SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2009-22

Co-integration Rank Testing under Conditional Heteroskedasticity

Giuseppe Cavaliere, Anders Rahbek and A.M.Robert Taylor

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Co-integration Rank Testing under Conditional Heteroskedasticity*

Giuseppe Cavaliere[a], Anders Rahbek[b] and A.M.Robert Taylor[c]

[a] Department of Statistical Sciences, University of Bologna

[b] Department of Economics, University of Copenhagen and CREATES

[c] School of Economics and Granger Centre for Time Series Econometrics, University of Nottingham

March 2009

Abstract

We analyse the properties of the conventional Gaussian-based co-integrating rank tests of Johansen (1996) in the case where the vector of series under test is driven by globally stationary, conditionally heteroskedastic (martingale difference) innovations. We first demonstrate that the limiting null distributions of the rank statistics coincide with those derived by previous authors who assume either i.i.d. or (strict and covariance) stationary martingale difference innovations. We then propose wild bootstrap implementations of the co-integrating rank tests and demonstrate that the associated bootstrap rank statistics replicate the first-order asymptotic null distributions of the rank statistics. We show the same is also true of the corresponding rank tests based on the i.i.d. bootstrap of Swensen (2006). The wild bootstrap, however, has the important property that, unlike the i.i.d. bootstrap, it preserves in the re-sampled data the pattern of heteroskedasticity present in the original shocks. Consistent with this, numerical evidence suggests that, relative to tests based on the asymptotic critical values or the i.i.d. bootstrap, the wild bootstrap rank tests perform very well in small samples under a variety of conditionally heteroskedastic innovation processes. An empirical application to the term structure of interest rates is given.

Keywords: Co-integration; trace and maximum eigenvalue rank tests; conditional heteroskedasticity; i.i.d. bootstrap; wild bootstrap.

J.E.L. Classifications: C30, C32.

*Correspondence to: Robert Taylor, School of Economics, University of Nottingham, Nottingham, NG7 2RD, U.K. *E-mail*: `robert.taylor@nottingham.ac.uk`

1 Introduction

In a recent paper, Gonçalves and Kilian (2004) argue that "... the failure of the i.i.d. assumption is well-documented in empirical finance ... many monthly macroeconomic variables also exhibit evidence of conditional heteroskedasticity." (2004,p.92); see Section 2 of Gonçalves and Kilian (2004) for detailed discussion and empirical evidence on this point. Gonçalves and Kilian (2004,2007) show that, so far as inference in stationary univariate autoregressive models is concerned, standard residual-based bootstraps based on an i.i.d. re-sampling scheme are invalid under conditional heteroskedasticity. They demonstrate that in such cases inference based on the wild bootstrap is asymptotically valid and delivers substantial improvements over both residual-based i.i.d. bootstrap tests and standard tests based on asymptotic critical values. Xu (2008) has recently shown that Gonçalves and Kilian's (2004) residual wild bootstrap is also valid in the presence of innovations with non-stationary volatility. Cavaliere and Taylor (2008) show that analogous properties also hold when using wild bootstrap methods in the context of the univariate unit root testing problem.

The trace and maximum eigenvalue co-integrating rank tests of Johansen (1996) are derived under the assumption of Gaussian i.i.d. innovations. In an important Monte Carlo study, Lee and Tse (1996) numerically examine the performance of the rank tests in the presence of GARCH errors. They find that (for the case of non-integrated GARCH errors) the rank tests show a tendency to over-reject the null hypothesis of no co-integration but that this is ameliorated, other things being equal, as the sample size is increased. These findings are consistent with Rahbek, Hansen and Dennis (2002) [RHD] who demonstrate that the assumption required on the innovations can be considerably weakened to that of a (strict and second-order) stationary and ergodic vector martingale difference sequence (with constant unconditional variance and satisfying certain mild regularity conditions) without altering the asymptotic null distributions of the rank statistics. In this paper we first show that these limiting null distributions remain valid in the less restrictive case of globally stationary, conditionally heteroskedastic shocks satisfying certain moment conditions. Moreover, we show that the pseudo maximum likelihood [PML] estimator of the error correction model which assumes Gaussian i.i.d. disturbances remains consistent under these weaker conditions.

Although, the standard rank tests based on asymptotic critical values therefore remain asymptotically valid in the presence of conditionally heteroskedastic shocks, the construction of these tests does not utilise sample information relating to any conditional heteroskedasticity present. Given this result, and the observation of Gonçalves and Kilian (2004) that conditional heteroskedasticity is a relatively common occurrence in macroeconomic and financial time series, it is clearly important and practically relevant to also consider bootstrap testing procedures in the multivariate time series setting which are asymptotically valid in the presence of conditional heteroskedasticity. We therefore develop bootstrap versions of the standard co-integrating rank tests. Our approach builds on the residual-based bootstrap co-integrating rank tests of van Giersbergen (1996), Harris and Judge (1998), Mantalos and Shukur (2001), Trenkler

(2009) and, most notably, Swensen (2006), all of which assume that the innovations are independent and identically distributed (i.i.d.).

Our proposed bootstrap tests are based on the wild bootstrap re-sampling scheme, since, unlike the other bootstrap schemes noted above, this replicates in the re-sampled data the pattern of heteroskedasticity present in the original shocks. The wild bootstrap scheme we use has also been considered in the co-integration rank testing scenario by Cavaliere, Rahbek and Taylor (2007) [CRT] in the fundamentally different scenario where the innovations display globally non-stationary volatility; that is, cases where the *unconditional* variance of the innovation vector varies over time in a systematic fashion. CRT demonstrate that in such cases, under the assumption of an absence of any conditional heteroskedasticity, the conventional co-integrating rank statistics do not have the same form as given in Johansen (1996), rather they depend on nuisance parameters relating to the underlying volatility process. They demonstrate, however, that the wild bootstrap rank statistics can replicate this limit distribution, to first order. Consequently, although the wild bootstrap algorithm we use here is the same as that in CRT, it is being used in the context of a quite different statistical model.

We show that wild bootstrap co-integrating rank statistics replicate the first-order asymptotic null distributions of the rank statistics, such that the corresponding bootstrap tests are asymptotically valid, in the presence of conditionally heteroskedastic innovations. The same is shown to be true of the corresponding i.i.d. bootstrap tests of Swensen (2006). It is not our aim in this paper to establish that the wild bootstrap provides a superior approximation to the conventional asymptotic approximation or to the i.i.d. bootstrap approximation. Rather we detail a less restrictive set of conditions than is adopted in the extant literature under which both the asymptotic test and both the wild and i.i.d. bootstrap approaches are asymptotically valid. However, since the wild bootstrap incorporates sample information on the conditional heteroskedasticity where present, one might anticipate that the wild bootstrap would provide a superior approximation to that provided by the asymptotic and i.i.d. bootstrap approximations which do not incorporate such sample information. Simulation evidence for a variety of conditionally heteroskedastic innovation models is supportive of this view. Taken together, the results in this paper coupled with those in CRT demonstrate that the wild bootstrap is a very powerful and useful tool, able to handle time-dependent behaviour in both the conditional and unconditional variance of the innovations. The question of whether there are conditions under which the wild bootstrap approach will provide asymptotic refinements is left for future research.

The paper is organized as follows. Section 2 outlines our reference co-integrated conditionally heteroskedastic VAR model, while section 3 establishes the large sample behaviour of the standard rank statistics and the MLE of the parameters from this model. Our wild bootstrap-based approach, which also incorporates a sieve procedure using the (consistently) estimated coefficient matrices from the co-integrated VAR model, is outlined in Section 4. The first-order asymptotic validity of both this approach and that based on the i.i.d. re-sampling bootstrap rank tests of Swensen (2006) are demonstrated. In Section 5, the finite sample properties of the tests are explored through Monte Carlo methods and compared with the standard asymptotic tests and with the i.i.d. bootstrap tests, for a variety of conditionally heteroskedastic error processes. In section 6 we apply our tests to bond market data from several major economies. Section 7 concludes. All proofs are contained in the Appendix.

In the following '$\xrightarrow{w}$' denotes weak convergence, '$\xrightarrow{p}$' convergence in probability, and '$\xrightarrow{w}_p$' weak convergence in probability (Giné and Zinn, 1990; Hansen, 1996), in each case as the sample size diverges to positive infinity; $\mathbb{I}(\cdot)$ denotes the indicator function and '$x := y$' ('$x =: y$') indicates that x is defined by y (y is defined by x); $\lfloor\cdot\rfloor$ denotes the integer part of its argument. The space spanned by the columns of any $m \times n$ matrix A is denoted as $\mathsf{col}(A)$; if A is of full column rank $n < m$, then $A_\perp$ denotes an $m \times (m-n)$ matrix of full column rank satisfying $A_\perp' A = 0$. For any square matrix, A, $|A|$ is used to denote the determinant of A, $\|A\|$ the norm $\|A\|^2 := \mathsf{tr}\{A'A\}$, where $\mathsf{tr}\{A\}$ denotes the trace of A, and $\rho(A)$ its spectral radius (that is, the maximal modulus of the eigenvalues of A). For any vector, x, $\|x\|$ denotes the usual Euclidean norm, $\|x\| := (x'x)^{1/2}$.

2 The Conditionally Heteroskedastic Co-integration Model

We consider the following VAR(k) model in error correction format:

$$\Delta X_t = \Pi X_{t-1} + \Psi U_t + \mu D_t + \varepsilon_t,\ t = 1, ..., T \tag{2.1}$$

where: X_t and ε_t are $p \times 1$, $U_t := \left(\Delta X_{t-1}', ..., \Delta X_{t-k+1}'\right)'$ is $p(k-1) \times 1$ and $\Psi := (\Gamma_1, ..., \Gamma_{k-1})$, where $\{\Gamma_i\}_{i=1}^{k-1}$ are $p \times p$ lag coefficient matrices and the impact matrix $\Pi := \alpha\beta'$ where α and β are full column $p \times r$ matrices, $r \leq p$. The term D_t collects all deterministic components, and in this paper we focus on the leading case of a linear trend, $D_t := (1, t)'$, with associated coefficients $\mu := (\mu_1', \mu_2')'$. The initial values, $\mathbb{X}_0 := \left(X_0', ..., X_{-k+1}'\right)'$, are taken to be fixed.

Throughout the paper the process in (2.1) is assumed to satisfy the following assumptions.

Assumption 1: (a) *All of the characteristic roots associated with (2.1), that is the solutions to the characteristic equation* $A(z) := (1-z) I_p - \alpha\beta' z - \Gamma_1 z (1-z) - \cdots - \Gamma_{k-1} z^{k-1} (1-z) = 0$, *lie either outside the unit circle or are equal to unity;* (b) $\det(\alpha_\perp' \Gamma \beta_\perp) \neq 0$, *with* $\Gamma := I_p - \Gamma_1 - \cdots - \Gamma_{k-1}$.

Assumption 2: *The innovations* $\{\varepsilon_t\}$ *form a martingale difference sequence with respect to the filtration* $\mathcal{F}_t$, *where* $\mathcal{F}_{t-1} \subseteq \mathcal{F}_t$ *for* $t = ..., -1, 0, 1, 2, ...$, *satisfying: (i) the global homoskedasticity condition:*

$$\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\left(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}\right) \xrightarrow{p} \Sigma > 0, \tag{2.2}$$

and (ii) $\mathrm{E}\left\|\varepsilon_t\right\|^4 \leq K < \infty$.

Remark 2.1. While Assumption 1 is standard in the co-integration testing literature, Assumption 2 is not. Assumption 2 implies that ε_t is a serially uncorrelated, potentially conditionally heteroskedastic process. This contrasts with the assumption that ε_t is i.i.d. as made in Johansen (1996) and Swensen (2006). Moreover, and in contrast to RHD, condition (i) of Assumption 2 imposes neither strict nor second-order stationarity on ε_t, but rather imposes a so-called *global stationarity* or *global homoskedasticity* condition; see e.g. Davidson (1994,pp.454-455). In particular, this condition allows the conditional (and, hence,[1] unconditional) variance of ε_t to change over time, provided that it is asymptotically stable over all possible fixed fractions of the data; that is, provided

$$\frac{1}{T\left(s'-s\right)} \sum_{t=\lfloor Ts\rfloor+1}^{\lfloor Ts'\rfloor} \mathrm{E}\left(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}\right) \xrightarrow{p} \Sigma \tag{2.3}$$

for all $s' < s \in [0,1]$. This framework therefore allows for, among other things, stable (G)ARCH models with initial values that are not drawn from the invariant distribution, and models which exhibit seasonal heteroskedasticity. An example of the latter is given by the case where ε_t satisfies $\mathrm{E}\left(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}\right) = \mathrm{E}\left(\varepsilon_t\varepsilon_t'\right) = \sum_{i=1}^{S}\Sigma^{(i)}d_{it}$, $t = 1,...,T$, where the d_{it}, $i = 1,...,S$, are standard seasonal dummies (S being the number of seasons) which is clearly not a covariance stationary process (unless $\Sigma^{(1)} = \cdots = \Sigma^{(S)}$) but is nonetheless globally homoskedastic because condition (2.3) holds. Notice, however, that (2.3) is not in general satisfied in the non-stationary volatility setting of CRT, where the right member of (2.3) depends on s', s.

Remark 2.2. Under Assumption 2, a multivariate functional central limit theorem [FCLT] as in Brown (1971, Theorem 3) applies to ε_t; *viz*,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor T\cdot\rfloor} \varepsilon_t \xrightarrow{w} W\left(\cdot\right), \tag{2.4}$$

where W is a Brownian motion with covariance matrix Σ, the latter defined as in (2.2). This result follows from Assumption 2(i) and since the finite fourth order moment requirement in Assumption 2(ii) implies the Lindeberg-type condition

$$T^{-1}\Sigma_{t=1}^{T}\mathrm{E}\left(\left\|\varepsilon_t\right\|^2 \cdot \mathbb{I}\left\{\left\|\varepsilon_t\right\| > \delta\sqrt{T}\right\}\middle|\, \mathcal{F}_{t-1}\right) \xrightarrow{p} 0\ .$$

Assumption 2 also ensures that conditions (5) and (6) in Hannan and Heyde (1972, Theorem 1) hold, implying that for any linear process s_t of the form $s_t := \sum_{i=0}^{\infty}\theta_i\varepsilon_{t-i}$ with $\sum_{i=0}^{\infty}\left\|\theta_i\right\| < \infty$, the empirical average, $T^{-1}\sum_{i=1}^{T}s_i$, and empirical autocovariances, $T^{-1}\sum_{t=1}^{T}s_ts_{t+k}'$, converge in probability to zero and $\sum_{i=0}^{\infty}\theta_i\Sigma\theta_{i+k}'$, respectively.

[1]Specifically, the condition in (2.2) coupled with the assumption of finite fourth order moments implies that the unconditional variances of ε_t, $t = 1,...,T$, satisfy $\lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\mathrm{E}\left(\varepsilon_t\varepsilon_t'\right) = \Sigma$.

Remark 2.3. The conditions in Assumption 2 ensure that a FCLT applies to the MDS, $\{\varepsilon_t\}$, and that the product moments converge, as detailed in Remark 2.2. Both the convergence in (2.2) and the convergence of the product moments would also be implied by assuming geometric ergodicity of the $\{\varepsilon_t\}$ sequence, since the law of large numbers applies to functions of geometrically ergodic processes; see Jensen and Rahbek (2007) for details. Geometric ergodicity is satisfied for a rich class of (G)ARCH processes; see, for example, the discussion in Kristensen and Rahbek (2005, 2009) and the references therein.

For unknown parameters α, β, Ψ, μ, and when α and β are $p \times r$ matrices, not necessarily of full rank, (2.1) denotes our conditionally heteroskedastic co-integrated VAR model, which we denote as $H(r)$. The model may then be written in the compact form

$$Z_{0t} = \alpha\beta^{*\prime}Z_{1t} + \delta Z_{2t} + \varepsilon_t \tag{2.5}$$

with $Z_{0t} := \Delta X_t$, and the remaining terms defined according to the following three leading cases for the deterministic terms (see, e.g., Johansen, 1996, p.81):

(i) $\mu D_t = 0$ in (2.1), which implies that $Z_{1t} := X_{t-1}$, $Z_{2t} := U_t$, $\beta^* = \beta$ and $\delta = \Psi$ (no deterministic components);

(ii) $\mu D_t = \mu_1 = \alpha\rho_1'$ in (2.1), which implies that $Z_{1t} := (X_{t-1}', 1)'$, $Z_{2t} := U_t$, $\beta^* = (\beta', \rho_1')'$ and $\delta = \Psi$ (restricted constant);

(iii) $\mu D_t = \mu_1 + \mu_2 t$ with $\mu_2 = \alpha\rho_2'$ in (2.1), which implies that $Z_{1t} := (X_{t-1}', t)'$, $Z_{2t} := (U_t', 1)'$, $\beta^* = (\beta', \rho_2')'$ and $\delta = (\Psi, \mu_1)$ (restricted linear trend).

3 Pseudo LR Tests

As is standard, let $M_{ij} := T^{-1}\sum_{t=1}^{T} Z_{it}Z_{jt}'$, $i, j = 0, 1, 2$, with Z_{it} defined as in (2.5), and let $S_{ij} := M_{ij.2} := M_{ij} - M_{i2}M_{22}^{-1}M_{2j}$, $i, j = 0, 1$. Under the auxiliary assumption of i.i.d. Gaussian disturbances, the pseudo Gaussian likelihood function depends on the vector $\theta^{PML} := (\alpha, \beta, \Psi, \mu, \Sigma)$, with μD_t satisfying one of the three cases considered at the end of the previous section (throughout we also apply the usual norming or identification as in Johansen, 1996, section 13.2). We denote the corresponding PML estimator as $\hat{\theta}^{PML} := (\hat{\alpha}, \hat{\beta}, \hat{\Psi}, \hat{\mu}, \hat{\Sigma})$. Write the maximized (pseudo) log-likelihood under $H(r)$, say $\ell(r)$, as

$$\ell(r) = -\frac{T}{2}\log|S_{00}| - \frac{T}{2}\sum_{i=1}^{r}\log\left(1 - \hat{\lambda}_i\right) \tag{3.1}$$

where $\hat{\lambda}_1 > \ldots > \hat{\lambda}_p$, solve the eigenvalue problem

$$\left|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}\right| = 0. \tag{3.2}$$

The pseudo LR (PLR) test for $H(r)$ vs $H(p)$ then rejects for large value of the statistic

$$Q_r := -2\left(\ell\left(r\right)-\ell\left(p\right)\right) = -T\sum_{i=r+1}^{p}\log\left(1-\hat{\lambda}_i\right). \tag{3.3}$$

We now demonstrate the validity of the following theorem concerning the limiting null distribution of the Q_r statistic under conditional heteroskedasticity of the form specified in Assumption 2. To keep the presentation simple we consider, for the present, the case of no deterministics in the model and in the estimation (so that $\hat{\mu}$ is omitted from the definition of $\hat{\theta}^{PML}$ above). This will be subsequently relaxed in Remark 2.5.

Theorem 1 *Let $\{X_t\}$ be generated as in* (2.1) *under Assumptions 1 and 2, with $\mu = 0$. Then, under the hypothesis $H(r)$,*

$$Q_r \xrightarrow{w} \mathsf{tr}(\mathcal{Q}_B) =: Q_{r,\infty} \tag{3.4}$$

where

$$\mathcal{Q}_B := \int_0^1 (dB(u))B(u)'\left(\int_0^1 B(u)B(u)'du\right)^{-1}\int_0^1 B(u)(dB(u))' \tag{3.5}$$

with $B(\cdot)$ a $(p-r)$-variate standard Brownian motion.

Remark 3.1. The representation for the limiting null distribution of Q_r given in (3.4) coincides with that given in Johansen (1996) for the case of independent Gaussian innovations and in RHD for covariance stationary martingale difference innovations.

Remark 3.2. The result in Theorem 1 can be generalized to cover the two additional cases for the deterministic component considered just below (2.5). It is an entirely straightforward extension of the result in Theorem 1 to establish that in such a case the asymptotic null distribution of Q_r is given by (3.4) but now with $\mathcal{Q}_B := \mathsf{tr}(\int(dB(u))F(u)'\left(\int F(u)F(u)'\right)^{-1}\times\int F(u)(dB(u))')$, where B is as defined in Theorem 1, and F is a function of B whose precise form depends on the deterministic term. More specifically, decomposing B as $B := (B_1', B_2)'$, where B_2 is one-dimensional and using the notation $a|b := a(\cdot) - \int a(s)b(s)'ds(\int b(s)b(s)'ds)^{-1}b(\cdot)$ to denote the projection residuals of a onto b:

(i) if $\mu D_t = 0$ in (2.1), then $F := B$, as in Theorem 1;

(ii) if $\mu D_t = \alpha\rho_1'$ in (2.1), then $F := (B', 1)'$;

(iii) if $\mu D_t = \mu_1 + \alpha\rho_2' t$ in (2.1), then $F := (B', u|1)'$.

Remark 3.3. The preceding discussion extends to the so-called maximum eigenvalue test; that is, a PLR test based for $H(r)$ vs $H(r+1)$. As is well known, this test rejects for large values of the statistic $Q_{r,\max} := -2\left(\ell\left(r\right)-\ell\left(r+1\right)\right) = -T\log(1-\hat{\lambda}_{r+1})$, see, for example, Equation (6.19) of Johansen (1996). It then follows trivially from the preceding results that the null asymptotic distribution of $Q_{r,\max}$ corresponds to the distribution of the maximum eigenvalue of the real symmetric random matrix $\mathcal{Q}_B$.

Remark 3.4. As in Johansen (1996), under $H(r)$, the r largest eigenvalues solving (3.2), $\hat{\lambda}_1, \ldots, \hat{\lambda}_r$, converge in probability to positive numbers, while $T\hat{\lambda}_{r+1}, \ldots, T\hat{\lambda}_p$ are of $O_p(1)$. Consequently, the PLR test based on either Q_r or $Q_{r,\max}$ will be consistent at rate $O_p(T)$ if the true co-integration rank is, say, $r_0 > r$. This implies, therefore, that the sequential approach to determining the co-integration rank[2] outlined in Johansen (1996) will still lead to the selection of the correct co-integrating rank with probability $(1-\xi)$ in large samples, as in the i.i.d. Gaussian case. The same results also hold under cases (ii) and (iii) of Remark 3.2.

We conclude this section by demonstrating that even though based on a mis-specified model[3] the PML estimator, $\hat{\theta}^{PML}$, is consistent. This will turn out to be a key property needed to establish the validity of the bootstrap PLR tests we propose in section 4.

Theorem 2 *Under the conditions of Theorem 1,* $T^{1/2}(\hat{\beta} - \beta) \xrightarrow{p} 0$. *Moreover,* $\hat{\alpha} \xrightarrow{p} \alpha$, $\hat{\Psi} \xrightarrow{p} \Psi$, *and* $\hat{\Sigma} \xrightarrow{p} \Sigma$.

Remark 3.5. Theorem 2 shows that in the presence of conditional heteroskedasticity of the form specified in Assumption 2, the PML estimators of α, β, Σ and Ψ remain consistent. Under cases (ii) and (iii) of Remark 3.2 it can additionally be shown that $\hat{\mu}$, the PML estimator of μ, also remains consistent.

4 Bootstrap PLR Tests

In section 4.1 we first outline our wild bootstrap algorithm. Subsequently in section 4.2 we show that because, as was shown in the previous section, we can still consistently estimate α, β, μ and Ψ in the presence of conditional heteroskedasticity, (asymptotically) pivotal null p-values can be obtained using wild bootstrap re-sampling methods, regardless of whether conditional heteroskedasticity is present or not in the shocks. In section 4.3 we then demonstrate that the i.i.d. bootstrap rank tests of Swensen (2006) share the same large sample properties as the wild bootstrap.

The re-sampling algorithm discussed in section 4.1 draws on the wild bootstrap literature (see, *inter alia*, Wu, 1986; Liu, 1988; Mammen, 1993). In the context of the present problem, we focus our primary attention on the wild bootstrap scheme because, unlike the i.i.d. residual re-sampling schemes used for other bootstrap co-integration

[2] This procedure starts with $r = 0$ and sequentially raises r by one until for $r = \hat{r}$ the test statistic $Q_{\hat{r}}$ (or $Q_{\hat{r},\max}$) does not exceed the ξ level critical value for the test.

[3] The likelihood being used in (3.1) is not the correct likelihood for the model in (2.5), unless $\varepsilon_t \sim NIID(\mathbf{0}, \mathbf{\Sigma})$.

tests proposed in the literature; see, e.g., Swensen (2006) and, in the univariate ($p = 1$) case, Inoue and Kilian (2002), Paparoditis and Politis (2003), Park (2003), the wild bootstrap replicates the pattern of heteroskedasticity present in the original shocks, and, hence, preserves the temporal ordering in the conditional heteroskedasticity. The wild bootstrap might therefore be expected to deliver improved finite sample size properties relative to the standard and i.i.d. bootstrap rank tests in the presence of conditional heteroskedasticity. The simulation results presented in section 5 support this conjecture.

4.1 The Wild Bootstrap Algorithm

Let us start by considering the problem of testing the null hypothesis $H(r)$ against $H(p)$, $r < p$. Swensen (2006, section 2) discusses at length a way of implementing a bootstrap version of the well known trace test in this case. Here we extend his approach by modifying his re-sampling scheme in order to account the presence of conditional heteroskedasticity by means of the wild bootstrap. Implementation of the wild bootstrap requires us only to estimate the VAR(k) model under $H(p)$ (i.e., the unrestricted VAR) and under $H(r)$.

Let $\hat{\Psi} := (\hat{\Gamma}_1, ..., \hat{\Gamma}_{k-1})$ denote the (unrestricted) PML estimate of Ψ from the model under $H(p)$; the corresponding unrestricted residuals are denoted by $\hat{\varepsilon}_t$, $t = 1, ..., T$. In addition, let $\hat{\alpha}, \hat{\beta}$ denote the (restricted) PML estimates of α,β under the null hypothesis $H(r)$. The bootstrap algorithm we consider in this section requires that the roots of the equation $|\hat{A}^*(z)| = 0$ are either one or are outside the unit circle, where

$$\hat{A}^*(z) := (1 - z) I_p - \hat{\alpha}\hat{\beta}' z - \hat{\Gamma}_1 (1 - z) z - ... - \hat{\Gamma}_{k-1} (1 - z) z^{k-1} ;$$

moreover, we also require that $|\hat{\alpha}'_{\perp}\hat{\Gamma}\hat{\beta}_{\perp}| \neq 0$, ($\hat{\Gamma} := I_p - \hat{\Gamma}_1 - ... - \hat{\Gamma}_{k-1}$). While the latter condition is always satisfied in practice, if the former condition is not met, then the bootstrap algorithm cannot be implemented, because the bootstrap samples may become explosive; cf. Swensen (2006, Remark 1). However, in such cases any estimated root which has modulus greater than unity could be shrunk to have modulus strictly less than unity; cf. Burridge and Taylor (2001,p.73).

The following steps constitute our wild bootstrap algorithm, which coincides with Algorithm 1 of CRT. We outline the procedure for the trace statistic, $Q(r)$. The maximum eigenvalue statistic, $Q_{r,\max}$ for $H(r)$ vs $H(r + 1)$ can be bootstrapped in the same way, replacing Q_r^b with $Q_{r,\max}^b := -2\left(\ell^b(r) - \ell^b(r+1)\right)$ in Steps 3 and 4 of Algorithm 1.

Algorithm 1 (Wild Bootstrap Co-integration Test)

Step 1: *Generate T bootstrap residuals ε_t^b, $t = 1, ..., T$, according to the device*

$$\varepsilon_t^b := \hat{\varepsilon}_t w_t \tag{4.1}$$

where $\{w_t\}_{t=1}^T$ *denotes an independent* $N(0,1)$ *scalar sequence;*

Step 2: *Construct the bootstrap sample recursively from*

$$\Delta X_t^b := \hat{\alpha}\hat{\beta}' X_{t-1}^b + \hat{\Gamma}_1 \Delta X_{t-1}^b + ... + \hat{\Gamma}_{k-1} \Delta X_{t-k+1}^b + \varepsilon_t^b,\ t = 1, ..., T, \tag{4.2}$$

with initial values, $X_t^b := 0,\ t = -k+1, ..., 0.$

Step 3: *Using the bootstrap sample,* $\{X_t^b\}$*, obtain the bootstrap test statistic,* $Q_r^b := -2\left(\ell^b(r) - \ell^b(p)\right)$*, where* $\ell^b(r)$ *and* $\ell^b(p)$ *denote the bootstrap analogues of* $\ell(r)$ *and* $\ell(p)$*, respectively;*

Step 4: *Bootstrap p-values are then computed as,* $p_{r,T}^b := 1 - G_{r,T}^b(Q_r)$*, where* $G_{r,T}^b(\cdot)$ *denotes the conditional (on the original data) cumulative distribution function (cdf) of* Q_r^b*.*

Remark 4.1. The key feature of the wild bootstrap is Step 1, where the bootstrap shocks, ε_t^b, in (4.1) are generated by multiplying the residuals $\hat{\varepsilon}_t$ by a *scalar* IID(0,1) sequence. This allows the bootstrap shocks to replicate the pattern of heteroskedasticity present in the original shocks since, conditionally on $\hat{\varepsilon}_t$, ε_t^b is independent over time with zero mean and variance matrix $\hat{\varepsilon}_t\hat{\varepsilon}_t'$. Also, conditionally on the data, the bootstrap partial sum $T^{-1/2}\sum_{i=1}^{\lfloor Tu \rfloor} \varepsilon_t^b = T^{-1/2}\sum_{i=1}^{\lfloor Tu \rfloor} \hat{\varepsilon}_t w_t$ has mean zero, independent increments and variance $T^{-1}\sum_{t=1}^{\lfloor Tu \rfloor} \hat{\varepsilon}_t\hat{\varepsilon}_t' = u\Sigma + o_p(1)$, Σ the average conditional variance; cf. Remark 2.1. Finally notice that, due to the normality assumption on w_t, the bootstrap partial sum is (conditionally on the original data) exact Gaussian[4].

Remark 4.2. Observe that, due to the (exact) invariance of Q_r with respect to μ, we need not add an estimate of the estimated deterministic component, μD_t, to the right member of (4.2) as is done in, for example, Swensen (2006). Moreover, since Q_r is similar (exact similar under cases (ii) and (iii) of Remark 3.2 and asymptotically similar under case (i)) with respect to the initial values we may set these to zero in our recursive scheme. As an alternative to (4.2) one could use the recursion

$$\Delta X_t^b := \hat{\alpha}\hat{\beta}' X_{t-1}^b + \hat{\Gamma}_1 \Delta X_{t-1}^b + ... + \hat{\Gamma}_{k-1} \Delta X_{t-k+1}^b + \hat{\mu} D_t + \varepsilon_t^b,\ t = 1, ..., T$$

with initial values, $X_t^b := X_t,\ t = -k+1, ..., 0$. In the restricted trend case, $\hat{\mu} := (\hat{\mu}_1', \hat{\mu}_2')'$ with $\hat{\mu}_1$ and $\hat{\mu}_2$ the PML estimates of μ_1 and μ_2 obtained from the model estimated under $H(r)$ and $H(p)$, respectively, while in the restricted constant case $\hat{\mu} := \hat{\mu}_1$, with $\hat{\mu}_1$ the PML estimates of μ_1 obtained from the model estimated under $H(r)$; cf. Swensen (2006). In unreported Monte Carlo simulations we found no discernible

[4] As discussed in Remark 4.3 of CRT, we also investigated whether improved small sample accuracy could be obtained by replacing the Gaussian distribution used for generating the pseudo-data in (4.1) by an asymmetric distribution. Like CRT we found no discernible differences between the finite sample properties of the bootstrap rank tests based on the Gaussian distribution, Mammen's (1993) two-point distribution or the Rademacher distribution, also consistent with evidence reported in Table 5 of Gonçalves and Kilian (2004,p.105).

differences between the finite sample properties of these two approaches and so we have adopted the simpler of the two.

Remark 4.3. As detailed in Remark 4.4 of CRT, the unknown cdf, $G^b_{r,T}(\cdot)$, required in Step 4 of Algorithm 1 can approximated through numerical simulation. This is done by generating N (conditionally) independent bootstrap statistics, $Q^b_{n:r}$, $n = 1, ..., N$, and then computing the p-value as $\tilde{p}^b_{r,T} := N^{-1} \sum_{n=1}^{N} \mathbb{I}\left(Q^b_{n:r} > Q_r\right)$, and is such that $\tilde{p}^b_{r,T} \overset{a.s.}{\rightarrow} p^b_{r,T}$ as $N \rightarrow \infty$. For further discussion of the wild bootstrap procedure outlined in Algorithm 1 we refer the reader to the discussion given in Section 4.1 of CRT.

4.2 Asymptotic Theory for the Wild Bootstrap

The asymptotic validity of the wild bootstrap method outlined in Algorithm 1 under conditional heteroskedasticity is now established in Theorem 3. In order to keep our presentation simple, we demonstrate our result for the case of no deterministic variables. The equivalence of the first-order limiting null distributions of the Q^b_r and Q_r statistics can also be shown to hold for cases (ii) and (iii) of Remark 3.2. Again this is a straightforward extension of the results in Theorem 3 and is omitted in the interests of brevity.

Theorem 3 *Let the conditions of Theorem 1 hold. Then, under the null hypothesis* $H(r)$, $Q^b_r \overset{w}{\rightarrow}_p Q_{r,\infty}$. *Moreover,* $p^b_{r,T} \overset{w}{\rightarrow} U[0,1]$.

Remark 4.4. A comparison of the result for Q^b_r in Theorem 3 with that given for Q_r in Theorem 1 demonstrates the usefulness of the wild bootstrap: as the number of observations increases, the wild bootstrapped statistic has the same first-order null distribution as the original test statistic. Consequently, the bootstrap p-values are (asymptotically) uniformly distributed under the null hypothesis, leading to tests with (asymptotically) correct size in the presence of conditional heteroskedasticity of the form given in Assumption 2.

Remark 4.5. It can be shown that the sequential procedure of Johansen (1996), see footnote 1, employed using the wild bootstrap Q^b_r, $r = 0, ..., p-1$, test statistics is consistent in the sense that it correctly selects the true co-integrating rank with probability $(1 - \xi)$ in large samples (ξ denoting the nominal significance level used in each test in the procedure) in the presence of conditional heteroskedasticity satisfying Assumption 2. The same can also be shown to be the case for the corresponding sequential procedure based on the i.i.d. boostrap approach of Swensen (2006). See the accompanying working paper, Cavaliere *et al.* (2009), for further details on this together with Monte Carlo simulation evidence into the finite sample performance of the sequential procedures based on the asymptotic and bootstrap tests under a variety of conditionally heteroskedastic models.

Remark 4.6. Given the results in Theorem 3, it follows straightforwardly that the limiting null distribution of the bootstrap maximum eigenvalue statistic, $Q^b_{r,\max}$, coincides with that given in Remark 3.3, so that again our wild bootstrap procedure will deliver (asymptotically) correctly sized maximum eigenvalue co-integration tests under the conditions of Theorem 3. The results of Remark 4.5 also apply for the sequential procedure based on the bootstrap maximum eigenvalue statistic.

4.3 Swensen's i.i.d. Bootstrap

The i.i.d. bootstrap method outlined in Swensen (2006) follows the same steps as the wild bootstrap method outlined above in section 4.1, except that Step 1 of Algorithm 1 is replaced by the following:

Step 1: *Generate T bootstrap residuals ε^s_t, $t = 1, ..., T$, as independent draws with replacement from the centred residuals $\{\hat{\varepsilon}_t - T^{-1}\sum_{i=1}^T \hat{\varepsilon}_i\}_{t=1}^T$.*

The algorithm for the i.i.d. bootstrap rank tests then continues exactly as in Algorithm 1, but using the centred[5] i.i.d. bootstrap residuals, ε^s_t, in place of the wild bootstrap residuals, ε^b_t. We denote the resulting i.i.d. bootstrap rank statistic by Q^s_r and the associated i.i.d. bootstrap p-value as $p^s_{r,T}$. The same conditions on the roots of the equation $|\hat{A}^*(z)| = 0$ as were required for the wild bootstrap must also hold here, as must the condition that $|\hat{\alpha}'_\perp \hat{\Gamma} \hat{\beta}_\perp| \neq 0$. Again any estimated root with modulus greater than unity may again be shrunk to have modulus strictly less than unity.

Under the (homoskedastic) assumption that $\varepsilon_t \sim$ i.i.d.$(\mathbf{0}, \Sigma)$ with finite fourth moments, Swensen (2006) demonstrates that the i.i.d. bootstrap rank statistic Q^s_r replicates the first-order asymptotic null distribution of the standard trace statistic, Q_r of (3.3). In Theorem 4 we now establish that the i.i.d. bootstrap method of Swensen (2006) remains asymptotically valid under the weaker conditionally heteroskedastic conditions placed on the innovations in this paper. This result is demonstrated for the case of no deterministic variables. The equivalence of the first-order limiting null distributions of the Q^s_r and Q_r statistics under cases (ii) and (iii) of Remark 3.2 is again a straightforward extension of the results in Theorem 4.

Theorem 4 *Let the conditions of Theorem 1 hold. Then, under the null hypothesis $H(r)$, $Q^s_r \xrightarrow{w}_p Q_{r,\infty}$. Moreover, $p^s_{r,T} \xrightarrow{w} U[0,1]$.*

Remark 4.7. As discussed at the end of Section 4.1, the cdf of Q^s_r used in Step 4 of the bootstrap algorithm can again be approximated through numerical simulation. Moreover, an i.i.d. bootstrap analogue of the maximum eigenvalue statistic can also be obtained in an obvious way. Again it follows immediately from the results in Theorem 4 that this statistic has the same limiting null distribution as that given for $Q_{r,\max}$ in Remark 3.3.

[5] Notice that if the estimated unrestricted VAR contains a constant, then $T^{-1}\sum_{t=1}^T \hat{\varepsilon}_t = 0$ and, hence, the residuals would not need to be centred prior to re-sampling.

5 Finite Sample Simulations

In this section we use Monte Carlo simulation methods to compare the finite sample size and power properties of the PLR co-integration rank test of Johansen (1996) with its wild bootstrap version proposed in Section 4 together with the corresponding i.i.d. bootstrap test of Swensen (2006). The simulation model we consider generalises that used by previous authors in that we are allowing for conditional heteroskedasticity in the innovation process driving the VAR model.[6]

In sections 5.1, and 5.2 we follow Johansen (2002) and Swensen (2006), and consider as our simulation DGP an $I(1)$, possibly co-integrated, VAR(1) process of dimension p. We allow the dimension of the VAR process to vary over $p = 2, ..., 5$, and consider both the case of no co-integration ($r = 0$) [section 5.1], and of a single co-integrating vector ($r = 1$) [section 5.2]. In section 5.3 we will subsequently report results for $r = 0$ in a VAR(2) model, thereby also investigating the finite sample impact of higher-order serial correlation.

The DGP considered in section 5.1 is the multivariate martingale process,

$$\Delta X_t = \varepsilon_t, \quad t = 1, ..., T \tag{5.1}$$

initialised at $X_0 = \mathbf{0}$, while a generalisation of this DGP to the non-co-integrated VAR(2) case is detailed in section 5.3. In section 5.2, we report results for the co-integrated VAR(1) model

$$\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t, \quad t = 1, ..., T$$

where α and β are $p \times 1$ vectors; following Johansen (2002) and Swensen (2006), we consider the parameter combinations, $\beta := (1, 0, ..., 0)'$ and $\alpha := (a_1, a_2, 0, ..., 0)'$. This leads to the model

$$\begin{aligned} \Delta X_{1,t} &= a_1 X_{1,t-1} + \varepsilon_{1,t} \\ \Delta X_{2,t} &= a_2 X_{1,t-1} + \varepsilon_{2,t} \\ \Delta X_{i,t} &= \varepsilon_{i,t}, \qquad i = 3, ..., p. \end{aligned} \tag{5.2}$$

In our reported simulations we set $a_1 = a_2 = -0.4$, as in Swensen (2006, Table 2). The initial value of the stationary component in (5.2), $X_{1,0}$, is drawn from the corresponding invariant distribution, while the remaining components are initialised at zero.

[6]Complementary results, comparing the properties of the sequential approach of Johansen (1996) when applied using the asymptotic PLR test and the two bootstrap analogue methods are reported in the accompanying working paper, Cavaliere *et al.* (2009). These results show that a sequential method based on wild bootstrap PLR tests works well under conditional heteroskedasticity when $r = 0$, avoiding a strong tendency to over-estimate the co-integrating rank displayed by the analogous procedures based on the i.i.d. bootstrap PLR and asymptotic PLR tests. The simulations in Cavaliere *et al.* (2009) also highlight the encouraging result that the wild bootstrap Q_0^b test does not lose power against $r = 1$ relative to the other tests, despite, as will be shown in what follows, displaying far superior size properties than the other tests in the presence of conditional heteroskedasticity; cf. Tables 1 and 2.

In both (5.1) and (5.2), $\varepsilon_t := (\varepsilon_{1,t}, ..., \varepsilon_{p,t})'$ is a p-dimensional martingale difference sequence with respect to the filtration $\mathcal{F}_t := \sigma(\varepsilon_t, \varepsilon_{t-1}, ...)$. Following van der Weide (2002), we assume that ε_t may be written as the linear map

$$\varepsilon_t = \Lambda e_t \tag{5.3}$$

where Λ is an invertible $p \times p$ matrix which is constant over time, while the p components of $e_t := (e_{1,t}, ..., e_{p,t})'$ are independent across $i = 1, ..., p$. In the case where the individual components follow a standard $GARCH(1,1)$ process (as is the case with Models A and B below), van der Weide (2002) refers to ε_t as a $GO\text{-}GARCH(1,1)$ process. Notice that, by definition, the PLR statistic does not depend on the matrix Λ, as the eigenvalue problem in (3.2) has the same eigenvalues upon re-scaling (as can be seen by simply pre- and post-multiplying by Λ^{-1} in (3.2)). This allows us to set $\Lambda = I_p$ in the simulations, without loss of generality.

Within the context of (5.3) we consider for the individual components of e_t the univariate innovation processes and parameter configurations used in Section 4 of Gonçalves and Kilian (2004), to which the reader is referred for further discussion. These are as follows:

- **Model A** is a standard $GARCH(1,1)$ process driven by standard normal innovations of the form $e_{it} = h_{it}^{1/2} v_{it}$, $i = 1, ..., p$, where v_{it} is i.i.d. $N(0,1)$, independent across i, and $h_{it} = \omega + d_0 e_{it-1}^2 + d_1 h_{it-1}$, $t = 0, ..., T$. Results are reported for $(d_0, d_1) \in \{(0.5, 0.0), (0.3, 0.65), (0.2, 0.79), (0.05, 0.94)\}$.

- **Model B** is the same as Model A except that the v_{it}, $i = 1, ..., p$, are independent i.i.d. t_5 (normalised to unit variance) variates.

- **Model C** is the exponential $GARCH(1,1)$ ($EGARCH(1,1)$) model of Nelson (1991) with $e_{it} = h_{it}^{1/2} v_{it}$, $\ln(h_{it}) = -0.23 + 0.9 \ln(h_{it-1}) + 0.25[|v_{it-1}^2| - 0.3 v_{it-1}]$, with $v_{it} \sim$ i.i.d. $N(0,1)$, independent across $i = 1, ..., p$.

- **Model D** is the asymmetric $GARCH(1,1)$ ($AGARCH(1,1)$) model of Engle (1990) with $e_{it} = h_{it}^{1/2} v_{it}$, $h_{it} = 0.0216 + 0.6896 h_{it-1} + 0.3174[e_{it-1} - 0.1108]^2$, with $v_{it} \sim$ i.i.d. $N(0,1)$, independent across $i = 1, ..., p$.

- **Model E** is the $GJR - GARCH(1,1)$ model of Glosten *et al.* (1993) with $e_{it} = h_{it}^{1/2} v_{it}$, $h_{it} = 0.005 + 0.7 h_{it-1} + 0.28[|e_{it-1}| - 0.23 e_{it-1}]^2$, with $v_{it} \sim$ i.i.d. $N(0,1)$, independent across $i = 1, ..., p$.

- **Model F** is the first-order AR stochastic volatility model: $e_{it} = v_{it} \exp(h_{it})$, $h_{it} = \lambda h_{it-1} + 0.5 \xi_{it}$, with $(\xi_{it}, v_{it}) \sim$ i.i.d. $N(0, \text{diag}(\sigma_\xi^2, 1))$, independent across $i = 1, ..., p$. Results are reported for $(\lambda, \sigma_\xi) = \{(0.936, 0.424), (0.951, 0.314)\}$.

The reported simulations were programmed using the **rndKMn** function of Gauss 7.0. All experiments were conducted using $10,000$ replications. The sample sizes were chosen within the set $\{50, 100, 200\}$ and the number of replications used in the wild bootstrap algorithm was set to 399. All tests were conducted at the nominal 0.05 significance level. For the reasons outlined on page 12 of RHD, relating to similarity with respect to initial values (see also Nielsen and Rahbek, 2000), the VAR model was fitted with a restricted constant (i.e. deterministic case (ii) of Remark 3.2), when calculating all of the tests. For the standard PLR tests we employed asymptotic critical values as reported in Table 15.2 of Johansen (1996).

We have shown that the standard PLR Q_r test of Johansen (1996), together with the wild bootstrap Q_r^b test outlined in section 4.1 and the i.i.d. bootstrap Q_r^s test of Swensen (2006) are all asymptotically valid under conditional heteroskedastiticy of the form given in Assumption 2. However, and unlike the wild bootstrap re-sampled data in (4.1), the i.i.d. re-sampled data will clearly not preserve the temporal ordering in the conditional heteroskedasticity present in the original data. We would therefore expect its finite sample performance to be quite similar to that of the asymptotic tests and to not perform as well as the wild bootstrap tests in the presence of conditional heteroskedasticity.

5.1 The Non-Co-Integrated Model ($r = 0$)

Table 1 reports the finite sample (empirical) size properties of both the standard PLR test, Q_0, and its wild and i.i.d. bootstrap analogue tests, Q_0^b and Q_0^s respectively, for $H(0) : r = 0$ against $H(p) : r = p$, for $p = 2, ..., 5$, in the presence of conditional heteroskedasticity of the types outlined above.

Table 1 about here

Under constant conditional variances (the cases where $d_0 = d_1 = 0$ in Models A and B) it can be seen from the first two panels of Table 1 that both the Q_0^b and Q_0^s tests display finite sample sizes which are closer to the nominal level than the standard Q_0 test based on asymptotic critical values (the wild bootstrap can, however, be a little undersized); for example, in the case of Model A for $p = 5$, while the standard PLR test has size of 8.1% for $T = 100$, the corresponding wild and i.i.d. bootstrap tests have size of 4.4% and 4.7% respectively.

It is, however, where the innovation process displays conditional heteroskedasticity that the benefits of the wild bootstrap over the other tests become clear. The results in Table 1 show that both the Q_0 and Q_0^s tests can display quite unreliable size properties, even for samples as large as $T = 200$, in the presence of conditional heteroskedasticity. In contrast, the size properties of our wild bootstrap PLR test, Q_0^b, seem largely satisfactory throughout.

The size distortions seen in the Q_0 and Q_0^s tests are generally worse, other things being equal, the higher is the VAR dimension, p. For example, in the case of Model A with $d_0 = 0.3$, $d_1 = 0.65$ and $T = 200$, the Q_0 and Q_0^s have size of 10% and 9.3%, respectively, for $p = 2$ rising to 13.9% and 10.9%, respectively, for $p = 5$. In contrast, here the Q_0^b test has size of 5.6% and 5.7% for $p = 2$ and $p = 5$, respectively. The precise model of conditional heteroskedasticity can also make quite a substantial difference to the size properties of the tests. For example, comparing the results for Models A and B, we see that t_5 innovations tend to cause rather less size inflation than is seen for standard normal innovations. Of all the models considered, it is the autoregressive stochastic volatility case, Model F, which has the strongest impact on the size of the tests. The two parameter configurations both imply relatively strong serial dependence in the conditional variance of the innovation process (although in both cases the process does formally satisfy Assumption 2). Here the standard PLR test, Q_0, displays size of between around 20% to 40% depending on p and the parameter configuration, while the i.i.d. bootstrap test, Q_0^s, performs only slightly better. Although the wild bootstrap test, Q_0^b, does also show a degree of over-size under Model F, it still represents an enormous improvement on the size properties of the other tests. Moreover, what size distortions there are in the wild bootstrap tests are ameliorated, other things equal, as the sample size is increased. Notice that this last observation is not the case for the Q_0 and Q_0^s tests where the size distortions *increase* as the sample size increases. Very significant over-sizing, although not as bad as for Model F, is also seen for the Q_0 and Q_0^s tests in each of Models C, D and E. Again here the wild bootstrap test is much better behaved throughout.

5.2 The Co-Integrated Model ($r = 1$)

Consider next the results in Table 2 for the empirical sizes of the standard PLR Q_1 test and its i.i.d. and wild bootstrap analogues. The results here are very much in line with those seen in Table 1 with the standard PLR and its i.i.d. bootstrap analogue test not displaying anything like adequate size control in the presence of conditional heteroskedasticity. The observed size distortions again worsen, others things being equal, as p is increased. Again the worst distortions are seen in these tests under Model F, with serious over-size problems also seen under Models C, D and E. For the $GO - GARCH(1,1)$ case (Models A and B) the observed size distortions are again generally smaller under t_5 innovations than $N(0,1)$ innovations. In contrast to the standard and i.i.d. bootstrap PLR tests, the wild bootstrap PLR test displays very good size control throughout, with size only ever exceeding 7% in the case of Model F, where although still a little over-sized it does, nonetheless, still represent a massive improvement over the other tests.

Table 2 about here

5.3 The Non-Co-Integrated VAR(2) Model

To conclude this section, and following Johansen (2002, p.1940), we report some additional results investigating the finite sample behaviour under the null hypothesis of tests for $\Pi = 0$ in the VAR(2) model:

$$\Delta X_t = \Pi X_{t-1} + \Gamma_1 \Delta X_{t-1} + \varepsilon_t, \quad t = 1, ..., T \tag{5.4}$$

where $\Gamma_1 = \xi I_p$ with $-1 < \xi < 1$. This model is an interesting extension of the conditionally heteroskedastic VAR(1) model considered in sections 5.1 and 5.2 because it allows for higher-order serial correlation. As regards the innovation term, ε_t, we again considered each of Models A-F, reporting results for a subset of the parameter configurations reported for Models A, B and F in sections 5.1 and 5.2.[7] A restricted constant was again included in the estimated model. Finally, we initialized the process by drawing ΔX_0 from its corresponding invariant distribution.

In Table 3 we consider first the case where $\xi = 0.5$, which allows for a moderate degree of higher-order stationary serial correlation in the process. Results are reported for both the standard PLR test and its wild and i.i.d. bootstrap analogue tests for $H(0) : r = 0$ against $H(p) : r = p$, for $p = 2, ..., 5$.

Table 3 about here

In general, it can be seen from the results in Table 3 that higher-order stationary serial correlation tends to inflate the finite sample size of the standard PLR test, Q_0, further above its nominal level, relative to the corresponding results for the VAR(1) case in Table 1. This is true in both the conditionally homoskedastic and conditionally heteroskedastic cases. Both bootstrap tests also display a degree of finite sample over-size. However, the size distortions seen in the bootstrap tests are much smaller than those observed for the Q_0 test, and in general the wild bootstrap Q_0^b test displays smaller size distortions than the i.i.d. bootstrap Q_0^s test. To illustrate, for $p = 4$ in the i.i.d. innovations case (Model A with $d_0 = d_1 = 0$) for a sample of size $T = 50$ the Q_0, Q_0^b and Q_0^s tests have size of 41.5%, 8.5% and 8.9%, respectively, as compared to 8.7%, 4.3% and 4.8%, respectively, for the VAR(1) case in Table 1. For all three tests the observed over-sizing is smaller for $T = 200$ than for $T = 50$. Indeed, both bootstrap tests display size close to the nominal level when $T = 200$. As a second example, under Model C for $p = 5$, the Q_0, Q_0^b and Q_0^s tests have size of 73%, 12.5% and 15.6%, respectively, for $T = 50$ (22.5 %, 6.5% and 10.2%, respectively, for $T = 200$) compared with 20%, 7.1% and 10.1%, respectively, (14.7%, 5.7% and 11.2%, respectively) in the corresponding VAR(1) model.

Table 4 about here

The condition that $-1 < \xi < 1$ ensures that the process X_t is $I(1)$ and so does not violate Assumption 1. However, as ξ tends towards 1, so X_t will increasingly resemble an $I(2)$ process for a given (finite) sample size and, as a consequence, the rejection probability of the asymptotic test will tend towards to unity, rendering the asymptotic $I(1)$ critical values inappropriate.[8] We can therefore investigate the impact on the

[7] This was done in the interests of space, the additional results qualitatively adding very little to what is reported.

[8] In this case, approaches based on a Bartlett correction tend to also perform badly, primarily because the correction factor over-corrects, leading to tests with size close to zero. See e.g. Johansen (2002,p.1941).

behaviour of the bootstrap tests in this 'near $I(2)$' case by considering the effects of $\xi = 0.9$. These results are reported in Table 4. As would be expected, stationary roots close to unity cause further oversizing in the standard asymptotic test Q_0, relative to both the corresponding results for the VAR(1) case in Table 1 and to those for the moderate serially correlated case of Table 3. Although both bootstrap tests are also significantly oversized, they offer a major improvement over the asymptotic test. Again, the wild bootstrap Q_0^b test behaves better than the i.i.d. bootstrap Q_0^s test. To illustrate, consider the case of i.i.d. innovations for $p = 4$. For $T = 50$, the Q_0, Q_0^b and Q_0^s tests have size of 93.6%, 31.7% and 35.4%, as compared to 8.7%, 4.3% and 4.8%, respectively, for the VAR(1) case in Table 1, and to 73%, 12.5% and 15.6%, respectively, for the moderate serial correlation case in Table 3. The degree of over-size ameliorates when the sample size is increased to $T = 200$, again as would be expected. However, while the size of the Q_0 test for $\xi = 0.9$ remains worryingly high in this case at 46.4%, both bootstrap tests display reasonably decent sizes with that of Q_0^b the better of the two at 11.2%. While undoubtedly not perfect, this is certainly a huge improvement on the size of the standard Q_0 test. As for case of Model C with $p = 5$, see the discussion above relating to Table 3, for $T = 200$ the size of the Q_0 test is 75.6%, while that of Q_0^b is 17.7%.

Overall, both bootstrap tests deal much better with higher-order serial correlation than does the standard Q_0 test. Moreover, and as with the results in Table 1 for the VAR(1) case, in the VAR(2) case the results in Tables 3 and 4 show that the wild bootstrap Q_0^b test again displays far more robust finite sample size properties than either the Q_0 or the Q_0^s test in the presence of conditional heteroskedasticity. Indeed, for the VAR(2) case with $\xi = 0.9$ reported in Table 4, the Q_0^b test displays almost no variation in its size properties across the different models of volatility reported, other things held equal.

6 Empirical application

In this section we illustrate the methods discussed in this paper with a short application to the term structure of interest rates; see Campbell and Shiller (1987) for an early reference. According to traditional theory, aside from a constant or stationary risk premium, long-term interest rates are an average of current and expected future short term rates over the life of the investment. Hence, provided interest rates are well described as $I(1)$ variables, bond rates at different maturities should be driven by a single common stochastic trend, with the spreads between rates at different maturities being stationary. Although early studies tend to corroborate this view, see, for example, Hall *et al.* (1992), more recent research, based on broader sets of maturities, suggests that yields are better characterised by *more* than one common trend, reflecting possible non-stationarities in the risk premia and additional risk factors, such as the slope and curvature of the yield curve; see, e.g., Diebold, Ji and Li (2007) and Giese (2006).

We consider monthly interest rate data from the United States, Canada, the United

Kingdom, and Japan, taken from the OECD/MEI database. For each country a single long-run interest rate, L_t, and a variety of short-run rates, S_{it}, were used in the co-integration analysis. Specifically, these were as follows. United States (1978:1–2002:12): L_t = government composite bond yield (> 10 years); S_{1t} = federal funds rate; S_{2t} = prime rate; S_{3t} = rate on certificates of deposit; S_{4t} = US dollar in London, 3-month deposit rate. Canada (1982:6–2002:12): L_t = benchmark bond yield (10 years); S_{1t} = official discount rate; S_{2t} = overnight money market rate; S_{3t} = rate on 90-day deposits. United Kingdom (1978:1–2002:12): L_t = yield on 10-year government bonds; S_{1t} = London clearing banks rate; S_{2t} = overnight interbank rate; S_{3t} = rate on 3-month interbank loans. Japan (1989:1–2002:12): L_t = yield on interest bearing government bonds (10 years); S_{1t} = official discount rate; S_{2t} = un-collateralized overnight rate; S_{3t} = rate on 90-day certificates of deposit.

For each country let $X_t := (L_t, S_{1t}, ..., S_{p-1,t})'$, where $p = 4$ for all but the U.S. where $p = 5$. As is standard, we fit a VAR model for X_t with restricted intercept; that is, $D_{2t} = 0$ and $D_{1t} = 1$ in (2.5). The VAR was estimated using Gaussian maximum likelihood under the assumption of constant volatility; cf. Section 2. For each country the number of lags, k, was estimated using the BIC: for the U.K., Japan and the U.S. $k = 2$ was chosen, while for Canada $k = 1$ obtained. For each country the residuals from the fitted VAR(k) model were subjected to both single-equation and vector diagnostic tests against non-normality, GARCH(1,1), and general heteroskedasticity (using White's test both with and without cross-variable terms).[9] In the case of the U.K. and the U.S. all of the single-equation and vector tests rejected at the 1% level. For Canada this was also the case, except that two of the single equation GARCH(1,1) were not significant. For Japan, all of the vector tests rejected at the 1% level, as did all of the single-equation normality tests. However, none of the GARCH(1,1) tests were significant, while White's single-equation tests delivered three (two) out of four significant outcomes at the 1% level when cross-variable terms were (were not) included. In summary, the interest rate data for all of the countries considered display (to varying degrees) statistically significant evidence of heteroskedasticity.

Table 4 about here

Table 4 reports the results of the standard, wild and i.i.d. bootstrap co-integration rank tests for each country. For the standard tests (asymptotic) p-values were computed as suggested in MacKinnon, Haug and Michelis (1999). For both of the bootstrap methods the number of bootstrap replications was set to 399.

For each country, the standard sequential procedure detects two co-integrating relations at any conventional significance level, with a third co-integration relation being significant at the 10% level (with a p-value of 0.08) in the case of the U.S. data. The same conclusions are drawn using the corresponding procedure based on the i.i.d. bootstrap tests of Swensen (2006), except that the third co-integrating vector in the case of the U.S. is deemed insignificant at the 10% level (with a p-value of 0.12). In line

[9] The complete set of diagnostic test results can be obtained from the authors on request.

with what would be expected from the Monte Carlo simulation results in section 5 for series displaying a significant degree of heteroskedasticity, the wild bootstrap-based procedure consistently delivers a higher p-value for a given hypothesised co-integrating rank. For both the U.K. and Canada this does not lead us to a different conclusion on the co-integrating rank (of two) as was drawn from the standard and i.i.d. bootstrap tests. However, for both Japan and the U.S. only one co-integrating vector is uncovered by the wild bootstrap procedure, implying the presence of four common trends in the five-dimensional U.S. system, and three common trends in the four-dimensional Japanese system.

These results all therefore contradict the traditional view of the expectation hypothesis of the term structure, suggesting the presence of additional risk factors, since the hypothesis of $p-1$ stationary relations (p being the number of interest rates considered) is never accepted, thereby providing further support in favour of recent multi-factor theories of the term structure; see, for example, Diebold, Ji and Li (2007). It is worth noting, however, that in the case of the U.S. data the p-value for testing $p-2$ against $p-1$ co-integrating relations is 12% using the asymptotic test and 15% using the i.i.d. bootstrap test. For the wild bootstrap this p-value rises sharply to 62%. The case of the U.S. data shows the biggest differences between the wild bootstrap procedure and those based on either the asymptotic test or the i.i.d. bootstrap tests of Swensen (2006). Given the significant heteroskedasticity found in the U.S. data (indeed the outcomes of the diagnostic test statistics were consistently much larger for the U.S. than for the other countries considered) the inferences from the wild bootstrap-based procedure would appear to be the most reliable.

7 Conclusions

In this paper we have demonstrated that the conventional co-integration rank tests of Johansen (1996) retain their usual limiting null distributions in the case where the innovations follow a globally stationary, conditionally heteroskedastic (martingale difference) process. We have also proposed wild bootstrap-based implementations of the co-integration rank tests in order to exploit the information in sample on the conditional heteroskedasticity, where present. As with any bootstrap procedure, no tables of critical values are required as the procedure automatically delivers a p-value for the hypothesis being tested. Both our proposed wild bootstrap scheme and the i.i.d. bootstrap scheme of Swensen (2006) were demonstrated to deliver rank statistics which share the same first-order limiting null distributions as the corresponding standard rank statistic. Monte Carlo evidence presented suggests that the proposed wild bootstrap co-integrating rank tests perform very well in finite samples, being considerably more robust than both the standard PLR tests based on asymptotic critical values and i.i.d. residual-based bootstrap analogues of the PLR tests, when the innovations are conditionally heteroskedasticity. That the i.i.d. residual-based bootstrap test is significantly better sized than the asymptotic test is in line with results established in the stationary linear regression case which show that the i.i.d. bootstrap can lead to asymptotic refinements even if the heteroskedasticity affecting the original data is not replicated into the bootstrap shocks (see Liu, 1988).

We conclude with a suggestion for further research. The analysis in this paper has been conducted under the assumption that the vector of time series under investigation are each either I(0) or I(1). This rules out the possibility of near-integration amongst the series, as is considered in, *inter alia*, Elliott (1998) and Pesavento (2004). An analysis of the bootstrap (P)LR co-integration rank tests under near-integration is beyond the scope of the present paper but would constitute an important and worthwhile extension of the results presented here. We note that the co-integrating rank selection procedure outlined in Cheng and Phillips (2008,2009) does consistently estimate the co-integrating rank under near-integration.

Acknowledgements

We are grateful to Peter Phillips, Søren Johansen, Anders Swensen and Carsten Trenkler for their helpful and constructive comments on earlier versions of this paper. We also thank Steve Leybourne for providing us with the data used in section 6. Parts of this paper were written while Cavaliere and Taylor both visited CREATES, whose hospitality is gratefully acknowledged. CREATES, Center for Research in Econometric Analysis of Time Series, is funded by The Danish National Research Foundation. Cavaliere and Rahbek thank the Danish Social Sciences Research Council, Project Number 2114-04-001, for continuing financial support. Cavaliere also acknowledges financial support from MIUR PRIN 2007 grants.

A Appendix

This section contains the proofs of the main theorems given in the paper. Proofs for Theorems 1 and 2 are collected in section A.1. The proof of the validity of the wild bootstrap co-integration test is reported in section A.2, while the corresponding result for the i.i.d. bootstrap test of Swensen (2006) is detailed in section A.3.

A.1 Proof of Theorems 1 and 2

Under the stated assumptions, the process X_t has the representation below in Lemma A.1 which is essential for the proofs of Lemmas A.2 and A.3. Lemma A.1 generalises the usual Granger-type representation in Johansen (1996) in that, rather than being *i.i.d.*, the ε_t sequence is now, by assumption, a (possibly non-stationary) MDS.

Lemmas A.2 and A.3 immediately imply that the proofs of Theorem 11.1 and Lemma 13.1 in Johansen (1996) hold, establishing Theorem 1 and 2 respectively. □

Lemma A.1 *Under the conditions of Theorem 1,*

$$X_t = C\sum_{i=1}^{t}\varepsilon_i + S_t + C_0. \tag{A.1}$$

Here the $(p\times p)$ *−dimensional matrices* $C := \beta_\perp\left(\alpha'_\perp\Gamma\beta_\perp\right)^{-1}\alpha'_\perp$ *and* $C_0 := C(I_p, -\Psi)\mathbb{X}_0$. *Define the* $(r+p(k-1))$*-dimensional autoregressive process* $\mathbb{X}_{\beta t}$ *where* $\mathbb{X}_{\beta t} := \beta' X_t$ *for* $k = 1$, *and otherwise,* $\mathbb{X}_{\beta t} := \left(X'_t\beta, \Delta X'_t, ..., \Delta X'_{t-k+1}\right)'$. *Then the p-dimensional process* $S_t := (\alpha, \Psi)\, Q\mathbb{X}_{\beta t}$, *where* $\mathbb{X}_{\beta t}$ *has the MA(∞) representation,* $\mathbb{X}_{\beta t} = \Phi(L)\,\eta_t = \sum_{i=0}^{\infty}\Phi^i\eta_{t-i}$. *Here* $\eta_t := (\beta, I_p, 0, ..., 0)'\,\varepsilon_t$ *and the spectral radius of* Φ *is smaller than one;* $\rho(\Phi) < 1$. *The* $(r+p(k-1))\times(r+p(k-1))$ *dimensional matrix* Q *is non-singular.*

Proof: With $\mathbb{X}_t := \left(X'_t, ..., X'_{t-k+1}\right)'$ the system can be written in companion form as,

$$\Delta\mathbb{X}_t = \mathbb{A}\mathbb{B}'\mathbb{X}_{t-1} + e_t \tag{A.2}$$

with $e_t := (\varepsilon'_t, 0, ..., 0)'$, $\mathbb{X}_0$ fixed and

$$\mathbb{A} := \begin{pmatrix} \alpha & \Gamma_1 & \Gamma_2 & ... & \Gamma_{k-1} \\ 0 & I_p & 0 & ... & 0 \\ 0 & 0 & I_p & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & I_p \end{pmatrix} \quad \mathbb{B} := \begin{pmatrix} \beta & I_p & 0 & ... & 0 \\ 0 & -I_p & I_p & ... & 0 \\ 0 & 0 & -I_p & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & -I_p \end{pmatrix}. \tag{A.3}$$

Note that with $\mathbb{X}_{\beta t} := \mathbb{B}'\mathbb{X}_t$, $\Phi := \left(I_{r+p(k-1)} + \mathbb{B}'\mathbb{A}\right)$, then $\mathbb{X}_{\beta t} = \Phi\mathbb{X}_{\beta t-1} + \mathbb{B}'e_t$. By Assumption 1, $\rho(\Phi) < 1$ and $\mathbb{X}_{\beta t}$ has the stated MA(∞) representation. Standard arguments and recursions give,

$$\mathbb{X}_t = \mathbb{C}\sum_{i=1}^{t} e_i + \mathbb{S}_t + \mathbb{C}\mathbb{X}_0 \tag{A.4}$$

where $\mathbb{C} := \mathbb{B}_\perp(\mathbb{A}'_\perp\mathbb{B}_\perp)^{-1}\mathbb{A}'_\perp$, and $\mathbb{S}_t := \mathbb{A}\left(\mathbb{B}'\mathbb{A}\right)^{-1}\mathbb{X}_{\beta t}$. As $X_t = (I_p, 0, ..., 0)\,\mathbb{X}_t$, the results in Lemma A.1 hold with $S_t = (I_p, 0, ..., 0)\,\mathbb{S}_t = (\alpha, \Psi)\, Q\mathbb{X}_{\beta t}$, $Q := (\mathbb{B}'\mathbb{A})^{-1}$. Noting that,

$$\mathbb{A}_\perp = (I_p, -\Gamma_1, ..., -\Gamma_{k-1})'\,\alpha_\perp, \quad \mathbb{B}_\perp = (I_p, ..., I_p)'\,\beta_\perp$$

the various expressions follow by simple algebraic identities. □

Let $\Omega_{\beta\beta} := \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T}\beta' Z_{1t}Z'_{1t}\beta$, $\Omega_{\beta i} := \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T}\beta' Z_{1t}Z'_{it}$ for $i = 0, 2$, and $\Omega_{ij} := \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} Z_{it}Z'_{jt}$, $i, j = 0, 2$. By Lemma A.1, these are well-defined as infinite sums in terms of exponentially decaying coefficients. E.g., since $\rho(\Phi) < 1$,

$$\Omega_{\beta 0} = \beta'(\alpha, \Psi)\, Q\sum_{i=0}^{\infty}\left[\Phi^i(\beta, I_p, 0, ..., 0)'\,\Sigma\,(\beta, I_p, 0, ..., 0)\,\Phi^{i\prime}\right](\alpha, \Psi)'.$$

In terms of these moment matrices we have the following results.

Lemma A.2 *Under the conditions of Theorem 1, and as* $T \to \infty$*,*

$$S_{00} \xrightarrow{p} \Sigma_{00},\ \beta' S_{10} \xrightarrow{p} \Sigma_{\beta 0} \text{ and } \beta' S_{11}\beta \xrightarrow{p} \Sigma_{\beta\beta} \tag{A.5}$$

where $\Sigma_{ij} = \Omega_{ij} - \Omega_{i2}\Omega_{22}^{-1}\Omega_{2j}$*,* $i, j = 0, 1, \beta$*. Moreover, the following identities hold,*

$$\Sigma_{00} = \alpha\Sigma_{\beta 0} + \Sigma,\ \Sigma_{0\beta} = \alpha\Sigma_{\beta\beta} \tag{A.6}$$

and

$$\Sigma_{00}^{-1} - \Sigma_{00}^{-1}\alpha(\alpha'\Sigma_{00}^{-1}\alpha)^{-1}\alpha'\Sigma_{00}^{-1} = \alpha_\perp(\alpha_\perp'\Sigma\alpha_\perp)^{-1}\alpha_\perp'. \tag{A.7}$$

PROOF: Consider $\beta' S_{10} = \beta' M_{10} - \beta' M_{12} M_{22}^{-1} M_{20}$. Using Lemma A.1 and the fact that, by definition,

$$\Delta X_t = \alpha\beta' X_{t-1} + \Psi U_t + \varepsilon_t = (\alpha, \Psi)\, \mathbb{X}_{\beta t-1} + \varepsilon_t, \tag{A.8}$$

the first term equals,

$$\beta' M_{10} = \frac{1}{T}\sum_{t=1}^{T} \beta' X_{t-1}\Delta X_t' = \frac{1}{T}\sum_{t=1}^{T} \beta' X_{t-1}\left((\alpha, \Psi)\, \mathbb{X}_{\beta t-1} + \varepsilon_t\right)'.$$

As mentioned in section 2, the strong law of large numbers in Hannan and Heyde (1972) can be applied by Assumption 2 and the fact that the coefficients Φ^i in the representation for $\mathbb{X}_{\beta t}$ in are exponentially decreasing by Lemma A.1. We then obtain directly that:

$$\beta' M_{10} \xrightarrow{p} \Omega_{\beta 0} := \beta'(\alpha, \Psi)\, Q \sum_{i=0}^{\infty} \left[\Phi^i (\beta, I_p, 0, ..., 0)' \,\Sigma\, (\beta, I_p, 0, ..., 0)\, \Phi^{i\prime}\right] (\alpha, \Psi)'.$$

Likewise, the terms $\beta' M_{12}$, M_{22} and M_{20} converge in probability and we conclude that

$$\beta' S_{10} \xrightarrow{p} \Sigma_{\beta 0} := \Omega_{\beta 0} - \Omega_{\beta 2}\Omega_{22}^{-1}\Omega_{20}\ .$$

Identical arguments lead to the other results in (A.5).

The identities in (A.6) follow by post-multiplying (A.8) by (the transpose of) $\beta' X_{t-1}, \Delta X_t$ and U_t respectively, taking averages and applying the law of large numbers as above, and solving the resulting system of equations. To prove the identity in (A.7) use the projection identity

$$I_p = \Sigma_{00}^{-1}\alpha(\alpha'\Sigma_{00}^{-1}\alpha)^{-1}\alpha' + \alpha_\perp(\alpha_\perp'\Sigma_{00}\alpha_\perp)^{-1}\alpha_\perp'\Sigma_{00}$$

and $\alpha_\perp'\Sigma_{00} = \alpha_\perp'\Sigma$; see (A.6). □

Lemma A.3 *Define the $(p-r)$-dimensional process,*

$$G(u) := \beta_\perp' CW(u), \tag{A.9}$$

where $W(\cdot)$ is a p-dimensional Browian motion with covariance Σ. Then under the conditions of Theorem 1, as $T \to \infty$,

$$\frac{1}{\sqrt{T}} \beta_\perp' X_{\lfloor Tu \rfloor} \xrightarrow{w} G(u) \tag{A.10}$$

$$\beta_\perp' S_{10} \alpha_\perp = \beta_\perp' S_{12} \alpha_\perp \xrightarrow{w} \int_0^1 G(s)\, dW(s)' \alpha_\perp \tag{A.11}$$

$$\frac{1}{T} \beta_\perp' S_{11} \beta_\perp \xrightarrow{w} \int_0^1 G(s) G(s)' ds \tag{A.12}$$

and furthermore,

$$\sqrt{T} \beta' S_{10} \alpha_\perp = \sqrt{T} \beta' S_{1\varepsilon} \alpha_\perp \xrightarrow{w} N_{r \times p-r}(0, \Sigma_{\beta\beta} \otimes \alpha_\perp' \Sigma \alpha_\perp) \tag{A.13}$$

$$\beta' S_{11} \beta_\perp \in O_p(1). \tag{A.14}$$

Proof: The result in (A.10) holds by using the FCLT in Brown (1971) (see the discussion in Remark 2.2) applied to ε_t as Lemma A.1 implies directly that $\beta_\perp' X_{\lfloor T \cdot \rfloor} = \beta_\perp' C \sum_1^{\lfloor T \cdot \rfloor} \varepsilon_t + o_p(\sqrt{T})$. To prove (A.11) note that

$$\beta_\perp' S_{1\varepsilon} = \beta_\perp' M_{1\varepsilon} - \beta_\perp' M_{12} M_{22}^{-1} M_{2\varepsilon}$$

where $M_{1\varepsilon} := T^{-1} \sum_{t=1}^T \Delta X_t \varepsilon_t'$. Consider first $\beta_\perp' M_{1\varepsilon}$ and use the representation of X_t given in (A.1) to see that

$$\beta_\perp' M_{1\varepsilon} = \frac{1}{T} \left(\beta_\perp' C \sum\nolimits_{t=1}^T (\sum\nolimits_{i=1}^{t-1} \varepsilon_i) \varepsilon_t' + \beta_\perp' \sum\nolimits_{t=1}^T S_{t-1} \varepsilon_t' + \beta_\perp' C_0 \sum\nolimits_{t=1}^T \varepsilon_t' \right)$$

which by Hansen (1992), the LLN and the fact that ε_t and ε_{t-1} are uncorrelated, weakly converges to $\int_0^1 G(s)\, dW(s)'$. Next, $M_{\varepsilon 2} := T^{-1} \sum_{t=1}^T \varepsilon_t U_t'$ tends to zero in probability by the law of large numbers. Since $\beta_\perp' M_{12} \in O_p(1)$ and M_{22} converges in probability by the law of large numbers, we conclude that (A.11) holds. The result in (A.12) follows immediately from (A.10) and the continuous mapping theorem. Finally (A.13) holds by applying the central limit theorem to the MDS $\beta' X_{t-1} \varepsilon_t'$, rewriting $S_{1\varepsilon}$ as above. □

A.2 Proof of Theorem 3

While our results are new and generalize the results in Swensen (2006), we closely follow the sequence of arguments in Swensen (2006). As there we use P^* to denote the bootstrap probability and likewise E^* to denote expectation under P^*. Thus, as in Swensen (2006, proof of Proposition 1), the weak convergence in probability result in Theorem 3, $Q_r^b \xrightarrow{w}_p Q_{r,\infty}$, can be shown to hold by using Lemmas A.6 and A.7

below. These extend Lemmas A.2 and A.3 in the proof of Theorem 1 to the case of the wild bootstrap data. Specifically, Lemmas A.4, A.5, A.7 and A.6 below extend and generalize Lemmas 1, S1 and S2 used in Swensen (2006, proof of Proposition 1) for IID bootstrap shocks.

Establishing that $Q_r^b \xrightarrow{w}_p Q_{r,\infty}$ implies $G_{r,T}^b(\cdot) \to G_{r,\infty}(\cdot)$, uniformly in probability, where $G_{r,\infty}$ denotes the cumulative distribution function of $Q_{r,\infty}$. Then, using the same arguments as in the proof of Theorem 5 in Hansen (2000b), it is entirely straightforward to prove that $p_{r,T}^b \xrightarrow{w} U[0,1]$ given the foregoing results. This completes the proof.

We now move to establishing the intermediate lemmas referred to above, establishing a Granger-type representation and an invariance principle for the bootstrap data, analogous to those given for the original data in Lemmas A.1 and A.3 respectively.

Lemma A.4 *Under the conditions of Theorem 1,*

$$X_t^b = \hat{C}\sum_{i=1}^{t}\varepsilon_i^b + T^{1/2}R_t^b$$

where

$$\begin{aligned}\hat{C} &:= (I_p,0,...,0)\,\widehat{\mathbb{B}}_\perp(\widehat{\mathbb{A}}'_\perp\widehat{\mathbb{B}}_\perp)^{-1}\widehat{\mathbb{A}}' = \hat{\beta}_\perp(\hat{\alpha}'_\perp\hat{\Gamma}\hat{\beta}_\perp)^{-1}\hat{\alpha}'_\perp \\ R_t^b &:= (\hat{\alpha},\hat{\Psi})(\widehat{\mathbb{B}}'\widehat{\mathbb{A}})^{-1}\sum_{i=0}^{t-1}\hat{\Phi}^i(T^{-1/2}\widehat{\mathbb{B}}'e_{t-i}^b)\end{aligned}$$

and where for all $\eta > 0$, $P^*\left(\max_{t=1,...,T}\left\|R_t^b\right\| > \eta\right) \to 0$ *in probability as* $T \to \infty$.

Proof: From the proof of Lemma A.1 with $\mathbb{X}_t^b := \left(X_t^{b\prime},...,X_{t-k+1}^{b\prime}\right)'$ and $\mathbb{X}_0^b := 0$ we find directly as in (A.4) that $X_t^b = (I_p,0,...,0)\,\mathbb{X}_t^b$ has the representation,

$$X_t^b = \hat{C}\sum_{i=1}^{t}\varepsilon_i^b + T^{1/2}R_t^b$$

where $\hat{C}$ and R_t^b as as defined in Lemma A.1, and where $\hat{\Phi} := (I_{pk}+\widehat{\mathbb{B}}'\widehat{\mathbb{A}})$ and $\hat{\Psi} := (\hat{\Gamma}_1,...,\hat{\Gamma}_{k-1})$. Note that in the definition of R_t^b the sum is *not* infinite as the bootstrap residuals are defined for $t \geq 1$ only. The matrices $\widehat{\mathbb{A}}$ and $\widehat{\mathbb{B}}$ are defined as $\mathbb{A}$,$\mathbb{B}$ of (A.3) with α and β replaced by the corresponding estimators $\hat{\alpha},\hat{\beta}$, and $e_t^b := \left(\varepsilon_t^{b\prime},0,...,0\right)'$. The proof is then completed along the same lines as the proof of Lemma A.4 in CRT; see the accompanying working paper, Cavaliere *et al.* (2009), for details. □

Lemma A.5 *Under the conditions of Theorem 1,*

$$S_T^b(\cdot) := \frac{1}{T^{1/2}}\sum_{t=1}^{\lfloor T\cdot\rfloor}\varepsilon_t^b \xrightarrow{w}_p W(\cdot)\ .$$

Proof: Paralleling arguments made in the proof of Lemma A.5 in CRT, the stated result follows on establishing pointwise convergence for $T^{-1}\sum_{t=1}^{\lfloor Tu \rfloor} \varepsilon_t \varepsilon_t' \to u\Sigma$, which indeed follows by the law of large numbers. □

Lemma A.6 *Let $G(\cdot)$ be defined as in (A.9). Then under the conditions of Theorem 1,*

$$\tfrac{1}{\sqrt{T}}\hat{\beta}_\perp' X_{\lfloor Tu \rfloor}^b \overset{w}{\to}_p G(u) \tag{A.15}$$

$$\hat{\beta}_\perp' S_{10}^b \alpha_\perp = \hat{\beta}_\perp' S_{12}^b \alpha_\perp \overset{w}{\to}_p \int_0^1 G(s)\, dW(s)' \alpha_\perp \tag{A.16}$$

$$\tfrac{1}{T}\hat{\beta}_\perp' S_{11}^b \hat{\beta}_\perp \overset{w}{\to}_p \int_0^1 G(s)G(s)'ds \tag{A.17}$$

and furthermore,

$$\sqrt{T}\hat{\beta}' S_{10}^b \hat{\alpha}_\perp = \sqrt{T}\hat{\beta}' S_{1\varepsilon}^b \hat{\alpha}_\perp \overset{w}{\to}_p N_{r\times p-r}(0, \Sigma_{\beta\beta} \otimes \alpha_\perp' \Sigma \alpha_\perp) \tag{A.18}$$

$$\hat{\beta}' S_{11}^b \hat{\beta} \in O_{p^*}(1) \tag{A.19}$$

in probability as $T \to \infty$.

Proof: Applying Lemma A.4 and Lemma A.5, the results hold as in Lemma S2 of Swensen (2006). □

Lemma A.7 *Under the conditions of Theorem 3,*

$$P^*\left(\left\|S_{00}^b - \Sigma_{00}\right\| > \eta\right) \to 0 \tag{A.20}$$

$$P^*\left(\left\|S_{01}^b\hat{\beta} - \Sigma_{0\beta}\right\| > \eta\right) \to 0 \tag{A.21}$$

$$P^*\left(\left\|\hat{\beta}' S_{11}^b \hat{\beta} - \Sigma_{\beta\beta}\right\| > \eta\right) \to 0 \tag{A.22}$$

in probability as $T \to \infty$.

Proof: In the interests of brevity, we only provide a proof of (A.20) here. Proofs of (A.21) and (A.22) can be obtained on request. Notice that $S_{00}^b = M_{00}^b - M_{02}^b \left(M_{22}^b\right)^{-1} M_{20}^b$ where the M_{ij}^b are the product moments in terms of the bootstrap data. Hence, as noted in Swensen (2006), (A.20) follows by establishing that $P^*\left(\left\|M^b - \Sigma_M\right\| > \eta\right) \to 0$, where

$$M := \frac{1}{T}\sum_{t=1}^{T} \Delta\mathbb{X}_t \Delta\mathbb{X}_t', \; M^b := \frac{1}{T}\sum_{t=1}^{T} \Delta\mathbb{X}_t^b \Delta\mathbb{X}_t^{b\prime} \text{ and } \Sigma_M := \operatorname*{plim}_{T\to\infty} M$$

with $\mathbb{X}_t := \left(X_t', ..., X_{t-k+1}'\right)'$ and $\mathbb{X}_t^b := \left(X_t^{b\prime}, ..., X_{t-k+1}^{b\prime}\right)'$. By Lemma A.1, $\mathbb{X}_{\beta t} = \sum_{i=0}^{\infty} \Phi^i \eta_{t-i}$ and, hence, (A.2), implies that

$$\Delta\mathbb{X}_t = \mathbb{A}\sum_{i=1}^{\infty} \Phi^{i-1}(\beta, I, 0, ..., 0)' \varepsilon_{t-i} + (I, 0, ..., 0)' \varepsilon_t := \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}. \tag{A.23}$$

Similarly, $\Delta\mathbb{X}_t^b = \sum_{i=0}^{t-1} \hat{\theta}_i \varepsilon_{t-i}^b$, $\varepsilon_t^b = \hat{\varepsilon}_t w_t$. As previously noted in the proof of Lemma A.4, for sufficiently large T, $\|\Phi^i\|, \left\|\hat{\Phi}^i\right\| < c\lambda^i$ for some generic constant $c > 0$, $0 < \lambda < 1$, uniformly in i. In particular, the coefficients θ_i and $\hat{\theta}_i$ are exponentially decreasing. Next recall that $\Sigma_M = \sum_{i=0}^{\infty} \theta_i \Sigma \theta_i'$, and observe that with $\Sigma_{M^b} := \mathrm{E}^*\left(M^b\right)$,

$$\left\|M^b - \Sigma_M\right\| \leq \left\|M^b - \Sigma_{M^b}\right\| + \left\|\Sigma_{M^b} - \Sigma_M\right\|.$$

To see that $\|\Sigma_{M^b} - \Sigma_M\|$ tends to zero in probability rewrite first Σ_{M^b} as:

$$\begin{aligned}\Sigma_{M^b} &= \mathrm{E}^*\left(\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\hat{\theta}_i\varepsilon_{t-i}^b\right)\left(\sum_{i=0}^{t-1}\hat{\theta}_i\varepsilon_{t-i}^b\right)'\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\hat{\theta}_i\hat{\varepsilon}_{t-i}\hat{\varepsilon}_{t-i}'\hat{\theta}_i'\right) \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{T-t}\hat{\theta}_i\hat{\varepsilon}_t\hat{\varepsilon}_t'\hat{\theta}_i'\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{\infty}\hat{\theta}_i\hat{\varepsilon}_t\hat{\varepsilon}_t'\hat{\theta}_i'\right) - V_{1T},\end{aligned}$$

where $V_{1T} := \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=T-t+1}^{\infty}\hat{\theta}_i\hat{\varepsilon}_t\hat{\varepsilon}_t'\hat{\theta}_i'\right) = o_p(1)$. To see this, use the fact that $\theta_i = A\Phi^i B$, where A and B are constant matrices, see (A.23), and $\hat{\theta}_i = \hat{A}\hat{\Phi}^i\hat{B}$. In particular, for sufficiently large T, $\left\|\hat{\theta}_i\right\| \leq c\lambda^i$, uniformly in i, and the result holds as $\mathrm{E}\|\varepsilon_t\|^4 < K < \infty$ and $\sum_{i=T-t+1}^{\infty}\lambda^{T-i} \to 0$ as $T \to \infty$. Next, observe that

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{\infty}\hat{\theta}_i\hat{\varepsilon}_t\hat{\varepsilon}_t'\hat{\theta}_i'\right) - \Sigma_M \\ &= \left(\sum_{i=0}^{\infty}\hat{\theta}_i\left(\frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t'\right)\hat{\theta}_i' - \sum_{i=0}^{\infty}\hat{\theta}_i\Sigma\hat{\theta}_i'\right) + \left(\sum_{i=0}^{\infty}\hat{\theta}_i\Sigma\hat{\theta}_i' - \Sigma_M\right) \\ &=: V_{2T} + V_{3T}.\end{aligned}$$

It then follows that, as $T \to \infty$,

$$\|V_{2T}\| \leq \left\|\sum_{i=0}^{\infty}\left(\hat{\theta}_i \otimes \hat{\theta}_i\right)\right\|\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t' - \Sigma\right\| \xrightarrow{p} 0$$

by the result that $T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t' \xrightarrow{p} \Sigma$ (see Theorem 2), and because $\left\|\sum_{i=0}^{\infty}\left(\hat{\theta}_i \otimes \hat{\theta}_i\right)\right\|$ is of order one. Also,

$$\mathsf{vec}\left(V_{3T}\right) = \left(\sum_{i=0}^{\infty}\left(\hat{\theta}_i \otimes \hat{\theta}_i\right) - \sum_{i=0}^{\infty}\left(\theta_i \otimes \theta_i\right)\right)\mathsf{vec}\left(\Sigma\right) \xrightarrow{p} 0,$$

using, as above, the fact that $\theta_i = A\Phi^i B$ and $\hat{\theta}_i = \hat{A}\hat{\Phi}^i\hat{B}$.

Finally, consider the term $\left\|M^b - \Sigma_{M^b}\right\|$. We have

$$\begin{aligned} M^b &= \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\hat{\theta}_i\varepsilon_{t-i}^b\right)\left(\sum_{i=0}^{t-1}\hat{\theta}_i\varepsilon_{t-i}^b\right)' \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\hat{\theta}_i\varepsilon_{t-i}^b\left(\hat{\theta}_i\varepsilon_{t-i}^b\right)'\right) + \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i,j=0,i\neq j}^{t-1}\hat{\theta}_i\varepsilon_{t-i}^b\varepsilon_{t-j}^{b\prime}\hat{\theta}_j'\right) \\ &=: M_1^b + M_2^b. \end{aligned}$$

First, notice that

$$M_1^b - \Sigma_{M^b} = \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\hat{\theta}_i\hat{\varepsilon}_{t-i}\hat{\varepsilon}_{t-i}'\hat{\theta}_i'\kappa_{t-i}\right)$$

with $\kappa_t := (w_t^2 - 1)$ an i.i.d. process with mean zero and finite moments of all order. Now, since

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\mathsf{vec}\left(\hat{\theta}_i\hat{\varepsilon}_{t-i}\hat{\varepsilon}_{t-i}'\hat{\theta}_i'\kappa_{t-i}\right)\right) &= \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\kappa_{t-i}\left(\hat{\theta}_i\otimes\hat{\theta}_i\right)\mathsf{vec}\left(\hat{\varepsilon}_{t-i}\hat{\varepsilon}_{t-i}'\right)\right) \\ &= \frac{1}{T}\sum_{t=1}^{T}\kappa_t\sum_{i=0}^{T-t}\left(\hat{\theta}_i\otimes\hat{\theta}_i\right)\mathsf{vec}\left(\hat{\varepsilon}_t\hat{\varepsilon}_t'\right), \end{aligned}$$

it therefore follows that,

$$\begin{aligned} P^*\left(\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\hat{\theta}_i\hat{\varepsilon}_{t-i}\hat{\varepsilon}_{t-i}'\hat{\theta}_i'\kappa_{t-i}\right)\right\| > \delta\right) &\leq \frac{1}{T^2\delta^2}\sum_{t=1}^{T}\mathrm{E}^*\left\|\kappa_t\sum_{i=0}^{T-t}\left(\hat{\theta}_i\otimes\hat{\theta}_i\right)\mathsf{vec}\left(\hat{\varepsilon}_t\hat{\varepsilon}_t'\right)\right\|^2 \\ &\leq \frac{\mathrm{E}\left(\kappa_t^2\right)}{T\delta^2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\|\sum_{i=0}^{T-t}\left(\hat{\theta}_i\otimes\hat{\theta}_i\right)\mathsf{vec}\left(\hat{\varepsilon}_t\hat{\varepsilon}_t'\right)\right\|^2\right). \end{aligned}$$

Thus, with $c_T = c + o_p(1)$,

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\left(\sum_{i=0}^{T-t}\hat{\theta}_i\otimes\hat{\theta}_i\right)\mathsf{vec}\left(\hat{\varepsilon}_t\hat{\varepsilon}_t'\right)\right\|^2 \leq \frac{c_T}{T}\sum_{t=1}^{T}\left\|\mathsf{vec}\left(\hat{\varepsilon}_t\hat{\varepsilon}_t'\right)\right\|^2$$

which converges in probability as ε_t has bounded fourth order moment. This establishes the result that $M_1^b - \Sigma_{M^b} = o_p(1)$. It can similarly be shown that $M_2^b = o_p(1)$, which completes the proof. □

A.3 Proof of Theorem 4

We proceed as in the proof of Theorem 3. Specifically, we establish that the results in Lemmas A.4, A.5, A.7 and A.6 also hold for the i.i.d. bootstrap. Without causing confusion, we now denote by P^* the i.i.d. bootstrap probability and likewise E^* denotes expectation under P^*. Objects with a superscript s in what follows are understood to be the i.i.d. bootstrap analogues of the corresponding wild bootstrap quantities with a superscript b.

Consider first the analogue of Lemma A.4.

Lemma A.8 *Under the conditions of Theorem 1, the i.i.d. bootstrap data satisfy,*

$$X_t^s = \hat{C} \sum_{i=1}^{t} \varepsilon_i^s + T^{1/2} R_t^s$$

where for all $\eta > 0$, $P^* \left(\max_{t=1,\ldots,T} \| R_t^* \| > \eta \right) \to 0$ *in probability as* $T \to \infty$.

Proof: The arguments are identical to the proof of Lemma A.4 apart from the final evaluation of $P^* \left(T^{-1/2} \max_{t=1,\ldots,T} \left\| \varepsilon_t^b \right\| > \eta \right)$ in the i.i.d. case. Using that under i.i.d. bootstrap,

$$\mathrm{E}^* \left(\varepsilon_t^{s\prime} \varepsilon_t^s \right)^2 = \frac{1}{T} \sum_{t=1}^{T} \left(\hat{\varepsilon}_t^\prime \hat{\varepsilon}_t \right)^2 ,$$

one finds,

$$P^* \left(T^{-1/2} \max_{t=1,\ldots,T} \| \varepsilon_t^s \| > \eta \right) \leq \frac{1}{\eta^4 T^2} \sum_{t=1}^{T} \mathrm{E}^* \left(\varepsilon_t^{s\prime} \varepsilon_t^s \right)^2 = \frac{1}{\eta^4 T^2} \sum_{t=1}^{T} \left(\hat{\varepsilon}_t^\prime \hat{\varepsilon}_t \right)^2 = O_p \left(\frac{1}{T} \right) \xrightarrow{p} 0$$

□

That Lemmas A.5 and A.7 hold for the i.i.d. bootstrap case holds by Lemma S2 of Swensen (2006). Finally, we need the analogue of Lemma A.7 for the i.i.d. case:

Lemma A.9 *For the i.i.d. bootstrap and under the conditions of Theorem 4,*

$$P^* \left(\| S_{00}^s - \Sigma_{00} \| > \eta \right) \to 0 \tag{A.24}$$

$$P^* \left(\left\| S_{01}^s \hat{\beta} - \Sigma_{0\beta} \right\| > \eta \right) \to 0 \tag{A.25}$$

$$P^* \left(\left\| \hat{\beta}^\prime S_{11}^s \hat{\beta} - \Sigma_{\beta\beta} \right\| > \eta \right) \to 0 \tag{A.26}$$

in probability as $T \to \infty$.

Proof: Proceed as in the proof of Lemma A.7 to reach the identical inequality:

$$\| M^s - \Sigma_M \| \leq \| M^s - \Sigma_{M^s} \| + \| \Sigma_{M^s} - \Sigma_M \| .$$

For evaluation of the last term, re-write Σ_{M^s} as:

$$\Sigma_{M^s} = \mathrm{E}^* \left(\frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \hat{\theta}_i \varepsilon_{t-i}^s \right) \left(\sum_{i=0}^{t-1} \hat{\theta}_i \varepsilon_{t-i}^s \right)' \right) = \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \sum_{j=0}^{t-1} \hat{\theta}_i \mathrm{E}^* \left(\varepsilon_{t-i}^s \varepsilon_{t-j}^{s\prime} \right) \hat{\theta}_j' \right)$$
$$= \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \hat{\theta}_i \hat{\Sigma}_T \hat{\theta}_i' \right)$$

where $\hat{\Sigma}_T = \frac{1}{T} \sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$, and making use of the fact that ε_t^s are conditionally independent. Re-write again,

$$\Sigma_{M^s} = \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \hat{\theta}_i \hat{\Sigma}_T \hat{\theta}_i' \right) = \sum_{i=0}^{\infty} \hat{\theta}_i \hat{\Sigma}_T \hat{\theta}_i' - \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=T-t+1}^{\infty} \hat{\theta}_i \hat{\Sigma}_T \hat{\theta}_i' \right). \tag{A.27}$$

The last term tends to zero by the arguments in the proof of Lemma A.7 for $V_{1T} \xrightarrow{p} 0$ and using the result that $\hat{\Sigma}_T \xrightarrow{p} \Sigma$ by consistency. Likewise, the first term in (A.27) tends in probability to Σ_M as desired. This holds by rewriting it as $V_{2T} + V_{3T}$, these objects defined analogously as in the proof of Lemma A.7, and using the arguments there to show that $V_{2T} \to 0$, while $V_{3T} \to \Sigma$ in probability.

Turning to the final term $\|M^s - \Sigma_{M^s}\|$, we have that

$$M^s = \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \hat{\theta}_i \varepsilon_{t-i}^s \left(\hat{\theta}_i \varepsilon_{t-i}^s \right)' \right) + \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i,j=0, i \neq j}^{t-1} \hat{\theta}_i \varepsilon_{t-i}^s \varepsilon_{t-j}^{s\prime} \hat{\theta}_j' \right) =: M_1^s + M_2^s.$$

First, observe that, $M_1^s - \Sigma_{M^s} = \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \hat{\theta}_i \left(\varepsilon_{t-i}^s \varepsilon_{t-i}^{s\prime} - \hat{\Sigma}_T \right) \hat{\theta}_i' \right)$. Using the vec$(\cdot)$ operator and interchanging summation,

$$\frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \mathsf{vec} \left(\hat{\theta}_i \left(\varepsilon_{t-i}^s \varepsilon_{t-i}^{s\prime} - \hat{\Sigma}_T \right) \hat{\theta}_i' \right) \right) = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=0}^{T-t} \left(\hat{\theta}_i \otimes \hat{\theta}_i \right) \mathsf{vec} \left(\varepsilon_t^s \varepsilon_t^{s\prime} - \hat{\Sigma}_T \right).$$

Hence,

$$P^* \left(\left\| \frac{1}{T} \sum_{t=1}^{T} \left(\sum_{i=0}^{t-1} \hat{\theta}_i \left(\varepsilon_{t-i}^s \varepsilon_{t-i}^{s\prime} - \hat{\Sigma}_T \right) \hat{\theta}_i' \kappa_{t-i} \right) \right\| > \delta \right)$$
$$\leq \frac{1}{T^2 \delta^2} \sum_{t=1}^{T} \mathrm{E}^* \left\| \sum_{i=0}^{T-t} \left(\hat{\theta}_i \otimes \hat{\theta}_i \right) \mathsf{vec} \left(\varepsilon_t^s \varepsilon_t^{s\prime} - \hat{\Sigma}_T \right) \right\|^2.$$

Thus, with $c_T = c + o_p(1)$,

$$\frac{1}{T} \sum_{t=1}^{T} \mathrm{E}^* \left\| \left(\sum_{i=0}^{T-t} \hat{\theta}_i \otimes \hat{\theta}_i \right) \mathsf{vec} \left(\varepsilon_t^s \varepsilon_t^{s\prime} - \hat{\Sigma}_T \right) \right\|^2 \leq \frac{c_T}{T} \sum_{t=1}^{T} \mathrm{E}^* \left\| \mathsf{vec} \left(\varepsilon_t^s \varepsilon_t^{s\prime} - \hat{\Sigma}_T \right) \right\|^2.$$

Use next that,

$$\mathrm{E}^* \left\| \mathsf{vec}\left(\varepsilon_t^s \varepsilon_t^{s\prime} - \hat{\Sigma}_T\right)\right\|^2 = \mathrm{E}^* tr\left(\varepsilon_t^s \varepsilon_t^{s\prime} - \hat{\Sigma}_T\right)^2 = \frac{1}{T}\sum_{t=1}^{T} tr\left(\hat{\varepsilon}_t \hat{\varepsilon}_t' - \hat{\Sigma}_T\right)^2$$

which converges in probability as a result of the assumption that ε_t has bounded fourth order moment. This establishes the result that $M_1^s - \Sigma_{M^s} = o_p(1)$. Similarly $M_2^s = o_p(1)$, which completes the proof. □

References

Basawa, I.V., A.K. Mallik, W.P. McCormick, J.H. Reeves and R.L. Taylor (1991), Bootstrapping unstable first-order autoregressive processes, *Annals of Statistics* 19, 1098-1101.

Brown, B.M. (1971), Martingale central limit theorems, *The Annals of Mathematical Statistics* 42, 59-66.

Burridge, P. and A.M.R. Taylor (2001), On regression-based tests for seasonal unit roots in the presence of periodic heteroscedasticity, *Journal of Econometrics* 104, 91–117.

Campbell, J.Y. and R.J. Shiller (1987), Cointegration and tests of present value models, *Journal of Political Economy* 95, 1062—1088.

Cavaliere G. and A.M.R. Taylor (2008), Bootstrap unit root tests for time series with non-stationary volatility, *Econometric Theory* 24, 43-71.

Cavaliere, G., A. Rahbek and A.M.R. Taylor (2007), Testing for cointegration in vector autoregressions with non-stationary volatility, Granger Center Discussion Paper 07/02. Forthcoming in *Journal of Econometrics*

Cavaliere, G., A. Rahbek and A.M.R. Taylor (2009), Co-integration Rank Testing under Conditional Heteroskedasticity, Granger Center Discussion Paper 09/02.

Cheng, X. and P.C.B. Phillips, (2009), Semiparametric Cointegrating Rank Selection, *The Econometrics Journal*, forthcoming.

Cheng, X. and P.C.B. Phillips, (2008), Cointegrating Rank Selection in Models with Time-Varying Variance, Cowles Foundation Discussion Paper Number 1688.

Diebold, F.X., L. Ji and C. Li (2007), A three-factor yield curve model: non-affine structure, systematic risk sources, and generalized duration, forthcoming in L.R. Klein (ed.), *Macroeconomics, Finance and Econometrics: Essays in Memory of Albert Ando*, Cheltenham, U.K.: Edward Elgar.

Elliott, G. (1998), The robustness of cointegration methods when regressors almost have unit roots, *Econometrica* 66, 149-58, 1998.

Engle, R.F. (1990), Discussion: stock market volatility and the crash of '87, *Review of Financial Studies* 3, 103-106.

Giné, E. and J. Zinn (1990), Bootstrapping general empirical measures, *Annals of Probability* 18, 851–869.

Giese J. (2006), Level, slope, curvature: characterising the yield curve's derivatives in a cointegrated VAR model, Working paper, Nuffield College.

Glosten, L.R., R. Jaganathan and D.E. Runkle (1993), On the relation between the expected value and the volatility of nominal excess returns on stocks, *Journal of Finance* 48, 1779-1801.

Gonçalves, S. and L. Kilian (2004), Bootstrapping autoregressions with conditional heteroskedasticity of unknown form, *Journal of Econometrics* 123, 89-120.

Gonçalves, S. and L. Kilian (2007), Asymptotic and bootstrap inference for $AR(\infty)$ processes with conditional heteroskedasticity, *Econometric Reviews* 26, 609-641.

Hall, A.D., H. Anderson and C.W.J. Granger (1992), A cointegration analysis of treasury bill yields, *Review of Economics and Statistics* 74, 116–26.

Hannan, E.J. and C.C. Heyde (1972), On limit theorems for quadratic functions of discrete time series, *The Annals of Mathematical Statistics* 43, 2058-2066.

Hansen, B.E. (1992), Convergence to stochastic integrals for dependent heterogeneous processes, *Econometric Theory* 8, 489–500.

Hansen, B.E. (1996), Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* 64, 413–430.

Hansen, B.E. (2000a), Sample splitting and threshold estimation, *Econometrica* 68, 575–603.

Hansen, B.E. (2000b), Testing for structural change in conditional models, *Journal of Econometrics* 97, 93–115.

Harris, R.I.D. and G. Judge (1998), Small Sample Testing for Cointegration Using the Bootstrap Approach, *Economics Letters* 58, 31–37.

Inoue, A. and L. Kilian (2002), Bootstrapping autoregressive processes with possible unit roots, *Econometrica* 70, 377–391.

Jensen, S.T. and A. Rahbek (2007), On the law of large numbers for (geometrically) ergodic Markov chains, *Econometric Theory* 23, 761–766

Johansen, S. (1996), *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford: Oxford University Press.

Johansen, S. (2002), A small sample correction of the test for cointegrating rank in the vector autoregressive model, *Econometrica* 70, 1929–1961.

Kristensen D. and A. Rahbek (2005), Asymptotics of the QMLE for a class of ARCH(q) Models, *Econometric Theory* 21, 946–961.

Kristensen D. and A. Rahbek (2009), Aymptotics of the QMLE for General ARCH(q) Models, *Journal of Time Series Econometrics*, forthcoming.

Lee, T.-H. and Y. Tse (1996), Cointegration tests with conditional heteroskedasticity, *Journal of Econometrics* 73, 401-410.

Liu, R.Y. (1988), Bootstrap procedure under some non-iid models, *Annals of Statistics* 16, 1696-1708.

MacKinnon, J.G., A.A. Haug, and L. Michelis (1999), Numerical distribution functions of likelihood ratio tests for cointegration, *Journal of Applied Econometrics*14, 563–577.

Mammen, E. (1993), Bootstrap and wild bootstrap for high dimensional linear models, *Annals of Statistics* 21, 255–285.

Mantalos, P. and G. Shukur (2001), Bootstrapped Johansen Tests for Cointegration Relationships: A Graphical Analysis, *Journal of Statistical Computation and Simulation* 68, 351–371.

Nelson, D.B. (1991), Conditional heteroskedasticity in asset returns: a new approach, *Econometrica* 59, 347-369.

Nielsen, B. and A. Rahbek (2000), Similarity issues in cointegration analysis, *Oxford Bulletin of Economics and Statistics* 62, 5-22.

Paparoditis, E.P. and D.N. Politis (2003), Residual-based block bootstrap for unit root testing, *Econometrica* 71, 813-855.

Park, J.Y. (2003), Bootstrap unit root tests, *Econometrica* 71, 1845–1895.

Pesavento, E. (2004), An analytical evaluation of the power of tests for the absence of cointegration, *Journal of Econometrics* 122, 349–384.

Rahbek, A., E. Hansen and J.G. Dennis (2002), ARCH innovations and their impact on cointegration rank testing. Downloadable from `http://www.math.ku.dk/~rahbek/`

Swensen, A.R. (2006), Bootstrap algorithms for testing and determining the cointegration rank in VAR models, *Econometrica* 74, 1699–1714.

Trenkler, C. (2009), Bootstrapping systems cointegration tests with a prior adjustment for deterministic terms, *Econometric Theory* 25, 243–269.

van der Weide, R. (2002), GO-GARCH: a multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics* 17, 549-564.

van Giersbergen, N.P.A. (1996), Bootstrapping the Trace Statistics in VAR Models: Monte Carlo Results and Applications, *Oxford Bulletin of Economics and Statistics* 58, 391–408.

Wu, C.F.J. (1986), Jackknife, bootstrap, and other resampling methods, *Annals of Statistics* 14, 1261–1295.

Xu, K.-L. (2008), Bootstrapping autoregression under non-stationary volatility, *Econometrics Journal* 11, 1-26.

Table 1: Size of Standard and Bootstrap PLR Tests for Rank = 0 Against Rank = p. True Rank is 0.

			$p=2$			$p=3$			$p=4$			$p=5$		
Model A: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
d_0	d_1	T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
0.0	0.0	50	6.3	5.7	4.9	7.0	5.1	4.9	8.7	4.3	4.8	12.1	3.7	4.8
		100	5.3	4.9	4.6	6.6	5.1	5.5	7.3	5.0	5.1	8.1	4.4	4.7
		200	5.3	4.6	4.8	6.1	5.3	5.2	6.5	4.9	4.9	6.9	4.4	4.7
0.5	0.0	50	9.9	7.2	7.9	10.7	6.6	7.8	13.6	6.3	8.5	17.6	6.2	9.1
		100	7.3	5.3	6.5	9.8	6.3	7.9	11.2	6.3	8.3	12.6	5.4	8.2
		200	6.6	4.8	5.9	8.3	5.3	7.2	8.7	5.2	6.7	9.6	4.3	6.8
0.3	0.65	50	10.2	6.8	8.3	12.6	7.2	9.6	14.8	6.4	9.4	18.1	6.2	9.5
		100	9.9	5.6	8.5	12.3	6.5	10.3	13.7	6.2	10.7	14.8	6.3	10.0
		200	10.0	5.6	9.3	10.7	5.2	9.5	12.1	5.6	10.0	13.9	5.7	10.9
0.2	0.79	50	9.3	6.6	7.6	11.2	7.1	8.3	13.8	5.9	8.2	16.2	5.5	7.7
		100	9.9	5.6	8.7	11.4	6.4	9.8	13.1	6.2	9.9	14.0	5.5	9.2
		200	10.8	5.5	10.1	12.2	5.4	11.0	12.8	5.5	10.7	13.5	5.6	10.6
0.05	0.94	50	6.5	5.9	5.2	7.6	5.5	5.2	9.3	4.6	5.3	12.3	4.2	4.9
		100	5.8	4.9	5.2	7.0	5.4	5.6	8.1	5.2	5.8	8.8	4.4	4.9
		200	6.5	5.1	5.9	7.2	5.1	6.5	7.2	5.0	5.5	7.9	4.9	5.5
Model B: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. t_5, $i=1,...,p$														
d_0	d_1	T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
0.0	0.0	50	6.6	5.2	5.1	8.0	4.9	5.8	9.3	4.4	5.6	12.8	3.6	5.0
		100	5.7	4.9	5.0	6.3	4.7	4.9	6.7	4.1	5.0	8.2	4.4	4.8
		200	5.5	4.7	5.0	5.8	4.6	4.6	6.3	4.8	4.9	6.5	3.8	4.4
0.5	0.0	50	8.5	6.0	6.8	11.3	6.4	7.9	12.7	5.6	8.2	15.9	4.8	7.4
		100	7.3	5.3	6.4	8.4	5.2	6.8	9.5	5.1	6.9	12.0	5.1	7.2
		200	6.5	5.0	5.6	6.9	4.7	5.9	8.1	4.9	6.4	8.6	4.3	6.1
0.3	0.65	50	8.7	5.8	7.1	11.0	6.2	7.8	12.6	5.9	7.9	15.8	4.9	7.2
		100	7.5	5.1	6.5	9.2	5.5	7.7	10.4	5.5	7.7	12.4	5.6	7.5
		200	7.2	5.2	6.6	8.2	5.2	7.1	9.5	5.1	7.4	10.2	4.7	7.2
0.2	0.79	50	8.0	5.6	6.4	10.5	6.0	7.6	11.7	5.2	7.3	14.5	4.7	6.4
		100	7.2	5.4	6.2	8.9	5.4	7.3	9.7	5.3	7.3	11.2	5.4	7.2
		200	7.1	5.0	6.2	8.3	5.0	7.0	8.7	5.1	7.4	9.7	4.5	6.6
0.05	0.94	50	6.9	5.2	5.3	8.9	5.4	6.2	9.9	4.6	5.9	13.1	3.9	5.2
		100	5.9	5.0	5.3	7.1	5.0	5.9	7.4	4.6	5.4	8.9	4.5	5.4
		200	5.8	4.7	5.4	6.5	4.6	5.5	7.2	5.1	5.8	7.2	4.0	4.8
Model C: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $\ln(h_{i,t})=-0.23+0.9\ln(h_{i,t-1})+0.25[\lvert v_{i,t-1}^2\rvert-0.3v_{i,t-1}]$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
		50	11.0	7.2	9.2	13.8	7.9	10.5	17.2	7.3	10.6	20.0	7.1	10.1
		100	10.3	5.6	9.2	12.9	6.6	10.7	14.6	6.3	11.1	16.8	7.0	11.8
		200	9.7	5.3	9.1	11.5	5.9	10.1	13.6	5.4	11.2	14.7	5.7	11.2
Model D: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.0216+0.6896h_{i,t-1}+0.3174[\varepsilon_{i,t-1}-0.1108]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
		50	11.8	6.8	9.9	14.4	7.9	11.1	16.9	7.6	11.3	19.8	6.6	10.3
		100	12.7	6.2	11.5	15.0	6.9	13.1	16.6	6.9	13.1	18.7	6.5	13.2
		200	13.9	5.6	13.0	17.0	6.0	15.0	17.9	6.3	15.3	20.2	6.4	16.1
Model E: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.005+0.7h_{i,t-1}+0.28[\lvert\varepsilon_{i,t-1}\rvert-0.23\varepsilon_{i,t-1}]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
		50	10.7	6.7	9.0	13.1	7.3	10.1	15.5	6.6	10.1	18.2	6.1	9.5
		100	11.1	5.7	10.0	13.0	6.3	11.2	14.9	6.2	11.8	16.1	5.8	11.4
		200	12.0	4.9	11.3	14.1	5.7	12.5	16.0	5.6	13.6	16.9	5.4	13.8
Model F: $\varepsilon_{i,t}=v_{i,t}\exp(h_{i,t})$, $h_{i,t}=\lambda h_{i,t-1}+0.5\xi_{i,t}$, $(\xi_{i,t},v_{i,t})\sim$ i.i.d. $N(0,\mathrm{diag}(\sigma_\xi^2,1))$, $i=1,...,p$														
λ	σ_ξ	T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
0.936	0.424	50	19.3	8.4	16.9	24.5	9.1	19.1	29.3	9.7	20.8	35.0	11.0	22.2
		100	21.3	6.8	19.1	26.8	8.5	23.2	32.2	8.7	26.3	35.4	9.5	27.0
		200	22.0	6.8	20.1	27.3	7.6	24.6	32.7	7.8	28.1	37.1	7.9	30.8
0.951	0.314	50	16.5	7.1	13.7	20.0	8.2	16.3	24.0	8.4	16.5	28.1	9.0	17.3
		100	17.5	6.5	15.6	22.2	7.4	19.2	25.4	7.9	20.8	28.0	8.7	21.5
		200	18.6	6.6	17.2	22.8	6.7	20.5	25.9	6.6	22.2	30.5	7.7	24.9

Table 2: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p=2$, True Rank is 0.

			Q-based			Q^b-based			Q^s-based		
		$r=$	0	1	2	0	1	2	0	1	2
Model A: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$											
d_0	d_1	T									
0.0	0.0	50	93.7	5.5	0.8	94.3	4.6	1.1	95.1	3.9	1.0
		100	94.7	5.0	0.4	95.1	4.2	0.6	95.4	4.1	0.5
		200	94.7	5.0	0.3	95.4	3.9	0.7	95.2	4.3	0.5
0.5	0.0	50	90.1	9.2	0.7	92.8	6.1	1.1	92.1	6.8	1.1
		100	92.7	6.7	0.6	94.7	4.6	0.7	93.5	5.7	0.8
		200	93.4	6.0	0.6	95.2	4.1	0.6	94.1	5.2	0.7
0.3	0.7	50	89.8	9.1	1.0	93.2	5.5	1.2	91.7	7.0	1.3
		100	90.1	9.1	0.7	94.4	5.0	0.6	91.5	7.4	1.1
		200	90.0	9.1	0.9	94.4	4.9	0.7	90.7	7.9	1.4
0.2	0.8	50	90.7	8.3	1.0	93.4	5.4	1.2	92.4	6.4	1.2
		100	90.1	9.1	0.9	94.4	4.8	0.8	91.3	7.5	1.3
		200	89.2	9.7	1.0	94.5	4.9	0.6	89.9	8.6	1.5
0.1	0.9	50	93.5	5.8	0.7	94.1	4.7	1.3	94.8	4.1	1.2
		100	94.2	5.5	0.3	95.1	4.1	0.8	94.8	4.7	0.6
		200	93.5	6.1	0.4	94.9	4.5	0.7	94.1	5.1	0.8
Model B: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. t_5, $i=1,...,p$											
d_0	d_1	T									
0.0	0.0	50	93.4	6.0	0.6	94.8	4.2	1.1	94.9	4.2	1.0
		100	94.3	5.4	0.3	95.1	4.3	0.7	95.0	4.4	0.6
		200	94.5	5.0	0.5	95.3	4.1	0.6	95.0	4.2	0.8
0.5	0.0	50	91.5	7.7	0.8	94.0	4.8	1.2	93.2	5.8	1.0
		100	92.7	6.9	0.4	94.7	4.6	0.7	93.6	5.6	0.8
		200	93.5	6.1	0.5	95.0	4.4	0.6	94.4	4.8	0.8
0.3	0.7	50	91.3	8.0	0.7	94.2	4.7	1.1	92.9	6.1	1.0
		100	92.5	7.0	0.5	94.9	4.5	0.6	93.5	5.6	0.9
		200	92.8	6.5	0.6	94.8	4.6	0.7	93.4	5.7	0.9
0.2	0.8	50	92.0	7.3	0.8	94.4	4.5	1.1	93.6	5.3	1.1
		100	92.8	6.7	0.5	94.6	4.7	0.7	93.8	5.5	0.7
		200	92.9	6.5	0.6	95.0	4.4	0.6	93.8	5.5	0.8
0.1	0.9	50	93.1	6.2	0.7	94.8	4.2	1.0	94.7	4.4	0.9
		100	94.1	5.5	0.4	95.0	4.4	0.7	94.7	4.5	0.7
		200	94.2	5.3	0.5	95.3	4.0	0.7	94.6	4.5	0.8
Model C: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $\ln(h_{i,t})=-0.23+0.9\ln(h_{i,t-1})+0.25[\lvert v_{i,t-1}^2\rvert-0.3v_{i,t-1}]$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$											
		50	89.0	9.8	1.1	92.8	6.1	1.1	90.8	7.7	1.5
		100	89.7	9.4	0.9	94.4	4.9	0.7	90.8	8.0	1.2
		200	90.3	9.0	0.8	94.7	4.5	0.7	90.9	7.9	1.2
Model D: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.0216+0.6896h_{i,t-1}+0.3174[\varepsilon_{i,t-1}-0.1108]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$											
		50	88.2	10.6	1.2	93.2	5.5	1.3	90.1	8.2	1.7
		100	87.3	11.7	1.1	93.8	5.4	0.8	88.5	9.8	1.6
		200	86.1	12.6	1.3	94.4	4.9	0.7	87.0	11.1	2.0
Model E: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.005+0.7h_{i,t-1}+0.28[\lvert\varepsilon_{i,t-1}\rvert-0.23\varepsilon_{i,t-1}]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$											
		50	89.3	9.6	1.2	93.3	5.6	1.2	91.0	7.6	1.5
		100	88.9	10.1	1.0	94.3	5.1	0.7	90.0	8.8	1.2
		200	88.0	10.7	1.3	95.1	4.1	0.8	88.7	9.7	1.6
Model F: $\varepsilon_{i,t}=v_{i,t}\exp(h_{i,t})$, $h_{i,t}=\lambda h_{i,t-1}+0.5\xi_{i,t}$, $(\xi_{i,t},v_{i,t})\sim$ i.i.d. $N(0,\mathrm{diag}(\sigma_\xi^2,1))$, $i=1,...,p$											
λ	σ_ξ	T									
0.936	0.424	50	80.7	16.8	2.5	91.6	7.3	1.1	83.1	14.0	2.9
		100	78.7	19.0	2.4	93.2	5.9	1.0	80.9	16.4	2.6
		200	78.0	19.9	2.1	93.2	6.2	0.6	79.9	17.7	2.4
0.951	0.314	50	83.5	14.4	2.1	92.9	6.0	1.1	86.3	11.4	2.3
		100	82.5	15.6	1.9	93.5	5.5	1.0	84.4	13.2	2.3
		200	81.4	16.9	1.7	93.4	6.1	0.6	82.8	15.2	2.0

Table 3: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p=3$, True Rank is 0.

			Q-based				Q^b-based				Q^s-based			
		$r=$	0	1	2	3	0	1	2	3	0	1	2	3
Model A: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}, h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}, v_{i,t}\sim i.i.d.\ N(0,1), i=1,...,p$														
d_0	d_1	T												
0.0	0.0	50	93.0	6.1	0.7	0.1	94.9	4.2	0.6	0.3	95.1	4.0	0.7	0.2
		100	93.4	6.0	0.6	0.1	94.9	4.4	0.5	0.2	94.5	4.8	0.6	0.1
		200	93.9	5.5	0.6	0.0	94.7	4.6	0.5	0.1	94.8	4.5	0.5	0.2
0.5	0.0	50	89.3	9.7	0.9	0.1	93.4	5.6	0.7	0.3	92.2	7.0	0.6	0.2
		100	90.2	8.8	0.9	0.1	93.7	5.5	0.6	0.2	92.1	6.9	0.8	0.2
		200	91.7	7.5	0.7	0.1	94.7	4.7	0.5	0.1	92.8	6.3	0.6	0.3
0.3	0.65	50	87.4	11.2	1.1	0.3	92.8	5.9	1.1	0.2	90.4	8.1	1.1	0.4
		100	87.7	11.1	1.0	0.2	93.5	5.7	0.6	0.3	89.7	8.8	1.1	0.3
		200	89.3	9.8	0.8	0.1	94.8	4.6	0.4	0.2	90.5	8.4	0.9	0.2
0.2	0.79	50	88.8	10.0	0.9	0.2	92.9	5.9	0.9	0.3	91.7	7.2	0.8	0.3
		100	88.6	10.1	1.1	0.2	93.6	5.4	0.8	0.2	90.2	8.4	1.1	0.3
		200	87.8	11.0	1.0	0.2	94.6	4.6	0.6	0.2	89.0	9.5	1.1	0.4
0.05	0.94	50	92.4	6.6	0.8	0.2	94.5	4.5	0.6	0.4	94.8	4.3	0.7	0.2
		100	93.0	6.2	0.6	0.2	94.6	4.5	0.6	0.2	94.4	4.9	0.6	0.2
		200	92.8	6.6	0.4	0.1	94.9	4.5	0.5	0.2	93.5	5.9	0.4	0.2
Model B: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}, h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}, v_{i,t}\sim i.i.d.\ t_5, i=1,...,p$														
d_0	d_1	T												
0.0	0.0	50	92.0	7.3	0.6	0.2	95.1	4.2	0.5	0.2	94.2	5.1	0.5	0.3
		100	93.7	5.7	0.4	0.1	95.3	4.0	0.5	0.2	95.1	4.2	0.4	0.2
		200	94.2	5.2	0.6	0.0	95.4	4.0	0.5	0.1	95.4	4.1	0.5	0.1
0.5	0.0	50	88.7	10.3	0.8	0.1	93.6	5.4	0.8	0.2	92.1	6.8	0.9	0.2
		100	91.6	7.7	0.5	0.2	94.8	4.4	0.7	0.2	93.2	6.0	0.7	0.2
		200	93.1	6.4	0.5	0.1	95.3	4.2	0.4	0.1	94.1	5.3	0.4	0.2
0.3	0.65	50	89.0	9.9	0.9	0.1	93.8	5.4	0.6	0.2	92.2	6.9	0.7	0.2
		100	90.8	8.3	0.7	0.2	94.5	4.6	0.6	0.3	92.3	6.7	0.8	0.3
		200	91.8	7.5	0.6	0.1	94.8	4.6	0.4	0.2	92.9	6.2	0.6	0.3
0.2	0.79	50	89.5	9.4	1.0	0.1	94.0	5.2	0.6	0.2	92.4	6.5	0.9	0.3
		100	91.1	7.9	0.8	0.2	94.6	4.6	0.6	0.3	92.7	6.2	0.7	0.3
		200	91.7	7.6	0.6	0.1	95.0	4.4	0.5	0.2	93.0	6.2	0.5	0.3
0.05	0.94	50	91.1	8.1	0.7	0.1	94.6	4.6	0.5	0.3	93.8	5.4	0.6	0.2
		100	92.9	6.5	0.4	0.2	95.0	4.2	0.5	0.3	94.1	5.1	0.4	0.3
		200	93.5	6.0	0.5	0.1	95.4	4.1	0.4	0.1	94.5	4.8	0.4	0.2
Model C: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}, \ln(h_{i,t})=-0.23+0.9\ln(h_{i,t-1})+0.25[\lvert v_{i,t-1}^2\rvert-0.3v_{i,t-1}], v_{i,t}\sim i.i.d.\ N(0,1), i=1,...,p$														
		50	86.2	12.1	1.4	0.3	92.1	6.5	1.0	0.4	89.5	8.9	1.2	0.4
		100	87.1	11.6	1.2	0.1	93.4	5.7	0.7	0.2	89.3	9.1	1.4	0.3
		200	88.5	10.4	1.0	0.1	94.1	5.4	0.3	0.2	89.9	8.7	1.0	0.4
Model D: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t},\ h_{i,t}=0.0216+0.6896h_{i,t-1}+0.3174[\varepsilon_{i,t-1}-0.1108]^2,\ v_{i,t}\sim$ i.i.d. $N(0,1),\ i=1,...,p$														
		50	85.6	12.6	1.5	0.3	92.1	6.5	1.1	0.4	88.9	9.4	1.2	0.5
		100	85.0	13.3	1.5	0.2	93.1	5.9	0.8	0.2	86.9	11.3	1.4	0.4
		200	83.0	15.0	1.8	0.3	94.0	5.2	0.6	0.2	85.0	12.6	1.9	0.5
Model E: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t},\ h_{i,t}=0.005+0.7h_{i,t-1}+0.28[\lvert\varepsilon_{i,t-1}\rvert-0.23\varepsilon_{i,t-1}]^2,\ v_{i,t}\sim$ i.i.d. $N(0,1),\ i=1,...,p$														
		50	86.9	11.6	1.3	0.2	92.7	6.1	0.9	0.3	89.9	8.5	1.0	0.6
		100	87.0	11.7	1.2	0.2	93.7	5.5	0.6	0.1	88.8	9.7	1.2	0.3
		200	85.9	12.8	1.2	0.1	94.3	5.1	0.5	0.1	87.5	10.7	1.5	0.3
Model F: $\varepsilon_{i,t}=v_{i,t}\exp(h_{i,t}),\ h_{i,t}=\lambda h_{i,t-1}+0.5\xi_{i,t},\ (\xi_{i,t},v_{i,t})\sim$ i.i.d. $N(0,\mathrm{diag}(\sigma_\xi^2,1)),\ i=1,...,p$														
λ	σ_ξ													
0.936	0.424	50	75.5	20.7	3.4	0.4	90.9	7.5	1.3	0.3	80.9	15.9	2.6	0.6
		100	73.2	22.3	3.8	0.6	91.5	7.4	0.9	0.2	76.8	19.1	3.2	0.8
		200	72.7	23.6	3.3	0.4	92.4	6.7	0.7	0.2	75.4	21.4	2.5	0.7
0.951	0.314	50	80.0	17.2	2.4	0.4	91.8	6.7	1.2	0.3	83.7	13.6	2.1	0.6
		100	77.8	18.8	3.0	0.5	92.6	6.4	0.9	0.2	80.8	15.9	2.5	0.8
		200	77.2	20.0	2.4	0.4	93.3	6.1	0.5	0.1	79.5	17.6	2.3	0.6

Table 4: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p = 4$, True Rank is 0.

			Q-based					Q^b-based					Q^s-based				
		$r =$	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Model A: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}, v_{i,t} \sim i.i.d.\ N(0,1), i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	91.3	7.7	0.9	0.1	0.0	95.7	3.7	0.5	0.1	0.1	95.2	4.1	0.5	0.1	0.0
		100	92.7	6.4	0.7	0.2	0.0	95.0	4.2	0.6	0.2	0.0	94.9	4.3	0.6	0.2	0.0
		200	93.5	5.9	0.5	0.1	0.0	95.1	4.2	0.5	0.1	0.1	95.1	4.2	0.5	0.1	0.0
0.5	0.0	50	86.4	12.0	1.3	0.2	0.0	93.7	5.4	0.6	0.2	0.2	91.5	7.4	0.8	0.2	0.1
		100	88.8	10.0	1.0	0.2	0.0	93.7	5.5	0.4	0.2	0.1	91.7	7.1	0.7	0.3	0.1
		200	91.3	8.0	0.6	0.1	0.0	94.8	4.7	0.4	0.1	0.0	93.3	5.9	0.6	0.1	0.0
0.3	0.65	50	85.2	13.1	1.4	0.2	0.0	93.6	5.4	0.7	0.2	0.1	90.6	8.0	1.0	0.3	0.2
		100	86.3	12.0	1.5	0.2	0.0	93.8	5.3	0.7	0.2	0.1	89.3	9.2	1.3	0.2	0.1
		200	87.9	10.9	1.0	0.2	0.0	94.4	4.9	0.6	0.1	0.0	90.0	8.8	1.0	0.2	0.1
0.2	0.79	50	86.2	12.1	1.4	0.2	0.0	94.1	4.9	0.6	0.3	0.1	91.8	6.9	0.8	0.3	0.1
		100	86.9	11.5	1.3	0.2	0.0	93.8	5.3	0.8	0.1	0.0	90.1	8.3	1.2	0.3	0.1
		200	87.2	11.4	1.2	0.1	0.1	94.5	4.8	0.6	0.1	0.0	89.3	9.3	1.2	0.1	0.1
0.05	0.94	50	90.7	8.2	1.0	0.1	0.1	95.4	3.9	0.4	0.2	0.1	94.7	4.5	0.5	0.2	0.1
		100	91.9	7.3	0.7	0.1	0.0	94.8	4.5	0.4	0.3	0.1	94.2	4.9	0.6	0.2	0.0
		200	92.8	6.5	0.6	0.1	0.0	95.0	4.4	0.5	0.1	0.0	94.5	4.8	0.6	0.1	0.0
Model B: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}, v_{i,t} \sim i.i.d.\ t_5, i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	90.7	8.2	0.9	0.2	0.0	95.6	3.6	0.6	0.1	0.1	94.4	4.9	0.5	0.2	0.0
		100	93.3	5.8	0.6	0.2	0.1	95.9	3.4	0.5	0.1	0.1	95.0	4.4	0.4	0.2	0.1
		200	93.7	5.6	0.6	0.0	0.0	95.2	4.2	0.6	0.0	0.0	95.1	4.2	0.7	0.1	0.0
0.5	0.0	50	87.3	11.2	1.1	0.3	0.0	94.4	4.8	0.6	0.2	0.0	91.8	7.2	0.7	0.2	0.1
		100	90.5	8.4	0.9	0.1	0.1	94.9	4.3	0.6	0.1	0.1	93.1	5.9	0.7	0.2	0.1
		200	91.9	7.4	0.6	0.1	0.0	95.1	4.4	0.4	0.1	0.1	93.6	5.8	0.4	0.2	0.1
0.3	0.65	50	87.4	10.9	1.4	0.2	0.0	94.1	5.1	0.6	0.1	0.1	92.1	6.8	0.8	0.2	0.1
		100	89.6	9.2	0.9	0.1	0.1	94.5	4.6	0.6	0.1	0.1	92.3	6.5	0.8	0.2	0.1
		200	90.5	8.6	0.7	0.1	0.0	94.9	4.5	0.5	0.1	0.1	92.6	6.6	0.6	0.1	0.1
0.2	0.79	50	88.3	10.1	1.3	0.2	0.1	94.8	4.5	0.5	0.1	0.1	92.7	6.3	0.7	0.2	0.1
		100	90.3	8.6	0.9	0.1	0.1	94.7	4.5	0.6	0.1	0.1	92.7	6.1	0.8	0.2	0.2
		200	91.3	7.8	0.6	0.2	0.0	94.9	4.4	0.5	0.1	0.1	92.6	6.5	0.7	0.1	0.1
0.05	0.94	50	90.1	8.8	0.9	0.2	0.0	95.4	3.8	0.6	0.1	0.1	94.1	5.1	0.6	0.2	0.0
		100	92.6	6.5	0.7	0.2	0.1	95.4	3.9	0.5	0.1	0.1	94.6	4.5	0.6	0.2	0.1
		200	92.8	6.5	0.6	0.0	0.0	94.9	4.5	0.5	0.1	0.0	94.2	5.0	0.7	0.0	0.0
Model C: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, \ln(h_{i,t}) = -0.23 + 0.9 \ln(h_{i,t-1}) + 0.25[\lvert v_{i,t-1}^2 \rvert - 0.3 v_{i,t-1}], v_{i,t} \sim i.i.d.\ N(0,1), i = 1, ..., p$																	
		50	82.8	15.1	1.8	0.2	0.0	92.7	6.3	0.8	0.2	0.1	89.4	9.0	1.1	0.4	0.1
		100	85.4	12.8	1.6	0.2	0.0	93.7	5.3	0.7	0.2	0.1	88.9	9.4	1.2	0.3	0.1
		200	86.4	12.1	1.3	0.2	0.0	94.6	4.6	0.6	0.1	0.0	88.8	9.7	1.2	0.2	0.0
Model D: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t},\ h_{i,t} = 0.0216 + 0.6896 h_{i,t-1} + 0.3174[\varepsilon_{i,t-1} - 0.1108]^2,\ v_{i,t} \sim \text{i.i.d. } N(0,1),\ i = 1, ..., p$																	
		50	83.1	14.6	2.0	0.3	0.0	92.4	6.4	0.8	0.3	0.0	88.7	9.7	1.2	0.3	0.0
		100	83.4	14.1	2.1	0.3	0.1	93.1	5.9	0.8	0.1	0.1	86.9	10.9	1.8	0.4	0.1
		200	82.1	15.4	2.2	0.2	0.1	93.7	5.4	0.6	0.2	0.0	84.7	13.0	1.9	0.3	0.1
Model E: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t},\ h_{i,t} = 0.005 + 0.7 h_{i,t-1} + 0.28[\lvert\varepsilon_{i,t-1}\rvert - 0.23 \varepsilon_{i,t-1}]^2,\ v_{i,t} \sim \text{i.i.d. } N(0,1),\ i = 1, ..., p$																	
		50	84.5	13.6	1.7	0.2	0.0	93.4	5.6	0.8	0.2	0.0	89.9	8.7	1.1	0.2	0.1
		100	85.1	13.1	1.4	0.2	0.1	93.8	5.3	0.7	0.1	0.1	88.2	10.1	1.3	0.2	0.1
		200	84.0	14.1	1.7	0.1	0.0	94.4	5.0	0.4	0.2	0.0	86.4	11.8	1.6	0.2	0.0
Model F: $\varepsilon_{i,t} = v_{i,t} \exp(h_{i,t}),\ h_{i,t} = \lambda h_{i,t-1} + 0.5 \xi_{i,t},\ (\xi_{i,t}, v_{i,t}) \sim \text{i.i.d. } N(0, \text{diag}(\sigma_\xi^2, 1)),\ i = 1, ..., p$																	
λ	σ_ξ	T															
0.936	0.424	50	70.7	23.9	4.4	0.8	0.2	90.3	8.2	1.1	0.2	0.2	79.2	16.8	2.9	0.7	0.4
		100	67.8	26.0	5.2	0.8	0.2	91.3	7.5	0.9	0.2	0.1	73.7	21.3	4.0	0.7	0.3
		200	67.3	26.8	5.0	0.8	0.1	92.2	6.9	0.8	0.0	0.0	71.9	23.3	3.9	0.6	0.3
0.951	0.314	50	76.0	20.0	3.4	0.5	0.1	91.6	7.0	1.0	0.2	0.2	83.5	13.7	2.2	0.4	0.2
		100	74.6	20.8	3.9	0.4	0.2	92.1	6.7	1.0	0.1	0.1	79.2	16.9	3.1	0.5	0.3
		200	74.1	21.1	4.0	0.6	0.1	93.4	5.5	0.8	0.2	0.1	77.8	18.2	3.2	0.6	0.2

Table 5: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p = 5$, True Rank is 0.

			Q-based					Q^b-based					Q^s-based				
		$r =$	0	1	2	3	4, 5	0	1	2	3	4, 5	0	1	2	3	4, 5
Model A: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}, v_{i,t} \sim i.i.d.\ N(0,1),\ i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	87.9	10.6	1.2	0.3	0.1	96.3	3.1	0.4	0.2	0.0	95.2	4.1	0.6	0.1	0.0
		100	91.9	7.3	0.6	0.1	0.0	95.6	4.0	0.4	0.0	0.0	95.3	4.3	0.3	0.1	0.0
		200	93.1	6.1	0.7	0.1	0.0	95.6	3.7	0.6	0.1	0.0	95.3	4.0	0.5	0.1	0.0
0.5	0.0	50	82.4	15.3	2.0	0.3	0.0	93.8	5.4	0.6	0.2	0.0	90.9	8.3	0.8	0.1	0.0
		100	87.4	11.0	1.4	0.1	0.0	94.6	4.7	0.6	0.1	0.1	91.8	7.1	0.8	0.2	0.1
		200	90.4	8.8	0.8	0.1	0.0	95.7	3.9	0.4	0.0	0.0	93.2	6.0	0.6	0.1	0.0
0.3	0.65	50	81.9	15.3	2.5	0.2	0.1	93.8	5.2	0.9	0.1	0.1	90.5	8.1	1.2	0.1	0.0
		100	85.2	12.8	1.5	0.4	0.1	93.7	5.4	0.6	0.1	0.1	90.0	8.7	0.9	0.3	0.1
		200	86.1	12.3	1.4	0.2	0.0	94.3	5.1	0.6	0.1	0.0	89.1	9.5	1.2	0.2	0.1
0.2	0.79	50	83.8	13.8	2.1	0.3	0.1	94.5	4.5	0.8	0.2	0.0	92.3	6.4	1.0	0.3	0.1
		100	86.0	12.1	1.5	0.3	0.1	94.5	4.8	0.5	0.1	0.0	90.8	7.9	1.0	0.2	0.1
		200	86.5	11.8	1.5	0.2	0.0	94.4	4.8	0.7	0.1	0.0	89.4	9.0	1.3	0.3	0.0
0.05	0.94	50	87.7	10.8	1.2	0.3	0.1	95.8	3.5	0.5	0.2	0.0	95.1	4.1	0.6	0.2	0.0
		100	91.2	7.8	0.8	0.1	0.1	95.6	3.8	0.5	0.1	0.0	95.1	4.3	0.6	0.1	0.0
		200	92.1	7.1	0.7	0.1	0.0	95.1	4.3	0.5	0.1	0.0	94.5	4.9	0.6	0.1	0.0
Model B: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}, v_{i,t} \sim i.i.d.\ t_5,\ i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	87.2	11.4	1.1	0.2	0.1	96.4	3.2	0.2	0.1	0.1	95.0	4.5	0.3	0.1	0.1
		100	91.8	7.3	0.8	0.1	0.0	95.6	3.9	0.3	0.1	0.0	95.2	4.3	0.4	0.0	0.1
		200	93.5	5.9	0.5	0.0	0.0	96.2	3.4	0.3	0.0	0.0	95.6	4.0	0.3	0.0	0.1
0.5	0.0	50	84.1	13.9	1.7	0.2	0.1	95.2	4.0	0.6	0.1	0.1	92.6	6.6	0.7	0.1	0.1
		100	88.0	10.6	1.2	0.1	0.0	94.9	4.4	0.5	0.1	0.0	92.8	6.4	0.7	0.1	0.0
		200	91.4	7.8	0.7	0.1	0.0	95.7	3.8	0.4	0.1	0.0	93.9	5.4	0.6	0.1	0.0
0.3	0.65	50	84.2	13.8	1.7	0.2	0.1	95.1	4.0	0.6	0.1	0.1	92.8	6.0	1.0	0.1	0.1
		100	87.6	11.1	1.1	0.1	0.1	94.4	5.0	0.4	0.1	0.0	92.5	6.6	0.7	0.1	0.1
		200	89.8	9.3	1.0	0.0	0.0	95.3	4.2	0.4	0.1	0.0	92.8	6.4	0.7	0.1	0.0
0.2	0.79	50	85.5	12.5	1.6	0.3	0.1	95.3	3.9	0.6	0.1	0.1	93.6	5.3	0.9	0.1	0.1
		100	88.8	9.9	1.2	0.1	0.0	94.6	4.8	0.4	0.1	0.1	92.8	6.3	0.7	0.1	0.1
		200	90.3	8.7	1.0	0.1	0.0	95.5	4.1	0.4	0.1	0.0	93.4	6.0	0.6	0.1	0.0
0.05	0.94	50	86.9	11.5	1.3	0.2	0.1	96.1	3.4	0.3	0.1	0.1	94.8	4.7	0.4	0.1	0.0
		100	91.1	8.0	0.9	0.1	0.0	95.5	3.9	0.5	0.1	0.0	94.6	4.7	0.6	0.1	0.0
		200	92.8	6.6	0.6	0.0	0.0	96.0	3.7	0.3	0.1	0.0	95.2	4.3	0.4	0.1	0.0
Model C: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, \ln(h_{i,t}) = -0.23 + 0.9 \ln(h_{i,t-1}) + 0.25[\lvert v_{i,t-1}^2 \rvert - 0.3 v_{i,t-1}], v_{i,t} \sim i.i.d.\ N(0,1),\ i = 1, ..., p$																	
		50	80.0	16.9	2.7	0.3	0.1	92.9	5.9	0.9	0.1	0.1	89.9	8.7	1.1	0.2	0.1
		100	83.2	14.7	1.8	0.3	0.1	93.0	5.9	0.9	0.2	0.0	88.2	10.1	1.2	0.4	0.1
		200	85.3	13.0	1.6	0.1	0.0	94.3	5.0	0.6	0.1	0.0	88.8	9.9	1.2	0.2	0.0
Model D: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t},\ h_{i,t} = 0.0216 + 0.6896 h_{i,t-1} + 0.3174[\varepsilon_{i,t-1} - 0.1108]^2,\ v_{i,t} \sim$ i.i.d. $N(0,1),\ i = 1, ..., p$																	
		50	80.2	16.5	2.7	0.5	0.1	93.4	5.4	0.8	0.3	0.1	89.7	8.3	1.6	0.4	0.1
		100	81.3	15.9	2.4	0.4	0.1	93.5	5.6	0.7	0.2	0.0	86.8	11.1	1.6	0.4	0.1
		200	79.8	17.1	2.8	0.3	0.1	93.6	5.3	0.9	0.1	0.1	83.9	13.3	2.2	0.5	0.1
Model E: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t},\ h_{i,t} = 0.005 + 0.7 h_{i,t-1} + 0.28[\lvert \varepsilon_{i,t-1} \rvert - 0.23 \varepsilon_{i,t-1}]^2,\ v_{i,t} \sim$ i.i.d. $N(0,1),\ i = 1, ..., p$																	
		50	81.8	15.2	2.6	0.4	0.1	93.9	4.9	0.9	0.2	0.1	90.5	7.8	1.4	0.2	0.1
		100	83.9	13.9	1.9	0.3	0.1	94.2	5.1	0.6	0.1	0.1	88.6	9.8	1.3	0.2	0.1
		200	83.1	14.7	1.8	0.3	0.0	94.6	4.6	0.5	0.1	0.0	86.2	11.7	1.6	0.3	0.1
Model F: $\varepsilon_{i,t} = v_{i,t} \exp(h_{i,t}),\ h_{i,t} = \lambda h_{i,t-1} + 0.5 \xi_{i,t},\ (\xi_{i,t}, v_{i,t}) \sim$ i.i.d. $N(0, \mathrm{diag}(\sigma_\xi^2, 1)),\ i = 1, ..., p$																	
λ	σ_ξ	T															
0.936	0.424	50	65.0	26.4	6.9	1.3	0.3	89.0	8.9	1.6	0.4	0.1	77.8	17.0	3.9	1.0	0.3
		100	64.6	27.8	6.2	1.2	0.2	90.5	7.9	1.2	0.3	0.1	73.0	21.4	4.6	0.8	0.2
		200	62.9	28.7	7.0	1.3	0.1	92.1	6.6	1.1	0.2	0.0	69.2	23.9	5.6	1.1	0.2
0.951	0.314	50	71.9	22.0	5.0	1.0	0.2	91.0	7.4	1.2	0.2	0.1	82.7	13.8	2.7	0.6	0.2
		100	72.0	22.5	4.5	0.9	0.2	91.3	7.3	1.2	0.2	0.0	78.5	17.6	3.0	0.8	0.2
		200	69.5	24.8	4.8	0.8	0.1	92.3	6.7	0.8	0.2	0.0	75.1	20.4	3.5	0.9	0.1

Table 6: Size of Standard and Bootstrap PLR Tests for Rank = 0 Against Rank = p. True Rank is 1.

			$p=2$			$p=3$			$p=4$			$p=5$		
Model A: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
d_0	d_1	T	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s
0.0	0.0	50	5.2	5.7	4.9	5.2	4.7	4.1	6.2	4.0	3.7	6.1	2.6	2.8
		100	5.8	5.6	5.5	5.9	5.2	5.0	6.9	5.2	5.0	7.3	4.6	4.6
		200	5.5	5.6	5.0	5.6	5.0	4.9	5.3	4.5	4.5	6.8	4.6	4.6
0.5	0.0	50	6.3	5.8	5.9	7.6	5.4	6.0	9.0	5.1	5.7	9.0	3.6	4.5
		100	6.4	6.0	6.0	7.9	5.5	6.6	8.5	5.2	6.2	10.6	5.8	7.4
		200	5.1	4.9	5.0	7.3	5.4	6.5	7.8	5.1	6.1	8.8	4.8	6.5
0.3	0.65	50	6.8	5.9	6.2	7.7	5.5	6.0	9.6	5.1	6.2	9.9	4.0	5.0
		100	7.7	5.8	7.4	10.2	5.9	8.2	11.2	6.2	8.6	12.3	5.6	8.6
		200	7.5	5.5	7.4	9.4	5.1	8.7	10.5	5.0	8.7	12.7	5.4	9.8
0.2	0.79	50	6.5	5.8	6.1	7.8	5.6	6.1	8.8	4.9	5.2	9.6	4.0	4.4
		100	8.0	5.9	7.5	10.1	5.7	8.0	10.6	5.5	8.3	12.1	5.1	8.1
		200	7.9	5.4	7.8	10.4	6.0	9.2	11.2	5.3	9.3	12.6	5.4	10.2
0.05	0.94	50	5.2	5.5	5.0	5.8	4.7	4.4	6.2	4.1	3.7	6.5	2.7	2.9
		100	6.1	5.6	5.9	6.9	5.1	5.6	7.1	5.3	5.2	7.9	4.8	4.9
		200	5.8	5.7	5.7	6.8	5.1	5.7	6.6	4.6	5.4	7.4	4.7	5.4
Model B: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. t_5, $i=1,...,p$														
d_0	d_1	T	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s
0.0	0.0	50	5.1	6.0	4.6	6.5	5.4	4.7	6.7	4.3	3.9	7.3	2.7	3.3
		100	5.6	5.6	5.1	6.4	5.3	5.5	6.7	4.6	4.9	6.8	3.8	4.2
		200	4.9	4.6	4.3	5.6	4.6	4.6	6.4	4.6	4.8	6.6	4.1	4.7
0.5	0.0	50	6.2	6.3	5.6	7.8	5.6	6.0	7.8	4.5	5.0	9.9	3.6	4.4
		100	6.3	6.2	5.8	6.9	5.2	6.0	8.7	5.4	6.4	9.3	5.0	6.4
		200	5.5	4.8	4.9	6.5	5.2	5.5	7.4	4.7	5.9	7.9	4.7	5.6
0.3	0.65	50	6.5	6.4	6.2	7.8	5.7	6.4	8.2	4.5	5.2	9.7	3.4	4.7
		100	6.6	5.9	6.1	7.7	5.3	6.5	9.0	5.4	6.7	10.0	4.9	6.4
		200	5.9	4.9	5.4	7.2	5.0	6.1	8.3	4.9	6.8	9.0	4.8	6.6
0.2	0.79	50	6.4	6.3	5.9	7.5	5.6	6.0	7.6	4.4	5.0	9.0	3.3	4.3
		100	6.5	6.0	6.0	7.7	5.3	6.3	8.8	5.3	6.5	9.3	4.4	6.0
		200	5.8	4.5	5.6	7.4	4.9	6.3	8.4	5.0	6.9	9.0	4.7	6.5
0.05	0.94	50	5.7	6.1	5.0	6.7	5.1	5.0	6.4	4.0	4.0	7.7	3.0	3.5
		100	6.0	5.5	5.6	6.9	5.4	5.8	7.1	4.9	5.3	7.5	3.8	4.7
		200	5.5	4.8	5.0	6.2	4.8	5.0	6.9	5.0	5.6	7.3	4.3	5.2
Model C: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $\ln(h_{i,t})=-0.23+0.9\ln(h_{i,t-1})+0.25[\lvert v_{i,t-1}^2\rvert-0.3v_{i,t-1}]$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s
		50	7.0	5.9	6.5	8.2	5.8	6.5	10.5	5.4	7.0	11.0	4.0	5.2
		100	7.7	5.9	7.2	10.3	6.3	8.8	11.9	5.7	9.0	13.9	6.6	9.6
		200	7.6	5.6	7.2	9.7	5.9	8.6	11.3	5.8	9.6	13.3	6.1	10.5
Model D: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.0216+0.6896h_{i,t-1}+0.3174[\varepsilon_{i,t-1}-0.1108]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s
		50	7.8	6.2	7.2	9.3	5.6	7.5	11.0	5.5	7.3	12.1	4.5	6.2
		100	9.8	6.5	9.2	13.2	6.9	11.2	14.0	6.2	11.3	15.7	5.5	11.3
		200	10.5	5.9	10.2	14.1	6.3	12.6	15.6	5.7	13.4	17.9	5.9	15.1
Model E: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.005+0.7h_{i,t-1}+0.28[\lvert\varepsilon_{i,t-1}\rvert-0.23\varepsilon_{i,t-1}]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s
		50	7.5	6.1	6.9	8.9	5.7	6.9	10.6	5.1	6.3	11.0	4.2	5.8
		100	9.2	6.2	8.5	11.7	6.5	9.9	12.6	5.9	10.1	13.5	5.5	9.5
		200	9.3	5.8	9.1	12.4	6.1	11.5	14.0	5.9	11.9	14.8	5.6	12.2
Model F: $\varepsilon_{i,t}=v_{i,t}\exp(h_{i,t})$, $h_{i,t}=\lambda h_{i,t-1}+0.5\xi_{i,t}$, $(\xi_{i,t},v_{i,t})\sim$ i.i.d. $N(0,\mathrm{diag}(\sigma_\xi^2,1))$, $i=1,...,p$														
λ	σ_ξ	T	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s	Q_1	Q_1^b	Q_1^s
0.936	0.424	50	10.5	7.0	10.1	15.1	7.1	11.7	19.4	7.8	13.8	21.8	7.6	13.3
		100	12.4	6.5	11.6	19.6	7.6	16.9	24.0	8.2	19.6	28.0	8.8	20.8
		200	12.0	6.1	11.5	18.9	6.6	17.0	25.5	7.4	21.9	30.3	7.9	25.4
0.951	0.314	50	9.2	6.7	9.0	12.7	7.0	10.5	15.3	6.6	10.6	17.9	6.3	10.7
		100	11.3	6.8	10.7	16.9	7.2	14.4	19.8	7.5	16.5	22.7	7.4	16.4
		200	10.9	5.5	10.4	16.2	6.1	14.3	21.0	6.5	17.9	25.3	7.1	20.7

Table 7: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p = 2$, True Rank is 1.

			Q-based			Q^b-based			Q^s-based		
			$r=0$	$r=1$	$r=2$	$r=0$	$r=1$	$r=2$	$r=0$	$r=1$	$r=2$
Model A: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$											
d_0	d_1	T									
0.0	0.0	50	9.4	85.4	5.2	14.8	79.5	5.7	12.0	83.1	4.8
		100	0.0	94.2	5.8	0.0	94.4	5.6	0.0	94.5	5.5
		200	0.0	94.5	5.5	0.0	94.4	5.6	0.0	95.0	5.0
0.5	0.0	50	9.8	83.9	6.3	18.0	76.4	5.6	12.5	81.7	5.8
		100	0.0	93.6	6.4	0.3	93.8	6.0	0.0	94.0	6.0
		200	0.0	94.9	5.1	0.0	95.1	4.9	0.0	95.0	5.0
0.3	0.65	50	12.5	80.7	6.8	21.2	73.1	5.8	14.7	79.0	6.2
		100	0.2	92.0	7.7	1.4	92.8	5.8	0.3	92.3	7.4
		200	0.0	92.5	7.5	0.1	94.4	5.5	0.0	92.6	7.4
0.2	0.79	50	14.3	79.3	6.5	22.3	72.0	5.7	17.0	76.9	6.1
		100	0.3	91.7	8.0	1.8	92.4	5.9	0.3	92.2	7.5
		200	0.0	92.1	7.9	0.1	94.6	5.4	0.0	92.2	7.8
0.05	0.94	50	11.7	83.1	5.2	16.6	77.9	5.5	14.2	80.8	5.0
		100	0.0	93.9	6.1	0.2	94.2	5.6	0.0	94.0	5.9
		200	0.0	94.2	5.8	0.0	94.3	5.7	0.0	94.3	5.7
Model B: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}$, $v_{i,t} \sim$ i.i.d. t_5, $i = 1, ..., p$											
d_0	d_1	T									
0.0	0.0	50	10.2	84.6	5.1	16.1	78.1	5.9	13.5	81.9	4.6
		100	0.0	94.4	5.6	0.1	94.2	5.6	0.0	94.8	5.1
		200	0.0	95.1	4.9	0.0	95.4	4.6	0.0	95.7	4.3
0.5	0.0	50	10.4	83.4	6.2	17.9	75.9	6.2	13.7	80.8	5.6
		100	0.0	93.7	6.3	0.3	93.6	6.2	0.1	94.1	5.8
		200	0.0	94.5	5.5	0.0	95.2	4.8	0.0	95.1	4.9
0.3	0.65	50	11.3	82.2	6.5	19.1	74.6	6.3	14.7	79.1	6.2
		100	0.1	93.3	6.6	0.6	93.5	5.9	0.1	93.7	6.1
		200	0.0	94.1	5.9	0.0	95.1	4.9	0.0	94.6	5.4
0.2	0.79	50	12.0	81.7	6.4	19.0	74.8	6.2	15.2	78.9	5.9
		100	0.1	93.4	6.5	0.6	93.4	6.0	0.2	93.9	6.0
		200	0.0	94.2	5.8	0.0	95.5	4.5	0.0	94.4	5.6
0.05	0.94	50	11.4	82.9	5.7	17.3	76.7	6.0	14.4	80.7	5.0
		100	0.1	93.9	6.0	0.3	94.2	5.5	0.1	94.3	5.6
		200	0.0	94.5	5.5	0.0	95.2	4.8	0.0	95.0	5.0
Model C: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $\ln(h_{i,t}) = -0.23 + 0.9 \ln(h_{i,t-1}) + 0.25[\lvert v_{i,t-1}^2 \rvert - 0.3 v_{i,t-1}]$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$											
		50	12.2	80.8	7.0	21.7	72.5	5.8	14.8	78.6	6.5
		100	0.2	92.1	7.7	1.4	92.7	5.9	0.3	92.5	7.2
		200	0.0	92.4	7.6	0.1	94.3	5.6	0.0	92.8	7.2
Model D: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.0216 + 0.6896 h_{i,t-1} + 0.3174[\varepsilon_{i,t-1} - 0.1108]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$											
		50	14.7	77.5	7.8	24.5	69.7	5.9	17.3	75.5	7.2
		100	0.6	89.7	9.8	3.0	90.5	6.5	0.7	90.1	9.2
		200	0.0	89.5	10.5	0.4	93.7	5.9	0.0	89.7	10.2
Model E: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.005 + 0.7 h_{i,t-1} + 0.28[\lvert \varepsilon_{i,t-1} \rvert - 0.23 \varepsilon_{i,t-1}]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$											
		50	13.5	79.0	7.5	22.7	71.3	6.0	16.2	77.0	6.8
		100	0.5	90.3	9.2	2.5	91.3	6.2	0.5	91.0	8.5
		200	0.0	90.7	9.3	0.2	94.0	5.8	0.0	90.9	9.1
Model F: $\varepsilon_{i,t} = v_{i,t} \exp(h_{i,t})$, $h_{i,t} = \lambda h_{i,t-1} + 0.5 \xi_{i,t}$, $(\xi_{i,t}, v_{i,t}) \sim$ i.i.d. $N(0, \mathrm{diag}(\sigma_\xi^2, 1))$, $i = 1, ..., p$											
λ	σ_ξ	T									
0.936	0.424	50	16.0	73.5	10.5	29.5	64.5	6.0	19.3	70.8	9.9
		100	1.3	86.3	12.4	8.7	85.0	6.3	1.9	86.5	11.6
		200	0.0	88.0	12.0	1.2	92.7	6.0	0.0	88.5	11.5
0.951	0.314	50	15.9	74.9	9.2	27.8	66.2	6.0	18.9	72.3	8.8
		100	0.9	87.8	11.3	6.3	87.0	6.7	1.3	88.0	10.7
		200	0.0	89.1	10.9	0.7	93.8	5.5	0.0	89.6	10.4

Table 8: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p = 3$, True Rank is 1.

			Q-based				Q^b-based				Q^s-based			
		$r =$	0	1	2	3	0	1	2	3	0	1	2	3
Model A: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$														
d_0	d_1	T												
0.0	0.0	50	28.5	66.3	4.8	0.4	40.3	55.0	3.8	0.9	36.7	59.3	3.4	0.6
		100	0.2	93.9	5.5	0.5	0.4	94.3	4.5	0.7	0.3	94.7	4.4	0.6
		200	0.0	94.4	5.2	0.4	0.0	95.0	4.5	0.5	0.0	95.1	4.4	0.5
0.5	0.0	50	25.5	66.9	6.8	0.8	39.8	54.9	4.3	1.0	32.4	61.7	5.0	0.9
		100	0.8	91.3	7.3	0.6	2.1	92.3	4.7	0.8	1.1	92.4	5.7	0.9
		200	0.0	92.7	6.7	0.6	0.0	94.6	4.6	0.8	0.0	93.5	5.7	0.8
0.3	0.65	50	25.9	66.4	6.8	0.8	39.8	54.9	4.3	1.1	31.6	62.5	4.9	1.0
		100	1.8	88.0	9.1	1.1	5.1	89.0	5.0	1.0	2.4	89.4	6.9	1.3
		200	0.0	90.6	8.6	0.8	0.2	94.7	4.3	0.8	0.0	91.3	7.6	1.2
0.2	0.79	50	27.3	64.9	7.0	0.7	38.8	55.7	4.5	0.9	32.6	61.4	5.2	0.8
		100	2.5	87.4	9.0	1.1	6.1	88.3	4.7	0.9	3.1	89.0	6.6	1.3
		200	0.0	89.6	9.4	1.0	0.2	93.8	5.1	0.9	0.0	90.8	7.9	1.3
0.05	0.94	50	27.7	66.5	5.3	0.5	37.6	57.9	3.5	1.0	34.2	61.4	3.7	0.7
		100	0.8	92.3	6.3	0.6	1.9	93.1	4.3	0.8	1.6	92.9	4.8	0.8
		200	0.0	93.2	6.4	0.4	0.0	94.9	4.4	0.7	0.0	94.3	5.1	0.6
Model B: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}$, $v_{i,t} \sim$ i.i.d. t_5, $i = 1, ..., p$														
d_0	d_1	T												
0.0	0.0	50	28.6	64.9	5.9	0.6	40.9	53.8	4.1	1.2	36.8	58.5	3.8	0.9
		100	0.5	93.1	5.7	0.7	1.1	93.7	4.2	1.0	0.8	93.7	4.7	0.8
		200	0.0	94.4	5.1	0.5	0.0	95.4	4.0	0.7	0.0	95.4	4.1	0.5
0.5	0.0	50	26.6	65.6	7.2	0.5	40.7	53.9	4.3	1.0	34.7	59.3	5.1	0.9
		100	0.5	92.6	6.2	0.7	2.0	92.8	4.2	1.1	1.0	93.0	5.0	0.9
		200	0.0	93.5	6.0	0.5	0.0	94.8	4.7	0.4	0.0	94.5	5.0	0.5
0.3	0.65	50	27.3	64.8	7.1	0.7	41.0	53.5	4.4	1.1	34.3	59.4	5.2	1.2
		100	0.9	91.5	7.0	0.7	2.7	91.9	4.4	1.0	1.2	92.3	5.6	0.9
		200	0.0	92.8	6.6	0.6	0.0	95.0	4.3	0.7	0.0	93.9	5.3	0.8
0.2	0.79	50	27.9	64.6	6.8	0.7	40.3	54.4	4.4	1.0	34.9	59.2	4.8	1.1
		100	1.0	91.3	7.0	0.7	2.8	91.9	4.4	0.9	1.3	92.3	5.6	0.8
		200	0.0	92.6	6.9	0.5	0.0	95.1	4.2	0.6	0.0	93.7	5.7	0.6
0.05	0.94	50	28.5	64.8	6.1	0.6	40.5	54.5	4.0	1.0	36.1	58.9	4.0	0.9
		100	0.7	92.4	6.3	0.6	1.7	92.9	4.3	1.2	1.0	93.1	5.0	0.9
		200	0.0	93.8	5.7	0.5	0.0	95.2	4.2	0.6	0.0	95.0	4.4	0.6
Model C: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $\ln(h_{i,t}) = -0.23 + 0.9 \ln(h_{i,t-1}) + 0.25[\lvert v_{i,t-1}^2 \rvert - 0.3 v_{i,t-1}]$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$														
		50	25.0	66.8	7.2	1.0	39.8	54.5	4.5	1.1	31.3	62.2	5.4	1.1
		100	1.7	88.0	9.3	1.0	5.4	88.4	5.3	1.0	2.3	88.8	7.7	1.2
		200	0.0	90.3	9.0	0.7	0.3	93.8	5.1	0.8	0.0	91.4	7.5	1.1
Model D: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.0216 + 0.6896 h_{i,t-1} + 0.3174[\varepsilon_{i,t-1} - 0.1108]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$														
		50	26.6	64.1	8.4	1.0	40.1	54.4	4.6	0.8	31.6	61.0	6.3	1.1
		100	3.2	83.6	11.6	1.5	8.9	84.2	5.6	1.3	3.9	84.8	9.4	1.9
		200	0.0	85.9	12.5	1.6	0.7	93.0	5.1	1.1	0.0	87.3	10.8	1.9
Model E: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.005 + 0.7 h_{i,t-1} + 0.28[\lvert \varepsilon_{i,t-1} \rvert - 0.23 \varepsilon_{i,t-1}]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$														
		50	27.0	64.1	7.9	1.0	40.6	53.9	4.5	1.0	31.9	61.2	5.6	1.3
		100	2.8	85.5	10.4	1.3	7.9	85.7	5.5	1.0	3.4	86.6	8.2	1.7
		200	0.0	87.6	11.1	1.3	0.5	93.3	5.5	0.6	0.0	88.5	10.0	1.5
Model F: $\varepsilon_{i,t} = v_{i,t} \exp(h_{i,t})$, $h_{i,t} = \lambda h_{i,t-1} + 0.5 \xi_{i,t}$, $(\xi_{i,t}, v_{i,t}) \sim$ i.i.d. $N(0, \mathrm{diag}(\sigma_\xi^2, 1))$, $i = 1, ..., p$														
0.936	0.424	50	23.4	61.5	13.1	2.0	42.0	51.4	5.3	1.3	28.7	59.7	9.6	2.1
		100	3.7	76.7	17.0	2.6	18.2	74.5	6.2	1.1	5.2	77.9	14.1	2.8
		200	0.1	81.0	16.9	2.0	2.3	91.1	5.8	0.8	0.1	82.9	14.6	2.4
0.951	0.314	50	24.9	62.5	11.0	1.7	41.1	52.4	5.1	1.4	29.9	59.7	8.4	2.1
		100	3.5	79.6	14.7	2.2	14.7	78.3	5.6	1.4	4.5	81.2	11.8	2.6
		200	0.1	83.8	14.7	1.5	1.5	92.4	5.2	0.8	0.1	85.6	12.3	2.0

Table 9: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p = 4$, True Rank is 1.

			Q-based					Q^b-based					Q^s-based				
		$r =$	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Model A: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	42.5	51.2	5.5	0.7	0.1	60.3	35.7	3.1	0.6	0.3	56.1	40.2	3.0	0.4	0.2
		100	2.2	90.9	6.3	0.6	0.0	4.6	90.3	4.3	0.6	0.3	3.6	91.4	4.3	0.5	0.2
		200	0.0	94.7	4.8	0.4	0.0	0.0	95.5	3.9	0.5	0.1	0.0	95.5	4.0	0.4	0.1
0.5	0.0	50	36.6	54.3	8.0	0.8	0.3	58.0	37.1	4.1	0.5	0.3	48.5	45.9	4.8	0.5	0.4
		100	2.9	88.6	7.5	0.9	0.1	7.4	87.4	4.5	0.6	0.2	4.3	89.4	5.5	0.6	0.1
		200	0.0	92.2	7.2	0.5	0.1	0.0	94.8	4.6	0.4	0.1	0.0	93.9	5.4	0.5	0.1
0.3	0.65	50	34.5	55.9	8.4	0.8	0.3	54.6	40.4	4.0	0.5	0.4	45.5	48.3	5.3	0.6	0.3
		100	4.9	83.9	9.7	1.3	0.2	12.6	81.1	5.3	0.6	0.3	7.5	84.0	7.2	0.9	0.5
		200	0.0	89.5	9.2	1.1	0.2	0.4	94.6	4.3	0.5	0.2	0.0	91.2	7.5	1.0	0.2
0.2	0.79	50	35.7	55.5	7.5	1.0	0.3	51.8	43.4	3.9	0.5	0.3	45.2	49.7	4.2	0.6	0.3
		100	6.5	82.9	9.2	1.3	0.2	14.1	80.4	4.6	0.6	0.3	8.9	82.8	6.9	1.1	0.4
		200	0.0	88.8	9.9	1.0	0.2	0.4	94.3	4.5	0.6	0.3	0.0	90.7	7.9	1.0	0.4
0.05	0.94	50	39.5	54.3	5.2	0.8	0.1	56.0	39.9	3.3	0.5	0.3	51.0	45.4	2.9	0.4	0.3
		100	4.2	88.7	6.1	0.9	0.1	7.5	87.2	4.3	0.8	0.2	6.1	88.7	4.2	0.7	0.3
		200	0.0	93.4	5.9	0.7	0.0	0.0	95.4	4.0	0.5	0.2	0.0	94.6	4.6	0.5	0.2
Model B: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}$, $v_{i,t} \sim$ i.i.d. t_5, $i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	43.4	49.9	5.7	0.8	0.1	61.9	34.1	3.1	0.7	0.2	57.0	39.2	3.0	0.8	0.1
		100	2.7	90.6	5.9	0.6	0.1	5.9	89.5	3.9	0.5	0.3	4.3	90.8	4.1	0.5	0.2
		200	0.0	93.6	5.9	0.4	0.1	0.0	95.4	4.1	0.4	0.1	0.0	95.2	4.1	0.5	0.2
0.5	0.0	50	39.6	52.6	6.8	0.9	0.1	58.8	36.9	3.5	0.8	0.2	51.6	43.4	4.1	0.7	0.2
		100	3.0	88.3	7.9	0.7	0.1	8.1	86.5	4.7	0.5	0.3	4.9	88.7	5.6	0.6	0.2
		200	0.0	92.6	6.6	0.7	0.0	0.0	95.2	4.2	0.4	0.1	0.0	94.1	5.1	0.6	0.1
0.3	0.65	50	39.6	52.2	7.2	0.8	0.1	58.4	37.2	3.4	0.7	0.3	51.5	43.4	4.1	0.8	0.2
		100	3.7	87.2	7.9	0.9	0.2	9.7	84.9	4.5	0.6	0.3	6.0	87.3	5.8	0.7	0.3
		200	0.0	91.7	7.6	0.7	0.1	0.0	95.0	4.3	0.5	0.1	0.0	93.2	5.9	0.8	0.2
0.2	0.79	50	40.4	52.1	6.5	0.9	0.2	58.6	37.2	3.2	0.8	0.3	52.3	42.8	4.0	0.7	0.2
		100	4.4	86.8	7.5	1.0	0.2	10.0	84.7	4.4	0.7	0.3	6.6	86.9	5.4	0.7	0.3
		200	0.0	91.6	7.7	0.6	0.1	0.1	95.0	4.3	0.5	0.1	0.0	93.1	6.3	0.5	0.1
0.05	0.94	50	42.3	51.3	5.4	0.7	0.2	60.1	36.1	2.9	0.6	0.3	54.2	41.8	3.2	0.5	0.3
		100	3.9	89.1	6.2	0.7	0.1	7.9	87.2	4.0	0.6	0.3	5.7	88.9	4.5	0.5	0.3
		200	0.0	93.1	6.3	0.4	0.2	0.0	95.0	4.3	0.5	0.1	0.0	94.4	4.9	0.4	0.2
Model C: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $\ln(h_{i,t}) = -0.23 + 0.9 \ln(h_{i,t-1}) + 0.25[\lvert v_{i,t-1}^2 \rvert - 0.3 v_{i,t-1}]$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$																	
		50	33.9	55.7	9.0	1.2	0.2	54.2	40.6	4.1	0.7	0.3	44.4	48.6	5.8	0.8	0.4
		100	4.9	83.3	10.5	1.3	0.1	13.7	80.6	4.7	0.6	0.4	7.1	83.9	7.6	1.0	0.4
		200	0.0	88.7	10.1	1.0	0.2	0.5	93.7	5.0	0.6	0.3	0.0	90.4	8.4	1.0	0.3
Model D: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.0216 + 0.6896 h_{i,t-1} + 0.3174[\varepsilon_{i,t-1} - 0.1108]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$																	
		50	33.3	55.8	9.3	1.4	0.3	51.6	43.1	4.2	0.9	0.3	42.4	50.4	5.9	1.1	0.3
		100	7.4	78.6	12.0	1.7	0.3	18.0	75.8	5.0	0.9	0.4	9.7	79.0	9.3	1.3	0.6
		200	0.1	84.4	13.5	1.8	0.2	1.6	92.7	4.7	0.7	0.3	0.1	86.5	11.1	1.8	0.5
Model E: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.005 + 0.7 h_{i,t-1} + 0.28[\lvert \varepsilon_{i,t-1} \rvert - 0.23 \varepsilon_{i,t-1}]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$																	
		50	34.5	54.9	9.2	1.1	0.3	52.4	42.6	3.9	0.7	0.3	43.4	50.3	5.3	0.7	0.3
		100	6.7	80.7	10.7	1.5	0.4	16.9	77.2	4.7	0.9	0.3	9.4	80.4	8.2	1.3	0.6
		200	0.1	86.0	12.3	1.4	0.3	0.9	93.2	5.0	0.5	0.3	0.1	88.0	10.2	1.3	0.4
Model F: $\varepsilon_{i,t} = v_{i,t} \exp(h_{i,t})$, $h_{i,t} = \lambda h_{i,t-1} + 0.5 \xi_{i,t}$, $(\xi_{i,t}, v_{i,t}) \sim$ i.i.d. $N(0, \mathrm{diag}(\sigma_\xi^2, 1))$, $i = 1, ..., p$																	
λ	σ_ξ	T															
0.936	0.424	50	27.9	52.7	16.1	3.0	0.4	49.6	42.9	6.3	0.9	0.3	36.0	50.2	10.9	2.2	0.6
		100	7.9	68.1	20.3	3.2	0.4	26.7	65.4	6.4	1.2	0.3	10.9	69.5	16.2	2.7	0.7
		200	0.2	74.2	22.0	2.9	0.6	4.9	87.8	6.5	0.6	0.2	0.4	77.7	18.5	2.6	0.8
0.951	0.314	50	30.0	54.7	12.7	2.2	0.4	49.9	43.7	5.3	0.8	0.3	38.5	51.0	8.4	1.6	0.6
		100	7.8	72.4	17.1	2.4	0.4	23.9	68.7	6.2	0.9	0.3	10.5	73.0	14.0	1.7	0.8
		200	0.2	78.9	18.1	2.5	0.4	3.0	90.5	5.8	0.6	0.1	0.2	81.9	15.2	2.0	0.7

Table 10: Standard and Bootstrap Sequential Procedures for Selecting the Co-integration Rank. $p = 5$, True Rank is 1.

			Q-based					Q^b-based					Q^s-based				
		$r =$	0	1	2	3	4, 5	0	1	2	3	4, 5	0	1	2	3	4, 5
Model A: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}, v_{i,t} \sim i.i.d.\ N(0,1), i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	51.5	42.4	5.3	0.6	0.1	75.4	22.1	2.1	0.2	0.2	71.0	26.3	2.3	0.3	0.1
		100	7.2	85.5	6.6	0.6	0.1	15.5	80.0	3.8	0.5	0.2	12.9	82.5	4.0	0.5	0.1
		200	0.0	93.2	6.0	0.7	0.1	0.0	95.4	4.0	0.4	0.2	0.0	95.4	4.0	0.5	0.1
0.5	0.0	50	44.7	46.3	7.9	0.9	0.2	71.4	25.0	3.1	0.3	0.2	63.1	32.4	3.8	0.4	0.3
		100	7.7	81.7	9.4	1.0	0.2	19.9	74.3	5.0	0.6	0.2	12.8	79.8	6.5	0.7	0.2
		200	0.0	91.2	8.0	0.7	0.1	0.0	95.2	4.3	0.3	0.1	0.0	93.5	5.8	0.5	0.1
0.3	0.65	50	41.5	48.6	8.7	1.1	0.1	65.6	30.4	3.4	0.5	0.1	57.4	37.7	4.2	0.6	0.1
		100	10.3	77.5	10.5	1.6	0.2	24.6	69.9	4.6	0.7	0.2	15.5	75.8	7.3	1.1	0.3
		200	0.0	87.2	11.5	1.1	0.2	1.0	93.6	4.7	0.6	0.1	0.1	90.1	8.5	1.1	0.2
0.2	0.79	50	41.8	48.5	8.4	1.0	0.3	64.0	32.0	3.3	0.5	0.2	56.2	39.3	3.7	0.5	0.2
		100	12.2	75.7	10.5	1.3	0.3	25.4	69.5	4.2	0.7	0.2	17.6	74.4	6.8	1.0	0.3
		200	0.2	87.2	11.1	1.4	0.2	1.3	93.3	4.8	0.4	0.2	0.3	89.5	8.7	1.2	0.3
0.05	0.94	50	47.6	45.8	5.6	0.7	0.2	70.6	26.7	2.1	0.4	0.2	64.9	32.3	2.4	0.3	0.1
		100	10.4	81.7	7.0	0.7	0.1	19.3	75.9	4.0	0.6	0.1	16.1	79.1	4.2	0.5	0.1
		200	0.0	92.6	6.7	0.6	0.1	0.1	95.2	4.2	0.4	0.1	0.0	94.6	4.8	0.4	0.2
Model B: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, h_{i,t} = \omega + d_0 \varepsilon_{i,t-1}^2 + d_1 h_{i,t-1}, v_{i,t} \sim i.i.d.\ t_5, i = 1, ..., p$																	
d_0	d_1	T															
0.0	0.0	50	49.8	42.9	6.4	0.8	0.1	74.6	22.7	2.3	0.3	0.1	69.1	27.7	2.8	0.3	0.1
		100	7.8	85.4	6.1	0.6	0.1	16.6	79.6	3.5	0.2	0.1	13.5	82.3	3.8	0.4	0.1
		200	0.0	93.4	6.0	0.6	0.1	0.0	95.9	3.8	0.3	0.1	0.0	95.3	4.2	0.4	0.1
0.5	0.0	50	43.6	46.5	8.9	0.7	0.2	71.1	25.4	3.0	0.3	0.2	63.1	32.6	3.9	0.3	0.2
		100	8.1	82.7	8.2	0.9	0.2	19.5	75.5	4.4	0.5	0.2	13.4	80.2	5.7	0.5	0.2
		200	0.0	92.0	7.4	0.5	0.1	0.2	95.1	4.3	0.3	0.1	0.0	94.4	5.1	0.3	0.2
0.3	0.65	50	43.1	47.2	8.8	0.8	0.1	70.6	26.0	2.8	0.3	0.2	61.8	33.6	4.2	0.4	0.1
		100	8.7	81.3	9.0	0.9	0.2	20.0	75.1	4.4	0.4	0.1	14.1	79.5	5.7	0.6	0.2
		200	0.0	91.0	8.1	0.7	0.2	0.3	94.9	4.2	0.5	0.1	0.0	93.3	5.8	0.7	0.1
0.2	0.79	50	44.0	47.0	8.0	0.8	0.1	70.0	26.8	2.7	0.4	0.1	62.4	33.4	3.7	0.4	0.1
		100	8.8	81.9	8.2	1.0	0.1	20.1	75.5	3.9	0.3	0.1	14.8	79.2	5.3	0.6	0.2
		200	0.0	91.0	8.2	0.7	0.1	0.2	95.0	4.2	0.5	0.1	0.0	93.4	5.7	0.6	0.2
0.05	0.94	50	47.8	44.5	6.9	0.7	0.1	72.6	24.4	2.5	0.3	0.2	66.3	30.3	2.8	0.4	0.1
		100	9.0	83.5	6.8	0.5	0.1	17.9	78.2	3.4	0.3	0.1	14.9	80.4	4.2	0.4	0.1
		200	0.0	92.7	6.6	0.7	0.1	0.1	95.6	3.9	0.4	0.0	0.0	94.8	4.5	0.5	0.1
Model C: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}, \ln(h_{i,t}) = -0.23 + 0.9 \ln(h_{i,t-1}) + 0.25[\lvert v_{i,t-1}^2 \rvert - 0.3 v_{i,t-1}], v_{i,t} \sim i.i.d.\ N(0,1), i = 1, ..., p$																	
		50	39.4	49.6	9.6	1.2	0.2	65.0	31.0	3.3	0.5	0.2	55.5	39.4	4.3	0.6	0.2
		100	9.2	76.9	12.1	1.7	0.1	23.9	69.6	5.6	0.8	0.2	14.4	76.0	8.3	1.1	0.3
		200	0.1	86.6	11.9	1.2	0.2	0.8	93.2	5.3	0.6	0.2	0.1	89.4	9.2	0.9	0.4
Model D: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.0216 + 0.6896 h_{i,t-1} + 0.3174[\varepsilon_{i,t-1} - 0.1108]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$																	
		50	38.8	49.1	10.2	1.6	0.3	61.0	34.5	3.4	0.8	0.3	52.7	41.2	5.0	0.9	0.3
		100	12.1	72.2	13.6	1.8	0.3	28.0	66.5	4.7	0.6	0.1	17.0	71.7	9.9	1.0	0.3
		200	0.2	81.9	15.2	2.4	0.3	3.3	90.8	4.9	0.7	0.3	0.4	84.5	12.5	1.9	0.7
Model E: $\varepsilon_{i,t} = h_{i,t}^{1/2} v_{i,t}$, $h_{i,t} = 0.005 + 0.7 h_{i,t-1} + 0.28[\lvert \varepsilon_{i,t-1} \rvert - 0.23 \varepsilon_{i,t-1}]^2$, $v_{i,t} \sim$ i.i.d. $N(0,1)$, $i = 1, ..., p$																	
		50	39.9	49.1	9.4	1.2	0.4	62.7	33.2	3.4	0.5	0.2	53.9	40.3	4.8	0.7	0.3
		100	11.4	75.1	11.7	1.6	0.3	26.3	68.2	4.7	0.6	0.2	16.2	74.3	8.3	0.9	0.3
		200	0.2	85.0	12.7	1.8	0.3	2.2	92.2	4.6	0.7	0.2	0.3	87.5	10.4	1.4	0.5
Model F: $\varepsilon_{i,t} = v_{i,t} \exp(h_{i,t})$, $h_{i,t} = \lambda h_{i,t-1} + 0.5 \xi_{i,t}$, $(\xi_{i,t}, v_{i,t}) \sim$ i.i.d. $N(0, \text{diag}(\sigma_\xi^2, 1))$, $i = 1, ..., p$																	
λ	σ_ξ	T															
0.936	0.424	50	28.8	49.4	16.9	3.9	1.0	55.7	37.2	5.6	1.2	0.3	41.0	45.8	10.3	2.2	0.7
		100	11.8	60.2	22.7	4.7	0.6	34.3	57.1	7.2	1.2	0.2	16.3	62.9	17.0	2.9	0.8
		200	0.5	69.2	25.2	4.5	0.7	8.6	83.5	6.7	0.9	0.3	0.8	73.8	20.8	3.7	0.9
0.951	0.314	50	32.8	49.3	14.3	2.9	0.6	57.6	36.4	4.6	1.1	0.4	45.4	43.9	8.3	1.8	0.5
		100	12.3	65.0	18.9	3.2	0.6	32.6	60.1	6.2	0.9	0.2	17.2	66.4	13.4	2.5	0.4
		200	0.3	74.4	21.7	2.8	0.7	6.1	86.9	5.8	0.9	0.4	0.8	78.5	17.5	2.5	0.7

Table 11: Size of Standard and Bootstrap PLR Tests for Rank = 0 Against Rank = p. True Rank is 0. VAR(2) Case.

			$p=2$			$p=3$			$p=4$			$p=5$		
Model A: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
d_0	d_1	T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
0.0	0.0	50	12.2	7.3	6.8	21.4	6.5	6.9	41.5	8.5	8.9	70.3	9.7	11.5
		100	8.9	6.0	6.2	12.5	5.6	5.9	18.9	5.2	5.9	32.0	6.5	7.4
		200	7.0	4.9	5.3	8.5	5.2	5.6	10.9	4.8	5.1	15.8	5.2	5.3
0.3	0.65	50	16.3	8.0	10.1	27.0	8.5	11.0	46.2	9.4	11.2	72.2	12.0	14.6
		100	12.9	6.7	9.4	17.3	7.1	10.1	25.4	7.0	10.6	38.1	8.6	12.5
		200	10.5	5.9	8.8	13.4	5.8	9.4	16.1	5.7	9.4	22.9	6.3	10.8
Model B: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=\omega+d_0\varepsilon_{i,t-1}^2+d_1h_{i,t-1}$, $v_{i,t}\sim$ i.i.d. t_5, $i=1,...,p$														
d_0	d_1	T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
0.0	0.0	50	12.3	6.2	6.7	22.0	6.9	7.6	41.3	7.7	8.7	70.2	10.2	11.1
		100	8.4	5.0	5.7	12.2	5.5	6.1	18.7	5.8	6.8	32.3	5.7	6.5
		200	7.3	5.4	5.6	8.7	5.1	5.5	10.9	5.2	5.8	16.8	5.9	6.4
0.3	0.65	50	14.3	7.2	8.0	24.6	8.0	9.6	44.7	8.6	10.5	72.4	11.0	12.6
		100	10.7	5.6	7.6	14.2	6.1	7.8	22.2	6.3	8.4	35.4	6.6	9.3
		200	9.0	5.9	7.4	11.0	5.9	7.5	13.5	5.2	7.3	19.8	6.1	8.1
Model C: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $\ln(h_{i,t})=-0.23+0.9\ln(h_{i,t-1})+0.25[\lvert v_{i,t-1}^2\rvert-0.3v_{i,t-1}]$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
		50	16.4	8.3	10.3	27.7	8.9	11.4	47.4	10.4	12.5	73.0	12.5	15.6
		100	13.3	6.7	9.1	18.0	7.0	10.6	25.5	7.6	11.4	39.9	8.5	12.6
		200	10.8	5.9	8.6	13.2	6.0	9.2	16.7	6.0	9.4	22.5	6.5	10.2
Model D: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.0216+0.6896h_{i,t-1}+0.3174[\varepsilon_{i,t-1}-0.1108]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
		50	17.9	8.9	11.8	29.8	9.8	12.3	47.6	10.6	13.5	73.1	12.7	15.9
		100	16.0	7.3	12.4	21.0	7.8	12.9	29.7	8.0	13.8	42.8	9.3	14.8
		200	15.0	6.2	12.5	18.8	6.7	14.4	22.9	6.6	14.6	29.7	7.4	16.2
Model E: $\varepsilon_{i,t}=h_{i,t}^{1/2}v_{i,t}$, $h_{i,t}=0.005+0.7h_{i,t-1}+0.28[\lvert\varepsilon_{i,t-1}\rvert-0.23\varepsilon_{i,t-1}]^2$, $v_{i,t}\sim$ i.i.d. $N(0,1)$, $i=1,...,p$														
		T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
		50	17.0	8.3	10.8	28.5	9.1	11.5	45.9	10.3	13.1	72.5	12.2	15.7
		100	14.4	7.0	10.8	19.9	7.4	12.2	27.4	7.4	12.4	40.9	8.2	13.3
		200	12.9	6.4	11.1	16.6	6.5	12.6	20.0	6.2	12.5	26.5	6.2	13.0
Model F: $\varepsilon_{i,t}=v_{i,t}\exp(h_{i,t})$, $h_{i,t}=\lambda h_{i,t-1}+0.5\xi_{i,t}$, $(\xi_{i,t},v_{i,t})\sim$ i.i.d. $N(0,\mathrm{diag}(\sigma_\xi^2,1))$, $i=1,...,p$														
λ	σ_ξ	T	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s	Q_0	Q_0^b	Q_0^s
0.951	0.314	50	21.7	8.8	14.1	33.0	10.6	16.1	52.2	13.4	19.8	76.0	16.5	22.7
		100	19.9	8.1	15.8	26.9	8.7	18.0	36.6	9.5	19.2	49.6	11.7	21.3
		200	16.9	5.9	13.8	22.6	6.7	16.9	30.2	7.3	20.7	37.2	8.1	21.4

Table 12: Standard and Bootstrap Co-integration Tests: UK, Japan, Canada and the U.S.

Country		Q_r Statistics					Asymptotic p-values					Wild Bootstrap p-values					I.I.D. Bootstrap p-values				
UK	$r =$	0	1	2	3		0	1	2	3		0	1	2	3		0	1	2	3	
		154.88	67.83	10.65	0.98		0.00	0.00	0.58	0.95		0.00	0.00	0.76	0.98		0.00	0.00	0.60	0.95	
Japan	$r =$	0	1	2	3		0	1	2	3		0	1	2	3		0	1	2	3	
		101.86	40.19	10.50	3.68		0.00	0.01	0.59	0.46		0.00	0.20	0.86	0.75		0.00	0.04	0.71	0.51	
Canada	$r =$	0	1	2	3		0	1	2	3		0	1	2	3		0	1	2	3	
		248.50	74.65	15.84	6.11		0.00	0.00	0.18	0.18		0.00	0.00	0.33	0.31		0.00	0.00	0.20	0.26	
USA	$r =$	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
		138.66	60.04	33.32	17.47	3.15	0.00	0.01	0.08	0.12	0.55	0.02	0.36	0.51	0.62	0.90	0.00	0.01	0.12	0.15	0.62

Research Papers 2009

2009-8: Torben B. Rasmussen: Jump Testing and the Speed of Market Adjustment

2009-9: Dennis Kristensen and Andrew Ang: Testing Conditional Factor Models

2009-10: José Fajardo and Ernesto Mordecki: Skewness Premium with Lévy Processes

2009-11: Lasse Bork: Estimating US Monetary Policy Shocks Using a Factor-Augmented Vector Autoregression: An EM Algorithm Approach

2009-12: Konstantinos Fokianos, Anders Rahbek and Dag Tjøstheim: Poisson Autoregression

2009-13: Peter Reinhard Hansen and Guillaume Horel: Quadratic Variation by Markov Chains

2009-14: Dennis Kristensen and Antonio Mele: Adding and Subtracting Black-Scholes: A New Approach to Approximating Derivative Prices in Continuous Time Models

2009-15: Charlotte Christiansen, Angelo Ranaldo and Paul Söderllind: The Time-Varying Systematic Risk of Carry Trade Strategies

2009-16: Ingmar Nolte and Valeri Voev: Least Squares Inference on Integrated Volatility and the Relationship between Efficient Prices and Noise

2009-17: Tom Engsted: Statistical vs. Economic Significance in Economics and Econometrics: Further comments on McCloskey & Ziliak

2009-18: Anders Bredahl Kock: Forecasting with Universal Approximators and a Learning Algorithm

2009-19: Søren Johansen and Anders Rygh Swensen: On a numerical and graphical technique for evaluating some models involving rational expectations

2009-20: Almut E. D. Veraart and Luitgard A. M. Veraart: Stochastic volatility and stochastic leverage

2009-21: Ole E. Barndorff-Nielsen, José Manuel Corcuera and Mark Podolskij: Multipower Variation for Brownian Semistationary Processes

2009-22: Giuseppe Cavaliere, Anders Rahbek and A.M.Robert Taylor: Co-integration Rank Testing under Conditional Heteroskedasticity