G. Cavaliere, L. Fanelli, P. Paruolo

Determining the number of cointegrating relations under rank constraints

2001/17

UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

http://eco.uninsubria.it

Printed in Italy in July 2001
Università degli Studi dell'Insubria
Via Ravasi, 2, 21100 Varese, Italy

Tests for cointegration rank and choice of the alternative

Giuseppe Cavaliere*, Luca Fanelli*, Paolo Paruolo†

August 31, 2005

Abstract

This paper discusses likelihood-ratio (LR) tests on the cointegrating (CI) rank which consider any possible dimension of the CI rank under the alternative. The trace test and lambda-max test are obtained as special cases. Limit quantiles for all the tests in the class are derived. It is found that any of these tests can be used to construct an estimator of the CI rank, with no differences in asymptotic properties when the alternative is fixed.

The properties of the class of tests are investigated by local asymptotic analysis, a simulation study and two empirical illustrations. It is found that all the tests in the class have comparable power, which deteriorates substantially as the number of random walks increases. For one dimensional alternatives sufficiently far from the null, the trace tests is dominated by other tests in the class; this is in line with expectations based on the results of Andrews (1996, Econometrica) for LR tests in a stationary regression setup, when alternatives are one-sided.

The tests in the class can also be arranged to give a constrained estimator of the CI rank, that restricts the minimum number of common trends. We find that mis-specification of the minimum number of common trends implies that the correct CI rank is selected with 0 limit probability. As a consequence, no value of the CI rank should be left untested, unless it can be excluded beyond any reasonable doubt.

Keywords: Cointegration rank, Likelihood ratio, Asymptotic power, Unit roots, Brownian motion.

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

1 Introduction

Likelihood ratio tests for the determination of cointegration rank in VAR processes are used extensively in applications. These tests were derived in Johansen (1988,

*Dipartimento di Scienze Statistiche, Università di Bologna, Italy.

†Corresponding Author. Department of Economics, University of Insubria, Via Monte Generoso 71, 21100 Varese, Italy. Email: paolo.paruolo@uninsubria.it; tel +39 0332 395500, fax +39 0332 395509.

1991, 1996) and admit many variants depending on the treatment of the intercept and the linear trend, see Lütkepohl, Saikkonen and Trenkler (2001) for a review. For any choice of deterministic components, two LR statistics are usually considered, the so-called trace test and the lambda-max test.

The trace test compares a given CI rank with the dimension of the process, where a CI rank equal to the dimension implies that the process is stationary. The lambda-max test, instead, compares a given CI rank with the closest larger integer value. In this paper we discuss the class of LR tests obtained comparing a given CI rank with any greater integer value. This class includes the trace test and lambda-max test as special cases, as well as some new statistics.

We derive appropriate asymptotic critical values for all the tests in the class. The procedure for CI rank determination in Johansen (1992), based on the trace test, is here re-considered using any of these tests as a building block. It is shown that the asymptotic properties of the resulting CI rank estimator are unchanged. This finding extends a similar result of Paruolo (2001a) for the lambda-max statistics.

The tests within the class are compared through local asymptotic analysis and a finite-sample simulation study. The local asymptotics of the trace test and lambda-max test have been compared in Paruolo (2001a) for the case of no deterministics and in Lütkepohl, Saikkonen and Trenkler (2001) for various choices of deterministic terms. In this paper, while extending the comparison to all the members of the class, we calculate the local asymptotic power with a number of unit root processes up to 6, instead of 3 as in the previous references.

We find that the power of all the tests deteriorates significantly with the increase of the number of unit roots. When comparing the powers of the tests in the class, it is found that the local powers do not differ dramatically. However, for local alternatives sufficiently far from the null and lying in a smaller dimensional subspace, the power of the trace test is dominated by the one of other members of the class.

This observation is in line with analystical results for the LR test in a classical regression setup for one-sided alternatives, as described in Andrews (1996). He showed that the LR test maximizes a weighted average of the power over the alternative when the alternative is sufficiently far from the null. This behavior is similar to the results presented here for the local power, although the theory developed in Andrews (1996) does not cover the present non-stationary case.

Any test in the class can also be arranged to give a constrained estimator of the CI rank, which restricts the minimum number of common trends. Several authors have entertained the idea that some values for the CI rank may be excluded on the basis of prior information. For instance, Saikkonen and Lütkepohl (2000, eq. (2.3) and following lines) argue that a CI rank equal to the dimension of the process (i.e. one that would imply that the process is stationary) "can often be ruled out on the basis of prior information on the data and variables".[1]

We investigate the properties of the constrained estimator of CI rank. If the restrictions on the minimum number of common trends are true, the estimator has the same asymptotic properties as the unconstrained estimator for fixed alternatives. On the other hand, under mis-specification, we show that the constrained estimator

[1] Horvath and Watson (1995) and Paruolo (2001b) consider CI rank determiantion under restrictions on the cointegration space instead than its dimension.

selects the (incorrect) entertained maximum CI rank with limit probability 1. We hence advise not to use the constrained estimator, unless there is no reasonable doubt about the minimum number of common trends.

A Monte Carlo analysis is performed in order to evaluate how well the asymptotics matches small-sample behavior the tests. In the Monte Carlo design, the I(1) data generating processes (DGP) are selected in order to verify (i) the validity of the local asymptotics (ii) the validity of the fixed alternative asymptotics (iii) the performance of the CI rank estimators.

It is found that, in most cases, sample sizes close to 100 are large enough to find a close resemblance of the local and fixed asymptotics with the finite sample results. For the same sample size, the sampling performance of the CI rank estimator is found to be reasonable, when the DGP has local asymptotic parameters far enough from the null. However, when the stationary autoregressive roots are close to unity, one finds that all the tests are still oversized for sample size close to 200.

The paper includes also two empirical illustrations; we reconsider a quarterly money demand system for the Euro area estimated in Brand and Cassola (2004), and an annual per-capita consumption model relative to four Italian regions considered in Cavaliere et al. (2005). As shown in these illustrations, all the LR tests can be computed from standard output of CI software. In both illustrations one finds similar inferences on the CI rank for the unconstrained estimator based on the trace test and the constrained estimator which assumes at least one common trend, even though the latter gives better support to the predictions of the underlying economic theory.

The rest of the paper is organized as follows. Section 2 illustrates the model and the statistical analysis. The asymptotics under the null are reported in Section 3; these are used to derive the properties of the CI rank estimators in Section 4. Section 5 derives the asymptotic local power which is calculated by simulation in Section 6. Section 7 reports the Monte Carlo simulation of the finite sample properties of the tests, while Section 8 contains the empirical illustrations. Section 9 concludes. All proofs are placed in the Appendix.

In the following "$\rightarrow^w$" denotes weak convergence and "$\rightarrow^p$" convergence in probability; "$x := y$" indicates that x is defined by y. $\mathrm{col}(a)$ is the space spanned by the columns of the matrix a.

2 Cointegration rank tests

In this section we present the various LR tests considered in the paper. We follow the notation of Johansen (1996) and Paruolo (2001a). Consider the following cointegrated $I(1)$ VAR process in equilibrium correction form:

$$\Delta X_t = \alpha\beta' X_{t-1} + \Psi U_t + \mu D_t + \varepsilon_t \tag{1}$$

where X_t and ε_t are $p \times 1$, $U_t := (\Delta X_{t-1}, ..., \Delta X_{t-q+1})'$ is $p(q-1) \times 1$, D_t is a vector of deterministic terms, ε_t is p-variate i.i.d. Gaussian, $\varepsilon_t \sim N(0, \Omega)$ and α and β are full column $p \times r$ matrices, $r \leq p$.

The process (1) is assumed to satisfy Granger's representation theorem, see Johansen (1996), Theorem 4.2. Specifically we label as '$I(1, r)$ condition' the following three conditions: (a) all the characteristic roots associated with (1) be outside the unit circle or equal to 1; (b) the AR impact matrix has rank r, i.e. it can be written as $-\alpha\beta'$ for α and β are full column $p \times r$ matrices; (c) $\det \alpha'_{\perp}\Gamma\beta_{\perp} \neq 0$, with $\Gamma := I_p - \Psi\left(i_{q-1} \otimes I_p\right)$, i_{q-1} and I_p being respectively a $(q-1) \times 1$ vector of ones and an identity matrix of dimension p.

Eq. (1) for unknown parameters α, β, Ψ, μ, Ω, when α and β are $p \times j$ matrices not necessarily of full rank and Ω is positive definite, denotes a model. We assume that the deterministic part can be partitioned into $D_t := (D'_{1t} : D'_{2t})'$ and $\mu := (\mu_1 : \mu_2)$ where $\mu_1 = \alpha\rho'_1$ is the part of the coefficient of the deterministic terms that is constrained to be in $\operatorname{col}(\alpha)$. We indicate the corresponding model as $H(j)$, which can be put in the format

$$Z_{0t} = \alpha\beta^{*\prime} Z_{1t} + \mu_2 Z_{2t} + \varepsilon_t \tag{2}$$

with $Z_{0t} := \Delta X_t$, $Z_{1t} := (X'_{t-1} : D'_{1t})'$, $Z_{2t} := (U'_t : D'_{2t})'$, $\beta^* := (\beta' : \rho'_1)'$. If D_{it} is set equal to 0, it is understood that D_{it} is dropped from the definition of Z_{it}, $i = 1, 2$.

Given a sample $\{X_t\}_{t=1}^{T}$, let $\ell(j)$ indicate the maximized Gaussian log-likelihood of model $H(j)$; the LR test of $H(j)$ within $H(s)$, for $j < s$ can be written as (see Johansen, 1996)

$$LR(j, s) := -2(\ell(j) - \ell(s)) = -T \sum_{i=j+1}^{s} \log(1 - \widehat{\lambda}_i). \tag{3}$$

where $\widehat{\lambda}_i$ is the i-th largest solution of the eigenvalue problem

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

and where $S_{ij} := M_{ij.2} := M_{ij} - M_{i2} M_{22}^{-1} M_{2j}$, $M_{ij} := T^{-1} \sum_{t=1}^{T} Z_{it} Z'_{jt}$.

Setting $s = j+1$ in (3) one obtains the lambda-max test, ${}^{\max}Q(j) := LR(j, j+1)$, whereas setting $s = p$ one obtains the trace statistic ${}^{\mathrm{tr}}Q(j) := LR(j, p)$. Values of s different from $j+1$ and p give other test statistics which form the class of interest of the paper. Consider the test of $H(j)$ against $H(p-m)$, where m can be interpreted as the assumed minimum number of common trends. The corresponding LR test statistic is given by

$${}^{m}Q(j) := LR(j, p-m) = -T \sum_{i=j+1}^{p-m} \log(1 - \widehat{\lambda}_i). \tag{4}$$

With this notation ${}^{\mathrm{tr}}Q(j) = {}^{0}Q(j)$ and ${}^{\max}Q(j) = {}^{p-j-1}Q(j)$. It is simple to express the ${}^{m}Q(j)$ statistic (4) as a function of ${}^{\mathrm{tr}}Q(j)$:

$${}^{m}Q(j) = LR(j, p) - LR(p-m, p) = {}^{\mathrm{tr}}Q(j) - {}^{\mathrm{tr}}Q(p-m). \tag{5}$$

Alternatively, ${}^{m}Q(j)$ can be written as a function of the lambda-max statistics

$${}^{m}Q(j) = \sum_{i=j}^{p-m-1} LR(i, i+1) = \sum_{i=j}^{p-m-1} {}^{\max}Q(i). \tag{6}$$

Any of the formulas (4), (5), (6) can be used to calculate the ${}^{m}Q(j)$ statistic from standard output of CI software, which typically reports (possibly a subset of) the 3 sets of values $\{\widehat{\lambda}_i\}$, $\{{}^{\mathrm{tr}}Q(i)\}$, $\{{}^{\max}Q(i)\}$.

The limit distribution of statistics (4) under the null given by the $I(1,r)$ condition, $r < p - m$, is derived in the following section.

3 Asymptotics under the null

In this section we derive asymptotics under the null. Let $eig_i\,(A)$ indicate the i-th largest eigenvalue of a real square symmetric matrix A. Let $B(u)$ indicate a standard Brownian motion of dimension $p-r$, decomposed as $B(u) := (B_1(u)', B_2(u))'$, where $B_2(u)$ is one-dimensional. In the following all integrals are from 0 to 1 and we use the notation $(a|b) := a(u) - \int a(s)b(s)'ds(\int b(s)b(s)'ds)^{-1}b(u)$. Moreover

$$N(F,B) := \int (dB)\, F' \left(\int FF'\right)^{-1} \int F\,(dB)',$$

where here and hereafter the argument u in processes like $B(u)$ is usually omitted for brevity.

Theorem 1 *Let the $I(1,r)$ condition hold, $r < p - m$; then the asymptotic distribution of ${}^{m}Q(r) := LR\,(r, p-m)$ is given by*

$$ {}^{m}Q(r) := LR\,(r, p-m) \xrightarrow{w} Z_m := \sum_{i=1}^{p-r-m} eig_i\,(N(F,B)) \qquad (7)$$

where F depends on B and on the deterministic term as described in Johansen (1996)*, Theorems* 6.1 *and* 6.2*. Specifically*

1. *if $D_t = 0$ in (1) and $D_{1t} = D_{2t} = 0$ in (2), then $F := B$;*

2. *if $\mu D_t = \mu_1 \in \mathrm{col}\,(\alpha)$ in (1), and $D_{1t} = 1$, $D_{2t} = 0$ in (2), then $F := (B', 1)'$;*

3. *if $\mu D_t = \mu_1 \notin \mathrm{col}\,(\alpha)$ in (1), and $D_{1t} = 0$, $D_{2t} = 1$ in (2), then $F := (B_1', u|1)'$;*

4. *if $\mu D_t = \mu_1 + \mu_2 t$, $\mu_2 \in \mathrm{col}\,(\alpha)$ in (1), and $D_{1t} = t$, $D_{2t} = 1$ in (2), then $F := (B', u|1)'$;*

5. *if $\mu D_t = \mu_1 + \mu_2 t$, $\mu_2 \notin \mathrm{col}\,(\alpha)$ in (1), and $D_{1t} = 0$, $D_{2t} = (1 : t)'$ in (2), then $F := (B_1', u^2|u, 1)'$.*

Moreover if $j < r$, ${}^{m}Q(j) := LR\,(j, p-m) = O_p(T)$.

Theorem 1 shows that the asymptotic distribution of (4) under the null depends both on $p - r$ and m, where $p - r$ is the number of common stochastic trends of the system, and m is the number of assumed common trends under the alternative. The limit distribution Z_m differ from the ones already tabulated e.g. in Johansen

(1996), and need to be simulated. This is performed in Section 6 along with the estimation of asymptotic local power.

It can be noted that the results in parts 3, 4, 5 of Theorem 1 do not change if seasonal dummies are included both in D_t in DGP (1) and in D_{2t} in model (2), provided the dummies are orthogonalized with respect to the constant, i.e. are of the type $d_{it} := 1(t \bmod s = i) - s^{-1}$, where s is the number of seasons and in $1(\cdot)$ is the indicator function.

Observe that for $m = 0$ the limit distribution (7) reduces to the sum of all the eigenvalues of N, i.e. one obtains the trace of N, which is the usual form of the limit distribution of the trace test. For $m = p - r - 1$, then the sum in the limit distribution (7) involves just the largest eigenvalue of N, which is the usual form of the limit distribution of the lambda-max test.

Before analyzing the local power properties of the tests, we consider the properties of the procedure for CI rank determination when the trace test is substituted with any of the ${}^mQ(j)$ tests.

4 Properties of CI rank estimators

The present section derives the properties of two CI rank estimators based on the ${}^mQ(j)$ tests. The first estimators, indicated by $\widetilde{r}_s$, simply uses ${}^mQ(j)$ tests in place of the trace test in the procedure proposed by Johansen (1992). The second estimator, indicated as $\widehat{r}_m$, uses the ${}^mQ(j)$ tests to build a constrained estimator, that cannot exceed a pre-specified value $p - m$, where m is the constraint on the minimum number of common trends.

Specifically, let $\widetilde{r}_s$ be defined as the smallest integer j for which ${}^mQ(j)$ does not reject when m is chosen as $m = \max(p - s - j, 0)$ and s is fixed; the ${}^mQ(j)$ tests in this sequence sum s consecutive values of $-T \ln(1 - \widehat{\lambda}_i)$, where $\widehat{\lambda}_i$ are defined equal to 0 for $i > p$. When $s = 1$, the building block ${}^mQ(j)$ is the lambda-max test, and $\widetilde{r}_1$ is the selection procedure based on the lambda-max test considered in Paruolo (2001a), denoted there by ${}^{\max}_{h}\widehat{r}$.

Formally, $\widetilde{r}_s$ is chosen according to the rule

$$\widetilde{r}_s := \min_{0 \le j < p} \{ j : {}^mQ(j) \le c_\varphi(p - j, m), \quad m = \max(p - s - j, 0) \} \tag{8}$$

where $c_\varphi(p - j, m)$ indicates the $1 - \varphi$ quantile of the limit distribution (7) and $\min \varnothing := p$. Note that $0 \le j < p$ and no constraint is placed on the minimum number of the common trends.

We next consider the constrained estimator that restricts the minimum number of common trends m, that is fixed in the sequence of tests. We define $\widehat{r}_m$ as the smallest integer j for which ${}^mQ(j)$ does not reject, where $0 \le j < p - m$. Formally, $\widehat{r}_m$ is chosen according to the rule

$$\widehat{r}_m := \min_{0 \le j < p - m} \{ j : {}^mQ(j) \le c_\varphi(p - j, m) \} \tag{9}$$

where $c_\varphi(p - j, m)$ indicates the $1 - \varphi$ quantile of the limit distribution (7), $\varphi \in (0, 1)$, and $\min \varnothing := p - m$. Note that this selection procedure does not consider values for the common trends inferior to m.

It is easy to prove that Propositions 1 and 2 in Paruolo (2001a) also hold for $\widehat{r}_m$ and $\widetilde{r}_s$. Here we just report the analogue to Proposition 1 for space constraints.

Theorem 2 *Let the $I(1,r)$ assumption hold and $T \to \infty$. Then* $\Pr(\widetilde{r}_s < r) \to 0$, $\Pr(\widetilde{r}_s = r) \to 1 - \varphi$. *Moreover if $r < p - m$ then* $\Pr(\widehat{r}_m < r) \to 0$, $\Pr(\widehat{r}_m = r) \to 1 - \varphi$.

The key property in establishing Theorem 2 is that ${}^{m}Q(j)$ diverges if $j < r$. Since $\widehat{\lambda}_1, ..., \widehat{\lambda}_r$ are $O_p(1)$, one has ${}^{\max}Q(j) := -T \ln(1 - \widehat{\lambda}_{j+1}) = O_p(T)$ for $j < r$. This is the same reason that ensures that ${}^{m}Q(j) = \sum_{i=j}^{p-m-1} {}^{\max}Q(i) = O_p(T)$ in the same circumstances, see (6). We hence observe that all the tests ${}^{m}Q(j)$ can be arranged to give the same asymptotic properties for $\widetilde{r}_s$.

We next discuss behavior of the constrained estimator $\widehat{r}_m$ under mis-specification.

Theorem 3 *Let the $I(1,r)$ assumption hold and let $r > p-m$, i.e. $p-r < m$; then as T diverges* $\Pr(\widehat{r}_m = p - m) \to 1$.

Theorem 3 highlights the dangers of incorrectly imposing at least m common trends when indeed fewer common trends drive the system. In this case all the tests in the sequence $\widehat{r}_m$ would reject in the limit, and the CI rank is estimated equal to its (assumed) maximum value $p - m$ with probability 1. We hence advise not to impose a minimum number of common trends without testing, unless the untested CI rank can be excluded beyond any reasonable doubt.

5 Asymptotics under a local alternative

In this section the local power of ${}^{m}Q(j)$ tests are analyzed. The local power of the trace test has been investigated by Johansen (1991b), and the one of the lambda-max test in Paruolo (2001a) for the case no deterministic terms, $D_t = 0$ in (1) and (2). Lütkepohl, Saikkonen and Trenkler (2001) considered the power of the trace and lambda-max tests in the presence of deterministic terms.

We analyze the asymptotic power of the tests under the local alternative that the product $\alpha\beta'$ is replaced by

$$\alpha\beta' + T^{-1}\alpha_1\beta_1' \tag{10}$$

where α, β are $p \times r$ full column rank matrices and α_1, β_1 are $p \times 1$ vectors, chosen not to lie in $\mathrm{col}(\alpha)$ and $\mathrm{col}(\beta)$ respectively. For simplicity, the local alternative (10) is chosen to be one-dimensional (i.e. α_1 and β_1 are chosen to be vectors), and we limit attention to the case of no deterministic terms.

Let $f := f_1$ and $g := f_2$ be defined as follows

$$\begin{aligned} f &:= f_1 := \beta_1' C \alpha_1 \\ g^2 &:= f_2^2 := \left(\alpha_1'\alpha_\perp \left(\alpha_\perp'\Omega\alpha_\perp\right)^{-1} \alpha_\perp'\alpha_1\right)\left(\beta_1' C \Omega C' \beta_1\right) - f^2, \end{aligned} \tag{11}$$

where C is the moving average impact matrix, $C = \beta_{\perp}\left(\alpha'_{\perp}\Gamma\beta_{\perp}\right)^{-1}\alpha'_{\perp}$. Let also B be a standard $(p-r)$-variate Brownian motion partitioned as $B = (B_1, B_2, B_3)$, with B_3 of dimension $(p-r-2) \times 1$; moreover let K be the diffusion $K := (K_1, K_2, B_3)'$ where $K_i(s) := B_i(s) + f_i \int_0^s K_1(u)\,du$, $i = 1, 2$.[2]

Theorem 4 *If $|eig_i(I_r + \beta'\alpha)| < 1$ and $r < p - m$, then the asymptotic distribution of ${}^mQ(r)$ is given by*

$$ {}^mQ(r) := LR(r, p-m) \xrightarrow{w} V_m := \sum_{i=1}^{p-r-m} eig_i\left(N(K, K)\right). $$

Observe that the distribution of V_m depends on f, g, $p-r$ and m. The quantiles of the asymptotic distribution under the null from Theorem 1 and the local power derived from Theorem 4 are calculated in the following section.

6 Quantiles simulation

Asymptotic quantiles of the distributions of Z_m and V_m in Theorems 1 and 4 may be calculated by Monte Carlo simulation. The asymptotic critical values, i.e. the 95% quantiles of Z_m, were simulated by substituting the limiting Brownian process with a random walk over $T = 2500$ segments and using 10^5 replications. 95% quantiles for the 5 different specifications of the deterministic component D_t of Theorem 1 are reported in the Tables 1-5.[3] Calculations were performed with Gauss versions 3.2.

[Tables 1-5 approximately here]

In order to compute the asymptotic local powers, we simulated the distribution of V_m using discretization of the Ornstein-Uhlenbeck process K as in Johansen (1996) Chapter 15. Again we chose $T = 2500$ segments and performed 10^5 replications.

Values of f and g which define the local alternative where selected in order to allow comparison with previous simulation results reported in Johansen (1991b) and Paruolo (2001a). The selected values of f and g were the points in the set $\mathbb{F}_0 \times \mathbb{G}_0$, where $\mathbb{F}_0 := \{0, -3, -6, -9, -12, -15, -18, -21, -24\}$, and $\mathbb{G}_0 := \{0, 6, 12, 18, 24\}$, while values of $p-r$ were chosen between 2 and 6. The values of ($p-r$, f, g) considered here extend the ones used in Johansen (1991b) and Paruolo (2001a), Lütkepohl, Saikkonen and Trenkler (2001), who considered $p - r \leq 3$. The selected significance level was 5%.

In order to calculate simulation standard errors for rejection frequencies taking the estimation of the critical values into account, we used the results in Paruolo (2002). The obtained standard errors are considerably larger than the ones that do not account for the variability induced by the estimation of critical values.

[2] Note that K_1 is an Ornstein-Uhlenbeck (OU) process, see e.g. Karatzas and Shreve (1988), p. 358, which is mean-reverting for $f < 0$. K_2 is given by the sum of a Brownian motion and an integrated OU process; therefore for f close to 0 and $g > 0$, K_2 behaves like an integrated Brownian motion.

[3] 50%, 75%, 80%, 90%, 97.5% and 99% quantiles can be downloaded from the internet address http://www2.stat.unibo.it/cavaliere/cvalues/cvalues.html for $p - r = 1, ..., 11$ and all possible values of m.

We also followed the discussion of MC design in Paruolo (2002) in order to reduce the MC standard errors for the power differences between the ${}^{m}Q$ tests and the ${}^{\mathrm{tr}}Q$. In particular the same realizations of the random walks were used to simulate the asymptotic distribution of all tests (i) both under the null hypothesis and (ii) under the alternative. This technique allows to reduce the MC variance for the difference of the powers of the ${}^{m}Q$ and the ${}^{\mathrm{tr}}Q$ tests; the obtained variance reduction was around 3 fold.

[Fig. 1 – 4, Table 6 approximately here]

Most of the results are reported graphically for brevity.[4] In the following we indicate by $\widehat{\pi}_m$ the rejection frequency of test ${}^{m}Q$. Minimum and maximum values across f, g, m of the power difference $100\left(\widehat{\pi}_m - \widehat{\pi}_{\mathrm{tr}}\right)$ are given in Table 6 for different values of $p-r$; the power differences $100\left(\widehat{\pi}_m - \widehat{\pi}_{\mathrm{tr}}\right)$ and the power function $\widehat{\pi}_{\mathrm{tr}}$ are also reported in Fig. 1-4. In these graphs insignificant differences $\widehat{\pi}_m - \widehat{\pi}_{\mathrm{tr}}$ are indicated by empty circles.

We collect observations about the power in the following remarks.

(a). The power of all the tests deteriorates with the increase of the number of unit roots. The power for $(f, g) = (-24, 24)$ is over 90% for $p - r = 3$ for all the tests in the class, while at the same value of (f, g) and $p - r = 6$ it is only around 50%. This ‘curse of dimensionality’ is already documented in Johansen (1991b) for $p-r = 1, 2, 3$; however the actual extent of its magnitude at higher values of $p - r$ is novel.

(b). When comparing the powers of the tests in the class, it is found that the local power does not differ dramatically. Table 6 in fact shows that the minimum and maximum values of $100\left(\widehat{\pi}_m - \widehat{\pi}_{\mathrm{tr}}\right)$ are -5.3% and 9% respectively.

(c). For local alternatives sufficiently far from the null, the power of the trace test is dominated by the ones of the other members of the class. This observation is in line of the theoretical results for the LR tests in classical situations, as described in Andrews (1996). He considered the case where the parameter space is restricted under the alternative, such as in (multivariate) one-sided tests. This situation is similar to the present local asymptotics, where the local alternative, indexed by (f, g) is one-sided.[5]

Andrews (1996) showed that the LR test maximizes a weighted average of the power over the alternative when the alternative is sufficiently far from the null. A LR test which averages over a larger region of the alternative should present worse power with respect to another LR test which averages across a smaller region. Because the alternative is one-dimensional (α_1 and β_1 are vectors), the ${}^{\max}Q$ test is expected to maximize the power for alternatives sufficiently far from the null. This appears to be the case in Fig. 1-4 for high values of (f, g).

[4] The complete set of results is available on the authors' webpages.

[5] Obviously the one-sided nature of the alternative is preserved also when considering the bijective transformation in Lütkepohl, Saikkonen and Trenkler (2001) of (f, g) into (l, d), where $l := f^2 + g^2$, $d := -\left(f^2/\left(f^2+g^2\right)\right)^{1/2}$.

7 Finite sample properties

In this section we present a finite-sample simulation study. This study is undertaken in order to ascertain how well the asymptotics (both under the null and under a local alternative) can describe finite samples properties of the ${}^{m}Q$ tests.

Similarly to previous simulations of the finite sample power, see Lütkepohl, Saikkonen and Trenkler (2001) and reference therein, we consider a VAR(1) as DGP. Unlike previous studies, however, we let dimension p vary from 3 to 7. For simplicity we consider the case of no deterministics, i.e. $D_{1t} = D_{2t} = 0$. Similarly to Nielsen (2004), we choose $T \in \mathbb{T} := \{24, 48, 96, 192\}$.

We apply the invariance properties described in Paruolo (2005) in order to reduce the VAR(1) design dimension. By invariance the design can be reduced to the form

$$\left\{ \begin{array}{c} X_{1t} = JX_{1t-1} + \upsilon_{1t} \\ \Delta X_{2t} = \kappa X_{1t-1} + \upsilon_{2t} \\ \Delta X_{3t} = \upsilon_{3t} \end{array} \right. , \qquad \mathrm{var} \begin{pmatrix} \upsilon_{1t} \\ \upsilon_{2t} \\ \upsilon_{3t} \end{pmatrix} = \begin{pmatrix} \Phi & 0 & 0 \\ 0 & I_j & 0 \\ 0 & 0 & I_{p-r-j} \end{pmatrix} \tag{12}$$

where X_{1t} is $r \times 1$, X_{2t} is $j \times 1$, $j \leq r$ and X_{3t} is $(p - r - j) \times 1$, and the covariance matrix var (υ_t) has been partitioned conformably. The matrix J is a Jordan matrix of dimension r, κ is upper triangular, with positive elements on the main diagonal; j can vary from 0 to r. Here we implicitly assume $r < p - r$, which is the case for the designs used in the simulations. Finally Φ is a covariance matrix, where first element on the main diagonal is equal to 1.

Let DGP_r indicate the group of designs with a specified CI rank r. The following 3 cases are considered:

(a). DGP_0 : The group contains one single design where $\Delta X_t = \upsilon_t$, var $(\upsilon_t) = I_p$.

(b). DGP_1 : Here $\Phi = 1$ and hence var $(\upsilon_t) = I_p$. The designs are indexed by two scalar parameters, J and κ. The choice of values for J and κ and the MC results are discussed in Subsection 7.1. This class contains 42 designs.

(c). DGP_2 : This class is indexed by J (of dimension 2×2), κ (of dimension $2 \times j$, $j = 0, 1, 2$) and a by correlation coefficient τ in Φ. The design and the results for this class are discussed in Subsection 7.2. This class contains 297 designs.

Therefore a total of 340 designs was considered, which cover all cases in DGP_0 and DGP_1, and a large part of the possible cases in DGP_2. For each design we performed 10^4 replications. We only considered the nominal 5% level for all the tests. Calculations were performed in Gauss, versions 3.6 and 6.0. Not all the simulation results are reported here for space constraints; the complete set is available on the authors' webpages.

The results of the single design in DGP_0 showed that for $T = 192$ empirical sizes for all the ${}^{m}Q(0)$ tests are close to the nominal 5% level. This simple case is not discussed further. The following two subsections describe the designs and the results for the classes DGP_1 and DGP_2.

7.1 CI rank equal to 1

We choose J in the stationary region $-1 < J < 1$. In order to mimic the local asymptotic analysis, we set $J = 1 + f/T$ for some values of $f \in \mathbb{F} := \mathbb{F}_0 \backslash \{0\} := \{-3, -6..., -24\}$, where the value 0 is excluded because this would violate the stationarity requirement $-1 < J < 1$ and change the CI rank. Table 7 reports some of the chosen values of J, where bold entries in the table indicate that the corresponding value of f is in $\mathbb{F}$. The left panel can be read along the columns for the asymptotics under the null, and along the major diagonal for the local asymptotics.

Bold entries in the table corresponds to values in the set $\mathbb{F}$. The values $f = -3$ and $f = -24$ can be traced in 3 designs for increasing T ($f = -3$ corresponds to $T = 24, 48, 96$, $f = -24$ to $T = 48, 96, 192$), while the values $f = -6$ and $f = -12$ can be traced in all 4 designs for increasing T.

[Table 7 here]

We also add the values $-\frac{1}{2}$ and 0 as possible values for J; the resulting set of design values for J are $\mathbb{J} := \{J_i\} := \left\{-\frac{1}{2}, 0, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \frac{31}{32}\right\}$, for a total of 7 entries. In the following we partition $\mathbb{J} = \mathbb{J}_1 \cup \mathbb{J}_2$ where $\mathbb{J}_1 := \left\{-\frac{1}{2}, 0, \frac{1}{2}\right\}$ contains values of J far from unity and $\mathbb{J}_2 := \left\{\frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \frac{31}{32}\right\}$ collects values close to unity.

κ is also a (nonnegative) scalar. We set $\kappa = g/T$ for some values of $g \in \mathbb{G}_0 := \{0, 6, 12, 18, 24\}$. We choose the values of κ in the heading of the right panel of Table 7, where entries in the panel indicate the corresponding value of g. Bold entries in the right panel correspond to values g in the set $\mathbb{G}_0$. Hence the values $g = 3$ and $g = 24$ can be traced in 3 designs for increasing T ($g = 3$ corresponds to $T = 24, 48, 96$, $g = 24$ to $T = 48, 96, 192$), while the values $g = 6$ and $g = 12$ can be traced in all 4 designs for increasing T. The resulting values of κ are collected in the set $\mathbb{K} := \{\kappa_j\} := \left\{0, \frac{1}{32}, \frac{1}{16}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}\right\}$, for a total of 6 entries. The class DGP_1 has thus a total of $6 \cdot 7 = 42$ designs, each one corresponding to $(J, \kappa) = (J_i, \kappa_j)$ for $i = 1, ..., 7$, $j = 1, ..., 6$. We indicate the corresponding DGP as $\mathrm{DGP}_1(i, j)$, where the ordering of i and j is the natural one in $\mathbb{J}$ and $\mathbb{K}$.

A selection of results concerning asymptotics under the null are presented in Tables 8–10. Table 8 reports results for $\mathrm{DGP}_1(3, 1)$, with $J = 1/2$, $\kappa = 0$, for $p = 5$. The entries of the table are percentage rejection frequencies of the tests ${}^mQ(i)$, $i = 0, 1$, and $m = 0, 1, 2$. These values of i are the relevant ones for the CI rank estimators of Section 4.

Table 8 represents a typical outcome for designs with $J \in \mathbb{J}_1$. It can be seen that for $T = 192$ all 3 tests ${}^mQ(0)$ reject with almost probability 1, and that all ${}^mQ(1)$ tests are undersized for small samples, while the size converges to the nominal level at $T = 192$. The results in this table hence are in line with predictions from the asymptotics under the null.

One can observe a moderate improvement in using ${}^1Q$, ${}^2Q$ or ${}^3Q$ in place of the trace test ${}^0Q$. At $T = 96$ the power of ${}^mQ(0)$ improves monotonically from 73% to 81% when going from $m = 0$ to $m = 2$, while the size of all the ${}^mQ(1)$ tests are close to the nominal level of 5%.

[Table 8 here]

A different situation occurs when J is local to unity, i.e. when $J \in \mathbb{J}_2$. In this case the stationary root is close to 1, and hence one can consider the process as having CI rank 0 with 1 autoregressive root close to 1, in the sense of the local asymptotic hypothesis, see eq. (10). Specifically in this case one has $\alpha = \beta = 0$, $\alpha_1 = (T(J-1), T\kappa, 0_{p-2})'$, $\beta_1 = (1, 0_{p-1})$ and $C = \Omega = I_p$. Substituting into (11) one finds $f = T(J-1)$ and $g = T\kappa$, as planned in the choice of design.

[Tables 9 and 10 here]

Tables 9 and 10 present results for $J = 31/32$ and $\kappa = 0$ or $\kappa = 1/2$ respectively. For ease of exposition the tables report also the values of f and g implied by the values and J, κ. In Table 9 $\kappa = g = 0$; notice that all ${}^mQ(0)$ reject the null very infrequently, $m = 0, 1, 2, 3$. This would make the CI rank estimators select $r = 0$ with probability approximately 80%, 90%, 94%, 95% for all the 4 tests as T grows from 28 to 192.

Note also that the calculation of the power function in Section 6 shows that the asymptotic local power at $p - r = 5$, $f = -6$, $g = 0$ is about 5% for all the 4 tests ${}^mQ(0)$. This matches the finite sample frequency of rejections, which approaches 5% for all the 4 tests when T grows to 192. Note that increasing T will eventually make f diverge to $-\infty$, thus approaching the limit distribution under the null I(1,1) reported in the last line.

The reason why this convergence is very slow in this case is associated to the fact that the asymptotic local power function increases very slowly in f for $g = 0$. This is also apparent comparing Tables 9 and 10, where in the latter $\kappa = 1/2$ and hence $g = T\kappa$ is also increasing with T. In this case the 4 tests ${}^mQ(0)$ reject with frequency that approaches 1 when $T = 192$. The price of a root close to unity is given by the fact that the tests ${}^mQ(1)$ are still oversized at $T = 192$.

We next consider wether the local asymptotic analysis of Section 5 is able to capture the relevant features of the finite sample results simulations. In Table 11 we collect the frequency of rejection of the trace test ${}^0Q(0)$ for the various combinations of coefficients described in Table 7 that give rise to values of $f \in \mathbb{F}$ and $g \in \mathbb{G}_0$; here $p = 3$. We also report the limit asymptotic power calculated in Section 6 under the label $T = \infty$.

It can be seen that the local asymptotic distribution is able to capture the main features of the finite sample properties at $T = 96$. Table 12 reports the same analysis for the test ${}^2Q$. Again the main features of the finite sample behavior are captured by the local asymptotics. Note that ${}^2Q$ is only marginally more powerful than ${}^0Q$ for large values of f, g.

The analysis of Tables 11- 12 (and of similar ones not reported for brevity) shows the usefulness of the local asymptotics for predicting the small sample behavior of the tests. It is found that differences in power are modest, as predicted by the local asymptotics.

[Tables 11, 12 and 13 here.]

Table 13 reports the behavior of the trace test when p is 5; we compared these results with Table 11 in order to find the effects of increasing the number of random walks $p - r$. Again the fit of the local asymptotics to the finite sample distributions is good. As predicted by the local asymptotic analysis, the power is lower than in the case $p = 3$, i.e. one looses power as more and more random walks are added to the process.

7.2 CI rank equal to 2

In this subsection we describe the structure of the designs in DGP_2, which are characterized by different values of the triplet J, κ, Φ. Not all possible configurations of J are considered here, due to space limitations. We choose

$$J = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$$

with λ_1, λ_2 real, non necessarily equal. This allows to have λ_2 close to unity and λ_1 sufficiently far away in order to calculate local power of CI rank test with the null of $r \leq 1$ versus higher rank.[6]

Because J is diagonal, we can apply further simplifications described in Paruolo (2005), thus treating Φ as a correlation matrix, indexed by a scalar correlation coefficient τ. Finally κ is a 2×2 matrix; we thus have

$$\mathrm{var}(\upsilon_t) = \begin{pmatrix} 1 & \tau & 0 \\ \tau & 1 & 0 \\ 0 & 0 & I_{p-2} \end{pmatrix}, \qquad \kappa = \begin{pmatrix} \kappa_{11} & \kappa_{12} \\ 0 & \kappa_{22} \end{pmatrix}$$

and $\kappa_{ii} \geq 0$.

In order to choose values for J, observe that $-1 < \lambda_1, \lambda_2 < 1$; moreover one can choose $\lambda_1 \leq \lambda_2$ because the process is invariant with respect to switches in λ_1, λ_2 given that Φ is a correlation matrix. We choose $\lambda_1 \in \mathbb{J}_1 := \left\{-\frac{1}{2}, 0, \frac{1}{2}\right\}$, $\lambda_2 \in \mathbb{J} := \left\{-\frac{1}{2}, 0, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \frac{31}{32}\right\}$ where we select pairs $l := (\lambda_1, \lambda_2)$ with $\lambda_1 \leq \lambda_2$. The resulting set $\mathbb{L} := \{l_i\}$ of pairs $l := (\lambda_1, \lambda_2)$ has cardinality 18 and it is shown in Table 14 with the ordering used in the following. 15 designs in $\mathbb{L}$ have $\lambda_1 < \lambda_2$ and 3 have $\lambda_1 = \lambda_2 =: \lambda$, with $J = \lambda I_2$ is a diagonal matrix; the latter are labelled l_1, l_2, l_3.

[Tables 14 and 15 here]

In Paruolo (2005) it is shown that τ can be always be taken to be non-negative for a diagonal J matrix, which is the case considered here. Moreover it is shown there that for the scalar configurations l_1, l_2, l_3 of J ($J = \lambda I_2$) one can set $\tau = 0$. For non-scalar J we take $\tau \in \{0, 0.5\}$, where for later reference we indicate $\tau_1 = 0$, $\tau_2 = 0.5$. Table 15 lists the configurations we consider for $k := \mathrm{vech}(\kappa')$.

The designs included in DGP_2 are obtained as follows. The first 27 designs correspond to a scalar J (l_1, l_2, l_3) and $\tau = 0$; they are given by $(l, \tau, k) = (l_i, 0, k_h)$ for $i = 1, 2, 3$ and $h = 1, ..., 9$. The remaining 270 designs are characterized by $\lambda_1 < \lambda_2$, and are given by $(l, \tau, k) = (l_i, \tau_j, k_h)$ for $i = 4, ..., 18$, $j = 1, 2$ and $h = 1, ..., 9$. We indicate the corresponding designs as $\mathrm{DGP}_2(i, j, h)$, where the ordering of i, j, h is the one described above.

For the analysis of the local asymptotics, some of the 297 designs in DGP_2 can be interpreted as local to the case of CI rank $r = 1$. This is obtained when λ_2 is close to 1, i.e. when $\lambda_2 \in \mathbb{J}_2 := \{\frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \frac{31}{32}\}$; note that λ_1 is always chosen far from 1, i.e. $\lambda_1 \in \mathbb{J}_1$.

[6] This choice of J excludes local-to-I(2) processes and stationary processes with complex AR roots.

In order to compare the finite sample results with the appropriate local asymptotics, we calculate the parameters f and g of the local asymptotics for the designs with $\lambda_2 \in \mathbb{J}_2$. In order to do so, we set $r = 1$, $\alpha = (\lambda_1 - 1\colon 0\colon \kappa_{11}\colon 0'_{p-3})'$, $\beta = (1\colon 0'_{p-1})'$, $\alpha_1 = (0\colon T(\lambda_2 - 1)\colon T\kappa_{12}\colon T\kappa_{22}\colon 0'_{p-4})'$, $\beta_1 = (0\colon 1\colon 0'_{p-2})'$, where 0_n indicates a $n \times 1$ zero vector. We substitute the values of α, β, α_1, β_1 and $\operatorname{var}(\upsilon_t)$ in place of Ω in (11), and we find $f = T(\lambda_2 - 1)$,

$$g = \left((T\kappa_{22})^2 + \frac{(T\kappa_{12}(\lambda_1 - 1) + f\kappa_{11}\tau)^2}{(1-\tau^2)\,\kappa_{11}^2 + (\lambda_1 - 1)^2} \right)^{\frac{1}{2}}. \tag{13}$$

Observe that one obtains f in $\mathbb{F}$ for the same pairs of values (J, T) reported in the left panel in Table 7, simply substituting J in the headings of the Table with λ_2. The expression (13) for g involves all the remaining design coefficients λ_1, τ, κ_{11}, κ_{12}, κ_{22}. Note that $\kappa = 0$ implies $g = 0$, so that designs (l_i, τ_j, k_1) all have $g = 0$. Similarly if $\tau = 0$, also $\kappa_{12} = \kappa_{22} = 0$ implies $g = 0$; this corresponds to designs (l_i, τ_1, k_4). In the following we use eq. (13) to calculate the value of g for each design with $\lambda_2 \in \mathbb{J}_2$ of in DGP_2.

As for DGP_1, we present evidence first results for the designs with $-1/2 \le \lambda_1 \le \lambda_2 \le 1/2$, where the stationary roots are far from 1. As a worst case we present results for $\lambda_1 = 1/2$, $\lambda_2 = 3$, $\tau = 0$, $k = k_9$, which is $\mathrm{DGP}_2(3, 1, 9)$.[7] Table 16 presents the percentage of rejections of ${}^mQ(i)$ for $m = 0, 1, 2$ and $i = 0, 1, 2$, when $p = 5$. These are the relevant tests in the CI rank estimators described in Section 4. It can be seen that all tests diverge for $i = 0, 1$ at $T = 96$ and give approximately the same number of rejections for all $T \in \mathbb{T}$. The size of the ${}^mQ(2)$ tests converges to 5% with a similar rate for different m, $m = 0, 1, 2$.

[Tables 16, 17 and 18 here]

We next consider cases where λ_2 is close to 1. Consider for example $\mathrm{DGP}_2(18, 1, 1)$ with $\lambda_1 = 1/2$, $\lambda_2 = 31/32$, $\tau = 0$, $k = 0$. Table 17 presents the percentage of rejections of ${}^mQ(i)$ for $m = 0, 1, 2$ and $i = 0, 1, 2$, when $p = 5$. It can be seen that all ${}^mQ(0)$ tests tend to diverge at $T = 192$ but the ${}^mQ(1)$ fails to do so, for all $m = 0, 1, 2$. This is parallel to the finding in Table 9 for $\mathrm{DGP}_1(7, 1)$. From the local power analysis we know that the bigger the g parameter, the higher the power. In this type of design $g = 0$, and the test has low power for all $T \in \mathbb{T}$. Note that, from Section 6, the simulated local asymptotic power of ${}^mQ$ for $p - r = 4$, $f = -6$, $g = 0$ is approximately 5% for all ${}^mQ$ tests, $m = 0, 1, 2$.

This is also apparent when we compare Table 17 with Table 18, which reports results for $\mathrm{DGP}_2(18, 2, 9)$, where l is the same and k now takes the non-zero value $(1, 1, 1)'$. The table presents the percentage of rejections of ${}^mQ(i)$ for $m = 0, 1, 2, 3$, $i = 0, 1, 2$, $p = 5$. It can be seen that all ${}^mQ(0)$ *and* ${}^mQ(1)$ tests diverge for all $m = 0, 1, 2$. Note that now the value of g exceeds 200 at $T = 192$.

Thus, despite a stationary root λ_2 very close to 1, a parameter g far away from 0 gives enough power to the ${}^mQ(1)$ tests to reject. The price of a root λ_2 very close to 1 is the fact that the ${}^mQ(2)$ tests are still oversized at $T = 192$, with a size close to 13% instead of 5%.

[7] We define as worst case the one where the size of the ${}^mQ$ tests is furthest away from the nominal size at $T = 192$.

Summarizing, in all the situations described above, the finite sample properties of all the different ${}^mQ$ tests appear to be very similar, for varying m. There is rather close fit between the small sample behavior and the asymptotics, both under the null and under the local alternative. When the stationary roots are sufficiently away from 1, the asymptotic nominal size appears to be reliable. In the same situations, the CI rank estimators described in Section 4 appear to be working as predicted by Theorem 2.

8 Two illustrative examples

This section provides two empirical illustrations. The first one deals with a money demand model for the Euro area and it is reported in Subsection 8.1. The second one, presented in Subsection 8.2, refers to consumption risk-sharing among four Italian regions. Both applications illustrate how unrestricted and restricted estimation of CI rank can be performed.

8.1 Money demand system for Euro area

We consider the money demand model of Brand and Cassola (2004), to which we refer for detailed explanations and references on how variables have been constructed and aggregated across European countries. Let $m_t - p_t$, y_t, s_t, l_t and Δp_t indicate respectively the log of the real stock of money, the log of real gross domestic product (GDP), the short-term and long-term interest rates and the inflation rate, with p_t being the implicit GDP deflator.[8] These variables are often regarded as I(1) in developed economies, so one may wish to impose that there is at least one common trend in the system.

The economic long run relations predicted for the system $X_t := (m_t - p_t, y_t, s_t, l_t, \Delta p_t)'$ are relatively uncontroversial, as they appear in most theoretical and applied models of transmission mechanisms of monetary policy. Three cointegrating relations should characterize the system: a money demand relation relating the real stock of money to real GDP and a measure of the opportunity cost of holding money; a Fisher inflation parity connecting the long-term interest rate and inflation; a relation among interest rates over the yield curve consistently with the expectation hypothesis.

A VAR with unrestricted constant and $q = 2$ lags was estimated by Brand and Cassola (2004) over the 1980:1-1999:3 period, using initial observations to account for lags.[9] The results of the CI test are summarized in the left panel of Table 19, which reports the values the trace test ${}^0Q(j)$, with the corresponding 5% asymptotic

[8] The source of the data is the European Central Bank (ECB) and the aggregate time-series cover the period 1980:1-1999:3. The nominal stock of money is measured as the log of M3 (m_t); the log of the price level (p_t) is obtained from the (seasonally adjusted) GDP deflator and the (annualized) inflation rate is computed as $\Delta p_t := 4(p_t - p_{t-1})$; income is measured as the log of real GDP (y_t); short term rates (s_t) are 3-month money market interest rates and long-term interest rates (l_t) are 10-year government bond yields or close substitutes.

[9] We re-estimated the model using PcFiml 10.0. Usual diagnostic tests give no indication of residuals misspecification or parameters instability for the VAR(2). Moreover, tests for the presence of I(2) components suggest that there is no I(2) component in the system.

critical values taken from Table 3, column $m = 0$. The ${}^0Q(j)$ statistic selects 3 cointegrating relations at the 5% significance level, albeit the existence of the third cointegration vector is not clear-cut, because the ${}^0Q(j)$ statistics is approximately equal to its 5% critical value.

Table 19 displays also the corresponding ${}^1Q(j)$ statistic and 5% critical values taken from Table 3, column $m = 1$. Using expression (5), ${}^1Q(j) := {}^{\text{tr}}Q(j) - {}^{\text{tr}}Q(4)$ for $j := 0, 1, 2, 3$, ${}^1Q(j)$ can be easily calculated from the trace statistics. The ${}^1Q(j)$ test gives the same inference as the trace test, although ${}^1Q(2)$ is greater that the 5% critical value, giving better support to the choice of $r = 3$.

The three deviations from the long run equilibrium identified and estimated in Brand and Cassola (2004), see their Table 4, $\widehat{\beta}'X_t = (v_{1t},\ v_{2t},\ v_{3t})'$, are given by $v_{1t} = m_t - p_t - 1.33y_y + 1.608l_t$ (money demand), $v_{2t} = \Delta p_t - 0.67l_t$ (Fisher-type parity) and $v_{3t} = l_t - s_t$ (interest rates spread) and lead to a LR test for over-identifying restrictions equal to $\chi^2(3) = 1.47$ with a p-value of 0.69. This is in line with economic expectations.

8.2 Consumption risk-sharing among Italian regions

Consider p regions of a given country, area or monetary union, attempting to insure their consumption streams against idiosyncratic income fluctuations. According to risk-sharing theory, see e.g. Canova and Ravn (1996), under complete markets, perfect factor mobility and with utility maximizing agents endowed with constant relative risk aversion, expected lifetime utilities and exogenous stochastic output processes, the optimal consumption allocation over time and across regions can be represented as

$$c_t^{*i} = \theta^i c_t^1 + \xi^i t \ + \eta_t^i \quad , \quad i = 2, ..., p, \tag{14}$$

where c_t^i is the log of per capita consumption in region i, c_t^{*i} is the corresponding optimal level, c_t^1 is the log of per capita consumption of the leader region (e.g. the one with highest per capita GDP or population level), θ^i is a parameter measuring heterogeneity of relative risk aversion between the leader and the i-th region, ξ^i is a trend parameter which is different from zero if the regions 1 and i discount at different rates, and η_t^i is a (possibly stationary) stochastic term depending on the preference shocks of the two regions[10].

Eq. (14) describes optimal risk sharing against long term income fluctuations and allows heterogeneity in preference parameters. The theory has strong implications on interregional consumption dynamics. It is widely recognized that per capita consumption data can be approximated as non-stationary I(1) processes over relatively long span of data; eq. (14) suggests that deviations of actual from optimal per capita consumption, $c_t^i - c_t^{*i}$, should be transitory for risk sharing to hold. Put differently, net of preference shocks, c_t^i and c_t^1 should be cointegrated, possibly around a linear trend, with cointegration vector $(1 : -\theta^i)'$, see e.g. Cavaliere et al. (2005) and references therein.

Under perfect risk sharing, the vector $X_t = (c_t^1\!:\ c_t^2\!:\ ....\!:\ c_t^p)'$ should therefore embody $p-1$ cointegrating relations, i.e. a single common stochastic trend. Also in

[10] Equation (14) can be obtained also under incomplete markets provided a (benevolent) social planner maximizes collective utility, see e.g. Canova and Ravn (1996).

this case one may use the constrained CI rank estimator that assumes the existence at least of one common trend.

We consider the consumption data of four regions of central Italy: Lazio, Marche, Umbria and Abruzzo.[11] One would expect to identify three cointegrating relations of the form $c_t^i - \theta^i c_t^1 - \xi^i t$, $i = 2, 3, 4$ in $X_t = (c_t^1 \ , c_t^2 \ , c_t^3 \ , c_t^4)'$, where we label Lazio as region 1, Marche, Abruzzo and Umbria as regions 2, 3 and 4. [12]

The attractive feature of the regional risk-sharing model sketched above is that the preference parameters θ^i are allowed to vary across regions under long run consumption insurance; moreover, preference homogeneity can be tested as the hypothesis $\theta^i = 1$ on the CI space, after determining its rank.

A VAR(1) with unrestricted constant and a linear trend restricted to the cointegration space was estimated over the 1960-2001 period.[13] The results for ${}^0Q(j)$ and ${}^1Q(j)$ are summarized in the right panel of Table 19, with the corresponding 5% asymptotic critical values taken from Table 4, column $m = 1$.

The ${}^0Q(j)$ trace statistic selects 2 cointegrating relations at the 5% significance level, one less than expected under the hypothesis of full risk sharing. Therefore, at face value, the trace test rejects the full risk sharing hypothesis as 2 stochastic trends are detected in the 4 regions. The ${}^1Q(j)$ statistics can again be calculated as ${}^1Q(j) := {}^0Q(j) - {}^0Q(3)$ for $j := 0, 1, 2$, see the right panel of Table 19. It is now observed that ${}^1Q(j)$ selects 3 cointegrating relations at the 5% significance level, as predicted by economic theory.

The restricted estimates of the CI relations β are given by

$$\widehat{\beta}' X_t := \begin{pmatrix} c_t^2 - c_t^1 \\ c_t^3 - \underset{(0.03)}{1.46} c_t^1 + \underset{(0.0009)}{0.011} t \\ c_t^4 - c_t^1 + \underset{(0.0003)}{0.001} t \end{pmatrix}$$

where asymptotic standard errors are reported in parenthesis. The LR test for the 3 over-identifying restriction in this specification equals 2.07, which gives a p-value of 0.56 when compared with a $\chi^2(3)$; over-identifying restrictions are hence supported by the data.

All estimated preference parameters θ^i exhibit the correct sign and magnitude; one here finds some heterogeneity in preference parameters. Overall it seems that the four selected regions of central Italy share risks consistently with the predictions of the theory; moreover, results confirm those in Cavaliere et al. (2005) who find substantial long-run risk sharing and heterogeneity in the preference parameters among all twenty Italian regions, using a different approach.

[Table 19 approximately here]

[11] Annual per capita regional consumption data is taken from Cavaliere et al. (2005). This is constructed as the sum of regional household and government consumption at constant 1995 prices, divided by resident population at the end of the corresponding year. The source of the data is Svimez and the Italian Statistical Institute (ISTAT); annual logged observations cover the 1960-2001 period with 42 observations per region.

[12] Lazio is here chosen as the "leader" region of central Italy because of its higher demographic density and economic importance.

[13] Also in this case the usual diagnostic tests give no indication of residuals misspecification or parameters instability for the VAR(1).

9 Conclusions

This paper investigates LR CI rank statistics for any possible value of the CI rank under the alternative; this class includes the trace and lambda-max tests as special cases. We tabulate the limit quantiles for all the new tests in the class. We find that any of the tests can be used in constructing the CI rank estimator proposed in Johansen (1992), without changing the asymptotic properties.

We investigate the local power of the tests and we find a significant deterioration of the power curves for all the tests when the number of common trends increases. The power differences within the class are moderate; these differences appear to increase as the number of common trends increases and power deteriorates.

When considering one-dimensional alternatives, one observes a superior power of the lambda-max test for values sufficiently far from the null; this is in line with the prediction of classical results on LR tests for one-sided alternatives, see Andrews (1996).

The tests in the class can also be arranged to give a constrained estimator of the CI rank, which imposes restrictions on the minimum number of common trends. We show that mis-specification of the minimum number of common trends has adverse effects, because it makes the limit probability of selecting the correct rank equal to 0. We hence advise not to leave any value of the common trends untested, unless the untested value can be excluded beyond any reasonable doubt.

Appendix

Proof of Theorem 1. Following Johansen (1991a) one can prove that $LR(r, p-m) = T\sum_{i=r+1}^{p-m}\widehat{\lambda}_i + o_p(1)$ where the eigenvalues $(T\widehat{\lambda}_{r+1},...,T\widehat{\lambda}_p)$ converge weakly to $(\eta_1, ..., \eta_{p-r})$ with $\eta_i := \mathrm{eig}_i(N(F, B))$, while $(\widehat{\lambda}_1,...,\widehat{\lambda}_r) = O_p(1)$. The latter results implies that $-T\ln(1-\widehat{\lambda}_i) = O_p(T)$ when $i \leq r$. Since the eigenvalues are all distinct with probability one, the statement of the theorem follows by applying the continuous mapping theorem, see e.g. Billingsley (1968).

Proof of Theorem 2.

The results follow applying the proof of Proposition 1 in Paruolo (2001a) to $\widetilde{r}_s$ and $\widehat{r}_m$.

Proof of Theorem 3.

The statements follow as in Theorem 2.

Proof of Theorem 4.

The proof is the same as in Johansen (1991b, 1996) and Paruolo (2001a) noting that if the eigenvalues are distinct, they are continuous functions of the argument.

Acknowledgement

A previous version of this paper circulated as 'Determining the number of cointegrating relations under rank constraints', Quaderno della Facoltà di Economia 2001/17, University of Insubria, and was presented at the (EC)^2 conference 'Model selection and evaluation', Bologna, 13-14 December 2002. We thank (without implicating)

Karim Abadir (the editor) and two referees for helpful comments on the previous version of the paper. Partial financial support is gratefully acknowledged under Italian MIUR COFIN 2002 and 2004 grants. Paolo Paruolo also gratefully acknowledges partial financial support from University of Insubria FAR grants 2001–2005.

References

Andrews D.W.K. (1996). Admissibility of the likelihood ratio test when the parameter space is restricted under the alternative, *Econometrica* 64, 705-718.

Billingsley, P. (1968). *Convergence of probability measures*, New York: John Wiley and Sons.

Brand, C. and Cassola, N. (2004). A money demand system for euro area M3, *Applied Economics* 36, 817-838.

Canova, F., and Ravn, M. O. (1996), International consumption risk sharing, *International Economic Review* 37, 573–601.

Cavaliere, G. Fanelli, L., Gardini, A. (2005). Regional consumption dynamics and risk sharing in Italy, *International Review of Economics and Finance*, forthcoming.

Horvath M. T. K., Watson M. W. (1995). Testing for cointegration when some of the cointegrating vectors are prespecified, *Econometric Theory*, 11, 984–1014.

Karatzas, I. and S.E. Shreve (1988). *Brownian motion and stochastic calculus*, New York: Springer Verlag.

Johansen, S. (1991a). Estimation and hypothesis testing of cointegration vectors in Gaussian autoregressive models, *Econometrica* 59, 1551–1580.

Johansen, S. (1991b). The power function of the likelihood ratio test for cointegration, in J. Gruber (ed.) *Econometric decision models: new methods of modelling and applications*, New York: Springer Verlag, 323–335.

Johansen, S. (1992). Determination of cointegration rank in the presence of a linear trend, *Oxford Bulletin of Economics and Statistics*, 54, 383–397.

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

Lütkepohl H., P. Saikkonen and C. Trenkler (2001). Maximum eigenvalue versus trace tests for the cointegrating rank of a VAR process, *Econometrics Journal* 4, 287–310.

Nielsen B. (2004). On the distribution of likelihood ratio test statistics for cointegration rank, *Econometric Reviews* 23, 1–23.

Paruolo, P. (2001a). The power of Lambda Max, *Oxford Bulletin of Economics and Statistics*, 63, 395–403.

Paruolo P. (2001b). LR cointegration tests when some cointegrating relations are known, *Statististical Methods and Applications*, 10, 123-137.

Paruolo, P. (2002). On Monte Carlo estimation of relative power, *Econometrics Journal*, 5, 65-75.

Paruolo P. (2005). Design of vector autoregressive processes for invariant statistics, Quaderno della Facoltà di Economia 2005/6, University of Insubria.

Saikkonen, P. and H. Lütkepohl (2000). Testing for the cointegrating rank of a VAR process with an intercept. *Econometric Theory* 16, 373–406.

Figures and tables

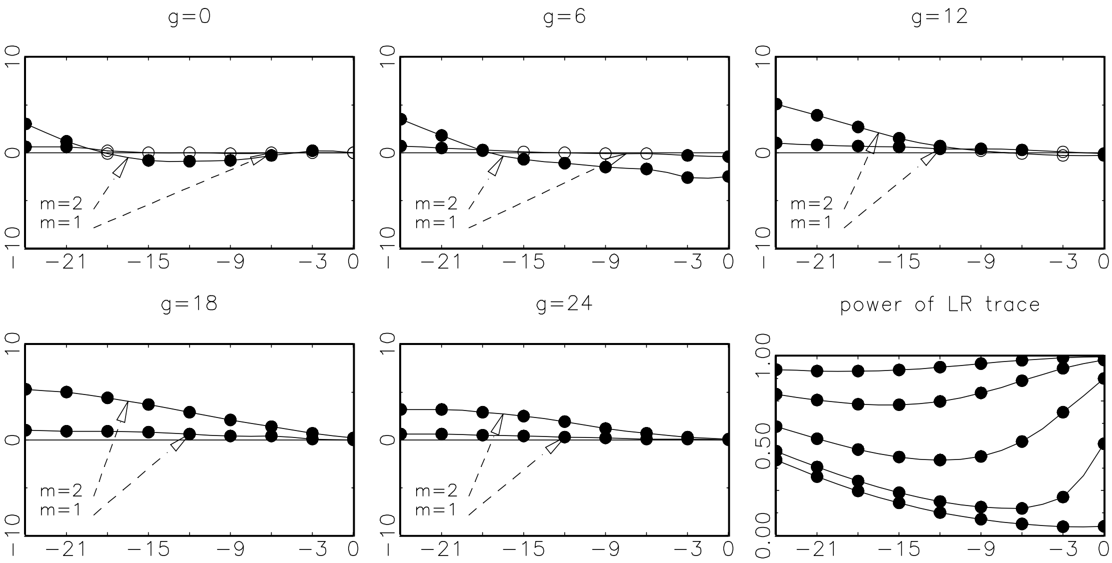

Figure 1: Power comparison for $p - r = 3$. f on x-axis; $n = 10^5$ replications, $T = 2500$. Lines obtained by quadratic interpolation. $m = 2$ corresponds to the ${}^{\max}Q$ test. Panels $g = 0$ to $g = 24$: percentage power difference $100(\widehat{\pi}_m - \widehat{\pi}_{\mathrm{tr}})$. Empty circle indicate insignificant power differences. Lower right panel: power function for ${}^{\mathrm{tr}}Q$. Power curves correspond to values of $g = 0$ (bottom line), 6, 12, 18 and 24 (top line).

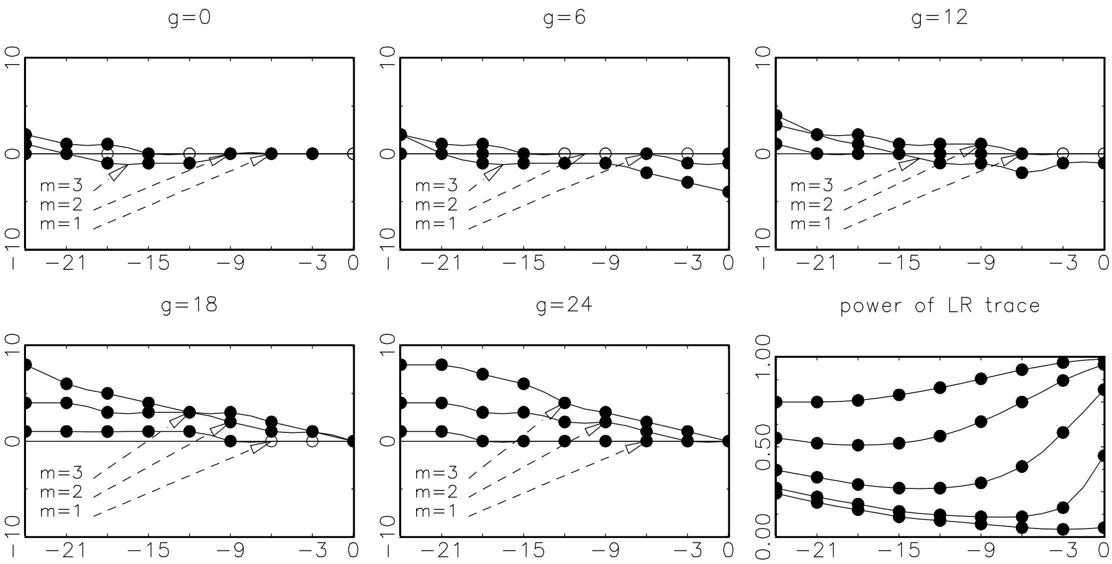

Figure 2: Power comparison for $p - r = 4$. $m = 3$ corresponds to the ${}^{\max}Q$ test. Panels $g = 0$ to $g = 24$: $100(\widehat{\pi}_m - \widehat{\pi}_{\text{tr}})$. Lower right panel: power of trace test ${}^{0}Q$. See caption of Fig. 1.

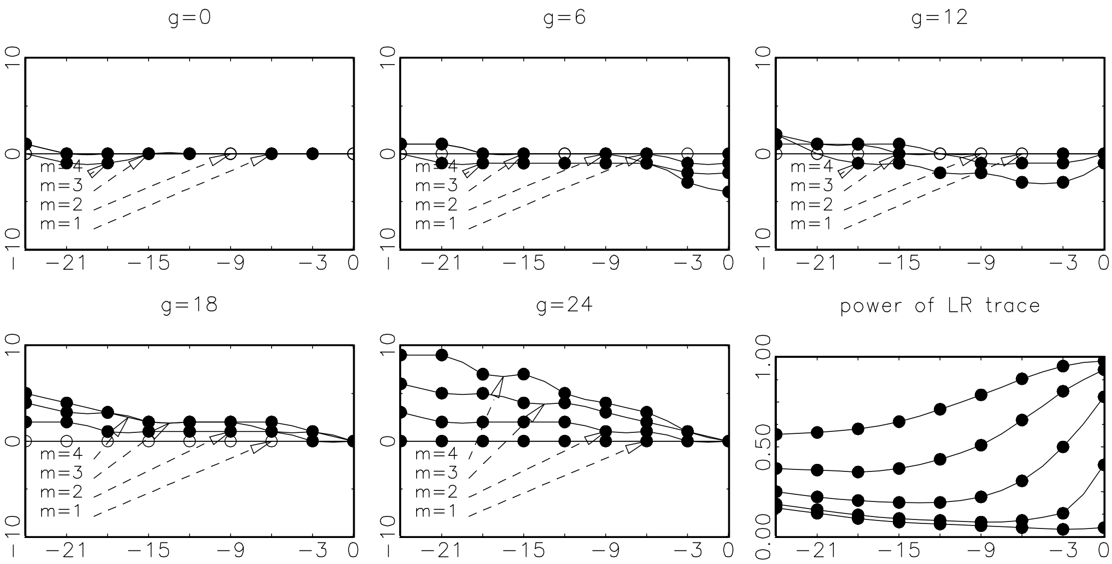

Figure 3: Power comparison for $p - r = 5$; $m = 4$ corresponds to the ${}^{\max}Q$ test. Panels $g = 0$ to $g = 24$: $100(\widehat{\pi}_m - \widehat{\pi}_{\mathrm{tr}})$. Lower right panel: power of trace test ${}^{0}Q$. See caption of Fig. 1.

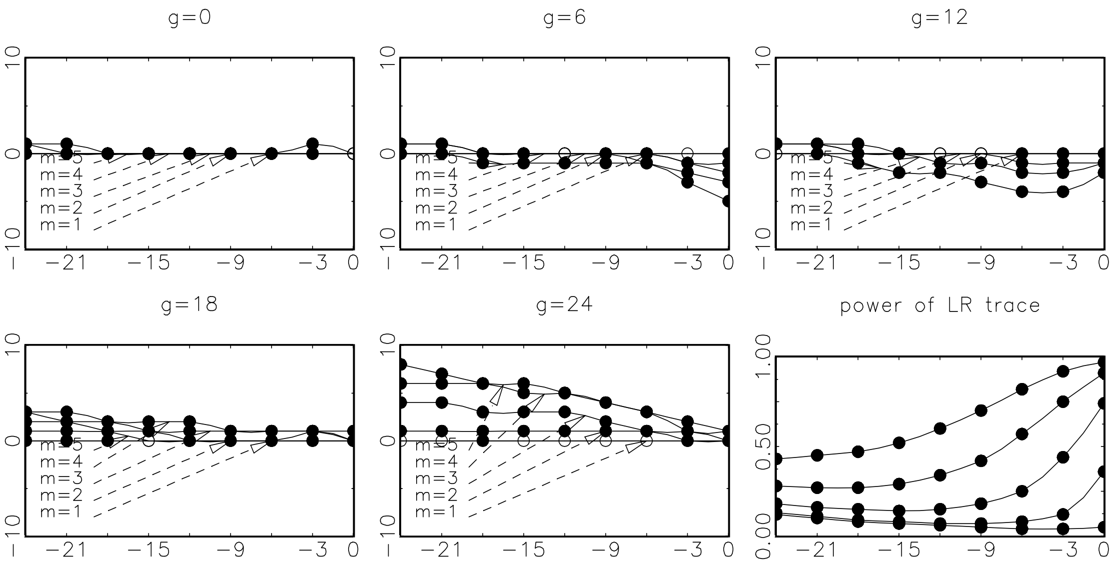

Figure 4: Power comparison for $p - r = 6$. $m = 5$ corresponds to the ${}^{\max}Q$ test. Panels $g = 0$ to $g = 24$: $100(\widehat{\pi}_m - \widehat{\pi}_{\mathrm{tr}})$. Lower right panel: power of trace test ${}^{0}Q$. See caption of Fig. 1.

$p-r$	$m=0$	1	2	3	4	5	6	7
1	4.156							
2	12.327	11.224						
3	24.286	23.500	17.790					
4	40.080	39.402	34.679	24.076				
5	59.819	59.217	54.953	46.063	30.284			
6	83.784	83.205	79.101	70.906	57.436	36.482		
7	111.400	110.803	106.687	99.038	86.832	68.754	42.653	
8	143.335	142.779	138.689	131.112	119.586	103.129	80.284	48.677

Table 1: 95% asymptotic quantiles for ${}^mQ$ for different values of $(p-r)$ and m. ${}^0Q$ is the trace test. Entries on the main diagonal, corresponding to $m = p-r-1$, are the ones of the lambda-max test. Case 1 in Theorem 1.

$p-r$	$m=0$	1	2	3	4	5	6	7
1	9.158							
2	20.287	15.962						
3	35.157	31.543	22.286					
4	53.945	50.536	42.819	28.506				
5	76.714	73.387	66.335	54.171	34.676			
6	103.574	100.207	93.442	82.277	65.448	40.833		
7	134.179	130.721	124.063	113.633	98.331	76.866	46.939	
8	169.042	165.458	158.894	148.709	134.337	114.581	88.248	53.051

Table 2: 95% asymptotic quantiles for ${}^mQ$. Case 2 in Theorem 1. See caption of Fig. 1.

$p-r$	$m=0$	1	2	3	4	5	6	7
1	3.807							
2	15.488	14.273						
3	29.782	28.849	21.076					
4	47.725	46.953	40.620	27.493				
5	69.553	68.836	63.160	52.129	33.771			
6	95.304	94.610	89.357	79.319	63.593	39.836		
7	125.162	124.468	119.216	109.984	95.697	75.129	46.109	
8	158.968	158.353	153.065	144.070	130.699	111.932	86.505	52.200

Table 3: 95% asymptotic quantiles for ${}^mQ$. Case 3 in Theorem 1. See caption of Fig. 1.

$p-r$	$m=0$	1	2	3	4	5	6	7
1	12.516							
2	25.844	19.375						
3	42.790	37.479	25.724					
4	63.630	58.799	48.975	31.936				
5	88.422	83.797	74.889	60.471	38.193			
6	117.324	112.729	104.281	91.073	71.798	44.323		
7	149.972	145.367	137.289	125.001	107.372	83.371	50.473	
8	186.676	182.087	174.029	162.255	145.804	123.703	94.697	56.524

Table 4: 95% asymptotic quantiles for ${}^mQ$. Case 4 in Theorem 1. See caption of Fig. 1.

$p-r$	$m=0$	1	2	3	4	5	6	7
1	3.820							
2	18.330	17.059						
3	35.025	33.983	24.255					
4	55.053	54.165	46.173	30.588				
5	78.994	78.158	70.971	57.880	37.054			
6	106.833	106.076	99.348	87.573	69.504	43.170		
7	138.629	137.886	131.399	120.336	104.140	81.188	49.406	
8	174.411	173.734	167.382	156.854	141.549	120.613	92.703	55.555

Table 5: 95% asymptotic quantiles for ${}^mQ$. Case 5 in Theorem 1. See caption of Fig. 1.

$p-r$	2	3	4	5	6
min	-1.5	-2.6	-4.0	-4.0	-5.0
max	1.5	5.3	8.0	9.0	8.0

Table 6: Absolute power differences $100(\widehat{\pi}_m - \widehat{\pi}_{tr})$; minimum and maximum calculated across f, g, m. $\widehat{\pi}_m$ is the estimated power of test ${}^mQ$.

T	$J=$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{15}{16}$	$\frac{31}{32}$	$\kappa=$	$\frac{1}{32}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$
24		**−12**	**−6**	**−3**	−1.5	−0.75		0.75	1.5	**3**	**6**	**12**
48		**−24**	**−12**	**−6**	**−3**	−1.5		1.5	**3**	**6**	**12**	**24**
96		−48	**−24**	**−12**	**−6**	**−3**		**3**	**6**	**12**	**24**	48
192		−96	−48	**−24**	**−12**	**−6**		**6**	**12**	**24**	48	96

Table 7: Left panel: $f = T(J-1)$; right panel: $g = T\kappa$; bold entries indicate values of f, g in the range of values used in the local asymptotic simulations.

T	${}^0Q(0)$	${}^0Q(1)$	${}^1Q(0)$	${}^1Q(1)$	${}^2Q(0)$	${}^2Q(1)$	${}^3Q(0)$	${}^3Q(1)$
24	23.50	3.83	23.44	3.60	23.56	3.08	22.22	1.85
48	28.19	3.22	28.20	3.30	28.91	2.82	29.42	1.99
96	73.23	4.80	73.62	4.78	76.86	4.83	81.61	4.22
192	99.97	5.44	99.97	5.52	99.99	5.48	99.99	5.09
∞	100	5	100	5	100	5	100	5

Table 8: DGP$_1$(3,1), $p = 5$. $J = 1/2$, $\kappa = 0$ % frequency of rejections. $T = \infty$ from the asymptotics under the null.

T	f	g	${}^0Q(0)$	${}^0Q(1)$	${}^1Q(0)$	${}^1Q(1)$	${}^2Q(0)$	${}^2Q(1)$	${}^3Q(0)$	${}^3Q(1)$
24	-0.75	0	19.10	2.68	18.99	2.63	18.98	2.19	17.95	1.30
48	-1.5	0	8.42	0.94	8.44	0.98	8.43	0.75	8.31	0.40
96	-3	0	6.06	0.42	6.10	0.44	6.15	0.33	5.94	0.31
192	-6	0	5.62	0.52	5.60	0.48	5.79	0.38	5.76	0.30
∞			100	5	100	5	100	5	100	5

Table 9: DGP$_1$(7,1), $p = 5$. $J = 31/32$, $\kappa = 0$. % frequency of rejections. $T = \infty$ from the asymptotics under the null.

T	f	g	${}^0Q(0)$	${}^0Q(1)$	${}^1Q(0)$	${}^1Q(1)$	${}^2Q(0)$	${}^2Q(1)$	${}^3Q(0)$	${}^3Q(1)$
24	-0.75	12	75.49	24.59	75.29	24.06	74.41	20.10	72.06	13.72
48	-1.5	24	95.63	21.97	95.67	21.88	95.86	18.44	96.11	13.67
96	-3	48	99.97	18.73	99.98	18.67	99.98	16.34	100	12.79
192	-6	96	100	14.12	100	14.15	100	12.85	100	10.37
∞			100	5	100	5	100	5	100	5

Table 10: DGP$_1$(7,6), $p = 5$. $J = 31/32$, $\kappa = 1/2$. % frequency of rejections. $T = \infty$ from the asymptotics under the null.

T	$f = -3$	-6	-12	-24	$f = -3$	-6	-12	-24
		$g = 0$				$g = 6$		
24	7.9	9.6	20.0	-	24.4	20.1	27.5	-
48	6.2	7.5	15.2	54.1	22.8	17.2	23.1	60.0
96	5.5	7.4	13.5	48.0	21.4	17.0	20.3	53.1
192	-	6.7	13.5	44.1	-	15.1	19.5	48.9
∞	4.8	6.3	12.7	42.0	21.3	15.0	18.8	46.9
		$g = 12$				$g = 24$		
24	66.6	54.4	50.6	-	-	-	-	-
48	67.3	52.0	45.9	72.7	98.5	96.8	93.5	95.0
96	68.5	53.0	45.0	66.2	98.6	97.0	93.6	93.9
192	-	51.8	42.8	63.2	-	97.6	93.7	93.2
∞	68.7	52.3	41.9	60.7	99.1	97.7	93.8	92.5

Table 11: DGP$_1$, $p = 3$, trace test ${}^0Q$: % frequency of rejections. $T = \infty$ from the simulation of the asymptotic distribution under the local alternative.

T	$f=-3$	-6	-12	-24	$f=-3$	-6	-12	-24
		$g=0$				$g=6$		
24	7.8	8.9	18.4	-	21.5	17.5	26.2	-
48	5.9	6.7	14.0	59.1	19.8	15.3	22.1	65.1
96	5.5	6.8	12.8	51.5	18.7	14.5	19.5	58.0
192	-	6.3	12.0	47.9	-	13.1	18.5	53.7
∞	4.9	6.0	11.8	45.0	18.7	13.3	17.6	50.4
		$g=12$				$g=24$		
24	64.8	53.2	51.7	-	-	-	-	-
48	66.3	51.3	46.2	77.9	98.8	97.7	95.5	97.2
96	67.3	52.6	45.8	71.6	98.9	97.8	95.4	96.7
192	-	52.0	43.7	68.3	-	98.1	95.6	96.2
∞	68.4	52.2	42.6	65.8	99.4	98.4	95.8	95.7

Table 12: DGP$_1$, $p = 3$, test ${}^2Q$: % frequency of rejections. $T = \infty$ from the simulation of the asymptotic distribution under the local alternative.

T	$f=-3$	-6	-12	-24	$f=-3$	-6	-12	-24
		$g=0$				$g=6$		
24	18.2	18.8	23.5	-	31.1	26.0	28.4	-
48	8.8	8.6	12.1	28.2	20.1	14.6	16.3	30.4
96	6.1	6.7	9.3	20.5	15.8	12.1	12.0	24.3
192	-	5.6	8.1	18.1	-	10.1	10.5	21.2
∞	4.3	4.7	6.6	15.8	13.4	8.7	8.8	17.7
		$g=12$				$g=24$		
24	61.1	48.2	41.8	-	-	-	-	-
48	53.1	36.5	27.9	39.8	93.0	86.0	73.1	69.4
96	51.1	33.1	23.3	31.6	93.6	86.3	70.8	62.5
192	-	31.2	21.4	28.1	-	87.6	71.5	59.9
∞	49.8	30.5	19.3	24.6	94.6	87.9	71.0	57.1

Table 13: DGP$_1$, $p = 5$, trace test ${}^0Q$: % frequency of rejections. $T = \infty$ from the simulation of the asymptotic distribution under the local alternative.

λ_1	$\lambda_2 =$	$-\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{15}{16}$	$\frac{31}{32}$
$-\frac{1}{2}$		l_1	l_4	l_5	l_6	l_7	l_8	l_9
0			l_2	l_{10}	l_{11}	l_{12}	l_{13}	l_{14}
$\frac{1}{2}$				l_3	l_{15}	l_{16}	l_{17}	l_{18}

Table 14: (λ_1, λ_2) pairs. l_1, l_2, l_3 correspond to $\lambda_1 = \lambda_2 =: \lambda$.

	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9
κ_{11}	0	0	0	1	1	1	1	1	1
κ_{12}	0	-1	1	0	-1	1	0	-1	1
κ_{22}	0	0	0	0	0	0	1	1	1

Table 15: k triplets included in the design.

T	f	g	${}^0Q(0)$	${}^0Q(1)$	${}^0Q(2)$	${}^1Q(0)$	${}^1Q(1)$	${}^1Q(2)$	${}^2Q(0)$	${}^2Q(1)$	${}^2Q(2)$
24	-12	26.29	99.49	58.76	10.96	99.53	58.74	1095	99.65	58.89	8.10
48	-24	52.58	100	93.14	8.85	100	93.49	907	100	95.09	8.38
96	-48	105.16	100	100	7.15	100	100	705	100	100	6.73
192	-96	210.33	100	100	6.12	100	100	596	100	100	5.56
∞			100	100	5	100	100	5	100	100	5

Table 16: $\text{DGP}_2(15,2,9)$, $p = 5$. % frequency of rejections. $T = \infty$ from the asymptotics under the null.

T	f	g	${}^0Q(0)$	${}^0Q(1)$	${}^0Q(2)$	${}^1Q(0)$	${}^1Q(1)$	${}^1Q(2)$	${}^2Q(0)$	${}^2Q(1)$	${}^2Q(2)$
24	-0.75	0	23.27	3.23	0.53	23.07	3.14	0.46	23.04	2.67	0.21
48	-1.5	0	27.06	3.06	0.40	27.30	3.01	0.39	28.06	2.52	0.20
96	-3	0	73.08	4.80	0.44	73.49	4.68	0.42	76.59	4.90	0.32
192	-6	0	99.99	5.90	0.46	100	5.93	0.42	100	6.13	0.22
∞			100	100	5	100	100	5	100	100	5

Table 17: $\text{DGP}_2(15,2,9)$, $p = 5$. % frequency of rejections. $T = \infty$ from the asymptotics under the null.

T	f	g	${}^0Q(0)$	${}^0Q(1)$	${}^0Q(2)$	${}^1Q(0)$	${}^1Q(1)$	${}^1Q(2)$	${}^2Q(0)$	${}^2Q(1)$	${}^2Q(2)$
24	-0.75	27.00	99.98	72.22	21.19	99.98	71.66	20.35	99.97	68.11	13.33
48	-1.5	54.01	100	96.82	20.96	100	96.86	20.37	100	97.12	15.10
96	-3	108.01	100	100	16.66	100	100	16.28	100	100	12.66
192	-6	216.02	100	100	12.98	100	100	12.74	100	100	11.00
∞			100	100	5	100	100	5	100	100	5

Table 18: $\text{DGP}_2(18,2,9)$, $p = 5$. % frequency of rejections. $T = \infty$ from the asymptotics under the null.

j	${}^0Q(j)$	5% cv	${}^1Q(j)$	5% cv	${}^0Q(j)$	5% cv	${}^1Q(j)$	5% cv
0	95.71	69.55	95.69	68.84	97.91	63.63	92.80	58.89
1	59.26	47.72	59.24	46.95	53.55	42.79	48.44	37.48
2	29.79	29.78	29.77	28.85	25.00	25.84	19.89	19.37
3	13.59	15.49	13.57	14.27	5.11	12.52	-	-
4	0.021	3.81	-	-				

Table 19: Two empirical illustrations. Left panel: Money demand system for Euro area M3. Right panel: Risk-sharing model for 4 Italian regions. Critical values (cv) for the trace test ${}^0Q(j)$ and ${}^1Q(j)$ are taken from Table 3.