Inference about the rank of cointegration of a locally trending VAR process

Guillaume Chevillon*

ESSEC Business School, Paris

and CREST-INSEE

September, 2008

Abstract

Standard tests for the rank of cointegration of a vector autoregressive (VAR) process present nonstandard distributions that are affected by the presence of deterministic trends. When it is known that the data generating process exhibits a linear trend, it is preferable to model it explicitly and use the corresponding form of the Likelihood Ratio (LR) or Lagrange Multiplier (LM) test statistics. In the presence of stochastic nonstationarity, such as in the cointegrated VAR, deterministic linear trends may be present in the data but with an effect that is not strong enough to be noticeable. Such a situation can be modeled by means of a cointegrated VAR process that exhibits a local linear trend that has same asymptotic magnitude as the common stochastic trends. We derive the properties of the LR and LM tests in this context. We show that whether the trend is orthogonal to the cointegrating vector has a major impact on the distributions. The LR statistics without unrestricted constant are mildlly powerful against the local alternative. The LR with unrestricted trend and the LM are robust towards local trends, but other specifications of the LR perform badly.

Keywords: Cointegration, Deterministic trend, Likelihood ratio, Lagrange Multiplier, Local trends.
JEL codes: C12, C32.

*Corresponding e-mail: chevillon@essec.fr; corresponding address: ESSEC, Avenue Bernard Hirsch, BP 105, 95021 Cergy-Pontoise Cedex, France. Research support from the Economics Department of Oxford University is gratefully acknowledged.

1 Introduction

There has been a recent renewed interest in designing testing strategies for unit roots that are robust to the possible presence or a linear trend, see Harvey, Leybourne and Taylor (2008) and the multivariate extension in Lütkepohl and Demetrescu (2008). These studies draw on the long established difficulties in distinguishing in finite samples between stochastic and deterministic trends, see e.g. see Sampson (1991) and Murray and Nelson (2000) for an empirical example. In a multivariate context, the difficulties are compounded by the presence of many nuisance parameters and many analyses have focused on their influences, see Hubrich, Lütkepohl and Saikkonen (2001) for an overview.

Also, in the joint occurrence of stochastic and deterministic trends, the latter can be restricted or not to lie within the space spanned by the cointegrating vectors. This is the reason why Perron and Campbell (1993) distinguished between "stochastic" and "deterministic" cointegration: only in the latter is the trend orthogonal to the cointegrating vector. In a simulation experiment, Toda (1994) showed that the likelihood ratio (LR) test (see Johansen, 1988 and 1991) can be strongly affected by nuisance parameters when a trend is also present. This realization has led H. Lütkepohl and P. Saikkonen to propose in a series of papers (in 1999 and 2000) a Lagrange Multiplier (LM) test which estimates the deterministic parameters under the null and proceeds to correct for them. In parallel, S. Johansen has suggested, also in a series of papers (in 2000 and 2002) a Bartlett correction for the LR test in finite samples. This correction works well in the presence of deterministic cointegration when, as in Nielsen and Rahbek (2000), the parameters are restricted so that similarity of the tests results. Unfortunately, in finite samples and in the presence of stochastic cointegration, Chevillon (2008) showed that the finite sample distributions are affected by the parameters of the deterministic components.

In view of these difficulties, Lütkepohl and Demetrescu (2008) have proposed extending the work by Harvey *et al.* (2008) to the vector autoregressive (VAR) process. This technique consists in estimating the two models with a deterministic trend restricted or not to lie in the space orthogonal to the cointegrating vector and reject the null if either statistic is significant. These authors show that their methodology compares advantageously to pretesting for the correct trend specification.

In this paper, we derive the distribution of the LR and LM tests in the presence of a, possibly misspecified, local deterministic trend. The latter has a parameter that is asymptotically vanishing at the rate $O\left(T^{-1/2}\right)$ so that both the stochastic and deterministic trends interact asymptotically. This allows us to analyze the robustness of the test for the rank of cointegration when a linear trend is present in the data but whose impact is small and possibly goes unnoticed. We show that very different behaviors result, depending on whether the data are stochastically or deterministically cointegrated. In a Monte Carlo simulation, we also observe that the LR statistic with a restricted trend is not as robust in finite samples as asymptotically. By contrast, the LR statistic that corrects for an unrestricted trend and the LM are, as expected, robust to the local trend. The LR without deterministic element exhibits some power against a misspecified local trend; hence it may be useful to use it as a specification tool by comparison with the other statistics.

The paper is organized as follows. Section 2 present the model and local asymptotic framework. We then derive in section 3 the distributions of the various statistics for the tests on the cointegration rank. A Monte Carlo analysis follows in section 4 and the last section concludes. An appendix collects the proofs. Throughout the paper, row vectors are denoted by $(a : b)$; also, for any $(p \times q)$ matrix $\boldsymbol{\alpha}$ of full rank, we define $\boldsymbol{\alpha}_\perp$ of dimension $p \times (p - q)$ such that $(\boldsymbol{\alpha} : \boldsymbol{\alpha}_\perp)$ is of full rank. We also let the generalized projection operator $\overline{\boldsymbol{\alpha}} = \boldsymbol{\alpha} (\boldsymbol{\alpha}'\boldsymbol{\alpha})^{-1}$.

2 The model

Consider a p-dimensional vector of variables $\mathbf{x}_t$ that admits a vector autoregressive representation of order k such that, for $t = 1, ..., T$,

$$\Delta \mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{x}_{t-i} + \boldsymbol{\epsilon}_t. \tag{1}$$

Assume that the disturbances $\boldsymbol{\epsilon}_t$ follow a martingale difference sequence with bounded fourth moments and variance covariance given by $\mathbf{\Omega} = \mathbf{\Sigma}\mathbf{\Sigma}'$ for some positive definite matrix $\mathbf{\Sigma}$. If $\mathbf{x}_t$ is $\mathsf{I}(1)$ and $\mathbf{\Pi}$ is of reduced rank q, then there exist $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ of order $(p \times q)$ such that $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ and that $\boldsymbol{\beta}'\mathbf{x}_t - \mathsf{E}\left[\boldsymbol{\beta}'\mathbf{x}_t\right]$ is stationary. $\mathbf{x}_t$ is then said to cointegrate, with cointegrating vector $\boldsymbol{\beta}$. We also let $\mathbf{x}_0 = \mathbf{0}$ in (1), although this is not an unconsequential assumption (see Müller and Elliott, 2003). We use the notation in Lütkepohl and Saikkonen (2000) and define $\mathbf{y}_t$ as the sum of $\mathbf{x}_t$ and of a deterministic trend which we assume local

$$\mathbf{y}_t = \mathbf{x}_t + \boldsymbol{\mu} + \boldsymbol{\psi} \frac{t}{\sqrt{T}} = \mathbf{x}_t + \mathbf{\Psi}\mathbf{d}_t \tag{2}$$

where $\mathbf{\Psi} = (\boldsymbol{\mu} : \boldsymbol{\psi})$ is a matrix of dimension $p \times 2$, $\mathbf{d}_t = \left(1 : T^{-1/2}t\right)'$. Then $\mathbf{y}_t$ admits the following moving average representation (see Johansen, 1995, theorem 4.2):

$$\mathbf{y}_t = \mathbf{C} \sum_{i=1}^{t} \boldsymbol{\epsilon}_i + \mathbf{\Psi}\mathbf{d}_t + \mathbf{C}_1 (L) \boldsymbol{\epsilon}_t + \mathbf{A}_t, \tag{3}$$

where L is the lag operator, $\mathbf{C} = \boldsymbol{\beta}_\perp \left(\boldsymbol{\alpha}_\perp' \mathbf{\Gamma} \boldsymbol{\beta}_\perp\right)^{-1} \boldsymbol{\alpha}_\perp'$, $\mathbf{\Gamma} = \mathbf{I}_p - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i$, the power series for $\mathbf{C}_1 (z)$ is convergent for $|z| < 1 + \delta$ for some $\delta > 0$ and $\mathbf{A}_t$ is a stationary process that depends on initial values such that $\boldsymbol{\beta}'\mathbf{A}_t = 0$.

In the model above we only consider local deterministic trends since our purpose is to study the robustness of the cointegration tests in the presence of potential deterministic misspecification.

This differs from analyses such as in Johansen (1995), chapter 14, Rahbek (1994) and Saikkonen and Lütkepohl (1999) where it is the power of the test for the rank of cointegration vis-à-vis a locally larger rank. The interaction between both local cointegration and local deterministic trends has been studied by Chevillon (2008) who focuses on the finite sample nonsimilarity of the Likelihood Ratio test with respect to the coefficients of the linear trend. Also, as noted by Lütkepohl and Saikkonen (2000), the reason why we represent the trend as additive is to preclude a potential quadratic trend in the data that may arise from the model

$$\Delta \mathbf{y}_t^{\dagger} = \mathbf{\Pi} \mathbf{y}_{t-1}^{\dagger} + \mathbf{\Psi} \left(T^{-1/2} : T^{-3/2} t \right)' + \boldsymbol{\epsilon}_t. \tag{4}$$

A well-known solution (see Johansen, 1994) is to restrict the linear trend to lie within the cointegrating space and let a constant enter unrestrictedly. Lütkepohl and Saikkonen (2000), expression (2.7), show that (1) rewrites as

$$\Delta \mathbf{y}_t = \boldsymbol{\nu}_T + \boldsymbol{\alpha} \left(\boldsymbol{\beta}' \mathbf{y}_{t-1} - \boldsymbol{\delta}_T (t-1) \right) + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\epsilon}_t \tag{5}$$

where $\boldsymbol{\nu}_T = -\mathbf{\Pi}\boldsymbol{\mu} + T^{-1/2}\mathbf{\Gamma}\boldsymbol{\psi}$ and $\boldsymbol{\delta}_T = T^{-1/2}\boldsymbol{\beta}'\boldsymbol{\psi}$. Unfortunately, such a specification does not explicitly model the linear trend that lies in the space spanned by $\boldsymbol{\alpha}_{\perp}$, i.e. that interacts with the common stochastic trends. This has significant impact in our analysis and motivates our choice of representation. We follow in this Lütkepohl and Saikkonen (2000).

The asymptotic distribution of $\mathbf{y}_t$ follows a straightforward multivariate extension of the random walk with a local drift as in Chevillon (2008) who draws on Haldrup and Hylleberg (1995) and Stock and Watson (1996). For this, we let, as usual, $[w]$ denote the integer part of w for any real scalar w. Define, then, $\mathbf{U}_T$ in $D^p[0,1]$, the space of $\mathbb{R}^p$-valued functions on the interval $[0,1]$ which are right continuous and have finite left limits (càdlàg). Hence $\forall r \in [0,1]$, $\mathbf{U}_T(r) = T^{-1/2} \sum_{i=0}^{[Tr]} \boldsymbol{\epsilon}_i \Rightarrow \mathbf{\Sigma W}(r)$, as $T \to \infty$, where '$\Rightarrow$' denotes weak convergence of the associated probability measure, and $\mathbf{W}$ is a standard Brownian motion on $C^p[0,1]$, the subspace of $D^p[0,1]$ of continuous functions.

Then $T^{-1/2}\mathbf{y}_{[Tr]}$ retains asymptotically the sum of both the stochastic and of the degenerate linear trend for, i.e. for $r \in [0,1]$:

$$T^{-1/2}\mathbf{y}_{[Tr]} \Rightarrow \mathbf{C\Sigma W}(r) + \boldsymbol{\psi} r \overset{def}{=} \mathbf{K}_{\boldsymbol{\psi},\mathbf{C\Sigma}}(r) \tag{6}$$

where $\mathbf{K}_{\boldsymbol{\psi},\mathbf{C\Sigma}}$ is a Brownian motion with drift.

Motivated by expressions (5) and (6), and by our interest in allowing for a local trend, we assume in the rest of the paper that $\mathbf{y}_0 = 0$, i.e. $\boldsymbol{\mu} = \mathbf{0}$.

3 Cointegration tests

In this section, we derive the distributions of test statistics in the presence of local trends. We aim to establish the properties of the types of tests for inference about the rank of cointegration of the vector $\mathbf{y}_t$. The null hypothesis is that of $\mathbf{\Pi}$ in (1) being of rank $q \leq p$. We assume that this null is true but that the model that is used is misspecified. Indeed the modeler assumes that the observables are $\{\mathbf{x}_t\}$ where she is in fact dealing with $\{\mathbf{y}_t\}$, hence erroneously assuming no deterministic terms in the DGP. For simplicity we let the modeler be mistaken about $\boldsymbol{\mu}$ being potentially non zero, but this matters less than the important aspect whether $\boldsymbol{\psi}$ is zero. In this setting, we are not analyzing the power of the test for the rank of cointegration, but its local robustness in the presence of misspecified deterministic trends: whether correct inference about the rank of $\mathbf{\Pi}$ is achieved. We refer to power against the local trend when the probability that the test statistic is significant tends to 1 under the local alternative.

3.1 Likelihood ratio test

Assume that the modeler wrongly assumes that the DGP follows equation (1) so that no deterministic component is included in the model. Then reduced rank regression of $\Delta \mathbf{x}_t$ on $\mathbf{x}_{t-1}$ corrected for the lagged differences leads to computing the likelihood ratio test statistic

$$-2 \log Q \left(H \left(q\right) | H \left(p\right)\right) = -T \sum_{i=q+1}^{p} \log \left(1 - \widehat{\lambda}_i\right) \tag{7}$$

where the eigenvalues λ_i are estimated as solutions to the problem $\left|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}\right| = \left|S\left(\lambda\right)\right| = 0$, with $S_{ij} = T^{-1} \sum_{i=1}^{T} R_{it} R_{jt}'$, $R_{it} = Z_{it} - M_{i2} M_{22}^{-1} Z_{2t}$, $M_{ij} = T^{-1} \sum_{i=1}^{T} Z_{it} Z_{jt}'$, $Z_{0t} = \Delta \mathbf{y}_t$, $Z_{1t} = \mathbf{y}_{t-1}$ and Z_{2t} is made of the stacked lagged differences of $\mathbf{y}_t$. Alternatively, the modeler may wish to use another of the LR statistics that have been proposed in the literature. The purpose of these is to take into account various assumptions about the deterministic terms that are present in the data. Under the null that $\mathbf{\Psi} = \mathbf{0}$, using another statistic than in expression (7) would imply a loss of power against an alternative rank of cointegration. But a modeler may be willing to trade in this loss for a gain in robustness against misspecifying $\mathbf{\Psi}$. Following the definitions in Johansen (1995), section 5.7, but with different notation, we define the statistic in (7) as LR for the hypothesis $\mathsf{H} : \left(rk\left(\mathbf{\Pi}\right), \mathbf{\Psi}\right) = \left(q, \mathbf{0}\right)$. Two main alternatives are (i) also including a constant in Z_{2t}, which provides the statistic LR_1, or (ii) including both a constant and a linear trend in Z_{2t}, thus yielding LR_2. The underlying rationale for such statistics is that $\mathbf{\Psi}$ may be nonzero. Then LR_1 should be robust against hypotheses such as $\mathsf{H}_1 : \mathbf{\Psi} = \left(\boldsymbol{\mu}, \boldsymbol{\psi}\right)$, $\boldsymbol{\psi} = \mathbf{0}$ and LR_2 against the presence of a linear trend. As is well known the latter assumptions might be too strong, hence the two other assumptions form: H^* or H_1^* in which cases Z_{1t} is augmented of a constant or a linear trend, respectively, and Z_{2t} contains, no deterministic terms or a constant, yielding LR^* and LR_1^*.

These assumptions are specific to the error correction form (4) where they consist in restricting, respectively, the drift (H^*) or the trend (H_1^*) to the cointegrating space.

Let the $(p-q)$-variate diffusion, for $r \in [0,1]$,

$$\mathbf{G}(r) = \mathbf{V}(r) + \begin{bmatrix} \mathbf{0}_{(p-q-1)\times 1} \\ \left(\overline{\boldsymbol{\psi}}_{\beta_\perp}' \mathbf{C}\boldsymbol{\Omega}'\mathbf{C}'\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)^{-1/2} r \end{bmatrix} \tag{8}$$

where $\mathbf{V}(r)$ is a standard Brownian motion of dimension $(p-q)$ and $\overline{\boldsymbol{\psi}} = \boldsymbol{\psi}\left(\boldsymbol{\psi}'\boldsymbol{\psi}\right)^{-1}$ and the notation $\boldsymbol{\psi}_{\beta_\perp}$ refers to the decomposition $\boldsymbol{\psi} = \boldsymbol{\psi}_\beta + \boldsymbol{\psi}_{\beta_\perp}$ alongside the two orthogonal supplements. We also define $\mathbf{G}_1(r) = \mathbf{G}(r) - \int_0^1 \mathbf{G}(u)\,du = \mathbf{V}(r) - \int_0^1 \mathbf{V}(u)\,du + \left[\overline{\boldsymbol{\psi}}_{\beta_\perp}' \mathbf{C}\boldsymbol{\Omega}\mathbf{C}'\overline{\boldsymbol{\psi}}_{\beta_\perp}\right]^{-1/2}(r-1/2)$, and $\mathbf{G}_2(r) = \mathbf{V}(r) - \mathbf{a}_v - \mathbf{b}_v r$ where $\mathbf{a}_v, \mathbf{b}_v$ are coefficients correction $\mathbf{V}$ for a constant and a trend. Finally, we denote by $\mathbf{J}_{\boldsymbol{\psi}_{\alpha_\perp}}(r) = \left(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}\boldsymbol{\alpha}_\perp\right)^{-1/2}\boldsymbol{\alpha}_\perp'\mathbf{K}_{\boldsymbol{\psi},\boldsymbol{\Sigma}}(r) = \mathbf{V}(r) + \left(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}\boldsymbol{\alpha}_\perp\right)^{-1/2}\boldsymbol{\alpha}_\perp'\boldsymbol{\psi}_{\alpha_\perp} r$.

Proposition 1 *Under the assumption that the DGP is generated as (1) and (2), with* $\mathbf{y}_0 = \mathbf{0}$, *then the asymptotic distributions of the Likelihood Ratio test statistics* LR_j^i, *for* $h \in \{\emptyset, 1, 2\}$ *and* $m \in \{\emptyset, *\}$, *with* $(h,m) \notin \{(\emptyset,\emptyset), (2,*)\}$, *is given by*

$$\begin{aligned} \mathsf{LR}^m &\Rightarrow tr\left\{\int_0^1 \mathbf{G}^m\left(d\mathbf{J}_{\boldsymbol{\psi}_{\alpha_\perp}}\right)'\left[\int_0^1 \mathbf{G}^m\mathbf{G}^{m\prime}\right]^{-1}\int_0^1\left(d\mathbf{J}_{\boldsymbol{\psi}_{\alpha_\perp}}\right)(\mathbf{G}^m)'\right\} \\ \mathsf{LR}_h^m &\Rightarrow tr\left\{\int_0^1 \mathbf{G}_h^m(d\mathbf{V})'\left[\int_0^1 \mathbf{G}_h^m\mathbf{G}_h^{m\prime}\right]^{-1}\int_0^1 (d\mathbf{V})(\mathbf{G}_h^m)'\right\},\ \textit{if } h \neq \emptyset \end{aligned} \tag{9}$$

for all $\boldsymbol{\psi}_\beta$ *under hypothesis* H_2 *and for* $\boldsymbol{\psi}_\beta = \mathbf{0}$ *under the other hypotheses.*

The presence of a local trend brings a major departure from the null inasmuch as the distributions of the LR statistics no longer depend only on the number of common trends $p-q$ irrespective of the number of cointegrating relations. As seen in proposition 1, the distributions depend on $\boldsymbol{\beta}'\boldsymbol{\psi}$ and $\boldsymbol{\beta}_\perp'\boldsymbol{\psi}$ which vary, at least in dimension, with q.

According to the proposition, the formulation of the statistic under hypothesis H is be locally non robust for all cases where $\boldsymbol{\psi} \neq \mathbf{0}$ except under H_2 where the distribution is asymptotically similar with respect to $\boldsymbol{\psi}$ (see also Nielsen and Rahbek, 2000). But the robustness of LR_2, and potentially LR_1^*, can come at the cost of a loss of power. The cases considered in this proposition are those that Perron and Campbell (1993) labeled "deterministic" cointegration and that can be treated using Johansen (1991). In the case of a linear trend which is not entirely contained in the space spanned by $\boldsymbol{\beta}_\perp$, i.e. in the presence of "stochastic" cointegration, the statistic LR_2 is not affected, but it is not the case under the other assumptions.

The above proposition does not cover hypotheses other than H_2 for stochastic cointegration. This is because the deterministic components are the underspecified. Indeed, for $\boldsymbol{\psi}_\beta \neq \mathbf{0}$, it is shown in the appendix that under assumption H, the statistic LR is the sum of the smallest $p-q$ eigenvalues of the matrix

$$\mathbf{S} = \mathbf{N}^{-1}\mathbf{M} + o_p(1)$$

where the matrices $\mathbf{N}$ and $\mathbf{M}$ are random. In particular

$$\mathsf{E}\left[\mathbf{N}\right] = -\begin{bmatrix} \frac{1}{6}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \\ \mathbf{0} \\ 1 \end{bmatrix}\begin{bmatrix} \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \\ \mathbf{0} \\ 1 \end{bmatrix}' + \begin{bmatrix} 0 \\ \overline{\boldsymbol{\gamma}}'\mathbf{C\Sigma} \\ \overline{\boldsymbol{\psi}}_{\beta_\perp}'\mathbf{C\Sigma} \end{bmatrix}\begin{bmatrix} 0 \\ \overline{\boldsymbol{\gamma}}'\mathbf{C\Sigma} \\ \overline{\boldsymbol{\psi}}_{\beta_\perp}'\mathbf{C\Sigma} \end{bmatrix}'$$

which shows that if $q = p$ then $\mathsf{E}\left[\mathbf{N}\right]$ is not invertible. The test statistic will diverge as the rank of $\boldsymbol{\Pi}$ in (1) increases. Also, when the dimension of $\boldsymbol{\psi}_\beta$ increases with q and then dominates $\overline{\boldsymbol{\psi}}_{\beta_\perp}'$ and $\mathbf{C\Sigma}$, so $\mathsf{E}\left[\mathbf{N}\right]$ becomes near singular and LR diverge for large q. We show in the appendix that the pattern is identical for LR^*.

The situation is worse for H_1 (and this also holds under H_1^*) : we show in the appendix that the matrix $\mathbf{N}$ becomes $\mathbf{N}_1$ a matrix with singular expectation. Hence LR_1 and LR_1^* are not robust at all to the presence of local trends that are not orthogonal to $\boldsymbol{\beta}$. Due to the complicated distribution that result, we observe the rejection rates via a Monte Carlo experiment in section 4.

3.2 Lagrange Multiplier test

Because of the difficulty associated with choosing the correction deterministic specification for the Likelihood Ratio test, Helmut Lütkepohl and Penti Saikkonen have proposed in a series of papers (Lütkepohl and Saikkonen, 2000, L&S henceforth, and Saikkonen and Lütkepohl, 1999, 2000a, 2000b) an alternative test which they deem a Lagrange-Multiplier test.

This LM test consists in estimating $\boldsymbol{\Psi}$ under the null hypothesis of q cointegrating relation and hence detrending $\mathbf{y}_t$ into $\widetilde{\mathbf{x}}_t = \mathbf{y}_t - \left(\widetilde{\boldsymbol{\mu}} : \widetilde{\boldsymbol{\psi}}\right)(1 : t)'$ and then testing for $\rho_* = 0$ in a feasible version of

$$\boldsymbol{\alpha}_\perp'\Delta\widetilde{\mathbf{x}}_t = \rho_*\boldsymbol{\beta}_\perp'\widetilde{\mathbf{x}}_{t-1} + \sum_{i=1}^{k-1}\boldsymbol{\Gamma}_{*,i}\Delta\widetilde{\mathbf{x}}_{t-i} + \boldsymbol{\epsilon}_{x,t}.$$

In the locally trending alternative, $\widetilde{\boldsymbol{\psi}}$ is not consistent since it is an estimator of $T^{-1/2}\boldsymbol{\psi}$. Under the null of q cointegrating relations and in the presence of a deterministic trend, L&S, theorem 1, derive the asymptotic distribution of the LM statistic as

$$\mathsf{LM} \Rightarrow tr\left\{\int_0^1 \mathbf{B}_* d\mathbf{B}_*' \left[\int_0^1 \mathbf{B}_*\mathbf{B}_*'\right]^{-1}\int_0^1 \left(d\mathbf{B}_*\right)\mathbf{B}_*'\right\} \tag{10}$$

with $\mathbf{B}_*$ a$(p-q)$-dimensional standard Brownian bridge.

Under the hypothesis of a local trend, although the estimators of the deterministic components are not consistent, detrending still proves effective. Indeed, consider the case of the estimator $\widetilde{\boldsymbol{\psi}}$ in lemma (A.3) of L&S (it is $\widetilde{\mu}_1$ in their notation). $\sqrt{T}\boldsymbol{\beta}_\perp'\left(\widetilde{\boldsymbol{\psi}} - T^{-1/2}\boldsymbol{\psi}\right) \xrightarrow{L} \boldsymbol{\beta}_\perp'\mathbf{C\Sigma B}(1)$. $\widetilde{\boldsymbol{\psi}}$ is in turn used to derive the distribution of the sample moments of

$$\mathbf{w}_t = \boldsymbol{\beta}_\perp'\widetilde{\mathbf{x}}_t = \boldsymbol{\beta}_\perp'\left(\mathbf{y}_t - \widetilde{\boldsymbol{\psi}}t\right) = \boldsymbol{\beta}_\perp'\mathbf{x}_t - \boldsymbol{\beta}_\perp' T\left(\widetilde{\boldsymbol{\psi}} - T^{-1/2}\boldsymbol{\psi}\right)\frac{t}{T}$$

hence, as in L&S, section A.2,

$$T^{-1/2}\mathbf{w}_{[Tr]} \xrightarrow{L} \boldsymbol{\beta}'_{\perp}\mathbf{C}\boldsymbol{\Sigma}\mathbf{B}_{*}(s)$$

and the procedure does provide the asymptotic distribution (10).

4 Monte Carlo

We observe the robustness of the tests via a Monte Carlo experiment where we compute the trace statistic of the LR test over 10,000 replications of the processes. In the simulations, we set $\boldsymbol{\Omega} = \mathbf{I}_p$, $k = 1$ and let, for a cointegration rank q, the vectors $\boldsymbol{\alpha} = \boldsymbol{\beta} = \left(\mathbf{I}_q : \mathbf{0}_{q\times(p-q)}\right)'$ where $\mathbf{I}_q$ is the q dimensional unit matrix and $\mathbf{0}_{m\times n}$ a $(m \times n)$-matrix of zeros. Then $\boldsymbol{\Pi} = diag(\mathbf{I}_q, \mathbf{0}_{(p-q)\times(p-q)})$ and we choose $\boldsymbol{\alpha}_{\perp} = \boldsymbol{\beta}_{\perp} = \left(\mathbf{0}_{(p-q)\times q} : \mathbf{I}_{(p-q)}\right)'$ and $\mathbf{C} = \boldsymbol{\beta}_{\perp}\left(\boldsymbol{\alpha}'_{\perp}\boldsymbol{\beta}_{\perp}\right)^{-1}\boldsymbol{\alpha}'_{\perp} = diag\left(\mathbf{0}_{q\times q}, \mathbf{I}_{(p-q)}\right)$. Hence, the process $\mathbf{x}_t$ consists of $p - q$ independent random walks and q white noises.

We simulate three different experiments: pure deterministic cointegration, pure stochastic (non-deterministic) cointegration and a hybrid, as in

$$\begin{aligned} A &: \boldsymbol{\psi} = \psi\left(\mathbf{0}'_q : \mathbf{1}'_{p-q}\right)' \\ B &: \boldsymbol{\psi} = \psi\left(\mathbf{1}'_q : \mathbf{0}'_{p-q}\right)' \\ C &: \boldsymbol{\psi} = \psi\mathbf{1}'_p \end{aligned}$$

for $\psi \in [0, 10]$, with $\mathbf{1}_{(p-q)}$ a vector $(1 : 1 : ...)'$ of dimension $p-q$. Hence, in experiment A, $\boldsymbol{\psi}_{\beta} = 0$, in $B, \boldsymbol{\psi}_{\beta_{\perp}} = 0$ and in C, $\boldsymbol{\beta}'\ \boldsymbol{\psi}$ and $\boldsymbol{\beta}'_{\perp}\boldsymbol{\psi}$ are both nonzero. In this setting, the coefficient of the trend in $\mathbf{G}$ in (8) is $\left(\overline{\boldsymbol{\psi}}'_{\beta_{\perp}}\mathbf{C}\boldsymbol{\Omega}'\mathbf{C}'\overline{\boldsymbol{\psi}}_{\beta_{\perp}}\right)^{-1/2} = (p-q)^{-1/2}$ in experiments A and C, and 0 in B.

We report in figures 1 to 3 the rejection probabilities of the null of q cointegrating relations out of $p = 6$ variables. In order to reduce Monte Carlo and finite sample variability, we compute the critical values under the corresponding null with $\psi = 0$. This allows to focus specifically on the robustness vis-à-vis the local trend. In the simulation, we let the sample sample vary from 100 to 400 but with very little impact on the conclusions; for this reason, we comment mostly the case $T = 100$.

As expected from proposition 1, the test that corresponds to the null H_2 is robust to local trends. This appears in all the figures as the rejection probability is independent on ψ for all experiments and nominal sizes. All the other statistics are affected by $\boldsymbol{\psi}$. Indeed, even LR_1^* which is asymptotically robust in experiment A is affected in finite samples. We comment on the various results in turn, except LR_2 which is robust to the local trend even in finite samples.

In the case $q = p - 1$ of a unique unit root in $\{\mathbf{x}_t\}$, the LR test under the null is the square of the corresponding Dickey-Fuller distribution. This is presented figure 1. First, in experiment A, two patterns appear. Under assumptions H and H^*, where the process is not corrected for a constant, the rate of rejection of the null increases with ψ: these statistics are not robust to

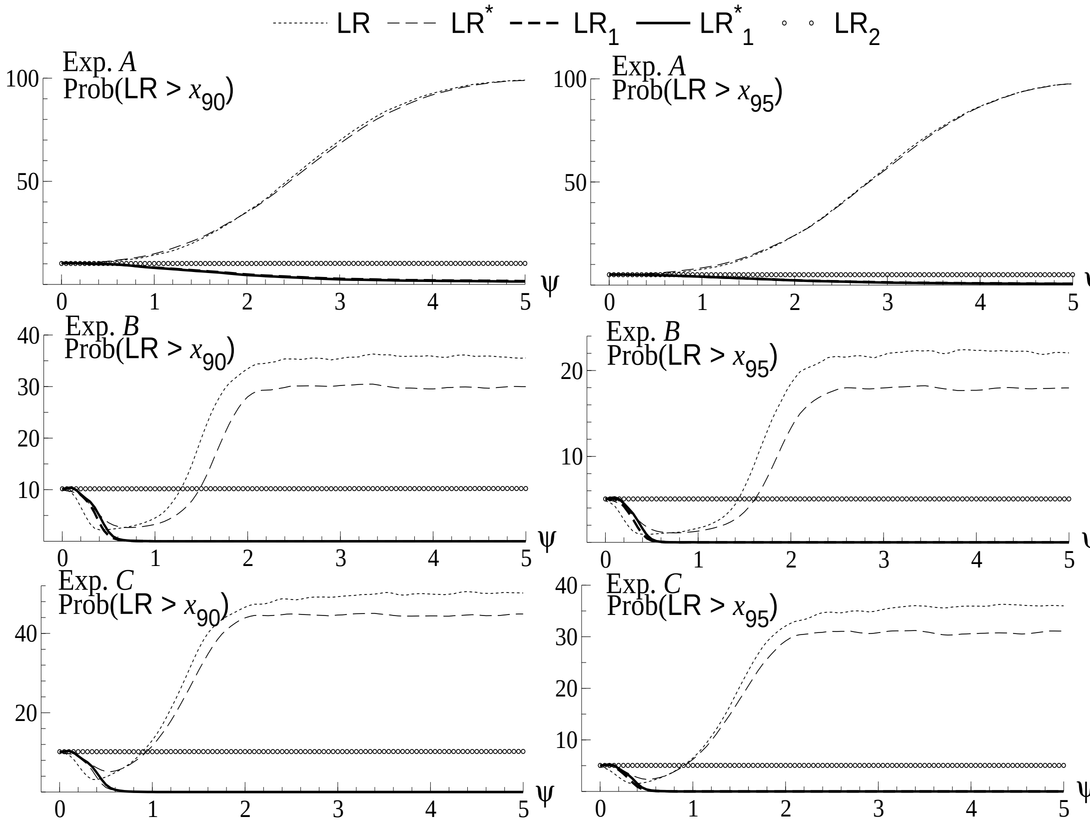

Figure 1: Rejection probabilities of the Likelihood Ratio test as a function of the local trend parameter. The process follows a $p = 6$ dimensional VAR(1) with cointegration rank of $q = 5$: there is a unique common trend. Critical values at sizes 10 (x_{90}) and 5 (x_{95}) are obtained from simulation (10,000 replication) of the distribution of test statistic under the null, with $\boldsymbol{\psi} = 0$. Experiments A, B and C are described section 4, depending on whether $\boldsymbol{\beta}'\boldsymbol{\psi} = 0$, $\boldsymbol{\beta}'_{\perp}\boldsymbol{\psi} = 0$ or neither. The sample size is 100.

misspecified local trends but powerful against the misspecification. The power tends to 100% as ψ increases. By contrast, LR_1 and LR_1^* are, as expected, more robust to the local trend. Yet, these statistics tend to under reject when ψ increases. This is a finite sample issue that is not resolved at $T = 400$.

In experiment B, $\boldsymbol{\psi}_\beta \neq 0$ so proposition 1 only applies to LR_2. As a result, we see that LR_1 and LR_1^* even more strikingly undereject the null: whereas in A the empirical size is half the nominal for $\psi = 2$, in experiment B it drops almost to zero for values of ψ as low as 0.6. Although LR_1^* has a size closer to the nominal than LR_1, this is only slightly so. It is understandable that LR_1 and LR_1^* present similar rejection probabilities since the linear trend is orthogonal to $\boldsymbol{\beta}$, i.e. also to $\boldsymbol{\alpha}$ here and the value added in robustness of LR_1^* over LR_1^* is only vis-à-vis the linear trend in $sp\left(\boldsymbol{\alpha}\right)$. The pattern for LR and LR^* differs greatly from that in A. These two tests also witness a drop in the rejection rate for small values of ψ; even more so for LR. But as ψ increases, test significance increases under both H and H^*. Yet, contrary to experiment A, the rejection rate does not tend to 100% as ψ increases but stabilizes below (depending on the statistic, at about 40% for a nominal size of 10% and 20 for a size of 5%). LR rejects more often than LR^* for large values of ψ, whereas it rejects less for low values of the local trend parameter.

Experiment C is clearly affected by the non zero $\boldsymbol{\psi}_\beta$ as the patterns resemble experiment B, yet with a slightly higher rejection probability overall.

As the number of common trend increases, the pattern vary. In particular, as figures 2 and 3 show, experiment C tends to resemble A more. This is consistent with q decreasing, and also the dimension of the space spanned by $\boldsymbol{\beta}$: $\boldsymbol{\psi}_{\beta_\perp}$ gets to dominate $\boldsymbol{\psi}_\beta$. In addition, experiment A then yields rejection probabilities which are more in line with the asymptotic distributions: LR_1 and LR_1^* become more robust and LR and LR^* more powerful (in the sense that they reject a model that is misspecified). We record the rates for ψ up to 10 in figure 3 because of experiment B. Indeed, as $p - q$ increases, the drop in the rejection rate of LR_1 and LR_1^* occurs for larger values of ψ; and LR and LR^* under reject over a wider range of ψ, crossing back the nominal size at values close to $\psi = p - q$. Also, in the presence of stochastic cointegration, the difference in rejection rates of H and H^* vanishes as q decreases. We omit for brevity the situation where $q = 0$, i.e. in the absence of cointegration. Simulations show that in experiment A, the pattern follows from the other cases considered: LR_1 and LR_1^* are more robust and LR and LR^* reject even more. Clearly A and C coincide in this case. Also, experiment B then implies that $\psi = 0$ since $\boldsymbol{\beta}_\perp = \mathbf{I}_p$.

The LM test being robust to local trends, we do not simulate it in this paper, results will be similar to LR_2. The only difference lies with respect to the respective power of LR_2 and LM with respect to other forms of misspecification, in particular viz. the rank of cointegration. Against alternatives with higher cointegration rank, L&S show that the LM test is locally more powerful than LR_2, but that LM is conservative when testing low ranks and hence should be preferred for

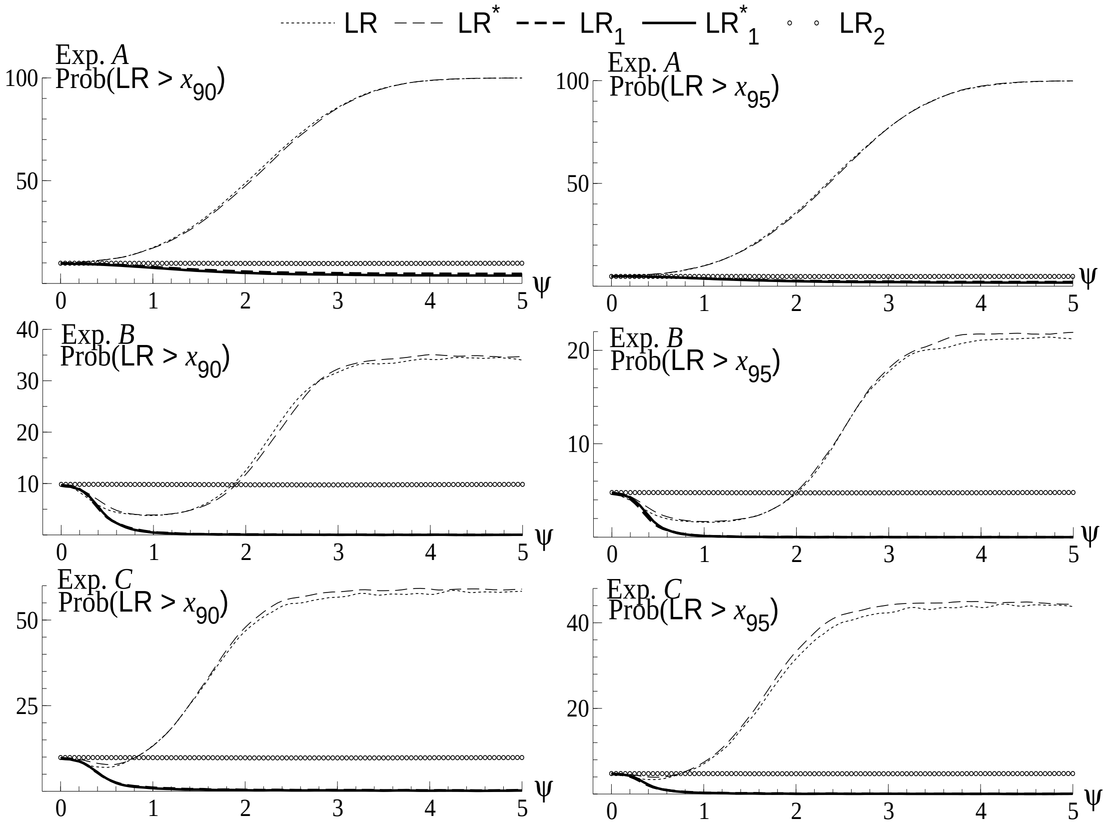

Figure 2: Rejection probabilities of the Likelihood Ratio test as a function of the local trend parameter. The process follows a $p = 6$ dimensional VAR(1) with cointegration rank of $q = 4$; i.e. there are 2 common stochastic trends. Critical values at sizes 10 (x_{90}) and 5 (x_{95}) are obtained from simulation (10,000 replication) of the distribution of test statistic under the null, with $\boldsymbol{\psi} = 0$. Experiments A, B and C are described section 4, depending on whether $\boldsymbol{\beta}'\boldsymbol{\psi} = 0$, $\boldsymbol{\beta}_\perp'\boldsymbol{\psi} = 0$ or neither. The sample size is 100.

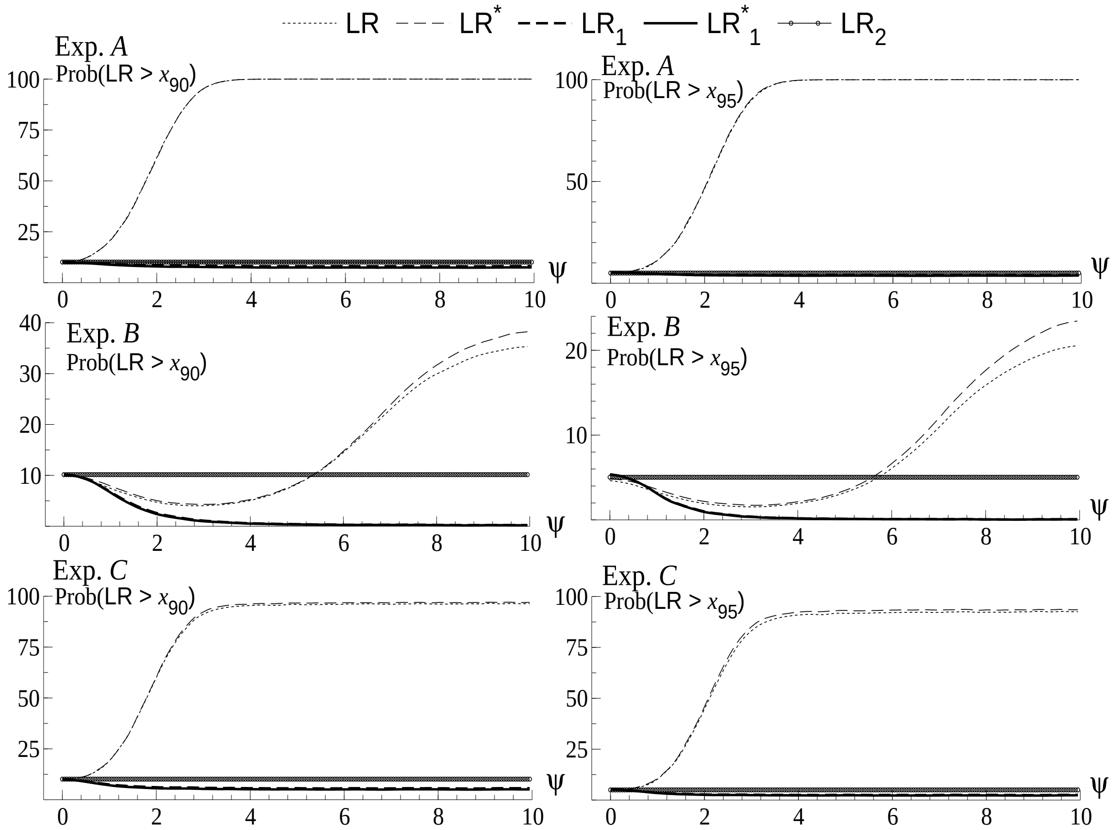

Figure 3: Rejection probabilities of the Likelihood Ratio test as a function of the local trend parameter. The process follows a $p = 6$ dimensional VAR(1) with cointegration rank of $q = 1$, i.e. with 5 common stochastic trends. Critical values at sizes 10 (x_{90}) and 5 (x_{95}) are obtained from simulation (10,000 replication) of the distribution of test statistic under the null, with $\boldsymbol{\psi} = 0$. Experiments A, B and C are described section 4, depending on whether $\boldsymbol{\beta}'\boldsymbol{\psi} = 0$, $\boldsymbol{\beta}'_{\perp}\boldsymbol{\psi} = 0$ or neither. The sample size is 100.

higher values of q. As for comparing with LR_1^*, Saikkonen and Lütkepohl (1999) show that the latter tends to be locally more powerful than LR_2, yet we show that it is not locally robust to linear trends. The recommendation that can be drawn from this analysis is hence to avoid LR_1^* in finite samples where the presence of a deterministic trend cannot be precluded on theoretical grounds, and especially if it is possibly not orthogonal to the cointegrating vector $\boldsymbol{\beta}$, i.e. in the presence of non-deterministic cointegration.

5 Conclusion

In this paper, we have studied the robustness of the test for cointegration in the cointegrated VAR process towards misspecified local linear trends. This situation arises when the data exhibit both stochastic and deterministic trends but the latter have a low magnitude that render them hardly noticeable and not significant. In this setting we have considered five versions of the likelihood ratio test and the Lagrange multiplier test of Lütkepohl and Saikkonen (2000). We have shown that both the LR with unrestricted trend (LR_2) and the LM statistics are asymptotically robust to the local trend when testing for the rank of cointegration in a stochastically cointegrated VAR (from the definition in Perron and Campbell, 1993). In such a setting where $\boldsymbol{\beta}'\boldsymbol{\psi} \neq \mathbf{0}$, using the test statistics LR_1 and LR_1^* leads to invalid inference as the rejection rate drops when the local trend coefficient increases. Yet, LR_1^* is locally robust to the trends when the data are deterministically cointegrated (i.e. $\boldsymbol{\beta}'\boldsymbol{\psi} = \mathbf{0}$), but we show that in finite samples it seriously undereject, even in the presence of small departures from zero trend, when the rank of cointegration is large. In practice, there is little gain in moving from LR_1 to LR_1^*. When the data are stochastically cointegrated with $\boldsymbol{\beta}_\perp'\boldsymbol{\psi} = \mathbf{0}$, both LR_1 and LR_1^* yield flawed inference.

As for the two statistic who never correct for a trend, LR and LR^*, they both show power against a local trend. Yet, when $\boldsymbol{\beta}'\boldsymbol{\psi} \neq \mathbf{0}$, they undereject in the presence of small departures and, although they gain power in the presence of larger departures from zero trend, they do not reject more than half the time. Discrepancies between inference based on LR and LR^* on the one side, and on LR_2 and LM on the other may hence be used as a diagnostic tool for the presence of a small deterministic trend.

Finally, in the general case of stochastic cointegration where the trend spans both $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_\perp$, it is the dimensions of the two subspaces that govern the distributions of the test statistics. Low cointegration ranks imply that the dimension of $sp(\boldsymbol{\beta})$ is low and hence a pattern close to that encountered when $\boldsymbol{\beta}'\boldsymbol{\psi} = \mathbf{0}$; by contrast high ranks of cointegration imply that non-deterministic cointegration prevails.

As a recommendation, given the bad performances of LR_1 and LR_1^*, these should only be used in moderately or largely sized samples and when there are strong reasons why deterministic cointegration must hold. On the other hand, it is known that when the trend is over-specified, the power of the tests for cointegration can prove low. Hence, although it proves useful to use a multivariate extension, as in Lütkepohl and Demetrescu (2008), to the combination of several statistics that was proposed by Harvey *et al.* (2008), it might prove preferable to use $(\mathsf{LR}_2, \mathsf{LM})$ or some other combination that does not include LR_1 and LR_1^*.

References

Chevillon, G. (2008). Inference in the presence of stochastic and deterministic trends. working paper, ESSEC Business School.

Haldrup, N., and Hylleberg, S. (1995). A note on the distribution of the least squares estimator of a random walk with drift: Some analytical evidence. *Economics Letters*, **48**, 221–8.

Harvey, D. I., Leybourne, S. J., and Taylor, A. M. R. (2008). Unit root testing in practice: dealing with uncertainty over the trend and initial condition (with commentaries and rejoinder). *Econometric Theory*, **forthcoming**.

Hubrich, K., Lütkepohl, H., and Saikkonen, P. (2001). A review of systems cointegration tests. *Econometric Reviews*, **20**, 247–318.

Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control*, **12**, 231–54.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, **59**, 1551–80.

Johansen, S. (1994). The role of the constant and linear terms in cointegration analysis of nonstationary time series. *Econometric Reviews*, **13**, 205–31.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector auto-regressive models*. Oxford: Oxford University Press.

Johansen, S. (2000). A bartlett correction factor for tests on the cointegrating relations. *Econometric Theory*, *16*(5), 740–78.

Johansen, S. (2002a). A small sample correction for the test of cointegrating rank in the vector autoregressive model. *Econometrica*, *70*(5), 1929–1961.

Johansen, S. (2002b). A small sample correction for the tests of hypotheses on the cointegrating vectors. *Journal of Econometrics*, *111*(2), 195–221.

Lütkepohl, H., and Demetrescu, M. (2008). Testing for the cointegrating rank of a vector autoregressive process with an uncertain deterministic trend term. mimeo, European University Institute.

Lütkepohl, H., and Saikkonen, P. (2000). Testing for the cointegrating rank of a var process with a time trend. *Journal of Econometrics*, **95**, 177–98.

Müller, U. K., and Elliott, G. (2003). Tests for unit roots and the initial condition. *Econometrica*, **71**, 1269–86.

Murray, C. J., and Nelson, C. R. (2000). The uncertain trend in U.S. GDP. *Journal of Monetary Economics*, **46**, 79–95.

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

Perron, P., and Campbell, J. Y. (1993). A note on Johansen's cointegration procedure when trends are present. *Empirical Economics*, **18**, 777–89.

Rahbek, A. (1994). The power of some cointegration tests. Department of theoretical statistics, University of Copenhagen.

Saikkonen, P., and Lütkepohl, H. (1999). Local power of likelihood ratio tests for the cointegrating rank of a var process. *Econometric Theory*, **15**, 50–78.

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

Saikkonen, P., and Lütkepohl, H. (2000b). Trend adjustment prior to testing for the cointegrating rank of a var process. *Journal of Time Series Analysis*, **21**, 435–56.

Sampson, M. (1991). The effect of parameter uncertainty on forecast variances and confidence intervals for unit root and trend stationary time-series models. *Journal of Applied Econometrics*, **6**, 67–76.

Stock, J. H., and Watson, M. W. (1996). Confidence sets in regressions with highly serially correlated regressors. mimeo, Princeton University.

Toda, H. Y. (1994). Finite sample properties of likelihood ratio tests for cointegrating rank when linear trends are present. *Review of Economics and Statistics*, **76**, 66–79.

6 Appendix

Proof of theorem 1.

6.1 Sample Moments

In the appendix, we assume that $\mathbf{x}_t$ admits a VAR(1) representation, i.e. $k = 1$ in (1). It is straightforward to extend the results to more general dynamics. We follow the lines of the proofs Johansen (1991). We first let $\boldsymbol{\psi} = \boldsymbol{\psi}_\beta + \boldsymbol{\psi}_{\beta_\perp}$ and choose $\boldsymbol{\gamma}$ orthogonal to $\boldsymbol{\beta}$ and $\boldsymbol{\psi}_{\beta_\perp}$ such that $\left(\boldsymbol{\beta} : \boldsymbol{\gamma} : \boldsymbol{\psi}_{\beta_\perp}\right)$ has full rank.

$$\begin{aligned} T^{-1/2}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' \mathbf{y}_{[Tr]} &\Rightarrow \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' \mathbf{K}_{\boldsymbol{\psi},\mathbf{C\Sigma}}(r) \\ &= \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' \mathbf{C\Sigma W}(r) + (\mathbf{0} : 1)' r \overset{def}{=} \mathbf{H}(r) \end{aligned}$$

and let also

$$\mathbf{H}_1(r) = \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' \mathbf{C\Sigma}\left(\mathbf{W}(r) - \int_0^1 \mathbf{W}(r)\,dr\right) + (\mathbf{0} : 1)'(r - 1/2)$$

and $\mathbf{H}_2(r) = \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' \mathbf{C\Sigma}(\mathbf{W}(r) - \mathbf{a} - \mathbf{b}r)$, where the coefficients are obtained by correcting $\mathbf{W}(r)$ for a constant and a linear trend. Now let the variance-covariance matrices

$$\mathsf{Var}\begin{bmatrix} \Delta\mathbf{x}_t \\ \boldsymbol{\beta}'\mathbf{x}_{t-1}\end{bmatrix} = \left[\begin{array}{cc|c} \boldsymbol{\Sigma}_{00} & \boldsymbol{\Sigma}_{0\beta} & \\ & & \Delta\mathbf{x}_{t-1}, ..., \Delta\mathbf{x}_{t-k+1} \\ \boldsymbol{\Sigma}_{\beta 0} & \boldsymbol{\Sigma}_{\beta\beta} & \end{array}\right]$$

which satisfy the relations in lemma 10.1 from Johansen (1995) (denoted lemma J-10.1, and we use similar notation in the following). In the remainder of the appendix, we do not write explicitly that the process is corrected for lagged values since it only affects the definition of Γ. Now recall that S_{ij} is the uncentered sample mean of $R_{it}R'_{jt}$ where $(R_{0t} : R_{1t})$ is $(\Delta\mathbf{y}_t : \mathbf{y}_{t-1})$ corrected for the deterministic terms present in the model under the null. We consider the hypotheses H, H_1 and H_2 in turn, noting that $T^{-2}\sum_{t=1}^T \frac{t}{\sqrt{T}}\frac{t}{\sqrt{T}} = \frac{1}{3} + \frac{1}{2T} + \frac{1}{6T^2}$ and $T^{-1}\sum_{t=1}^T \left[\frac{t-(T+1)/2}{\sqrt{T}}\right]^2 = \frac{T}{12} - \frac{1}{12T}$. Under the hypotheses, the model rewrites as (where we keep the notation ϵ_t although this may also corrected for a constant and possibly a trend):

$$\begin{aligned} \mathsf{H} &: R_{0t} = T^{-1/2}\left(\boldsymbol{\Gamma\psi} - \boldsymbol{\alpha\beta}'\boldsymbol{\psi}_\beta(t-1)\right) + \boldsymbol{\alpha\beta}'R_{1t} + \epsilon_t \\ \mathsf{H}_1 &: R_{0t} = -T^{-1/2}\boldsymbol{\alpha\beta}'\boldsymbol{\psi}_\beta(t - 1 - T/2) + \boldsymbol{\alpha\beta}'R_{1t} + \epsilon_t \\ \mathsf{H}_2 &: R_{0t} = \boldsymbol{\alpha\beta}'R_{1t} + \epsilon_t \end{aligned}$$

Hence the following results.

- First, under H, the residuals R_{it} are not corrected, hence, different limits result, depending on whether $\boldsymbol{\psi}_\beta = \mathbf{0}$:

  First if $\boldsymbol{\psi}_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $\boldsymbol{\beta}'S_{11}\boldsymbol{\beta} \xrightarrow{p} \boldsymbol{\Sigma}_{\beta\beta}$, $\boldsymbol{\beta}'S_{10} \xrightarrow{p} \boldsymbol{\Sigma}_{\beta 0}$ and

$$\begin{aligned} &\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'(S_{10} - S_{11}\boldsymbol{\beta\alpha}') \Rightarrow \textstyle\int_0^1 \mathbf{H}(r)\,d\mathbf{K}'_{\boldsymbol{\Gamma\psi},\mathbf{C\Sigma}}(r) \\ &\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{10} \Rightarrow \textstyle\int_0^1 \mathbf{H}(r)\,d\mathbf{K}'_{\boldsymbol{\psi},\mathbf{C\Sigma}}(r) \\ &T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) \Rightarrow \textstyle\int_0^1 \mathbf{HH}'dr \\ &\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11}\boldsymbol{\beta} = O_p(1) \end{aligned}$$

  and for $\boldsymbol{\psi}_\beta \neq \mathbf{0}$, $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $T^{-1}\boldsymbol{\beta}'S_{11}\boldsymbol{\beta} \xrightarrow{p} \frac{1}{3}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}'_\beta\boldsymbol{\beta}$, $\boldsymbol{\beta}'S_{10} \Rightarrow \boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 r d\mathbf{K}'_{\boldsymbol{\psi},\mathbf{C\Sigma}}(r)$ and

$$\begin{aligned} &T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'(S_{10} - S_{11}\boldsymbol{\beta\alpha}') \Rightarrow \left(\textstyle\int_0^1 r\mathbf{H}(r)\,dr\right)\boldsymbol{\alpha}'\boldsymbol{\beta}\boldsymbol{\psi}'_\beta \\ &\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{10} \Rightarrow \textstyle\int_0^1 \mathbf{H}(r)\,d\mathbf{K}'_{\boldsymbol{\psi},\mathbf{C\Sigma}}(r) \\ &T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) \Rightarrow \textstyle\int_0^1 \mathbf{HH}'dr \\ &T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11}\boldsymbol{\beta} \Rightarrow \left(\textstyle\int_0^1 r\mathbf{H}(r)\,dr\right)\boldsymbol{\psi}'_\beta\boldsymbol{\beta} \end{aligned}$$

- Under H_1,

  If $\boldsymbol{\psi}_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $\boldsymbol{\beta}' S_{11} \boldsymbol{\beta} \xrightarrow{p} \boldsymbol{\Sigma}_{\beta\beta}$, $\boldsymbol{\beta}' S_{10} \xrightarrow{p} \boldsymbol{\Sigma}_{\beta 0}$ and also

$$
\begin{array}{l}
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' (S_{10} - S_{11}\boldsymbol{\beta}\boldsymbol{\alpha}') \Rightarrow \int_0^1 \mathbf{H}_1 d\mathbf{K}'_{\mathbf{0},\boldsymbol{\Sigma}}(r) \\
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{10} \Rightarrow \int_0^1 \mathbf{H}_1 d\mathbf{K}'_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}(r) \\
T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11} \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) \Rightarrow \int_0^1 \mathbf{H}_1 \mathbf{H}_1' dr \\
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11}\boldsymbol{\beta} = O_p(1)
\end{array}
$$

  and for $\boldsymbol{\psi}_\beta \neq \mathbf{0}$, $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $T^{-1}\boldsymbol{\beta}' S_{11}\boldsymbol{\beta} \xrightarrow{p} \frac{1}{12}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}_\beta'\boldsymbol{\beta}$, $\boldsymbol{\beta}' S_{10} \Rightarrow \boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 \left(r - \frac{1}{2}\right) d\mathbf{K}'_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}(r)$ and

$$
\begin{array}{l}
T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' (S_{10} - S_{11}\boldsymbol{\beta}\boldsymbol{\alpha}') \Rightarrow \left(\int_0^1 \left(r - \frac{1}{2}\right) \mathbf{H}_1(r)\, dr\right) \boldsymbol{\alpha}'\boldsymbol{\beta}\boldsymbol{\psi}_\beta' \\
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{10} \Rightarrow \int_0^1 \mathbf{H}_1(r)\, d\mathbf{W}'(r)\, \boldsymbol{\Sigma}'\mathbf{C}' = \int_0^1 \mathbf{H}_1 d\mathbf{K}'_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}(r) \\
T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11} \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) \Rightarrow \int_0^1 \mathbf{H}_1 \mathbf{H}_1' dr \\
T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11}\boldsymbol{\beta} \Rightarrow \left(\int_0^1 \left(r - \frac{1}{2}\right) \mathbf{H}_1(r)\, dr\right) \boldsymbol{\psi}_\beta'\boldsymbol{\beta}
\end{array}
$$

- and finally, under H_2, whether or not $\boldsymbol{\psi}_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $\boldsymbol{\beta}' S_{11} \boldsymbol{\beta} \xrightarrow{p} \boldsymbol{\Sigma}_{\beta\beta}$, $\boldsymbol{\beta}' S_{10} \xrightarrow{p} \boldsymbol{\Sigma}_{\beta 0}$ and

$$
\begin{array}{l}
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' (S_{10} - S_{11}\boldsymbol{\beta}\boldsymbol{\alpha}') \Rightarrow \int_0^1 \mathbf{H}_2 d\mathbf{K}'_{\mathbf{0},\boldsymbol{\Sigma}}(r) \\
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{10} \Rightarrow \int_0^1 \mathbf{H}_2 d\mathbf{K}^{\times\prime}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}(r) \\
T^{-1}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11} \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) \Rightarrow \int_0^1 \mathbf{H}_2 \mathbf{H}_2' dr \\
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S_{11}\boldsymbol{\beta} = O_p(1)
\end{array}
$$

  where we denote by $\mathbf{K}^{\times\prime}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}(r)$ the detrended version of $\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}(r)$.

We complete the analysis above with the two hypotheses H^* and H_1^*. Note that, under H^*, $R_{2t} = 1$ and under $\mathsf{H}_1^* : R_{2t} = t - 1 - T/2$. Now let $(\boldsymbol{\Gamma\psi})_\alpha = \boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})^{-1}\boldsymbol{\beta}'\boldsymbol{\Gamma\psi}$ and $(\boldsymbol{\Gamma\psi})_{\alpha_\perp} = \boldsymbol{\Gamma\psi} - (\boldsymbol{\Gamma\psi})_\alpha$, then

$$
\begin{array}{lll}
\mathsf{H}^* & : & R_{0t} = -T^{-1/2}\boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta(t-1) + T^{-1/2}(\boldsymbol{\Gamma\psi})_{\alpha_\perp} + \boldsymbol{\alpha}\boldsymbol{\beta}' R_{1t} + T^{-1/2}(\boldsymbol{\Gamma\psi})_\alpha R_{2t} + \epsilon_t \\
\mathsf{H}_1^* & : & R_{0t} = \boldsymbol{\alpha}\boldsymbol{\beta}' R_{1t} - T^{-1/2}\boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta R_{2t} + \epsilon_t.
\end{array}
$$

We therefore derive the following properties of the sample moments under linear restrictions of the parameters.

- First under H^* : letting $R_{1t}^* = (R_{1t}' : R_{2t}')'$ and $\boldsymbol{\beta}^{*\prime} = \left(\boldsymbol{\beta}' : (\boldsymbol{\beta}'\boldsymbol{\alpha})^{-1}\boldsymbol{\beta}'\boldsymbol{\Gamma\psi}\right) = (\boldsymbol{\beta}' : \beta_2)$. We also define the vectors $\boldsymbol{\gamma}^{*\prime} = \left(\left[\gamma : \boldsymbol{\psi}_{\beta_\perp}\right]' : 0\right)$ and $\tau^{*\prime} = (\mathbf{0} : 1)$ such that $(\boldsymbol{\beta}^* : \boldsymbol{\gamma}^* : \tau^*)$ is of full rank $p+1$, then

$$
\left(T^{-1/2}\overline{\boldsymbol{\gamma}}^* : \overline{\tau}^*\right)' \begin{bmatrix} \mathbf{y}_{[Tr]} \\ 1 \end{bmatrix} \Rightarrow \left(T^{-1/2}\overline{\boldsymbol{\gamma}}^* : \overline{\tau}^*\right)' \begin{bmatrix} \mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}(r) \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{H}(r) \\ 1 \end{bmatrix} \stackrel{def}{=} \mathbf{H}^*(r)
$$

  if $\boldsymbol{\psi}_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $\boldsymbol{\beta}^{*\prime} S_{11}^* \boldsymbol{\beta}^* \xrightarrow{p} \boldsymbol{\Sigma}_{\beta\beta} + \beta_2\beta_2'$, $\boldsymbol{\beta}^{*\prime} S_{10}^* \xrightarrow{p} \boldsymbol{\Sigma}_{\beta 0}$ and

$$
\begin{array}{l}
\left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' (S_{10}^* - S_{11}^*\boldsymbol{\beta}^*\boldsymbol{\alpha}') \Rightarrow \int_0^1 \mathbf{H}^*(r)\, d\mathbf{K}'_{(\boldsymbol{\Gamma\psi})_{\alpha_\perp},\boldsymbol{\Sigma}}(r) \\
\left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' S_{10}^* \Rightarrow \int_0^1 \mathbf{H}^*(r)\, d\mathbf{K}'_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}(r) \\
T^{-1}\left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' S_{11} \left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right) \Rightarrow \int_0^1 \mathbf{H}^*(r)\, \mathbf{H}^{*\prime}(r)\, dr \\
\left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' S_{11}\boldsymbol{\beta} = o_p(T)
\end{array}
$$

and for $\boldsymbol{\psi}_\beta \neq \mathbf{0}$, then $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $T^{-1}\boldsymbol{\beta}^{*\prime} S_{11}^* \boldsymbol{\beta}^* \xrightarrow{p} \frac{1}{3}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}'_\beta\boldsymbol{\beta}$ and

$$\boldsymbol{\beta}^{*\prime} S_{10}^* \Rightarrow \boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 r d\mathbf{K}'_{\psi,\mathbf{C\Sigma}}(r)$$

with also

$$\begin{array}{l} T^{-1}\left(\overline{\gamma}^* : \sqrt{T}\overline{\tau}^*\right)'\left(S_{10}^* - S_{11}^*\boldsymbol{\beta}^*\boldsymbol{\alpha}'\right) \Rightarrow -\int_0^1 r\mathbf{H}^*(r)\,\boldsymbol{\psi}'_\beta\boldsymbol{\beta}\boldsymbol{\alpha}' dr \\ \left(\overline{\gamma}^* : \sqrt{T}\overline{\tau}^*\right)' S_{10}^* \Rightarrow \int_0^1 \mathbf{H}^*(r)\, d\mathbf{K}'_{\psi,\mathbf{C\Sigma}}(r) \\ T^{-1}\left(\overline{\gamma}^* : \sqrt{T}\overline{\tau}^*\right)' S_{11}^* \left(\overline{\gamma}^* : \sqrt{T}\overline{\tau}^*\right) \Rightarrow \int_0^1 \mathbf{H}^*(r)\,\mathbf{H}^{*\prime}(r)\, dr \\ T^{-1}\left(\overline{\gamma}^* : \sqrt{T}\overline{\tau}^*\right)' S_{11}^*\boldsymbol{\beta}^* \Rightarrow \int_0^1 r\mathbf{H}^*(r)\,\boldsymbol{\psi}'_\beta\boldsymbol{\beta} dr \end{array}$$

- And finally under H_1^*, letting $R_{1t}^* = (R'_{1t} : R'_{2t})'$ and $\boldsymbol{\beta}^{*\prime} = \left(\boldsymbol{\beta}' : \boldsymbol{\beta}'\boldsymbol{\psi}_\beta\right)$. We also define the matrices $\boldsymbol{\gamma}^{*\prime} = \left(\left[\boldsymbol{\gamma} : \boldsymbol{\psi}_{\beta_\perp}\right]' : 0\right)$ and $\tau^{*\prime} = (\mathbf{0} : 1)$ such that $(\boldsymbol{\beta}^* : \boldsymbol{\gamma}^* : \tau^*)$ is of full rank $p+1$, then

$$T^{-1/2}\left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)'\begin{bmatrix} \mathbf{y}_{[Tr]} \\ [Tr] \end{bmatrix} \Rightarrow \left(\overline{\gamma}^* : \overline{\tau}^*\right)'\begin{bmatrix} \mathbf{K}_{\psi,\mathbf{C\Sigma}}(r) \\ r \end{bmatrix} = \begin{bmatrix} \mathbf{H}(r) \\ r \end{bmatrix}$$

Hence, if $\boldsymbol{\psi}_\beta = \mathbf{0}$, $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}, \boldsymbol{\beta}^{*\prime} S_{11}^*\boldsymbol{\beta}^* \xrightarrow{p} \boldsymbol{\Sigma}_{\beta\beta}, \boldsymbol{\beta}^{*\prime} S_{10}^* \xrightarrow{p} \boldsymbol{\Sigma}_{\beta 0}$ and, letting $\mathbf{H}_1^*(r) = \left(\mathbf{H}'_1(r) : r - 1/2\right)'$

$$\begin{array}{l} \left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)'\left(S_{10}^* - S_{11}^*\boldsymbol{\beta}^*\boldsymbol{\alpha}'\right) \Rightarrow \int_0^1 \mathbf{H}_1^*(r)\, d\mathbf{K}'_{0,\boldsymbol{\Sigma}}(r) \\ \left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)' S_{10}^* \Rightarrow \int_0^1 \mathbf{H}_1^*(r)\, d\mathbf{K}'_{0,\mathbf{C\Sigma}}(r) \\ T^{-1}\left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)' S_{11}^*\left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right) \Rightarrow \int_0^1 \mathbf{H}_1^*(r)\,\mathbf{H}_1^{*\prime}(r)\, dr \\ \left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)' S_{11}^*\boldsymbol{\beta}^* = \left(o_p(T) : o_p(T)\right)' \end{array}$$

and for $\boldsymbol{\psi}_\beta \neq \mathbf{0}$, $S_{00} \xrightarrow{p} \boldsymbol{\Sigma}_{00}$, $T^{-2}\boldsymbol{\beta}^{*\prime} S_{11}^*\boldsymbol{\beta}^* \xrightarrow{p} \frac{1}{12}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}'_\beta\boldsymbol{\beta}$, $\boldsymbol{\beta}^{*\prime} S_{10}^* \xrightarrow{p} \boldsymbol{\Sigma}_{\beta 0}$,

$$\begin{array}{l} T^{-1}\left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)'\left(S_{10}^* - S_{11}^*\boldsymbol{\beta}^*\boldsymbol{\alpha}'\right) \Rightarrow -\int_0^1 \left(r - \frac{1}{2}\right)\mathbf{H}_1^*(r)\,\boldsymbol{\psi}'_\beta\boldsymbol{\beta}\boldsymbol{\alpha}' dr \\ \left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)' S_{10}^* \Rightarrow \int_0^1 \mathbf{H}_1^*(r)\, d\mathbf{K}'_{0,\mathbf{C\Sigma}}(r) \\ T^{-1}\left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)' S_{11}^*\left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right) \Rightarrow \int_0^1 \mathbf{H}_1^*(r)\,\mathbf{H}_1^{*\prime}(r)\, dr \\ T^{-3/2}\left(\overline{\gamma}^* : T^{-1/2}\overline{\tau}^*\right)' S_{11}\boldsymbol{\beta}^* \Rightarrow \int_0^1 \left(r - \frac{1}{2}\right)\mathbf{H}_1^*(r)\,\boldsymbol{\psi}'_\beta\boldsymbol{\beta} dr \end{array}$$

6.2 Trace statistic

We turn next to the asymptotic distribution of

$$|S(\lambda)| = \left|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}\right|$$

where the trace statistic is the sum of the $p-q$ smallest solutions to the equation $|S(\lambda)| = 0$ (except for the hypotheses H^* and H_1^* which we treat later). The matrix $\left(\boldsymbol{\beta} : \boldsymbol{\psi}_{\beta_\perp} : \boldsymbol{\gamma}\right)$ has full rank, then denote $\mathbf{A}_T = \left(\boldsymbol{\beta} : T^{-1/2}\boldsymbol{\psi}_{\beta_\perp} : T^{-1/2}\boldsymbol{\gamma}\right)$ such that we obtain distributions under the various hypotheses. We provide standardizing matrices at the end of the section.

- Under assumption H, if $\boldsymbol{\psi}_\beta = \mathbf{0}$

$$\left|\mathbf{A}_T' S(\lambda)\mathbf{A}_T\right| \Rightarrow \left|\lambda\begin{bmatrix}\boldsymbol{\Sigma}_{\beta\beta} & 0\\ 0 & \int_0^1 \mathbf{H}\mathbf{H}'dr\end{bmatrix} - \begin{bmatrix}\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta} & 0\\ 0 & 0\end{bmatrix}\right|$$
$$= \left|\lambda\boldsymbol{\Sigma}_{\beta\beta} - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right| \left|\lambda\int_0^1 \mathbf{H}\mathbf{H}'dr\right|$$

which has q positive roots given by $\left|\lambda\boldsymbol{\Sigma}_{\beta\beta} - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right|$. Now, consider

$$\begin{aligned}
&\left|\left(\boldsymbol{\beta} : \overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\left(\boldsymbol{\beta} : \overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\right| \\
= &\begin{vmatrix}\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta} & \boldsymbol{\beta}' S(\lambda)\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\\ \left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\boldsymbol{\beta} & \left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\end{vmatrix} \\
= &\left|\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta}\right| \left|\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' \left\{S(\lambda) - S(\lambda)\boldsymbol{\beta}\left[\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}' S(\lambda)\right\}\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\right|
\end{aligned}$$

where the first factor has no roots:

$$\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta} \to -\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}$$

and, letting $\rho = T\lambda$

$$\begin{aligned}
\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) &= \rho\int_0^1 \mathbf{H}\mathbf{H}'dr - \left(\int_0^1 \mathbf{H}(r)\, d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}(r)\,\mathbf{H}'(r) \\
\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\boldsymbol{\beta} &= O_p(\lambda) - \left(\int_0^1 \mathbf{H}(r)\, d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta} \\
&\Rightarrow \left(\int_0^1 \mathbf{H}(r)\, d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}
\end{aligned}$$

hence

$$\begin{aligned}
&\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' \left\{S(\lambda) - S(\lambda)\boldsymbol{\beta}\left[\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}' S'(\lambda)\right\}\left(\overline{\boldsymbol{\gamma}} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) \\
\Rightarrow\ &\rho\int_0^1 \mathbf{H}\mathbf{H}'dr - \left(\int_0^1 \mathbf{H}(r)\, d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}(r)\,\mathbf{H}'(r) \\
&+ \left(\int_0^1 \mathbf{H}(r)\, d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\left[\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right]^{-1}\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}(r)\,\mathbf{H}'(r) \\
=\ &\rho\int_0^1 \mathbf{H}\mathbf{H}'dr - \left(\int_0^1 \mathbf{H}(r)\, d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\left[\boldsymbol{\Sigma}_{00}^{-1} - \boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\left[\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right]^{-1}\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\right]\int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}(r)\,\mathbf{H}'(r)
\end{aligned}$$

where $\boldsymbol{\Sigma}_{00}^{-1} - \boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\left[\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right]^{-1}\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1} = \boldsymbol{\alpha}_\perp\left(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}\boldsymbol{\alpha}_\perp\right)^{-1}\boldsymbol{\alpha}_\perp'$. Noting that $S_{10}\boldsymbol{\alpha}_\perp = \left(S_{10} - S_{11}\boldsymbol{\beta}\boldsymbol{\alpha}'\right)\boldsymbol{\alpha}_\perp$The above expression is therefore equal to

$$\rho\int_0^1 \mathbf{H}\mathbf{H}'dr - \left(\int_0^1 \mathbf{H}(r)\, d\mathbf{K}_{\boldsymbol{\psi},\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\alpha}_\perp\left(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}\boldsymbol{\alpha}_\perp\right)^{-1}\boldsymbol{\alpha}_\perp'\int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\boldsymbol{\Sigma}}(r)\,\mathbf{H}'(r) \qquad (11)$$

Now for $\boldsymbol{\psi}_\beta \neq \mathbf{0}$, define the two matrices

$$\mathbf{N} = \begin{bmatrix} \frac{1}{3}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & \boldsymbol{\beta}'\boldsymbol{\psi}_\beta\left(\int_0^1 r\mathbf{H}'(r)\,dr\right) \\ \left(\int_0^1 r\mathbf{H}(r)\,dr\right)\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & \int_0^1 \mathbf{H}\mathbf{H}'dr \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} \left(\boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 rd\mathbf{K}_{\psi,\mathbf{C\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\left(\boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 rd\mathbf{K}_{\psi,\mathbf{C\Sigma}}'(r)\right)' & \left(\boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 rd\mathbf{K}_{\psi,\mathbf{C\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C\Sigma}}\mathbf{H}'(r) \\ \int_0^1 \mathbf{H}(r)\,d\mathbf{K}_{\psi,\mathbf{C\Sigma}}'(r)\,\boldsymbol{\Sigma}_{00}^{-1}\left(\boldsymbol{\Sigma}_{\beta 0} + \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 rd\mathbf{K}_{\psi,\mathbf{C\Sigma}}'(r)\right)' & \int_0^1 \mathbf{H}(r)\,d\mathbf{K}_{\psi,\mathbf{C\Sigma}}'(r)\,\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\psi,\mathbf{C\Sigma}}\mathbf{H}'(r) \end{bmatrix}$$

then

$$\left(\boldsymbol{\beta} \vdots \overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\left(\boldsymbol{\beta} \vdots \overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) = \lambda T\mathbf{N} - \mathbf{M} + o_p(1)$$

which shows that if $|\mathbf{N}| \neq 0$, then the test statistic has the same distribution as

$$Ttr\left\{S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}\right\} \Rightarrow tr\left\{\mathbf{N}^{-1}\mathbf{M}\right\}$$

Now recall the definition $\mathbf{H}(r) = \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'\mathbf{C\Sigma W}(r) + \left(\boldsymbol{\psi}_\beta'\overline{\gamma} : 1\right)' r$, then$\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'\mathbf{C\Sigma W}(r) + (\mathbf{0} : 1)' r \overset{def}{=} \mathbf{H}(r)$

$$\begin{aligned}
\mathsf{E}[\mathbf{N}] &= \begin{bmatrix} \frac{1}{3}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & \boldsymbol{\beta}'\boldsymbol{\psi}_\beta\left(\mathbf{0} : \frac{1}{2}\right) \\ \left(\mathbf{0} : \frac{1}{2}\right)'\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & \left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'\mathbf{C\Omega C}'\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) + \frac{1}{2}(\mathbf{0} : 1)'(\mathbf{0} : 1) \end{bmatrix} \\
&= -\begin{bmatrix} \frac{1}{6}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \\ \mathbf{0} \\ 1 \end{bmatrix}\begin{bmatrix} \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \\ \mathbf{0} \\ 1 \end{bmatrix}' + \begin{bmatrix} 0 \\ \overline{\gamma}'\mathbf{C\Sigma} \\ \overline{\boldsymbol{\psi}}_{\beta_\perp}'\mathbf{C\Sigma} \end{bmatrix}\begin{bmatrix} 0 \\ \overline{\gamma}'\mathbf{C\Sigma} \\ \overline{\boldsymbol{\psi}}_{\beta_\perp}'\mathbf{C\Sigma} \end{bmatrix}'
\end{aligned}$$

hence the expectation of the matrix has nonzero determinant as long as $q < p$.

- Under assumption H_1, if $\boldsymbol{\psi}_\beta = \mathbf{0}$

$$\left|\mathbf{A}_T' S(\lambda)\,\mathbf{A}_T\right| \Rightarrow \left|\lambda\boldsymbol{\Sigma}_{\beta\beta} - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right|\left|\lambda\int_0^1 \mathbf{H}_1\mathbf{H}_1'dr\right|$$

which has q positive roots given by $\left|\lambda\boldsymbol{\Sigma}_{\beta\beta} - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right|$. Now, consider

$$\left|\left(\boldsymbol{\beta} \vdots \overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\left(\boldsymbol{\beta} \vdots \overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\right| = \left|\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta}\right|\left|\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'\left\{S(\lambda) - S(\lambda)\boldsymbol{\beta}\left[\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}' S(\lambda)\right\}\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\right|$$

where the first factor has no roots as $\lambda \to 0$

$$\boldsymbol{\beta}' S(\lambda)\boldsymbol{\beta} \to -\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}$$

and,

$$\begin{aligned}
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right) &= \rho\int_0^1 \mathbf{H}_1\mathbf{H}_1'dr - \left(\int_0^1 \mathbf{H}(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\mathbf{0},\mathbf{C\Sigma}}(r)\,\mathbf{H}'(r) \\
\left(\overline{\gamma} : \overline{\boldsymbol{\psi}}_{\beta_\perp}\right)' S(\lambda)\boldsymbol{\beta} &= O_p(\lambda) - \left(\int_0^1 \mathbf{H}_1(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta} \\
&\Rightarrow \left(\int_0^1 \mathbf{H}_1(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}
\end{aligned}$$

hence

$$\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'\left\{S(\lambda)-S(\lambda)\boldsymbol{\beta}\left[\boldsymbol{\beta}'S(\lambda)\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}'S'(\lambda)\right\}\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)$$
$$\Rightarrow\quad \rho\int_0^1 \mathbf{H}_1\mathbf{H}_1'dr-\left(\int_0^1 \mathbf{H}_1(r)\,d\mathbf{K}_{\mathbf{0},\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\alpha}_\perp\left(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}\boldsymbol{\alpha}_\perp\right)^{-1}\boldsymbol{\alpha}_\perp'\int_0^1 d\mathbf{K}_{\mathbf{0},\boldsymbol{\Sigma}}(r)\,\mathbf{H}_1'(r)$$

Now, for $\boldsymbol{\psi}_\beta\neq\mathbf{0}$, as previously, let

$$\mathbf{N}_1 = \begin{bmatrix} \frac{1}{12}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & \boldsymbol{\beta}'\boldsymbol{\psi}_\beta\left(\int_0^1\left(r-\frac{1}{2}\right)\mathbf{H}_1'(r)\,dr\right) \\ \left(\int_0^1\left(r-\frac{1}{2}\right)\mathbf{H}_1(r)\,dr\right)\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & \int_0^1\mathbf{H}_1\mathbf{H}_1'dr \end{bmatrix}$$

$$\mathbf{M}_1 = \begin{bmatrix} \left(\boldsymbol{\Sigma}_{\beta0}+\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\int_0^1\left(r-\frac{1}{2}\right)d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\left(\boldsymbol{\Sigma}_{\beta0}+\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\int_0^1\left(r-\frac{1}{2}\right)d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)' & \left(\boldsymbol{\Sigma}_{\beta0}+\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\int_0^1\left(r-\frac{1}{2}\right)d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}\mathbf{H}_1'(r) \\ \int_0^1\mathbf{H}_1(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}'(r)\,\boldsymbol{\Sigma}_{00}^{-1}\left(\boldsymbol{\Sigma}_{\beta0}+\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\int_0^1\left(r-\frac{1}{2}\right)d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)' & \int_0^1\mathbf{H}_1(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}'(r)\,\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}\mathbf{H}_1'(r) \end{bmatrix}$$

where $\mathsf{E}[\mathbf{N}_1]$ has zero determinant since it is the sum of two outer products of matrices:

$$\mathsf{E}[\mathbf{N}_1]=\frac{1}{12}\begin{bmatrix}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\\ \mathbf{0}\\ 1\end{bmatrix}\begin{bmatrix}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\\ \mathbf{0}\\ 1\end{bmatrix}'+\frac{1}{6}\begin{bmatrix}0\\ \overline{\gamma}'\mathbf{C}\boldsymbol{\Sigma}\\ \overline{\boldsymbol{\psi}}_{\beta_\perp}'\mathbf{C}\boldsymbol{\Sigma}\end{bmatrix}\begin{bmatrix}0\\ \overline{\gamma}'\mathbf{C}\boldsymbol{\Sigma}\\ \overline{\boldsymbol{\psi}}_{\beta_\perp}'\mathbf{C}\boldsymbol{\Sigma}\end{bmatrix}'.$$

The trace statistic therefore tends to a random variable whose expectation can take any values.

- Under H_2, if $\boldsymbol{\psi}_\beta=\mathbf{0}$ or $\boldsymbol{\psi}_\beta\neq\mathbf{0}$

$$\left|\mathbf{A}_T'S(\lambda)\,\mathbf{A}_T\right|\Rightarrow\left|\lambda\boldsymbol{\Sigma}_{\beta\beta}-\boldsymbol{\Sigma}_{\beta0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right|\left|\lambda\int_0^1\mathbf{H}_2\mathbf{H}_2'dr\right|$$

which has q positive roots given by $\left|\lambda\boldsymbol{\Sigma}_{\beta\beta}-\boldsymbol{\Sigma}_{\beta0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right|$. Now, consider

$$\begin{aligned}&\left|\left(\boldsymbol{\beta}:\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'S(\lambda)\left(\boldsymbol{\beta}:\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\right|\\ =\;&\left|\boldsymbol{\beta}'S(\lambda)\boldsymbol{\beta}\right|\left|\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'\left\{S(\lambda)-S(\lambda)\boldsymbol{\beta}\left[\boldsymbol{\beta}'S(\lambda)\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}'S'(\lambda)\right\}\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)\right|\end{aligned}$$

where the first factor has no roots as $\lambda\to0$, since $\boldsymbol{\beta}'S(\lambda)\boldsymbol{\beta}\to-\,\boldsymbol{\Sigma}_{\beta0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}$, and,

$$\begin{aligned}\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'S(\lambda)\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right) &= \rho\int_0^1\mathbf{H}_2\mathbf{H}_2'dr-\left(\int_0^1\mathbf{H}_2(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\int_0^1 d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}(r)\,\mathbf{H}_2'(r)\\ \left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'S(\lambda)\boldsymbol{\beta} &= O_p(\lambda)-\left(\int_0^1\mathbf{H}_2(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\\ &\Rightarrow\left(\int_0^1\mathbf{H}_2(r)\,d\mathbf{K}_{\mathbf{0},\mathbf{C}\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\end{aligned}$$

hence

$$\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)'\left\{S(\lambda)-S(\lambda)\boldsymbol{\beta}\left[\boldsymbol{\beta}'S(\lambda)\boldsymbol{\beta}\right]^{-1}\boldsymbol{\beta}'S(\lambda)\right\}\left(\overline{\gamma}:\overline{\boldsymbol{\psi}}_{\beta_\perp}\right)$$
$$\Rightarrow\quad \rho\int_0^1 \mathbf{H}_2\mathbf{H}_2'dr-\left(\int_0^1 \mathbf{H}_2(r)\,d\mathbf{K}_{\mathbf{0},\boldsymbol{\Sigma}}'(r)\right)\boldsymbol{\alpha}_\perp\left(\boldsymbol{\alpha}_\perp'\boldsymbol{\Omega}\boldsymbol{\alpha}_\perp\right)^{-1}\boldsymbol{\alpha}_\perp'\int_0^1 d\mathbf{K}_{\mathbf{0},\boldsymbol{\Sigma}}(r)\,\mathbf{H}_2'(r)$$

- under H^*, with $\mathbf{A}_T^* = \left(\boldsymbol{\beta}^* : T^{-1/2}\overline{\boldsymbol{\gamma}}^* : \overline{\tau}^*\right)$, if $\boldsymbol{\psi}_\beta = \mathbf{0}$, then

$$\begin{aligned}
& \left|\mathbf{A}_T^{*\prime} S^* (\lambda) \mathbf{A}_T^*\right| \\
\Rightarrow\quad & \left|\lambda \begin{bmatrix} \boldsymbol{\Sigma}_{\beta\beta} + \beta_2\beta_2' & 0 \\ 0 & \int_0^1 \mathbf{H}^*\mathbf{H}^{*\prime} dr \end{bmatrix} - \begin{bmatrix} \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta} & 0 \\ 0 & 0 \end{bmatrix}\right| \\
=\quad & \left|\lambda\left(\boldsymbol{\Sigma}_{\beta\beta} + \beta_2\beta_2'\right) - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right| \left|\lambda \int_0^1 \mathbf{H}^*\mathbf{H}^{*\prime} dr\right|
\end{aligned}$$

which has q positive roots given by $\left|\lambda\left(\boldsymbol{\Sigma}_{\beta\beta} + \beta_2\beta_2'\right) - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right|$. Now, consider

$$\begin{aligned}
& \left|\left(\boldsymbol{\beta}^* : \overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' S^* (\lambda) \left(\boldsymbol{\beta} : \overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)\right| \\
=\quad & \left|\boldsymbol{\beta}^{*\prime} S^* (\lambda) \boldsymbol{\beta}^*\right| \left|\left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' \left\{S^* (\lambda) - S^* (\lambda) \boldsymbol{\beta}^* \left[\boldsymbol{\beta}^{*\prime} S^* (\lambda) \boldsymbol{\beta}^*\right]^{-1} \boldsymbol{\beta}^{*\prime} S^* (\lambda)\right\} \left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)\right|
\end{aligned}$$

where the first factor has no roots as $\lambda \to 0$ since $\boldsymbol{\beta}^{*\prime} S^* (\lambda) \boldsymbol{\beta}^* \to -\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}$ and, letting $\rho = T\lambda$

$$\begin{aligned}
\left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' S^* (\lambda) \left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right) \quad & \Rightarrow \quad \rho \int_0^1 \mathbf{H}^*\mathbf{H}^{*\prime} dr - \int_0^1 \mathbf{H}^* d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}' (r) \boldsymbol{\Sigma}_{00}^{-1} \int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}} (r) \mathbf{H}^{*\prime} \\
\left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' S^* (\lambda) \boldsymbol{\beta}^* \quad & \Rightarrow \quad \int_0^1 \mathbf{H}^* d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}' (r) \boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{\beta 0}
\end{aligned}$$

hence

$$\begin{aligned}
& \left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right)' \left\{S^* (\lambda) - S^* (\lambda) \boldsymbol{\beta}^* \left[\boldsymbol{\beta}^{*\prime} S^* (\lambda) \boldsymbol{\beta}^*\right]^{-1} \boldsymbol{\beta}^{*\prime} S^{*\prime} (\lambda)\right\} \left(\overline{\boldsymbol{\gamma}}^* : \sqrt{T}\overline{\tau}^*\right) \\
\Rightarrow\quad & \rho \int_0^1 \mathbf{H}^*\mathbf{H}^{*\prime} dr - \int_0^1 \mathbf{H}^* d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}}' (r) \left(\boldsymbol{\Sigma}_{00}^{-1} - \boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{\beta 0}\left(\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right)^{-1}\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\right) \int_0^1 d\mathbf{K}_{\boldsymbol{\psi},\mathbf{C}\boldsymbol{\Sigma}} (r) \mathbf{H}^{*\prime} \\
=\quad & \rho \int_0^1 \mathbf{H}^*\mathbf{H}^{*\prime} dr - \int_0^1 \mathbf{H}^* d\mathbf{K}_{(\boldsymbol{\Gamma}\boldsymbol{\psi})_{\alpha_\perp},\boldsymbol{\Sigma}}' (r) \boldsymbol{\alpha}_\perp \left(\boldsymbol{\alpha}_\perp' \boldsymbol{\Omega} \boldsymbol{\alpha}_\perp\right)^{-1} \boldsymbol{\alpha}_\perp' \int_0^1 d\mathbf{K}_{(\boldsymbol{\Gamma}\boldsymbol{\psi})_{\alpha_\perp},\boldsymbol{\Sigma}} (r) \mathbf{H}^{*\prime}
\end{aligned}$$

For $\boldsymbol{\psi}_\beta \neq \mathbf{0}$,

$$\left(T^{-1/2}\boldsymbol{\beta}^* : T^{-1/2}\overline{\boldsymbol{\gamma}}^* : \overline{\tau}^*\right)' S^* (\lambda) \left(T^{-1/2}\boldsymbol{\beta}^* : T^{-1/2}\overline{\boldsymbol{\gamma}}^* : \overline{\tau}^*\right) \Rightarrow \lambda \begin{bmatrix} \frac{1}{3}\boldsymbol{\beta}'\boldsymbol{\psi}_\beta\boldsymbol{\psi}_\beta'\boldsymbol{\beta} & \boldsymbol{\beta}'\boldsymbol{\psi}_\beta \int_0^1 r\mathbf{H}^{*\prime}(r)\, dr \\ \int_0^1 r\mathbf{H}^*(r)\boldsymbol{\psi}_\beta'\boldsymbol{\beta} dr & \int_0^1 \mathbf{H}^*(r)\mathbf{H}^{*\prime}(r)\, dr \end{bmatrix}$$

which implies a behavior similar to that under H.

- under H_1^*, with $\mathbf{A}_T^* = \left(\boldsymbol{\beta}^* : T^{-1/2}\overline{\boldsymbol{\gamma}}^* : T^{-1}\overline{\tau}^*\right)$, if $\boldsymbol{\psi}_\beta = \mathbf{0}$, then

$$\begin{aligned}
& \left|\mathbf{A}_T^{*\prime} S^* (\lambda) \mathbf{A}_T^*\right| \\
\Rightarrow\quad & \left|\lambda \begin{bmatrix} \boldsymbol{\Sigma}_{\beta\beta} & 0 \\ 0 & \int_0^1 \mathbf{H}_1^*\mathbf{H}_1^{*\prime} dr \end{bmatrix} - \begin{bmatrix} \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta} & 0 \\ 0 & 0 \end{bmatrix}\right| \\
=\quad & \left|\lambda\left(\boldsymbol{\Sigma}_{\beta\beta}\right) - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right| \left|\lambda \int_0^1 \mathbf{H}_1^*\mathbf{H}_1^{*\prime} dr\right|
\end{aligned}$$

which has q positive roots given by $\left|\lambda\boldsymbol{\Sigma}_{\beta\beta} - \boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}\right|$. Now, consider

$$\begin{aligned}
& \left|\left(\boldsymbol{\beta}^* : \overline{\boldsymbol{\gamma}}^* : T^{-1/2}\overline{\tau}^*\right)' S^* (\lambda) \left(\boldsymbol{\beta} : \overline{\boldsymbol{\gamma}}^* : T^{-1/2}\overline{\tau}^*\right)\right| \\
=\quad & \left|\boldsymbol{\beta}^{*\prime} S^* (\lambda) \boldsymbol{\beta}^*\right| \left|\left(\overline{\boldsymbol{\gamma}}^* : T^{-1/2}\overline{\tau}^*\right)' \left\{S^* (\lambda) - S^* (\lambda) \boldsymbol{\beta}^* \left[\boldsymbol{\beta}^{*\prime} S^* (\lambda) \boldsymbol{\beta}^*\right]^{-1} \boldsymbol{\beta}^{*\prime} S^* (\lambda)\right\} \left(\overline{\boldsymbol{\gamma}}^* : T^{-1/2}\overline{\tau}^*\right)\right|
\end{aligned}$$

where the first factor has no roots as $\lambda \to 0$ since $\boldsymbol{\beta}^{*\prime} S^{*}(\lambda)\boldsymbol{\beta}^{*} \to -\boldsymbol{\Sigma}_{\beta 0}\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{0\beta}$, and

$$\begin{aligned}
\left(\overline{\gamma}^{*} : T^{-1/2}\overline{\tau}^{*}\right)' S^{*}(\lambda)\left(\overline{\gamma}^{*} : T^{-1/2}\overline{\tau}^{*}\right) &\Rightarrow \rho\int_{0}^{1}\mathbf{H}_{1}^{*}\mathbf{H}_{1}^{*\prime}dr - \int_{0}^{1}\mathbf{H}_{1}^{*}d\mathbf{K}_{\mathbf{0},\mathbf{C\Sigma}}'(r)\,\boldsymbol{\Sigma}_{00}^{-1}\int_{0}^{1}d\mathbf{K}_{\mathbf{0},\mathbf{C\Sigma}}(r)\,\mathbf{H}_{1}^{*\prime} \\
\left(\overline{\gamma}^{*} : \sqrt{T}\overline{\tau}^{*}\right)' S^{*}(\lambda)\boldsymbol{\beta}^{*} &\Rightarrow \int_{0}^{1}\mathbf{H}_{1}^{*}d\mathbf{K}_{0,\mathbf{C\Sigma}}'(r)\,\boldsymbol{\Sigma}_{00}^{-1}\boldsymbol{\Sigma}_{\beta 0}
\end{aligned}$$

hence

$$\begin{aligned}
&\left(\overline{\gamma}^{*} : \sqrt{T}\overline{\tau}^{*}\right)'\left\{S^{*}(\lambda) - S^{*}(\lambda)\boldsymbol{\beta}^{*}\left[\boldsymbol{\beta}^{*\prime}S^{*}(\lambda)\boldsymbol{\beta}^{*}\right]^{-1}\boldsymbol{\beta}^{*\prime}S^{*\prime}(\lambda)\right\}\left(\overline{\gamma}^{*} : \sqrt{T}\overline{\tau}^{*}\right) \\
\Rightarrow\ &\rho\int_{0}^{1}\mathbf{H}_{1}^{*}\mathbf{H}_{1}^{*\prime}dr - \int_{0}^{1}\mathbf{H}_{1}^{*}d\mathbf{K}_{0,\boldsymbol{\Sigma}}'(r)\,\boldsymbol{\alpha}_{\perp}\left(\boldsymbol{\alpha}_{\perp}'\boldsymbol{\Omega}\boldsymbol{\alpha}_{\perp}\right)^{-1}\boldsymbol{\alpha}_{\perp}'\int_{0}^{1}d\mathbf{K}_{0,\boldsymbol{\Sigma}}(r)\,\mathbf{H}_{1}^{*\prime}
\end{aligned}$$

For $\boldsymbol{\psi}_{\beta} \neq \mathbf{0}$, then

$$T^{-1}\left(\boldsymbol{\beta}^{*} : \overline{\gamma}^{*} : T^{-1/2}\overline{\tau}^{*}\right)' S^{*}(\lambda)\left(\boldsymbol{\beta}^{*} : \overline{\gamma}^{*} : T^{-1/2}\overline{\tau}^{*}\right) \Rightarrow \lambda\begin{bmatrix} \frac{1}{12}\boldsymbol{\beta}'\boldsymbol{\psi}_{\beta}\boldsymbol{\psi}_{\beta}'\boldsymbol{\beta} & \boldsymbol{\beta}'\boldsymbol{\psi}_{\beta}\int_{0}^{1}(r-1/2)\,\mathbf{H}_{1}^{*\prime}(r)\,dr \\ \int_{0}^{1}\left(r-\frac{1}{2}\right)\mathbf{H}_{1}^{*}(r)\,\boldsymbol{\psi}_{\beta}'\boldsymbol{\beta}dr & \int_{0}^{1}\mathbf{H}_{1}^{*}(r)\,\mathbf{H}_{1}^{*\prime}(r)\,dr \end{bmatrix}$$

and again the determinant of this matrix has zero expectation., as under H_1.

- Define $\mathbf{V} = \left(\mathbf{V}_{\gamma}' : \mathbf{V}_{\psi}'\right)'$ with $\mathbf{V}_{\gamma} = \left[\overline{\gamma}'\mathbf{C\Omega C}'\overline{\gamma}\right]^{-1/2}\overline{\gamma}'\mathbf{C\Sigma W}$ and $\mathbf{V}_{\psi} = \left[\overline{\boldsymbol{\psi}}_{\beta_{\perp}}'\mathbf{C\Omega C}'\overline{\boldsymbol{\psi}}_{\beta_{\perp}}\right]^{-1/2}\overline{\boldsymbol{\psi}}_{\beta_{\perp}}'\mathbf{C\Sigma W}$. $\mathbf{V}$ is a standard Brownian Motion. Also, let

$$\begin{aligned}
\mathbf{G}(r) &= \begin{bmatrix} \left[\overline{\gamma}'\mathbf{C\Omega C}'\overline{\gamma}\right]^{-1/2} & 0 \\ 0 & \left[\overline{\boldsymbol{\psi}}_{\beta_{\perp}}'\mathbf{C\Omega C}'\overline{\boldsymbol{\psi}}_{\beta_{\perp}}\right]^{-1/2} \end{bmatrix}\mathbf{H}(r) = \mathbf{V}(r) + \begin{bmatrix} \mathbf{0} \\ \left[\overline{\boldsymbol{\psi}}_{\beta_{\perp}}'\mathbf{C\Omega C}'\overline{\boldsymbol{\psi}}_{\beta_{\perp}}\right]^{-1/2} r \end{bmatrix} \\
\mathbf{J}_{\boldsymbol{\psi}_{\alpha_{\perp}}}(r) &= \left(\boldsymbol{\alpha}_{\perp}'\boldsymbol{\Omega}\boldsymbol{\alpha}_{\perp}\right)^{-1/2}\boldsymbol{\alpha}_{\perp}'\mathbf{K}_{\boldsymbol{\psi},\boldsymbol{\Sigma}}(r) = \mathbf{V}(r) + \left(\boldsymbol{\alpha}_{\perp}'\boldsymbol{\Omega}\boldsymbol{\alpha}_{\perp}\right)^{-1/2}\boldsymbol{\alpha}_{\perp}'\boldsymbol{\psi}_{\alpha_{\perp}}r
\end{aligned}$$

The trace statistic therefore admits the following distributions

$$\mathsf{LR} \Rightarrow\ tr\left(\int_{0}^{1}\mathbf{G}d\mathbf{J}_{\boldsymbol{\psi}_{\alpha_{\perp}}}'\left[\int_{0}^{1}\mathbf{G}\mathbf{G}'\right]^{-1}\int_{0}^{1}\left(d\mathbf{J}_{\boldsymbol{\psi}_{\alpha_{\perp}}}\right)\mathbf{G}'\right)\quad if\ \boldsymbol{\psi}_{\beta} = 0 \tag{12}$$

Now, let $\mathbf{G}_{1}(r) = \mathbf{G}(r) - \int_{0}^{1}\mathbf{G}(u)\,du = \mathbf{V}(r) - \int_{0}^{1}\mathbf{V}(u)\,du + \left[\overline{\boldsymbol{\psi}}_{\beta_{\perp}}'\mathbf{C\Omega C}'\overline{\boldsymbol{\psi}}_{\beta_{\perp}}\right]^{-1/2}(r-1/2)$, and $\mathbf{G}_{2}(r) = \mathbf{V}(r) - \mathbf{a}_{v} - \mathbf{b}_{v}r$ where $\mathbf{a}_{v}, \mathbf{b}_{v}$ are coefficients correction $\mathbf{V}$ for a constant and a trend. Then

$$\begin{aligned}
\mathsf{LR}_{1} &\Rightarrow\ tr\left(\int_{0}^{1}\mathbf{G}_{1}d\mathbf{J}_{0}'\left[\int_{0}^{1}\mathbf{G}_{1}\mathbf{G}_{1}'\right]^{-1}\int_{0}^{1}(d\mathbf{J}_{0})\,\mathbf{G}_{1}'\right)\quad if\ \boldsymbol{\psi}_{\beta} = 0 \\
\mathsf{LR}_{2} &\Rightarrow\ tr\left(\int_{0}^{1}\mathbf{G}_{2}d\mathbf{J}_{0}'\left[\int_{0}^{1}\mathbf{G}_{2}\mathbf{G}_{2}'\right]^{-1}\int_{0}^{1}(d\mathbf{J}_{0})\,\mathbf{G}_{2}'\right)
\end{aligned}$$

and for the other two restricted hypotheses

$$\begin{aligned}
\mathsf{LR}^{*} &\Rightarrow\ tr\left(\int_{0}^{1}\mathbf{G}^{*}d\mathbf{J}_{(\boldsymbol{\Gamma\psi})_{\alpha_{\perp}}}'\left[\int_{0}^{1}\mathbf{G}^{*}\mathbf{G}^{*\prime}\right]^{-1}\int_{0}^{1}\left(d\mathbf{J}_{(\boldsymbol{\Gamma\psi})_{\alpha_{\perp}}}\right)\mathbf{G}^{*\prime}\right)\quad if\ \boldsymbol{\psi}_{\beta} = 0 \\
\mathsf{LR}_{1}^{*} &\Rightarrow\ tr\left(\int_{0}^{1}\mathbf{G}_{1}^{*}d\mathbf{J}_{0}'\left[\int_{0}^{1}\mathbf{G}_{1}^{*}\mathbf{G}_{1}^{*\prime}\right]^{-1}\int_{0}^{1}(d\mathbf{J}_{0})\,\mathbf{G}_{1}^{*\prime}\right)\quad if\ \boldsymbol{\psi}_{\beta} = 0
\end{aligned}$$