (0.10, 0.85)$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$, respectively, display ERFs of 9.2% and 8.7% for $T = 500$ and 10.4% and 9.9% for $T = 1000$ with Gaussian innovations, increasing to 28.2% and 32.4% for $T = 500$ and 40.4% and 46.2% for $T = 1000$ with Student-t innovations.

For non-zero values of θ_1, we observe qualitatively similar patterns in relation to the power properties of $LM_{\boldsymbol{d}}^{FGLS}$ as were reported in Table 2 in the homoskedastic case, albeit persistent GARCH-type behaviour in the innovations can be seen to clearly lower the finite sample power of $LM_{\boldsymbol{d}}^{FGLS}$ relative to the i.i.d. case, and particularly so when the conditional distribution of the innovations is heavy-tailed. This is of course consistent with Theorem 2 where it was shown that the asymptotic local power function of the $LM_{\boldsymbol{d}}^{FGLS}$ test depends on any nuisance parameters arising from conditional heteroskedasticity in the innovations. To illustrate, from Table 2 for $\pi_1 = \pi_2 = 0.4$, $\rho = 0$, $T = 500$, the power of $LM_{\boldsymbol{d}}^{FGLS}$ to detect $\theta_1 = 0.3$ ($\theta_1 = -0.3$) in the i.i.d. case is 57.4% (87.5%). However, from Table B.1, 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. Similarly, for $T = 1000$ in the previous example power is seen from Table 3 to be 98.7% (99.9%) in the Gaussian case, and 71.6% (81.1%) in the Student-t case. A comparison between the finite sample power of $LM_{\boldsymbol{d}}^{FGLS}$ and that of $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ is somewhat uninformative here because of the poor size control of the latter two tests under conditional heteroskedasticity.

5 Long-run Dynamics in Volume and Volatility

Understanding the linkages between return volatility, liquidity and trading activity has been an area of considerable research interest in the finance literature. We apply the multivariate testing approach developed in this paper to perform joint inference on the order of fractional integration of trading volume and return volatility for a sample of major stocks traded in the U.S. market. As part of this, we also investigate the hypothesis that these variables exhibit the same order of fractional integration.

A number of previous studies have investigated this hypothesis in trading volume and return volatility within a multivariate ARFIMA framework. No strong consensus has emerged across these studies which are based on a variety of methods of estimation and inference and employ a number of different observable variables to proxy the latent return volatility process. Bollerslev and Jubinski (1999) and Lobato and Velasco (2000) use semiparametric multivariate periodogram-based estimators in the frequency domain, proxying return volatility by absolute-valued returns. They conclude that, for most of the stocks analysed, the hypothesis that trading volume and return volatility share the same order of fractional integration cannot be rejected. However, Fleming and Kirby (2011) argue that the slow rate of convergence of periodogram-based estimators raises concerns about estimation efficiency. Consequently, they implement a parametric Gaussian quasi-maximum likelihood (QML) approach as in Nielsen (2004a) to estimate a bivariate FIVAR model, allowing for short-run dependencies, but under the assumption of conditional homoskedasticity. Moreover, Fleming and Kirby (2011) proxy return volatility using intra-day data with the aim of improving accuracy over the use of absolute-valued returns and reject the hypothesis of a common long memory coefficient in most cases.

Our testing procedure is expected to be useful here for two key reasons. First, as shown in Theorem 1, the FGLS-based test achieves the usual $\sqrt{T}$ rate of convergence in parametric testing, and is therefore expected to yield improved finite-sample power performance relative to periodogram-based estimators; see, for example, Tanaka (1999). This consideration addresses concerns surrounding efficiency raised by Fleming and Kirby (2011). Second, and arguably most importantly, our testing approach is valid in the presence of stationary conditionally time-varying second-order moments and heavy-tailed innovations, unlike the QML approach of Nielsen (2004a) used by Fleming and Kirby (2011).

5.1 Data

Our analysis focuses on 30 major U.S. stocks from the DJI. We analyse data sampled from 02/01/2003 to 31/12/2014. Unlike trading volume, return volatility cannot be directly observed. The literature has suggested a number of different estimation methods in increasing degree of accuracy, which we implement. The simplest approach uses absolute-valued returns computed from close-to-close daily prices. Unfortunately, this measure is known to be highly inefficient and subject to large estimation errors. More accurate estimates can be constructed building on intra-day information. Following Garman and Klass (1980), we also proxy daily return 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. Such range-based estimators produce more efficient estimates than absolute-valued returns computed from close-to-close prices (Parkinson (1980)) and, as discussed in Andersen and Bollerslev (1998), can be as efficient a measure of return volatility as realised volatility computed on the basis of three to four hour returns. The last estimator we consider is a realised variance measure computed from aggregating 5-minute squared continuously compounded returns over the trading session. Daily share volumes and high, low, opening and closing prices are obtained from CRSP. High-frequency prices necessary to compute realised variances are obtained from the NYSE Trade and Quote (TAQ) database. As is customary in this literature, we implement log-transforms in both trading volume and return volatility variables. Standard descriptive statistics for the aforementioned variables as well as a statistical analysis highlighting statistically significant evidence for the presence of time-varying second order moments in the data are presented in Tables C1 and C2 of the supplementary appendix.

5.2 Implementation Issues

In conducting our analysis of the long memory properties of log-trading volume and log-return volatility, hereafter denoted as $(d(vlm), d(\sigma))'$, a number of key implementation issues arise which we now detail.

First, we construct 99%, 95%, and 90% confidence sets for $(d\,(vlm)\,, d\,(\sigma))'$ by inverting the non-rejection regions of the multivariate test in a discrete grid search over the support $\Psi = [-0.2, 1.2] \times [-0.2, 1.2]$ (see Remark 7). More specifically, we evaluate $LM_{\boldsymbol{d}}^{FGLS}$ for any pair of values d_1 and d_2 in the grid sequence $\{-0.2, -0.1, ..., 1.1, 1.2\}$. Point estimates of the long memory parameter vector can also be obtained by minimising the value of $LM_{\boldsymbol{d}}^{FGLS}$ over Ψ; notice this estimate does not depend on the confidence level used. This method of point estimation has been used in the univariate context; see, for example, Gil-Alaña and Robinson (1997). We denote the resulting point estimates of the long memory parameter for log-trading volume and log-volatility as $\widehat{d}_{\min}\,(vlm)$ and $\widehat{d}_{\min}\,(\sigma)$, respectively.[5]

Second, to account for deterministic effects in the level of these series, we apply the OLS-based demeaning procedure described in Remark 8. While most papers do not consider deterministic trends as a stylised feature of return volatility, trading volume is widely accepted to exhibit trending paths conformable with increasing growth in the number of traders and trading activity; see Fleming and Kirby (2011) and references therein. Consequently, for the log-volatility measures, our main analysis is carried out by including a constant to capture a non-zero drift, as in Hassler et al. (2016), while in the case of log-volume we allow for a quadratic time trend polynomial of the form $\mu_t = \mu_0 + \mu_1\,(t/T) + \mu_2(t/T)^2$, as advocated by, among others, Luu and Martens (2003) and Fleming and Kirby (2011). Parameters in these functions are estimated through univariate OLS (see Remark 8), with the multivariate fractional integration test then computed on the resultant residuals.

Third, as discussed in Remark 9, we determine the lag length according to Schwert's rule, $p = \lfloor 4(T/100)^{1/4} \rfloor$. Given the large sample size involved, Schwert's rule ensures a relatively long lag length, so that the short-run component of log-volume and log-realised variance should be well captured in the auxiliary regression. Andersen et al. (2003) also adopt a relatively long lag length in estimating their FIVAR model for the realised volatility

[5]Numerical experiments investigating the finite sample accuracy (bias and MSE) of these estimates in the context of a bivariate model are reported in section B.2 of the supplementary appendix.

of exchange rates in order to maintain a conservative approach.[6]

5.3 Main Results

For each stock and for each volatility measure, Table 5 reports the resulting point estimates $\widehat{d}_{\min}(vlm)$ and $\widehat{d}_{\min}(\sigma)$. Table 5 also gives the *upper* and *lower* bounds of the corresponding 95% confidence ellipsoids formed as the vertical and horizontal projections of the confidence set onto the log-trading volume and log-volatility axes, respectively.[7] The columns headed "Common d" in Table 5 report the range of values $\overline{d}$ for which the null hypothesis $\mathrm{H}_0 : d(vlm) = d(\sigma) = \overline{d}$ cannot be rejected at the asymptotic 5% nominal significance level. If this region is non-empty, it shows the set of values along the 45-degree line contained within the 95% confidence ellipsoid; that is, those values of a common order of fractional integration for which the null cannot be rejected. Notice that, by construction, the resulting interval contains the true value of a common long memory parameter with an (asymptotic) probability not smaller than 95%. In addition, given $\widehat{d}_{\min}(vlm)$ and $\widehat{d}_{\min}(\sigma)$, we can compute the residuals from the multivariate FGLS regression and use these to estimate the contemporaneous correlation between the innovations to log-volume and a given return volatility measure; this estimate is denoted by $\widehat{\rho}_e$ in Table 5. Large values of $\widehat{\rho}_e$ are supportive of the usefulness of the multivariate approach we advocate. And, indeed, we see from Table 5 that this estimated correlation is generally quite large and positive.

Let us first discuss the results from the analysis of the joint dynamics of log-volume and log absolute returns. Consistent with previous literature, we observe that for most stocks considered our multivariate test rejects both the null hypothesis that the order of integration of the bivariate series is I(0) (such that both variables are weakly dependent) and the null hypothesis that it is I(1) (such that both series admit an autoregressive unit root). The only exceptions are INTC (Intel) and MSFT (Microsoft), for which the multivariate

[6]We also investigated the robustness of our main conclusions to the lag augmentation order used in the FGLS regression and to the inclusion of a deterministic time trend in connection with the return volatility measures. To that end, as in Fleming and Kirby (2011), we also looked at the case where a linear time trend was allowed for in return volatility and a low-order VAR(p) was fitted. Table C.3 in the supplementary appendix reports the main results from this analysis, focusing directly on log-volume and log-realised variance, with $p = 2$ and both with and without a linear time trend in return volatility. Here we also report the related results when p is chosen according to Schwert's rule and return volatility includes a deterministic time trend. While the results show some sensitivity to these variations in the estimated model, the main qualitative picture that emerges is essentially very similar to that discussed below.

[7]These bounds (projections) define a rectangular approximation to the true confidence interval ellipsoids, whose area cannot be smaller than that of the true ellipsoid. However, they have the advantage that they provide a summary measure which can easily be tabulated. The full set of confidence ellipsoids for each stock considered can be found in sections C.2.2-C.2.4 of the supplementary appendix.

test cannot reject the null hypothesis that log-volume is I(0) at the 5% level. Taking a simple average of the estimates $\widehat{d}_{\min}(vlm)$ and $\widehat{d}_{\min}(\sigma)$ across all stocks considered yields 0.41 and 0.39, respectively, essentially matching the "characteristic" value of 0.40 typically found in literature; see, Andersen et al. (2003). While for many of the stocks considered the point estimates of the vector of fractional exponents are below the non-stationary threshold, we note that for most of the stocks the respective confidence sets cover both the stationary and non-stationary regions of the parameter space, preventing us from drawing clear conclusions on the stationarity of the underlying series. This is a common finding in the realised-volatility modelling literature; see, e.g., Kellard and Sarantis (2010).

Reflecting the strong similarities seen between the estimates of the two long memory parameters in the bivariate system, the hypothesis that trading volume and return volatility are driven by a FIVAR model with the same fractional exponent can be rejected for only five of the stocks considered at the 5% level, which constitutes about 20% of the stocks in our sample. This is, however, considerably higher than the corresponding frequency found by Bollerslev and Jubinski (1999) who only reject for 8% of the series they considered, but is the same as Lobato and Velasco (2000) who also reject the null hypothesis of a common long memory parameter for 20% of the series they consider.[8] In their study of log-volume and log absolute returns, Fleming and Kirby (2011) reject the common long memory parameter null for 100% of the series they analyse. They attribute this to estimation bias in the QML-based inference they use yielding systematically larger parameter estimates for the long memory coefficient for trading volume, and conjecture that departures from normality in log absolute returns may be causing a pervasive effect on QML estimation; see Fleming and Kirby (2011, pp.1721-1722). Our Monte Carlo simulations in section 4 accord with this conjecture suggesting that the combination of persistent time-varying volatility and non-Gaussian features in the data can introduce sizeable biases into the QML-based methods of Nielsen (2004a, 2005) used by Fleming and Kirby (2011).

We now move to a discussion of the results relating to the use of the log-range and log-realised proxies for return volatility. The overall picture that emerges here 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) null hypotheses both being rejected at

[8] In making such comparisons it is important to note, however, that these authors use different sample data than we do involving different stocks and different time periods. In particular, the sample lengths considered in Lobato and Velasco (2000) are more than double those we consider and our findings of the same frequency of rejections of a common exponent as they do may reflect the greater efficiency of the methods used here.

the 5% level for all of the stocks considered. For log trading volume, some changes are seen relative to the results discussed previously relating to the use of log absolute returns.[9] Overall, the average value of $\widehat{d}_{\min}(vlm)$ taken across all of the stocks considered for the log-range and log-realised variance estimators is 0.45, reasonably similar to the 0.41 value reported above in connection with the use of log absolute returns. In contrast, the estimates of the order of integration of return volatility based on either the log-range or log-realised variance measures show a marked increase compared to log absolute returns. In both cases, the average value of $\widehat{d}_{\min}(\sigma)$ is 0.58 and the overall evidence is strongly suggestive that return volatility displays non-stationary dynamics over the period because the lower bounds of the confidence ellipsoids are not smaller than the 0.5 threshold for many of the stocks considered. Evidence of non-stationary fractionally integrated dynamics in realised volatility over this period (which includes the financial crisis) is consistent with the results reported by Hassler et al. (2016); see also Bandi and Perron (2006). Consequently, and because the persistence of log-realised variance tends to be greater than that of log-volume, the hypothesis that both variables share a common fractional exponent is rejected at the 5% level for a significantly larger proportion of the stocks considered, namely, 53.33% when using log-range and 63.33% when using log-realised variance.

It is well understood in the financial econometrics literature that measurement errors in absolute returns can cause bias in (univariate) long memory parameter estimation. Essentially, log absolute returns are subject to noisy additive measurement errors with large variability, which will make the underlying process appear less persistent than it really is, leading to downward-biased estimates of the true order of fractional integration; see, among others, Bollerslev and Wright (2000), Haldrup and Nielsen (2007), and Dalla (2015). This provides a straightforward and plausible explanation for the systematic differences seen in the long memory estimation results for the different return volatility measures reported in Table 5.[10] According to our results, the more efficient the estimate of return volatility used the higher the percentage of the stocks for which the null hypothesis of a common order of integration can be rejected. Essentially, downward biases in the estimation of the long memory parameter on absolute returns biases the tests to non-rejection of a common order

[9]Because we conduct joint estimation, and the innovations to the short-term component of volume and return volatility are strongly positively correlated, as reported in the column $\widehat{\rho}_e$ in Table 5, the estimates of the long memory parameter of log-volume would be expected to be somewhat sensitive to changes in the variable used as a proxy for return volatility.

[10]An alternative explanation, put forward by a referee, is that the VAR dynamics may be misspecified and, as a result, some of the high-frequency measurement error is picked up in the estimate of the fractional exponent. However, the relatively long lag length used should mitigate against this and, moreover, as discussed in footnote 6 the results appear relatively robust to the lag order specified.

of integration. Using more accurate return volatility measures reduces this estimation bias and leads to increased evidence that return volatility is more persistent than volume.

5.4 Fractional Cointegration

As noted in Remark 5, our FGLS-based LM test, like the LM test of Nielsen (2005), assumes the absence of fractional cointegration between the variables, and diverges if fractional cointegration is present. Given we reject the null hypothesis of a common order of integration for trading volume and return volatility for most of the stocks considered, we now also investigate the order of fractional integration of the series using the semiparametric approach of Nielsen and Shimotsu (2007) (NS henceforth), detailed in the supplementary appendix. NS's procedure allows us (under certain regularity conditions) to consistently estimate the cointegration rank of the series and, using the approach of Robinson and Yajima (2002), to test the null hypothesis that the elements of long memory vector, $\boldsymbol{d}$, are equal (although it is important to note that this is not a multivariate test as it is based on the univariate estimates of the fractional exponents). Denoting the statistic for the latter as T_0, NS show that $T_0 \xrightarrow{p} 0$ when the cointegration rank, r, is greater than zero (i.e., where the variables are cointegrated), whereas $T_0 \Rightarrow \chi^2_{(1)}$ when $r = 0$ (where the variables are not cointegrated) and the null of an equal order of integration holds on $\boldsymbol{d}$. NS argue that large values of the test statistic provide evidence against the hypothesis of a common order of integration, regardless of whether the underlying series are fractionally cointegrated or not.

Given that the highest frequency of rejections of a common fractional exponent occurred when using log-realised variances, we only report that case here. Consistent with the analysis in Table 5, we account for a deterministic drift in log-volatility and a polynomial time trend in log-volume by prior detrending of the data, using the two-stage exact local Whittle estimator in Shimotsu (2010). Following the empirical analysis in NS, we estimated $\boldsymbol{d}$ by setting $m_T = \lfloor T^{0.6} \rfloor$ and compute T_0 with $s_T = \log T$. Following NS we also use $m_{1T} = \lfloor T^{0.55} \rfloor$ and $v_T = m_{1T}^{-0.3}$ in the estimation of the cointegration rank. Table 6 reports the point estimates and 95% asymptotic confidence level estimates of $\boldsymbol{d}$, the T_0 test statistic and related p-values, the values of the function $L(u)$ used in the model selection procedure, and the estimates of the cointegration rank, $\widehat{r}_T$. The column $\widehat{r}_T^*$ reports the conditional estimates of r for the cases in which the hypothesis of a common order of integration cannot be rejected at the 5% nominal size level.

Three key features arise from this analysis. First, the results based on the NS test provide the same qualitative evidence as the tests based on FGLS estimation. There exists strong evidence of fractional integration in both series which is again suggestive that realised volatility is more persistent than trading volume. Accordingly, the hypothesis of an equal order of integration is rejected at any of the usual significance levels for the majority of the stocks in our sample. Second, in most cases, the FGLS and the NS tests agree on whether to reject or not the null hypothesis of a common order of fractional integration. In particular, all of the cases in which the FGLS test rejects at the 5% level correspond to stocks for which the NS test also rejects at this level. There are, however, 5 stocks for which NS rejects the null but the FGLS test does not, so the average rejection rate of the NS test taken across all of the stocks considered is slightly higher at 76.67%. Crucially, the p-values of the T_0 test in three of these 5 cases are only slightly below the 5% threshold, suggesting that the differences with the FGLS test are caused by only marginal differences in significance. Finally, the estimates of the cointegration rank suggest that volume and realised volatilty are not in general cointegrated, supporting the suitability of FIVAR-type modelling. In particular, fractional cointegration, indicated by $\widehat{r}_T = 1$ or $\widehat{r}_T^* = 1$, is found for only 13.33% of the stocks, but crucially these are all stocks for which neither the NS T_0 test nor our FGLS-based tests reject the null hypothesis of a common order of integration; that is, none of the rejections of a common order of integration seen with the FGLS-based test in Table 5 are associated with a non-zero estimate of the cointegration rank.

It is worth pointing out in conclusion that the assumptions on which the NS approach are based include the requirement of conditional homoskedasticity. To check how sensitive the NS test is to violations of this assumption, we conducted a small Monte Carlo experiment using the sample simulation DGP as in section 4. For samples of size $T = 500$ and $T = 1000$, $\rho = 0.8$ and persistent GARCH processes with $(\alpha, \beta) = (0.1, 0.85)$, the ERFs of the T_0 test under Gaussian innovations at the 5% nominal asymptotic level were 3.10% and 4.30%, respectively, suggesting approximately correct size. However, under Student-t innovations with 5 degrees of freedom, the respective ERFs were 15.5% and 24.5%. Although clearly oversized, these distortions are considerably smaller than those that were seen in the corresponding results in section 4 for the QML-based test of Nielsen (2005). These simulation results might help explain the differences seen between the results for the FGLS and T_0 tests in our empirical study whereby slightly more rejections of the null of a common order of integration are obtained when using the T_0 test, given that this test has a tendency to be somewhat oversized when the data display conditional heteroskedastcity and heavy-tailed behaviour, as is the case with the data in our empirical study.

6 Conclusions

We have proposed a new test for fractional integration in the context of a quite general FIVAR model which allows for conditional heteroskedasticity in the innovations, does not require the order of integration of the elements of the vector time series to coincide or to lie in a certain region (thereby allowing for both stationary and non-stationary dynamics) and does not assume a particular distribution for the innovations. To the best of our knowledge, none of the methods in the extant literature has achieved this degree of flexibility. Our approach is based on an LM-type test statistic using a heteroskedasticity-robust estimate of the variance matrix, and can be readily implemented using FGLS estimation in a regression-based context. We have demonstrated that our proposed test statistic has a standard χ^2 asymptotic null distribution, that the test exhibits non-trivial power to reject against a sequence of local alternatives, and that in the case of i.i.d. Gaussian errors the test is asymptotically locally efficient. Monte Carlo analysis was used to show that while our test is approximately correctly sized in finite samples of data exhibiting conditional heteroskedasticity and heavy-tailed features, extant tests in the literature which neglect conditional heteroskedasticity can be severely over-sized even for very large samples.

In an empirical case study we have used our proposed testing procedure to jointly infer the order of fractional integration of trading volume and return volatility in a sample of major stocks traded in the U.S. market. Return volatility was proxied by three different measures with increasing degrees of accuracy: absolute returns, a range-based estimator, and a realised variance computed over 5-minute returns. The evidence from the analysis based on the realised variance and range-based estimates delivered similar conclusions, namely, that for many stocks in the sample return volatility is more persistent than trading volume. On the other hand, the analysis based on log absolute returns showed that volume and return volatility share the same order of fractional integration. Because long memory estimation in absolute returns is known to be downward-biased, measurement errors in the data would seem to be a plausible explanation for the evidence from log absolute returns.

For applied work it is of interest that our conclusions based on the realised variance and range-based estimators of return volatility were very similar. While the former is a more efficient estimate of conditional variability, the latter seems to provide a reasonable enough level of accuracy such that the conclusions drawn from the data are not markedly different. This might be a useful observation in practice because for many applications the high-frequency intra-day data needed to construct realised variance and related measures is often not available, for example when considering small or illiquid markets. In the absence of intra-day data, but when information on high and low prices are available, inference based on range-based volatility estimates may still lead to reliable conclusions.

We finish with a suggestion for further research. Here we have proposed parametric FGLS-based multivariate fractional integration tests which, unlike other extant parametric tests, allow for conditionally heteroskedastic innovations. There are relatively few semi-parametric multivariate fractional integration tests in the literature, most notably Lobato and Robinson (1998), Lobato (1999), Marinucci and Robinson (2001) and Shimotsu (2007), all of which assume conditionally homoskedastic innovations. Investigating whether or not these tests remain asymptotically valid under conditional heteroskedasticity and, if so, comparing their finite sample performance with the tests developed in this paper is beyond the scope of the present paper but would constitute an interesting topic for further research.

Acknowledgments

We thank the Co-Editor, Eric Ghysels, and three anonymous referees for their helpful and constructive comments on an earlier version of this paper. Balboa and Rubia gratefully acknowledge financial support from the Spanish Ministry of Economy through project ECO2017-87069-P. Rodrigues gratefully acknowledges financial support from the Portuguese Science Foundation (FCT) through project PTDC/EGE-ECO/28924/2017, and (UID/ECO/00124/2013 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

References

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics 73*, 5–59.

Bandi, F. M. and B. Perron (2006). Long memory and the relation between implied and realized volatility. *Journal of Financial Econometrics 4*, 636–670.

Bollerslev, T. and D. Jubinski (1999). Equity trading volume and volatility: latent information arrivals and common long-run dependencies. *Journal of Business and Economic Statistics 17*, 9–21.

Bollerslev, T. and J. Wright (2000). Semiparametric estimation of long-memory volatility dependencies: the role of high-frequency data. *Journal of Econometrics 98*, 81–106.

Boswijk, H. P., G. Cavaliere, A. Rahbek, and A. M. R. Taylor (2016). Inference on cointegration parameters in heteroskedastic vector autoregressions. *Journal of Econometrics 192*, 64–85.

Breitung, J. and U. Hassler (2002). Inference on the cointegration rank in fractionally integrated processes. *Journal of Econometrics 110*, 167–185.

Cavaliere, G., M. O. Nielsen, and A. M. R. Taylor (2017). Quasi-maximum likelihood estimation and bootstrap inference in fractional time series models with heteroskedasticity of unknown form. *Journal of Econometrics 198*, 165–188.

Dalla, V. (2015). Power transformations of absolute returns and long memory estimation. *Journal of Empirical Finance 33*, 1–18.

Demetrescu, M., V. Kuzin, and U. Hassler (2008). Long memory testing in the time domain. *Econometric Theory 24*, 176–215.

Fleming, J. and C. Kirby (2011). Long memory in volatility and trading volume. *Journal of Banking and Finance 35*, 1714–1726.

Garman, M. B. and M. J. Klass (1980). On the estimation of security price volatilities from historical data. *The Journal of Business 53*, 67–78.

Gil-Alaña, L. A. and P. M. Robinson (1997). Testing of unit root and other nonstationary hypotheses in macroeconomic time series. *Journal of Econometrics 80*, 241–268.

Haldrup, N. and M. O. Nielsen (2007). Estimation of fractional integration in the presence of data noise. *Computational Statistics and Data Analysis 51*, 3100–3114.

Hassler, U., P. M. M. Rodrigues, and A. Rubia (2009). Testing for the general unit root hypothesis in the time domain. *Econometric Theory 25*, 1793–1828.

Hassler, U., P. M. M. Rodrigues, and A. Rubia (2016). Quantile regression for long memory testing: a case of realized volatility. *Journal of Financial Econometrics 25*, 693–724.

Kellard, N., C. D. and N. Sarantis (2010). Foreign exchange, fractional cointegration and the implied–realized volatility relation. *Journal of Banking and Finance 34*, 882–891.

LeBaron, B. and R. Yamamoto (2008). The impact of imitation on long memory in an order-driven market. *Eastern Economic Journal 34*, 504–517.

Lobato, I. N. (1999). A semiparametric two-step estimator in a multivariate long memory model. *Journal of Econometrics 90*, 129–153.

Lobato, I. N. and P. M. Robinson (1998). A nonparametric test for i(0). *The Review of Economic Studies 65*(3), 475–495.

Lobato, I. N. and C. Velasco (2000). Long memory in stock-market trading volume. *Journal of Business and Economic Statistics 18*, 410–427.

Lovcha, Y. and A. Perez-Laborda (2018). Monetary policy shocks, inflation persistence, and long memory. *Journal of Macroeconomics 55*, 117–127.

Luu, J. and M. Martens (2003). Testing the mixture of distributions hypothesis using "realized" volatility. *The Journal of Futures Markets 23*, 61–69.

Marinucci, D. and P. Robinson (2001). Semiparametric fractional cointegration analysis. *Journal of Econometrics 105*(1), 225 – 247.

Nielsen, F. S. (2011). Local whittle estimation of multi-variate fractionally integrated processes. *Journal of Time Series Analysis 32*, 317–335.

Nielsen, M. O. (2004a). Efficient inference in multivariate fractionally integrated time series models. *Econometrics Journal 7*, 63–97.

Nielsen, M. O. (2004b). Efficient likelihood inference in nonstationary univariate models. *Econometric Theory 20*, 116–146.

Nielsen, M. O. (2005). Multivariate lagrange multiplier tests for fractional integration. *Journal of Financial Econometrics 3*, 372–398.

Nielsen, M. O. and K. Shimotsu (2007). Determining the cointegrating rank in nonstationary fractional systems by the exact local whittle approach. *Journal of Econometrics 141*, 574–596.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business 53*, 61–65.

Robinson, P. M. (1994). Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association 89*, 1420–1437.

Robinson, P. M. (1995). Log-periodogram regression of time series with long range dependence. *The Annals of Statistics 23*, 1048–1072.

Robinson, P. M. (2003). *Long memory time series, in P.M. Robinson ed.,* Time Series with Long Memory. Oxford University Press: Oxford.

Robinson, P. M. and Y. Yajima (2002). Determination of cointegrating rank in fractional systems. *Journal of Econometrics 106*(2), 217–241.

Rossi, E. and P. S. de Magistris (2013). Long memory and tail dependence in trading volume and volatility. *Journal of Empirical Finance 22*, 94–112.

Schwert, G. W. (1989). Tests for unit roots: A monte carlo investigation. *Journal of Business and Economic Statistics 7*, 147–160.

Sela, R. J. and C. M. Hurvich (2009). Computationaly efficient methods for two multivariate fractionnaly integrated models. *Journal of Time Series Analysis 30*, 631–651.

Shimotsu, K. (2007). Gaussian semiparametric estimation of multivariate fractionally integrated processes. *Journal of Econometrics 137*, 277–310.

Shimotsu, K. (2010). Exact local whittle estimation of fractional integration with unknown mean and time trend. *Econometric Theory 26*, 501–540.

Shimotsu, K. and P. C. B. Phillips (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics 33*(4), 1890–1933.

Tanaka, K. (1999). The nonstationary fractional unit root. *Econometric Theory 15*, 549–582.

Yamamoto, R. (2011). Order aggressiveness, pre-trade transparency, and long memory in an order-driven market. *Journal of Economic Dynamics and Control 35*, 1938–1963.

Table 1. Empirical rejection frequencies (empirical sizes) for $\theta_1 = \theta_2 = 0$ at the 5% nominal asymptotic significance level of the $LM_{\boldsymbol{d}}^{FGLS}$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ tests, for different values for the contemporaneous correlation parameter ρ, the GARCH parameters (α, β), and sample lengths T. The innovations are drawn from either a multivariate normal distribution or a multivariate Student-t distribution with 5 degrees of freedom.

		Gaussian Innovations						Student-t Innovations					
		$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$
GARCH	ρ	$T = 500$			$T = 1000$			$T = 500$			$T = 1000$		
$\alpha = 0, \beta = 0$	0.0	0.051	0.045	0.050	0.048	0.048	0.048	0.053	0.049	0.048	0.051	0.047	0.048
	0.2	0.050	0.047	0.050	0.047	0.048	0.048	0.056	0.050	0.048	0.051	0.046	0.048
	0.4	0.053	0.050	0.050	0.044	0.045	0.048	0.056	0.050	0.048	0.050	0.043	0.048
	0.6	0.055	0.051	0.050	0.047	0.045	0.048	0.055	0.052	0.048	0.046	0.044	0.048
	0.8	0.055	0.050	0.050	0.048	0.049	0.048	0.056	0.052	0.048	0.047	0.043	0.048
$\alpha = 0.1, \beta = 0.5$	0.0	0.054	0.071	0.065	0.049	0.070	0.067	0.060	0.180	0.163	0.056	0.187	0.183
	0.2	0.053	0.072	0.065	0.048	0.068	0.066	0.060	0.179	0.165	0.056	0.189	0.179
	0.4	0.054	0.072	0.064	0.049	0.068	0.066	0.061	0.181	0.169	0.055	0.192	0.187
	0.6	0.055	0.074	0.065	0.049	0.071	0.067	0.059	0.189	0.175	0.057	0.204	0.197
	0.8	0.056	0.075	0.066	0.052	0.075	0.067	0.064	0.199	0.186	0.059	0.213	0.207
$\alpha = 0.1, \beta = 0.7$	0.0	0.053	0.080	0.070	0.048	0.080	0.071	0.060	0.219	0.196	0.056	0.238	0.228
	0.2	0.054	0.080	0.071	0.047	0.077	0.072	0.060	0.218	0.199	0.055	0.243	0.233
	0.4	0.054	0.080	0.070	0.048	0.079	0.073	0.058	0.221	0.205	0.054	0.247	0.240
	0.6	0.055	0.080	0.073	0.050	0.081	0.075	0.059	0.233	0.221	0.056	0.256	0.256
	0.8	0.056	0.080	0.074	0.054	0.084	0.077	0.063	0.242	0.246	0.060	0.269	0.276
$\alpha = 0.1, \beta = 0.8$	0.0	0.052	0.088	0.076	0.048	0.092	0.081	0.062	0.269	0.249	0.055	0.332	0.316
	0.2	0.052	0.088	0.079	0.047	0.088	0.079	0.059	0.275	0.255	0.053	0.334	0.322
	0.4	0.052	0.089	0.079	0.045	0.089	0.080	0.061	0.281	0.270	0.058	0.332	0.336
	0.6	0.056	0.091	0.082	0.049	0.095	0.085	0.065	0.296	0.294	0.057	0.347	0.362
	0.8	0.055	0.093	0.086	0.051	0.097	0.089	0.069	0.306	0.322	0.059	0.366	0.395
$\alpha = 0.1, \beta = 0.85$	0	0.053	0.103	0.089	0.050	0.113	0.093	0.064	0.337	0.319	0.056	0.475	0.450
	0.2	0.050	0.104	0.089	0.049	0.114	0.094	0.063	0.346	0.327	0.057	0.473	0.447
	0.4	0.054	0.103	0.090	0.048	0.117	0.096	0.066	0.350	0.342	0.056	0.468	0.467
	0.6	0.056	0.107	0.096	0.049	0.124	0.104	0.068	0.353	0.362	0.059	0.473	0.489
	0.8	0.058	0.114	0.105	0.052	0.121	0.111	0.071	0.360	0.388	0.062	0.483	0.528

Table 2. Empirical rejection frequencies at the 5% nominal asymptotic level for $\theta_2 = 0$ and the values $\theta_1 = 0$ (empirical size) and $|\theta_1| > 0$ (empirical power) of the $LM_{\boldsymbol{d}}^{FGLS}$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ tests for the correlation coefficient ρ, and sample length T. The short-run errors in the DGP obey VAR(1) dynamics with on-diagonal coefficients π_1 and π_2 and off-diagonal coefficients $\pi_{12}= \pi_{21}= 0$. The $LM_{\boldsymbol{d}}^{FGLS}$ statistic is computed from an augmented auxiliary regression with one lag of the dependent variable. $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ are computed from VAR(1) residuals. The innovations are drawn from a multivariate Gaussian distribution.

		$\pi_1 = \pi_2 = 0$						$\pi_1 = \pi_2 = 0.4$					
		$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$
θ_1	ρ	$T = 500$			$T = 1000$			$T = 500$			$T = 1000$		
-0.30	0.0	1.000	0.999	0.994	1.000	1.000	1.000	0.875	0.876	0.777	0.997	0.998	0.990
-0.25	0.0	0.987	0.982	0.946	1.000	1.000	0.999	0.720	0.718	0.591	0.967	0.972	0.932
-0.20	0.0	0.899	0.879	0.781	0.999	0.999	0.990	0.496	0.486	0.374	0.835	0.855	0.748
-0.15	0.0	0.644	0.626	0.493	0.941	0.934	0.868	0.289	0.288	0.214	0.558	0.590	0.465
-0.10	0.0	0.338	0.321	0.239	0.609	0.601	0.478	0.154	0.147	0.123	0.271	0.284	0.213
-0.05	0.0	0.118	0.112	0.091	0.183	0.179	0.134	0.074	0.070	0.067	0.099	0.098	0.085
0	0.0	0.057	0.051	0.054	0.052	0.050	0.049	0.053	0.045	0.053	0.050	0.048	0.046
0.05	0.0	0.121	0.112	0.091	0.179	0.174	0.142	0.080	0.079	0.069	0.100	0.097	0.083
0.10	0.0	0.321	0.311	0.243	0.584	0.569	0.472	0.145	0.143	0.119	0.267	0.270	0.205
0.15	0.0	0.600	0.581	0.498	0.888	0.876	0.814	0.272	0.254	0.198	0.479	0.454	0.360
0.20	0.0	0.812	0.788	0.701	0.983	0.978	0.957	0.392	0.321	0.265	0.674	0.597	0.507
0.25	0.0	0.926	0.895	0.831	0.998	0.996	0.988	0.496	0.358	0.296	0.793	0.652	0.576
0.30	0.0	0.969	0.942	0.902	1.000	0.999	0.997	0.574	0.323	0.263	0.859	0.613	0.546
-0.30	0.8	1.000	1.000	1.000	1.000	1.000	1.000	0.998	0.999	0.997	1.000	1.000	1.000
-0.25	0.8	1.000	1.000	1.000	1.000	1.000	1.000	0.985	0.990	0.969	1.000	1.000	1.000
-0.20	0.8	1.000	1.000	0.999	1.000	1.000	1.000	0.908	0.925	0.848	1.000	1.000	1.000
-0.15	0.8	0.983	0.981	0.945	1.000	1.000	0.999	0.678	0.705	0.572	1.000	1.000	0.999
-0.10	0.8	0.757	0.738	0.621	0.920	0.954	0.909	0.350	0.360	0.271	0.920	0.954	0.909
-0.05	0.8	0.243	0.234	0.179	0.369	0.463	0.369	0.124	0.124	0.098	0.369	0.463	0.369
0	0.8	0.058	0.050	0.054	0.053	0.088	0.080	0.057	0.056	0.053	0.053	0.088	0.080
0.05	0.8	0.245	0.233	0.180	0.372	0.448	0.370	0.117	0.125	0.100	0.372	0.448	0.370
0.10	0.8	0.712	0.700	0.596	0.906	0.929	0.884	0.324	0.339	0.253	0.906	0.929	0.884
0.15	0.8	0.951	0.943	0.904	0.996	0.997	0.995	0.567	0.559	0.464	0.996	0.997	0.995
0.20	0.8	0.994	0.993	0.982	1.000	1.000	1.000	0.743	0.692	0.594	1.000	1.000	1.000
0.25	0.8	0.999	0.998	0.997	1.000	1.000	1.000	0.845	0.774	0.691	1.000	1.000	1.000
0.30	0.8	1.000	1.000	0.999	1.000	1.000	1.000	0.898	0.756	0.679	1.000	1.000	1.000

Table 3. Empirical rejection frequencies at the 5% nominal asymptotic level for $\theta_2 = 0$ and $\theta_1 = 0$ (empirical size) and $|\theta_1| > 0$ (empirical power) of the $LM_{\boldsymbol{d}}^{FGLS}$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ tests for the correlation coefficient ρ. The short-run errors follow a VAR(1) model with $\pi_1 = \pi_2 = 0.4$ and GARCH innovations with parameters α and β. The $LM_{\boldsymbol{d}}^{FGLS}$ test statistic is computed from an augmented auxiliary regression with one lag of the dependent variable. $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ are computed from VAR(1) residuals. The sample length is $T = 500$.

$T = 500$		GARCH: $\alpha = 0.1, \beta = 0.8$						GARCH: $\alpha = 0.1, \beta = 0.85$					
		Gaussian			Student-t			Gaussian			Student-t		
θ_1	ρ	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$
-0.30	0.0	0.795	0.859	0.766	0.568	0.840	0.776	0.757	0.855	0.768	0.493	0.836	0.784
-0.25	0.0	0.625	0.718	0.599	0.419	0.724	0.644	0.591	0.716	0.612	0.365	0.730	0.673
-0.20	0.0	0.419	0.512	0.398	0.295	0.579	0.508	0.398	0.521	0.417	0.259	0.615	0.557
-0.15	0.0	0.250	0.328	0.251	0.192	0.434	0.384	0.240	0.341	0.268	0.165	0.490	0.450
-0.10	0.0	0.142	0.182	0.152	0.124	0.310	0.282	0.136	0.199	0.162	0.113	0.385	0.362
-0.05	0.0	0.076	0.102	0.084	0.073	0.218	0.207	0.074	0.121	0.098	0.074	0.294	0.291
0	0.0	0.051	0.072	0.066	0.056	0.186	0.180	0.048	0.083	0.076	0.057	0.252	0.257
0.05	0.0	0.074	0.102	0.085	0.074	0.202	0.192	0.074	0.117	0.097	0.070	0.265	0.272
0.10	0.0	0.136	0.164	0.127	0.118	0.241	0.232	0.131	0.173	0.137	0.105	0.306	0.310
0.15	0.0	0.251	0.266	0.213	0.192	0.321	0.291	0.241	0.277	0.222	0.162	0.360	0.351
0.20	0.0	0.358	0.334	0.278	0.284	0.391	0.350	0.345	0.344	0.281	0.242	0.420	0.409
0.25	0.0	0.461	0.377	0.312	0.377	0.432	0.395	0.449	0.381	0.318	0.310	0.456	0.436
0.30	0.0	0.543	0.346	0.284	0.463	0.418	0.388	0.527	0.353	0.294	0.373	0.447	0.430
-0.30	0.8	0.992	0.998	0.994	0.839	0.977	0.971	0.984	0.997	0.992	0.744	0.961	0.957
-0.25	0.8	0.954	0.981	0.961	0.729	0.931	0.916	0.928	0.973	0.950	0.614	0.902	0.896
-0.20	0.8	0.836	0.901	0.839	0.576	0.837	0.804	0.794	0.892	0.831	0.466	0.810	0.786
-0.15	0.8	0.588	0.693	0.583	0.371	0.658	0.610	0.545	0.686	0.587	0.294	0.642	0.632
-0.10	0.8	0.300	0.378	0.295	0.212	0.461	0.426	0.279	0.390	0.307	0.172	0.478	0.493
-0.05	0.8	0.121	0.159	0.128	0.096	0.271	0.276	0.112	0.176	0.150	0.086	0.325	0.367
0	0.8	0.054	0.077	0.0670	0.061	0.202	0.221	0.055	0.092	0.087	0.066	0.282	0.324
0.05	0.8	0.103	0.152	0.120	0.102	0.249	0.256	0.098	0.165	0.137	0.085	0.302	0.339
0.10	0.8	0.296	0.351	0.277	0.214	0.377	0.371	0.278	0.355	0.288	0.169	0.410	0.424
0.15	0.8	0.519	0.550	0.464	0.363	0.549	0.513	0.492	0.546	0.469	0.282	0.544	0.542
0.20	0.8	0.696	0.689	0.604	0.521	0.658	0.619	0.662	0.680	0.604	0.410	0.615	0.608
0.25	0.8	0.809	0.763	0.689	0.640	0.714	0.694	0.775	0.751	0.682	0.514	0.677	0.682
0.30	0.8	0.870	0.735	0.676	0.719	0.705	0.703	0.846	0.725	0.674	0.586	0.660	0.687

Table 4. Empirical rejection frequencies at the 5% nominal asymptotic level for $\theta_2 = 0$ and $\theta_1 = 0$ (empirical size) and $|\theta_1| > 0$ (empirical power) of the $LM_{\boldsymbol{d}}^{FGLS}$, $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ tests for the correlation coefficient ρ. The short-run errors follow a VAR(1) model with $\pi_1 = \pi_2 = 0.4$ and GARCH innovations with parameters α and β. The $LM_{\boldsymbol{d}}^{FGLS}$ test statistic is computed from an augmented auxiliary regression with one lag of the dependent variable. $LM_{\boldsymbol{d}}^{MLE}$ and $BH_{\boldsymbol{d}}$ are computed from VAR(1) residuals. The sample length is $T = 1000$.

$T = 1000$		GARCH: $\alpha = 0.1, \beta = 0.8$						GARCH: $\alpha = 0.1, \beta = 0.85$					
		Gaussian			Student-t			Gaussian			Student-t		
θ_1	ρ	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$	$LM_{\boldsymbol{d}}^{FGLS}$	$LM_{\boldsymbol{d}}^{MLE}$	$BH_{\boldsymbol{d}}$
-0.30	0.0	0.983	0.993	0.983	0.753	0.961	0.949	0.967	0.991	0.979	0.603	0.947	0.930
-0.25	0.0	0.916	0.959	0.920	0.632	0.915	0.884	0.891	0.956	0.914	0.481	0.904	0.879
-0.20	0.0	0.744	0.841	0.744	0.457	0.811	0.754	0.702	0.836	0.742	0.324	0.817	0.774
-0.15	0.0	0.471	0.597	0.481	0.287	0.645	0.592	0.437	0.603	0.489	0.219	0.688	0.655
-0.10	0.0	0.237	0.316	0.248	0.155	0.457	0.415	0.220	0.332	0.269	0.135	0.550	0.524
-0.05	0.0	0.093	0.132	0.107	0.080	0.302	0.286	0.093	0.150	0.124	0.073	0.446	0.424
0	0.0	0.051	0.072	0.066	0.058	0.243	0.237	0.053	0.089	0.083	0.056	0.397	0.387
0.05	0.0	0.091	0.122	0.098	0.077	0.262	0.247	0.084	0.139	0.112	0.073	0.398	0.391
0.10	0.0	0.246	0.284	0.220	0.176	0.385	0.353	0.226	0.293	0.228	0.131	0.472	0.456
0.15	0.0	0.437	0.461	0.376	0.288	0.508	0.470	0.414	0.464	0.384	0.213	0.558	0.539
0.20	0.0	0.635	0.603	0.523	0.434	0.607	0.557	0.607	0.607	0.531	0.322	0.631	0.596
0.25	0.0	0.756	0.653	0.576	0.588	0.670	0.625	0.729	0.652	0.578	0.433	0.668	0.639
0.30	0.0	0.830	0.620	0.559	0.673	0.664	0.623	0.811	0.621	0.555	0.516	0.659	0.649
-0.30	0.8	1.000	1.000	1.000	0.929	0.998	0.999	1.000	1.000	1.000	0.811	0.991	0.994
-0.25	0.8	1.000	1.000	1.000	0.877	0.999	0.992	0.997	1.000	1.000	0.715	0.974	0.980
-0.20	0.8	0.988	0.998	0.993	0.752	0.965	0.965	0.977	0.995	0.992	0.545	0.926	0.929
-0.15	0.8	0.893	0.951	0.906	0.555	0.876	0.860	0.846	0.940	0.898	0.387	0.833	0.838
-0.10	0.8	0.535	0.658	0.554	0.295	0.670	0.635	0.491	0.655	0.560	0.203	0.665	0.679
-0.05	0.8	0.164	0.228	0.180	0.112	0.386	0.391	0.155	0.254	0.207	0.091	0.488	0.535
0	0.8	0.058	0.083	0.080	0.057	0.271	0.296	0.058	0.104	0.099	0.058	0.404	0.462
0.05	0.8	0.175	0.228	0.180	0.113	0.349	0.352	0.162	0.246	0.196	0.088	0.441	0.487
0.10	0.8	0.510	0.578	0.477	0.307	0.591	0.569	0.471	0.574	0.485	0.206	0.584	0.611
0.15	0.8	0.822	0.848	0.780	0.542	0.771	0.760	0.781	0.839	0.768	0.366	0.711	0.725
0.20	0.8	0.940	0.940	0.908	0.714	0.873	0.860	0.916	0.928	0.897	0.504	0.806	0.816
0.25	0.8	0.977	0.969	0.949	0.819	0.911	0.903	0.964	0.961	0.943	0.632	0.833	0.856
0.30	0.8	0.993	0.958	0.943	0.885	0.900	0.908	0.987	0.946	0.930	0.716	0.840	0.867

Table 5. Results from the implementation of the FGLS multivariate test on log-volume and log-volatility (proxied by log-absolute returns, log-range estimator, and log-realised variance) for each stock in the sample. The columns headed: '$\widehat{d}_{\min}(vlm)$' and '$\widehat{d}_{\min}(\sigma)$' denote the values which minimise the test statistic; '95% CI' denote the bounds of the 95% confidence ellipsoid for the long memory parameter; when non-empty, 'Common d' reports the values along the 45 degree line contained in the 95% confidence ellipsoid (if empty, no common value is detected); '$\widehat{\rho}_e$' reports the sample cross-correlation coefficient of the residuals from the FGLS auxiliary regression given the values $\widehat{d}_{\min}(vlm)$ and $\widehat{d}_{\min}(\sigma)$.

	Log Absolute Returns						Log Range						Log RV					
Stock	$\widehat{d}_{\min}(vlm)$	95% CIB	$\widehat{d}_{\min}(\sigma)$	95% CIB	Common d	$\widehat{\rho}_e$	$\widehat{d}_{\min}(vlm)$	95% CIB	$\widehat{d}_{\min}(\sigma)$	95% CIB	Common d	$\widehat{\rho}_e$	$\widehat{d}_{\min}(vlm)$	95% CIB	$\widehat{d}_{\min}(\sigma)$	95% CIB	Common d	$\widehat{\rho}_e$
AAPL	0.62	[0.52,0.71]	0.38	[0.31,0.45]	-	0.41	0.65	[0.56,0.73]	0.59	[0.52,0.67]	[0.56,0.67]	0.63	0.66	[0.57,0.74]	0.64	[0.55,0.72]	[0.58,0.71]	0.62
AXP	0.49	[0.39,0.58]	0.46	[0.37,0.56]	[0.40,0.55]	0.37	0.53	[0.45,0.62]	0.61	[0.53,0.69]	[0.55,0.60]	0.53	0.53	[0.45,0.60]	0.61	[0.51,0.70]	[0.54,0.56]	0.49
BA	0.34	[0.23,0.44]	0.30	[0.22,0.39]	[0.25,0.39]	0.39	0.41	[0.30,0.50]	0.57	[0.48,0.66]	-	0.52	0.40	[0.30,0.51]	0.58	[0.46,0.69]	-	0.52
CAT	0.42	[0.29,0.54]	0.41	[0.32,0.50]	[0.32,0.50]	0.43	0.45	[0.33,0.56]	0.56	[0.46,0.65]	[0.49,0.55]	0.53	0.48	[0.37,0.59]	0.61	[0.50,0.70]	-	0.58
CSCO	0.35	[0.23,0.45]	0.38	[0.29,0.49]	[0.29,0.44]	0.36	0.39	[0.27,0.49]	0.57	[0.48,0.65]	-	0.48	0.41	[0.31,0.51]	0.58	[0.49,0.66]	-	0.53
CVX	0.52	[0.40,0.63]	0.37	[0.29,0.45]	-	0.30	0.50	[0.41,0.59]	0.53	[0.43,0.64]	[0.44,0.59]	0.47	0.50	[0.41,0.59]	0.57	[0.47,0.68]	[0.47,0.58]	0.43
DD	0.36	[0.22,0.48]	0.40	[0.30,0.50]	[0.30,0.47]	0.32	0.38	[0.26,0.50]	0.55	[0.44,0.64]	-	0.47	0.39	[0.28,0.50]	0.57	[0.46,0.68]	-	0.46
DIS	0.32	[0.18,0.45]	0.40	[0.30,0.52]	[0.31,0.45]	0.37	0.38	[0.26,0.50]	0.62	[0.52,0.71]	-	0.47	0.38	[0.27,0.49]	0.59	[0.49,0.68]	-	0.42
GE	0.47	[0.37,0.58]	0.40	[0.32,0.49]	[0.38,0.48]	0.42	0.52	[0.43,0.61]	0.58	[0.49,0.67]	[0.49,0.61]	0.54	0.52	[0.43,0.61]	0.59	[0.48,0.68]	[0.48,0.59]	0.51
GS	0.52	[0.39,0.65]	0.40	[0.31,0.48]	[0.40,0.44]	0.38	0.57	[0.47,0.68]	0.62	[0.53,0.71]	[0.53,0.68]	0.58	0.55	[0.46,0.65]	0.58	[0.48,0.68]	[0.48,0.65]	0.58
HD	0.54	[0.44,0.63]	0.43	[0.34,0.53]	[0.45,0.52]	0.35	0.57	[0.47,0.65]	0.61	[0.54,0.69]	[0.54,0.65]	0.46	0.56	[0.47,0.63]	0.62	[0.52,0.70]	[0.52,0.63]	0.47
IBM	0.36	[0.24,0.47]	0.30	[0.20,0.41]	[0.24,0.41]	0.38	0.39	[0.29,0.48]	0.54	[0.42,0.64]	-	0.47	0.41	[0.31,0.50]	0.55	[0.44,0.65]	-	0.47
INTC	0.18	[-0.03,0.34]	0.41	[0.30,0.54]	-	0.36	0.28	[0.10,0.42]	0.64	[0.53,0.73]	-	0.46	0.29	[0.12,0.42]	0.60	[0.51,0.69]	-	0.50
JNJ	0.44	[0.32,0.54]	0.35	[0.22,0.51]	[0.32,0.51]	0.36	0.46	[0.37,0.56]	0.58	[0.48,0.67]	-	0.49	0.46	[0.37,0.55]	0.62	[0.50,0.72]	-	0.44
JPM	0.56	[0.45,0.65]	0.47	[0.37,0.56]	[0.47,0.56]	0.38	0.59	[0.52,0.67]	0.62	[0.55,0.69]	[0.55,0.67]	0.56	0.58	[0.51,0.66]	0.61	[0.53,0.70]	[0.53,0.66]	0.56
KO	0.44	[0.30,0.57]	0.38	[0.29,0.47]	[0.32,0.47]	0.38	0.44	[0.30,0.56]	0.58	[0.50,0.66]	-	0.51	0.46	[0.33,0.58]	0.61	[0.52,0.69]	-	0.47
MCD	0.38	[0.25,0.52]	0.44	[0.33,0.56]	[0.33,0.51]	0.33	0.40	[0.29,0.52]	0.63	[0.55,0.71]	-	0.49	0.42	[0.30,0.54]	0.62	[0.53,0.70]	-	0.46
MMM	0.30	[0.11,0.47]	0.36	[0.25,0.46]	[0.28,0.45]	0.37	0.35	[0.20,0.48]	0.52	[0.42,0.61]	-	0.49	0.37	[0.24,0.50]	0.55	[0.45,0.65]	-	0.47
MRK	0.35	[0.17,0.50]	0.35	[0.27,0.44]	[0.27,0.44]	0.39	0.43	[0.24,0.56]	0.55	[0.46,0.64]	[0.48,0.56]	0.50	0.43	[0.28,0.55]	0.55	[0.45,0.64]	[0.48,0.54]	0.50
MSFT	0.16	[-0.11,0.37]	0.43	[0.35,0.54]	-	0.36	0.25	[0.03,0.44]	0.58	[0.47,0.68]	-	0.49	0.25	[0.06,0.43]	0.52	[0.41,0.62]	-	0.50
NKE	0.41	[0.27,0.53]	0.37	[0.28,0.47]	[0.30,0.47]	0.35	0.38	[0.26,0.49]	0.53	[0.44,0.62]	-	0.46	0.39	[0.28,0.50]	0.55	[0.46,0.63]	-	0.44
PFE	0.32	[0.04,0.52]	0.35	[0.26,0.46]	[0.26,0.46]	0.36	0.39	[0.21,0.55]	0.53	[0.42,0.64]	[0.48,0.54]	0.50	0.36	[0.19,0.52]	0.54	[0.42,0.66]	-	0.45
PG	0.38	[0.20,0.53]	0.36	[0.27,0.46]	[0.27,0.45]	0.34	0.40	[0.24,0.53]	0.47	[0.36,0.57]	[0.36,0.53]	0.45	0.39	[0.24,0.52]	0.50	[0.40,0.61]	[0.43,0.52]	0.43
TRV	0.40	[0.24,0.54]	0.45	[0.33,0.59]	[0.34,0.53]	0.27	0.47	[0.34,0.59]	0.66	[0.52,0.80]	-	0.42	0.47	[0.36,0.59]	0.62	[0.50,0.76]	-	0.43
UNH	0.26	[0.06,0.42]	0.35	[0.26,0.44]	[0.27,0.41]	0.37	0.39	[0.20,0.54]	0.59	[0.51,0.67]	-	0.54	0.39	[0.22,0.53]	0.61	[0.52,0.70]	-	0.52
UTX	0.43	[0.30,0.54]	0.37	[0.28,0.46]	[0.31,0.46]	0.30	0.45	[0.34,0.55]	0.57	[0.47,0.65]	-	0.46	0.44	[0.34,0.53]	0.57	[0.48,0.67]	-	0.44
V	0.73	[0.62,0.83]	0.43	[0.33,0.53]	-	0.30	0.71	[0.61,0.81]	0.62	[0.54,0.71]	[0.62,0.69]	0.42	0.73	[0.63,0.83]	0.63	[0.55,0.72]	[0.64,0.70]	0.44
VZ	0.40	[0.24,0.54]	0.41	[0.30,0.53]	[0.31,0.50]	0.32	0.44	[0.27,0.58]	0.60	[0.51,0.69]	-	0.48	0.46	[0.31,0.58]	0.61	[0.50,0.70]	-	0.42
WMT	0.38	[0.24,0.51]	0.37	[0.29,0.47]	[0.29,0.46]	0.40	0.44	[0.32,0.55]	0.54	[0.45,0.63]	[0.49,0.53]	0.50	0.44	[0.30,0.55]	0.57	[0.47,0.66]	-	0.51
XOM	0.48	[0.33,0.60]	0.30	[0.21,0.40]	-	0.30	0.50	[0.38,0.60]	0.53	[0.41,0.64]	[0.41,0.60]	0.47	0.49	[0.37,0.59]	0.52	[0.39,0.64]	[0.39,0.59]	0.46
Average	0.41		0.39			0.36	0.45		0.58			0.49	0.45		0.58			0.48
Rejection rate					20.00%						53.33%						63.33%	

Table 6. Results from the the Nielsen and Shimotsu (2007)-based approach. The columns headed: '$\widehat{d}\,(vlm)$' and '$\widehat{d}\,(\sigma)$' report point estimates from the two-stage univariate exact local Whittle estimator in Shimotsu and Phillips (2005); 'T_0' and 'p-value' report the test statistic for the null of a common order of integration and related p-values; '$L(u)$', $u = 0, 1$, report the objective function used to infer the cointegration rank in a model selection procedure; finally, '$\widehat{r}_T$' reports the estimated cointegration rank resulting from this criterion, and '$\widehat{r}^*_T$' reports the conditional estimates of r for the cases in which the null hypothesis of a common order of integration cannot be rejected at 5% level.

				Nielsen-Shimotsu Testing Approach						
Stock	$\widehat{d}\,(vlm)$	95% CI	$\widehat{d}\,(\sigma)$	95% CI	T_0	p-value	$L(0)$	$L(1)$	$\widehat{r}_T$	$\widehat{r}^*_T$
AAPL	0.62	[0.53,0.71]	0.48	[0.40,0.57]	5.37	0.02	-1.465	-1.337	0	-
AXP	0.52	[0.43,0.61]	0.63	[0.54,0.72]	3.90	0.05	-1.465	-1.444	0	-
BA	0.37	[0.28,0.46]	0.57	[0.48,0.66]	11.41	0.00	-1.465	-1.366	0	-
CAT	0.40	[0.31,0.48]	0.56	[0.47,0.65]	8.74	0.00	-1.465	-1.419	0	-
CSCO	0.34	[0.25,0.42]	0.52	[0.44,0.61]	10.46	0.00	-1.465	-1.416	0	-
CVX	0.52	[0.43,0.61]	0.61	[0.52,0.70]	2.78	0.10	-1.465	-1.455	0	0
DD	0.35	[0.26,0.44]	0.55	[0.46,0.64]	12.08	0.00	-1.465	-1.378	0	-
DIS	0.38	[0.29,0.47]	0.61	[0.52,0.70]	14.55	0.00	-1.465	-1.408	0	-
GE	0.55	[0.46,0.64]	0.60	[0.51,0.69]	0.90	0.34	-1.465	-1.468	1	1
GS	0.53	[0.44,0.62]	0.56	[0.47,0.65]	0.35	0.56	-1.465	-1.482	1	1
HD	0.60	[0.51,0.69]	0.57	[0.48,0.66]	0.33	0.57	-1.465	-1.380	0	0
IBM	0.39	[0.31,0.48]	0.58	[0.49,0.66]	10.83	0.00	-1.465	-1.460	0	-
INTC	0.19	[0.10,0.28]	0.54	[0.45,0.63]	33.64	0.00	-1.465	-1.301	0	-
JNJ	0.41	[0.33,0.50]	0.60	[0.51,0.69]	10.86	0.00	-1.465	-1.452	0	-
JPM	0.59	[0.50,0.68]	0.62	[0.53,0.71]	0.31	0.58	-1.465	-1.530	1	1
KO	0.40	[0.31,0.49]	0.61	[0.52,0.70]	13.03	0.00	-1.465	-1.350	0	-
MCD	0.35	[0.26,0.44]	0.60	[0.51,0.69]	16.74	0.00	-1.465	-1.258	0	-
MMM	0.36	[0.27,0.45]	0.54	[0.45,0.63]	10.87	0.00	-1.465	-1.447	0	-
MRK	0.33	[0.24,0.41]	0.52	[0.43,0.61]	11.35	0.00	-1.465	-1.368	0	-
MSFT	0.22	[0.14,0.31]	0.48	[0.39,0.56]	16.50	0.00	-1.465	-1.351	0	-
NKE	0.39	[0.30,0.48]	0.52	[0.43,0.60]	4.45	0.03	-1.465	-1.326	0	-
PFE	0.37	[0.28,0.46]	0.53	[0.44,0.62]	7.52	0.01	-1.465	-1.332	0	-
PG	0.36	[0.27,0.45]	0.48	[0.39,0.57]	3.89	0.05	-1.465	-1.323	0	-
TRV	0.41	[0.30,0.51]	0.59	[0.49,0.69]	7.34	0.01	-1.426	-1.399	0	-
UNH	0.33	[0.24,0.41]	0.56	[0.47,0.65]	16.39	0.00	-1.465	-1.338	0	-
UTX	0.41	[0.32,0.50]	0.58	[0.49,0.67]	9.01	0.00	-1.465	-1.420	0	-
VZ	0.39	[0.30,0.48]	0.57	[0.48,0.66]	8.22	0.00	-1.465	-1.248	0	-
V	0.64	[0.55,0.74]	0.62	[0.53,0.72]	0.11	0.73	-1.452	-1.305	0	0
WMT	0.40	[0.32,0.49]	0.57	[0.48,0.66]	7.82	0.01	-1.465	-1.392	0	-
XOM	0.47	[0.38,0.56]	0.53	[0.44,0.62]	1.06	0.30	-1.465	-1.493	1	1
Average	0.42		0.56						13.33%	13.33%
Rejection 95%							76.67%			

Supplementary Appendix

to

Multivariate Fractional Integration Tests allowing for Conditional Heteroskedasticity with an Application to Return Volatility and Trading Volume

by

M. Balboa, P.M.M. Rodrigues, A. Rubia and A.M. Robert Taylor

Date: January 18, 2021

Table of Contents

Summary of Contents

This supplement to our paper "Multivariate Fractional Integration Tests allowing for Conditional Heteroskedasticity with an Application to Return Volatility and Trading Volume" has three main parts. The first part, Appendix A, contains a number of preparatory lemmas and their proofs which are used to prove the main results, together with proofs of Theorems 1 and 2 in the paper. The second part, Appendix B, contains additional Monte Carlo results. The third part, Appendix C, includes additional data analysis related to the empirical application in section 5 of the paper.

Equation references (A.n), (B.n) and (C.n) for $n \geq 1$ refer to equations in in Appendices A, B and C, respectively, of this supplementary appendix. Other equation references are to the main paper. Additional references are included at the end of section A.2 of the supplement.

Appendix A - Technical Appendix

A.1 Preliminary Results

Before presenting the proofs of the main results in the paper, we first need to state and prove some preparatory Lemmas. To this end, consider the following additional notation. For an $(n \times 1)$ vector $\boldsymbol{A}$, $||\boldsymbol{A}||$ denotes the Euclidean vector norm, such that $||\boldsymbol{A}||^2 = \boldsymbol{A}'\boldsymbol{A}$. For an $(n \times m)$ matrix $\boldsymbol{A}$, $||\boldsymbol{A}||$ denotes the Euclidean matrix norm, $||\boldsymbol{A}||^2 = tr\left(\boldsymbol{A}'\boldsymbol{A}\right)$. The constants K and C are used throughout the proofs to refer to some generic strictly positive constant which does not depend on the sample size. The notation $\overset{a.s.}{\to}$ denotes almost surely convergence of a random sequence as the sample length is allowed to diverge to $+\infty$. Finally, throughout the proofs, we will use the superscripts '$*$' and '$**$' to denote truncated processes and their non-truncated counterpart, respectively, such as the truncated and non-truncated processes $z^{*}_{s,t-1,d_s} := \sum_{l=1}^{t-1} l^{-1}\varepsilon_{s,t-l,d_s}$ and $z^{**}_{s,t-1,d_s} := \sum_{l=1}^{\infty} l^{-1}\varepsilon_{s,t-l,d_s}$, respectively. This distinction is necessary because, while the statistics computed in the paper are constructed from the truncated variables, the asymptotic theory is developed with respect to the corresponding non-truncated processes. Lemmas A2 and A3 below show that the distinction between the two is asymptotically negligible so far as the limiting distribution theory for the statistics considered in this paper are concerned.

Lemma A1. *Let* $\boldsymbol{x}^{**}_{s,t-1,\boldsymbol{d}} := \left(z^{**}_{s,t-1,d_s}, \boldsymbol{\varepsilon}'_{t-1,\boldsymbol{d}}, ..., \boldsymbol{\varepsilon}'_{t-p,\boldsymbol{d}}\right)'$, *with* $z^{**}_{s,t-1,d_s} := \sum_{l=1}^{\infty} l^{-1}\varepsilon_{s,t-l,d_s}$, *and* ε_{s,t,d_s} *denoting the s-th element of* $\boldsymbol{\varepsilon}_{t,\boldsymbol{d}}$, $1 \leq s \leq k$. *Let* ν_{ij} *be the (i,j)-th element of* $\boldsymbol{\Sigma}^{-1}$. *Under Assumption 1 and* $\mathrm{H}_0 : \boldsymbol{\theta} = 0$, $\nu_{ij}\left(\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}}\right)/T \overset{a.s.}{\to} \boldsymbol{\Omega}_{\boldsymbol{A}ij}$, *with* $\boldsymbol{\Omega}_{\boldsymbol{A}ij} := \nu_{ij}E\left(\boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}}\boldsymbol{x}'^{**}_{j,t-1,\boldsymbol{d}}\right)$ *bounded, and bounded away from zero if* $\nu_{ij} \neq 0$.

Proof of Lemma A1. Denote $\boldsymbol{\lambda}_{t-1,\boldsymbol{d}} := \left(\boldsymbol{\varepsilon}'_{t-1,\boldsymbol{d}}, ..., \boldsymbol{\varepsilon}'_{t-p,\boldsymbol{d}}\right)'$, such that for any $1 \leq i, j \leq k$ we have

$$T^{-1}\, \boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}} = \left[\begin{array}{cc} T^{-1}\sum_{t=p^*+1}^{T} z^{**}_{i,t-1,d_i}z^{**}_{j,t-1,d_j} & T^{-1}\sum_{t=p^*+1}^{T} z^{**}_{i,t-1,d_i}\boldsymbol{\lambda}'_{t-1,\boldsymbol{d}} \\ T^{-1}\sum_{t=p^*+1}^{T} \boldsymbol{\lambda}_{t-1,\boldsymbol{d}}z^{**}_{j,t-1,d_j} & T^{-1}\sum_{t=p^*+1}^{T} \boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\boldsymbol{\lambda}'_{t-1,\boldsymbol{d}} \end{array}\right].$$

Under Assumption 1, $\mathrm{H}_0 : \boldsymbol{\theta} = 0$, and for any $1 \leq s \leq k$, $\left\{\boldsymbol{X}^{**}_{s,-1,\boldsymbol{d}}\right\}$ is a measurable function of a strictly stationary and ergodic process and is therefore also a strictly stationary and ergodic process, and so is $\left\{\boldsymbol{X}'^{**}_{s,-1,\boldsymbol{d}}\boldsymbol{X}^{**}_{s,-1,\boldsymbol{d}}\right\}$. The required result then follows from the Ergodic Theorem because $\boldsymbol{x}^{**}_{s,t-1,\boldsymbol{d}}$ is (uniformly) L_2-bounded, so the elements in $\left\{\boldsymbol{x}^{**}_{s,t-1,\boldsymbol{d}}\boldsymbol{x}'^{**}_{s,t-1,\boldsymbol{d}}\right\}$ have finite absolute expected values. To see this, first note that for any $1 \leq s \leq k$, there exists some finite $K > 0$ such that $E\left(z^{2**}_{s,t-1,d_s}\right) = \sum_{l=1}^{\infty}\omega^2_{sl}E\left(e^2_{s,t-l}\right) < K$, because $\omega_{sl} = O\left(1/l\right)$ and $\{\boldsymbol{e}_t\}$ is uniformly L_2-bounded under Assumption (A2). From this result, it follows from the Cauchy-Swcharz inequality that $E\left(\left|z^{**}_{i,t-1,d_i}z^{**}_{j,t-1,d_j}\right|\right) \leq$

$\sqrt{E\left(z^{2**}_{i,t-1,d_i}\right)}\sqrt{E\left(z^{2**}_{j,t-1,d_j}\right)} < K$, $1 \le i,j \le k$. Similarly, because $\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}$ is uniformly L_2-bounded under Assumption (A2), there exists some finite $C > 0$ for which $E\left\|\boldsymbol{\lambda}'_{t-1,\boldsymbol{d}} z^{**}_{s,t-1,d_s}\right\| \le \sqrt{E||\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}||^2}\sqrt{E\left(z^{2**}_{s,t-1,d_s}\right)} < C$ and $E\left\|\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\boldsymbol{\lambda}'_{t-1,\boldsymbol{d}}\right\| \le E\left\|\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\right\|^2 < C$. Consequently, the Ergodic Theorem ensures that $\nu_{ij}\left(\boldsymbol{X}'^{*}_{i,-1,\boldsymbol{d}}\boldsymbol{X}^{*}_{j,-1,\boldsymbol{d}}\right)/T \overset{a.s.}{\to} \nu_{ij}E\left(\boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}}\boldsymbol{x}'^{**}_{j,t-1,\boldsymbol{d}}\right)$. Finally, due to stationarity, $\boldsymbol{x}^{**}_{s,t-1,\boldsymbol{d}} = \sum_{l=1}^{\infty}\boldsymbol{\Gamma}_{sl}\boldsymbol{e}_{j-l}$ with $||\boldsymbol{\Gamma}_{sl}|| = O(1/l)$, so $\boldsymbol{\Omega}_{\boldsymbol{A}ij} = \nu_{ij}\sum_{l=1}^{\infty}\boldsymbol{\Gamma}_{il}\ \boldsymbol{\Sigma}\ \boldsymbol{\Gamma}'_{jl} < \infty$. Clearly, the condition $\boldsymbol{\Sigma} > \boldsymbol{0}$ rules out the degenerate case $E\left(\boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}}\boldsymbol{x}'^{**}_{i,t-1,\boldsymbol{d}}\right) = 0$, from which the required results follow. Furthermore, for $i = j$, $\boldsymbol{\Omega}_{\boldsymbol{A}ij} = \nu_{ii}\sum_{l=1}^{\infty}\boldsymbol{\Gamma}_{il}\ \boldsymbol{\Sigma}\ \boldsymbol{\Gamma}'_{il}$, and so $\boldsymbol{\Omega}_{\boldsymbol{A}ij}$ is positive definite (Davidson, 2000, Corollary 14.2.10, p.216). ■

Lemma A2. *Under Assumption 1 and* $\mathrm{H}_0 : \boldsymbol{\theta} = \boldsymbol{0}$, *for* $1 \le i,j \le k$, *it follows that,*
i) $T^{-1}\left\|\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}} - \boldsymbol{X}'^{*}_{i,-1,\boldsymbol{d}}\boldsymbol{X}^{*}_{j,-1,\boldsymbol{d}}\right\| = O_p\left(T^{-1/2}\right)$;
ii) $T^{-1/2}\left\|\left(\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}} - \boldsymbol{X}'^{*}_{i,-1,\boldsymbol{d}}\right)\boldsymbol{u}_j\right\| = O_p\left(\frac{\sqrt{\log T}}{\sqrt{T}}\right)$.

Proof of Lemma A2. For i), we can write $T^{-1}\left(\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}} - \boldsymbol{X}'^{*}_{i,-1,\boldsymbol{d}}\boldsymbol{X}^{*}_{j,-1,\boldsymbol{d}}\right)$ as

$$\left[\begin{array}{cc} T^{-1}\sum_{t=p^*+1}^{T}\left(z^{**}_{i,t-1,d_i}z^{**}_{j,t-1,d_j} - z^{*}_{i,t-1,d_i}z^{*}_{j,t-1,d_j}\right) & T^{-1}\sum_{t=p^*+1}^{T}\boldsymbol{\lambda}'_{t-1,\boldsymbol{d}}\left(z^{**}_{it-1,d_i} - z^{*}_{it-1,d_i}\right) \\ T^{-1}\sum_{t=p^*+1}^{T}\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\left(z^{**}_{j,t-1,d_j} - z^{*}_{j,t-1,d_j}\right) & \mathbf{0}_{kp\times kp} \end{array}\right].$$

Because $E\left(z^{**}_{s,t-1,d_s} - z^{*}_{s,t-1,d_s}\right)^2 = O(1/t)$ (cf. Demetrescu *et al.* 2008, Lemma B.1.), we have from the Cauchy-Schwarz inequality that,

$$\begin{aligned} E\left\|T^{-1}\sum_{t=p^*+1}^{T}\boldsymbol{\lambda}'_{t-1,\mathbf{d}}\left(z^{**}_{s,t-1,d_s} - z^{*}_{s,t-1,d_s}\right)\right\| &\le T^{-1}\sum_{t=p^*+1}^{T}\sqrt{E\left(||\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}||^2\right)}\sqrt{E\left(|z^{**}_{s,t-1,d_s} - z^{*}_{s,t-1,d_s}|^2\right)} \\ &= O\left(T^{-1}\sum_{t=p^*+1}^{T}\frac{1}{\sqrt{t}}\right) = O\left(T^{-1/2}\right). \end{aligned}$$

Hence, $T^{-1}\sum_{t=p^*+1}^{T}\boldsymbol{\lambda}'_{t-1,\boldsymbol{d}}\left(z^{**}_{it-1,d_i} - z^{*}_{it-1,d_i}\right) = O_p\left(T^{-1/2}\right)$ and $T^{-1}\sum_{t=p^*+1}^{T}\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\left(z^{**}_{j,t-1,d_j} - z^{*}_{j,t-1,d_j}\right)$ $= O_p\left(T^{-1/2}\right)$ by the Markov inequality. Next, write $z^{**}_{s,t-1,d_s} = z^{*}_{s,t-1,d_s} + b_{s,t-1}$, with $b_{s,t-1} := \sum_{l=t}^{\infty}\omega_{sl}e_{s,t-l}$. Because $\omega_{sl} = O(1/l)$ and $b_{s,t-1} = O_p\left(1/\sqrt{t}\right)$, it follows that,

$$z^{**}_{i,t-1,d_i}z^{**}_{j,t-1,d_j} = \left(z^{*}_{i,t-1,d_i} + b_{i,t-1}\right)\left(z^{*}_{j,t-1,d_j} + b_{j,t-1}\right) = z^{*}_{i,t-1,d_i}z^{*}_{j,t-1,d_j} + r_{ij,t-1}$$

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

$$E\left\|T^{-1}\sum_{t=p^*+1}^{T}\left(z^{**}_{i,t-1,d_i}z^{**}_{j,t-1,d_j} - z^{*}_{i,t-1,d_i}z^{*}_{j,t-1,d_j}\right)\right\| \le T^{-1}\sum_{t=p^*+1}^{T}E|r_{ij,t-1}| = O\left(T^{-1/2}\right) = o(1)$$

and the required result holds from the Markov inequality. For part b), note that the first element of the column vector $\left(\boldsymbol{X}'^{**}_{i,-1,d_i} - \boldsymbol{X}'^{*}_{i,-1,d_i}\right)\boldsymbol{u}_j$ is given by $T^{-1/2}\sum_{t=p^*+1}^{T} b_{i,t-1}e_{j,t}$, while all of the remaining elements are zero. Owing to the MDS property of $\{\boldsymbol{e}_t\}$ and the stationarity condition in Assumption (A3), together with the moment conditions in Assumptions (A2) and (A4) it follows that,

$$
\begin{aligned}
E\left(T^{-1/2}\sum_{t=p^*+1}^{T} b_{i,t-1}e_{j,t}\right)^2 &= T^{-1}\sum_{t=p^*+1}^{T} E\left(b_{i,t-1}^2 e_{j,t}^2\right) \\
&= T^{-1}\sum_{t=p^*+1}^{T}\sum_{l_1=t}^{\infty}\sum_{l_2=t}^{\infty}\omega_{il_1}\omega_{il_2}E\left(e_{j,t}^2 e_{i,t-l_1}e_{i,t-l_2}\right) = O\left(\frac{\log T}{T}\right)
\end{aligned}
$$

because

$$
\begin{aligned}
\sum_{l_1=t}^{\infty}\sum_{l_2=t}^{\infty}\omega_{il_1}\omega_{il_2}E\left(e_{j,t}^2 e_{i,t-l_1}e_{i,t-l_2}\right) &= \sum_{l_1=t}^{\infty}\omega_{il_1}^2 E\left(e_{j,t}^2 e_{i,t-l_1}^2\right) + \sum_{l_1=t}^{\infty}\sum_{\substack{l_2=t \\ l_2\neq l_1}}^{\infty}\omega_{il_1}\omega_{il_2}E\left(e_{j,t}^2 e_{i,t-l_1}e_{i,t-l_2}\right) \\
&= O\left(\sum_{l_1=t}^{\infty}\frac{1}{l_1^2}\right) + o\left(\sum_{\substack{l_1=t,l_2=t \\ l_2\neq l_1}}\frac{1}{l_1^2 l_2^2}\right) = O\left(1/t\right)
\end{aligned}
$$

given that $E\left(e_{j,t}^2 e_{i,t-l_1}^2\right) \leq \left(E\left(e_{j,t}^4\right)E\left(e_{i,t}^4\right)\right)^{1/2} < K$ for all t, and Assumption (A4) which implies that $E\left(e_{j,t}^2 e_{i,t-l_1}e_{i,t-l_2}\right) \leq E\left(|e_{j,t}e_{jt}e_{i,t-l_1}e_{i,t-l_2}|\right) = o\left(\frac{1}{l_1 l_2}\right)$ 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 then holds from the Markov inequality. ■

Lemma A3. *Let* $\widetilde{\boldsymbol{\Sigma}} = \{\widetilde{\sigma}_{ij}\}$ *denote the OLS estimator of* $\boldsymbol{\Sigma} = \{\sigma_{ij}\}$*, namely,* $\widetilde{\sigma}_{ij} = T^{-1}\widetilde{\boldsymbol{u}}_i'\widetilde{\boldsymbol{u}}_j$*,* $\widetilde{\boldsymbol{u}}_s := \boldsymbol{Y}_{s,d_s} - \boldsymbol{X}^*_{s,-1,\boldsymbol{d}}\widetilde{\boldsymbol{\beta}}_s$*, with* $\widetilde{\boldsymbol{\beta}}_s$ *denoting the OLS estimator of* $\boldsymbol{\beta}_s$ *in the corresponding equation. Then, under Assumption 1 and* $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$,
i) $\widetilde{\boldsymbol{\Sigma}} \xrightarrow{p} \boldsymbol{\Sigma}$;
ii) $T^{-1}\left\|\boldsymbol{X}'^{**}_{-1,\boldsymbol{d}}\left[\left(\widetilde{\boldsymbol{\Sigma}}^{-1} - \boldsymbol{\Sigma}^{-1}\right)\otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^{**}_{-1,\boldsymbol{d}}\right\| = O_p\left(T^{-1/2}\right)$;
iii) $T^{-1/2}\boldsymbol{X}'^{**}_{-1,\boldsymbol{d}}\left[\left(\widetilde{\boldsymbol{\Sigma}}^{-1} - \boldsymbol{\Sigma}^{-1}\right)\otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u} \xrightarrow{p} 0$.

Proof of Lemma A3. Part i) follows from the consistency of the equation-by-equation OLS estimator, $\widetilde{\boldsymbol{\beta}}_s$, under Assumption 1 and $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, which can be proved along the same lines as in Demetrescu *et al.* (2008). Part ii) follows from $\sqrt{T}$-consistency in i)

because $\boldsymbol{X}^{**}_{-1,\boldsymbol{d}}$ is uniformly L_2-bounded. Finally, for part *iii*), note that

$$T^{-1/2}\boldsymbol{X}'^{**}_{-1,\boldsymbol{d}}\left[\left(\widetilde{\boldsymbol{\Sigma}}^{-1}-\boldsymbol{\Sigma}^{-1}\right)\otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}=\begin{bmatrix} \sum_{s=1}^{k} T^{-1/2}\left(\widetilde{\nu}_{1s}-\nu_{1s}\right)\left(\frac{1}{T}\boldsymbol{X}'^{**}_{1,-1,\boldsymbol{d}}\boldsymbol{u}_s\right)\\ \sum_{s=1}^{k} T^{-1/2}\left(\widetilde{\nu}_{2s}-\nu_{2s}\right)\left(\frac{1}{T}\boldsymbol{X}'^{**}_{2,-1,\boldsymbol{d}}\boldsymbol{u}_s\right)\\ \vdots \\ \sum_{s=1}^{k} T^{-1/2}\left(\widetilde{\nu}_{ks}-\nu_{ks}\right)\left(\frac{1}{T}\boldsymbol{X}'^{**}_{k,-1,\boldsymbol{d}}\boldsymbol{u}_s\right)\end{bmatrix}$$

and the required result follows noting that $T^{-1/2}\left(\widetilde{\nu}_{ij}-\nu_{ij}\right)=O_p(1)$ for all $1\leq i,j\leq k$ from *i*) above, while $\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{u}_j/T\overset{a.s.}{\to}\boldsymbol{0}$ from the Ergodic Theorem, because $\left\{\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{u}_s\right\}$ is a strictly stationary and ergodic vector MDS, and $E||\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{u}_s||\leq\sqrt{E||\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}||^2E||\boldsymbol{u}_s||^2}<\infty$ under Assumption 1 and $\mathrm{H}_0:\boldsymbol{\theta}=\boldsymbol{0}$. ■

Lemma A4. *Define* $\boldsymbol{D}_{rsij}:=e_{r,t}e_{s,t}\boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}}\boldsymbol{x}'^{**}_{j,t-1,\boldsymbol{d}}$, *for all* $1\leq r,s,i,j\leq k$. *Under Assumption 1 and* $\mathrm{H}_0:\boldsymbol{\theta}=\boldsymbol{0}$, $E||\boldsymbol{D}_{rsij}||<\infty$.

Proof of Lemma A4. The proof follows from the Cauchy-Schwarz inequality given that $e_{i,t}\boldsymbol{x}^{**}_{j,t-1,\boldsymbol{d}}$ is (uniformly) L_2-bounded for any $1\leq i,j\leq k$.

$$E\left\|e_{i,t}\boldsymbol{x}^{**}_{j,t-1,\boldsymbol{d}}\right\|^2=E\left\|e_{i,t}\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\right\|^2+E\left(e^2_{i,t}z^{2**}_{j,t-1,d_j}\right)<K$$

because for some finite $C>0$,

$$E\left\|e_{i,t}\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\right\|\leq\sqrt{E\left(e^2_{i,t}\right)E\left\|\boldsymbol{\lambda}_{t-1,\boldsymbol{d}}\right\|^2}<C$$

and

$$\begin{aligned}E\left(e^2_{i,t}z^{2**}_{j,t-1,d_j}\right)&=\sum_{l_1=1}^{\infty}\sum_{l_2=1}^{\infty}\omega_{jl_1}\omega_{jl_2}E\left(e^2_{i,t}e_{j,t-l_1}e_{j,t-l_2}\right)\\&=\sum_{l_1=1}^{\infty}\omega^2_{jl_1}E\left(e^2_{i,t}e_{j,t-l_1}e_{j,t-l_1}\right)+\sum_{l_1=1}^{\infty}\sum_{\substack{l_2=1\\ l_2\neq l_1}}^{\infty}\omega_{jl_1}\omega_{jl_2}E\left(e^2_{i,t}e_{j,t-l_1}e_{j,t-l_2}\right)\end{aligned}$$

where $E\left(e^2_{i,t}e_{j,t-l_1}e_{j,t-l_1}\right)\leq E(e^4_{i,t})^{1/4}\times E\left(e^4_{j,t}\right)^{3/4}<K$ and, as in Lemma A2,

$$\begin{aligned}\sum_{l_1=1}^{\infty}\sum_{\substack{l_2=1\\ l_2\neq l_1}}^{\infty}\omega_{jl_1}\omega_{jl_2}E\left(e^2_{i,t}e_{j,t-l_1}e_{j,t-l_2}\right)&=O\left(\sum_{l_1=t}^{\infty}\frac{1}{l_1^2}\right)+o\left(\sum_{\substack{l_1=1,l_2=1\\ l_2\neq l_1}}^{\infty}\frac{1}{l_1^2l_2^2}\right)\\&=O(1).\end{aligned}$$

Consequently, $E\left\|e_{r,t}e_{s,t}\boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}}\boldsymbol{x}'^{**}_{j,t-1,\boldsymbol{d}}\right\|\leq\sqrt{E\left\|e_{r,t}\boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}}\right\|^2}\sqrt{E\left\|e_{r,t}\boldsymbol{x}^{**}_{j,t-1,\boldsymbol{d}}\right\|^2}<\infty$. ■

A.2 Proofs of Main Results

Proof of Theorem 1. Under Assumption 1 and $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, the FGLS estimator of $\boldsymbol{\beta}$ can be written as

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + \left(\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}} \left[\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{I}_{T-p^*} \right] \boldsymbol{X}^{*}_{-1,\boldsymbol{d}} \right)^{-1} \left(\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}} \left[\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{I}_{T-p^*} \right] \boldsymbol{u} \right).$$

Using Lemmas A2 and A3, we therefore have that,

$$\sqrt{T} \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \right) = \left(\frac{1}{T} \boldsymbol{X}'^{**}_{t-1,\boldsymbol{d}} \left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*} \right] \boldsymbol{X}^{**}_{-1,\boldsymbol{d}} \right)^{-1} \left(\frac{1}{\sqrt{T}} \boldsymbol{X}'^{**}_{-1,\boldsymbol{d}} \left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*} \right] \boldsymbol{u} \right) + o_p(1). \tag{A.1}$$

Recall that ν_{ij} denotes the (i,j)-th element of $\boldsymbol{\Sigma}^{-1}$, and define $\boldsymbol{A}^{**}_T := \frac{1}{T} \boldsymbol{X}'^{**}_{-1,\boldsymbol{d}} \left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*} \right] \boldsymbol{X}^{**}_{-1,\boldsymbol{d}}$, noting that $\boldsymbol{A}^{**}_T$ can be represented as a partitioned matrix with ij-block $\boldsymbol{A}^{**}_{Tij} = \nu_{ij} \boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}} \boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}} / T$. From Lemma A2, $\nu_{ij} \boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}} \boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}} / T \overset{a.s.}{\to} \boldsymbol{\Omega}_{\boldsymbol{A}ij}$ for all $1 \leq i,j \leq k$. Consequently, $\boldsymbol{A}^{**}_T \overset{a.s.}{\to} \boldsymbol{A}_{\boldsymbol{\beta}}$, where $\boldsymbol{A}_{\boldsymbol{\beta}}$ is a partitioned matrix with ij-th submatrix given by $\boldsymbol{\Omega}_{\boldsymbol{A}ij}$. Noting that the columns of $\boldsymbol{A}_{\boldsymbol{\beta}}$ cannot be written as linear combinations of the other elements, $\det(\boldsymbol{A}_{\boldsymbol{\beta}}) > 0$, and consequently

$$\left(\frac{1}{T} \boldsymbol{X}'^{**}_{-1,\boldsymbol{d}} \left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*} \right] \boldsymbol{X}^{**}_{-1,\boldsymbol{d}} \right)^{-1} \overset{a.s.}{\to} \boldsymbol{A}^{-1}_{\boldsymbol{\beta}}$$

by Slutsky's Theorem. We now discuss the asymptotic behaviour of the second term in (A.1). To this end, define the column vector $\boldsymbol{w}^{**}_{-1,\boldsymbol{d}} := \boldsymbol{X}'^{**}_{-1,\boldsymbol{d}} \left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*} \right] \boldsymbol{u}$, noting that

$$\boldsymbol{w}^{**}_{-1,\boldsymbol{d}} = \left(\sum_{s=1}^{k} \nu_{1s} \boldsymbol{X}'^{**}_{1,-1,\boldsymbol{d}} \boldsymbol{u}_s, \ \sum_{s=1}^{k} \nu_{2s} \boldsymbol{X}'^{**}_{2,-1,\boldsymbol{d}} \boldsymbol{u}_s, ..., \sum_{s=1}^{k} \nu_{ks} \boldsymbol{X}'^{**}_{k,-1,\boldsymbol{d}} \boldsymbol{u}_s \right)'$$

with $\boldsymbol{u}_s := (e_{s,p+1}, ..., e_{s,T})'$ under $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$. Given that $E\left(\boldsymbol{w}^{**}_{-1,\boldsymbol{d}} | \mathcal{F}_{t-1} \right) = 0$ and $\boldsymbol{w}^{**}_{-1,\boldsymbol{d}}$ is a measurable function of $\{\boldsymbol{e}_t\}$, $\left\{ \boldsymbol{w}^{**}_{-1,\boldsymbol{d}}, \mathcal{F}_t \right\}$ is a strictly stationary and ergodic vector MDS. The covariance matrix of $\boldsymbol{w}^{**}_{-1,\boldsymbol{d}}$ is $\boldsymbol{B}_{\boldsymbol{\beta}} := E\left(\boldsymbol{w}^{**}_{-1,\boldsymbol{d}} \boldsymbol{w}'^{**}_{-1,\boldsymbol{d}} \right)$, which can be represented as a partitioned matrix with ij-th block $\boldsymbol{\Omega}_{\boldsymbol{B}ij}$ given by

$$\boldsymbol{\Omega}_{\boldsymbol{B}ij} := \sum_{r=1}^{k} \sum_{s=1}^{k} \nu_{ir} \nu_{is} E\left(e_{r,t} e_{s,t} \boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}} \boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}} \right).$$

From Lemma A4, $E||e_{r,t} e_{s,t} \boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}} \boldsymbol{x}'^{**}_{j,t-1,\boldsymbol{d}}|| < \infty$, and consequently $\boldsymbol{\Omega}_{\boldsymbol{B}ij} < \infty$ for all $1 \leq i,j \leq k$, so $\boldsymbol{B}_{\boldsymbol{\beta}} < \infty$. Furthermore, the condition that $\boldsymbol{\Sigma}$ is positive definite trivially rules out the degenerate case $||\boldsymbol{\Omega}_{\boldsymbol{B}ii}|| = 0$, and so $\boldsymbol{B}_{\boldsymbol{\beta}}$ is bounded away from zero. Consequently, the Central Limit Theorem (CLT) for MDS (Davidson, 2000, Theorem 24.3) and the Cramér-Wold device deliver the result that

$$\frac{1}{\sqrt{T}}\boldsymbol{X}'^{**}_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}\Rightarrow\mathcal{N}\left(\mathbf{0},\boldsymbol{B}_{\boldsymbol{\beta}}\right),$$

and so we may conclude that,

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\right)\Rightarrow\mathcal{N}\left(\mathbf{0},\boldsymbol{A}_{\boldsymbol{\beta}}^{-1}\boldsymbol{B}_{\boldsymbol{\beta}}\boldsymbol{A}_{\boldsymbol{\beta}}^{-1}\right)$$

as required. ■

Proof of Theorem 2. We first prove the stated convergence result under the null hypothesis, which follows directly from Theorem 1 and the consistency of $\widehat{\boldsymbol{\Omega}}_{\boldsymbol{\beta}}:=\boldsymbol{A}_T^{*-1}\boldsymbol{B}_T^{*}\boldsymbol{A}_T^{*-1}$, with $\boldsymbol{A}_T^{*}:=\frac{1}{T}\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes\boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^{*}_{-1,\boldsymbol{d}}$ and $\boldsymbol{B}_T^{*}:=\frac{1}{T}\widehat{\boldsymbol{w}}^{*}_{-1,\boldsymbol{d}}\widehat{\boldsymbol{w}}'^{*}_{-1,\boldsymbol{d}}$, $\widehat{\boldsymbol{w}}^{*}_{-1,\boldsymbol{d}}:=\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes\boldsymbol{I}_{T-p^*}\right]\widehat{\boldsymbol{u}}$, $\widehat{\boldsymbol{u}}:=\boldsymbol{Y}_{t,\boldsymbol{d}}-\boldsymbol{X}^{*}_{-1,\boldsymbol{d}}\widehat{\boldsymbol{\beta}}$. As discussed previously, $\boldsymbol{A}_T^{**}\overset{a.s.}{\to}\boldsymbol{A}_{\boldsymbol{\beta}}$, and so Lemma A2 and the Asymptotic Equivalence Lemma (AEL) allow us to conclude that $\boldsymbol{A}_T^{*}\overset{p}{\to}\boldsymbol{A}_{\boldsymbol{\beta}}$. Using the $\sqrt{T}$-consistency result from Theorem 1, $\widehat{\boldsymbol{u}}:=\boldsymbol{u}+O_p\left(T^{-1/2}\right)$ and, therefore, $\widehat{\boldsymbol{w}}^{*}_{-1,\boldsymbol{d}}:=\boldsymbol{X}'^{*}_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1}\otimes\boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}+O_p\left(T^{-1/2}\right)$. Then, after consecutive application of Lemmas A2 and A3, we can write $\boldsymbol{B}_T^{*}=\boldsymbol{B}_T^{**}+o_p(1)$, where $\boldsymbol{B}_T^{**}:=\boldsymbol{w}^{**}_{-1,\boldsymbol{d}}\boldsymbol{w}'^{**}_{-1,\boldsymbol{d}}/T$ can be represented as a partitioned matrix with ij-th block

$$\boldsymbol{B}_{Tij}^{**}:=\sum_{r=1}^{k}\sum_{s=1}^{k}\nu_{ir}\nu_{is}\left[\frac{1}{T}\boldsymbol{X}'^{**}_{i,-1,\boldsymbol{d}}\boldsymbol{u}_r\boldsymbol{u}_s'\boldsymbol{X}^{**}_{j,-1,\boldsymbol{d}}\right]$$

such that

$$E\left(\boldsymbol{B}_{Tij}^{**}\right)=\sum_{r=1}^{k}\sum_{s=1}^{k}\nu_{ir}\nu_{is}\left(T^{-1}\sum_{t=p+1}^{T}E\left(e_{r,t}e_{s,t}\boldsymbol{x}^{**}_{i,t-1,\boldsymbol{d}}\boldsymbol{x}'^{**}_{j,t-1,\boldsymbol{d}}\right)\right)=\boldsymbol{\Omega}_{\boldsymbol{B}ij}$$

from the stationarity and the MDS property of $\{\boldsymbol{e}_t\}$. Because $\left\{\boldsymbol{w}^{**}_{-1,\boldsymbol{d}}\boldsymbol{w}'^{**}_{-1,\boldsymbol{d}}\right\}$ is strictly stationary, ergodic, and L_2-bounded by Lemma A6, the Ergodic Theorem ensures that $\boldsymbol{B}_T^{**}\overset{a.s.}{\to}\boldsymbol{B}_{\boldsymbol{\beta}}$, so the AEL implies $\boldsymbol{B}_T^{*}\overset{p}{\to}\boldsymbol{B}_{\boldsymbol{\beta}}$. By Slutsky's Theorem, $\sqrt{T}\boldsymbol{R}\widehat{\boldsymbol{\beta}}\Rightarrow\mathcal{N}\left(0,\boldsymbol{R}\boldsymbol{\Omega}_{\boldsymbol{\beta}}\boldsymbol{R}'\right)$, and since $\boldsymbol{R}\boldsymbol{\Omega}_{\boldsymbol{\beta}}\boldsymbol{R}'$ is symmetric and nonnegative, there exists an upper triangular matrix $\boldsymbol{L}$ such that $\boldsymbol{R}\boldsymbol{\Omega}_{\boldsymbol{\beta}}\boldsymbol{R}'=\boldsymbol{L}'\boldsymbol{L}$. Consequently, $\sqrt{T}\boldsymbol{L}^{-1\prime}\boldsymbol{R}\widehat{\boldsymbol{\beta}}\Rightarrow\mathcal{N}\left(0,\boldsymbol{I}_k\right)$, and, hence,

$$\begin{aligned}LM_{\boldsymbol{d}} &= T\left[\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right]'\left(\boldsymbol{R}\boldsymbol{\Omega}_{\boldsymbol{\beta}}\boldsymbol{R}'\right)^{-1}\left[\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right]+O_p\left(T^{-1/2}\right)\\ &= T\left[\boldsymbol{L}'^{-1}\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right]'\left[\boldsymbol{L}'^{-1}\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right]\Rightarrow\chi^2_{(k)}.\end{aligned}$$

We now establish the corresponding asymptotic convergence result under the local alternative $\mathrm{H}_{\boldsymbol{c}}:\boldsymbol{\theta}=\boldsymbol{c}/\sqrt{T}$, where at least one element of $\boldsymbol{c}$ is non-zero. Under Assumption 1 and $\mathrm{H}_{\boldsymbol{c}}$, we have that $\boldsymbol{Y}_{t,\boldsymbol{d}}=\boldsymbol{X}^{*}_{-1,\boldsymbol{d}}\boldsymbol{\beta}_0+\boldsymbol{u}_{\boldsymbol{\theta}}$ where $\boldsymbol{u}_{\boldsymbol{\theta}}\equiv\left(\boldsymbol{u}'_{\boldsymbol{\theta}1},...,\boldsymbol{u}'_{\boldsymbol{\theta}k}\right)$,

$\boldsymbol{u}_{\boldsymbol{\theta}s} := \boldsymbol{u}_s + \frac{1}{\sqrt{T}}\boldsymbol{X}^*_{s,-1,\boldsymbol{d}+\boldsymbol{\theta}}\boldsymbol{\psi}_{\boldsymbol{c}s}$, and

$$\boldsymbol{\psi}_{\boldsymbol{c}s} := (c_s, -\boldsymbol{\pi}_{s1} \odot \boldsymbol{c}', ..., -\boldsymbol{\pi}_{sp} \odot \boldsymbol{c}')'$$

for $1 \leq s \leq k$; see Tanaka (1999). In this context, the FGLS estimator is given by

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + \left(\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^*_{-1,\boldsymbol{d}}\right)^{-1}\left(\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\widetilde{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}_{\boldsymbol{\theta}}\right) + o_p(1).$$

Lemma A3i) applies under the alternative hypothesis because, although the OLS estimator is no longer consistent, it still follows that $\widehat{\boldsymbol{u}}_s = \boldsymbol{u}_s + O_p\left(T^{-1/2}\right)$ and, hence, $\widehat{\nu}_{ij} = \nu_{ij} + O\left(T^{-1/2}\right)$. Consequently,

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) = \left(\frac{1}{T}\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^*_{-1,\boldsymbol{d}}\right)^{-1}\left(\frac{1}{\sqrt{T}}\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}_{\boldsymbol{\theta}}\right) + o_p(1).$$

Define $\boldsymbol{\psi}_{\boldsymbol{c}} := (\boldsymbol{\psi}'_{\boldsymbol{c}1}, ..., \boldsymbol{\psi}'_{\boldsymbol{c}k})'$. Then,

$$\frac{1}{\sqrt{T}}\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}_{\boldsymbol{\theta}} = \frac{1}{T}\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^*_{-1,\boldsymbol{d}+\boldsymbol{\theta}}\boldsymbol{\psi}_{\boldsymbol{c}} + \frac{1}{\sqrt{T}}\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}$$

where we can show that,

$$\frac{1}{T}\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^*_{-1,\boldsymbol{d}+\boldsymbol{\theta}}\boldsymbol{\psi}_{\boldsymbol{c}} \xrightarrow{p} \boldsymbol{A}_{\boldsymbol{\beta}}\boldsymbol{\psi}_{\boldsymbol{c}}$$

which follows from the Ergodic Theorem and the AEL because

$$T^{-1}\left\|\left(\boldsymbol{X}'^*_{-1,\boldsymbol{d}} - \boldsymbol{X}'^{**}_{-1,\boldsymbol{d}+\boldsymbol{\theta}}\right)\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\left(\boldsymbol{X}^*_{-1,\boldsymbol{d}+\boldsymbol{\theta}} - \boldsymbol{X}^{**}_{-1,\boldsymbol{d}+\boldsymbol{\theta}}\right)\right\| = O_p\left(\frac{\log T}{\sqrt{T}}\right) = o_p(1).$$

Similarly,

$$\frac{1}{\sqrt{T}}\boldsymbol{X}'^{**}_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u} \Rightarrow \mathcal{N}\left(0, \boldsymbol{B}_{\boldsymbol{\beta}}\right).$$

from the CLT for MDS, given that

$$T^{-1/2}\left\|\left(\boldsymbol{X}'^{**}_{-1,\boldsymbol{d}} - \boldsymbol{X}'^{**}_{-1,\boldsymbol{d}+\boldsymbol{\theta}}\right)\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{u}\right\| = O_p\left(\frac{\log T}{\sqrt{T}}\right)$$

and, finally,

$$\frac{1}{T}\boldsymbol{X}'^*_{-1,\boldsymbol{d}}\left[\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{T-p^*}\right]\boldsymbol{X}^*_{-1,\boldsymbol{d}} \xrightarrow{p} \boldsymbol{A}_{\boldsymbol{\beta}}.$$

Consequently, under Assumption 1 and $\mathrm{H}_{\boldsymbol{c}} : \boldsymbol{\theta} = \boldsymbol{c}/\sqrt{T}$,

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) \Rightarrow \mathcal{N}\left(\boldsymbol{\psi}_{\boldsymbol{c}}, \boldsymbol{\Omega}_{\boldsymbol{\beta}}\right). \tag{A.2}$$

This combined with the result that $\boldsymbol{A}_T^* \xrightarrow{p} \boldsymbol{\Omega}_{\boldsymbol{A}}$ ensures that $\widehat{\boldsymbol{u}} = \boldsymbol{u} + O_p\left(T^{-1/2}\right)$ and $\boldsymbol{B}_T^* \xrightarrow{p} \boldsymbol{\Omega}_{\boldsymbol{B}}$. Finally, from (A.2) we have that $\sqrt{T}\boldsymbol{R}\widehat{\boldsymbol{\beta}} \Rightarrow \mathcal{N}\left(\boldsymbol{c}, \boldsymbol{R}\boldsymbol{\Omega}_{\boldsymbol{\beta}}\boldsymbol{R}'\right)$ and, hence, $\sqrt{T}\left[\boldsymbol{L}'^{-1}\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right] \Rightarrow \mathcal{N}\left(\boldsymbol{L}'^{-1}\boldsymbol{c}, \boldsymbol{I}_k\right)$, and so the result that $\widehat{\boldsymbol{\Omega}}_T \xrightarrow{p} \boldsymbol{\Omega}_{\boldsymbol{\beta}}$ implies that

$$\begin{aligned} LM_{\boldsymbol{d}} &= T\left[\boldsymbol{L}'^{-1}\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right]'\left[\boldsymbol{L}'^{-1}\boldsymbol{R}\widehat{\boldsymbol{\beta}}\right] + O_p\left(T^{-1/2}\right) \\ &\Rightarrow \chi^2_{(k,\xi)} \end{aligned}$$

with $\xi := (\boldsymbol{L}'^{-1}\boldsymbol{c})'(\boldsymbol{L}'^{-1}\boldsymbol{c})$, as required. ■

Additional Reference

Davidson, J. (2000) *Stochastic Limit Theory* (3rd Ed). Oxford University Press: Oxford.

Appendix B - Additional Monte Carlo Results

To provide further insights into the finite sample performance of the tests we again focus on the bivariate ($k = 2$) case where $\mathbf{y}_t \equiv (y_{1t}, y_{2t})'$, and consider the simulation DGP,

$$\begin{bmatrix} (1-L)^{d_1} & 0 \\ 0 & (1-L)^{d_2} \end{bmatrix} (\mathbf{y}_t - \boldsymbol{\mu}) = \varepsilon_t I(t \geq 1), \quad t = 1, ..., T. \tag{B.1}$$

Without loss of generality we set $d_1 = 0.3$ and $d_2 = 0.6$, which correspond to a (marginally) stationary and a nonstationary long-memory process, respectively, and report results for samples of length $T \in \{500, 1000\}$. We set $\boldsymbol{\mu} = \mathbf{0}$ for all experiments, except for those in Section B.3. With the exception of the results in section B.1 where unconditional heteroskedasticity is allowed for, the innovations $\{\boldsymbol{\varepsilon}_t\}$ are generated to exhibit time-varying conditional second-order moments according to the design

$$\boldsymbol{\varepsilon}_t = \begin{bmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{bmatrix} \boldsymbol{\eta}_t; \quad E(\boldsymbol{\eta}_t) = \mathbf{0}, \quad E(\boldsymbol{\eta}_t \boldsymbol{\eta}_t') =: \boldsymbol{\Omega}_\rho = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{B.2}$$

where $\boldsymbol{\eta}_t \equiv (\eta_{1t}, \eta_{2t})'$ is an i.i.d. vector drawn from either a multivariate Gaussian distribution or a (heavy-tailed) multivariate Student-t distribution with 5 degrees of freedom. The covariance matrix $\boldsymbol{\Omega}_\rho$ depends on the contemporaneous correlation coefficient ρ, whose value we vary among $\rho \in \{0,\ 0.2,\ 0.4,\ 0.6,\ 0.8\}$. The conditional variances $\{\sigma_{it}^2\}$ are driven by (normalised) stationary GARCH(1,1) processes characterised by:

$$\sigma_{it}^2 = (1 - \alpha - \beta) + \alpha e_{i,t-1}^2 + \beta \sigma_{i,t-1}^2, \ i = 1, 2 \tag{B.3}$$

with $\alpha, \beta \geq 0$ and $\alpha + \beta < 1$, such that $E(e_{it}^2) = 1$. For simplicity, we impose the same GARCH dynamics on the two series, focusing on GARCH parameter configurations that allow for varying degrees of persistence in the conditional variances 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. Non-zero values of these parameters induce serial dependencies in the short-run dynamics of the process, while $\rho \neq 0$ introduces cross sectional dependence in the innovations. Applied work with financial data routinely suggests both the presence of heavy tailed behaviour in the innovations and high persistence in the fitted GARCH model with $\alpha + \beta$ generally estimated to be relatively close to one. All reported simulation results are based on 5000 Monte Carlo replications.

B.1 Unconditional Heteroskedasticity

In Table B.1 we report results for the case where the innovations are homoskedastic, DGP1: $\sigma_{1t}^2 = \sigma_{2t}^2 = 1$, and for the case where there is a contemporaneous one-time break of equal magnitude in the variances of $\boldsymbol{\varepsilon}_t$. Regarding the latter, two heteroskedastic cases are considered: (i) an upward change in variance such that DGP2: $\sigma_{1t}^2 = \sigma_{1t}^2 = 1\mathbb{I}(t \leq \lfloor \tau T \rfloor) + 4\mathbb{I}(t > \lfloor (1-\tau)T \rfloor)$, and (ii) a downward change where DGP3: $\sigma_{1t}^2 = \sigma_{2t}^2 = 1\mathbb{I}(t \leq \lfloor \tau T \rfloor) + \frac{1}{4}\mathbb{I}(t > \lfloor (1-\tau)T \rfloor)$, where in each case $\mathbb{I}(\cdot)$ denotes the indicator function, taking the value one when its argument is true and zero otherwise, and $\tau \in \{1/3, 1/2, 2/3\}$ corresponds to the break fraction. DGP2 and DGP3 allow us to examine the impact of unconditional heteroskedasticity, both in isolation and in its interaction with ρ, on the finite sample size of the tests. In each of DGP2 and DGP3 a fourfold change in variance is seen which is likely to be of considerably larger magnitude than we might expect to see in practice, but serves to illustrate how the tests behave in the presence of large changes in unconditional volatility.

Table B.1: Empirical rejection frequencies of $LM_{\mathbf{d}}^{FGLS}$ under the null hypothesis and unconditional variance breaks

		Normal			Student-t(5)		
DGP	ρ	$\tau=1/3$	$\tau=1/2$	$\tau=2/3$	$\tau=1/3$	$\tau=1/2$	$\tau=2/3$
				$T=500$			
DGP1	0	0.053	0.050	0.056	0.055	0.056	0.054
	0.2	0.055	0.051	0.047	0.049	0.056	0.057
	0.4	0.051	0.049	0.048	0.048	0.051	0.057
	0.6	0.052	0.056	0.056	0.054	0.054	0.056
	0.8	0.055	0.056	0.051	0.054	0.053	0.059
DGP2	0	0.056	0.053	0.054	0.048	0.054	0.052
	0.2	0.057	0.054	0.055	0.058	0.058	0.046
	0.4	0.056	0.062	0.054	0.060	0.053	0.054
	0.6	0.053	0.054	0.050	0.059	0.052	0.053
	0.8	0.056	0.063	0.056	0.053	0.049	0.058
DGP3	0	0.060	0.052	0.058	0.049	0.052	0.056
	0.2	0.054	0.053	0.063	0.058	0.051	0.053
	0.4	0.057	0.062	0.053	0.055	0.056	0.051
	0.6	0.055	0.054	0.055	0.055	0.054	0.056
	0.8	0.056	0.052	0.058	0.051	0.059	0.050
				$T=1000$			
DGP1	0	0.047	0.048	0.052	0.056	0.052	0.050
	0.2	0.049	0.049	0.051	0.057	0.054	0.057
	0.4	0.050	0.053	0.053	0.049	0.055	0.051
	0.6	0.049	0.051	0.053	0.050	0.049	0.056
	0.8	0.057	0.055	0.047	0.047	0.053	0.049
DGP2	0	0.049	0.053	0.057	0.052	0.049	0.051
	0.2	0.054	0.056	0.057	0.051	0.047	0.054
	0.4	0.052	0.055	0.054	0.053	0.051	0.051
	0.6	0.053	0.047	0.051	0.053	0.052	0.054
	0.8	0.049	0.053	0.055	0.057	0.052	0.058
DGP3	0	0.055	0.057	0.053	0.053	0.051	0.048
	0.2	0.052	0.050	0.055	0.047	0.053	0.053
	0.4	0.052	0.052	0.053	0.049	0.050	0.053
	0.6	0.055	0.053	0.055	0.059	0.048	0.050
	0.8	0.049	0.049	0.055	0.050	0.048	0.052

B.2 Estimation Accuracy

To illustrate the gains that can be obtained by using the FGLS approach over equation-by-equation OLS estimation (we thank a referee for suggesting these experiments) we have performed a detailed Monte Carlo analysis into the finite sample bias and mean squared error [MSE] of the estimates of the fractional integration parameter vector $\boldsymbol{d} := (d_1, d_2)'$, computed as described in section 5.2 and in Hassler *et al.* (2009, Remark 2.7). The simulation DGP is as described in (B.1)-(B.3) with $d_1 = 0.3$ and $d_2 = 0.6$, with 5000 Monte Carlo replications. Tables B.2 and B.3 report the empirical average (taken across the 5000 Monte Carlo replications) of the estimates of the long memory parameters, computed as $\bar{d}_i^k := \frac{1}{5000}\sum_{j=1}^{5000} \hat{d}_{i,j}^k$, together with the corresponding empirical MSEs computed as,

$$MSE_i^k := \frac{1}{5000}\sum_{j=1}^{5000}(\hat{d}_{i,j}^k - d_i)^2, \tag{B.4}$$

where in each case $i = 1, 2$ and $k = FGLS, OLS$, and where $\hat{d}_{i,j}^k$ denotes the estimate of d_i in the jth, $j = 1, ..., 5000$, Monte Carlo replication based on either FGLS estimation ($k = FGLS$) or equation-by-equation OLS estimation ($k = OLS$).

Table B.2: Empirical Average and MSE of fractional exponent estimates. DGP (B.1)-(B.3), $T = 500$.

ρ	$\bar{d}_1^{\,FGLS}$	$\bar{d}_1^{\,OLS}$	$\bar{d}_2^{\,FGLS}$	$\bar{d}_2^{\,OLS}$	MSE_1^{FGLS}	MSE_1^{OLS}	MSE_2^{FGLS}	MSE_2^{OLS}
				i.i.d. errors				
0	0.2979	0.2976	0.5978	0.5978	0.0013	0.0013	0.0013	0.0013
0.2	0.2981	0.2945	0.5984	0.5998	0.0013	0.0014	0.0012	0.0012
0.4	0.2980	0.2839	0.5987	0.6043	0.0013	0.0017	0.0013	0.0012
0.6	0.2979	0.2603	0.5984	0.6121	0.0013	0.0033	0.0013	0.0013
0.8	0.2986	0.2020	0.5985	0.6336	0.0013	0.0119	0.0013	0.0022
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.5$				
0	0.2976	0.2973	0.5978	0.5979	0.0015	0.0015	0.0015	0.0015
0.2	0.2980	0.2942	0.5986	0.6002	0.0015	0.0016	0.0014	0.0014
0.4	0.2979	0.2838	0.5984	0.6038	0.0015	0.0019	0.0015	0.0014
0.6	0.2976	0.2601	0.5982	0.6119	0.0015	0.0035	0.0015	0.0014
0.8	0.2986	0.2033	0.5981	0.6328	0.0015	0.0118	0.0015	0.0023
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.7$				
0	0.2971	0.2969	0.5968	0.5968	0.0015	0.0015	0.0016	0.0016
0.2	0.2980	0.2945	0.5978	0.5990	0.0015	0.0016	0.0015	0.0015
0.4	0.2975	0.2835	0.5978	0.6035	0.0015	0.0019	0.0015	0.0015
0.6	0.2990	0.2618	0.5980	0.6119	0.0016	0.0034	0.0015	0.0015
0.8	0.2981	0.2015	0.5979	0.6323	0.0016	0.0123	0.0015	0.0023
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.8$				
0	0.2971	0.2969	0.5981	0.5982	0.0016	0.0016	0.0016	0.0016
0.2	0.2983	0.2949	0.5988	0.6003	0.0016	0.0017	0.0015	0.0015
0.4	0.2971	0.2833	0.5973	0.6028	0.0016	0.0020	0.0016	0.0015
0.6	0.2966	0.2602	0.5970	0.6110	0.0017	0.0037	0.0016	0.0015
0.8	0.2974	0.2040	0.5968	0.6311	0.0015	0.0118	0.0016	0.0023
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.85$				
0	0.2972	0.2970	0.5981	0.5981	0.0018	0.0018	0.0017	0.0017
0.2	0.2971	0.2935	0.5991	0.6002	0.0017	0.0018	0.0016	0.0016
0.4	0.2973	0.2832	0.5979	0.6031	0.0017	0.0021	0.0017	0.0016
0.6	0.2972	0.2627	0.5973	0.6105	0.0017	0.0035	0.0017	0.0016
0.8	0.2976	0.2061	0.5970	0.6301	0.0018	0.0117	0.0017	0.0024

Table B.3: Empirical Average and MSE of fractional exponent estimates. (B.1)-(B.3), $T = 1000$.

ρ	$\bar{d}_1^{\,FGLS}$	$\bar{d}_1^{\,OLS}$	$\bar{d}_2^{\,FGLS}$	$\bar{d}_2^{\,OLS}$	MSE_1^{FGLS}	MSE_1^{OLS}	MSE_2^{FGLS}	MSE_2^{OLS}
				i.i.d. errors				
0	0.2987	0.2985	0.5987	0.5987	0.0006	0.0006	0.0006	0.0006
0.2	0.2988	0.2952	0.5990	0.6004	0.0006	0.0007	0.0006	0.0006
0.4	0.2982	0.2830	0.5994	0.6049	0.0006	0.0010	0.0006	0.0006
0.6	0.2993	0.2593	0.5991	0.6130	0.0006	0.0025	0.0006	0.0007
0.8	0.2986	0.1901	0.5991	0.6341	0.0006	0.0136	0.0006	0.0017
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.5$				
0	0.2989	0.2987	0.5986	0.5986	0.0007	0.0007	0.0007	0.0007
0.2	0.2989	0.2952	0.5996	0.6009	0.0007	0.0008	0.0007	0.0007
0.4	0.2983	0.2834	0.5990	0.6046	0.0007	0.0011	0.0007	0.0007
0.6	0.2990	0.2585	0.5990	0.6129	0.0008	0.0028	0.0007	0.0008
0.8	0.2987	0.1921	0.5986	0.6333	0.0007	0.0131	0.0007	0.0017
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.7$				
0	0.2985	0.2984	0.5986	0.5986	0.0007	0.0007	0.0008	0.0008
0.2	0.2992	0.2957	0.5988	0.6000	0.0007	0.0008	0.0008	0.0007
0.4	0.2985	0.2834	0.5992	0.6047	0.0008	0.0011	0.0008	0.0007
0.6	0.2986	0.2587	0.5996	0.6136	0.0007	0.0028	0.0008	0.0009
0.8	0.2993	0.1933	0.5988	0.6332	0.0008	0.0129	0.0008	0.0017
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.8$				
0	0.2988	0.2987	0.5984	0.5984	0.0008	0.0008	0.0008	0.0008
0.2	0.2988	0.2952	0.5995	0.6007	0.0008	0.0009	0.0008	0.0008
0.4	0.2983	0.2835	0.5991	0.6045	0.0008	0.0012	0.0008	0.0008
0.6	0.2986	0.2593	0.5987	0.6123	0.0008	0.0028	0.0008	0.0009
0.8	0.2986	0.1945	0.5985	0.6324	0.0008	0.0128	0.0008	0.0017
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.85$				
0	0.2985	0.2984	0.5982	0.5982	0.0008	0.0008	0.0009	0.0009
0.2	0.2986	0.2953	0.5994	0.6005	0.0009	0.0009	0.0009	0.0009
0.4	0.2980	0.2839	0.5989	0.6041	0.0009	0.0012	0.0009	0.0008
0.6	0.2984	0.2603	0.5987	0.6118	0.0009	0.0028	0.0009	0.0009
0.8	0.2983	0.1972	0.5987	0.6314	0.0009	0.0123	0.0009	0.0018

We also computed the empirical MSE of the parameter estimates of β_i, $i = 1, 2$ in (6) resulting from a multivariate linear regression model as in (5) with $p = 0$, using FGLS and equation-by-equation OLS. Notice that in this restricted framework ($p = 0$) the β_i, $i = 1, 2$ can be seen as indicators of over ($\beta_i < 0$) or under differencing ($\beta_i > 0$) of the time series induced by the null hypothesis, as the null hypothesis implies that $\boldsymbol{\phi} = 0$ in (5) which in the restricted case is equivalent to $\beta_i = 0$, $i = 1, 2$.

Table B.4: Empirical MSE of the parameter estimates of the β_i, $i = 1, 2$ in (6), $\hat{\beta}_i$, $i = 1, 2$, computed by FGLS and OLS

ρ	$MSE^{FGLS}_{\hat{\beta}_1}$	$MSE^{OLS}_{\hat{\beta}_1}$	$MSE^{FGLS}_{\hat{\beta}_2}$	$MSE^{OLS}_{\hat{\beta}_2}$	$MSE^{FGLS}_{\hat{\beta}_1}$	$MSE^{OLS}_{\hat{\beta}_1}$	$MSE^{FGLS}_{\hat{\beta}_2}$	$MSE^{OLS}_{\hat{\beta}_2}$
	$T = 500$				$T = 1000$			
	GARCH: $\theta_1 = 0.1; \theta_2 = 0.5$							
0	0.0027	0.0027	0.0024	0.0024	0.0013	0.0013	0.0013	0.0013
0.2	0.0026	0.0027	0.0025	0.0026	0.0013	0.0013	0.0013	0.0013
0.4	0.0024	0.0027	0.0022	0.0025	0.0012	0.0013	0.0012	0.0013
0.6	0.0021	0.0026	0.0020	0.0025	0.0010	0.0013	0.0011	0.0013
0.8	0.0017	0.0027	0.0017	0.0026	0.0008	0.0013	0.0008	0.0013
	GARCH: $\theta_1 = 0.1; \theta_2 = 0.7$							
0	0.0029	0.0029	0.0028	0.0028	0.0016	0.0016	0.0016	0.0016
0.2	0.0029	0.0030	0.0027	0.0028	0.0015	0.0016	0.0015	0.0015
0.4	0.0027	0.0030	0.0025	0.0028	0.0015	0.0016	0.0014	0.0016
0.6	0.0023	0.0029	0.0022	0.0027	0.0013	0.0017	0.0012	0.0015
0.8	0.0020	0.0029	0.0019	0.0028	0.0011	0.0016	0.0011	0.0016
	GARCH: $\theta_1 = 0.1; \theta_2 = 0.8$							
0	0.0036	0.0036	0.0034	0.0034	0.0021	0.0021	0.0021	0.0021
0.2	0.0034	0.0035	0.0034	0.0035	0.0021	0.0022	0.0020	0.0020
0.4	0.0034	0.0038	0.0032	0.0035	0.0019	0.0021	0.0020	0.0021
0.6	0.0029	0.0036	0.0028	0.0034	0.0018	0.0022	0.0017	0.0021
0.8	0.0024	0.0035	0.0025	0.0034	0.0016	0.0022	0.0015	0.0021
	GARCH: $\theta_1 = 0.1; \theta_2 = 0.85$							
0	0.0042	0.0043	0.0041	0.0041	0.0029	0.0029	0.0029	0.0029
0.2	0.0042	0.0043	0.0042	0.0042	0.0031	0.0031	0.0030	0.0030
0.4	0.0039	0.0042	0.0039	0.0042	0.0030	0.0032	0.0027	0.0029
0.6	0.0036	0.0043	0.0036	0.0043	0.0027	0.0031	0.0026	0.0030
0.8	0.0032	0.0044	0.0031	0.0042	0.0023	0.0031	0.0023	0.0031

B.3 The Impact of Nonzero Means

To illustrate the impact of nonzero means on the finite sample size performance of the test procedure, we consider the following three cases: $\boldsymbol{\mu} = \mathbf{0}, \mathbf{5}, \mathbf{10}$ which correspond to 2×1 vectors of common elements (0, 5 and 10, respectively), and we use three different demeaning approaches: i) no demeaning (which we denote as μ_0); ii) demeaning only (which we denote as μ_1); and iii) demeaning and detrending (which we denote as μ_2).

Specifically, to account for a non-zero deterministic mean in the level of the series we use the demeaning process described in Robinson (1994); Demetrescu et al. (2008) and Hassler et al. (2016). Hence, to account for the nonzero means in (B.1) we regress the differences $(1-L)_{+}^{d_i} y_{it} := \sum_{j=0}^{t-1} \lambda_j(d_i) y_{it-j}$ on the variable $h_{t,d_i} := \sum_{j=0}^{t-1} \lambda_j(d_i)$, $t = 2, ..., T$, with $\{\lambda_j(d_i)\}$ as defined in (**??**) of the paper. Denote the resulting estimates $\widetilde{\mu}_i$, $i = 1, 2$, and the corresponding residuals as $\widetilde{\varepsilon}_{it,d_i} := (1-L)_{+}^{d_i} y_{it} - \widetilde{\mu}_i h_{t,d_i}$. One then redefines the ith element of the vector $\boldsymbol{\varepsilon}_{t,\mathbf{d}}$ to be $\widetilde{\varepsilon}_{it,d_i}$, $i = 1, 2$, and then proceeds as before to compute the respective test statistics; see Remark 8 for further details.

Table B.5: Impact of $\boldsymbol{\mu}$ on finite sample size performance. Normally distributed innovations. $T = 500$.

ρ	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{10}$
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.5$					
0	0.057	1.000	1.000	0.059	0.059	0.059	0.065	0.065	0.065
0.2	0.068	1.000	1.000	0.060	0.060	0.060	0.069	0.069	0.069
0.4	0.058	1.000	1.000	0.060	0.060	0.060	0.069	0.069	0.069
0.6	0.059	1.000	1.000	0.062	0.062	0.062	0.067	0.067	0.067
0.8	0.058	1.000	1.000	0.066	0.066	0.066	0.066	0.066	0.066
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.7$					
0	0.056	1.000	1.000	0.060	0.060	0.060	0.066	0.066	0.066
0.2	0.055	1.000	1.000	0.060	0.060	0.060	0.068	0.068	0.068
0.4	0.056	1.000	1.000	0.061	0.061	0.061	0.069	0.069	0.069
0.6	0.058	1.000	1.000	0.061	0.061	0.061	0.065	0.065	0.065
0.8	0.059	1.000	1.000	0.066	0.066	0.066	0.069	0.069	0.069
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.8$					
0	0.054	1.000	1.000	0.058	0.058	0.058	0.067	0.067	0.067
0.2	0.056	1.000	1.000	0.060	0.060	0.060	0.067	0.067	0.067
0.4	0.058	1.000	1.000	0.064	0.064	0.064	0.070	0.070	0.070
0.6	0.058	1.000	1.000	0.066	0.066	0.066	0.069	0.069	0.069
0.8	0.059	1.000	1.000	0.066	0.066	0.066	0.069	0.069	0.069
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.85$					
0	0.055	1.000	1.000	0.057	0.057	0.057	0.066	0.066	0.066
0.2	0.059	1.000	1.000	0.063	0.063	0.063	0.069	0.069	0.069
0.4	0.062	1.000	1.000	0.065	0.065	0.065	0.072	0.072	0.072
0.6	0.059	1.000	1.000	0.066	0.066	0.066	0.071	0.071	0.071
0.8	0.060	1.000	1.000	0.068	0.068	0.068	0.071	0.071	0.071

Note: $LM^{FGLS}_{\mathbf{d},\mu_i}$, $i = 0, 1, 2$ correspond to statistics computed from data which has not been demeaned (μ_0), data that has been demeaned (μ_1) and data which has been demeaned and detrended (μ_2).

Table B.6: Impact of $\boldsymbol{\mu}$ on finite sample size performance. Normally distributed innovations. $T = 1000$.

ρ	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{10}$
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.5$					
0	0.060	1.000	1.000	0.058	0.058	0.058	0.067	0.067	0.067
0.2	0.058	1.000	1.000	0.058	0.058	0.058	0.066	0.066	0.066
0.4	0.055	1.000	1.000	0.057	0.057	0.057	0.065	0.065	0.065
0.6	0.054	1.000	1.000	0.056	0.056	0.056	0.062	0.062	0.062
0.8	0.051	1.000	1.000	0.058	0.058	0.058	0.056	0.056	0.056
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.7$					
0	0.059	1.000	1.000	0.058	0.058	0.058	0.067	0.067	0.067
0.2	0.059	1.000	1.000	0.057	0.057	0.057	0.065	0.065	0.065
0.4	0.054	1.000	1.000	0.056	0.056	0.056	0.064	0.064	0.064
0.6	0.052	1.000	1.000	0.056	0.056	0.056	0.061	0.061	0.061
0.8	0.054	1.000	1.000	0.061	0.061	0.061	0.057	0.057	0.057
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.8$					
0	0.057	1.000	1.000	0.058	0.058	0.058	0.063	0.063	0.063
0.2	0.055	1.000	1.000	0.056	0.056	0.056	0.064	0.064	0.064
0.4	0.056	1.000	1.000	0.057	0.057	0.057	0.060	0.060	0.060
0.6	0.053	1.000	1.000	0.057	0.057	0.057	0.059	0.059	0.059
0.8	0.050	1.000	1.000	0.059	0.059	0.059	0.059	0.059	0.050
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.85$					
0	0.053	1.000	1.000	0.056	0.056	0.056	0.060	0.060	0.060
0.2	0.054	1.000	1.000	0.054	0.054	0.054	0.060	0.060	0.060
0.4	0.054	1.000	1.000	0.056	0.056	0.056	0.060	0.060	0.060
0.6	0.051	1.000	1.000	0.056	0.056	0.056	0.056	0.056	0.056
0.8	0.052	1.000	1.000	0.060	0.060	0.060	0.058	0.058	0.058

Note: See note under Table B.5

Table B.7: Impact of $\boldsymbol{\mu}$ on finite sample size performance. Student-t distributed innovations (5 degrees of freedom). $T = 500$.

ρ	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{10}$
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.5$					
0	0.065	1.000	1.000	0.066	0.066	0.066	0.070	0.070	0.070
0.2	0.063	1.000	1.000	0.064	0.064	0.064	0.073	0.073	0.073
0.4	0.064	1.000	1.000	0.066	0.066	0.066	0.076	0.076	0.076
0.6	0.062	1.000	1.000	0.065	0.065	0.065	0.075	0.075	0.075
0.8	0.061	1.000	1.000	0.070	0.070	0.070	0.071	0.071	0.071
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.7$					
0	0.060	0.998	1.000	0.065	0.065	0.065	0.069	0.069	0.069
0.2	0.062	0.997	1.000	0.063	0.063	0.063	0.072	0.072	0.072
0.4	0.063	0.997	1.000	0.064	0.064	0.064	0.073	0.073	0.073
0.6	0.064	0.998	1.000	0.063	0.063	0.063	0.073	0.073	0.073
0.8	0.064	0.999	1.000	0.068	0.068	0.068	0.070	0.070	0.070
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.8$					
0	0.063	0.970	0.997	0.064	0.064	0.064	0.067	0.067	0.067
0.2	0.065	0.969	0.996	0.063	0.063	0.063	0.070	0.070	0.070
0.4	0.065	0.969	0.995	0.065	0.065	0.065	0.072	0.072	0.072
0.6	0.069	0.973	0.995	0.065	0.065	0.065	0.070	0.070	0.070
0.8	0.069	0.982	0.997	0.072	0.072	0.072	0.074	0.074	0.074
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.85$					
0	0.057	0.627	0.870	0.059	0.059	0.059	0.064	0.064	0.064
0.2	0.060	0.621	0.862	0.061	0.061	0.061	0.065	0.065	0.065
0.4	0.059	0.620	0.860	0.062	0.062	0.062	0.069	0.069	0.069
0.6	0.064	0.641	0.868	0.063	0.063	0.063	0.068	0.068	0.068
0.8	0.064	0.699	0.892	0.065	0.065	0.065	0.069	0.069	0.069

Note: See note under Table B.5

Table B.8: Impact of $\boldsymbol{\mu}$ on finite sample size performance. Student-t distributed innovations (5 degrees of freedom). $T = 1000$.

ρ	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_0}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_1}$ $\boldsymbol{\mu}=\mathbf{10}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{0}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{5}$	$LM^{FGLS}_{\mathbf{d},\mu_2}$ $\boldsymbol{\mu}=\mathbf{10}$
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.5$					
0	0.052	1.000	1.000	0.056	0.056	0.056	0.059	0.059	0.059
0.2	0.053	1.000	1.000	0.056	0.056	0.056	0.057	0.057	0.057
0.4	0.057	1.000	1.000	0.058	0.058	0.058	0.059	0.059	0.059
0.6	0.056	1.000	1.000	0.058	0.058	0.058	0.056	0.056	0.056
0.8	0.051	1.000	1.000	0.059	0.059	0.059	0.053	0.053	0.053
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.7$					
0	0.057	0.998	1.000	0.059	0.059	0.059	0.063	0.063	0.063
0.2	0.053	0.998	1.000	0.056	0.056	0.056	0.057	0.057	0.057
0.4	0.054	0.998	1.000	0.056	0.056	0.056	0.058	0.058	0.058
0.6	0.055	0.999	1.000	0.059	0.059	0.059	0.057	0.057	0.057
0.8	0.055	0.999	1.000	0.058	0.058	0.058	0.059	0.059	0.059
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.8$					
0	0.059	0.962	0.997	0.064	0.064	0.064	0.065	0.065	0.065
0.2	0.062	0.959	0.997	0.063	0.063	0.063	0.064	0.064	0.064
0.4	0.057	0.961	0.996	0.061	0.061	0.061	0.063	0.063	0.063
0.6	0.059	0.965	0.997	0.059	0.059	0.059	0.061	0.061	0.061
0.8	0.060	0.979	0.998	0.063	0.063	0.063	0.065	0.065	0.065
				GARCH: $\theta_1 = 0.1; \theta_2 = 0.85$					
0	0.065	0.508	0.774	0.065	0.065	0.065	0.067	0.067	0.067
0.2	0.067	0.502	0.769	0.066	0.066	0.066	0.068	0.068	0.068
0.4	0.062	0.507	0.768	0.064	0.064	0.064	0.065	0.065	0.065
0.6	0.062	0.519	0.776	0.061	0.061	0.061	0.063	0.063	0.063
0.8	0.063	0.576	0.803	0.062	0.062	0.062	0.062	0.062	0.062

Note: See note under Table B.5

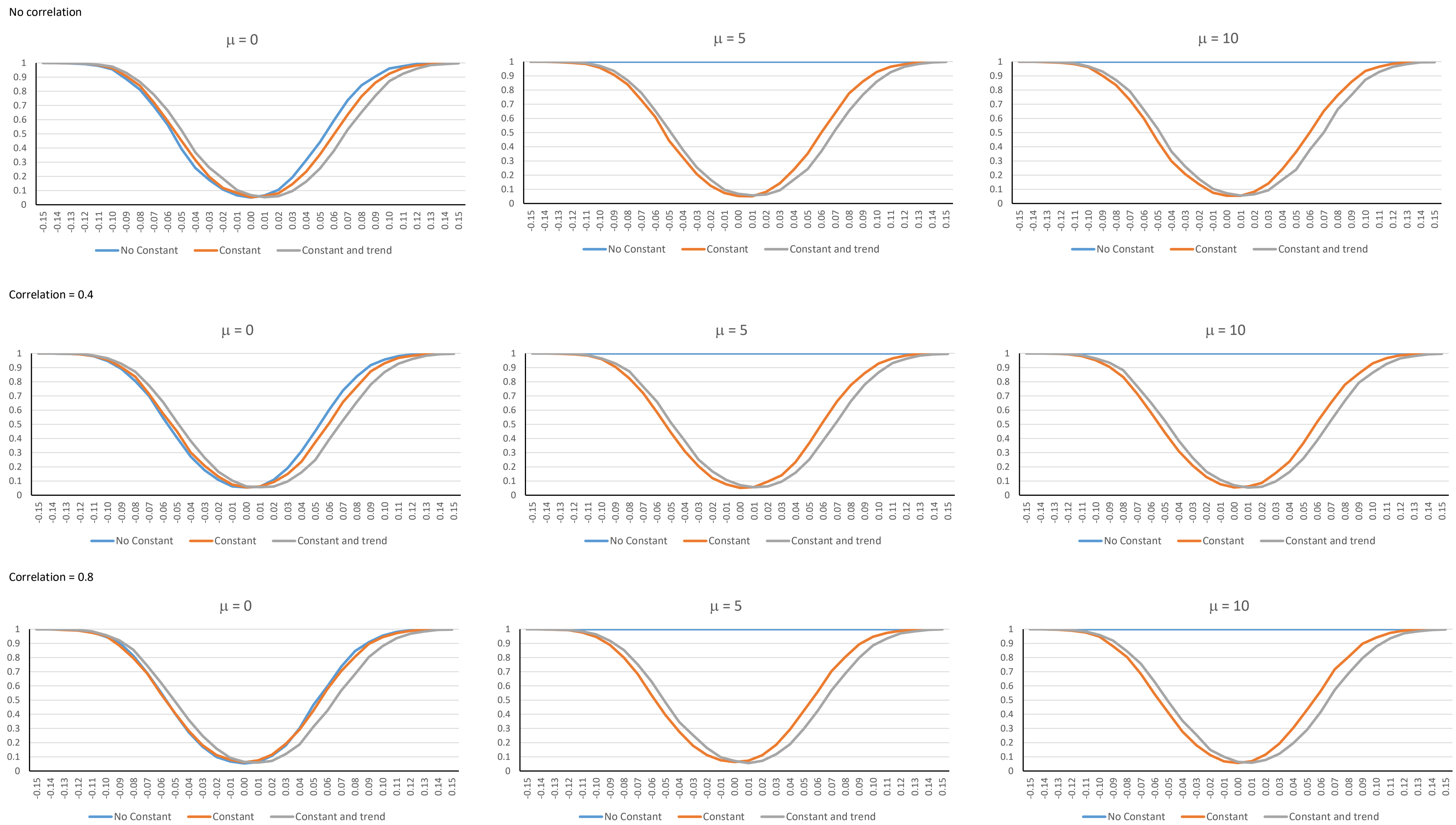


Figure B.1: Power of test when data is not demeaned or detrended, when data is demeaned and when data is demeaned and detrended using the approach described in Remark 8, and for $\boldsymbol{\mu} = \{\mathbf{0}, \mathbf{5}, \mathbf{10}\}$ and sample size $T = 500$.

B.4 Performance Under Fractional Cointegration

The data generation process (DGP) considered for investigating the impact of fractional cointegration is the same as that used in Nielsen (2005); that is,

$$y_{1t} = y_{2t} + u_t \quad \text{(B.5)}$$

$$\begin{bmatrix} (1-L)^{d-\theta} & 0 \\ 0 & (1-L)^d \end{bmatrix} \begin{bmatrix} u_t \\ y_{2t} \end{bmatrix} = \boldsymbol{\varepsilon}_t I(t \geq 1) \quad \text{(B.6)}$$

where

$$\boldsymbol{\varepsilon}_t \sim i.i.d.\ N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

For $\rho = 0$, y_{2t} is strictly exogenous whereas for $\rho \neq 0$, y_{2t} is endogenous. We consider $\rho \in \{0, 0.4, 0.8\}$.

Table B.9 reports empirical rejection frequencies for the $LM_{\mathbf{d}}^{FGLS}$, $LM_{\mathbf{d}}^{MLE}$ and the $BH_{\mathbf{d}}$ tests for data generated from (B.5)–(B.6) with $d = 0.6$ (without loss of generality) and for $\theta \in \{0, -0.01, -0.02, ..., -0.2\}$. The parameter θ measures the degree of fractional cointegration, with the case of no fractional cointegration corresponding to $\theta = 0$. All of the tests are run at the nominal 5% level.

Table B.9: Empirical Rejection Frequencies under Fractional Cointegration - $T = 500$

θ	ρ	$LM_{\mathbf{d}}^{FGLS}$	$LM_{\mathbf{d}}^{MLE}$	$BH_{\mathbf{d}}$	ρ	$LM_{\mathbf{d}}^{FGLS}$	$LM_{\mathbf{d}}^{MLE}$	$BH_{\mathbf{d}}$	ρ	$LM_{\mathbf{d}}^{FGLS}$	$LM_{\mathbf{d}}^{MLE}$	$BH_{\mathbf{d}}$	ρ	$LM_{\mathbf{d}}^{FGLS}$	$LM_{\mathbf{d}}^{MLE}$	$BH_{\mathbf{d}}$	ρ	$LM_{\mathbf{d}}^{FGLS}$	$LM_{\mathbf{d}}^{MLE}$	$BH_{\mathbf{d}}$
0.00	0	0.052	0.049	0.051	0.20	0.052	0.048	0.048	0.40	0.051	0.048	0.046	0.60	0.054	0.051	0.047	0.80	0.055	0.045	0.044
-0.01	0	0.055	0.057	0.051	0.20	0.062	0.064	0.054	0.40	0.056	0.060	0.059	0.60	0.055	0.064	0.054	0.80	0.064	0.075	0.066
-0.02	0	0.057	0.083	0.064	0.20	0.073	0.081	0.062	0.40	0.071	0.084	0.065	0.60	0.076	0.098	0.075	0.80	0.087	0.136	0.090
-0.03	0	0.079	0.113	0.081	0.20	0.082	0.130	0.090	0.40	0.083	0.134	0.097	0.60	0.092	0.158	0.109	0.80	0.117	0.242	0.165
-0.04	0	0.095	0.178	0.125	0.20	0.095	0.165	0.109	0.40	0.096	0.191	0.124	0.60	0.119	0.246	0.155	0.80	0.172	0.391	0.271
-0.05	0	0.123	0.245	0.154	0.20	0.115	0.239	0.160	0.40	0.139	0.275	0.177	0.60	0.172	0.350	0.233	0.80	0.247	0.566	0.416
-0.06	0	0.143	0.321	0.210	0.20	0.149	0.325	0.211	0.40	0.170	0.363	0.244	0.60	0.218	0.465	0.329	0.80	0.341	0.735	0.586
-0.07	0	0.188	0.431	0.291	0.20	0.189	0.434	0.291	0.40	0.218	0.501	0.340	0.60	0.273	0.606	0.459	0.80	0.434	0.854	0.728
-0.08	0	0.219	0.513	0.355	0.20	0.235	0.548	0.386	0.40	0.258	0.603	0.433	0.60	0.332	0.722	0.561	0.80	0.545	0.930	0.838
-0.09	0	0.263	0.625	0.446	0.20	0.287	0.647	0.463	0.40	0.328	0.709	0.530	0.60	0.415	0.839	0.690	0.80	0.664	0.974	0.924
-0.10	0	0.335	0.728	0.552	0.20	0.353	0.745	0.565	0.40	0.388	0.799	0.641	0.60	0.498	0.902	0.786	0.80	0.762	0.993	0.967
-0.11	0	0.400	0.811	0.641	0.20	0.405	0.826	0.652	0.40	0.466	0.876	0.741	0.60	0.588	0.952	0.869	0.80	0.828	0.998	0.989
-0.12	0	0.459	0.871	0.729	0.20	0.465	0.886	0.741	0.40	0.521	0.924	0.809	0.60	0.660	0.978	0.924	0.80	0.898	0.999	0.997
-0.13	0	0.532	0.923	0.800	0.20	0.553	0.932	0.829	0.40	0.611	0.961	0.888	0.60	0.739	0.992	0.960	0.80	0.929	1.000	0.999
-0.14	0	0.593	0.956	0.861	0.20	0.613	0.968	0.885	0.40	0.688	0.981	0.926	0.60	0.802	0.997	0.981	0.80	0.968	1.000	1.000
-0.15	0	0.651	0.978	0.913	0.20	0.667	0.978	0.928	0.40	0.749	0.993	0.961	0.60	0.852	0.999	0.991	0.80	0.984	1.000	1.000
-0.16	0	0.723	0.987	0.944	0.20	0.738	0.993	0.956	0.40	0.805	0.996	0.977	0.60	0.899	1.000	0.997	0.80	0.994	1.000	1.000
-0.17	0	0.777	0.995	0.973	0.20	0.811	0.997	0.975	0.40	0.851	0.999	0.990	0.60	0.939	1.000	0.999	0.80	0.997	1.000	1.000
-0.18	0	0.822	0.996	0.980	0.20	0.851	0.999	0.990	0.40	0.891	1.000	0.996	0.60	0.956	1.000	0.999	0.80	0.999	1.000	1.000
-0.19	0	0.871	0.999	0.993	0.20	0.885	1.000	0.994	0.40	0.918	1.000	0.998	0.60	0.975	1.000	1.000	0.80	0.999	1.000	1.000
-0.20	0	0.906	0.999	0.995	0.20	0.914	1.000	0.998	0.40	0.945	1.000	0.999	0.60	0.984	1.000	1.000	0.80	1.000	1.000	1.000

Appendix C - Empirical Results

C.1 Implementation of the Nielsen and Shimotsu (2007) Procedure

For a bivariate system, the NS test statistic, T_0, for testing the null hypothesis of equality in the long memory coefficients is defined as

$$T_0 := m_T \left(\boldsymbol{S}\ \widehat{\boldsymbol{d}}_T\right)' \left(\frac{1}{4}\boldsymbol{S}\ \boldsymbol{D}_T^{-1}\left(\boldsymbol{G}_T \odot \boldsymbol{G}_T\right)\boldsymbol{D}_T^{-1}\boldsymbol{S}' + \frac{1}{s_T}\right)^{-1}\left(\boldsymbol{S}\ \widehat{\boldsymbol{d}}_T\right) \tag{C.1}$$

where $\boldsymbol{S} := (1,-1)'$, $\widehat{\boldsymbol{d}}_T$ denotes an $\sqrt{m_T}-$consistent estimate of the long memory parameter vector, $\boldsymbol{G}_T = \{\widehat{g}_{ij}\}$, $i,j \in \{1,2\}$, is a consistent estimate of the spectral density of $\boldsymbol{\varepsilon}_t$ at the origin, $\boldsymbol{D}_T := diag\left(\widehat{g}_{11}, \widehat{g}_{22}\right)$, and m_T and s_T are positive sequences that diverge at a suitable rate as $T \to \infty$.

The cointegration rank can be consistently estimated through a model selection procedure based on the eigenvalues $\widehat{\delta}_i^*$ of the correlation matrix $\boldsymbol{P}_T := \boldsymbol{D}_T^{*-1/2}\boldsymbol{G}_T^*\boldsymbol{D}_T^{*-1/2}$, with $\boldsymbol{G}_T^*$ denoting an estimate of the spectral density of $\boldsymbol{\Delta}^{\boldsymbol{d}_T^*}(L)\,\boldsymbol{y}_t := \boldsymbol{\varepsilon}_t\left(\boldsymbol{d}_T^*\right)$ at the origin, with $\boldsymbol{d}_T^*$ denoting a k-vector with all entries equal to the sample mean of $\widehat{\boldsymbol{d}}_T$, and $\boldsymbol{D}_T^*$ defined analogously to $\boldsymbol{D}_T$. In particular, $\boldsymbol{G}_T^* := \sum_{j=1}^{m_{1T}} \Re\left[I_{\boldsymbol{\varepsilon}_t(\boldsymbol{d}^*)}\left(\lambda_j\right)\right]/m_{1T}$, where $I_{\boldsymbol{u}}\left(\lambda_j\right)$ is the periodogram of $\boldsymbol{u}$ evaluated at the fundamental frequencies $\lambda_j := 2\pi j/T$, m_{1T} is a bandwidth parameter, and $\Re[\cdot]$ denotes the real part of the argument. Given the eigenvalues $\widehat{\delta}_i^*$ and a suitable bandwith parameter v_T, the cointegration rank can be determined as

$$\widehat{r}_T = \arg\min_{u=0,1} L(u)\,, \quad L(u) := v_T\,(2-u) - \sum_{i=1}^{2-u}\widehat{\delta}_i^*. \tag{C.2}$$

C.2 Additional Empirical Results

C.2.1 Descriptive statistics and robustness checks

Table C.1. Descriptive statistics (mean, standard deviation, minimum, maximum, skewness and kurtosis) and Engle's LM test for ARCH effects, for the log-volume, log absolute returns, log range estimator and log-realized variance for each stock series considered.

		Log Trading Volume								**Log Absolute Returns**							
Ticker	Company	Mean	StdDev	Max	Min	Skew.	Kurt.	$LM_{(1)}$	$LM_{(5)}$	Mean	StdDev	Max	Min	Skew.	Kurt.	$LM_{(1)}$	$LM_{(5)}$
AAPL	Apple	16.70	0.74	19.06	14.14	-0.44	3.00	18.91***	32.79***	-4.62	1.16	-1.72	-10.98	-1.01	4.73	0.06	0.95
AXP	Amex	15.65	0.62	18.32	13.70	0.63	3.32	16.92***	17.89***	-4.87	1.23	-1.58	-9.15	-0.56	3.51	0.58	1.14
BA	Boeing Co	15.32	0.45	17.61	13.13	0.34	3.67	2.56	5.22	-4.83	1.11	-1.87	-10.20	-0.90	4.24	0.20	7.65
CAT	Caterpillar	15.47	0.64	18.03	12.87	-0.14	2.90	1.06	3.88	-4.78	1.17	-1.92	-9.34	-0.87	4.05	1.82	3.60
CSCO	Cisco Systems	17.69	0.41	20.15	15.78	0.25	4.92	76.00***	82.54***	-4.81	1.11	-1.82	-9.76	-0.69	3.55	0.01	1.53
CVX	Chevron	15.79	0.56	17.69	13.52	-0.35	3.00	34.83***	36.16***	-4.97	1.10	-1.57	-9.47	-0.91	4.35	3.56*	4.05
DD	DuPont Co.	15.41	0.47	17.26	13.35	0.19	3.04	6.01	8.18	-4.96	1.13	-2.17	-8.81	-0.74	3.64	0.21	3.27
DIS	Walt Disney Co.	16.01	0.44	18.56	14.20	0.44	4.05	20.03***	20.68***	-4.91	1.10	-1.83	-9.11	-0.66	3.48	0.00	1.16
GE	General Electrics	17.45	0.65	20.44	15.44	0.56	3.24	13.85***	15.61***	-4.97	1.11	-1.62	-9.02	-0.45	3.30	0.11	4.97
GS	Goldman Sachs	15.51	0.70	18.56	13.46	0.73	3.54	8.78***	9.59*	-4.77	1.21	-1.33	-9.97	-0.85	4.25	1.79	2.52
HD	Home Depot	16.16	0.53	18.36	14.41	0.29	2.89	22.61***	29.70***	-4.92	1.12	-1.96	-9.00	-0.66	3.39	0.02	3.96
IBM	IBM	15.55	0.44	17.24	14.16	0.29	3.29	31.17***	40.61***	-5.16	1.12	-2.16	-9.92	-0.84	4.18	1.81	4.91
INTC	Intel Corporation	17.79	0.43	19.55	16.26	-0.21	3.59	65.84***	68.98***	-4.74	1.08	-2.09	-12.43	-0.86	4.52	0.20	4.22
JNJ	Johnson & Johnson	16.06	0.42	18.40	14.06	0.33	3.85	39.50***	41.20***	-5.45	1.14	-2.10	-9.29	-0.73	3.58	0.06	2.69
JPM	JPMorgan Chase & Co.	16.81	0.75	19.20	14.11	0.30	2.51	6.41**	7.74	-4.80	1.22	-1.38	-8.73	-0.50	3.41	3.43*	5.99
KO	The Coca-Cola Co.	15.97	0.49	18.41	13.89	0.16	3.07	12.08***	15.27***	-5.30	1.07	-1.97	-8.90	-0.61	3.40	0.03	3.06
MCD	McDonald's Corporation	15.63	0.44	18.28	13.96	0.47	4.00	22.89***	26.32***	-5.15	1.12	-2.37	-10.35	-0.84	3.92	0.43	2.73
MMM	3M Co.	14.97	0.43	17.01	13.39	0.52	3.88	3.72*	15.75***	-5.18	1.15	-2.31	-9.69	-0.75	3.81	1.64	2.46
MRK	Merck & Co., Inc.	16.25	0.51	18.79	14.45	0.38	3.62	8.65***	10.84*	-4.99	1.09	-1.32	-8.70	-0.63	3.61	1.15	4.19
MSFT	Microsoft Corporation	17.83	0.42	20.20	16.22	0.11	4.08	18.08***	20.07***	-4.98	1.11	-1.68	-9.19	-0.63	3.42	0.10	5.40
NKE	Nike	14.67	0.57	16.89	12.74	0.03	3.14	24.61***	43.40***	-4.98	1.12	-2.07	-9.19	-0.70	3.81	1.21	4.27
PFE	Pfizer Inc.	17.31	0.50	19.49	15.18	0.26	3.23	16.91***	26.21***	-5.02	1.04	-2.19	-8.44	-0.58	3.25	0.07	3.31
PG	Procter & Gamble Co.	15.93	0.55	18.09	13.15	-0.26	3.51	30.12***	34.28***	-5.37	1.09	-2.28	-9.86	-0.69	3.65	0.35	0.83
TRV	Travelers Companies Inc	14.97	0.53	17.14	13.15	0.12	2.79	8.92***	9.56*	-5.01	1.21	-1.36	-9.25	-0.57	3.55	1.57	7.19
UNH	UnitedHealth Group Inc	15.56	0.63	17.76	13.38	-0.01	2.85	5.62**	8.59	-4.84	1.16	-1.06	-9.18	-0.79	3.95	0.01	2.94
UTX	United Technologies Corporation	15.10	0.49	16.86	12.82	0.03	3.14	8.73***	11.46**	-5.06	1.15	-1.99	-9.34	-0.82	3.80	4.34**	7.80
VZ	Verizon Communications Inc.	16.31	0.50	20.24	14.47	0.37	4.67	16.11***	29.33***	-5.09	1.08	-1.92	-8.52	-0.66	3.36	3.42*	4.73
V	Visa Inc	14.39	1.47	18.25	10.52	-0.45	2.10	52.43***	65.75***	-4.75	1.13	-1.90	-10.01	-0.84	4.20	0.77	4.49
WMT	Wal-Mart Stores Inc	16.23	0.52	18.39	14.56	0.23	2.92	15.58***	15.88***	-5.23	1.09	-2.20	-9.32	-0.78	3.78	0.76	1.56
XOM	Exxon Mobil Corporation	16.69	0.47	18.59	15.24	0.34	3.04	23.21***	24.30***	-5.05	1.12	-1.76	-9.15	-0.90	4.20	0.91	1.30

Note: ***, ** and * indicate significance at the 1%, 5% and 10% significance level, respectively; and $LM_{(1)}$ and $LM_{(5)}$, correspond to Engle's LM test results for ARCH effects using 1 and 5 lags of the squared residuals of an ARFIMA, respectively.

Table C.2. Descriptive statistics (mean, standard deviation, minimum, maximum, skewness and kurtosis) and Engle's LM test for ARCH effects, for the log-volume, log absolute returns, log range estimator and log-realized variance for each stock series considered.

		Log Range								**Log Realised Variance**								
Ticker	Company	Mean	StdDev	Max	Min	Skew.	Kurt.	$LM_{(1)}$	$LM_{(5)}$	Mean	StdDev	Max	Min	Skew.	Kurt.	$LM_{(1)}$	$LM_{(5)}$	Obs.
AAPL	Apple	-8.51	1.06	-3.40	-12.65	0.15	3.35	5.69**	7.92	-8.01	1.04	-3.78	-11.34	0.39	3.61	27.76***	27.87***	3008
AXP	Amex	-8.91	1.29	-3.35	-12.81	0.69	3.55	2.70*	3.35	-8.33	1.25	-2.75	-11.41	0.77	3.49	9.46***	16.26***	3000
BA	Boeing Co	-8.89	0.96	-4.65	-11.69	0.45	3.59	0.04	2.74	-8.33	0.95	-3.18	-10.75	0.84	4.32	5.81**	8.34	3013
CAT	Caterpillar	-8.71	1.00	-3.81	-11.23	0.55	3.68	8.26***	10.25*	-8.14	1.01	-4.14	-10.77	0.80	3.83	29.20***	30.47***	3009
CSCO	Cisco Systems	-8.73	0.94	-4.61	-12.35	0.26	3.61	1.22	11.94**	-8.12	0.92	-3.79	-10.92	0.74	4.27	110.06***	110.81***	2976
CVX	Chevron	-9.15	0.97	-4.50	-11.94	0.62	4.36	25.95***	30.13***	-8.54	0.96	-4.02	-11.37	0.66	3.95	16.64***	24.03***	3009
DD	DuPont Co.	-8.99	1.00	-4.65	-11.91	0.59	3.71	4.87**	11.68**	-8.43	0.98	-4.21	-11.04	0.77	3.80	11.22***	11.69**	2998
DIS	Walt Disney Co.	-9.00	1.00	-4.65	-12.09	0.54	3.85	0.93	3.01	-8.40	0.98	-2.74	-10.82	0.90	4.37	35.89***	39.85***	2988
GE	General Electrics	-9.13	1.17	-3.67	-12.36	0.80	4.15	3.46*	22.30***	-8.47	1.12	-3.18	-11.11	1.04	4.56	38.61***	43.33***	2968
GS	Goldman Sachs	-8.68	1.10	-2.89	-11.77	0.94	4.81	33.66***	39.78***	-8.15	1.06	-2.81	-10.68	1.04	4.59	24.55***	37.32***	3013
HD	Home Depot	-8.87	1.03	-3.74	-12.13	0.57	3.73	1.32	6.04	-8.28	1.00	-3.63	-10.84	0.80	3.93	26.33***	24.40***	2994
IBM	IBM	-9.44	0.95	-5.27	-12.90	0.67	4.37	9.22***	17.09***	-8.80	0.94	-3.64	-11.22	0.94	4.52	14.19***	20.07***	3011
INTC	Intel Corporation	-8.71	0.90	-5.03	-11.77	0.31	3.49	3.08**	19.55***	-8.08	0.90	-3.96	-10.77	0.69	3.83	36.79***	39.49***	2982
JNJ	Johnson & Johnson	-9.83	0.96	-4.85	-12.98	0.54	3.82	2.51	19.80***	-9.18	0.92	-4.75	-11.68	0.81	4.21	9.99***	18.23***	2991
JPM	JPMorgan Chase & Co.	-8.73	1.21	-3.80	-12.10	0.69	3.67	5.84**	10.67*	-8.17	1.18	-3.65	-10.96	0.82	3.60	22.66***	29.72***	3001
KO	The Coca-Cola Co.	-9.62	0.92	-5.09	-13.29	0.64	4.31	3.42*	15.68***	-9.00	0.91	-4.09	-11.64	1.00	4.98	6.90***	8.38	2980
MCD	McDonald's Corporation	-9.38	1.03	-3.67	-12.52	0.36	3.44	5.26**	13.32**	-8.75	1.00	-3.30	-11.59	0.68	4.12	14.50***	17.97***	2998
MMM	3M Co.	-9.39	0.96	-3.20	-12.20	0.58	4.31	1.07	4.38	-8.79	0.96	-4.17	-11.28	0.82	4.26	6.91***	9.67*	2999
MRK	Merck & Co., Inc.	-9.02	1.02	-4.15	-12.13	0.58	3.99	1.56	2.25	-8.45	0.95	-3.79	-10.82	0.92	4.51	5.63**	6.48	2999
MSFT	Microsoft Corporation	-9.06	0.95	-4.29	-12.01	0.62	3.98	4.81**	7.50	-8.46	0.92	-4.25	-10.88	0.84	4.13	25.01***	25.26***	2978
NKE	Nike	-9.02	0.99	-4.47	-11.59	0.63	3.81	3.06	4.08	-8.48	0.96	-3.93	-10.86	0.97	4.34	53.62***	63.74***	3004
PFE	Pfizer Inc.	-9.10	0.92	-4.31	-11.91	0.60	4.25	1.03	2.99	-8.44	0.87	-3.42	-10.77	0.90	4.83	2.12	5.37	2976
PG	Procter & Gamble Co.	-9.72	0.95	-2.23	-12.88	0.78	5.62	2.82*	16.19***	-9.07	0.88	-4.66	-11.39	0.99	4.86	4.87**	7.02	2989
TRV	Travelers Companies Inc	-9.08	1.27	-3.47	-11.98	0.86	3.88	13.77***	20.93***	-8.47	1.25	-3.40	-11.26	0.81	3.63	39.47***	44.31***	1963
UNH	UnitedHealth Group Inc	-8.68	1.07	-3.33	-12.01	0.69	3.83	7.45***	10.43**	-8.16	1.03	-3.69	-10.63	0.88	3.96	8.23***	8.87	3002
UTX	United Technologies Corporation	-9.21	0.95	-4.50	-12.02	0.46	3.91	1.94	4.79	-8.61	0.94	-3.74	-11.01	0.93	4.65	28.97***	37.35***	3008
VZ	Verizon Communications Inc.	-9.20	0.99	-4.17	-11.99	0.58	3.78	0.45	3.33	-8.61	0.95	-3.99	-11.28	0.75	4.09	21.79***	33.75***	2595
V	Visa Inc	-8.97	1.14	-4.67	-12.26	0.63	3.48	17.82***	23.87***	-8.26	1.05	-4.04	-11.37	0.68	3.30	12.72***	19.08***	2990
WMT	Wal-Mart Stores Inc	-9.46	0.93	-4.64	-12.74	0.46	4.04	5.04	7.43	-8.83	0.93	-3.82	-11.33	0.84	4.21	12.36***	28.28***	3000
XOM	Exxon Mobil Corporation	-9.28	0.96	-4.43	-12.04	0.58	4.34	1.80	10.45*	-8.64	0.92	-4.05	-11.10	0.85	4.67	132.51***	141.60***	3001

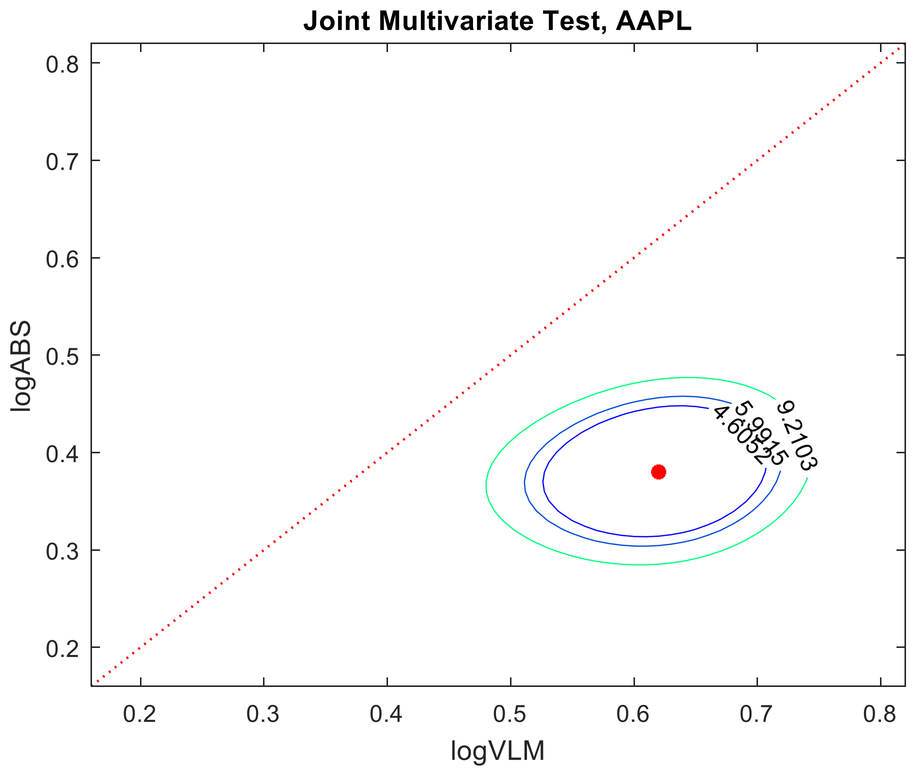
Note: ***, ** and * indicate significance at the 1%, 5% and 10% significance level, respectively; and $LM_{(1)}$ and $LM_{(5)}$, correspond to Engle's LM test results for ARCH effects using 1 and 5 lags of the squared residuals of an ARFIMA, respectively.

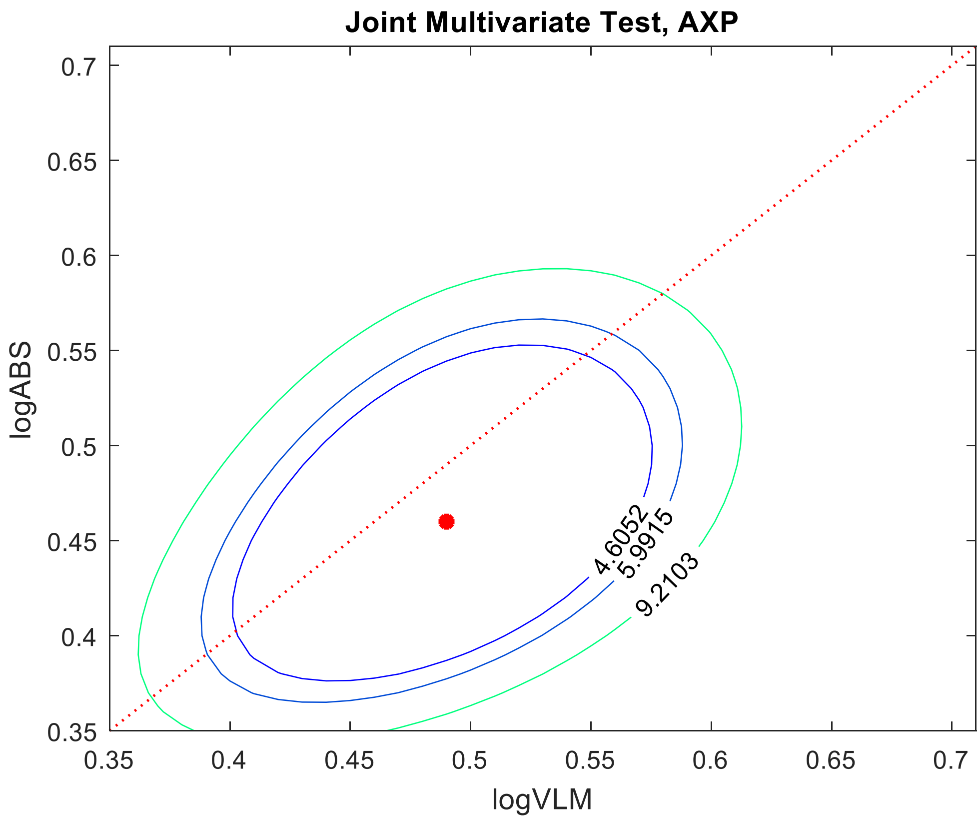
Table C.3. Robustness checks in the joint analysis on log-volume and log-realised variance against the choice of p and the inclusion of a time trend in volatility. Auxiliary regressions are augmented with p lags, with either p=*2* or p determined according to Schwert's rule.

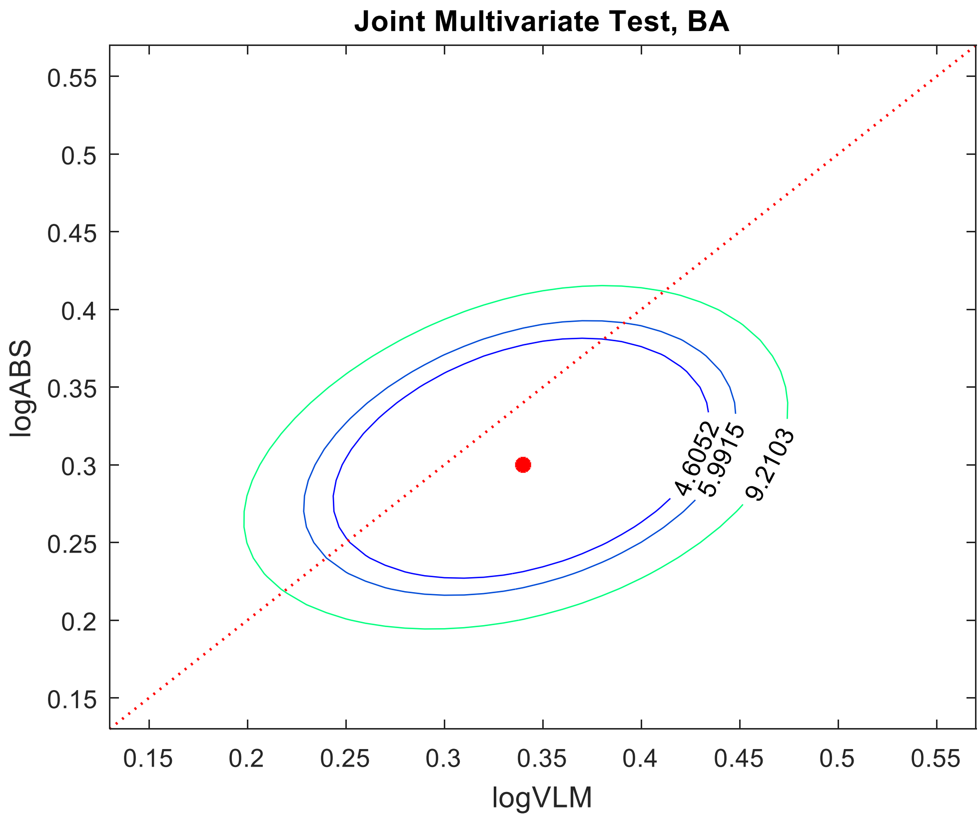
	$p = 2$, no linear trend				$p = 2$, linear trend				Schwert's rule, linear trend			
Stock	95 % CI VLM	95 % CIB RV	Common d	$\widehat{\rho}_e$	95 % CIB VLM	95 % CIB RV	Common d	$\widehat{\rho}_e$	95 % CIB	95 % CIB	Common d	$\widehat{\rho}_e$
AAPL	[0.49,0.57]	[0.45,0.53]	[0.49,0.53]	0.62	[0.49,0.57]	[0.43,0.51]	-	0.62	[0.56,0.73]	[0.51,0.70]	[0.56,0.70]	0.62
AXP	[0.38,0.47]	[0.44,0.52]	[0.45,0.47]	0.48	[0.38,0.47]	[0.44,0.51]	[0.45,0.47]	0.48	[0.45,0.60]	[0.50,0.70]	[0.52,0.57]	0.49
BA	[0.27,0.40]	[0.40,0.49]	-	0.52	[0.27,0.39]	[0.39,0.48]	-	0.52	[0.29,0.49]	[0.43,0.67]	-	0.52
CAT	[0.34,0.46]	[0.41,0.48]	[0.42,0.46]	0.58	[0.34,0.45]	[0.41,0.48]	[0.42,0.45]	0.58	[0.36,0.59]	[0.50,0.69]	-	0.58
CSCO	[0.28,0.40]	[0.38,0.46]	-	0.53	[0.27,0.39]	[0.36,0.44]	[0.38,0.39]	0.53	[0.30,0.49]	[0.45,0.63]	-	0.53
CVX	[0.38,0.48]	[0.43,0.51]	[0.43,0.48]	0.43	[0.38,0.48]	[0.42,0.51]	[0.43,0.48]	0.43	[0.41,0.59]	[0.47,0.68]	[0.47,0.58]	0.43
DD	[0.39,0.47]	[0.30,0.41]	-	0.45	[0.30,0.41]	[0.39,0.47]	-	0.45	[0.29,0.49]	[0.45,0.66]	-	0.46
DIS	[0.28,0.39]	[0.40,0.49]	-	0.42	[0.27,0.39]	[0.39,0.48]	-	0.42	[0.26,0.48]	[0.46,0.67]	-	0.42
GE	[0.38,0.47]	[0.43,0.52]	[0.45,0.46]	0.50	[0.37,0.47]	[0.43,0.52]	[0.44,0.46]	0.50	[0.42,0.60]	[0.46,0.67]	[0.46,0.59]	0.51
GS	[0.41,0.51]	[0.43,0.50]	[0.43,0.50]	0.58	[0.41,0.51]	[0.42,0.50]	[0.42,0.49]	0.58	[0.46,0.64]	[0.47,0.67]	[0.48,0.64]	0.58
HD	[0.36,0.47]	[0.42,0.49]	[0.43,0.47]	0.46	[0.35,0.46]	[0.40,0.48]	[0.41,0.46]	0.47	[0.45,0.62]	[0.48,0.68]	[0.49,0.62]	0.47
IBM	[0.31,0.41]	[0.39,0.47]	-	0.46	[0.30,0.41]	[0.38,0.47]	-	0.46	[0.30,0.49]	[0.42,0.64]	-	0.47
INTC	[0.23,0.38]	[0.38,0.46]	-	0.50	[0.23,0.38]	[0.37,0.45]	-	0.50	[0.12,0.40]	[0.47,0.66]	-	0.50
JNJ	[0.31,0.42]	[0.41,0.50]	-	0.44	[0.30,0.42]	[0.40,0.49]	-	0.44	[0.35,0.54]	[0.45,0.71]	-	0.44
JPM	[0.41,0.49]	[0.44,0.51]	[0.44,0.49]	0.55	[0.40,0.49]	[0.43,0.51]	[0.44,0.49]	0.55	[0.50,0.65]	[0.51,0.69]	[0.51,0.65]	0.56
KO	[0.32,0.44]	[0.40,0.48]	[0.41,0.44]	0.46	[0.32,0.44]	[0.39,0.47]	[0.40,0.44]	0.46	[0.33,0.57]	[0.51,0.68]	-	0.47
MCD	[0.28,0.41]	[0.38,0.46]	-	0.45	[0.26,0.40]	[0.34,0.42]	[0.36,0.40]	0.45	[0.28,0.53]	[0.45,0.67]	-	0.46
MMM	[0.26,0.40]	[0.37,0.45]	-	0.46	[0.26,0.40]	[0.37,0.45]	-	0.46	[0.23,0.50]	[0.44,0.65]	-	0.47
MRK	[0.31,0.44]	[0.38,0.46]	[0.38,0.44]	0.50	[0.31,0.44]	[0.37,0.45]	[0.37,0.44]	0.50	[0.27,0.55]	[0.44,0.63]	[0.46,0.54]	0.50
MSFT	[0.22,0.39]	[0.39,0.47]	-	0.50	[0.22,0.39]	[0.39,0.46]	-	0.50	[0.41,0.61]	[0.05,0.43]	-	0.50
NKE	[0.28,0.39]	[0.38,0.46]	-	0.43	[0.28,0.39]	[0.38,0.46]	-	0.43	[0.28,0.50]	[0.45,0.63]	-	0.44
PFE	[0.30,0.43]	[0.39,0.48]	[0.42,0.43]	0.45	[0.30,0.43]	[0.39,0.48]	[0.41,0.43]	0.45	[0.19,0.51]	[0.40,0.66]	-	0.45
PG	[0.28,0.42]	[0.38,0.47]	[0.40,0.41]	0.43	[0.28,0.42]	[0.38,0.47]	[0.39,0.41]	0.43	[0.24,0.52]	[0.39,0.60]	[0.41,0.52]	0.43
TRV	[0.35,0.49]	[0.45,0.56]	-	0.43	[0.34,0.48]	[0.41,0.54]	[0.43,0.47]	0.43	[0.38,0.73]	[0.33,0.50]	[0.59,0.70]	0.43
UNH	[0.28,0.42]	[0.40,0.48]	-	0.52	[0.28,0.42]	[0.40,0.48]	-	0.52	[0.22,0.53]	[0.52,0.70]	-	0.52
UTX	[0.31,0.42]	[0.41,0.49]	-	0.44	[0.31,0.42]	[0.40,0.49]	-	0.44	[0.34,0.53]	[0.46,0.66]	-	0.44
V	[0.51,0.60]	[0.43,0.51]	-	0.44	[0.50,0.59]	[0.41,0.49]	-	0.44	[0.61,0.82]	[0.49,0.70]	-	0.44
VZ	[0.32,0.46]	[0.44,0.52]	-	0.41	[0.32,0.46]	[0.42,0.51]	-	0.42	[0.31,0.58]	[0.47,0.68]	[0.48,0.57]	0.42
WMT	[0.28,0.40]	[0.36,0.44]	-	0.51	[0.27,0.39]	[0.34,0.42]	[0.36,0.39]	0.51	[0.29,0.54]	[0.42,0.62]	[0.44,0.53]	0.51
XOM	[0.33,0.45]	[0.38,0.48]	[0.39,0.45]	0.45	[0.33,0.45]	[0.38,0.48]	[0.38,0.45]	0.45	[0.36,0.59]	[0.37,0.64]	[0.38,0.59]	0.46
Average				0.48				0.48				0.48
Rejection Rate			56.67%				46.67%				56.67%	

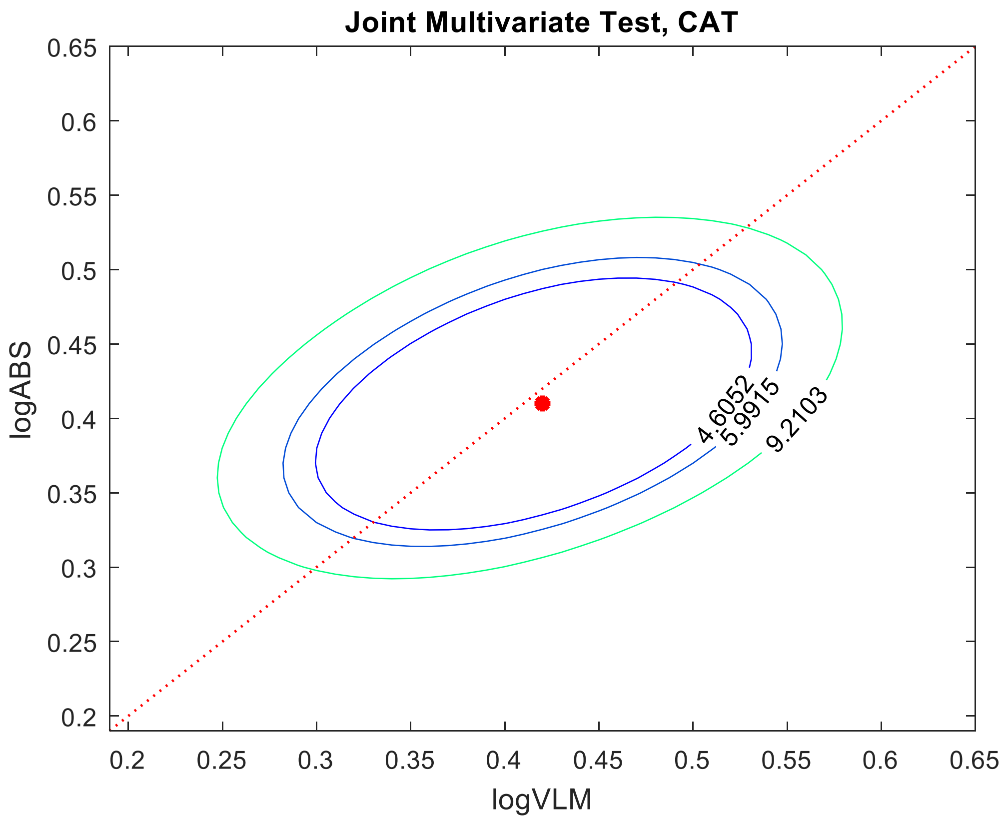
C.2.2 Confidence ellipsoids long memory coefficients for volatility and trading volume. Volatility proxy (σ): **absolute returns**.

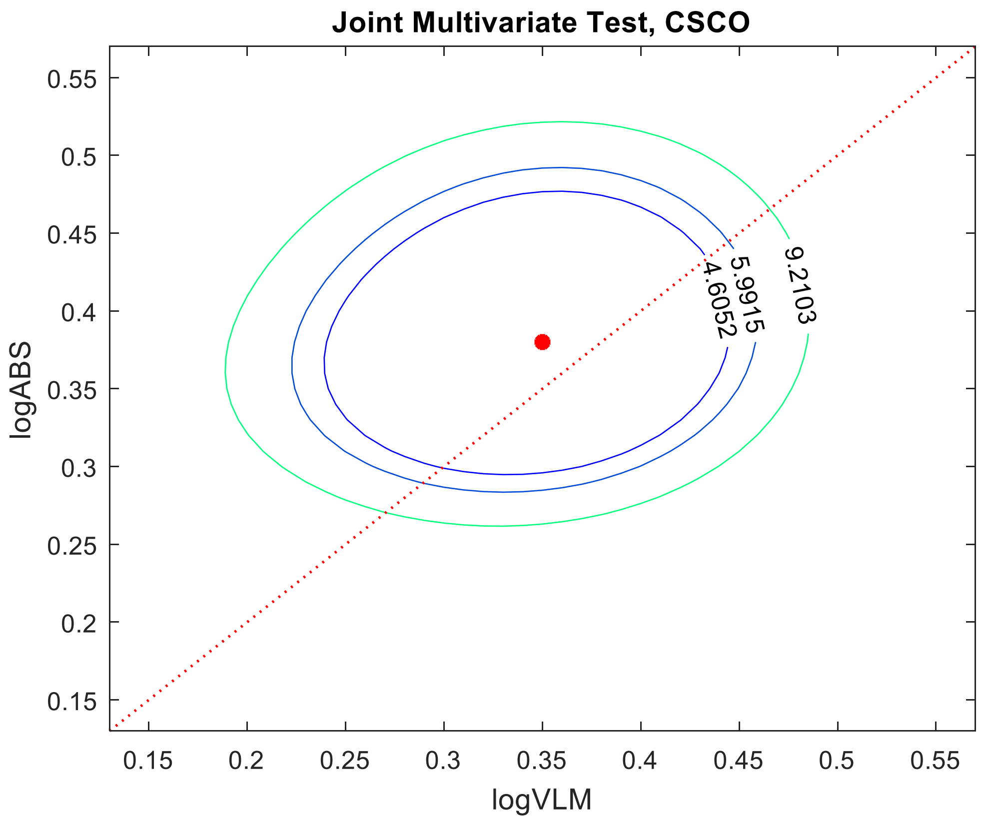
The figures that follow present 90%, 95%, and 99% confidence sets for the long memory parameters of log-trading volume and log-volatility of the stocks under analysis. The confidence sets are obtained from empirical level curves of the $LM_{\boldsymbol{d}}^{FGLS}$ test statistic evaluated at different values given the sample observations, for 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 $\hat{d}_{min}(\sigma)$. The red dashed line represents the 45-degree line.

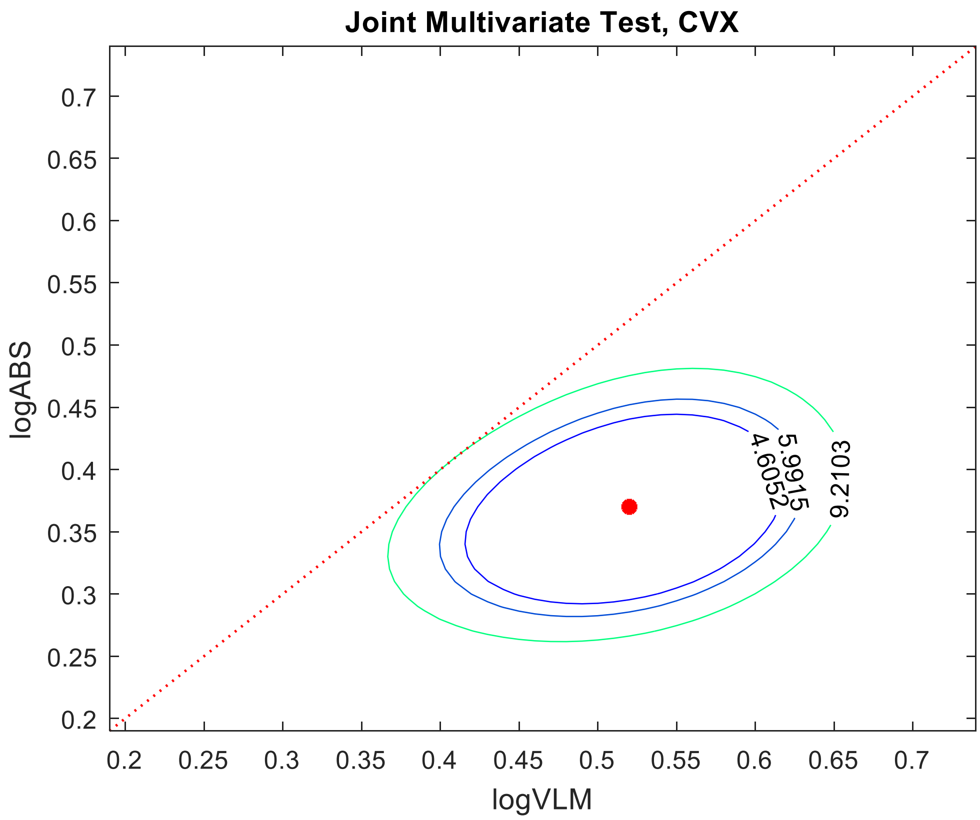


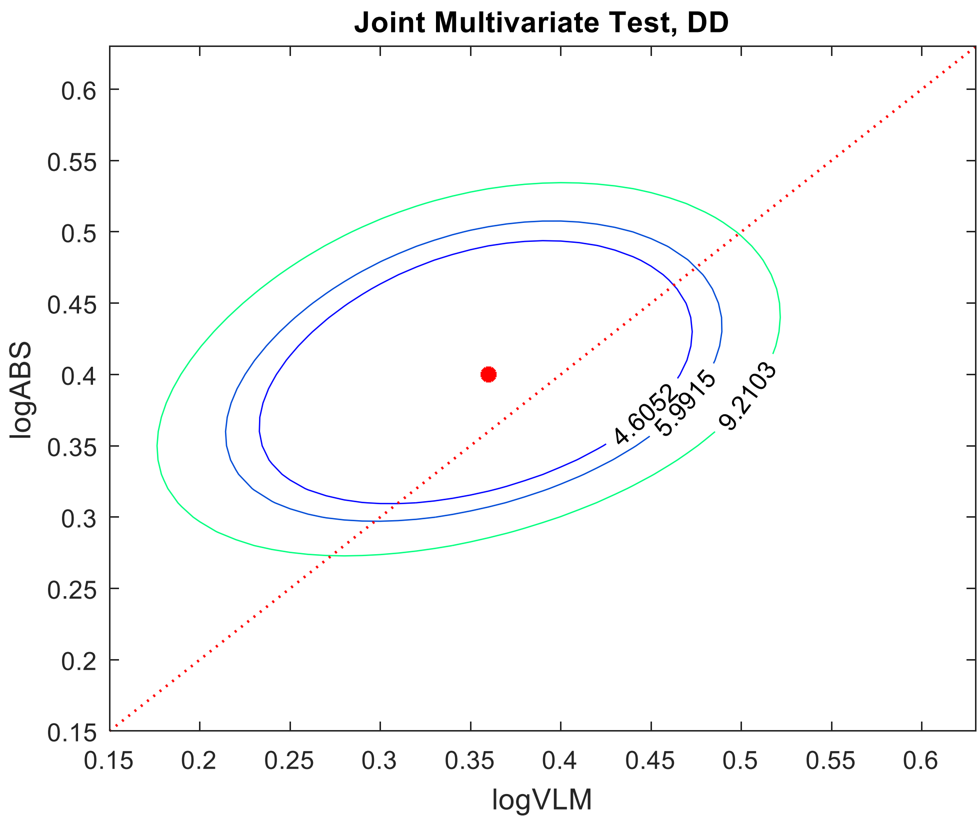

Joint Multivariate Test, AXP
logABS
logVLM
4.6052
5.9915
9.2103
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7

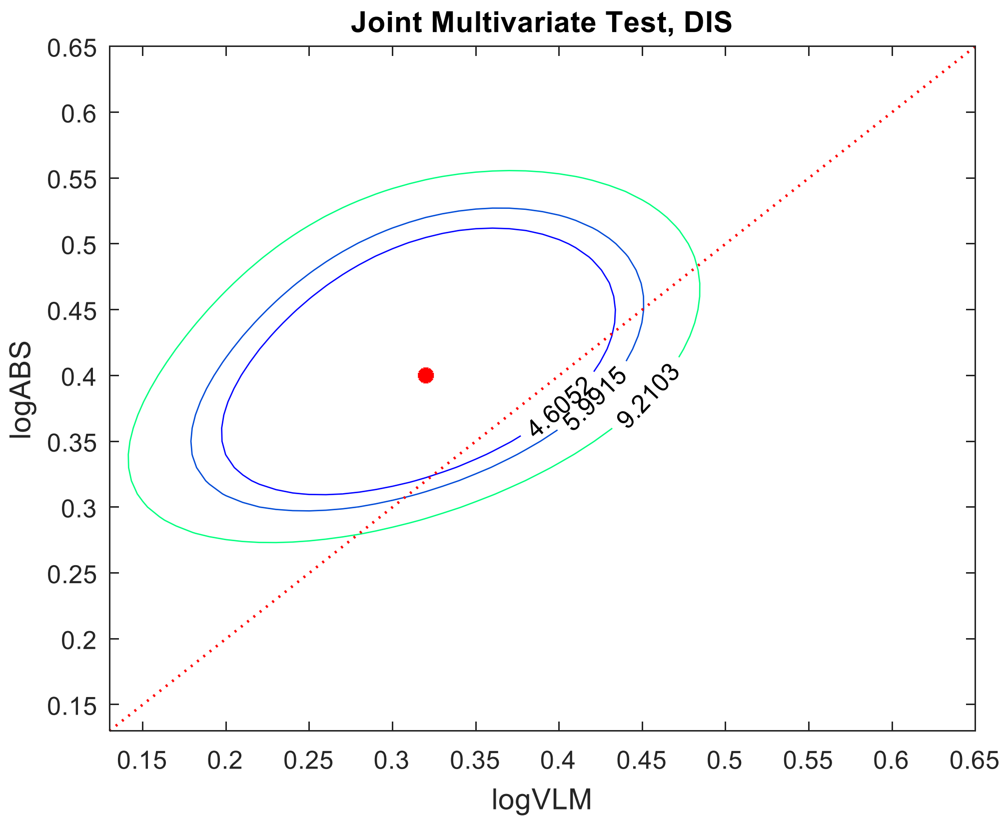

Joint Multivariate Test, BA
logABS
logVLM
4.6052
5.9915
9.2103
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55

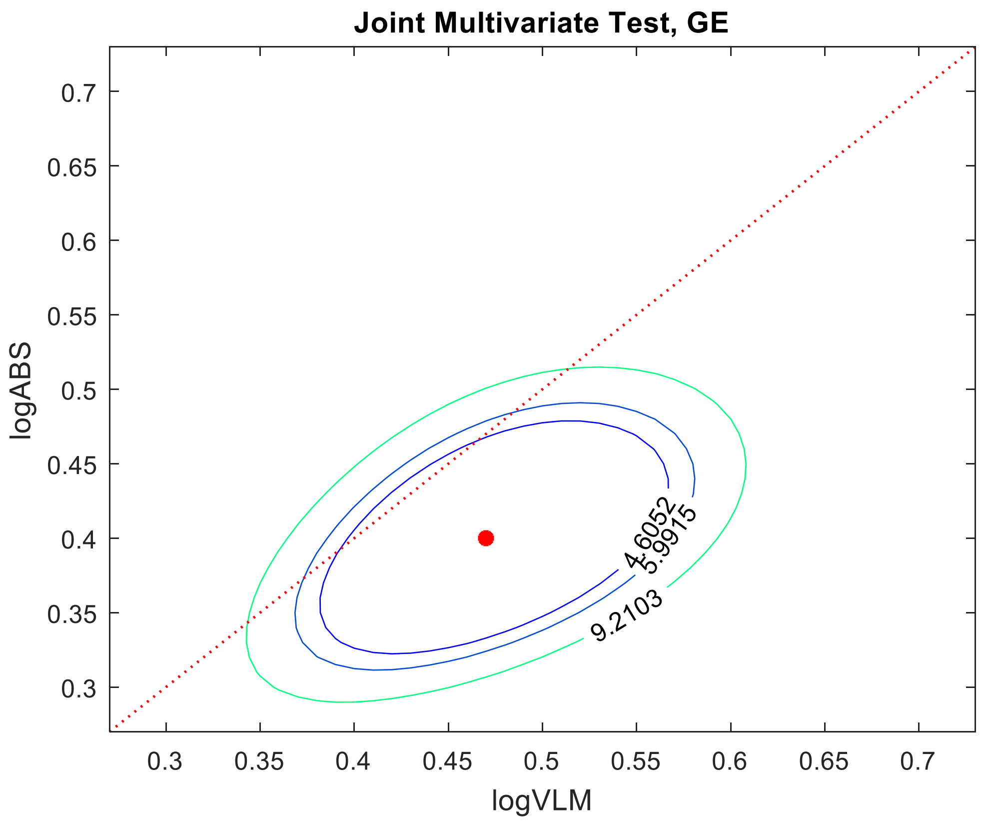

Joint Multivariate Test, CAT
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
logABS
4.6052
5.9915
9.2103
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
logVLM

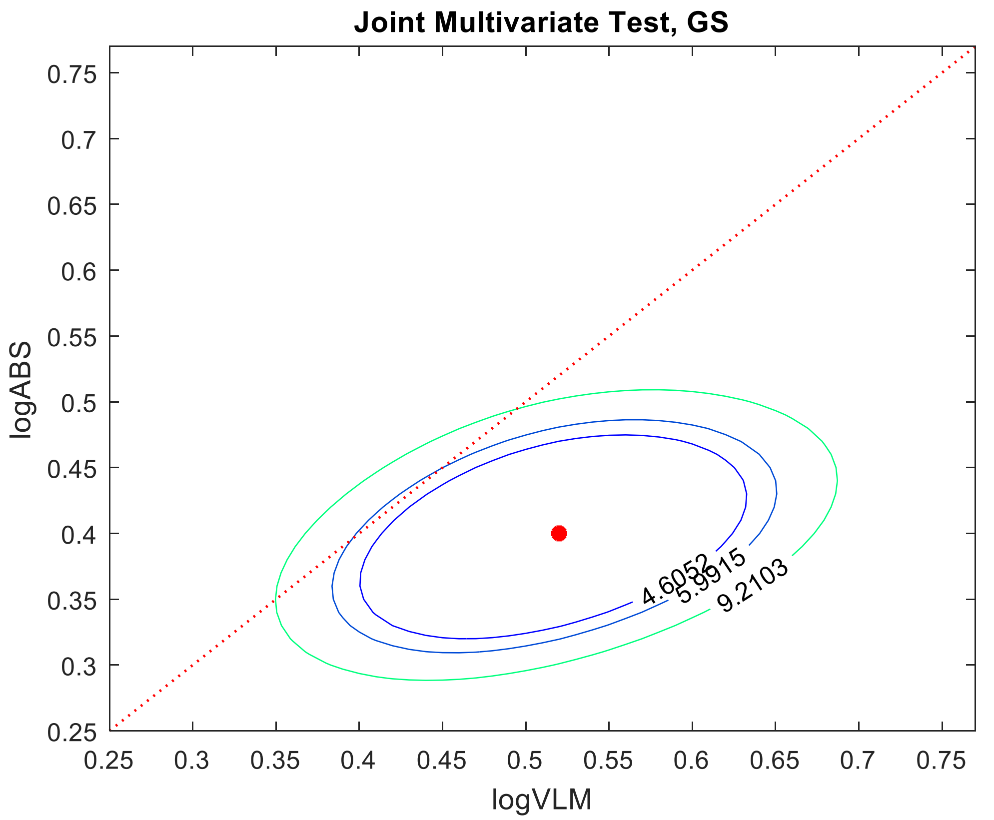

Joint Multivariate Test, CSCO
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
logABS
4.6052
5.9915
9.2103
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
logVLM

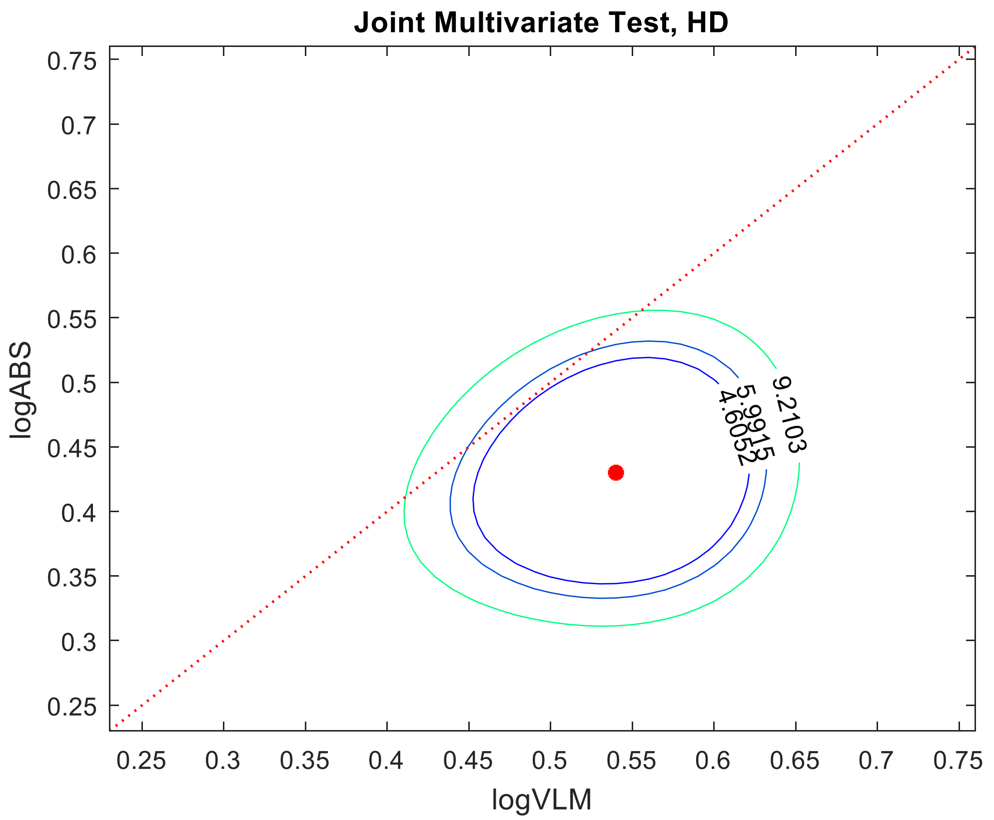

Joint Multivariate Test, CVX
logABS
logVLM
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
4.6052
5.9915
9.2103

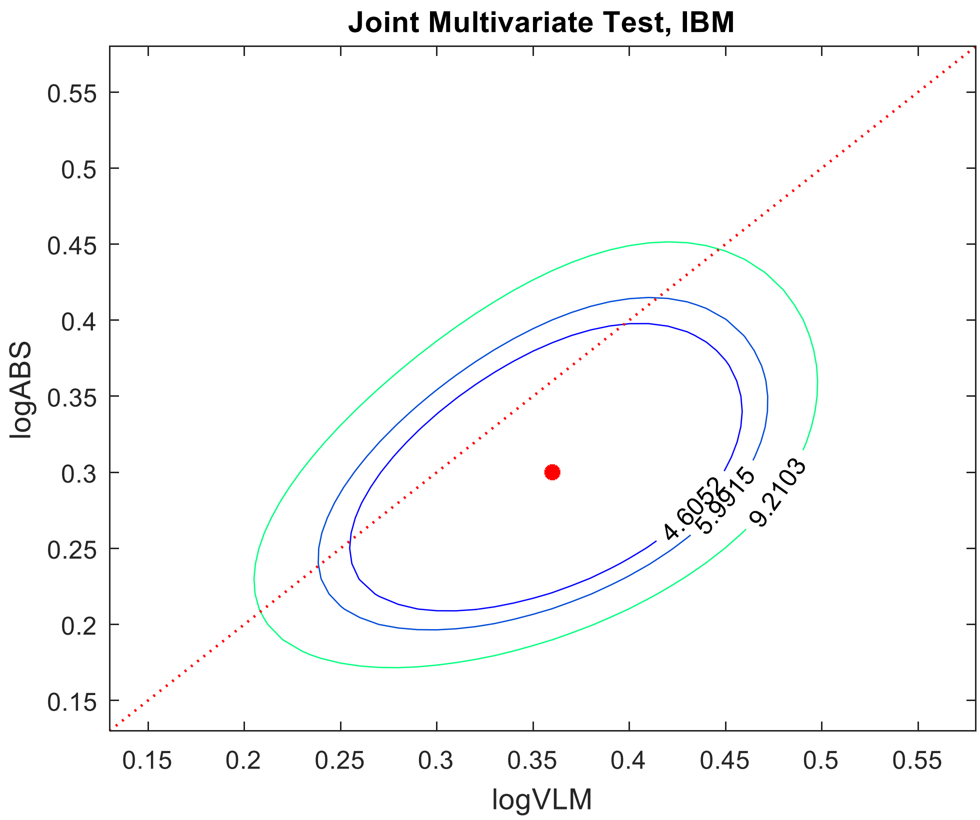

Joint Multivariate Test, DD
logABS
logVLM
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
4.6052
5.9915
9.2103

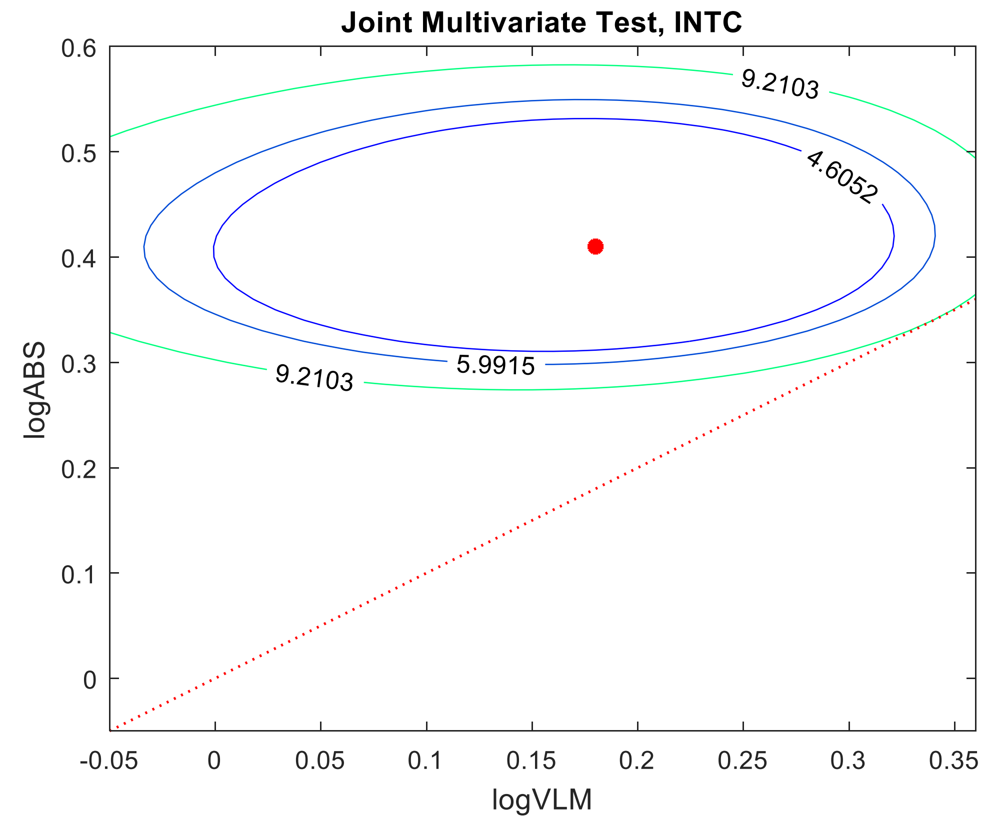

Joint Multivariate Test, DIS
logABS
logVLM
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
4.6052
5.9915
9.2103

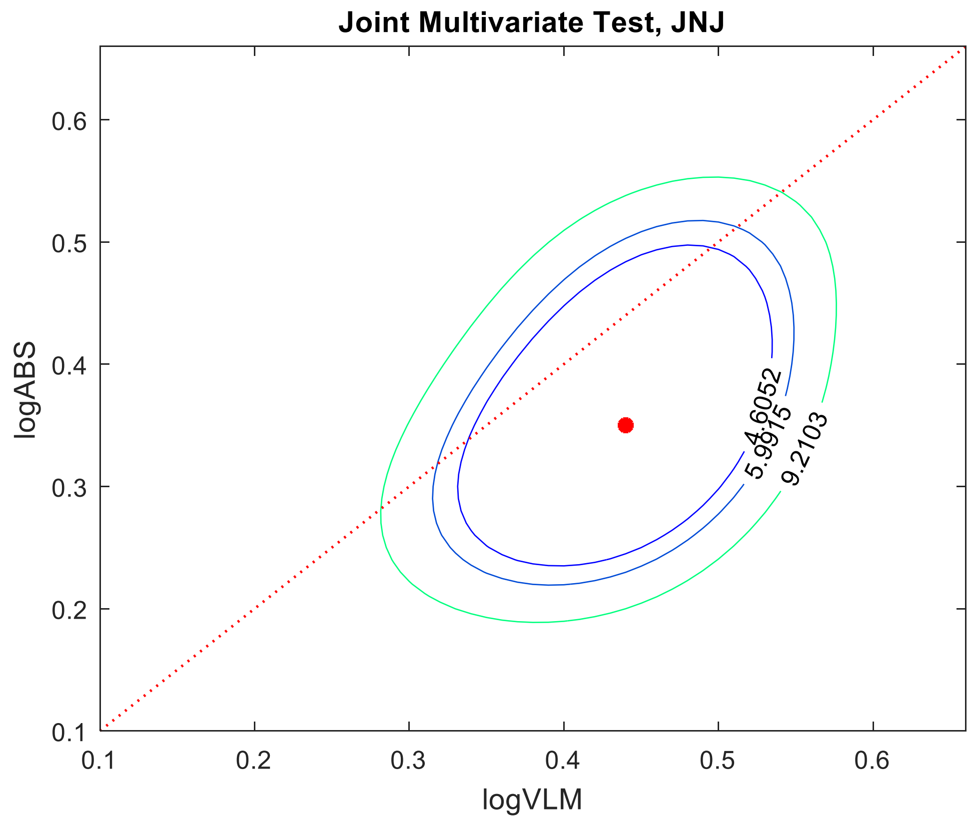

Joint Multivariate Test, GE
logABS
logVLM
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
4.6052
5.9915
9.2103


Joint Multivariate Test, GS
logABS
logVLM
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.25 0.3 0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7 0.75
4.6052
5.9915
9.2103

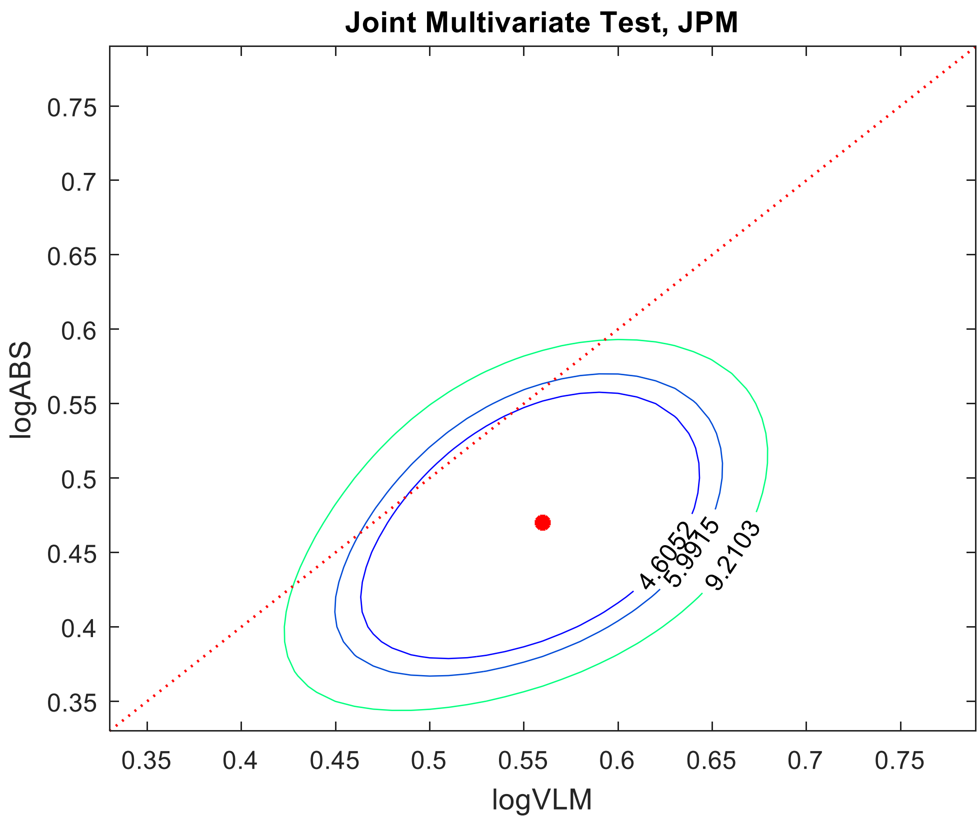

Joint Multivariate Test, HD
logABS
logVLM
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.25 0.3 0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7 0.75
4.6052
5.9915
9.2103


Joint Multivariate Test, IBM
logABS
logVLM
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
4.6052
5.9915
9.2103

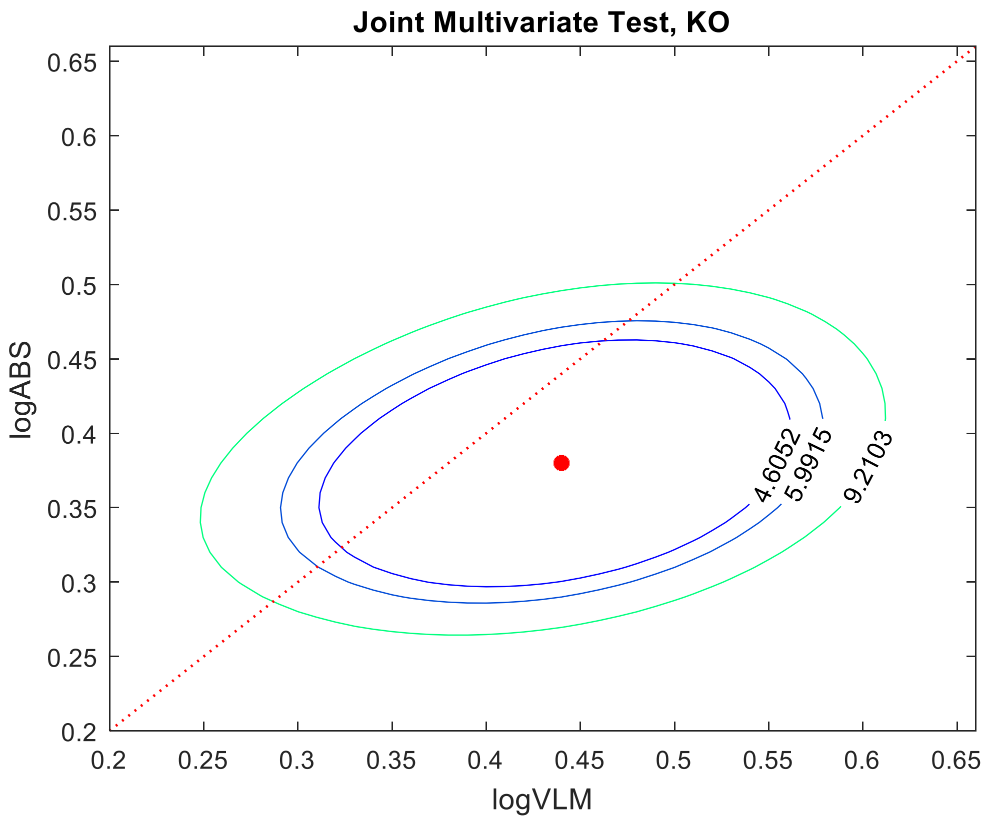

Joint Multivariate Test, INTC
logABS
logVLM
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.05
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
9.2103
4.6052
5.9915
9.2103

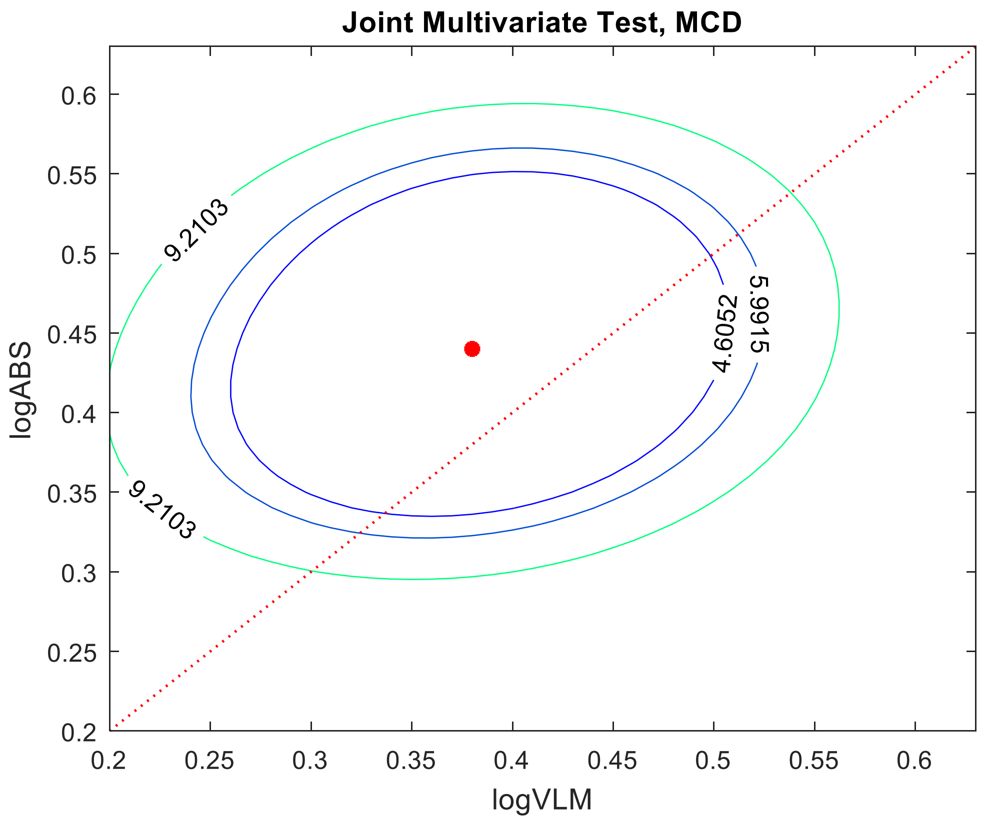


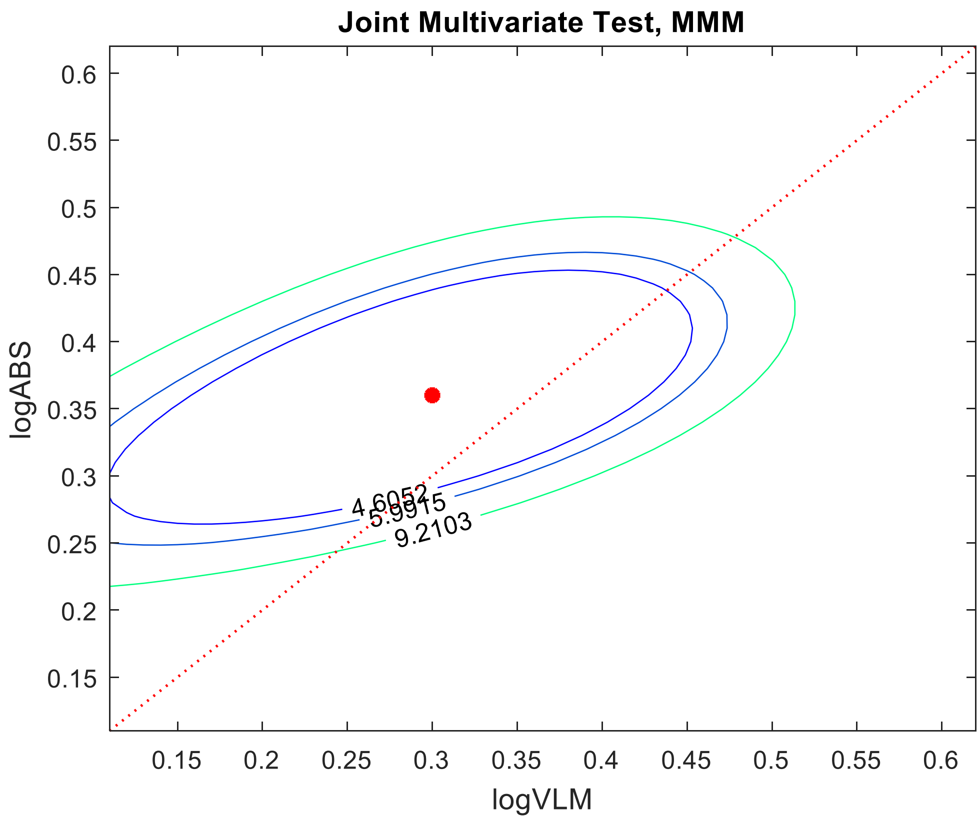
Joint Multivariate Test, JNJ
logABS
logVLM
0.1
0.2
0.3
0.4
0.5
0.6
4.6052
5.9915
9.2103

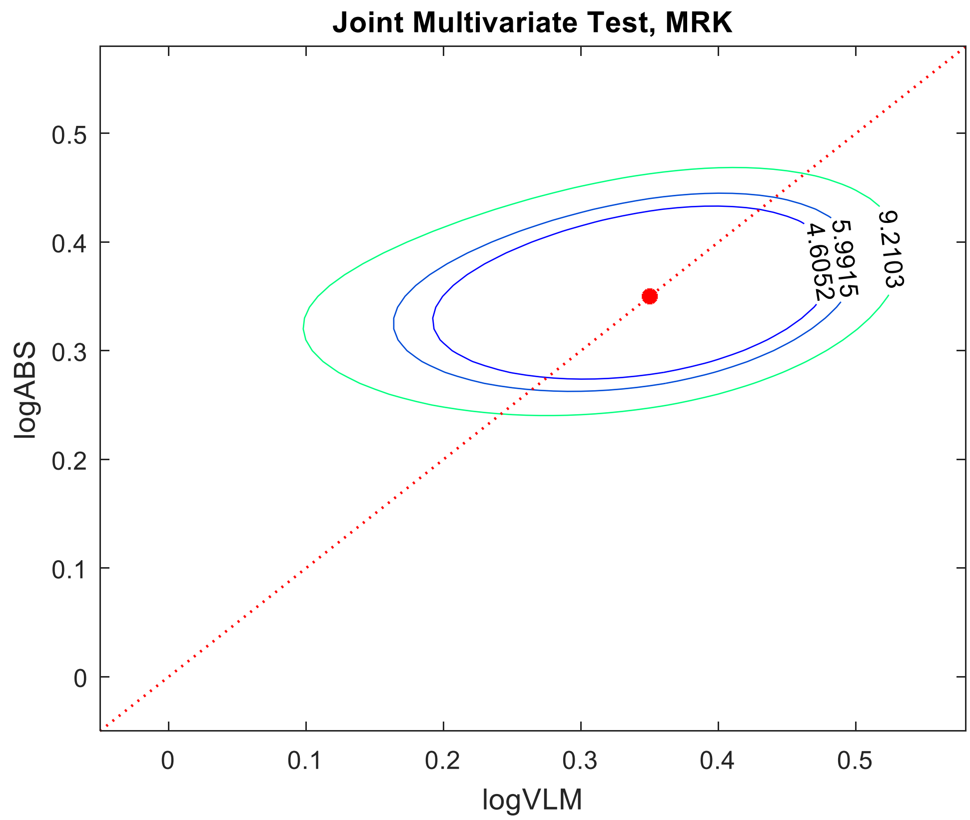


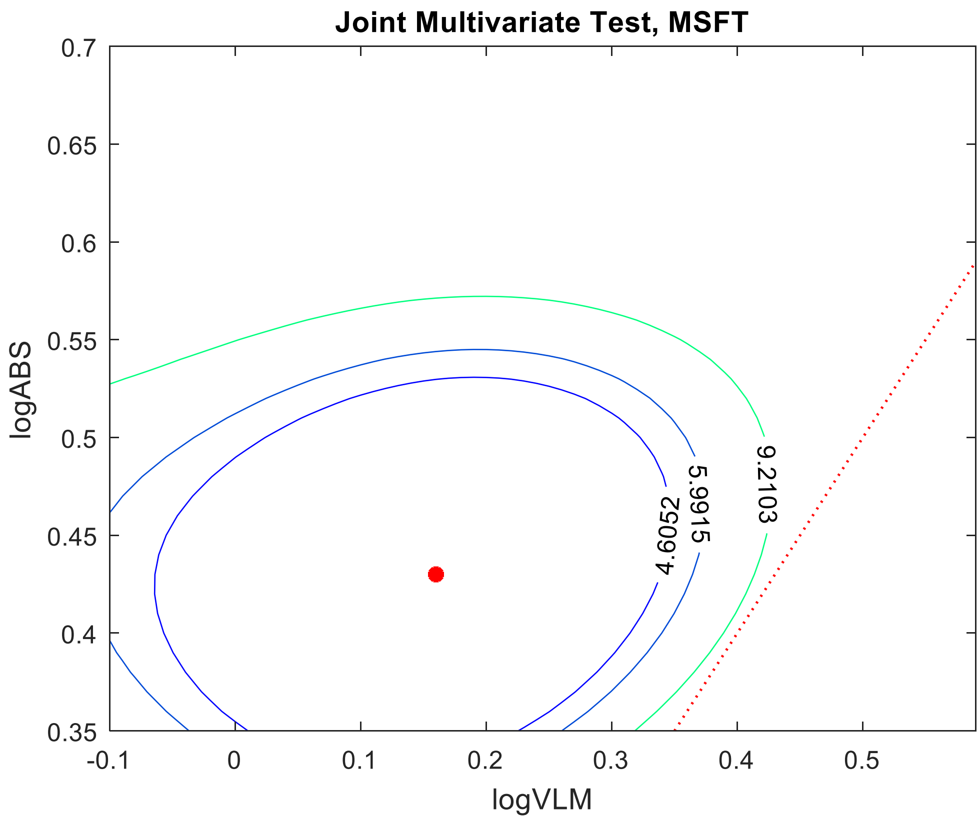
Joint Multivariate Test, JPM
logABS
logVLM
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
4.6052
5.9915
9.2103

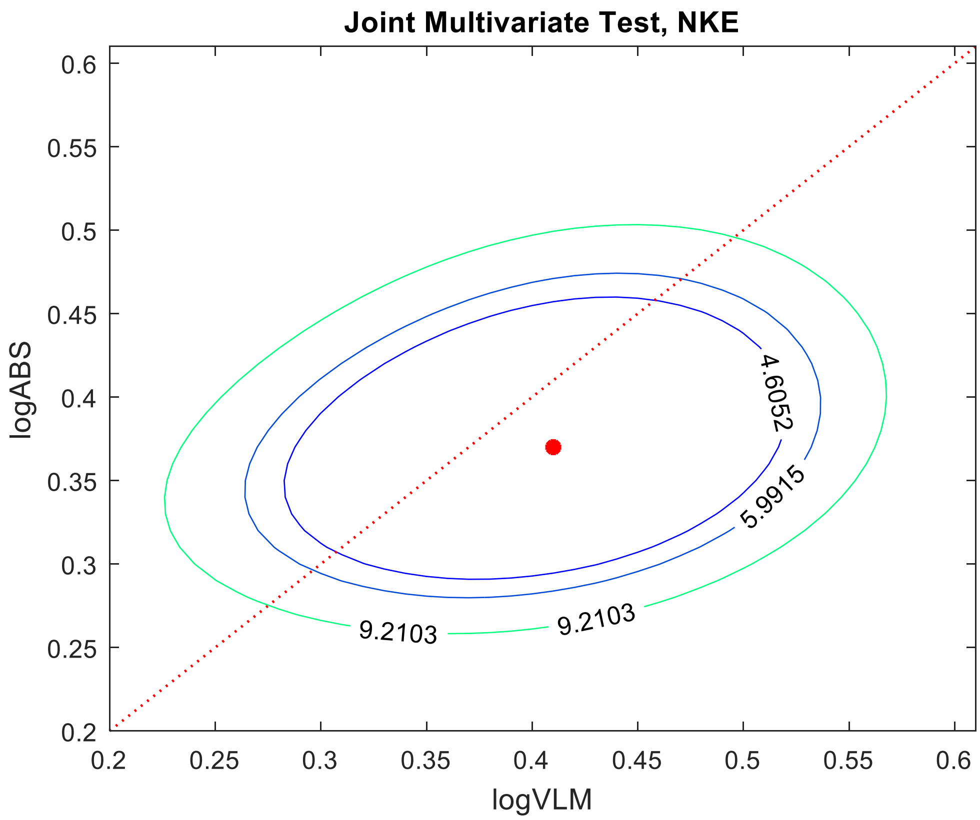

Joint Multivariate Test, KO
logABS
logVLM
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
4.6052
5.9915
9.2103

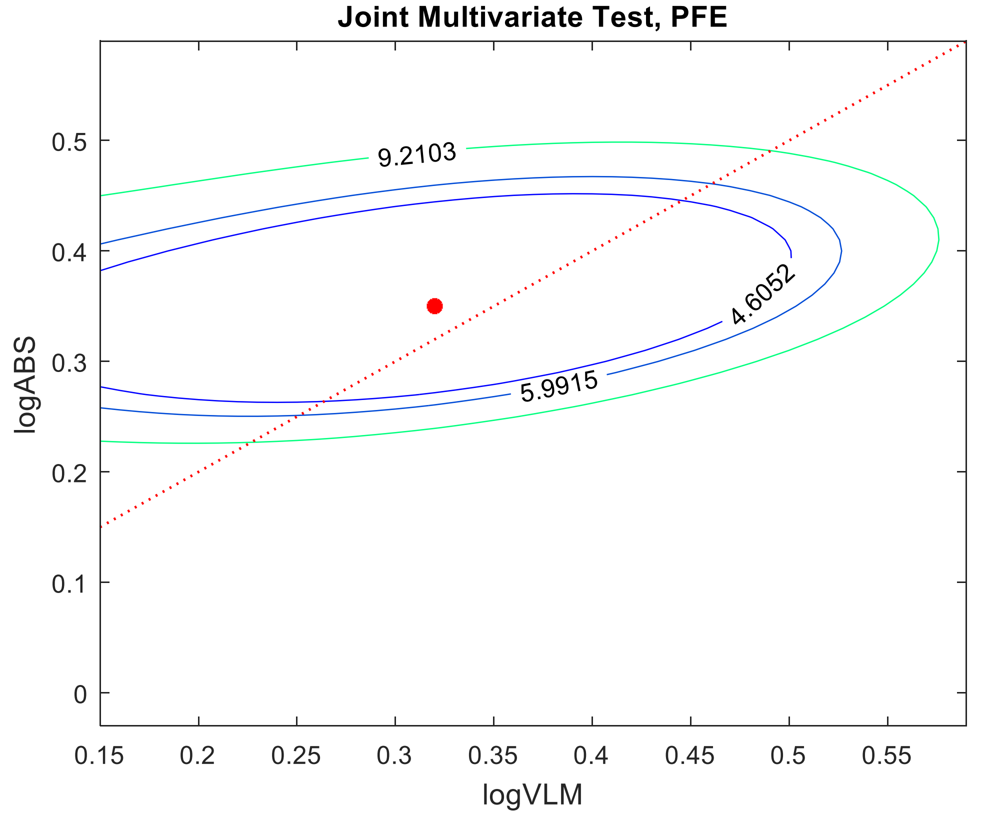

Joint Multivariate Test, MCD
logABS
logVLM
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
9.2103
9.2103
4.6052
5.9915


Joint Multivariate Test, MMM
logABS
logVLM
4.6052
5.9915
9.2103

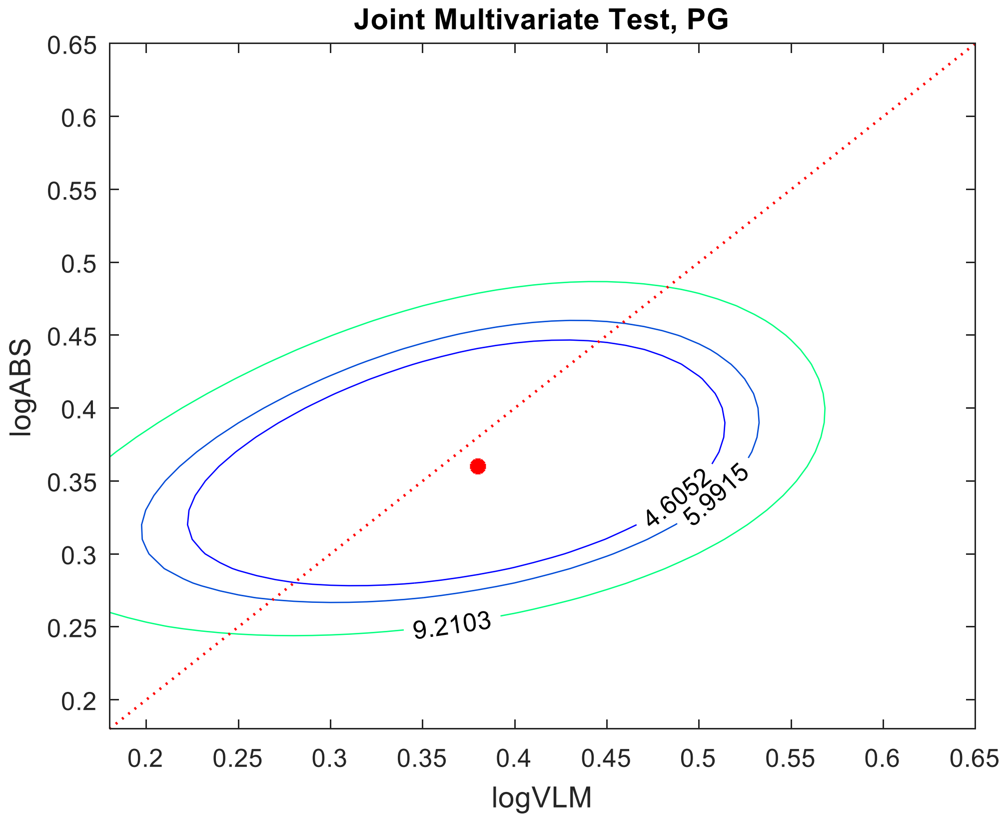

Joint Multivariate Test, MRK
logABS
logVLM
9.2103
5.9915
4.6052


Joint Multivariate Test, MSFT
logABS
logVLM
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
-0.1
0
0.1
0.2
0.3
0.4
0.5
4.6052
5.9915
9.2103

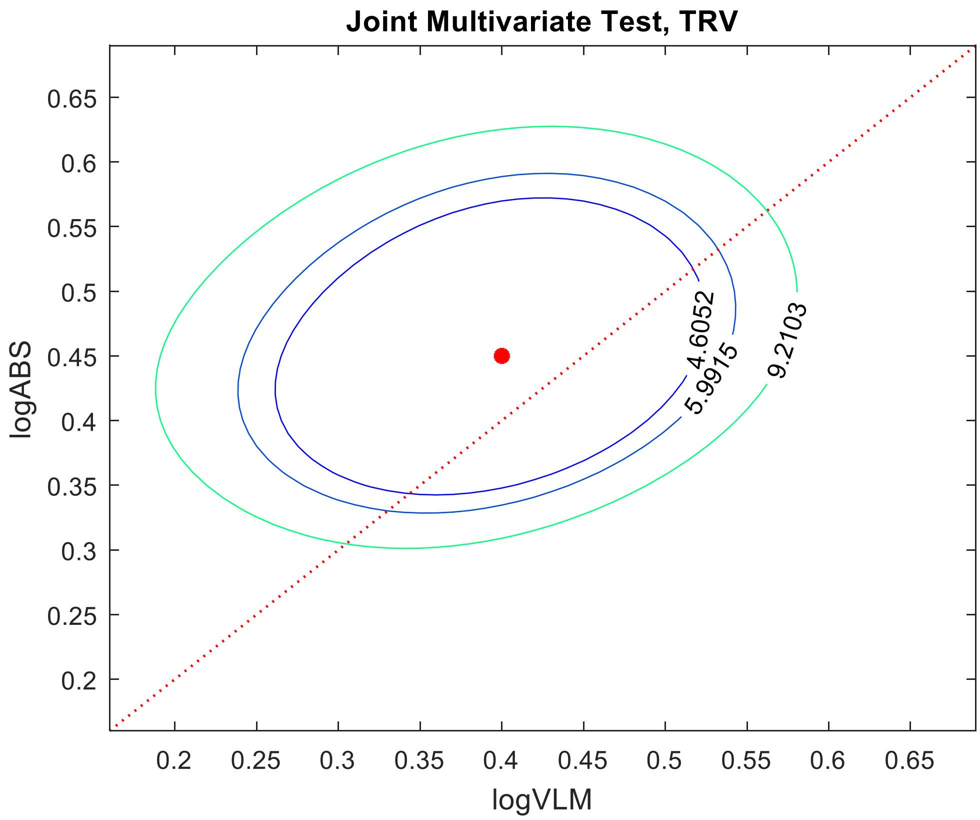

Joint Multivariate Test, NKE
logABS
logVLM
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
4.6052
5.9915
9.2103
9.2103

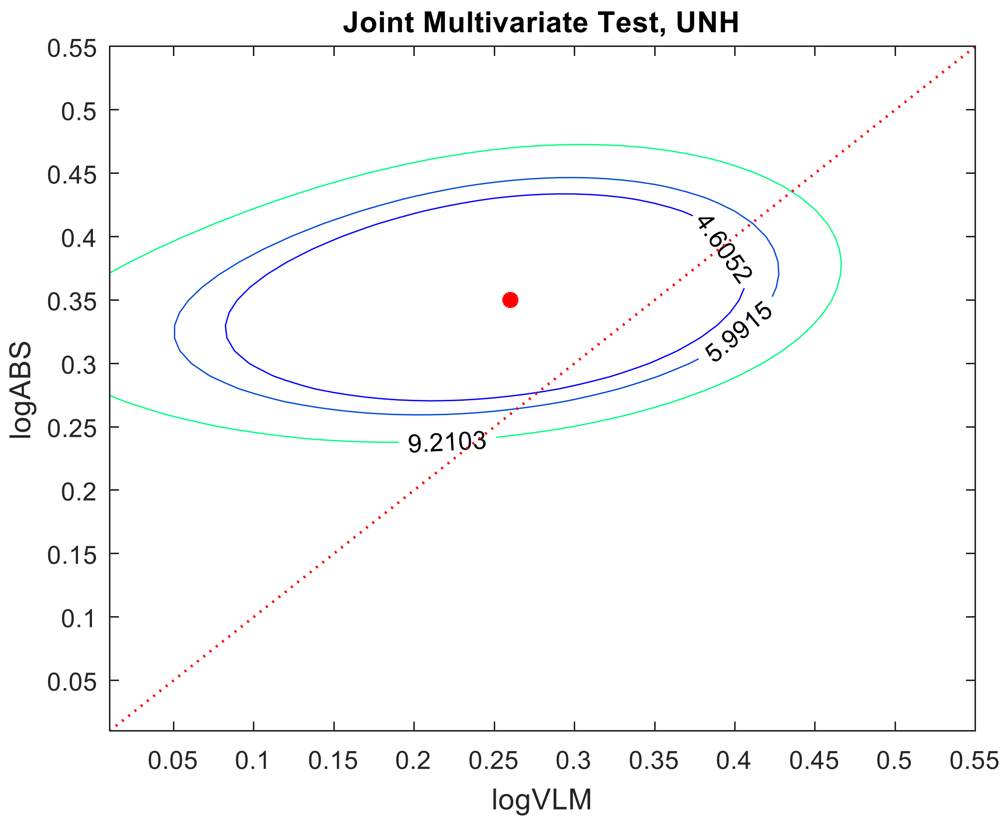


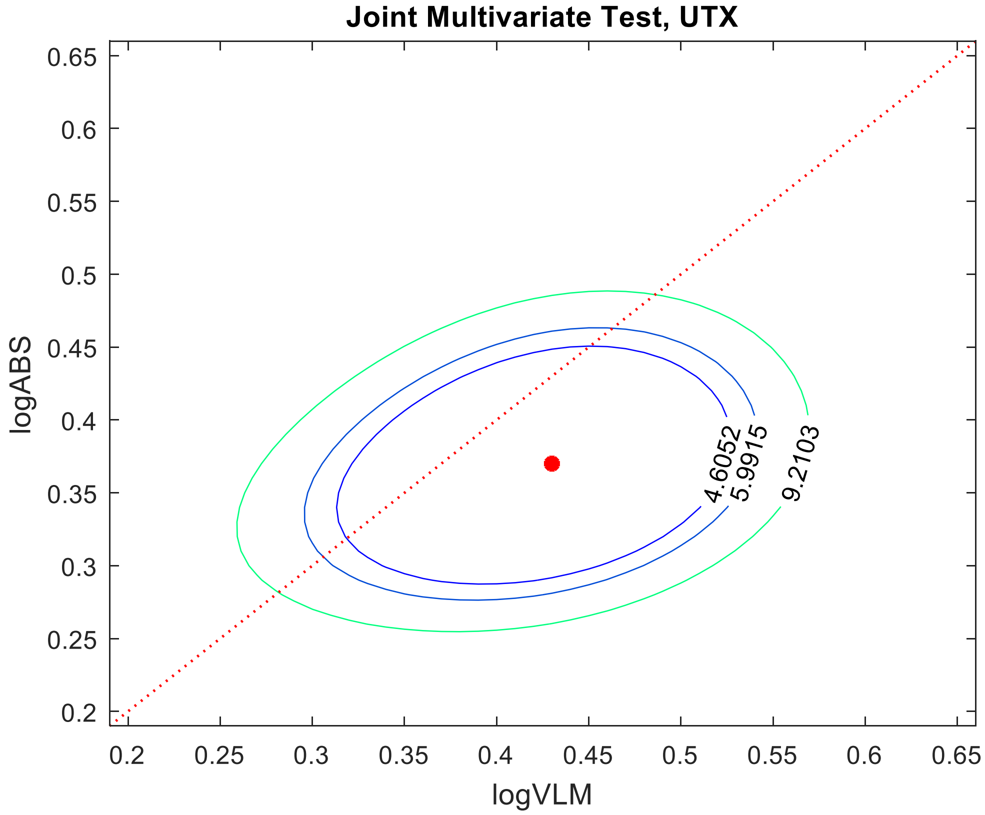
Joint Multivariate Test, PFE
9.2103
5.9915
4.6052
logABS
logVLM
0.5
0.4
0.3
0.2
0.1
0
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55

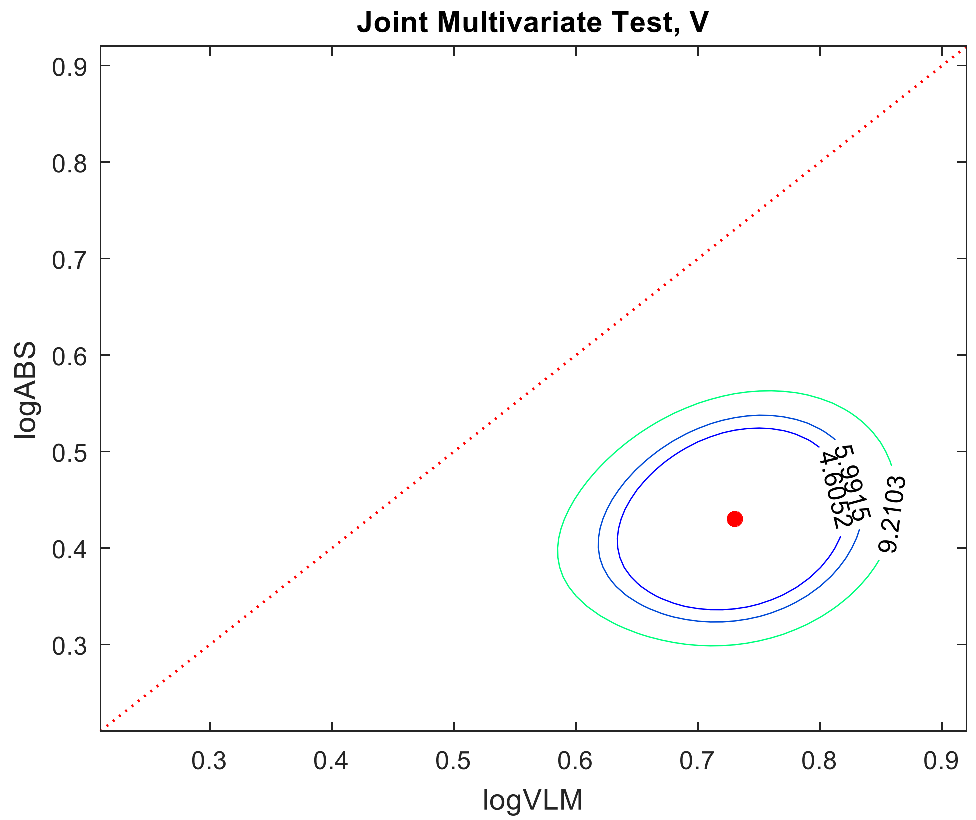


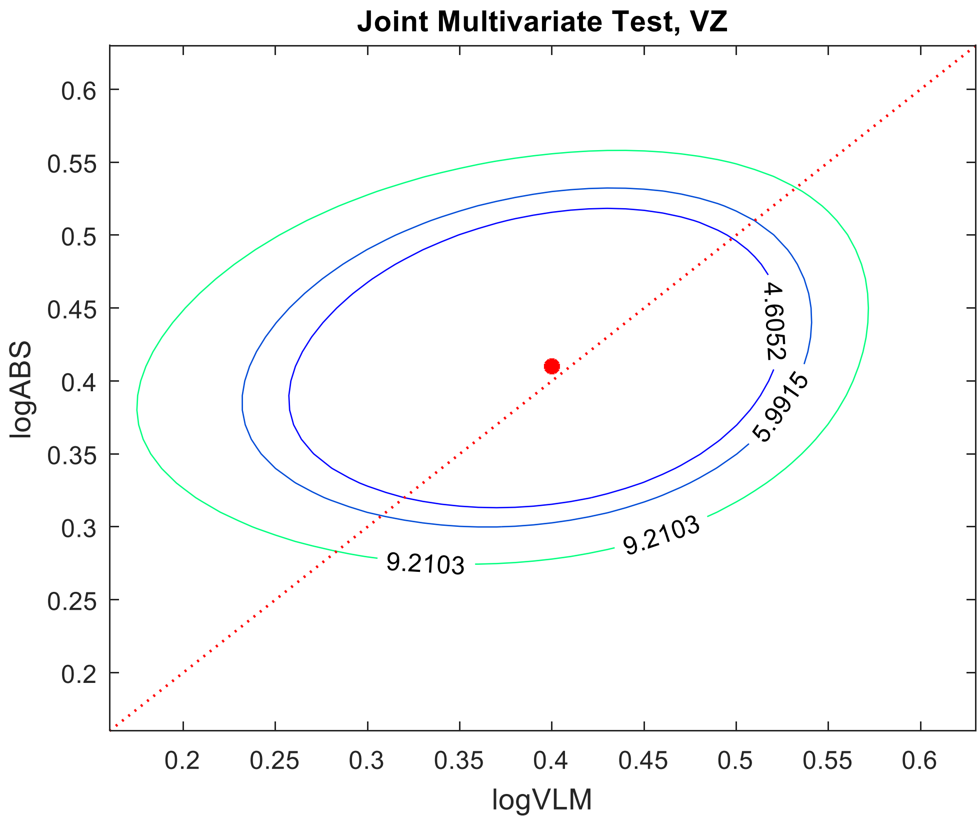
Joint Multivariate Test, PG
9.2103
5.9915
4.6052
logABS
logVLM
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65

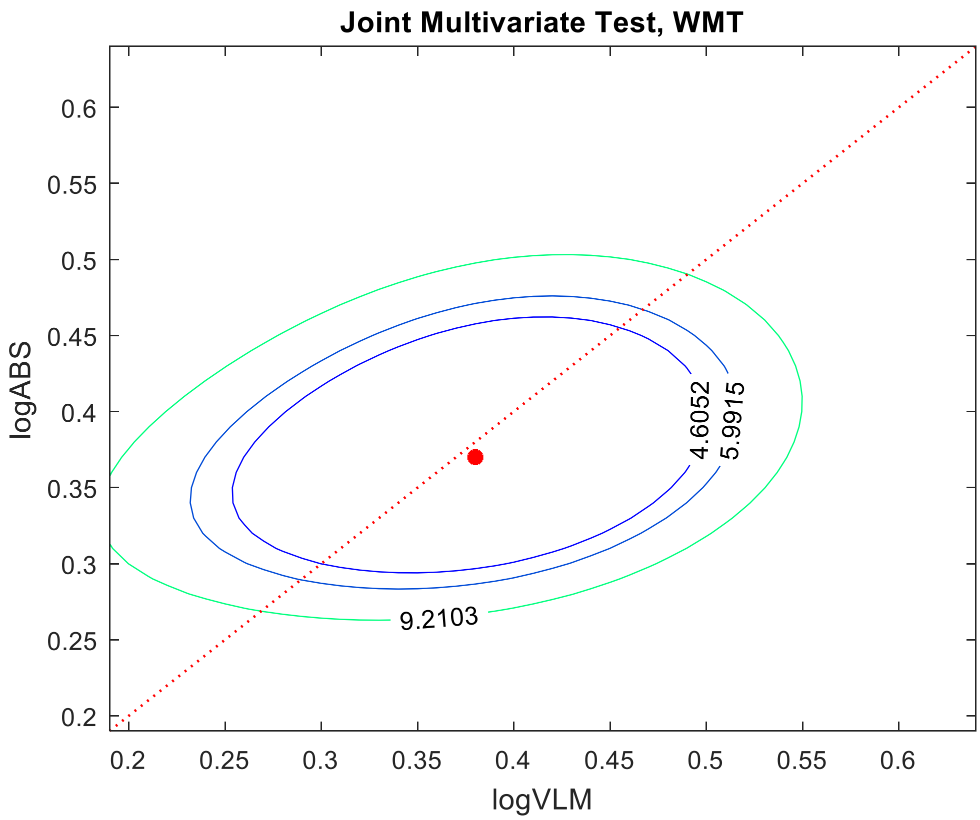

Joint Multivariate Test, TRV
logABS
logVLM
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
4.6052
5.9915
9.2103

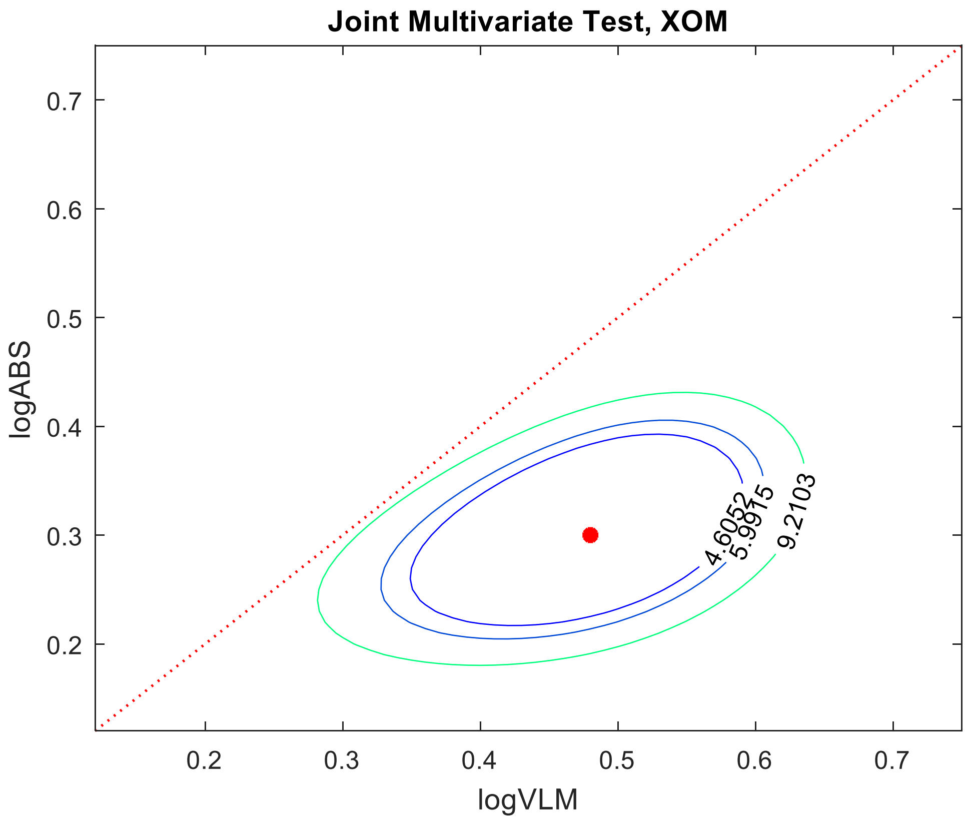

Joint Multivariate Test, UNH
logABS
logVLM
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
4.6052
5.9915
9.2103


Joint Multivariate Test, UTX
logABS
logVLM
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
4.6052
5.9915
9.2103


Joint Multivariate Test, V
logABS
logVLM
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.3
0.4
0.5
0.6
0.7
0.8
0.9
5.9915
4.6052
9.2103

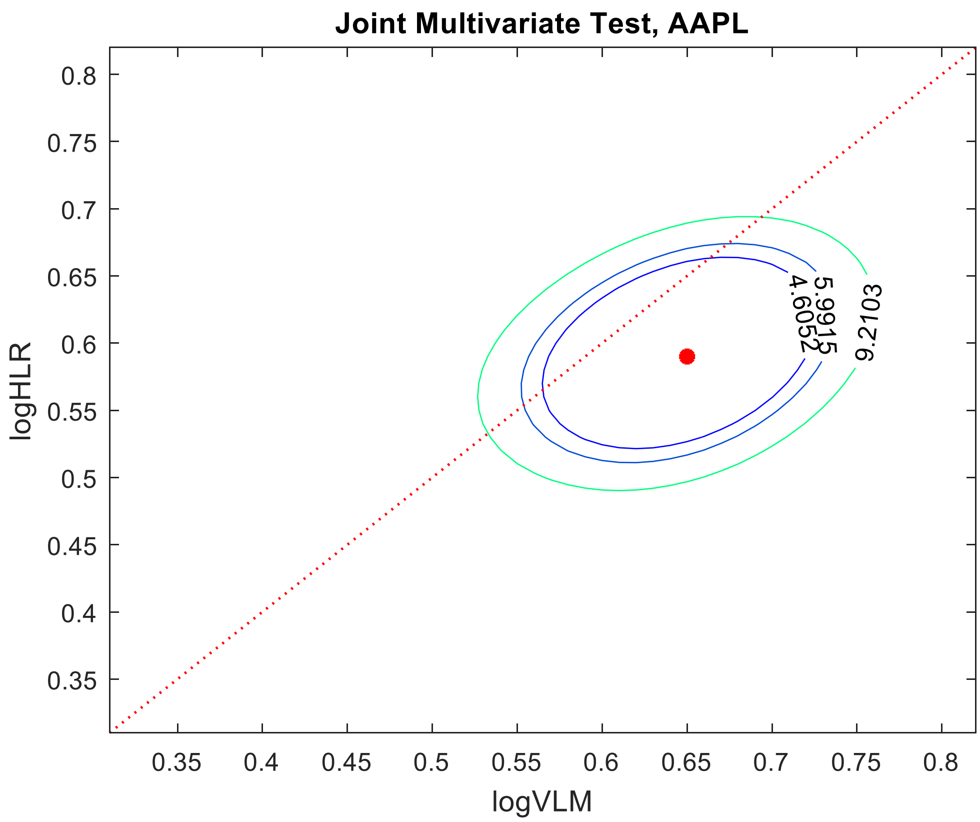

Joint Multivariate Test, VZ
logABS
logVLM
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
4.6052
5.9915
9.2103
9.2103

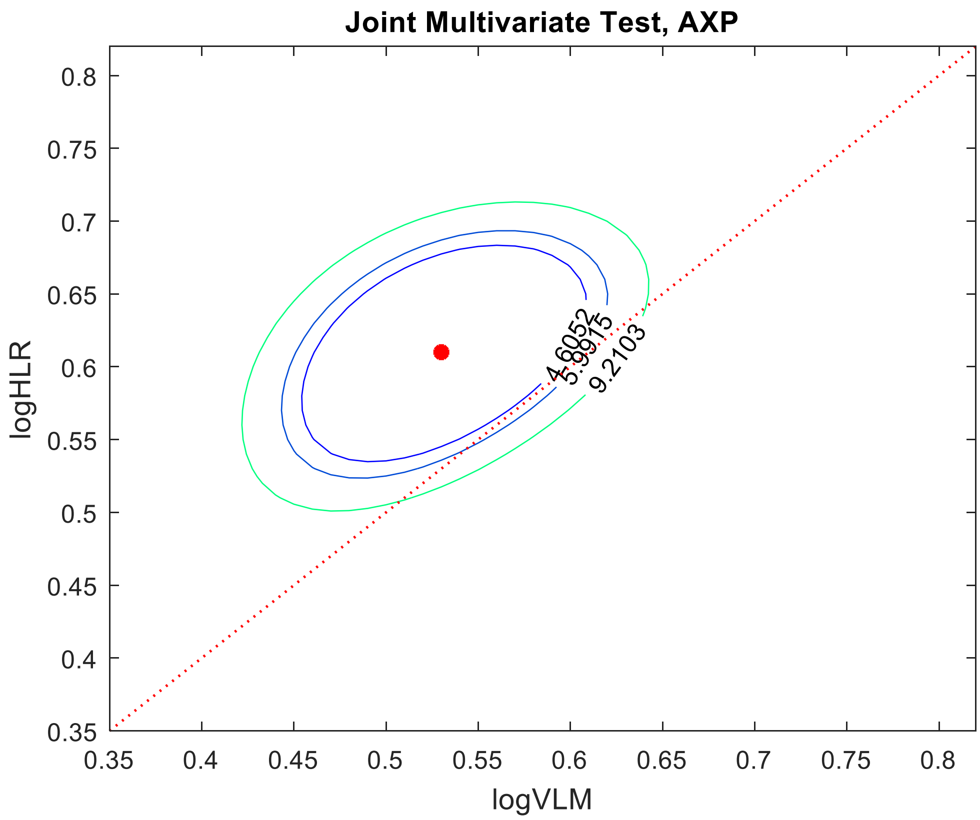

Joint Multivariate Test, WMT
logABS
logVLM
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
4.6052
5.9915
9.2103

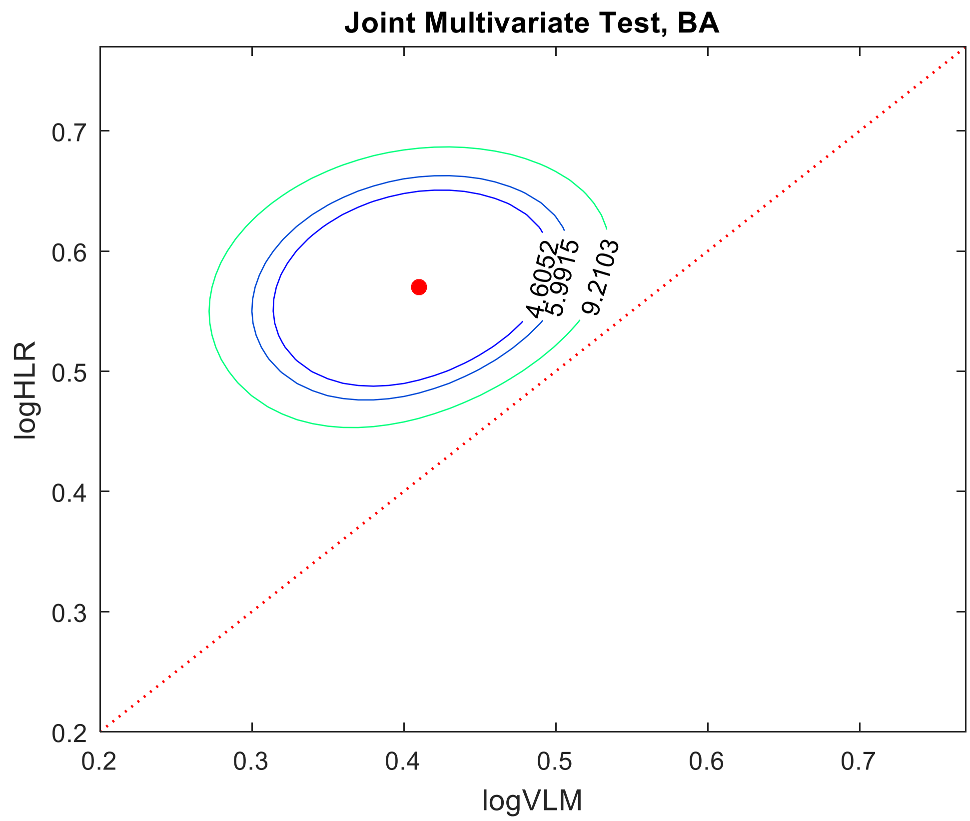

Joint Multivariate Test, XOM
logABS
logVLM
0.7
0.6
0.5
0.4
0.3
0.2
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

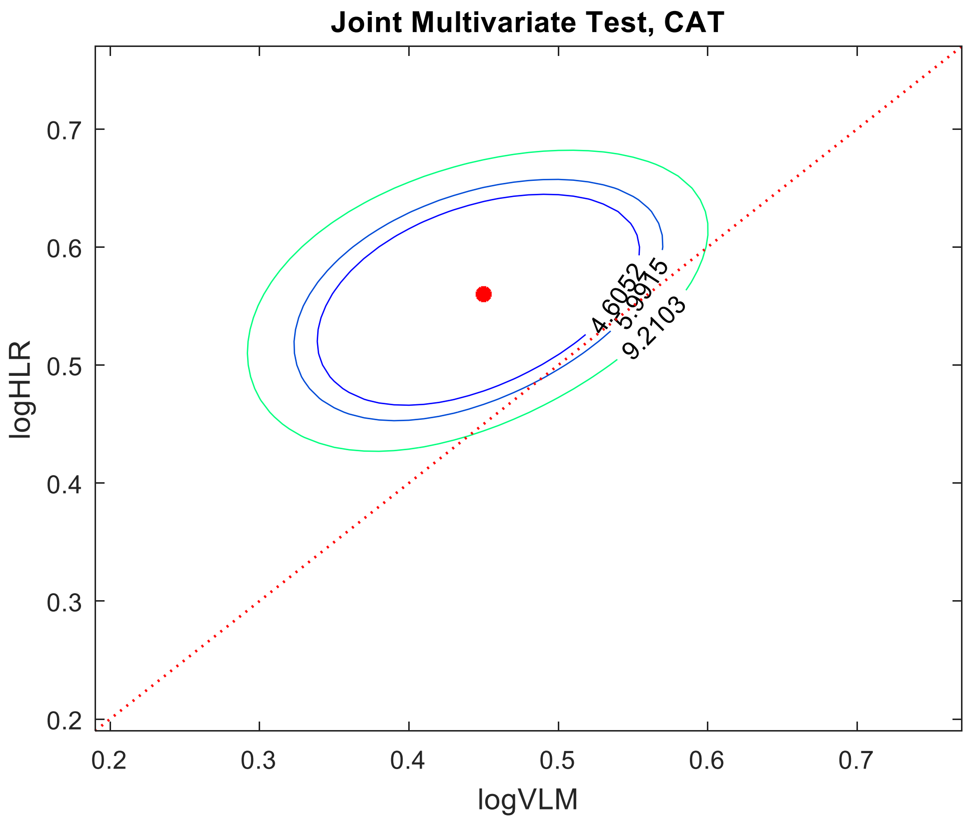
C.2.3 Confidence ellipsoids long memory coefficients for volatility and trading volume. Volatility proxy (σ): **GK measure**.

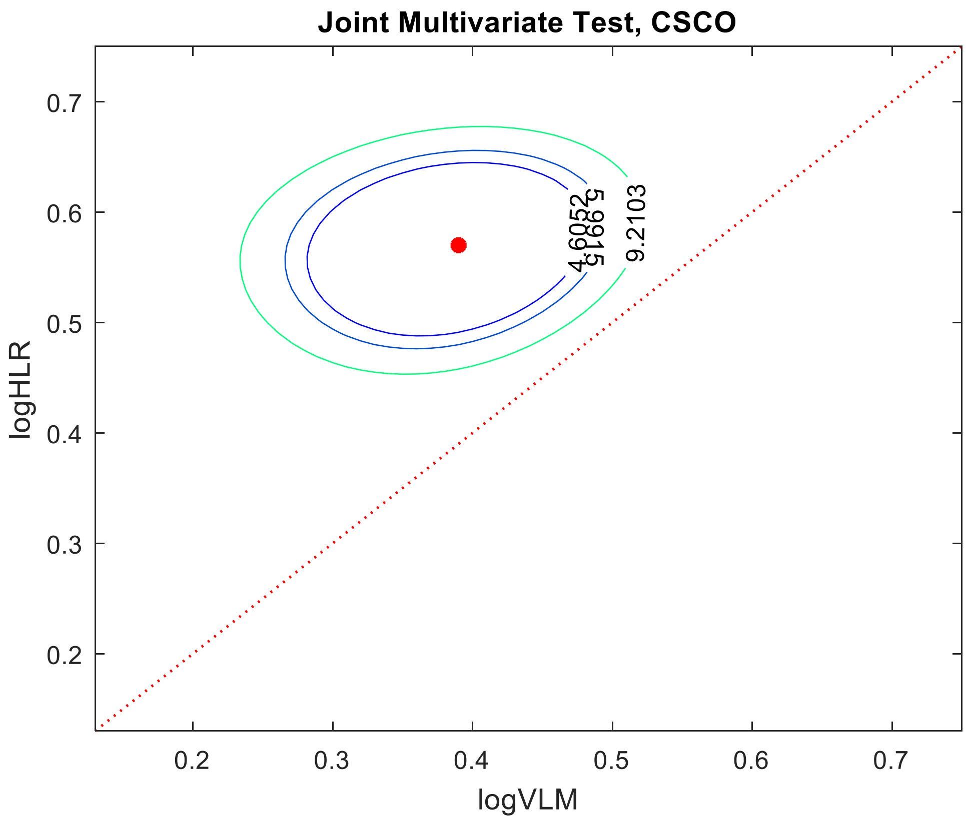
The figures that follow present 90%, 95%, and 99% confidence sets for the long memory parameters of log-trading volume and log-volatility of the stocks under analysis. The confidence sets are obtained from empirical level curves of the $LM_{\boldsymbol{d}}^{FGLS}$ test statistic evaluated at different values given the sample observations, for 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 $\hat{d}_{min}(\sigma)$. The red dashed line represents the 45-degree line.

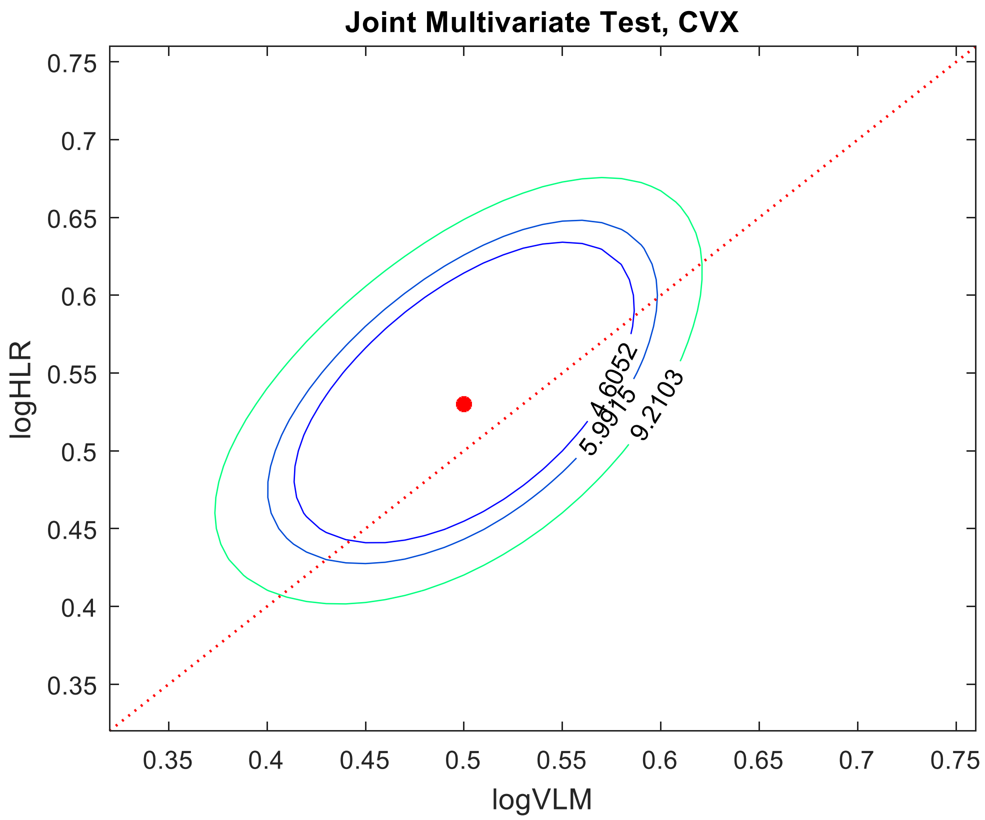


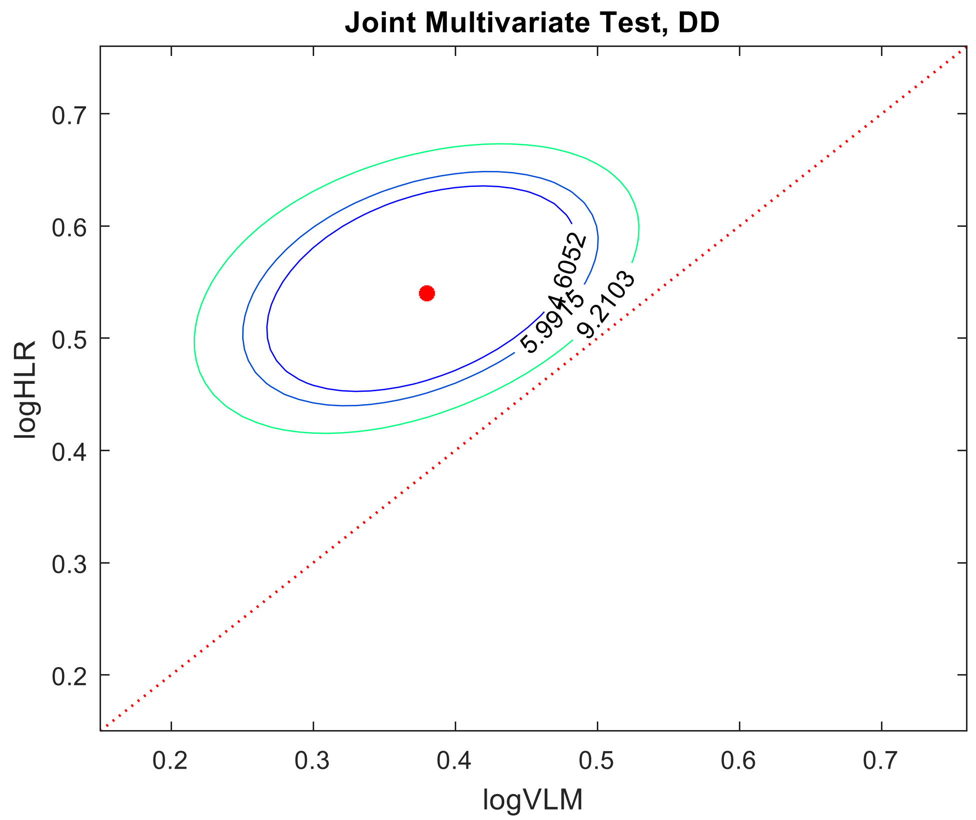

Joint Multivariate Test, AXP
logHLR
logVLM
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
4.6052
5.9915
9.2103

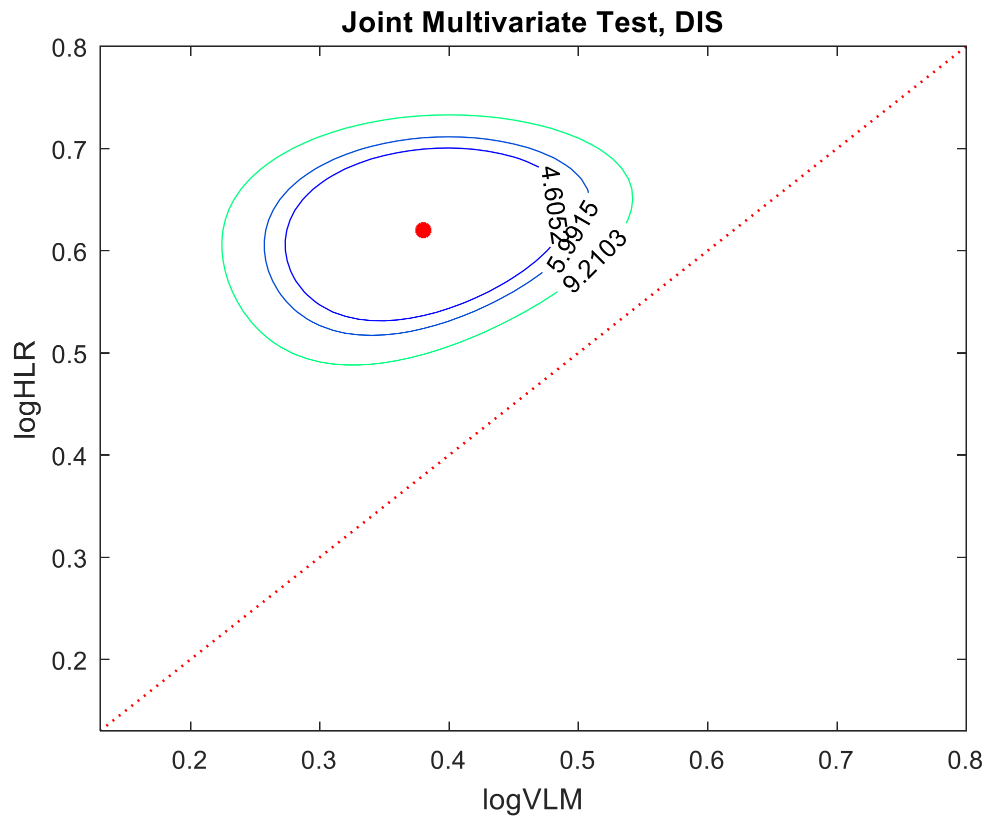

Joint Multivariate Test, BA
logHLR
logVLM
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

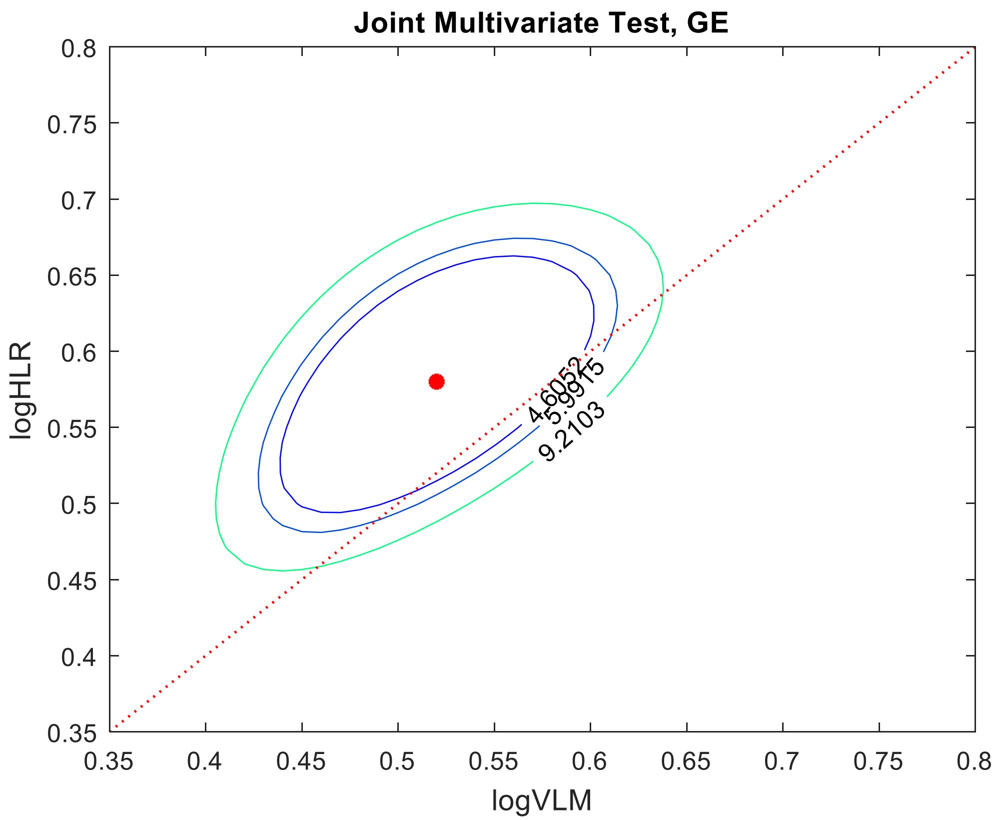

Joint Multivariate Test, CAT
logHLR
logVLM
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

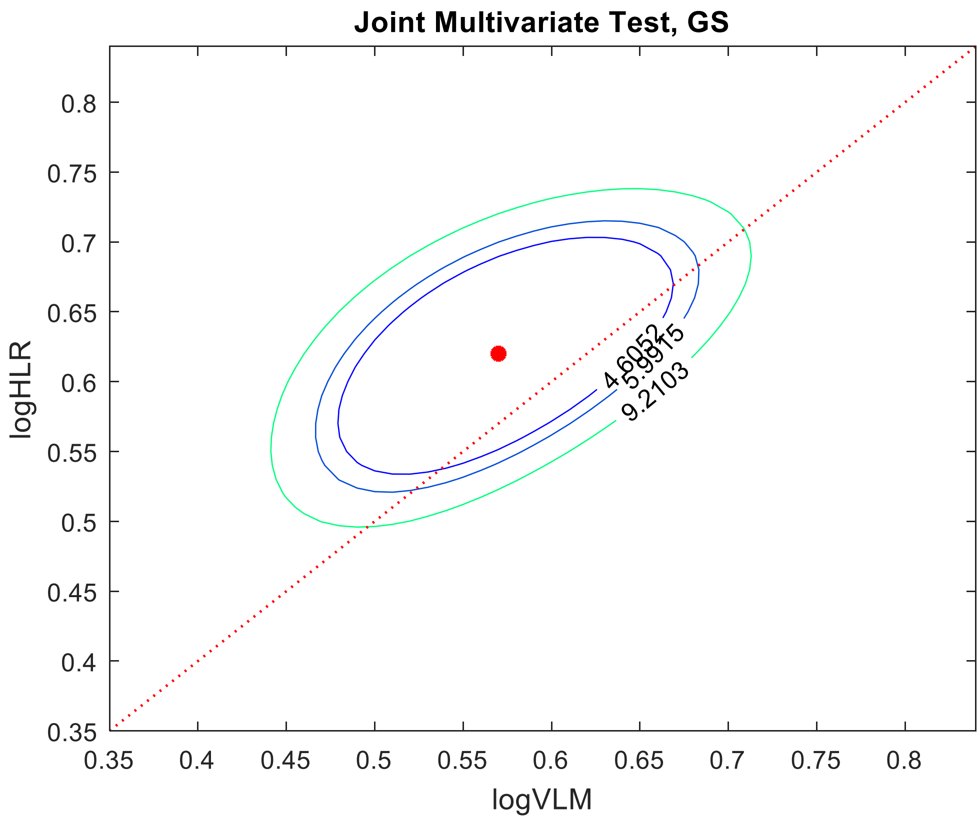

Joint Multivariate Test, CSCO
logHLR
logVLM
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

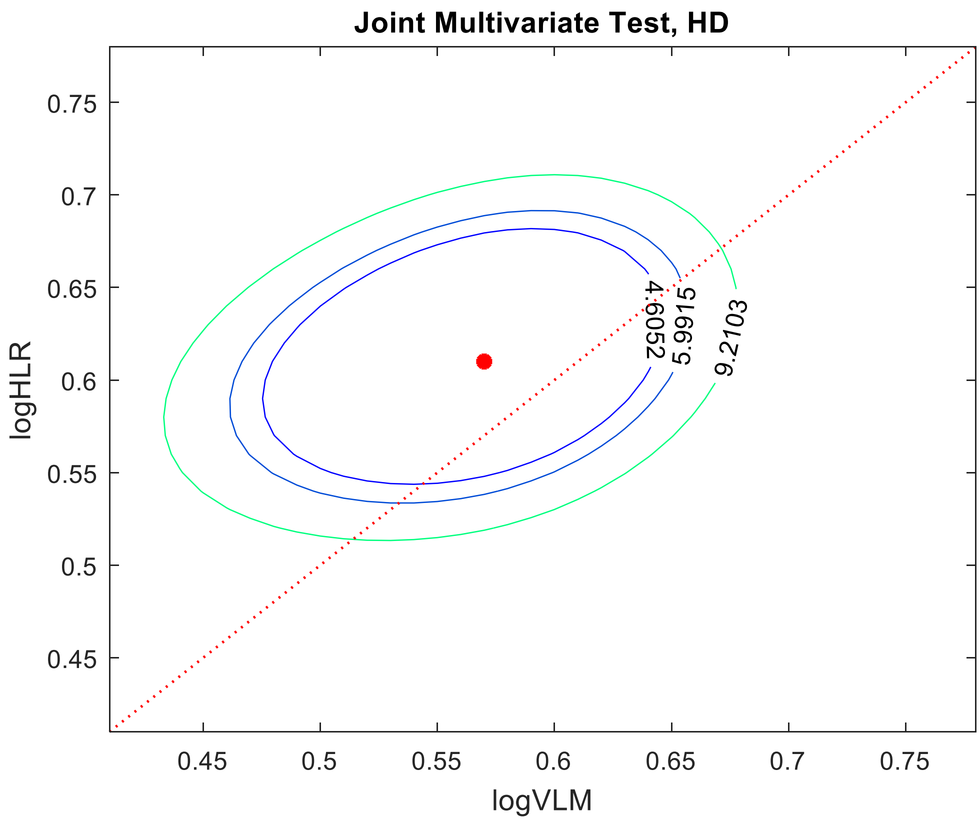

Joint Multivariate Test, CVX
logHLR
logVLM
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
4.6052
5.9915
9.2103

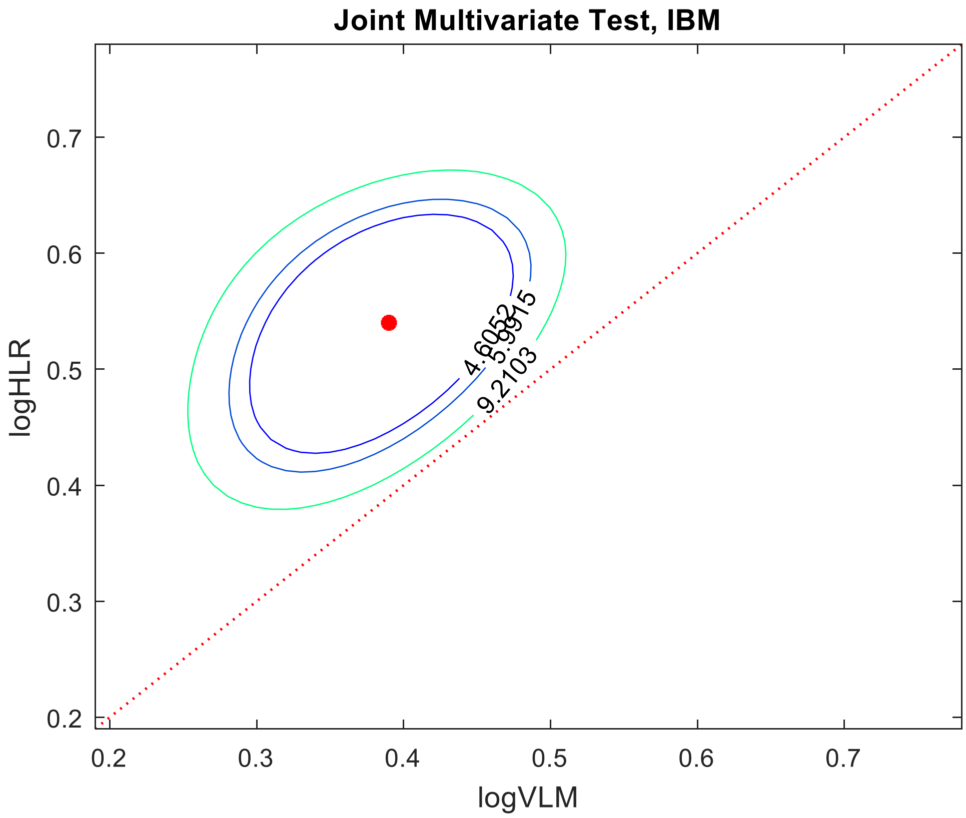

Joint Multivariate Test, DD
logHLR
logVLM
0.7
0.6
0.5
0.4
0.3
0.2
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

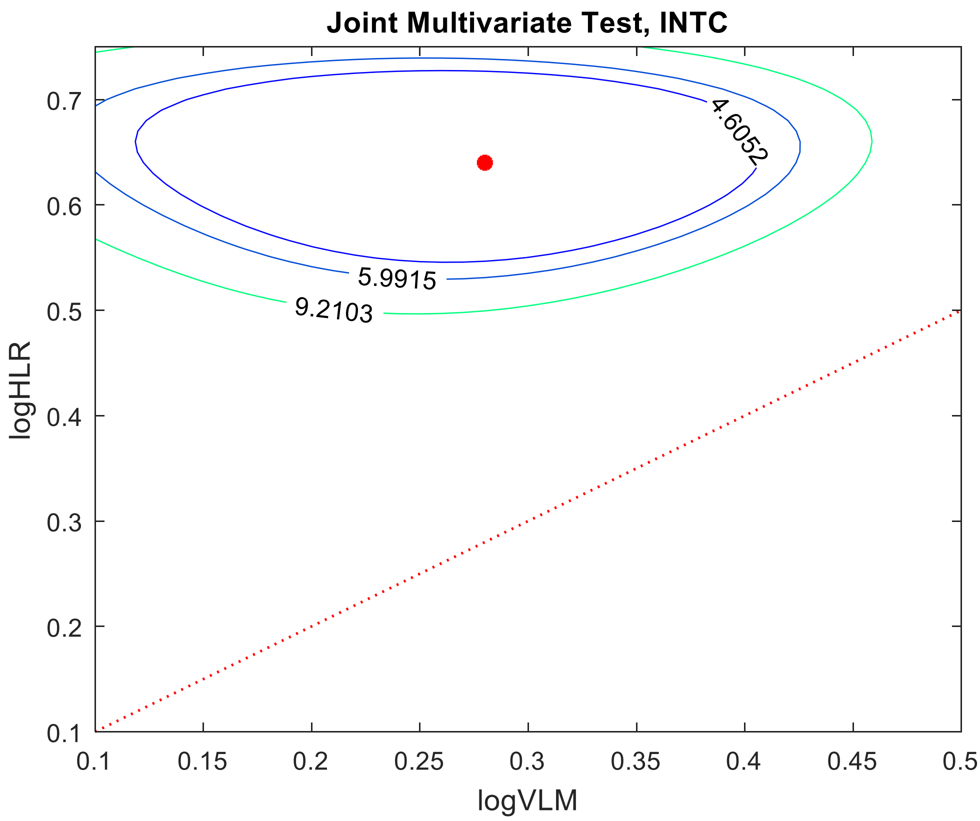

Joint Multivariate Test, DIS
logHLR
logVLM
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.2
0.3
0.4
0.5
0.6
0.7
0.8
4.6052
5.9915
9.2103

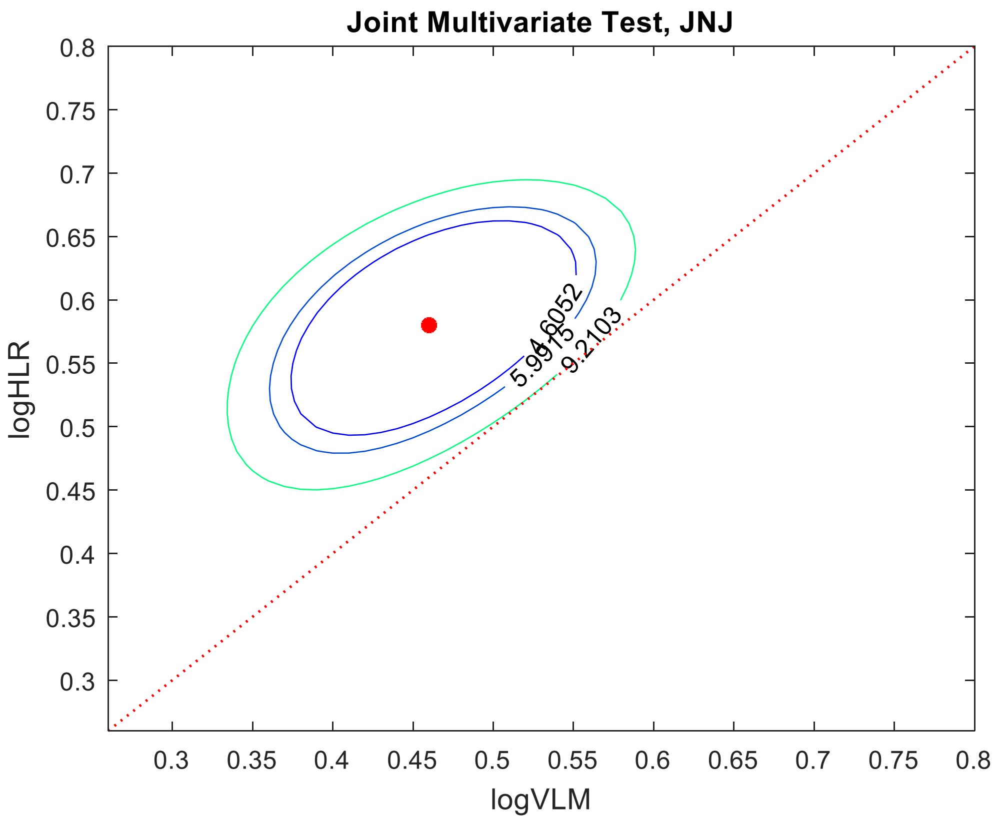

Joint Multivariate Test, GE
logHLR
logVLM
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
4.6052
5.9915
9.2103


Joint Multivariate Test, GS
logHLR
logVLM
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
4.6052
5.9915
9.2103

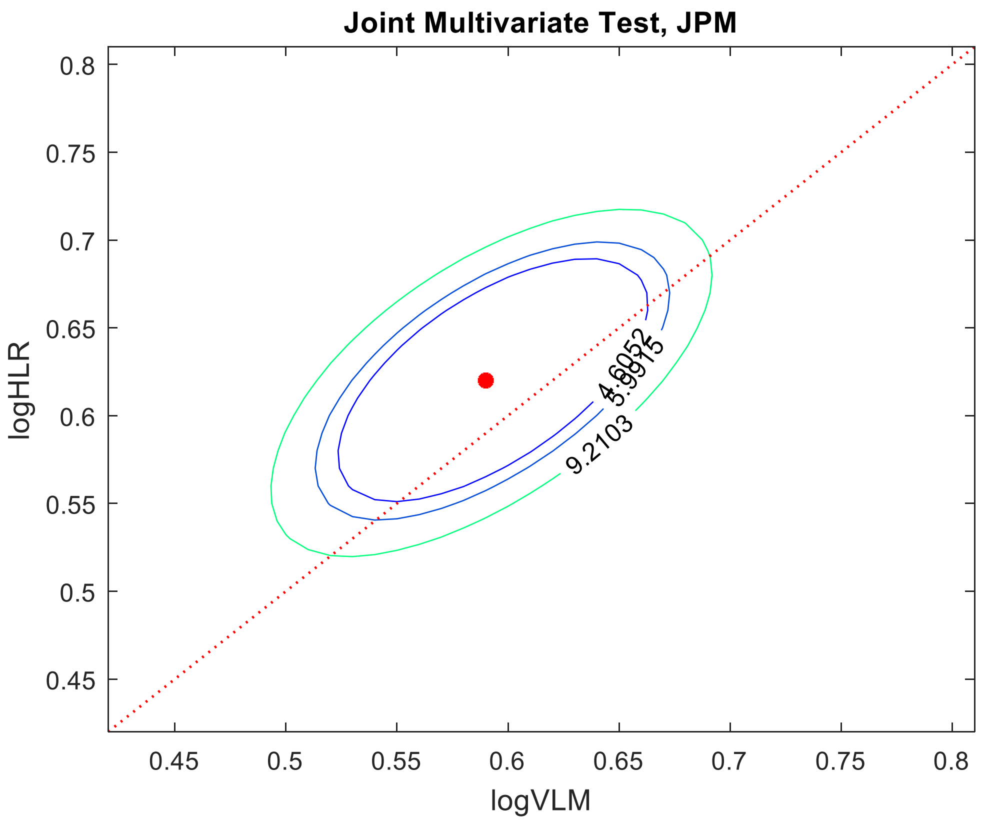

Joint Multivariate Test, HD
logHLR
logVLM
0.45
0.5
0.55
0.6
0.65
0.7
0.75
4.6052
5.9915
9.2103


Joint Multivariate Test, IBM
logHLR
logVLM
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

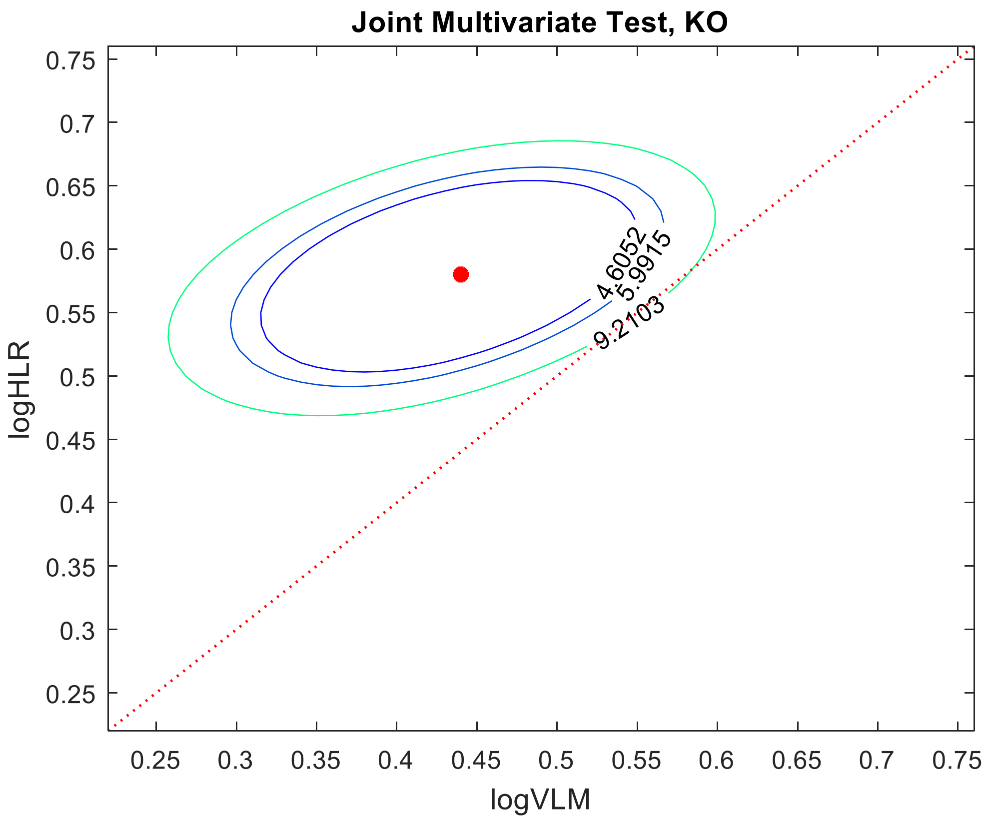

Joint Multivariate Test, INTC
logHLR
logVLM
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.6
0.7
4.6052
5.9915
9.2103

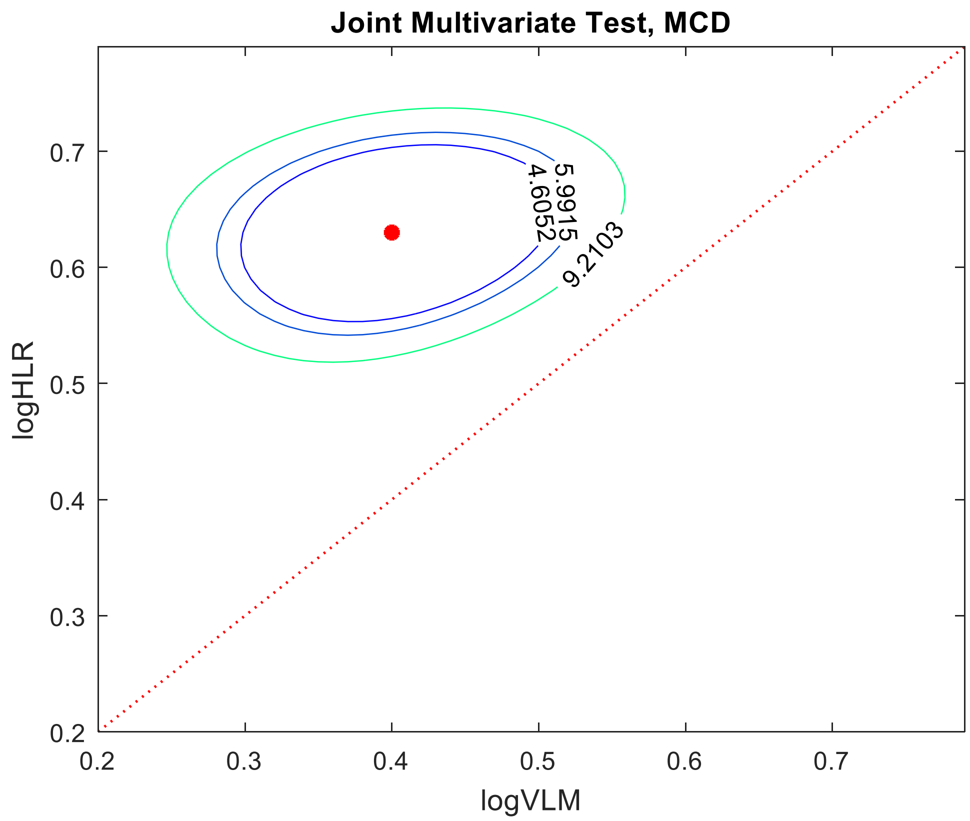


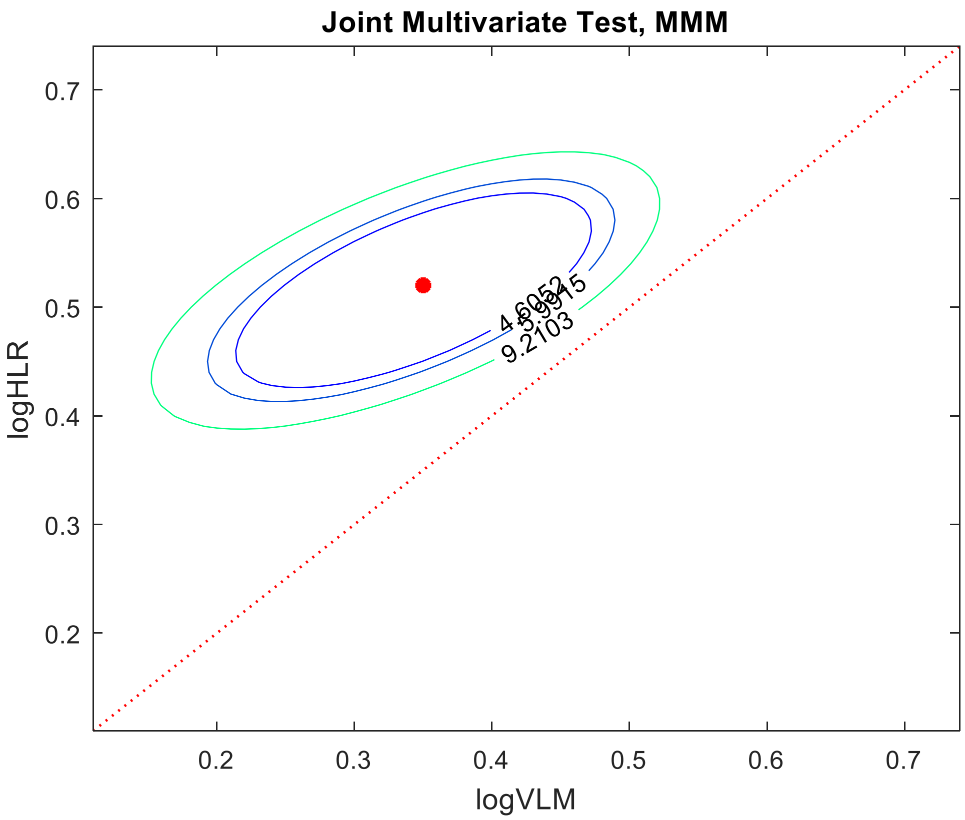
Joint Multivariate Test, JNJ
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
logHLR
4.6052
5.9915
9.2103
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
logVLM

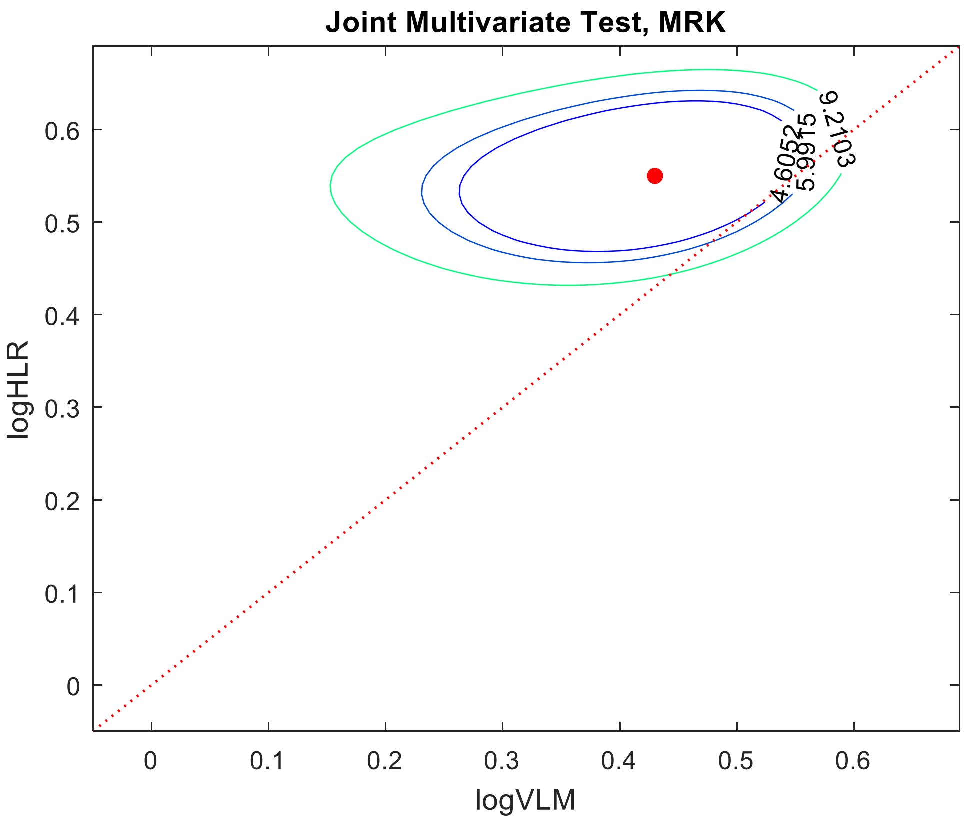


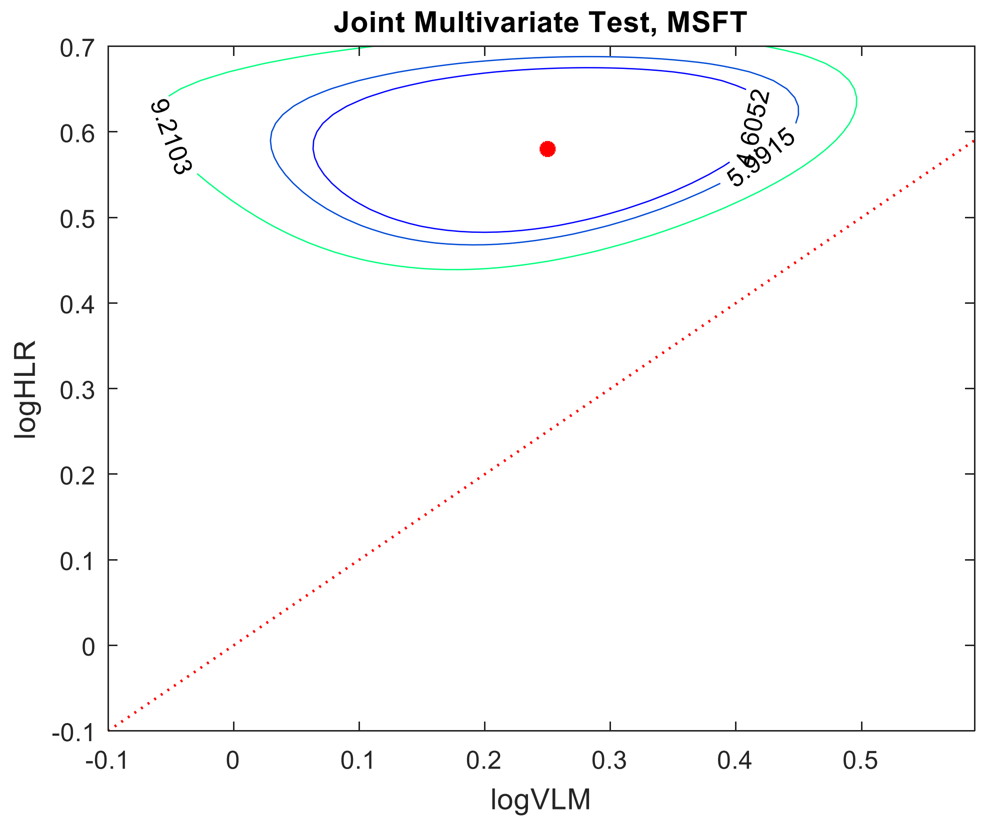
Joint Multivariate Test, JPM
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
logHLR
4.6052
5.9915
9.2103
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
logVLM

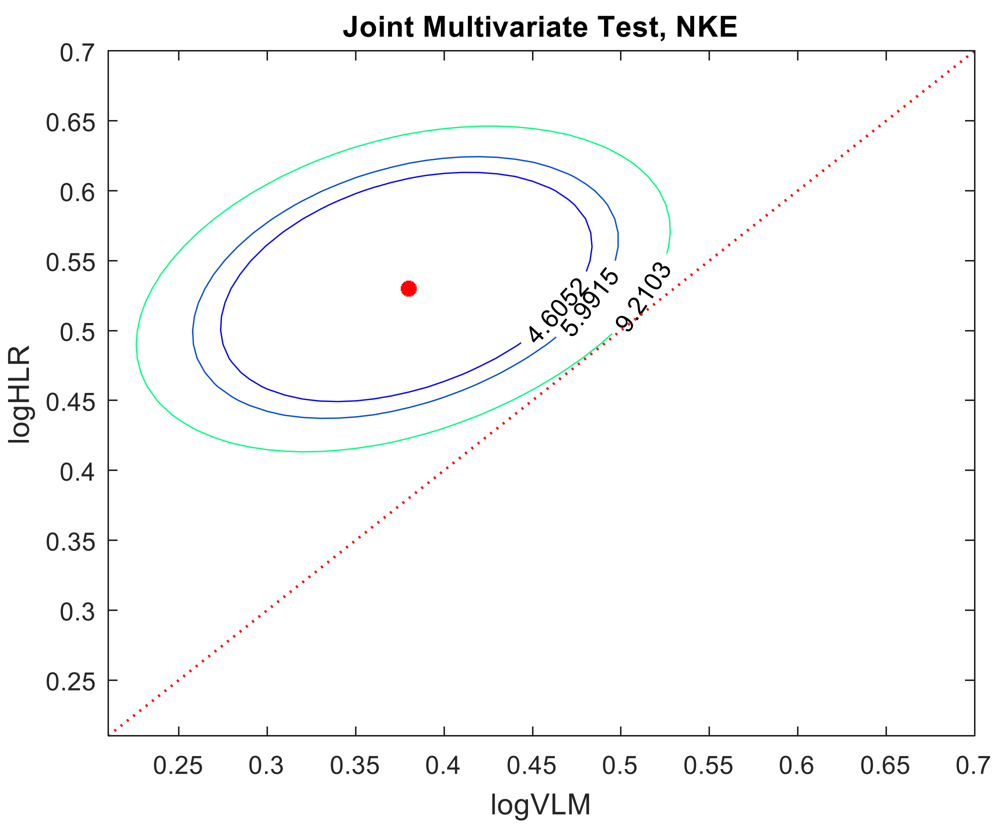

Joint Multivariate Test, KO
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
logHLR
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
logVLM
4.6052
5.9915
9.2103

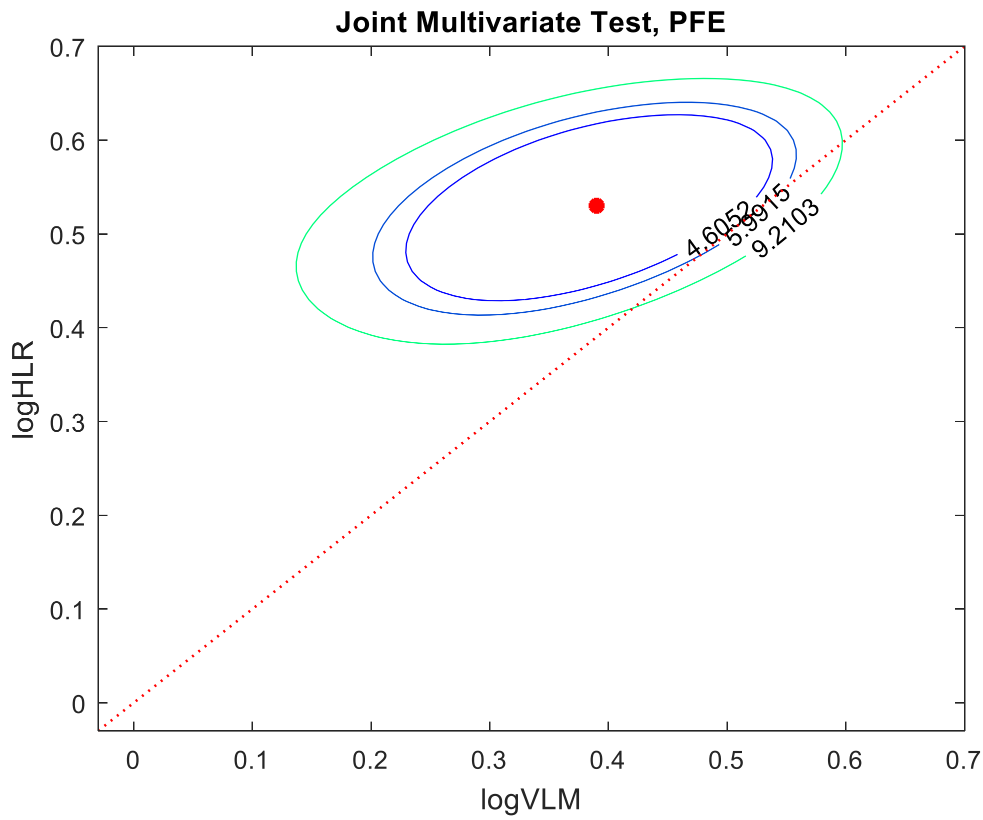

Joint Multivariate Test, MCD
0.7
0.6
0.5
0.4
0.3
0.2
logHLR
0.2
0.3
0.4
0.5
0.6
0.7
logVLM
5.9915
4.6052
9.2103

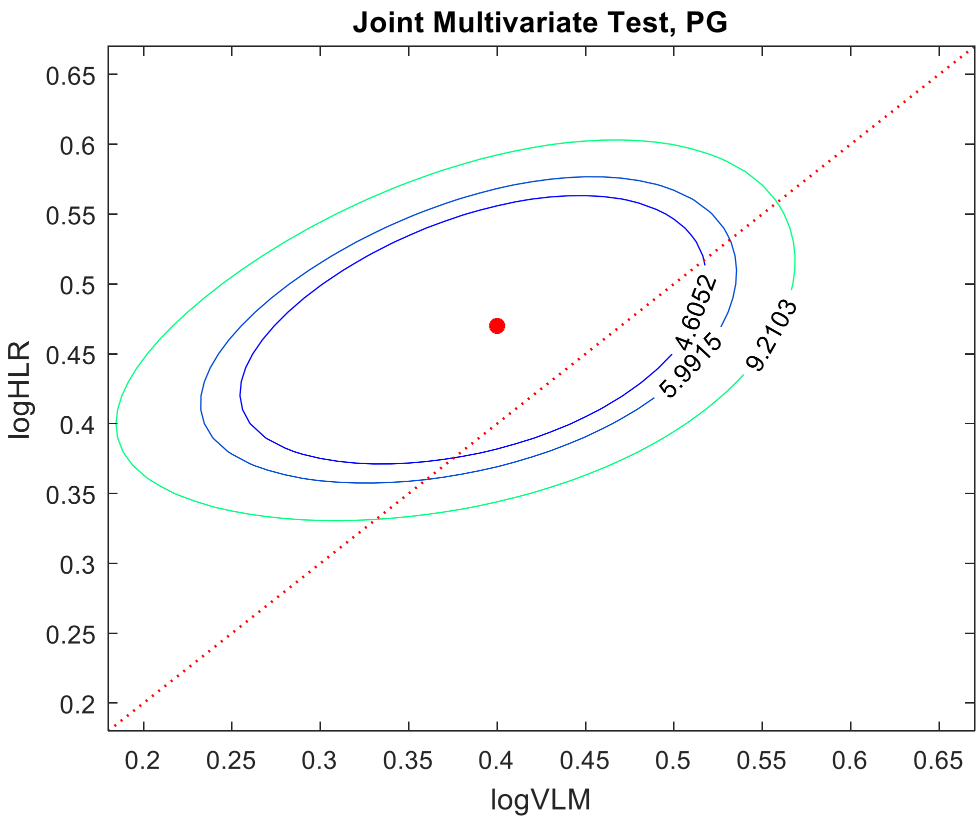

Joint Multivariate Test, MMM
logHLR
logVLM
4.6052
5.9915
9.2103
0.2
0.3
0.4
0.5
0.6
0.7

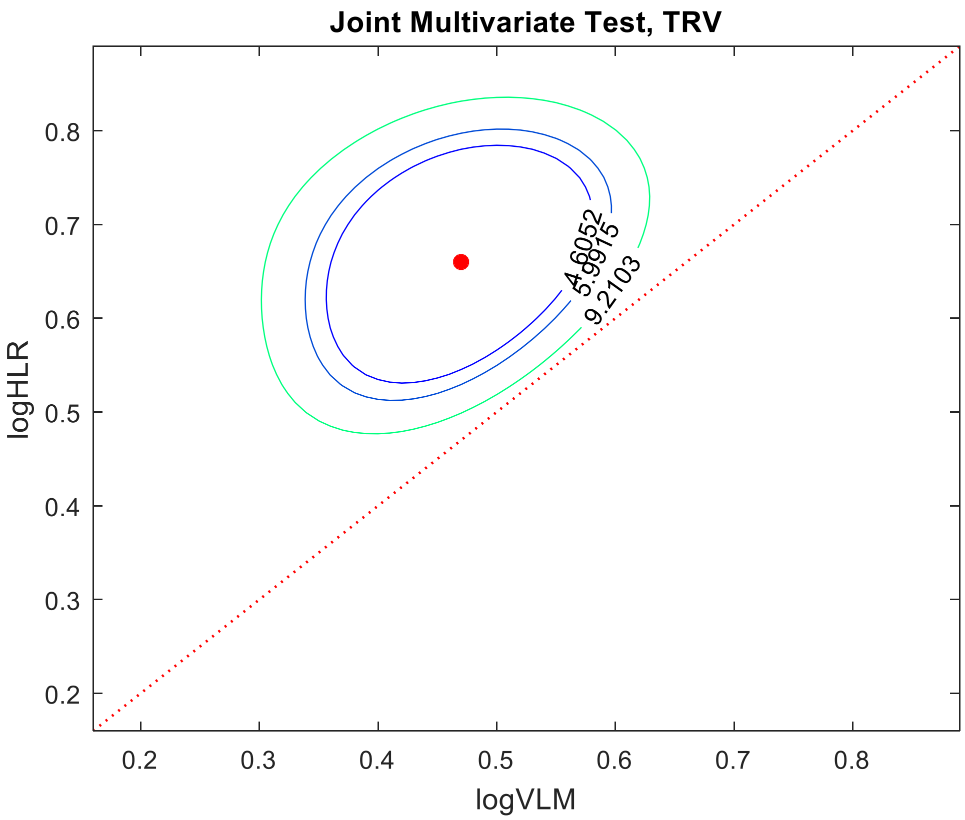

Joint Multivariate Test, MRK
logHLR
logVLM
4.6052
5.9915
9.2103
0
0.1
0.2
0.3
0.4
0.5
0.6

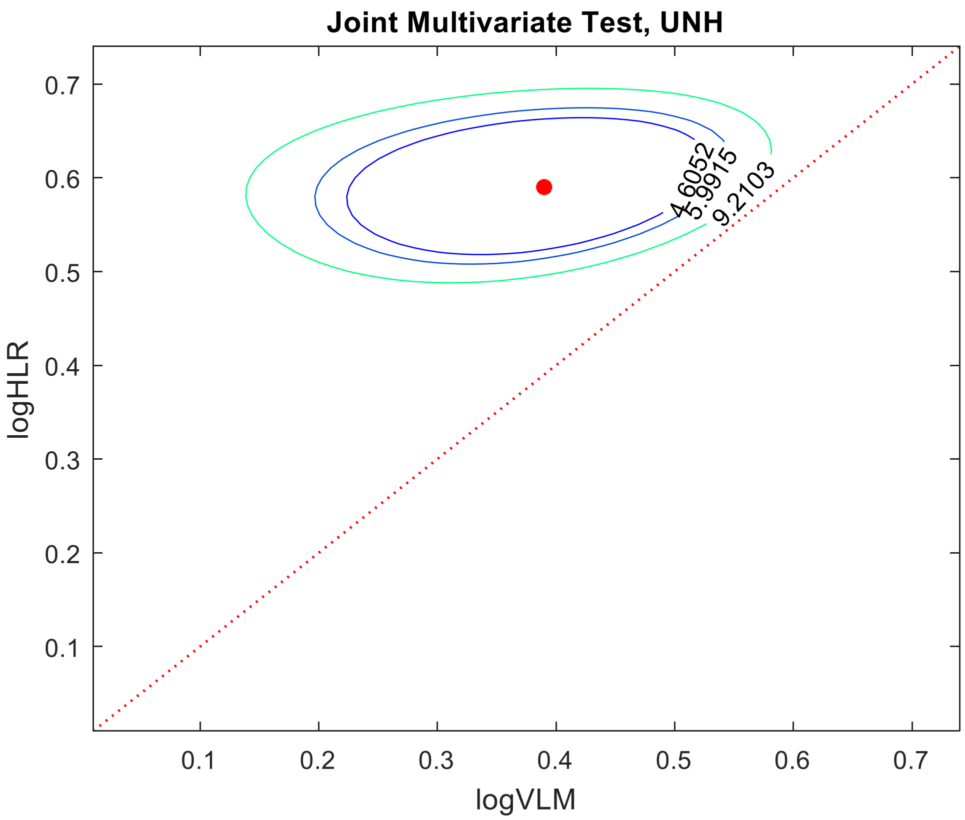

Joint Multivariate Test, MSFT
logHLR
logVLM
9.2103
5.9915
4.6052

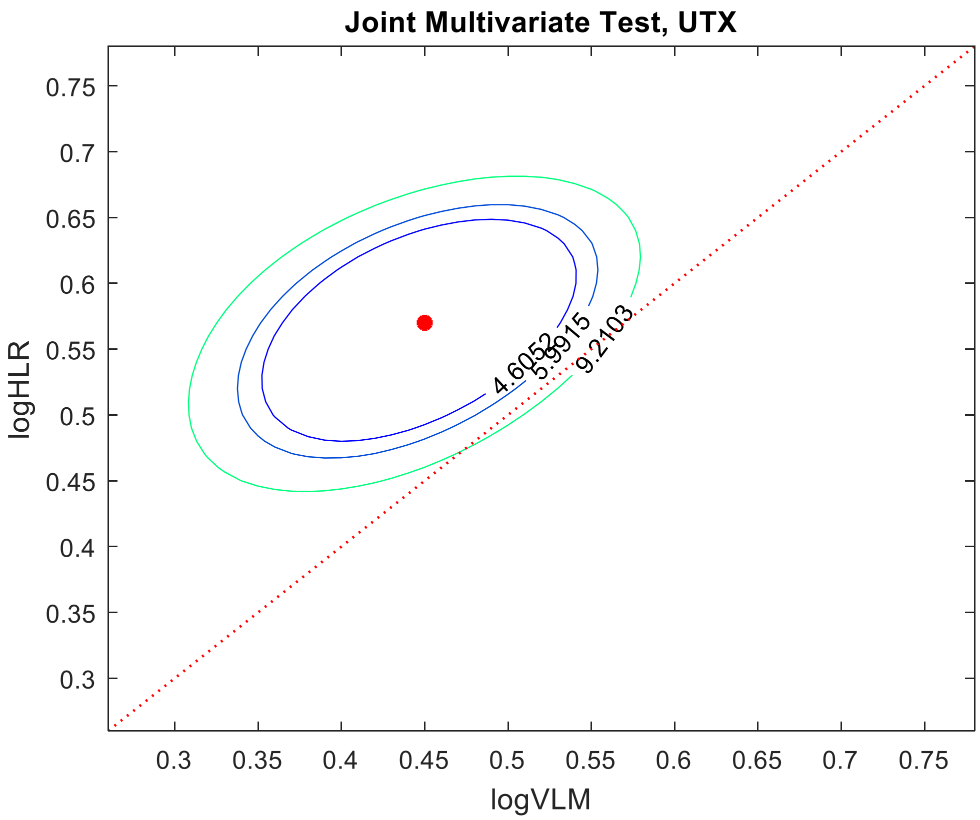

Joint Multivariate Test, NKE
logHLR
logVLM
4.6052
5.9915
9.2103

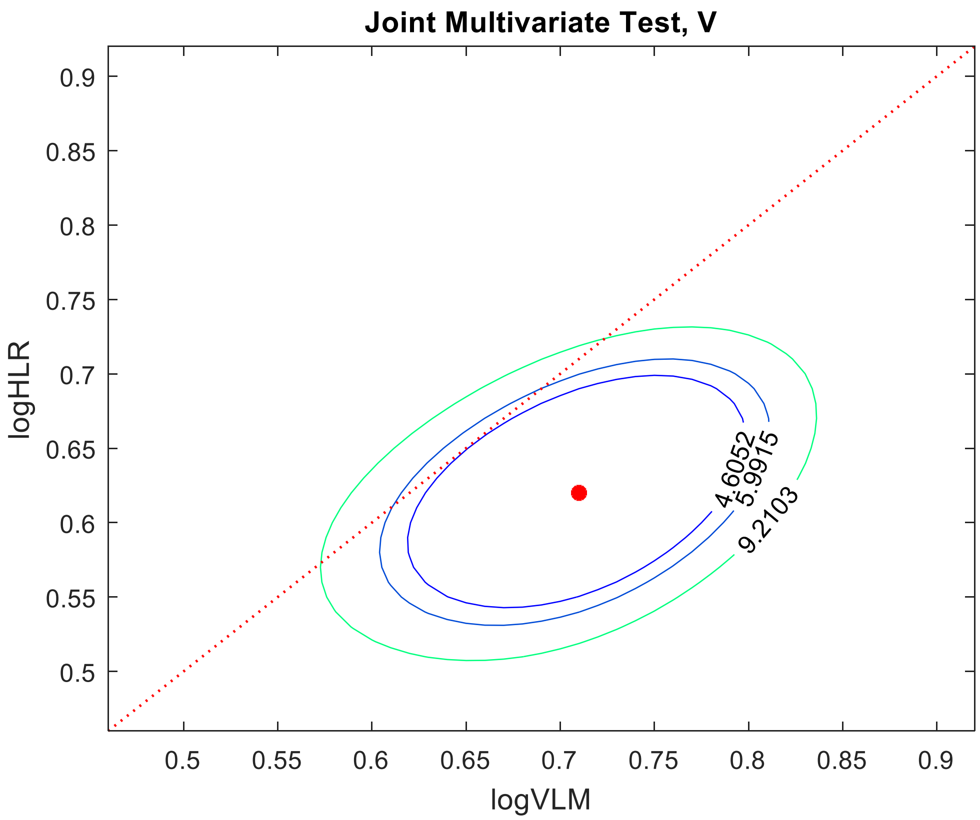

Joint Multivariate Test, PFE
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
logHLR
4.6052
5.9915
9.2103
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
logVLM

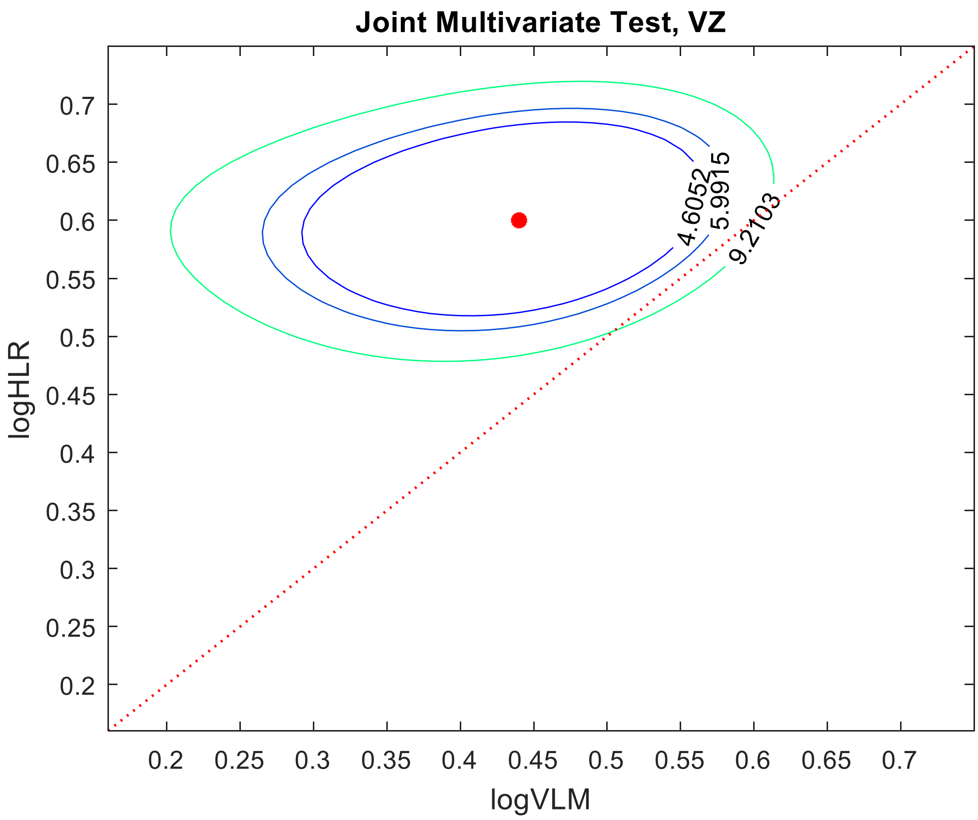

Joint Multivariate Test, PG
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
logHLR
4.6052
5.9915
9.2103
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
logVLM

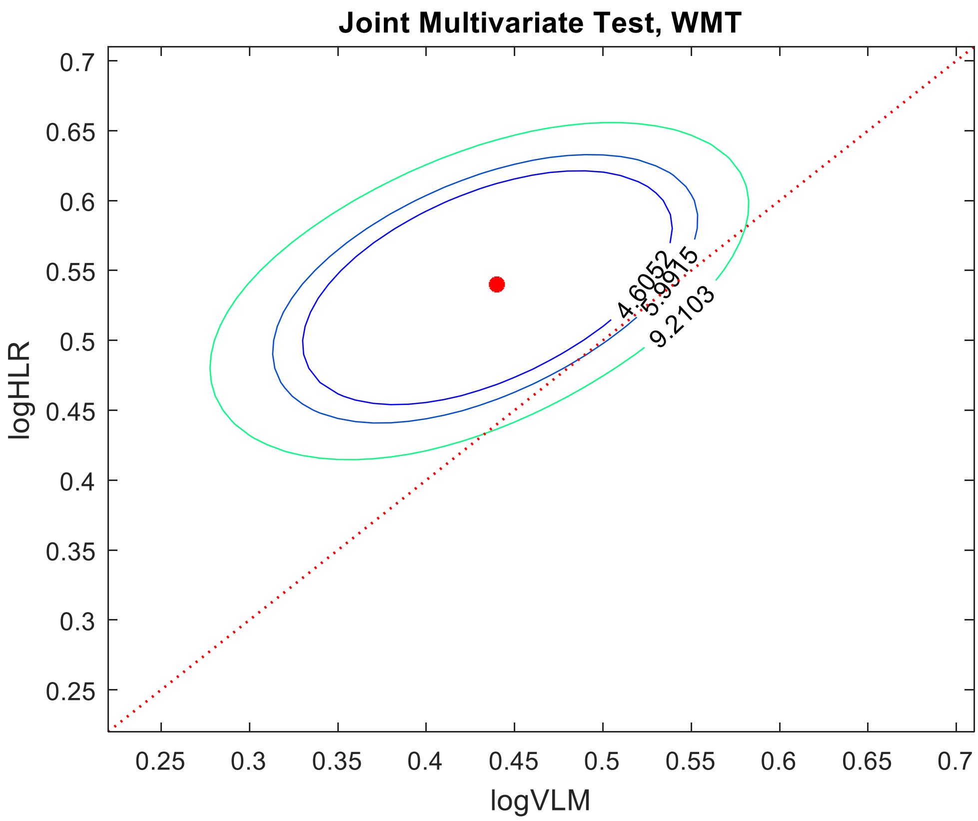

Joint Multivariate Test, TRV
4.6052
5.9915
9.2103
0.2
0.3
0.4
0.5
0.6
0.7
0.8
logVLM
logHLR

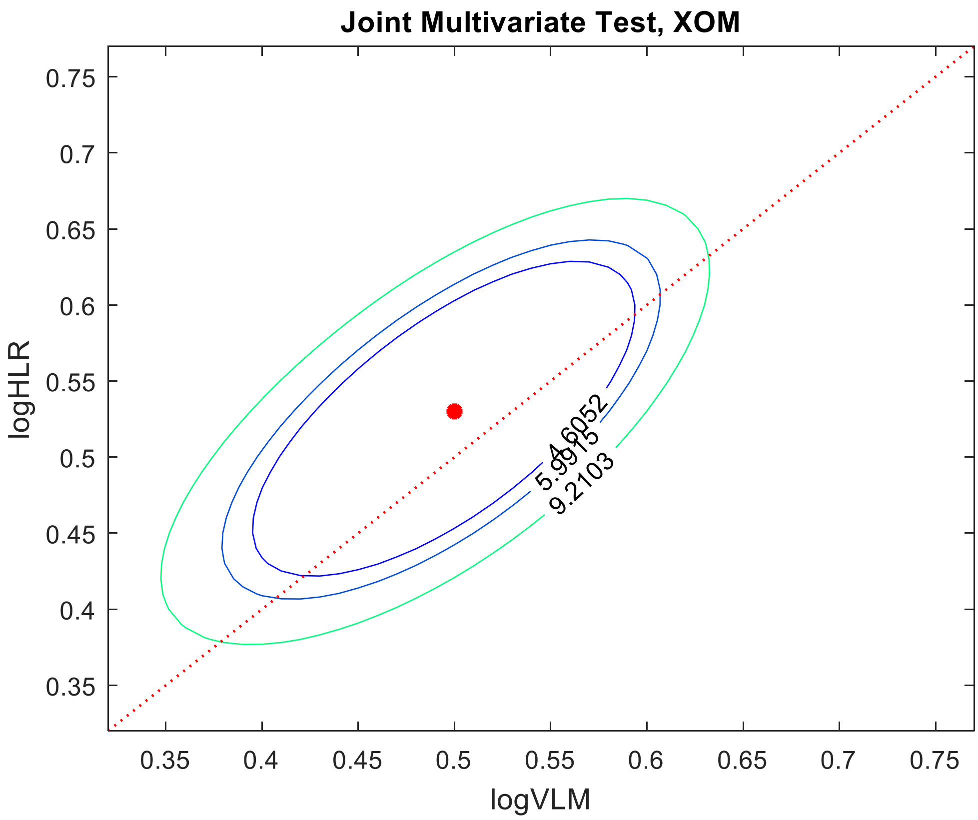

Joint Multivariate Test, UNH
4.6052
5.9915
9.2103
0.1
0.2
0.3
0.4
0.5
0.6
0.7
logVLM
logHLR


Joint Multivariate Test, UTX
logHLR
logVLM
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
4.6052
5.9915
9.2103


Joint Multivariate Test, V
logHLR
logVLM
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
4.6052
5.9915
9.2103


Joint Multivariate Test, VZ
4.6052
5.9915
9.2103
logHLR
logVLM
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7

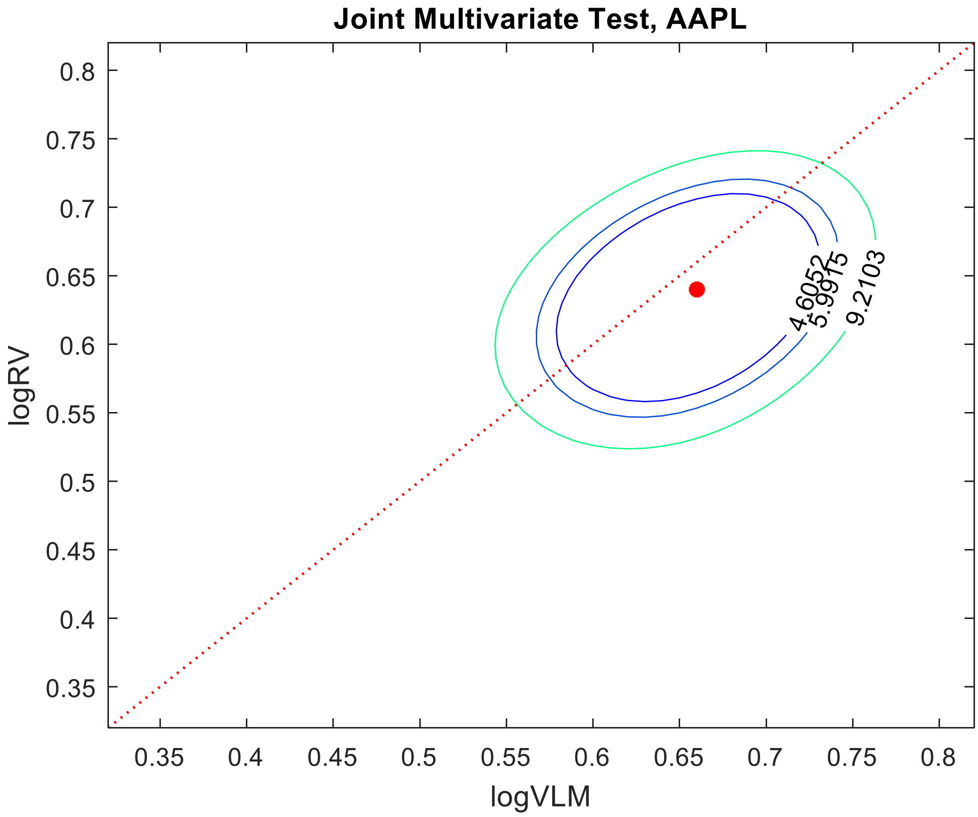

Joint Multivariate Test, WMT
4.6052
5.9915
9.2103
logHLR
logVLM
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7

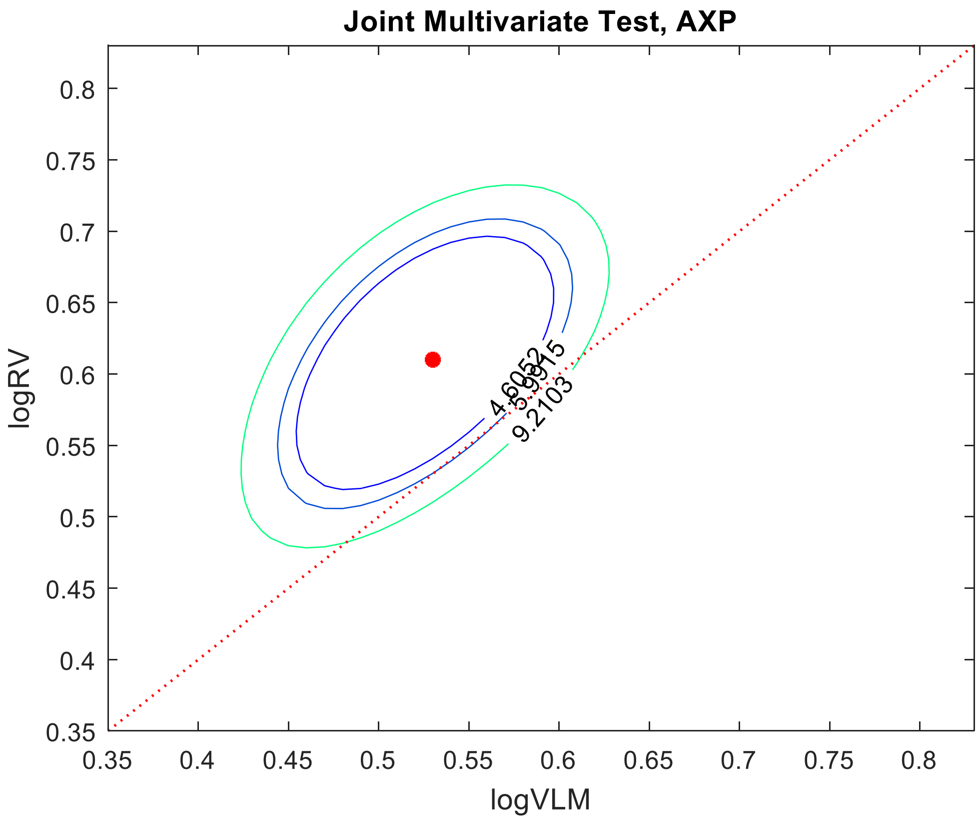


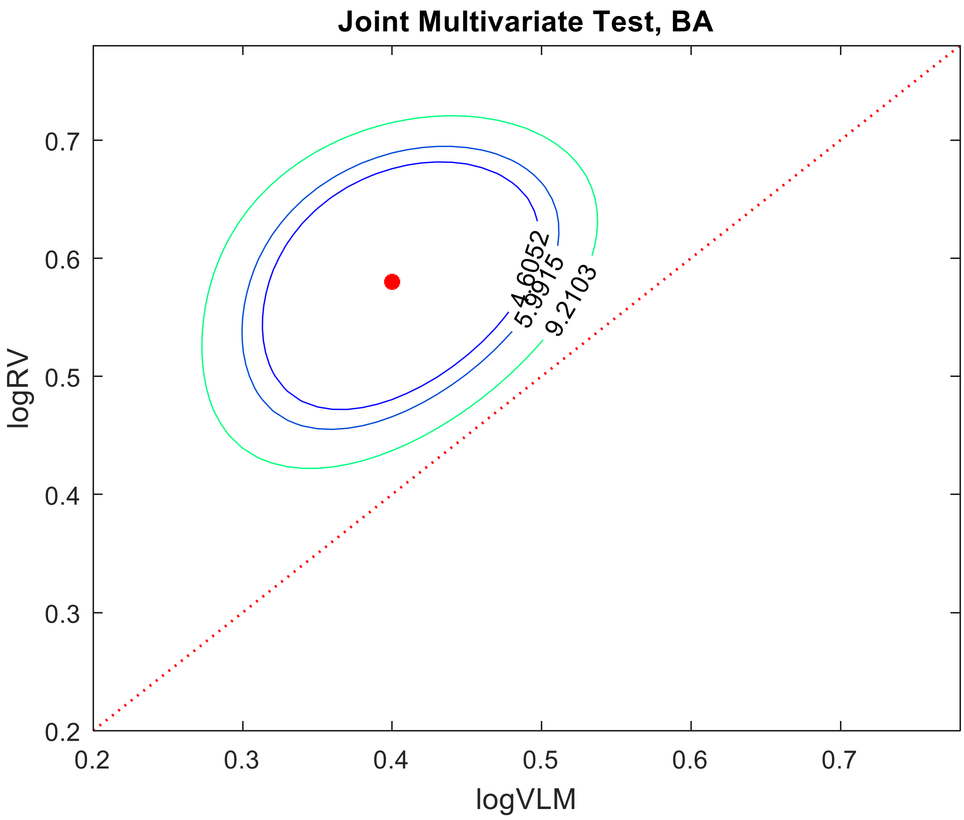
Joint Multivariate Test, XOM
logHLR
logVLM
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7 0.75
4.6052
5.9915
9.2103

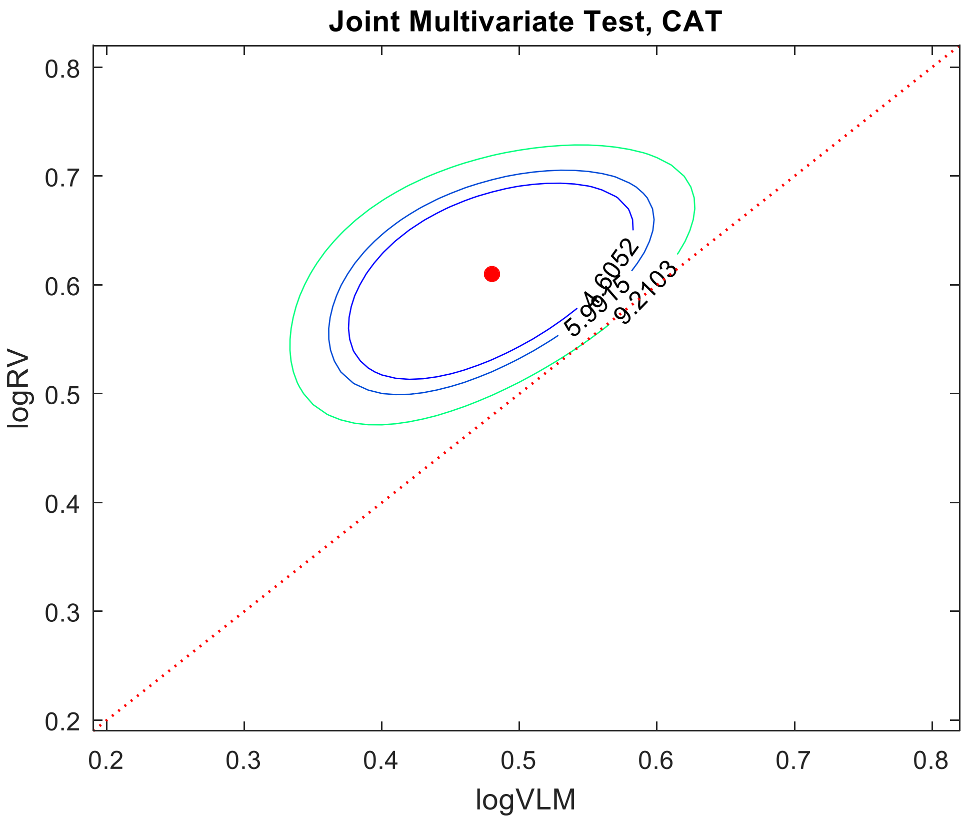
C.2.4 Confidence ellipsoids long memory coefficients for volatility and trading volume. Volatility proxy (σ): **realized variance**.

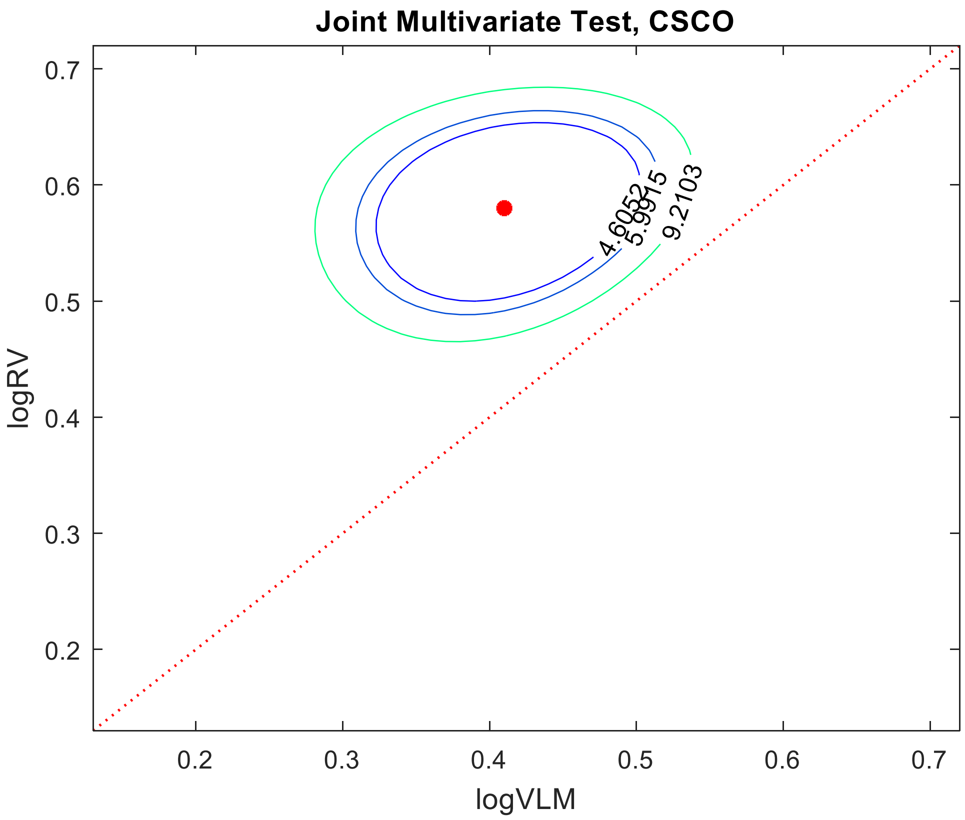
The figures that follow present 90%, 95%, and 99% confidence sets for the long memory parameters of log-trading volume and log-volatility of the stocks under analysis. The confidence sets are obtained from empirical level curves of the $LM_{\boldsymbol{d}}^{FGLS}$ test statistic evaluated at different values given the sample observations, for 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 $\hat{d}_{min}(\sigma)$. The red dashed line represents the 45-degree line.

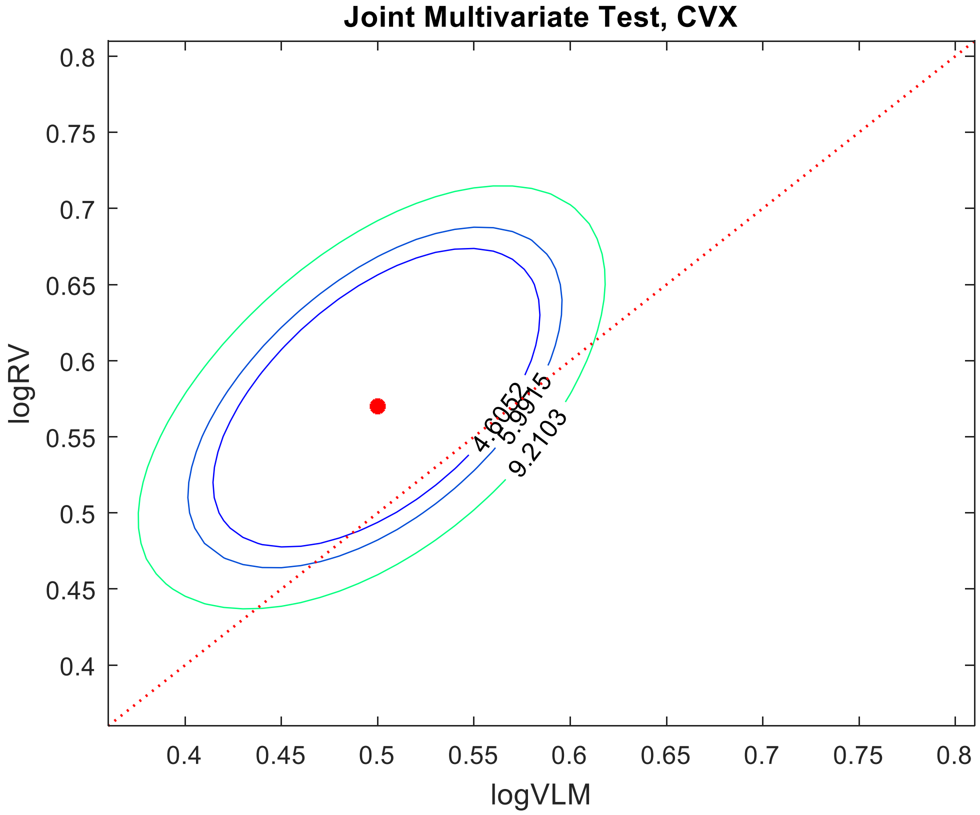


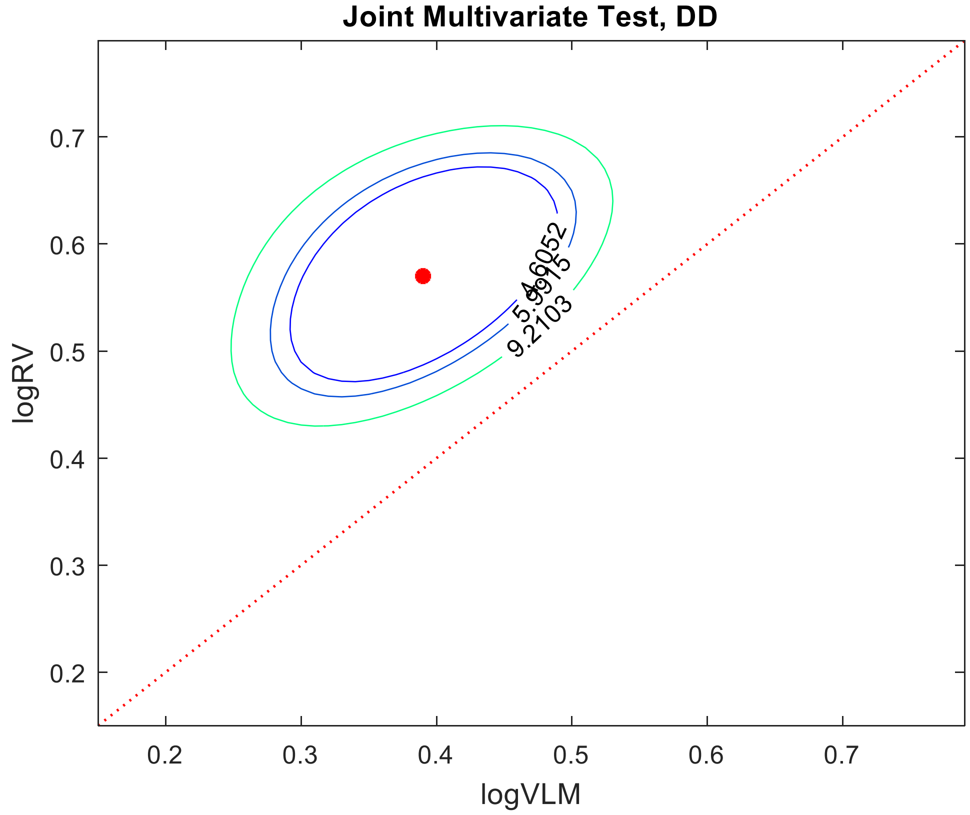

Joint Multivariate Test, AXP
logRV
logVLM
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
4.6052
5.9915
9.2103

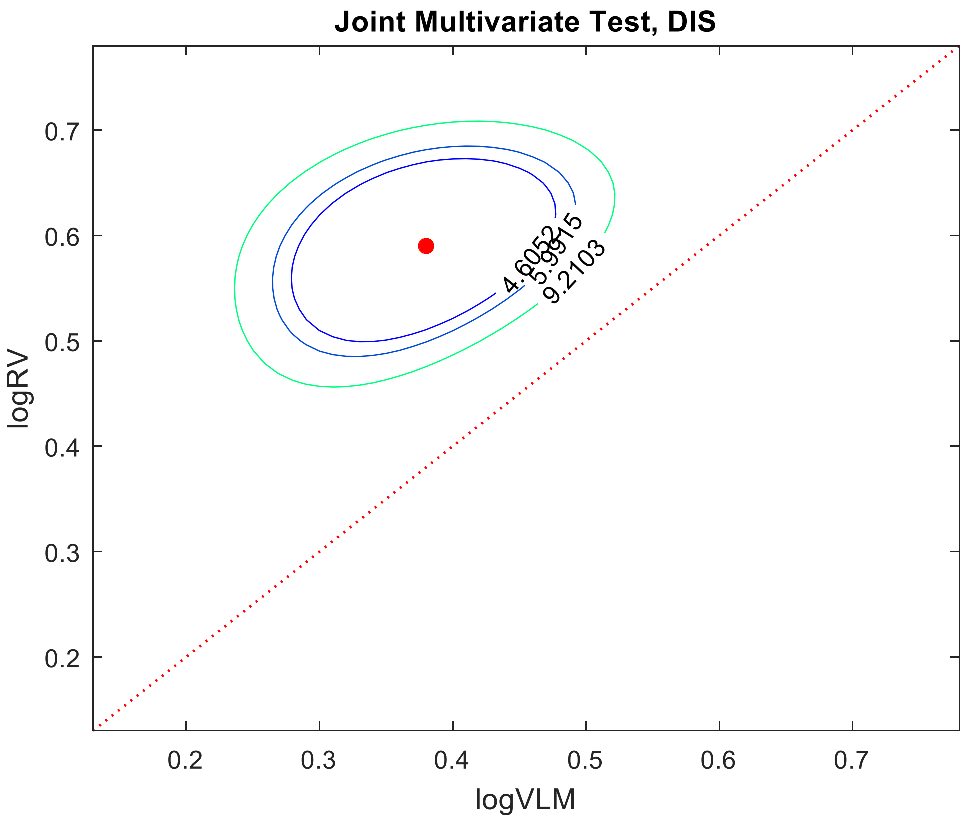

Joint Multivariate Test, BA
logRV
logVLM
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

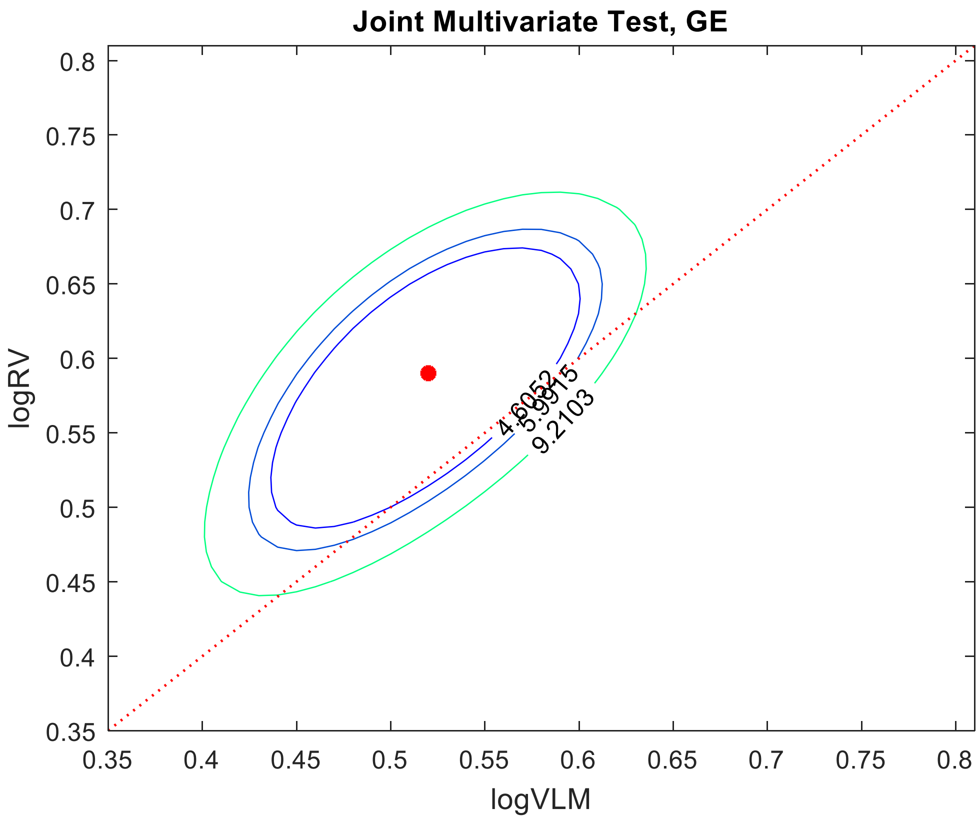

Joint Multivariate Test, CAT
logRV
logVLM
4.6052
5.9915
9.2103
0.2
0.3
0.4
0.5
0.6
0.7
0.8

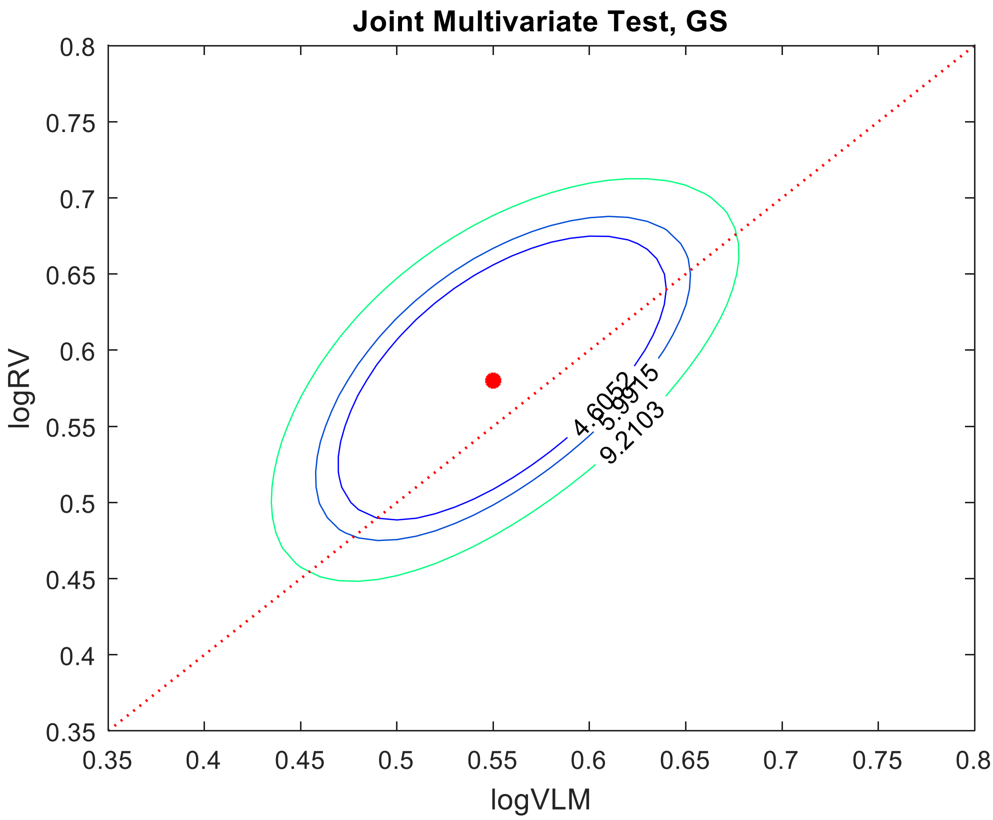

Joint Multivariate Test, CSCO
logRV
logVLM
4.6052
5.9915
9.2103
0.2
0.3
0.4
0.5
0.6
0.7

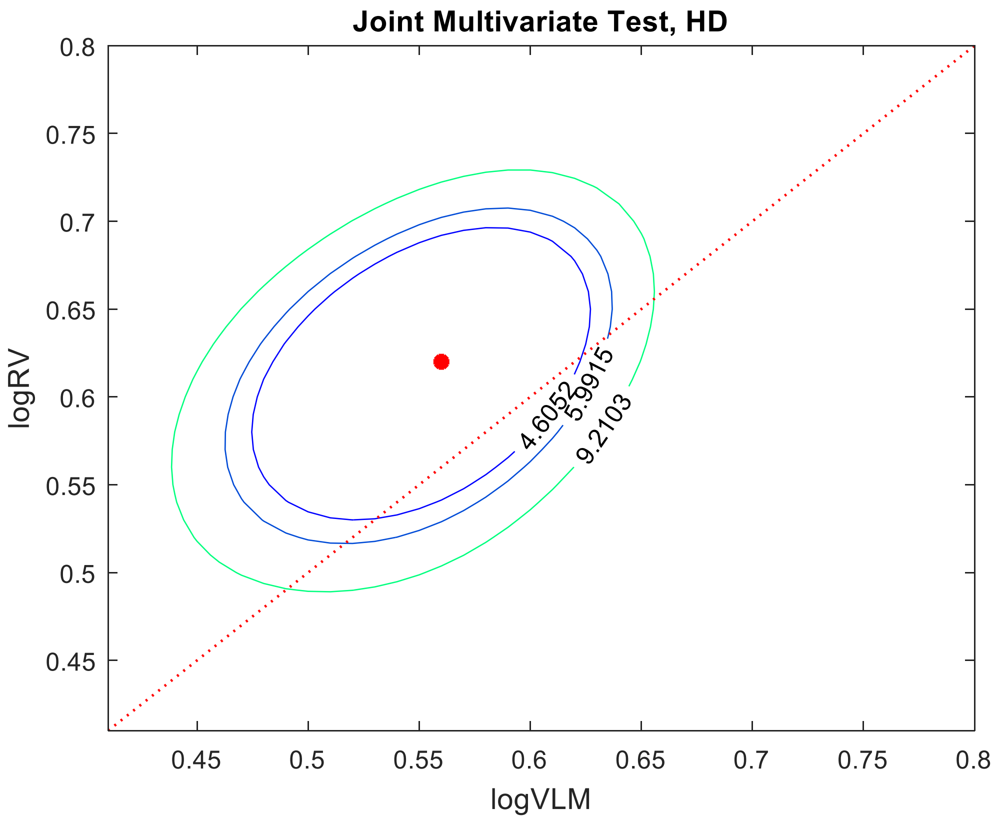

Joint Multivariate Test, CVX
logRV
logVLM
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
4.6052
5.9915
9.2103

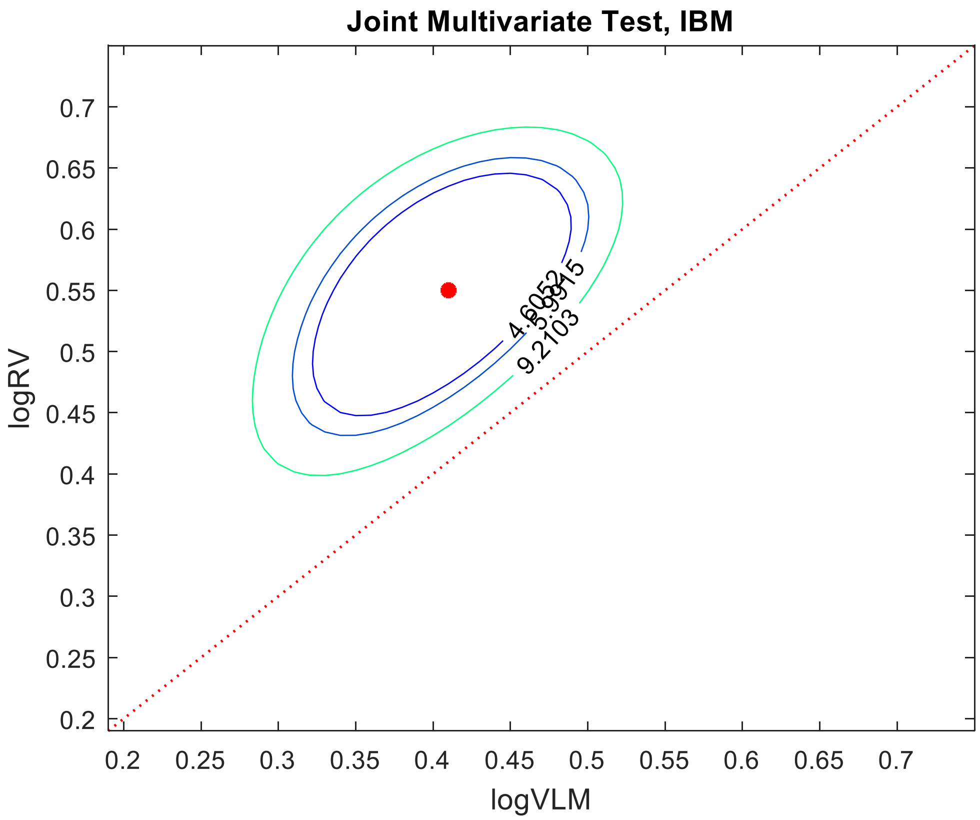

Joint Multivariate Test, DD
logRV
logVLM
0.7
0.6
0.5
0.4
0.3
0.2
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

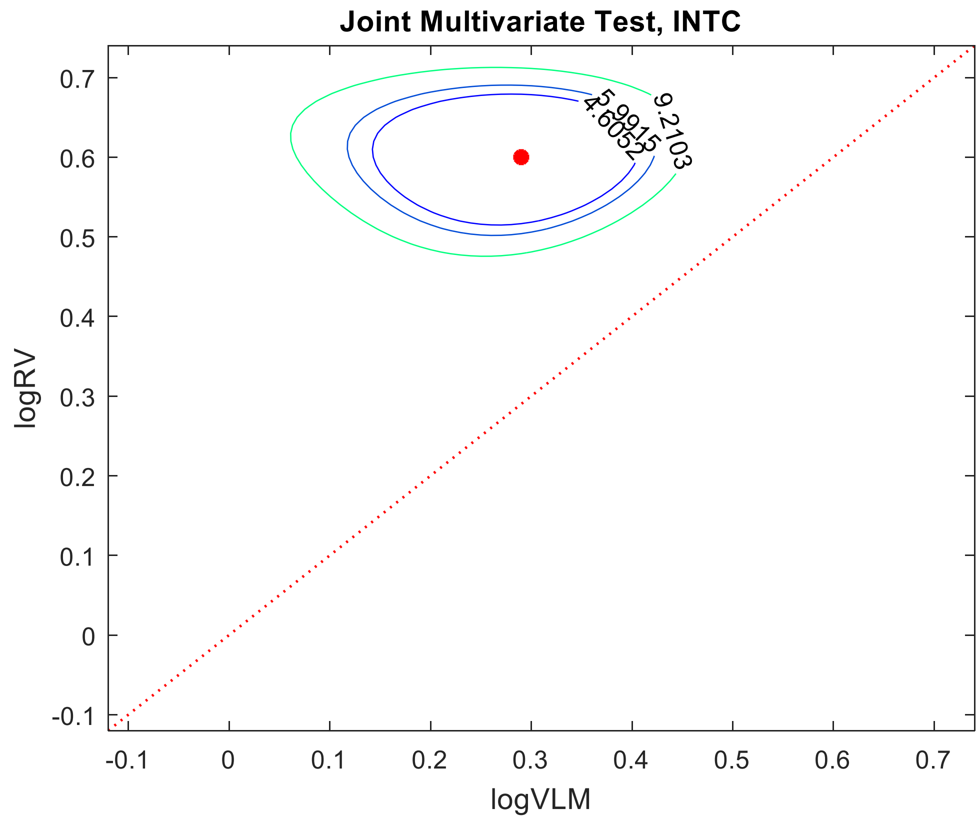

Joint Multivariate Test, DIS
logRV
logVLM
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

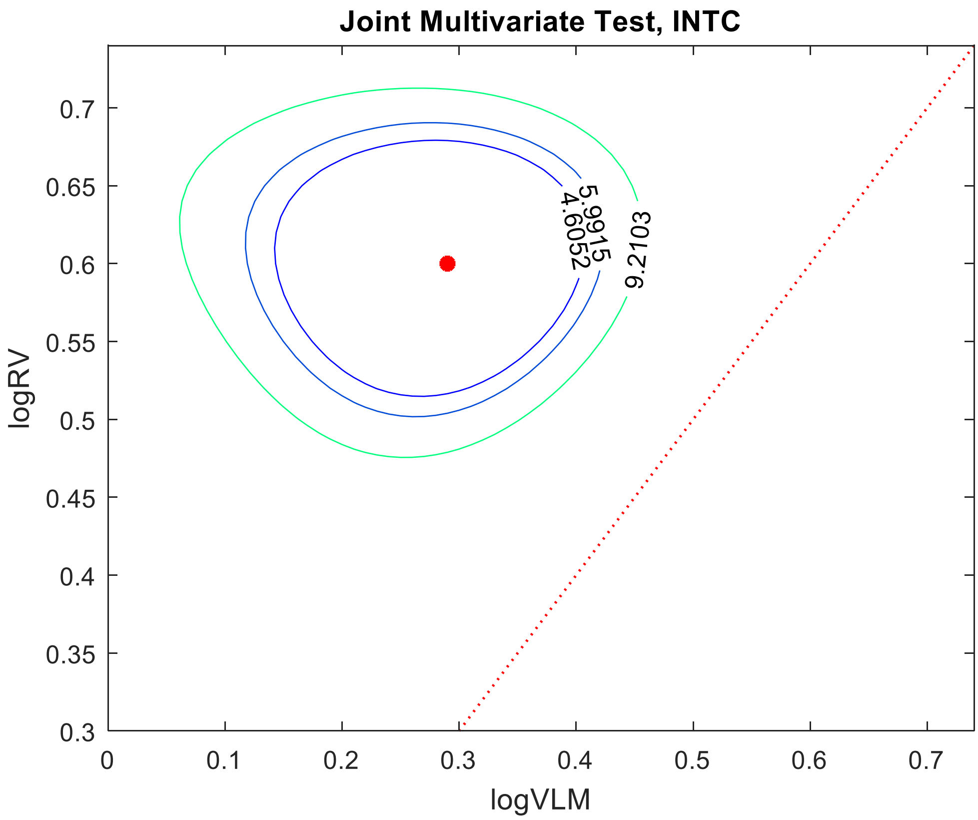

Joint Multivariate Test, GE
logRV
logVLM
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
4.6052
5.9915
9.2103

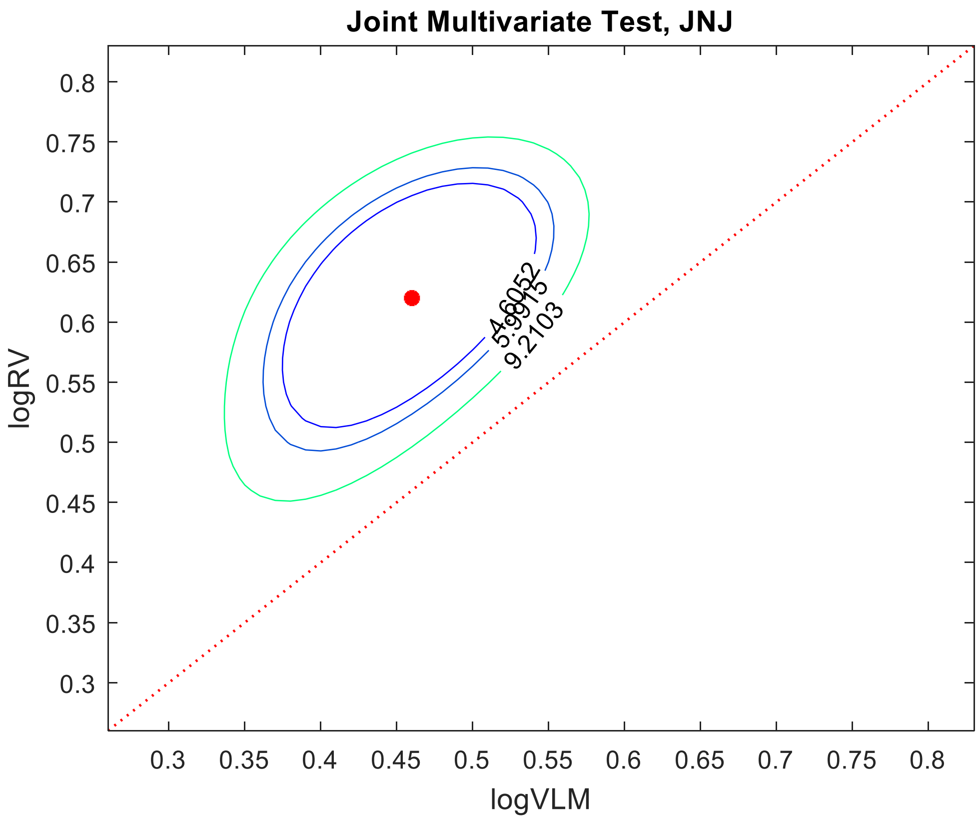

Joint Multivariate Test, GS
logRV
logVLM
4.6052
5.9915
9.2103
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8

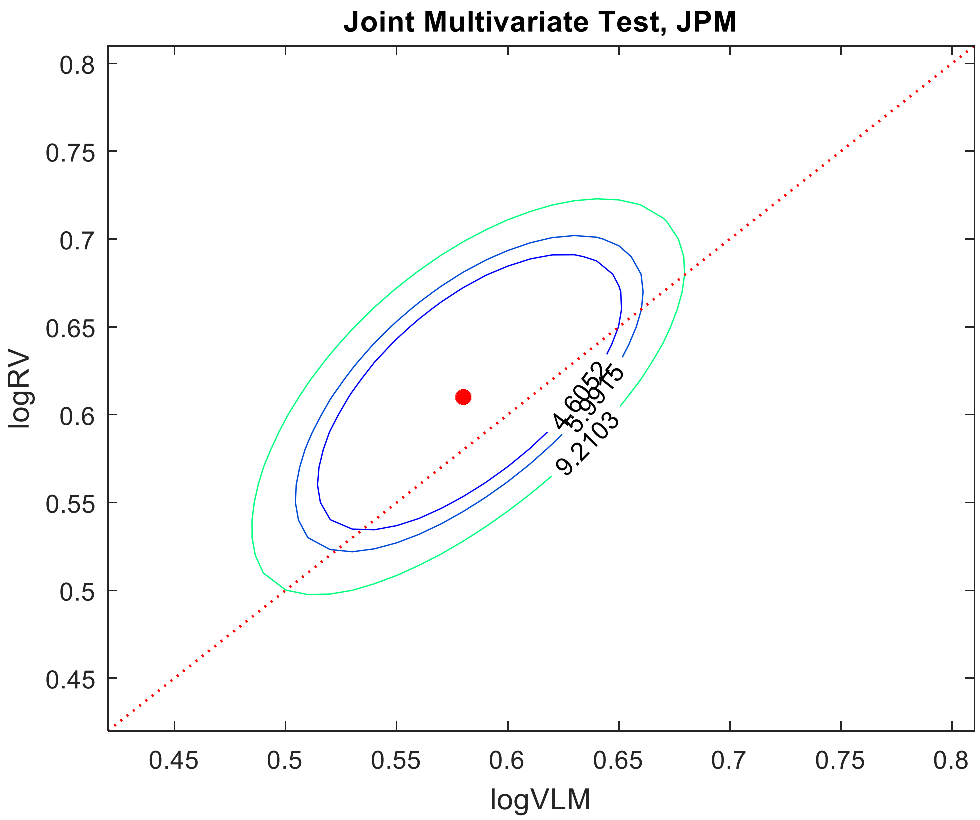

Joint Multivariate Test, HD
logRV
logVLM
4.6052
5.9915
9.2103
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8

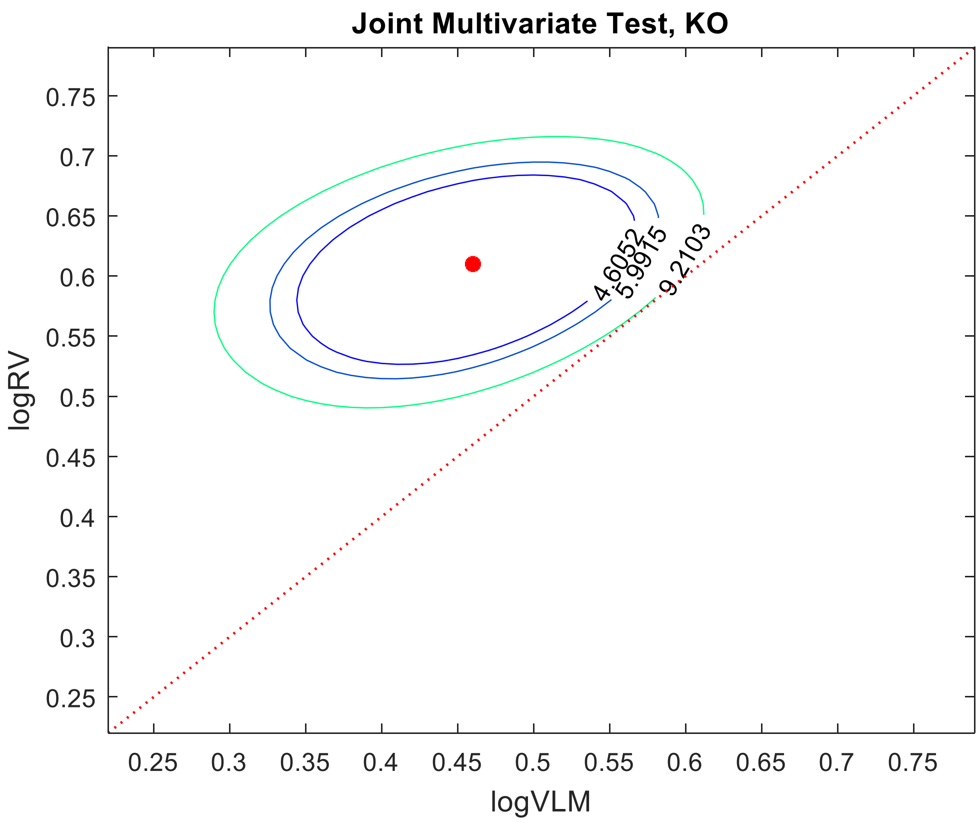

Joint Multivariate Test, IBM
logRV
logVLM
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
4.6052
5.9915
9.2103

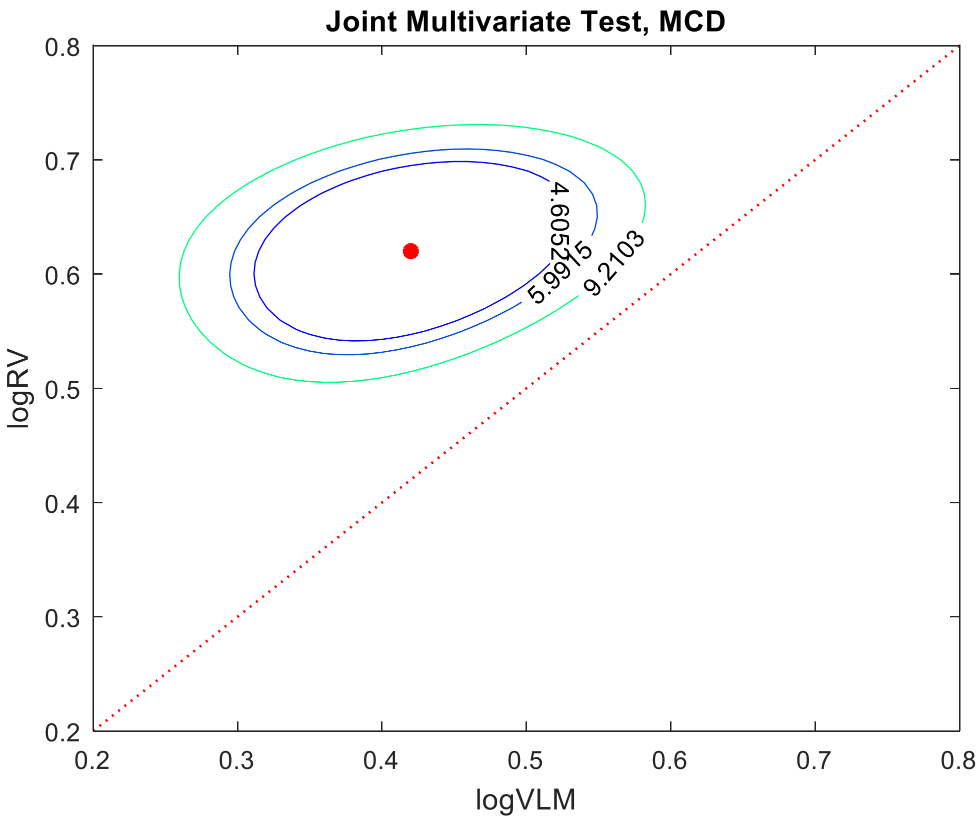

Joint Multivariate Test, INTC
logRV
logVLM
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

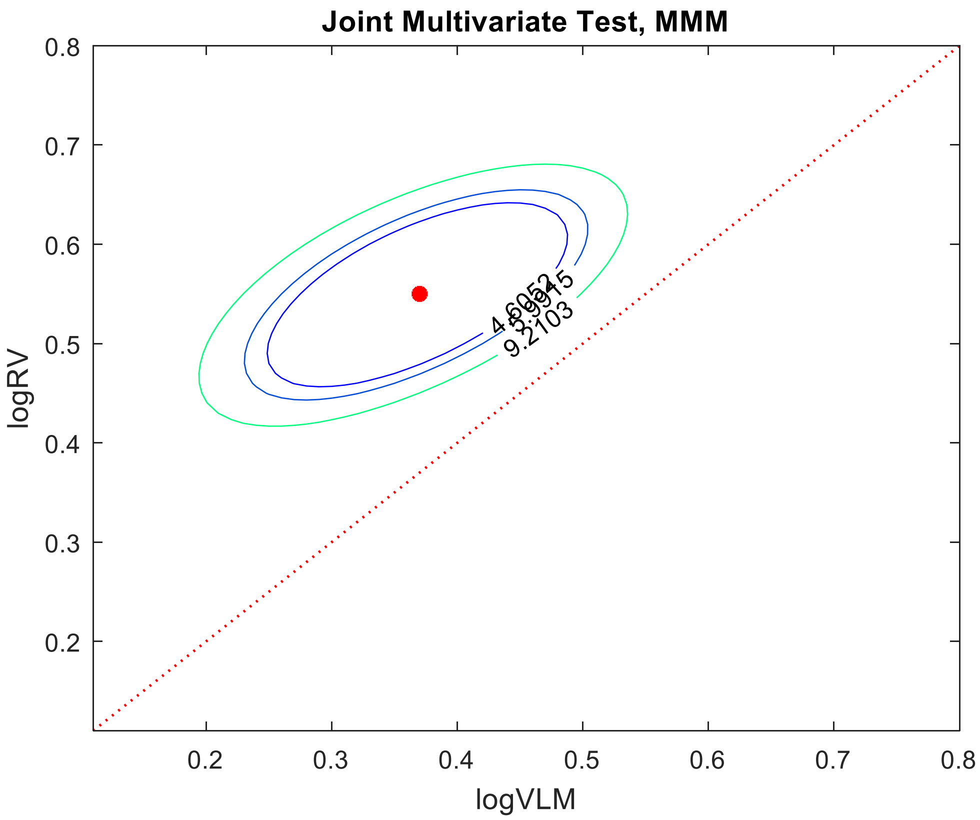

Joint Multivariate Test, INTC
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
logRV
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
logVLM
5.9915
4.6052
9.2103

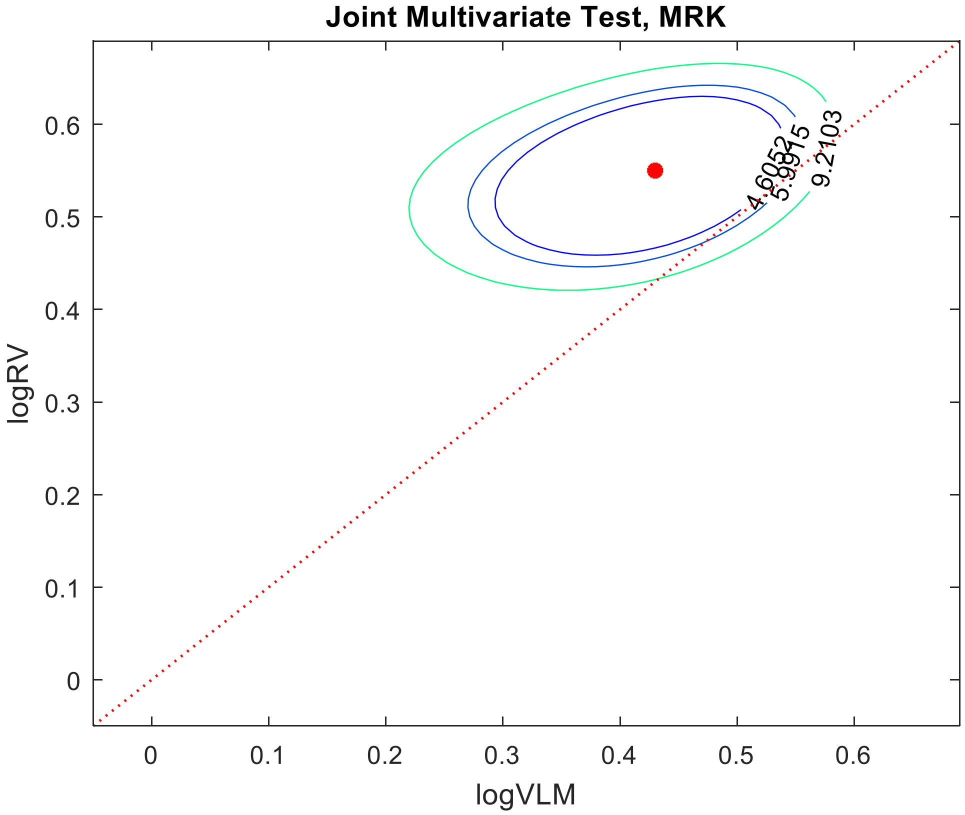

Joint Multivariate Test, JNJ
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
logRV
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
logVLM
4.6052
5.9915
9.2103


Joint Multivariate Test, JPM
logRV
logVLM
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
4.6052
5.9915
9.2103

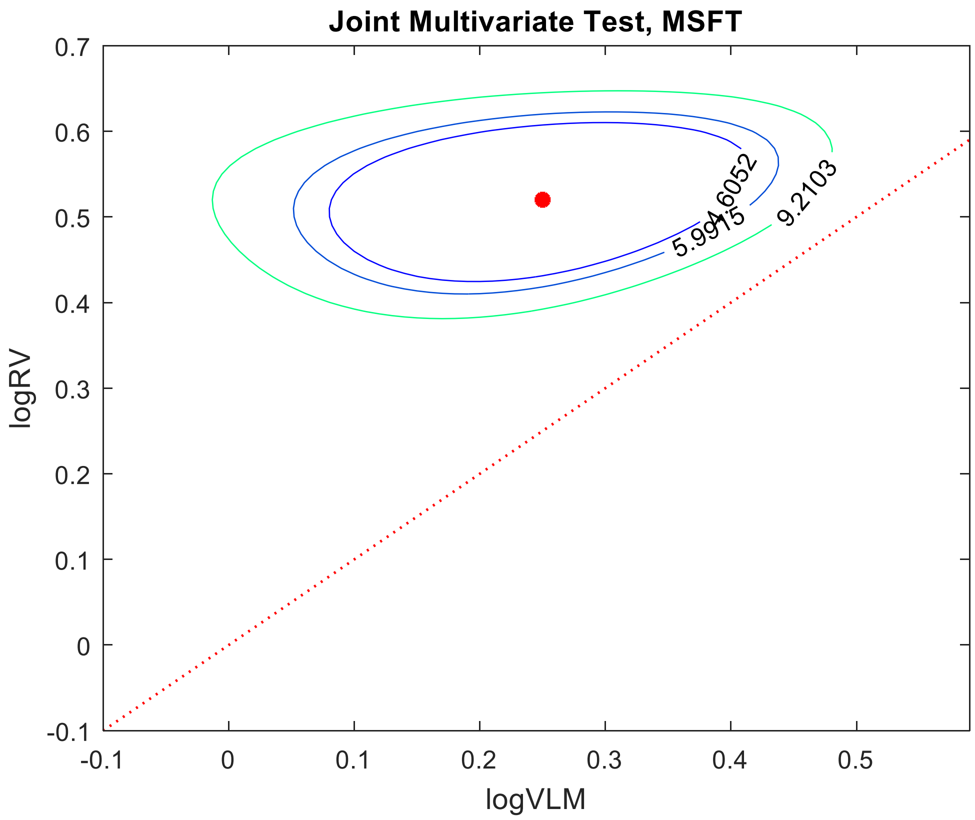

Joint Multivariate Test, KO
logRV
logVLM
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
4.6052
5.9915
9.2103


Joint Multivariate Test, MCD
logRV
logVLM
4.6052
5.9915
9.2103
0.2
0.3
0.4
0.5
0.6
0.7
0.8

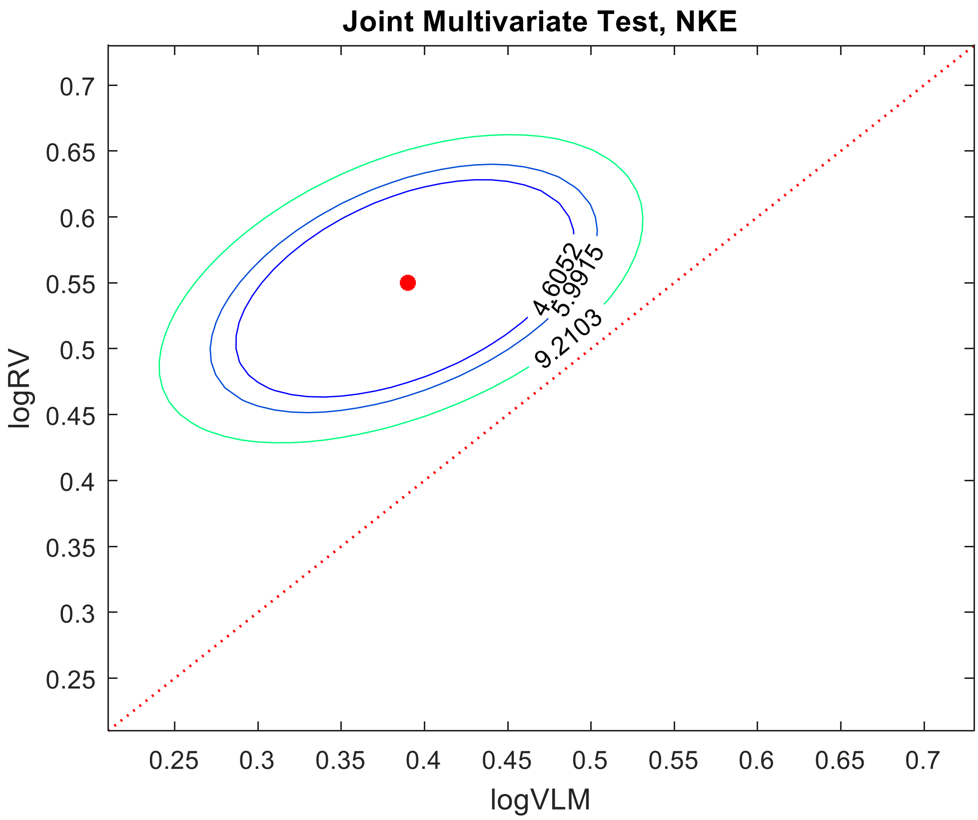

Joint Multivariate Test, MMM
logRV
logVLM
4.6052
5.9915
9.2103
0.2
0.3
0.4
0.5
0.6
0.7
0.8

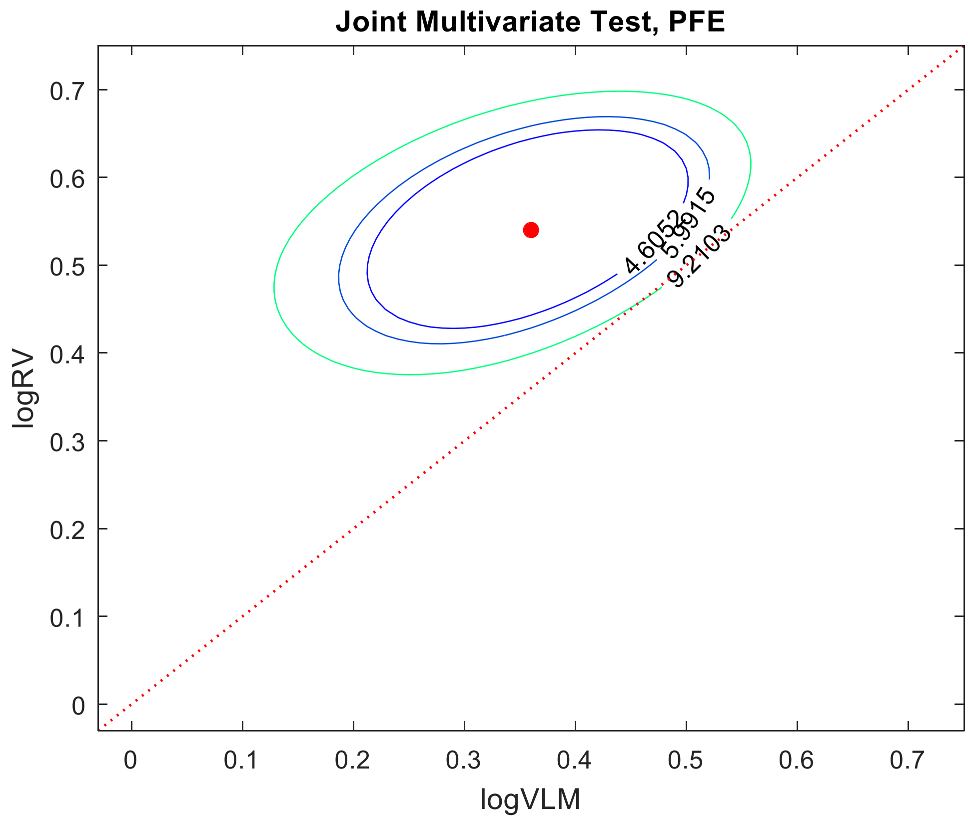


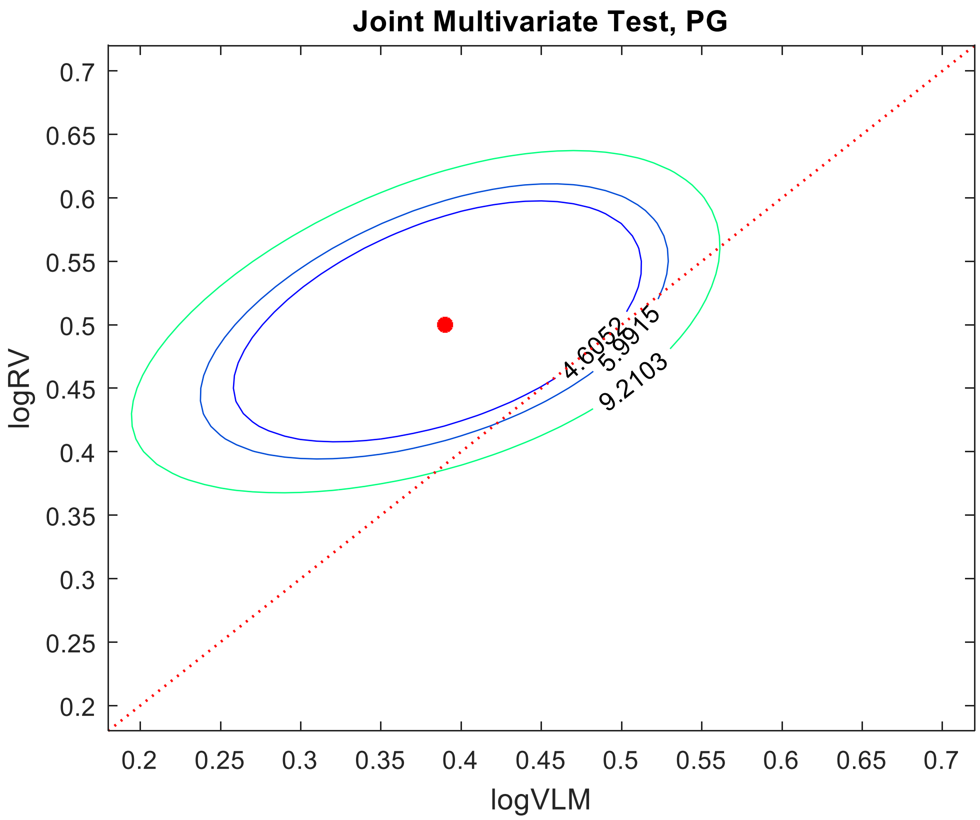
Joint Multivariate Test, MRK
logRV
logVLM
0.6
0.5
0.4
0.3
0.2
0.1
0
0.1
0.2
0.3
0.4
0.5
0.6
4.6052
5.9915
9.2103

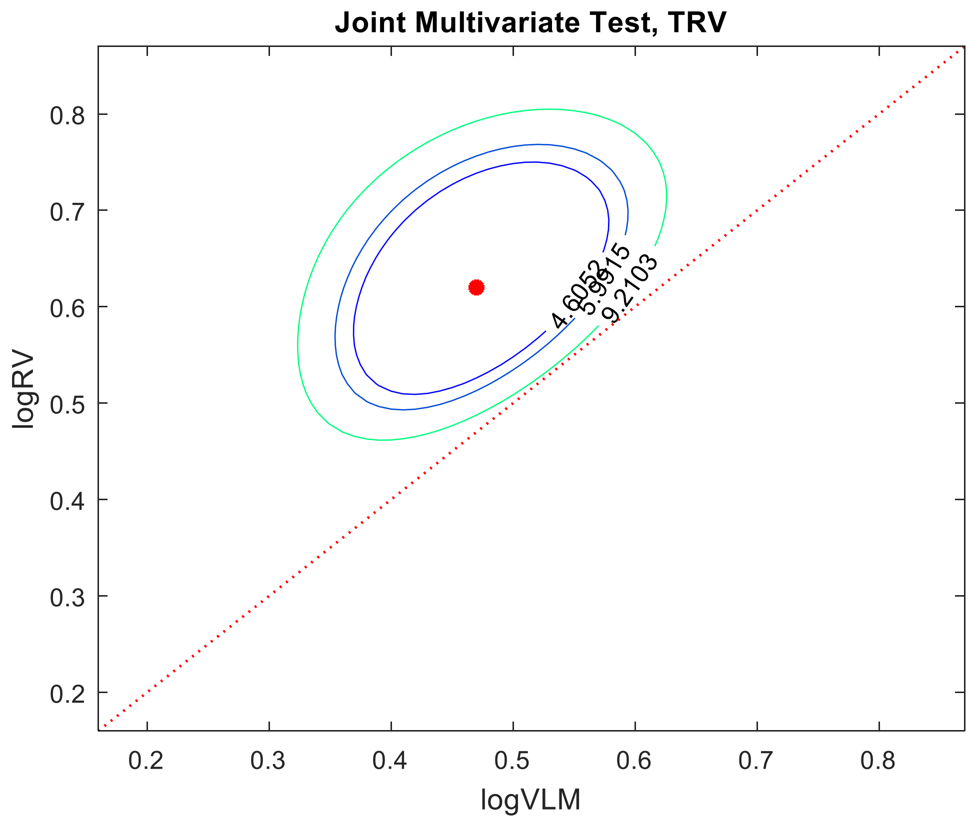


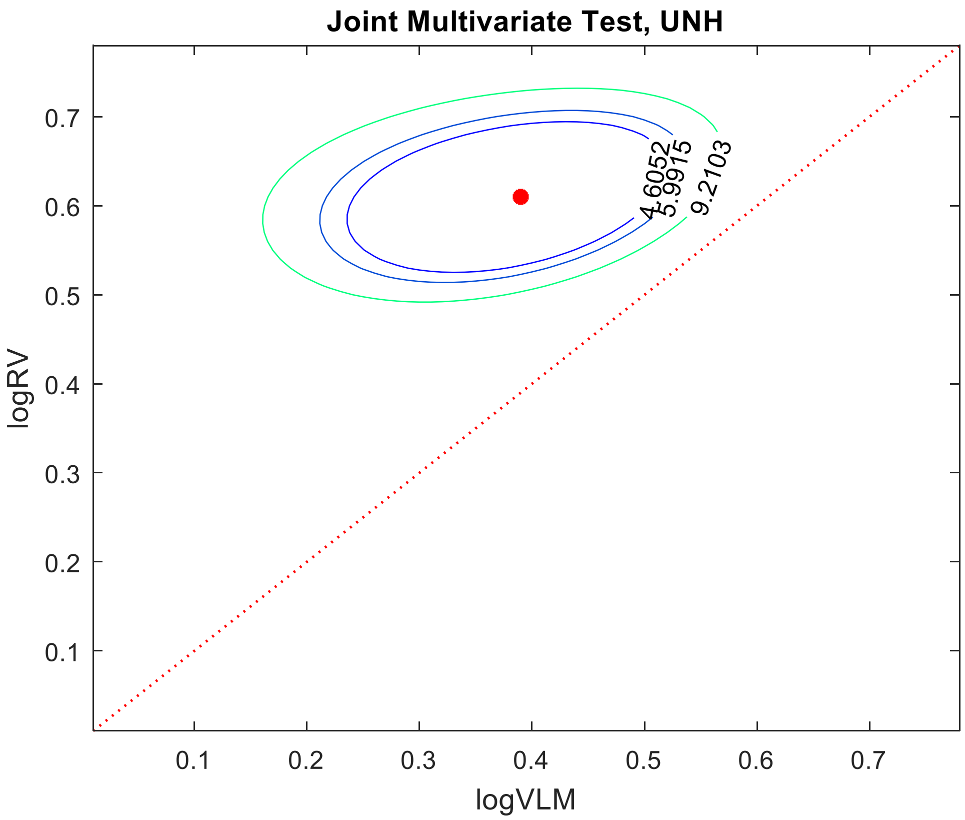
Joint Multivariate Test, MSFT
logRV
logVLM
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
0
0.1
0.2
0.3
0.4
0.5
4.6052
5.9915
9.2103

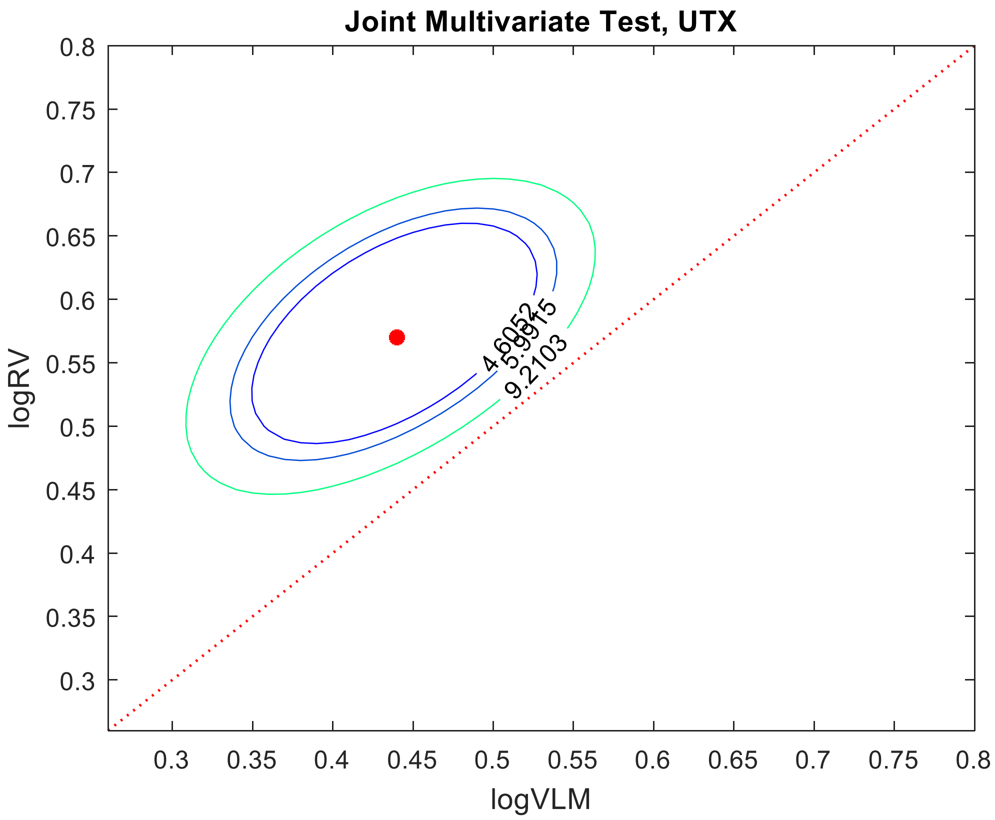

Joint Multivariate Test, NKE
logRV
logVLM
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
4.6052
5.9915
9.2103

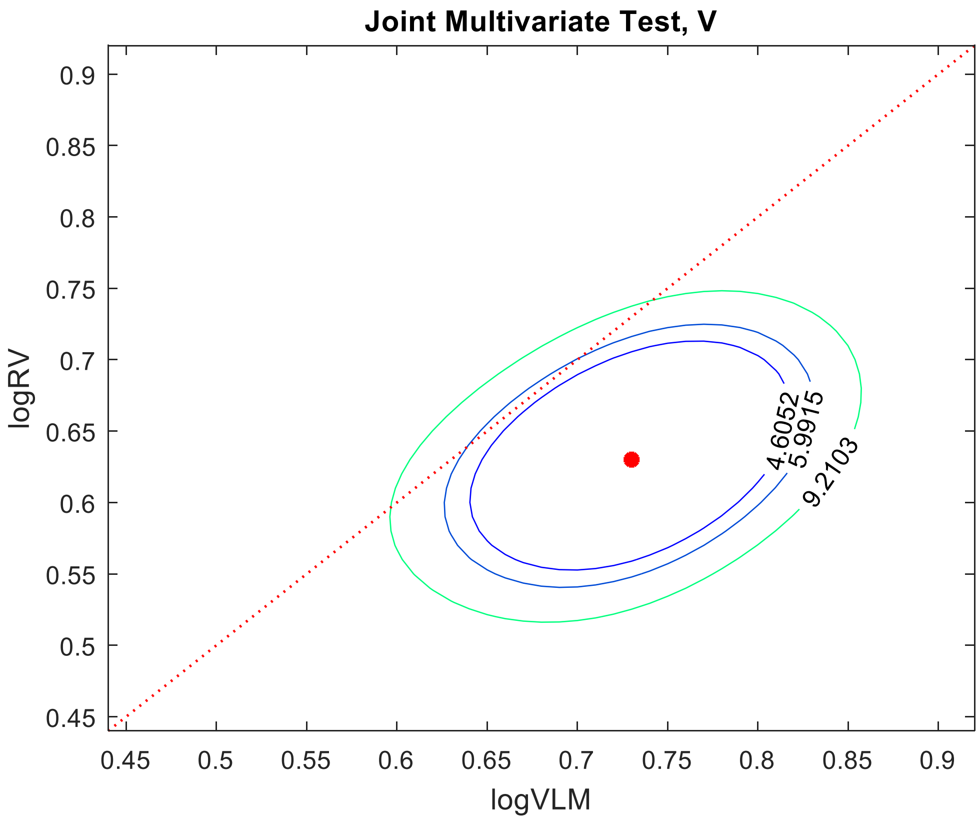

Joint Multivariate Test, PFE
logRV
logVLM
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103

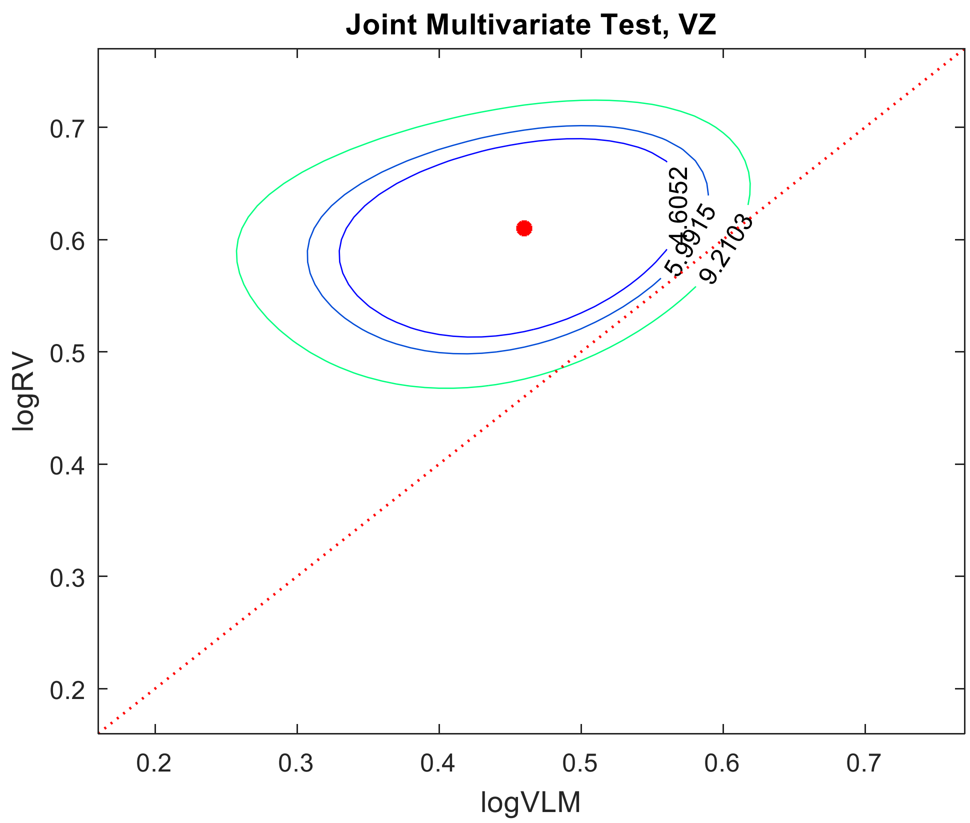

Joint Multivariate Test, PG
logRV
logVLM
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
4.6052
5.9915
9.2103

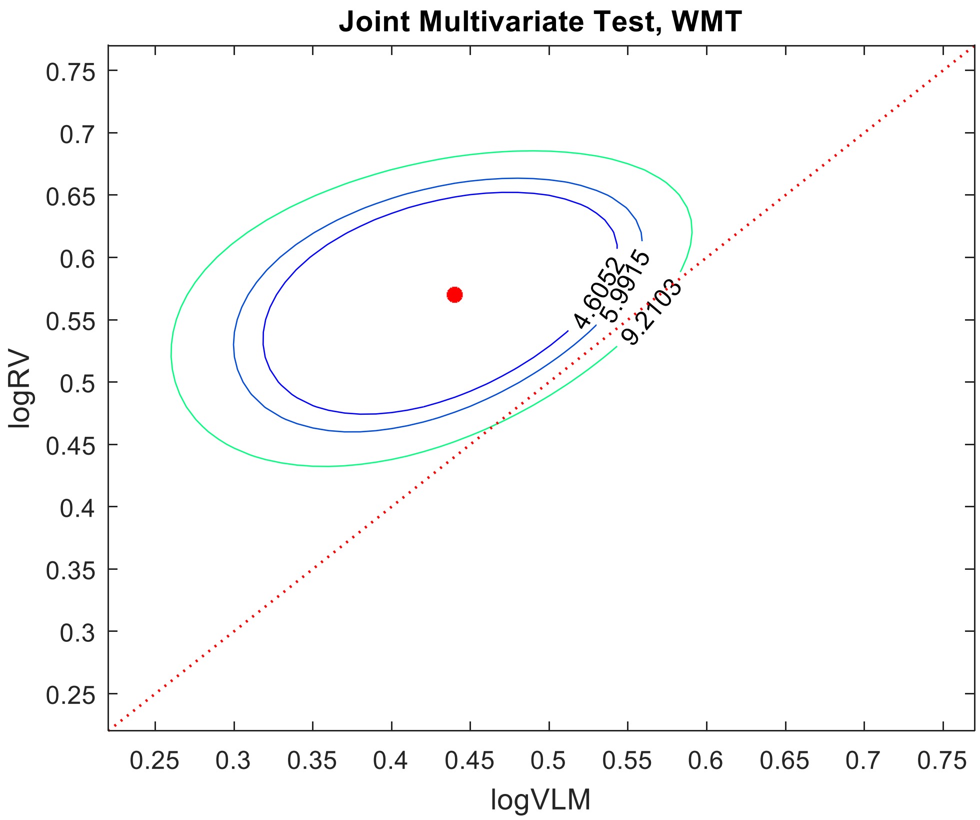

Joint Multivariate Test, TRV
logRV
logVLM
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.3
0.4
0.5
0.6
0.7
0.8
4.6052
5.9915
9.2103

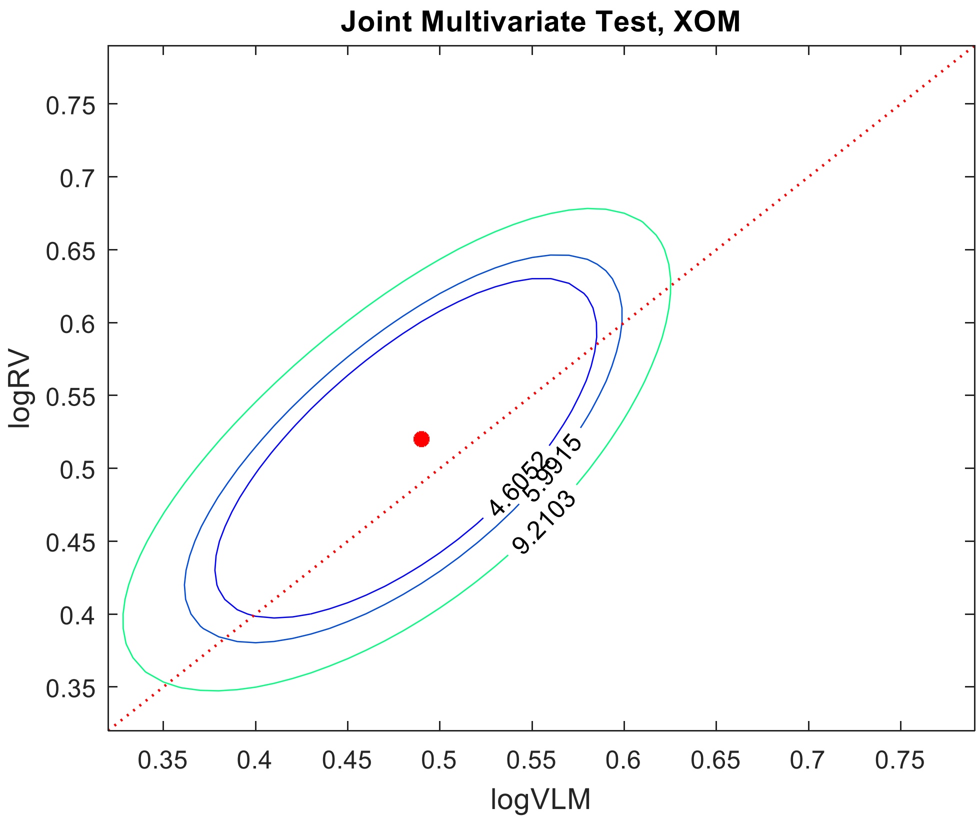

Joint Multivariate Test, UNH
logRV
logVLM
4.6052
5.9915
9.2103
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7


Joint Multivariate Test, UTX
logRV
logVLM
4.6052
5.9915
9.2103
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8


Joint Multivariate Test, V
logRV
logVLM
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
4.6052
5.9915
9.2103


Joint Multivariate Test, VZ
logRV
logVLM
0.7
0.6
0.5
0.4
0.3
0.2
0.2
0.3
0.4
0.5
0.6
0.7
4.6052
5.9915
9.2103


Joint Multivariate Test, WMT
logRV
logVLM
4.6052
5.9915
9.2103
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75


Joint Multivariate Test, XOM
logRV
logVLM
4.6052
5.9915
9.2103
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75