A BOOTSTRAP CO-INTEGRATION RANK TEST FOR PANELS OF VAR MODELS

LAURENT A. F. CALLOT
CREATES
AARHUS UNIVERSITY
LCALLOT@CREATES.AU.DK

Abstract. This paper proposes a sequential procedure to determine the co-integration rank of panels of co-integrated VAR models. The rank is defined as the number of co-integration vectors within an individual system and between that system and the rest of the panel. The method proposed by Pesaran (2006) is used to control for cross section dependence and reduce the dimension of the parameter space. A bootstrap procedure derived from the bootstrap rank test of Cavaliere et al. (2012) is used to compute the empirical distribution of the trace test statistics and construct a panel trace test statistic based on pooling of the individual *p-values*. The small sample properties of these tests are documented by means of Monte Carlo. An empirical application illustrates the usefulness of this tests.

JEL classification: C12, C32, C33

Keyword Co-integration, Rank test, Panel data, Bootstrap, Cross section dependence.

1. Introduction

This paper proposes a panel co-integration rank test statistic using a bootstrap procedure to compute $p-values$, and an associated sequential procedure to determine the cointegration rank in a Panel of Co-integrated Vector Autoregressive models (PCVAR). The co-integration rank is defined as the number of co-integrating relationships among the variables of a given unit of the panel, as well as between these variables and the rest of the panel. In an international macroeconomic setting this would correspond to the number of co-integrating relationships among the variables of country i and between country i and the rest of the world. It is of interest to be able to model the dynamics across units of the panel. Economic theory predicts many relations among variables from different countries, often driven by arbitrage for instance. For example, domestic and foreign prices are linked by purchasing power parity, interest rates by the interest rate parity and so forth.

Date: September 11, 2012.

I would like to thank Niels Haldrup, Anders Bredahl Kock, Johan Lyhagen, M. Hashem Pesaran, Anders Rahbek, and participants at the 2011 Nordic Econometric Meeting and The 2011 Econometric Society European Meeting for valuable comments and suggestions. All remaining errors and shortcomings are mine. Financial support by the Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation, is gratefully acknowledged.

Determination of the co-integration rank is important to understand the long-run dynamics of a system of variables. The sequential likelihood-based procedure by Johansen (1995) is frequently used to estimate the co-integration rank. This procedure is based on a likelihood ratio test, the so-called trace test, for the hypothesis that the true rank of the system, r_0, is less than or equal to r (the hypothesis is noted $\mathcal{H}(r)$) against the hypothesis that the system has full rank (*i.e.* $\mathcal{H}(p)$ in a system with p endogenous variables). The procedure tests $\mathcal{H}(0), \cdots, \mathcal{H}(p-1)$ sequentially until one of the hypotheses isn't rejected. If all are rejected, the system has full rank. The poor small sample performance of inference procedures based on asymptotic distribution of the likelihood ratio test have been documented (see among others Reinsel and Ahn (1992); Johansen (2002); Cavaliere et al. (2012)).

Bootstrap methods are increasingly used to compute empirical test statistic distributions that are more accurate than their asymptotic counterparts, thus yielding tests with small sample sizes closer to their nominal values. Cavaliere et al. (2012) propose a bootstrap algorithm to compute an empirical distribution of the likelihood ratio rank test statistic. They show that by estimating all the parameters of a co-integrated VAR under the hypothesis that the rank is equal to r, the resulting process is asymptotically I(1) with r co-integrating vectors (noted I(1,r)). Simulations show that this bootstrap procedure yields much more accurate sizes than the asymptotic alternatives. The main aim of this paper is to extend the procedure of Cavaliere et al. (2012) to panels of co-integrated VAR models.

Large panels are increasingly used in empirical economics to analyze data set composed of many countries, regions, industries or markets. This has led to a large literature on panels with large cross section (N) and a large time dimension (T). Two difficulties arise when working with large panels.

(1) Panels where the number of individual N is small relative to T can be estimated as a single model. The number of parameters in such models grows quadratically with the number of individuals. This is often referred to as the curse of dimensionality. In order to estimates panels with large N and T, one has to control the number of parameters.
(2) Many economic variables exhibit common patterns across individuals. When ignored in the modeling of the panel this translate in cross section dependence of the residuals, leading to biased inference.

Breitung and Pesaran (2008) review the literature on cointegration and rank testing in panels. Only a few procedures exist to test for multiple cointegration (i.e. a rank potentially greater than one) in multivariate systems. Larsson et al. (2001) adapts the Johansen (1995) likelihood based framework to panels VAR, using a standardized rank test statistic to obtain a normally distributed panel rank test statistic. The procedure by Larsson et al. (2001) has three major drawbacks: it ignores cross section dependence and consequently is potentially invalid, it requires the asymptotic distribution of the trace test statistic to be homogeneous across individuals, and it requires the two first moments of the asymptotic distribution of the trace test statistic to be simulated. The procedure proposed in this paper overcomes these problems.

Pesaran (2006) and Kapetanios et al. (2011) propose a method to break the curse of dimensionality and model cross section dependence in order to obtain residuals that are cross-sectionally uncorrelated. This method is extended to non-stationary VAR by Dees et al. (2007). It is based on the use of weighted cross section averages

of the data to construct proxy for the unobserved common factors assumed to cause cross section dependence. One of the useful side-effect of this method is that it allows to test jointly for the number of cointegration relations within the individual system and between the individual system and the rest of the panel.

In the next section I introduce the PCVAR model and show how it can be transformed into set of independent individual models. I then introduce the bootstrap algorithm and the panel trace test statistic and investigate its finite sample properties by means of a Monte Carlo experiment. In the last section an empirical application to interest rate dynamics in the Euro zone demonstrates the potential uses of this method.

2. The Model

Consider a panel vector autoregressive model of order k in error correction form:

$$\Delta Y_t = \tilde{\alpha}\tilde{\rho}' d_t + \tilde{\psi}\delta_t + \tilde{\alpha}\tilde{\beta}' Y_{t-1} + \sum_{l=1}^{k-1} \tilde{\Gamma}_l \Delta Y_{t-l} + \tilde{\epsilon}_t \tag{1}$$

The variables are stacked $Y_t := (Y'_{1,t}, \ldots, Y'_{N,t})'$ in an $[Np \times 1]$ vector. $Y_{i,t}$ is the $p \times 1$ vector of variables for unit i. If there is cointegration among the variables of the model, it is useful to write the model in error correction form to highlight the matrix of cointegration vectors $\tilde{\beta}$, of dimension $Np \times \tilde{r}$, and the matrix of adjustment parameters $\tilde{\alpha}$ of dimension $Np \times \tilde{r}$. The cointegration rank of the system, *i.e.* the number of cointegration vectors, is $\tilde{r}$. The parameter matrices $\tilde{\Gamma}_l$ are of dimension $Np \times Np$. $\tilde{\epsilon}_t$ is an $Np \times 1$ vector of independent Gaussian shocks to the system, with mean 0 and variance-covariance matrix $\tilde{\Omega}$. The characteristic polynomial associated with model 1 has $Np - \tilde{r}$ roots equal to 1 and all the other roots are outside the unit circle.

The deterministic variables are separated between a component laying in the cointegration space (d_t) and a component outside the cointegration space (δ_t). They are assumed to fall in one of the three following categories defined in Johansen (1988):

i) No deterministic components: $d_t = 0$, $\delta_t = 0$.
ii) Restricted constant: $d_t = 1$, $\delta_t = 0$.
iii) Restricted linear trend: $d_t = t$, $\delta_t = 1$.

The log-likelihood function for the model given in equation 1 is:

$$\mathcal{L}\left(\tilde{\alpha}, \tilde{\beta}, \tilde{\Gamma}_l, \tilde{\Omega}, \tilde{\rho}, \tilde{\psi}\right) = -\frac{T}{2} \ln\left(\det\left(\tilde{\Omega}\right)\right) - \frac{1}{2} \sum_t \left(\tilde{\epsilon}'_t \tilde{\Omega}^{-1} \tilde{\epsilon}_t\right) \tag{2}$$

The number of parameters in the model depends quadratically on N and p. Even for a system with a moderate number of individuals, the parameters of the model become impossible to estimate by maximum likelihood for conventional sizes of T. To estimate the model some reduction of the dimension of the problem is necessary. By controlling for cross section dependence we obtain a likelihood function that is separable and a reduced number of parameters to be estimated. In this paper residual cross section dependence is defined as a non-zero covariance of the residuals from pairs of units of the panel *i.e.* $E(\epsilon_{it}\epsilon_{jt}) \neq 0 \;\; i \neq j$.

In the following the approach pioneered by Pesaran (2006) and Dees et al. (2007) is discussed. It is based on augmenting the model with weighted cross sectional averages. The assumed DGP is similar to that of Dees et al. (2007).

Assumption 1. *The DGP of Y_i is given by the common factor model*

$$\begin{aligned} Y_{it} &= \delta_{i0} + \delta_{i1} t + \gamma_i \mathbf{f}_t + \xi_{it} \\ \Delta\xi_{it} &= \mathbf{\Psi}_i(L)\,\nu_{it}, \quad \nu_{it} \sim \mathcal{N}(0, I_p) \\ \Delta\mathbf{f}_t &= \mathbf{\Lambda}(L)\,\eta_t, \quad \eta_t \sim \mathcal{N}(0, I_p) \end{aligned}$$

where $\mathbf{f}_t$ is a $m_f \times 1$ vector of common unobserved factors, with γ_i the associated $p \times m_f$ matrix of individual loadings. $\mathbf{\Psi}_i(L) = \sum_{l=0}^{\infty} \Psi_l L^l$ and $\mathbf{\Lambda}_i(L) = \sum_{l=0}^{\infty} \Lambda_l L^l$ are lag polynomials composed of absolute summable matrices Ψ_l and Λ_l such that $var(\Delta\xi_{it})$ and $var(\Delta\mathbf{f}_t)$ are positive definite and bounded.

I further assume that $E(\nu_{it}\nu_{jt}) = 0$, that is, the idiosyncratic shocks are cross sectionally uncorrelated.

The data generating process in assumption 1 is a quite general common factor process allowing for m_f common factors and p individual variables integrated of order at most one, with the possibility of cointegration among the individual variables and between those and the common factors. The cross section dependence in the observed data stems entirely from the unobserved common factors.

I now discuss how the common factors can be approximated by the observed variables in order to control for cross section dependence. Construct weighted averages of the data as

$$Y_{it}^* = \mathbf{w}_i Y_t = \sum_{j=1}^{N} w_{ij} Y_{jt} \tag{3}$$

where the weights are defined as follows:

$$\begin{aligned} &i) \;\; w_{ii} = 0 && ii) w_{ij} \in]0,1[\quad \forall i \neq j \\ &iii) \sum_{j=1}^{N} w_{ij} = 1 && iv) w_{ij} = O\left(N^{-1}\right) \end{aligned}$$

The three first conditions ensure that the weights for individual i construct a weighted average of the $Y_{-i} := \{Y_j | j \neq i\}$. The last condition ensures that the average is not dominated by a single individual, so that idiosyncratic dynamics cancel out when N grows large. Dees et al. (2007) show the following

$$\begin{aligned} \sum_{j=1}^{N} w_{ij} Y_{jt} &= \sum_{j=1}^{N} w_{ij} \left(\delta_{j0} + \delta_{j1} t + \gamma_j \mathbf{f}_t + \xi_{jt}\right) \\ Y_{it}^* &= \delta_{i0}^* + \delta_{i1}^* t + \gamma_i^* \mathbf{f}_t + \xi_{it}^* \\ \mathbf{f}_t &\xrightarrow[N]{q.m.} \left(\gamma_i^{*\prime} \gamma_i^*\right)^{-1} \gamma_i^* \left(Y_{it}^* - \delta_{i0}^* - \delta_{i1}^* t - \xi_{it}^*\right) \end{aligned}$$

where $\xrightarrow[N]{q.m.}$ stands for convergence in quadratic mean when N grows large. Note that the weights may be time varying as long as they satisfy the conditions above, but for simplicity of exposition the weights won't be indexed by t. When the number of cross section units becomes large and under assumption 1 the common factors can be approximated by averages of the observed variables. By augmenting the model with weighted averages, the unobserved factors (and hence the cross section

dependence) can be controlled for and we have

$$\begin{aligned}\Delta Y_{i,t} = \alpha_i \rho_i' d_{it} + \psi_i \delta_{it} \quad &+ \quad \alpha_i \beta_i' \left(Y_{i,t-1}', Y_{i,t-1}^{*'}\right)' + \Lambda_{i,0} \Delta Y_{i,t}^* \\ &+ \quad \sum_{l=1}^{k} \Gamma_{i,l} \Delta \left(Y_{i,t-l}', Y_{i,t-l}^{*'}\right)' + \epsilon_{i,t} \end{aligned} \tag{4}$$

where:

$$\mathrm{Cov}(\epsilon_{it}' \epsilon_{jt}) = \mathtt{0_p} \quad \text{for } i \neq j$$

By this transformation we obtain a model with cross sectionally independent innovations. The model for individual i given in (4) is not subject to the curse of dimensionality in the sense that the number of parameters per equation is not a function of N. All the while, it still relates Y_{it} to every other variables in the panel through the weighted averages Y_{it}^*. This provides some interesting properties to the model and deserves further inspection. Consider the vector of length $2p$:

$$Z_{it} \quad = \quad \left(Y_{i,t-1}', Y_{i,t-1}'^{*}\right)' = W_i Y_t$$

The W_i matrix is composed of a first block of p rows with a unit matrix of dimension p between the columns ip and $(i+1)p-1$ and zeros elsewhere. The second block of p rows is composed of unit matrices multiplied by a weight scalar as defined above, except between column ip and $(i+1)p-1$ where it is equal to zero:

$$W_i = \begin{bmatrix} \mathtt{0_p} & \dots & \mathtt{0_p} & \mathtt{I_p} & \mathtt{0_p} & \dots & \mathtt{0_p} \\ w_{i1}\mathtt{I_p} & \dots & w_{ii-1}\mathtt{I_p} & \mathtt{0_p} & w_{ii+1}\mathtt{I_p} & \dots & w_{iN}\mathtt{I_p} \end{bmatrix} \tag{5}$$

This matrix multiplied on Y_t returns a vector of length $2p$ with Y_{it} and the corresponding weighted average Y_{it}^* stacked. Similarly, define the matrix W_{i0}:

$$W_{i0} = \begin{bmatrix} w_{i1}\mathtt{I_p} & \dots & w_{ii-1}\mathtt{I_p} & \mathtt{0_p} & w_{ii+1}\mathtt{I_p} & \dots & w_{iN}\mathtt{I_p} \end{bmatrix} \tag{6}$$

This matrix applied to Y_t returns the ith weighted average Y_{it}^* such that: $Y_{it}^* = W_{i0} Y_t$. Define also $W = [W_1', \dots, W_N']'$, and similarly $W_0 = [W_{10}', \dots, W_{N0}']'$. These matrices are crucial in the formulation of the model, since they provide a link between every unit in the panel while reducing the dimension of the parameter space.

From the individual model given in equation 4 the model for the full panel can be recovered by stacking:

$$\Delta Y_t = \alpha \beta' W Y_{t-1}, + \Lambda_0 W_0 \Delta Y_t + \sum_{l=1}^{k} \Gamma_l \Delta W Y_{t-l} + \alpha \rho' d_t + \psi \delta_t + \epsilon_t \tag{7}$$

where, by construction:

$$\alpha = \begin{bmatrix} \alpha_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_N \end{bmatrix} \beta = \begin{bmatrix} \beta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \beta_N \end{bmatrix} \tag{8}$$

Similarly, the lag matrices Γ_l, Λ_0, and the variance covariance matrix for the full panel are also block diagonal:

$$\Omega \quad = \quad E[\epsilon_t' \epsilon_t] = \begin{bmatrix} \Omega_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \Omega_{NN} \end{bmatrix} \tag{9}$$

The log likelihood of the full panel given in equation 7 is the standard log likelihood function of the Gaussian VAR:

$$\begin{aligned}\mathcal{L}(\alpha,\beta,\Lambda_0,\Gamma_l,\Omega,\rho,\psi) &= -\frac{T}{2}\ln\det(\Omega)-\frac{1}{2}\sum_t\left(\epsilon_t'\Omega^{-1}\epsilon_t\right)\\ &= -\frac{T}{2}\sum_{i=1}^{N}\ln\det(\Omega_{ii})-\frac{1}{2}\sum_{i=1}^{N}\sum_t\left(\epsilon_{it}'\Omega_{ii}^{-1}\epsilon_{it}\right)\\ &= \sum_{i=1}^{N}\mathcal{L}_i(\alpha_i,\beta_i,\Lambda_{0i},\Gamma_{li},\Omega_i,\rho_i,\psi_i)\end{aligned}$$

The dimensions of the parameter matrices of the panel given in equation 7 are functions of N, but are sparse with a known sparsity pattern. Since the transformed panel is, by construction, not subject to residual cross section dependence, its likelihood function is the sum of the likelihood functions of the individual models. This permits independent estimation of the parameters of the individual models by maximum likelihood. The parameters of the full panel can be recovered by manipulation of the estimated parameters of the individual models. In each individual model the weighted averages Y^* are treated as weakly exogenous. However, every variable is endogenous in the full panel. Furthermore, it allows for immediate feedback between the variables of different units through the Λ_0 matrix.

Finally, we can reformulate the model in levels:

$$\begin{aligned}(I_{Np}-\Lambda_0W_0)\,\Delta Y_t &= \alpha\beta'WY_{t-1}+\sum_{l=1}^{k}\Gamma_l\Delta WY_{t-l}+\alpha\rho'd_t+\psi\delta_t+\epsilon_t\\ AY_t &= (A+\alpha\beta'W)\,Y_{t-1}+\sum_{l=1}^{k}\Gamma_l\Delta WY_{t-l}+\alpha\rho'd_t+\psi\delta_t+\epsilon_t\\ Y_t &= (A)^{-1}(A+\alpha\beta'W)\,Y_{t-1}+\sum_{l=1}^{k}(A)^{-1}\Gamma_l\Delta WY_{t-l}\\ &\quad+(A)^{-1}\alpha\rho'd_t+(A)^{-1}\psi\delta_t+(A)^{-1}\epsilon_t\end{aligned}\tag{10}$$

where $A=I_{Np}-\Lambda_0W_0$ is full rank. In the next section this recursion will be used to generate bootstrap data for the full panel.

3. Bootstrap Panel Rank Test

A widely used method to determine the co-integration rank of a system is the likelihood ratio test of Johansen (1995) for the hypothesis that the true co-integration rank r_0 is less than or equal to r against the hypothesis that the system has full rank. For a VAR with p variables the test statistics are computed as $Q_{r,T}:=-T\sum_{i=r+1}^{p}\log\left(1-\hat{\lambda}_i\right)$, where $\hat{\lambda}_1>\cdots>\hat{\lambda}_p$ are the solutions of the following eigenvalue problem:

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

The matrix S_{ij} are defined as $S_{ij}:=T^{-1}\sum_{t=1}^{T}R_{it}R_{jt}'$, with R_{0t} and R_{1t} being ΔY_{t-1} and (Y_{t-1},d_t) corrected for the short run dynamics $(\Delta Y_{t-1},\cdots,\Delta Y_{t-k+1})$ and the short run deterministic component δ_t. The test rejects for large values of the $Q_{r,T}$ statistics, that is, if some λ_i is close to 1.

This likelihood ratio test is embedded in a sequential procedure to select the rank of a VAR. The procedure starts by testing the hypothesis that $r = 0$, then $r \leq 1$ up to $r \leq p-1$, until the test fails to reject one of the hypotheses. If $r = p-1$ is rejected, the rank is set to p. This paper will construct a panel version of the sequential procedure using the bootstrap to compute *p-values* associated with the rank test statistics.

There are two main motivations for using a bootstrap based procedure to construct a panel cointegration rank test:

(1) The bootstrap is expected to yield more accurate finite sample performances than an approach based on the asymptotic distribution of the trace statistic. Cavaliere et al. (2012) shows that it is the case in co-integrated VAR models Since the tests presented in this paper are based on pooling individual tests, it likely to be the case that the bootstrap delivers more accurate inference than an approach based on asymptotic distributions.

(2) The trace test statistic's asymptotic distribution depends on the deterministic components as well as the number of weakly exogenous variables. This distribution needs to be simulated for every specification of model. The bootstrap procedure does not require simulation of the asymptotic distribution of the trace test statistic for the model under investigation, but instead provides an approximation of the cumulative density function of the statistics conditional on the original data. This enables us to propose a test that can accommodate a wide degree of heterogeneity in the structure of the individual models.

Assume a co-integrated VAR of the form of (4) with a co-integration rank r_0. Let the parameters of this model be estimated under an assumed rank $r < r_0$. Denote these parameters $\hat{\theta}^r := \left\{\hat{\alpha}^r, \hat{\beta}^r, \hat{\Gamma}_l^r, \hat{\rho}^r, \hat{\delta}^r\right\}$. One important result of Cavaliere et al. (2012) is to show that bootstrap data generated using the estimated parameters $\hat{\theta}^r$ are integrated of order at most 1 and have r co-integrating relations in the limit. This is an asymptotic result; in finite samples the $I(1, r)$ condition could be violated and thus it needs to be checked, which is the purpose of the second step of the algorithm below.

Before stating the algorithm and the panel bootstrap rank test statistics, some assumptions on the dynamics of the model have to be made.

Assumption 2. *The co-integration rank of model 4 is equal to $r_0 \in 0, \cdots, p$ for all $i \in 1, \cdots, N$, that is, we assume an identical co-integration rank for every individual model of the panel, so that the number of co-integration relations in the panel is equal to Nr_0.*

Assumption 2 is not a testable assumption, but since every individual model contains the same set of variables it seems to be reasonable to postulate that they also posses the same number of long run equilibrium relations.

Assumption 3. *The innovations follow an* i.i.d. *sequence $\{\epsilon_{it}\}$ satisfying:*

$$i) E(\epsilon_{it}) = 0 \qquad iii) E||\epsilon_{it}||^4 \leq K < \infty$$
$$ii) E(\epsilon_{it}\epsilon_{it}') = \Omega_i \qquad iv) E(\epsilon_{it}'\epsilon_{jt}) = 0_p,\ i \neq j.$$

Points i) to iii) of assumption 3 are standard and ensure well behaved residuals. iv) assumes a block diagonal structure of the variance covariance matrix.

Assumption 4. *The dynamics of every individual model satisfy the following:*

(1) The characteristic polynomial of the model described by equation 4 has r_0 roots on the unit circle and $p - r_0$ roots outside.
(2) α_i and β_i have full column rank r_0
(3) $\det(\alpha'_{i,\perp}(I_p - \sum_l \Gamma_{i,l} W)(W'\beta_i)_\perp) \neq 0$

These assumptions ensure that every individual system is composed of variables of integration order at most 1, with r_0 co-integrating relations among them.

Assumption 5. *Distinct eigenvalues: The limiting non-zero eigenvalue solutions of problem 11 are distinct.*

I now introduce the bootstrap algorithm to implement the likelihood ratio test for the hypothesis $\mathcal{H}_(r)$.

Algorithm 1. *(1) Under the assumption that $r_0 = r$, estimate by Gaussian quasi-maximum likelihood the set of parameters θ_i^r and the residuals $\hat{\epsilon}_{r,i,t}$ of equation (4) for each individual model. Compute the individual trace test statistic $Q_{r,i}$ by solving the eigenvalue problem (11).*

(2) Verify that the characteristic polynomial $|\hat{A}^r(z)| = 0$ with:

$$\hat{A}^r(z) = (1-z)I_{Np} - \hat{\alpha}^r \hat{\beta}^{r'} W z - \sum_{l=1}^{k-1} \Gamma_l^{r'} W (1-z) z^l$$

has $Np - Nr_0$ roots equal to 1 and the remaining roots outside the unit circle. If this is the case proceed to the next step, otherwise exit the algorithm.

(3) Recenter the residuals $\hat{\epsilon}^c_{r,i,t} = \hat{\epsilon}_{r,i,t} - \frac{1}{T}\sum_t \hat{\epsilon}_{r,i,t}$ and construct a bootstrap matrix of residuals by resampling from $\hat{\epsilon}^c_{r,i,t}$. Define the bootstrap residuals as $\hat{\epsilon}^\dagger_{r,i,t}$. Using the parameter matrix $\hat{\theta}^r$ and the bootstrap residuals, generate B bootstrap samples for the full panel using the recursion given in (10).

(4) Using the bootstrap sample $Y_t^\dagger$, apply the procedure in the previous section to transform the panel in a set of N independent models. Solve the eigenvalue problem 11 for each individual model and compute the LR statistics $Q^{\dagger,b}_{r,i,T} = -T\sum_{v=r+1}^{p} \log\left(1 - \hat{\lambda}^{\dagger,b}_{v,i}\right)$ where $b \in 1, \cdots, B$ indexes the bootstrap samples.

(5) Compute the $p-value$ for every individual unit: $p^\dagger_{i,r} = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}_{\{Q_{i,r} < Q^{\dagger,b}_{i,r}\}}$ and pool them to obtain a bootstrap panel test statistics:

$$\bar{P}_r = \frac{\sum_{i=1}^{N}\left(-2log(p^\dagger_{i,r}) - 2\right)}{\sqrt{4N}}$$

The asymptotic behavior of this test is discussed in theorem 1.

Remark 1. The trace test statistics $Q_{r,i,T}$ is similar with respect to ρ_i and asymptotically similar with respect to ψ_i. That is, the rejection frequency of the test does not depend on the value of ρ_i nor, asymptotically, on the value of ψ_i (see Nielsen and Rahbek (2000) and Cavaliere et al. (2012), remark 2). Hence the recursion used in step 3 of algorithm 1 can be used excluding the estimated deterministic terms $\hat{\alpha}_i^r \hat{\rho}'_{it} d_t + \hat{\psi}_i \delta_{it}$ and setting the initial values $Y_{i,1-k} = \cdots = Y_{i,0} = 0$.

Theorem 1. *Under assumptions 2 to 5, and if $r = r_0$, we have:*

$$\bar{P}_r \xrightarrow[(N,T)_j]{w} \mathcal{N}(0,1)$$

with no restrictions on the relative rate of convergence of N and T.

The $\bar{P}$ statistic is based on the Fisher test, as discussed in Maddala and Wu (1999). We rely on large N asymptotics for pooling and the fact that the logarithm of the individual $p-values$ (multiplied by -2) is χ^2 distributed. We cannot sum the statistics directly since the moments of this sum would be a function of N. However, we can use a central limit theorem to aggregate the individual $p-values$ and obtain a panel statistic.

Remark 2. The $p-values$ associated with each test statistics $Q_{r,i,T}$ are computed independently based on the B bootstrap statistics $Q^{\dagger,b}_{r,i,T}$ generated with algorithm 1. The trace test statistics is known to have different limiting distribution for different d_t, δ_t, and the number of weakly exogenous variables. The $p-values$ obtained at step 5 of algorithm 1 are *i.i.d.* so that the panel statistic is still valid when different deterministic components are used for different individuals. This is particularly useful when dummy variables are required to account for an exceptional event happening in one individual model.

I now introduce the sequential procedure to determine the co-integration rank of the panel.

Algorithm 2. *Initialize the algorithm with $r = 0$.*

(1) Perform steps 1 to 5 of algorithm 1.
(2) If $p^{\dagger}\left(\bar{P}_r\right)$ exceeds the selected significance level, set $\hat{r}_0 = r$. Otherwise repeat the previous steps for $r+1$ if $r+1 < p$, else set $\hat{r}_0 = p$.

The sequential rank test procedure yields an estimator of the co-integration rank with the following properties:

Theorem 2. *Let $\hat{r}$ be the rank selected by algorithm 2 and η the selected significance level. Then under assumptions 2, 3, 4, and 5:*

$$\begin{aligned}
\lim_{(N,T)_j \to \infty} P(\hat{r} = r) &= 0, \forall r < r_0 \\
\lim_{(N,T)_j \to \infty} P(\hat{r} = r_0) &= \begin{cases} 1-\eta \text{ if } r_0 < p \\ 1 \text{ if } r_0 = p \end{cases} \\
\lim_{(N,T)_j \to \infty} \sup_{r \in \{r_0+1, \cdots, p\}} P(\hat{r} = r) &\leq \eta \text{ for } r > r_0
\end{aligned}$$

The remark below considers the case where assumption 2 is violated. Suppose the panel can be split into two groups of individual with different co-integration ranks:

Remark 3. If assumption 2 is violated such that for a set of units J $r_0 = r_J$ while for J^c, $r_0 = r_{J^c}$ with $r_J > r_{J^C}$ then $P(\hat{r_0} = r_{J^c}) \to 0$ and $P(\hat{r_0} = r_J) \to 1-\eta$. This is due to the fact that the hypothesis used in the sequential procedure is of the form $\mathcal{H}(r) : r_0 \leq r$.

When the co-integration rank of the individual models of the panel are heterogeneous, the panel rank test will asymptotically select the largest co-integration rank.

Testing restrictions on the cointegration vectors β_i and the adjustment matrices α_i can help detect the violation of assumption 2.

4. Monte Carlo

In order to assess the performance of the panel rank test, I conduct a Monte Carlo experiment. The first data generating process considered in these experiments is reduced to the bare minimum: no short run dynamics nor deterministic components are included. The data is generated according to recursion 10 with uniform weights ($w_{ij} = 1/(N-1) \ \ \forall i \neq j$), a unit covariance matrix, and no short run dependence between ΔY_{it} and ΔY_{it}^* (*i.e.* $\Lambda_0 = 0$). The model becomes

$$Y_t = (I_{Np} + \alpha\beta' W)\, Y_{t-1} + \epsilon_t$$

Experiment A considers a system with two variables per individual unit, a single co-integration vector and no cross sectional co-integration. I set $\alpha_i' = [-0.4, 0.4]$ and $\beta_i' = [1, -1, 0, 0]$. The value of the adjustment parameters is similar to that used in Cavaliere et al. (2012). The significance level is 0.05, and all initial values are set to 0. Simulation results are reported in table 1. It reports the selection frequency of the panel co-integration rank test $\bar{P}_r$ used in the sequential procedure described in algorithm 2 for different sample length (T) and cross section dimension (N). It appears from these results that the test performs very well in selecting the correct rank even for samples of only 100 observations. It is also striking that the test is oversized, in the sense that the selection rate of the correct rank is often higher than the 95% that would be expected, and more so when N and T increase.

	N =5			N =10		
T	r =0	r =1	r =2	r =0	r =1	r =2
100	0.0000	0.9790	0.0210	0.0000	0.9810	0.0190
200	0.0000	0.9850	0.0150	0.0000	0.9920	0.0080
500	0.0000	0.9860	0.0140	0.0000	0.9950	0.0050
1000	0.0000	0.9970	0.0030	0.0000	0.9970	0.0030
	N =20			N =50		
	r =0	r =1	r =2	r =0	r =1	r =2
100	0.0000	0.9990	0.0010	0.0000	0.9900	0.0100
200	0.0000	1.0000	0.0000	0.0000	0.9950	0.0050
500	0.0000	0.9990	0.0010	0.0000	1.0000	0.0000
1000	0.0000	0.9970	0.0030	0.0000	1.0000	0.0000

Table 1. $\bar{P}_r$ selection frequency, experiment A, 1000 Monte Carlo replications, 199 bootstrap iterations.

Experiment B considers a larger system with 3 variables per unit and 2 co-integration vectors: 1 among the variables of a given unit, the other between these variables and the rest of the system. I set

$$\alpha_i' = \begin{bmatrix} -0.4 & -0.4 & 0.4 \\ -0.4 & 0 & 0 \end{bmatrix} \beta_i' = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

	N =5				N =10			
T	r =0	r =1	r =2	r =3	r =0	r =1	r =2	r =3
100	0.0000	0.2450	0.6000	0.1550	0.0000	0.0300	0.8950	0.0750
200	0.0000	0.0000	0.8750	0.1250	0.0000	0.0000	0.9500	0.0500
500	0.0000	0.0000	0.9300	0.0700	0.0000	0.0000	0.9850	0.0150
1000	0.0000	0.0000	0.9300	0.0700	0.0000	0.0000	1.0000	0.0000
	N =20				N =50			
T	r =0	r =1	r =2	r =3	r =0	r =1	r =2	r =3
100	0.0000	0.0000	0.9700	0.0300	0.0000	0.0000	0.9750	0.0250
200	0.0000	0.0000	0.9800	0.0200	0.0000	0.0000	1.0000	0.0000
500	0.0000	0.0000	0.9900	0.0100	0.0000	0.0000	1.0000	0.0000
1000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	1.0000	0.0000

TABLE 2. $\bar{P}_r$ selection frequency, experiment B, 1000 Monte Carlo replications, 199 bootstrap iterations.

We also introduce some short run dynamics by setting the contemporaneous dependency $\Lambda_{0i} = -0.5 \times \mathtt{I}_p$ and $\Gamma_{1i} = (0.5 \times \mathtt{I}_p, \mathtt{0}_p)$ for all i, so that the recursion used to generate the data is

$$\begin{aligned} Y_t = &\left(I_{Np} - \Lambda_0 W_0\right)^{-1} \left(I_{Np} + \alpha\beta' W + \Gamma_1 W\right) Y_{t-1} \\ &- \left(I_{Np} - \Lambda_0 W_0\right)^{-1} \Gamma_1 W Y_{t-2} + \left(I_{Np} - \Lambda_0 W_0\right)^{-1} \epsilon_t \end{aligned}$$

As in the previous experiments the results of experiment B, reported in table 2 show that the sequential procedure reveals the correct rank (the hypotheses $r_0 = 0$ and $r_0 \leq 1$ are rejected) very often, except when N and T are small (5 and 100 respectively). As in the previous experiment, the test rejects hypotheses where $r < r_0$ in every instances when T is greater than 100. The test is oversized, selecting the true rank in more than 99% of the cases.

5. Application

To illustrate the use of the sequential panel cointegration rank test, I propose to apply the test to a set of interest rates and inflation variables from the Euro Zone. In 1993 the Maastricht Treaty was ratified and enforced by a number of European states. It provided for the establishment of a common currency among countries of the European Union that was to be launched on 1 January 1999. Interest rates reacted to this fact by converging but rates of inflation did not converge as can be seen in figure 1, where CPI is the consumer price index, LIR the 10-year state obligation rate and SIR the 3-month interest rate. The convergence of nominal interest rates is predicted by a simple no-arbitrage conditions. However, the sustained divergence in inflation rate ensured that real interest rates were not converging in the Euro zone, as figure 2 illustrates, with $DCPI$ the monthly inflation rate, and $RSIR$ the real 3-month interest rate (monthly).

Given the degree of integration of capital markets in the Euro zone, it would be unreasonable to investigate the dynamics of each country separately without taking into account the dependencies among them. The procedure described above

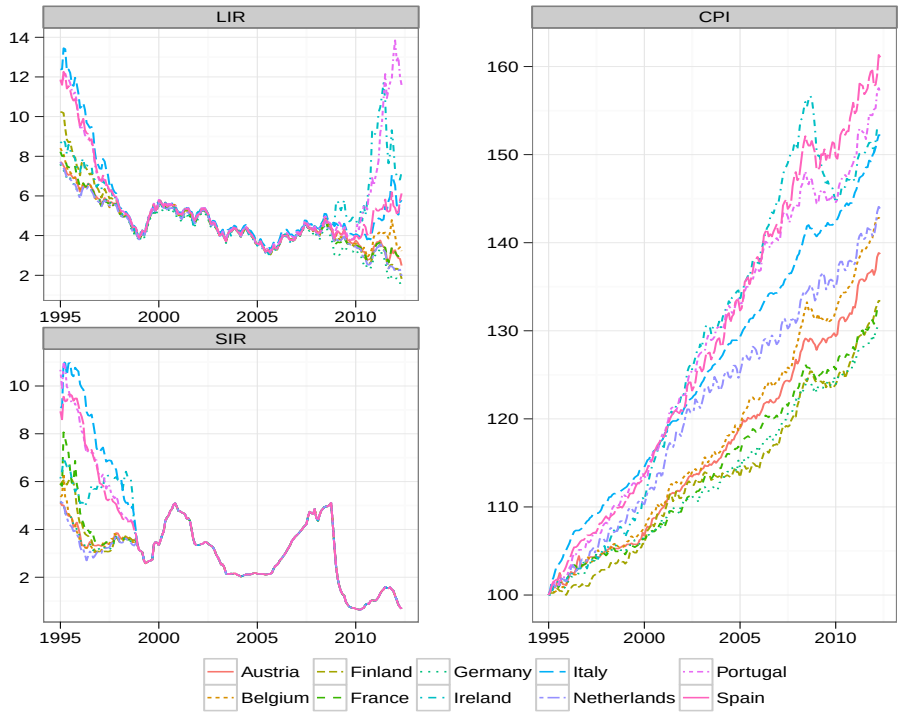

FIGURE 1. Interest Rates and CPI

to model cross section dependence will enable us to explicitly model the dependency between a member country and the rest of the currency area. The purpose of this application is to investigate whether some long run equilibrium (cointegration) exists between domestic and foreign interest rates and inflation rates in the Euro zone.

The data set covers 10 of the 11 founding countries of the Euro zone (Luxembourg being excluded) sampled from January 1995 to December 2011. Data were retrieved from the OECD statistical database in August 2012. The data is composed of three monthly series for short and long interest rates (3 month EURIBOR, noted r_{it}, and 10 years government bonds, noted b_{it} respectively) as well as the inflation rate (noted π_{it}) based on the harmonized consumer price index for each country (indexed by i), The weights used to construct the cross section averages are based on the ECB's *weight of the countries within the euro area* series.

We start by investigating systems comprising the short and long term interest rate (in *per cent* per month) as well as the monthly change in the consumer price index. The model is specified with a restricted constant and one lagged first difference of the dependent variables. Inclusion of seasonal dummies was tried, but this did not modify the results significantly. The individual model is of the form:

$$\Delta Y_{i,t} = \alpha_i d_i + \alpha_i \beta_i' \left(Y_{i,t-1}', Y_{i,t-1}^{*\prime} \right)' + \Lambda_{0i} Y_{i,t-1}^{*} + \Gamma_{1i} \Delta \left(Y_{i,t-1}', Y_{i,t-1}^{*\prime} \right)' \tag{12}$$

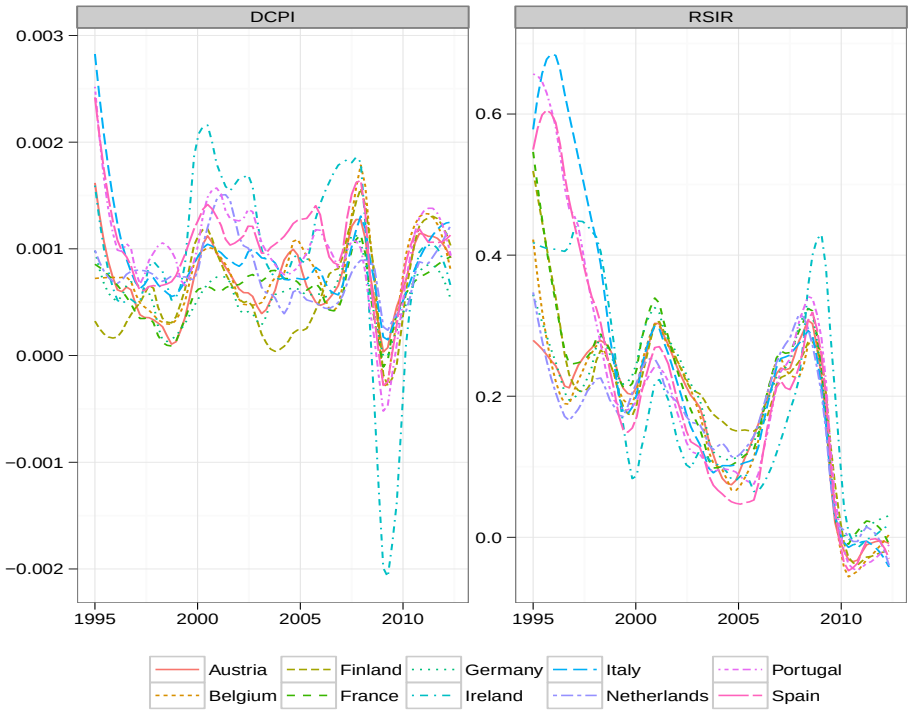

FIGURE 2. Real Short Interest Rate and Monthly Inflation (smoothed)

with $Y'_{i,t} = [r_{it}, b_{it}, \pi_{it}]$, and d_i being a constant restricted to lie in the cointegration space.

The individual rank tests reported in Table 3 show that the sequential procedure selects a variety of co-integration ranks: 5 countries are found to have rank two, 3 to have rank one and two to have rank three. The specification tests reported at the bottom of Table 3 show that there is no residual autocorrelation and that cross section dependence has been properly captured except for the b_t series where there is still residual cross section dependence. The panel statistics reported in Table 4 selects clearly a rank of 2, strongly rejecting the hypothesis that the rank is 0 or 1. This result is consistent with the interest rate parity condition.

We then investigate a system where both interest rates have been transformed from nominal to real by subtracting the inflation rate. Each individual system is now composed of 4 variables, the two domestic real interest rates and the two foreign real interest rates. The results of the panel cointegration rank tests are reported in Table 5. The cointegration rank of this new system is found to be 1, indicating that real interest rates (domestic and foreign, long and short) share a common stochastic trend.

Country	Rank	$\mathcal{H}(0)$	$\mathcal{H}(1)$	$\mathcal{H}(2)$	$p(\mathcal{H}(0))$	$p(\mathcal{H}(1))$	$p(\mathcal{H}(2))$
		Likelihood Ratio				$p-value$	
Austria	2	159.72	47.40	7.25	0.001	0.004	0.868
Belgium	3	214.50	88.07	28.22	0.001	0.001	0.031
Finland	2	190.94	64.74	15.57	0.001	0.001	0.206
France	2	140.62	45.63	11.64	0.001	0.015	0.691
Germany	1	117.23	10.80	2.39	0.031	0.988	0.994
Ireland	2	160.06	49.32	6.74	0.001	0.006	0.821
Italy	1	174.21	31.05	4.48	0.001	0.358	0.982
Netherlands	1	145.91	30.54	9.74	0.001	0.274	0.808
Portugal	2	155.42	50.54	5.78	0.001	0.005	0.932
Spain	3	168.99	52.93	23.85	0.001	0.004	0.023

Specification Tests

	Cross Section Dependence			Autocorrelation
	r_t	b_t	π_t	Ljung-Box
Statistics	0.434	2.822	0.300	-0.672
$p-value$	0.332	0.002	0.382	0.274

TABLE 3. $p-values$ for the individual cointegration rank tests, 999 Bootstrap iterations.

	$\mathcal{H}(0)$	$\mathcal{H}(1)$	$\mathcal{H}(2)$
Likelihood Ratio	17.596	10.058	-0.050
$p-value$	0.00	0.00	0.52

TABLE 4. Panel Cointegration Rank Test, 999 Bootstrap iterations

	$\mathcal{H}(0)$	$\mathcal{H}(1)$
Likelihood Ratio	11.060	-0.022
$p-value$	0.00	0.51

TABLE 5. Panel Cointegration Rank Test, 999 Bootstrap iterations

6. Conclusion

The aim of this paper was to construct a sequential bootstrap co-integration rank test procedure for panels of co-integrated VAR models with large cross-section dimension and cross section dependence. I adapt the bootstrap sequential rank test procedure by Cavaliere, Rahbek, and Taylor (2012) to panel data and use results by Dees et al. (2007) to control for cross section dependence while reducing the dimension of the parameter space relative to that of the unrestricted panel VAR. I introduced a bootstrap panel test statistic, the $\bar{P}_r$ statistics. The $\bar{P}_r$ statistics allows for a great amount of heterogeneity across models both in terms of parameter values and in the specification of deterministic components. A Monte Carlo experiment shows that this test performs very closely to its nominal size in finite samples while being often oversized. The test relies on the assumption that the co-integration

rank is homogeneous across units of the panel. Further work would involve finding a way to test this hypothesis.

References

Breitung, J. and M. Pesaran (2008): "Unit roots and cointegration in panels," *The Econometrics of Panel Data*, 279–322.

Cavaliere, G., A. Rahbek, and A. Taylor (2012): "Bootstrap Determination of the Co-Integration Rank in Vector Autoregressive Models," *Econometrica*, 80, 1721–1740.

Dees, S., F. Mauro, M. Pesaran, and L. Smith (2007): "Exploring the international linkages of the euro area: a global VAR analysis," *Journal of Applied Econometrics*, 22, 1–38.

Johansen, S. (1988): "Statistical analysis of cointegration vectors," *Journal of economic dynamics and control*, 12, 231–254.

——— (1995): *Likelihood-based inference in cointegrated vector autoregressive models*, Oxford University Press, USA.

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

Kapetanios, G., M. Pesaran, and T. Yamagata (2011): "Panels with nonstationary multifactor error structures," *Journal of Econometrics*, 160, 326–348.

Larsson, R., J. Lyhagen, and M. Lothgren (2001): "Likelihood-based cointegration tests in heterogeneous panels," *The Econometrics Journal*, 4, 109–142.

Maddala, G. and S. Wu (1999): "A comparative study of unit root tests with panel data and a new simple test," *Oxford Bulletin of Economics and statistics*, 61, 631–652.

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

Pesaran, M. (2006): "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica*, 74, 967–1012.

Reinsel, G. and S. Ahn (1992): "Vector Autoregressive Models With Unit Roots And Reduced Rank Structure: Estimation. Likelihood Ratio Test, And Forecasting," *Journal of Time Series Analysis*, 13, 353–375.

Proof. Proof of Theorem 1:

Consider the recursion used in step 3 of algorithm 1:

$$\begin{aligned} Y_t = (A)^{-1}(A+\alpha\beta' W)Y_{t-1}, + \sum_{l=1}^{k}(A)^{-1}\Gamma_l \Delta W Y_{t-l} \\ + (A)^{-1}\alpha\rho' d_t + (A)^{-1}\psi\delta_t + (A)^{-1}\epsilon_t \end{aligned} \tag{13}$$

with $A = I_{Np} - \Lambda_0$. As shown in section 2, this recursion is equivalent to N independent models of the form:

$$\begin{aligned} \Delta Y_{i,t} = \alpha_i\rho_i' d_{it} + \psi_i\delta_{it} \quad &+ \quad \alpha_i\beta_i'\left(Y_{i,t-1}', Y_{i,t-1}^{*'}\right)' + \Lambda_{i,0}\Delta Y_{i,t}^* \\ &+ \quad \sum_{l=1}^{k}\Gamma_{i,l}\Delta\left(Y_{i,t-l}', Y_{i,t-l}^{*'}\right) + \epsilon_{i,t} \end{aligned} \tag{14}$$

where $Y_{i,t}^*$ is constructed as in (3). Using the estimated parameters $\hat{\theta}_i^r$ to generate data according to 13 and for any $r \leq r_0$, lemma 1 of Cavaliere et al. (2012) shows that as $T \to \infty$, $\hat{\theta}_i^r \to \theta_i^r$ where θ_i^r satisfy the I(1,r) condition for all $i \in 1, \cdots, N$.

By proposition 2 in Cavaliere et al. (2012), $Q_{r,i,T}^{\dagger} \xrightarrow{w.p.} tr(Q_{i,r,\infty})$ where

$$Q_{r,i,\infty} := \int_0^1 dB_{p-r}(u)F_{p-r}(u)'\left(\int_0^1 dF_{p-r}(u)F_{p-r}(u)'du\right)\int_0^1 dF_{p-r}(u)B_{p-r}(u)'$$

with B_{p-r} being a standard Brownian motion of dimension $p-r$ and:

$$\begin{aligned} F_{p-r} &:= B_{p-r} && \text{no deterministics} \\ F_{p-r} &:= (B_{p-r}', 1)' && \text{restricted constant} \\ F_{p-r} &:= (B_{p-r}', u|1)' && \text{restricted linear trend} \end{aligned}$$

where $x|y$ denotes the projection of the residuals of x on y. Since $Q_{r,i,T} \to \infty$ if $r < r_0$ and $Q_{r_0,i,T} \xrightarrow{w} tr\left(Q_{i,r_0,\infty}\right)$ it follows that for $p_{i,r}^{\dagger} = \frac{1}{B}\sum_{b=1}^{B}\mathbb{1}_{\{Q_{i,r,T} < Q_{i,r,T}^{\dagger,b}\}}$ we have:

$$\begin{aligned} p^{\dagger}(Q_{i,r,T}) &\xrightarrow{p} 0 && \text{if } r < r_0 \\ p^{\dagger}(Q_{i,r,T}) &\sim \mathcal{U}[0,1] && \text{if } r = r_0 \end{aligned}$$

The logarithm of the $p-values$ (scaled by -2) are χ^2 distributed and can be pooled across individuals:

$$\bar{P}_r = \frac{\sum_{i=1}^{N}\left(-2log(p_{i,r}^{\dagger}) - 2\right)}{\sqrt{4N}}$$

and by central limit theorem it follows that:

$$\begin{aligned} \bar{P}_r &\xrightarrow[(N,T)_j]{p} -\infty \\ \bar{P}_r &\xrightarrow[(N,T)_j]{w} \mathcal{N}(0,1) \end{aligned}$$

□

Proof. Proof of Theorem 2: Let η be the chosen significance level of the LR test. To prove theorem 2, three cases have to be considered:

(1) $\hat{r} < r_0$: from theorem 1, $\bar{P}_r \xrightarrow[(N,T)_j]{} -\infty$.

(2) $\hat{r} = r_0$: from theorem 1, $\bar{P}_r \xrightarrow[(N,T)_j]{w} \mathcal{N}(0,1)$. In this case, if $r_0 < p$ it follows that $P(\hat{r} = r_0) = 1 - \eta$. If $r_0 = p$, and since $P(\hat{r} < r_0) \xrightarrow[(N,T)_j]{p} 0$ it follows that $P(\hat{r} = r_0) \xrightarrow[(N,T)_j]{} 1$.

(3) $\hat{r} > r_0$ and $r_0 < p$: $\sup_{p \geq r > r_0} P(\hat{r} = r) \xrightarrow[(N,T)_j]{} \eta$ follows from the previous point.

□