

Robust Cointegration Testing in the Presence of Weak Trends, with an Application to the Human Origin of Global Warming

Guillaume Chevillon*

ESSEC Business School

June, 2014

Abstract

Standard tests for the rank of cointegration of a vector autoregressive process present distributions that are affected by the presence of deterministic trends. We consider the recent approach of Demetrescu et al. (2009) who recommend testing a composite null. We assess this methodology in the presence of trends (linear or broken) whose magnitude is small enough not to be always detectable at conventional significance levels. We model them using local asymptotics and derive the properties of the test statistics. We show that whether the trend is orthogonal to the cointegrating vector has a major impact on the distributions but that the test combination approach remains valid. We apply of the methodology to the study of cointegration properties between global temperatures and the radiative forcing of human gas emissions. We find new evidence of Granger Causality.

Keywords: Cointegration, Deterministic trend, Likelihood ratio, Local trends, Global Warming.

JEL codes: C12, C32, Q54.

*Corresponding e-mail: chevillon@essec.edu; corresponding address: ESSEC Business School, Avenue Bernard Hirsch, BP 105, 95021 Cergy-Pontoise Cedex, France. I wish to thank Stéphane Grégoir, David Hendry, Bent Nielsen, and Robert Taylor for their valuable comments on an earlier version of the paper. I am grateful to Robert Kaufmann and David Stern for sharing their data with me. I also thank two anonymous referees, the Editor and Associate Editor, participants at the NBER/NSF Time Series Conference, as well as Nottingham and Brown University seminars. Research Support from CREST 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 of a linear trend, see Harvey, Leybourne and Taylor (2008) and the multivariate extension in Demetrescu, Lütkepohl and Saikkonen (2009). These studies draw on the long established difficulties in distinguishing in finite samples between stochastic and deterministic trends, see e.g. Sampson (1991), and Murray and Nelson (2000). 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. Perron and Campbell (1993) distinguish between “stochastic” and “deterministic” cointegration: only for the latter is the trend orthogonal to the cointegrating vector. In a simulation experiment, Toda (1994) shows that the Likelihood Ratio (LR) test (see Johansen, 1988, 1991 and 1994) 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. The bootstrap has also been suggested for testing the presence of cointegration in finite samples (see e.g. Li and Maddala, 1997, or Chang, Park and Song, 2002). It has been applied to the LR test, first, by simulation (see e.g. van Giesbergen, 1996, Harris and Judge, 1998, and Mantalos and Shukur, 2001), and then, analytically by Swensen (2006), Trenkler (2009) and Cavaliere, Rahbek and Taylor (2012). Yet, it appears that the issue of how to handle deterministic trends is not entirely settled: indeed Cavaliere, Taylor and Trenkler (2013) show that some techniques perform better than others. Their simulations show that the approach of Cavaliere et al. (2012) that does not rely on estimating deterministic components (since the statistics being bootstrapped are invariant to them) performs overall better than those of Swensen (2006) and Trenkler (2009) for the LR statistic, but that Trenkler’s algorithm is preferable for the LM test.

In view of the difficulties associated with disentangling stochastic and deterministic trends, Demetrescu, Lütkepohl and Saikkonen (2009) have proposed extending the work by Harvey, Leybourne and Taylor (2008) to the vector autoregressive (VAR) process. The technique they recommend consists in estimating two models with a deterministic trend that is either–or not–restricted to lie in the space orthogonal to the cointegrating vector; the test rejects the null if one of the statistics is significant. These authors show that their ‘combination’ methodology compares advan-

tageously to pretesting for the correct trend specification. In particular, when there is uncertainty about the presence and position of a deterministic trend, their simulations show that the technique selects the ‘correct’ cointegration rank most often.

The question we address here concerns the behavior of the combination method when the trend is hardly noticeable so its absence cannot be rejected at conventional significance levels. We assess whether the approach recommended by Demetrescu et al. (2009) still holds then. This situation may occur when the power of the t -test for the null of no linear trend is low, say less than 50%. Consider for instance a random walk with drift τ driven by normal innovations with variance σ^2 and where the estimator of τ is the mean of the differenced process over a sample of T observations. The power of the test for the null that $\tau = 0$ is about 50% when the noncentrality $\sqrt{T}\tau/\sigma$ is equal to the critical value (close to 2 in general for a 95% two-sided test, see Section 7.1 in the Appendix). This is for instance the magnitude of the values that Diebold and Senhadji (1996, Table 1) report for tests of the presence of a drift in the difference of the log U.S. Gross National Product (GNP) using four distinct measures over 1875-1993.

A convenient way to analyze the impact of “hardly noticeable” trends in this context is to derive asymptotic distributions where the parameter of the deterministic trend vanishes at the rate $O(T^{-1/2})$, as in the famous Pitman drift, so that both the stochastic and deterministic trends interact asymptotically. We consider in this paper the robustness to such *local trends* of the test for the rank of cointegration and show that very different behaviors result, depending on whether the data are stochastically or deterministically cointegrated.

Why do we assess the recommendation of Demetrescu et al. (2009) in the presence of local trends? The rationale behind their choice of test statistics is that they combine (i) one which is robust to trends but inefficient in their absence and (ii) another which is inconsistent in the presence of trends. Hence in the presence of a trend, the latter statistic terminates (by not rejecting the null) the conventional sequential cointegration rank testing strategy too early. The authors therefore ensure this does not happen in the overall testing strategy by combining the two tests. The results of these authors hold in the presence of deterministic trends whose presence is testable, but they show that their procedure is still preferable to pretesting for the trend. Clearly if pretesting is not available (e.g. under a local trend) then, *a fortiori*, the test combination approach should be recommended.

We wish to assess their results as they could require modifying under local trends. Our question here is twofold: under a local trend, (i) is the “inconsistent” statistic still inconsistent so its impact on the cointegration rank selection is limited, and (ii) is there a better choice of statistics to combine? The contribution of this paper is to show that although the distribution of the LR statistics differ under the alternative when the trend is considered local from when it is not, the recommendation of Demetrescu et al. (2009) still holds. Indeed, the LR statistics which correspond

to the models with an unrestricted constant (LR_μ in our notation below) or with a restricted linear trend (LR_T^*) achieve, here also, their desirable properties: the former is more efficient under a well specified model and we show it to be conservative under the local alternative; the latter is robust to the presence of trends. There does not seem to be test combinations that achieve better properties than the Demetrescu et al. proposal. We also extend the linear deterministic trend setting to allow for a stochastic trend driven in part by Bernoulli innovations with local asymptotic parameter. This allows us to show that the Demetrescu et al. proposal applies to a larger class of *weak* trends that include in particular a form of broken trend.

In order to assess the relevance of the *weak trends* framework (which we understand as comprising local linear trends), we provide an application of our results to the question of the human origin of global warming. We show that global temperatures and greenhouse gas emissions of human origin are integrated and we test whether they cointegrate. The presence of a weak trend constitutes an hypothesis which is difficult to rule out and which has been the subject of a rich and hotly debated literature. Our results point towards the presence of cointegration without requiring that we take a stance on the presence and position of a (possibly broken) deterministic trend. In this context, we test that human emissions Granger cause global temperatures.

The paper is organized as follows. Section 2 present the models for weak trends. 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 where we confirm the results of Demetrescu et al. (2009) in the presence of weak trends. We provide an empirical analysis of the human origin of global warming in Section 5. An appendix collects the proofs. Throughout the paper, row vectors are denoted by $(a : b)$; also, for any $(p \times q)$ matrix α of full rank $q \leq p$, we define α_\perp of dimension $p \times (p - q)$ such that $(\alpha : \alpha_\perp)$ is of full rank p . We also let the generalized projection operator $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$. The space spanned by linear combinations of the columns of a matrix α is denoted by $sp(\alpha)$; the rank of a matrix \mathbf{A} is written $rk(\mathbf{A})$.

2 The models

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

$$\Delta \mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{x}_{t-i} + \epsilon_t. \quad (1)$$

Assume that the disturbances ϵ_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 $I(1)$ and $\mathbf{\Pi}$ is of reduced rank q , then there exist α and β of order $(p \times q)$ such that $\mathbf{\Pi} = \alpha \beta'$ and $\beta' \mathbf{x}_t - \mathbf{E}[\beta' \mathbf{x}_t]$ is stationary. \mathbf{x}_t is then said to cointegrate, with cointegrating vector β . We also let

$\mathbf{x}_0 = \mathbf{0}$ in (1). Although this is not an inconsequential assumption (see Müller and Elliott, 2003), Ahlgren and Juselius (2012) show that it does not constitute the choice that is artificially most favorable to this paper's power analyses.

2.1 Local linear trend

We define \mathbf{y}_t as the sum of \mathbf{x}_t and of a deterministic trend which we assume *local*, i.e. satisfying the definition:

$$\mathbf{y}_t = \mathbf{x}_t + \mu + \psi \frac{t}{\sqrt{T}} = \mathbf{x}_t + \mathbf{\Psi} \mathbf{d}_{T,t} \quad (2)$$

where $\mathbf{\Psi} = (\mu : \psi)$ is a matrix of dimension $p \times 2$, $\mathbf{d}_{T,t} = (1 : T^{-1/2}t)'$. Throughout the paper, we rely on triangular arrays that are parameterized by the sample size T ; for notational ease, we do not write explicitly this dependence and let $\mathbf{d}_t = \mathbf{d}_{T,t}$. Then \mathbf{y}_t admits the following moving average representation : $\mathbf{y}_t = \mathbf{C} \sum_{i=1}^t \epsilon_i + \mathbf{\Psi} \mathbf{d}_t + \mathbf{C}_1(L) \epsilon_t + \mathbf{A}$ (Johansen, 1995), where L is the lag operator, $\mathbf{C} = \beta_\perp (\alpha'_\perp \mathbf{\Gamma} \beta_\perp)^{-1} \alpha'_\perp$, $\mathbf{\Gamma} = \mathbf{I}_p - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i$, the power series for $\mathbf{C}_1(z)$ is convergent for $|z| < 1 + \delta$ for some $\delta > 0$ and $\beta' \mathbf{A} = 0$. Expressions (1) and (2) together imply that \mathbf{y}_t admits the equilibrium correction:

$$\Delta \mathbf{y}_t = -\mathbf{\Pi} \mu + T^{-1/2} \mathbf{\Gamma} \psi + \alpha \left(\beta' \mathbf{y}_{t-1} - T^{-1/2} \beta' \psi (t-1) \right) + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{y}_{t-i} + \epsilon_t. \quad (3)$$

Here, we only consider deterministic trends that are *local* since our purpose is to study the robustness of the cointegration tests to the presence of potential deterministic misspecification. We follow in this Lütkepohl and Saikkonen (2000) and differ from analyses such as in Johansen (1995), chapter 14, Rahbek (1994) and Saikkonen and Lütkepohl (1999) whose focus is on the power of the test for the rank of cointegration vis-à-vis a locally larger rank.

The process \mathbf{y}_t , suitably scaled and interpolated, admits a weak limit that follows from a straightforward multivariate extension of the random walk with a local drift, drawing on Haldrup and Hylleberg (1995) and Stock and Watson (1996).¹ Define \mathbf{U}_T in $D^p[0,1]$, such that $\forall r \in [0,1]$, $\mathbf{U}_T(r) = T^{-1/2} \sum_{i=0}^{[Tr]} \epsilon_i \Rightarrow \mathbf{\Sigma} \mathbf{W}(r)$, as $T \rightarrow \infty$, where \mathbf{W} is a standard Brownian motion on $C^p[0,1]$. Then $T^{-1/2} \mathbf{y}_{[Tr]}$ retains asymptotically the sum of both the stochastic and of the degenerate linear trend, i.e. for $r \in [0,1]$:

$$T^{-1/2} \mathbf{y}_{[Tr]} \Rightarrow \mathbf{C} \mathbf{\Sigma} \mathbf{W}(r) + \psi r \stackrel{def}{=} \mathbf{K}_{\psi, \mathbf{C} \mathbf{\Sigma}}(r) \quad (4)$$

where $\mathbf{K}_{\psi, \mathbf{C} \mathbf{\Sigma}}$ is a Brownian motion with drift. Motivated by expressions (3) and (4), and by our interest in allowing for a local trend, we assume in the rest of the paper that $\mathbf{y}_0 = \mathbf{0}$, i.e. $\mu = \mathbf{0}$.

¹We let, as usual, $[w]$ denote the integer part of w for any real scalar w ; $D^p[0,1]$ is the space of \mathbb{R}^p -valued functions on the interval $[0,1]$ which are right continuous and have finite left limits (càdlàg); $C^p[0,1]$ is the subspace of $D^p[0,1]$ of continuous functions; and ' \Rightarrow ' denotes weak convergence of the associated probability measure.

2.2 Alternative weak broken trends

We also consider a slightly different setting of *weak* broken trends since the recent literature has shown that they can be confused with stochastic trends (and hence local linear trends), see e.g. Pretis and Hendry (2012). These alternative settings yield processes that differ in finite samples from the local linear trend presented previously. Yet they, purposefully, coincide all asymptotically. These extensions introduce the notion of *weak trends* and hence show that our theoretical results accommodate rich finite sample dynamics.

To present the alternative trends we consider, we introduce random-level and random-slope processes (see Harvey and Todd, 1983) with infrequent shocks as in Balke and Fomby (1991).

Assumption 1 *We define the following processes for $t = 1, \dots, T$ and $p \in \mathbb{N}$:*

(i) $\Delta\delta_t = \mathbf{q}_t\nu_t$, $\delta_0 = \mathbf{0}$, \mathbf{q}_t is a diagonal matrix process with diagonal that follows an i.i.d. Bernoulli ($T^{-1/2}\psi$), $\psi \in \mathbb{R}^p$ and ν_t is an i.i.d p -vector normal process with unit expectation and variance-covariance $\mathbf{\Omega}_\nu$ that is non-degenerate and of full rank.

(ii) $\Delta\delta_t^* = \mathbf{q}_t^*\nu_t^*$, $\delta_0^* = T^{-1/2}\psi$, \mathbf{q}_t^* is a diagonal matrix with diagonal that follows an i.i.d. Bernoulli ($T^{-1/2}\pi$), $\pi \in \mathbb{R}^p$ and ν_t^* is i.i.d normal with zero expectation and variance-covariance $T^{-1}\mathbf{\Omega}_\nu$.

Processes δ_t and δ_t^* define multivariate random walks with non-Gaussian innovations and, in the case of δ_t , a nonzero stochastic drift (since $\delta_t = \mathbf{q}_t + \delta_{t-1} + \mathbf{q}_t(\nu_t - 1)$). Their local asymptotic formulation is borrowed from Diebold and Inoue (2001) and Perron and Qu (2007, 2010) who — in a different parameterization where the parameters of Bernoulli processes are $O(T^{-1})$ and $E[\delta_t] = E[\delta_t^*] = 0$ — show that irregular breaks can yield processes with properties akin to those of fractionally integrated processes. Elliott and Müller (2006) also consider related processes, but with a parameterization that ensures the process behaves as a pure stochastic trend (with zero drift). Here, by contrast, Assumption 1 implies that the following processes exhibit local linear trends:

$$\begin{cases} \delta_t = \psi \frac{t}{\sqrt{T}} + O_p(T^{1/4}) \\ \sum_{i=1}^t \delta_i^* = \psi \frac{t}{\sqrt{T}} + O_p(T^{1/4}) \end{cases}$$

We define the processes

$$\begin{cases} \mathbf{z}_t = \mathbf{x}_t + \delta_t; \\ \mathbf{w}_t = \mathbf{x}_t + \sum_{i=1}^t \delta_i^*, \end{cases} \quad (5)$$

Since, for a fixed T , $\Delta\delta_t$ is covariance stationary, we refer to δ_t and $\sum_{i=1}^t \delta_i^*$ as I(1) and I(2)

Bernoulli trends. The processes asymptotically satisfy:²

$$\begin{cases} T^{-1/2} \mathbf{z}_{[Tr]} \Rightarrow \mathbf{K}_{\psi, \mathbf{C}\Sigma}(r) \\ T^{-1/2} \mathbf{w}_{[Tr]} \Rightarrow \mathbf{K}_{\psi, \mathbf{C}\Sigma}(r) \end{cases} \quad (6)$$

Figure 1 presents realizations of univariate processes (x_t, y_t, z_t, w_t) where $T = 100, 250$ and 1000 and $\sigma_\nu^2 = 1$. The third column shows the I(1) and I(2) Bernoulli trends δ_t and $\sum_i \delta_i^*$ together with the local linear trend t/\sqrt{T} . The figure exemplifies why we referred to δ_t and $\sum_i \delta_i^*$ as broken trends. The shifts in δ_t occur irregularly and the process appears piecewise constant, but the magnitude of the drift asymptotically renders the process comparable to a local linear trend. The local trend in the piecewise affine $\sum_i \delta_i^*$ exhibits stochastically shifting slopes.

In the following we use the term *weakly* trending to refer to the behaviors of \mathbf{y}_t , \mathbf{z}_t and \mathbf{w}_t . The three parametric models of local deterministic trends provide asymptotically comparable properties so a wide variety of finite sample behaviors can be accommodated within our framework (the models presented above do not constitute an exhaustive class).

3 Cointegration tests

In this section, we derive the distributions of test statistics when the process exhibits a weak trend and show that Proposition 3.1 of Demetrescu et al. (2009) which justifies their combination of tests also holds here. The null hypothesis is that of $\mathbf{\Pi}$ in (1) being of rank $q < p$, and we allow for different models for deterministic terms. We assess the robustness of the test statistics in the presence of deterministic misspecification. We therefore do not treat deterministic specification as part of a composite null that would also include the rank of cointegration. Trends are seen as a nuisance for inference on cointegration.

If the modeler assumes that the DGP exhibits no deterministic components, then reduced rank regression of $\Delta \mathbf{y}_t$ on \mathbf{y}_{t-1} corrected for the lagged differences leads to computing the likelihood ratio (LR) test statistic

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

where the eigenvalues λ_i are estimated as solutions to the problem $|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = |S(\lambda)| = 0$, 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 $\Delta \mathbf{y}_t$. We define the statistic in (7)

²These are solutions to:

$$\begin{cases} \Delta \mathbf{z}_t = T^{-1/2} \left(\Delta \delta_t - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \delta_{t-i} \right) + \alpha \left(\beta' \mathbf{z}_{t-1} - T^{-1/2} \beta' \delta_{t-1} \right) + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{z}_{t-i} + \epsilon_t; \\ \Delta \mathbf{w}_t = T^{-1/2} \left(\delta_t^* - \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \delta_{t-i}^* \right) + \alpha \left(\beta' \mathbf{w}_{t-1} - T^{-1/2} \beta' \sum_{i=1}^{t-1} \delta_i^* \right) + \sum_{i=1}^{k-1} \mathbf{\Gamma}_i \Delta \mathbf{w}_{t-i} + \epsilon_t. \end{cases}$$

as LR for the hypothesis $H_0 : rk(\mathbf{\Pi}) \leq q$ under model M that is given by (1) and (2) with $\Psi = \mathbf{0}$. Alternatively, the modeler may wish to use another of the LR statistics that have been proposed by S. Johansen to achieve robustness against $\Psi \neq \mathbf{0}$. The two main methods consist in (i) also including a constant in Z_{2t} , which provides the statistic LR_μ , or (ii) including both a constant and a linear trend in Z_{2t} , thus yielding LR_τ . Then LR_μ should be robust to $\Psi = (\mu, \psi)$, $\psi = \mathbf{0}$ (model M_μ) and LR_τ against the presence of a linear trend (M_τ). As is well known, deterministic terms lying within or outside the cointegration space impact processes differently. Hence the models which specify that, in the error correction form, the drift (model M_μ^*) or the trend (M_τ^*) are restricted to the cointegration space. These lead to the statistics LR_μ^* , or LR_τ^* respectively, for which, Z_{1t} is augmented of a constant, or a linear trend; Z_{2t} contains no deterministic terms, or a constant.

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

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

where $\mathbf{V}(r)$ is a standard Brownian motion of dimension $(p - q)$ and the notation ψ_{β_\perp} refers to the decomposition $\psi = \psi_\beta + \psi_{\beta_\perp}$ alongside the two orthogonal complements: $\psi_\beta = \overline{\beta} \beta' \psi$ and $\psi_{\beta_\perp} = \overline{\beta}_\perp \beta'_\perp \psi$. Let subscripts μ and τ denote correction for a mean, or a mean and trend, i.e. $\mathbf{V}_\mu = \mathbf{V}(r) - \int_0^1 \mathbf{V}(u) du$ and $\mathbf{V}_\tau(r) = \mathbf{V}(r) - \mathbf{a}_v - \mathbf{b}_v r$ where $\mathbf{a}_v, \mathbf{b}_v$ are the coefficients of the correction of \mathbf{V} for a constant and a trend. Augmented vectors are $\mathbf{V}^* = (\mathbf{V}'(r) : r)'$, with corresponding mean or trend correction, e.g. $\mathbf{V}_\mu^* = (\mathbf{V}'_\mu(r) : r - 1/2)'$. We also let $\mathbf{J}_{\psi_{\alpha_\perp}}(r) = (\alpha'_\perp \boldsymbol{\Omega} \alpha_\perp)^{-1/2} \alpha'_\perp \mathbf{K}_{\psi, \Sigma}(r) = \mathbf{V}(r) + (\alpha'_\perp \boldsymbol{\Omega} \alpha_\perp)^{-1/2} \alpha'_\perp \psi r$; $\mathbf{J}_{(\Gamma\psi)_{\alpha_\perp}}$ follows a similar definition. When $\psi_\beta \neq 0$ we project β onto the space spanned by the trend and let $\xi = \beta' \psi_\beta$. We then define the $(p - q + 1)$ dimensional standard Brownian motion $\mathbf{U} = \left(\boldsymbol{\Sigma}'_\xi \alpha'_\perp \boldsymbol{\Sigma}_{00} \alpha_\perp \boldsymbol{\Sigma}_\xi \right)^{-1/2} \boldsymbol{\Sigma}'_\xi \alpha'_\perp \boldsymbol{\Sigma}_{00}^{1/2} \mathbf{W}(r)$ where $\boldsymbol{\Sigma}_\xi$ is such that $(\boldsymbol{\Sigma}_{0\beta} \xi_\perp)_\perp = \alpha_\perp \boldsymbol{\Sigma}_\xi$. Finally, we let $\mathbf{I}_{\psi_{\alpha_\perp}}(r) = \mathbf{U}(r) + \left(\boldsymbol{\Sigma}'_\xi \alpha'_\perp \boldsymbol{\Sigma}_{00} \alpha_\perp \boldsymbol{\Sigma}_\xi \right)^{-1/2} \boldsymbol{\Sigma}'_\xi \alpha'_\perp \psi r$. We are now ready to state our main result.

Proposition 1 *Under the assumption that the DGP is generated as (1) and (2), with $\mathbf{y}_0 = \mathbf{0}$, the asymptotic distributions of the Likelihood Ratio test statistics are given by:*

if $\psi_\beta = \mathbf{0}$,

$$\begin{aligned} LR &\Rightarrow tr \left\{ \int_0^1 d\mathbf{J}_{\psi_{\alpha_\perp}} \mathbf{G}' \left[\int_0^1 \mathbf{G} \mathbf{G}' \right]^{-1} \int_0^1 \mathbf{G} d\mathbf{J}_{\psi_{\alpha_\perp}} \right\}, \\ LR_\mu &\Rightarrow tr \left\{ \int_0^1 d\mathbf{V} \mathbf{G}'_\mu \left[\int_0^1 \mathbf{G}_\mu \mathbf{G}'_\mu \right]^{-1} \int_0^1 \mathbf{G}_\mu d\mathbf{V}' \right\}. \end{aligned}$$

If $\psi_\beta \neq \mathbf{0}$, and $q > 1$, LR and LR_μ tend to the sum of the $p - q$ smallest eigenvalues of the

$(p - q + 1)$ -dimensional matrices

$$\begin{aligned} & \left[\begin{pmatrix} \mathbf{0} \\ \bar{\xi}' \Sigma_{\beta 0} \end{pmatrix}' + \int d\mathbf{I}_{\psi_{\alpha \perp}} \mathbf{G}^{*'} \right] \left(\int \mathbf{G}^* \mathbf{G}^{*'} \right)^{-1} \left[\int \mathbf{G}^* d\mathbf{I}'_{\psi_{\alpha \perp}} + \begin{pmatrix} \mathbf{0} \\ \bar{\xi}' \Sigma_{\beta 0} \end{pmatrix} \right] \\ & \left[\begin{pmatrix} \mathbf{0} \\ \bar{\xi}' \Sigma_{\beta 0} \end{pmatrix}' + \int d\mathbf{U} \mathbf{G}'_{\mu} \right] \left(\int \mathbf{G}'_{\mu} \mathbf{G}'_{\mu} \right)^{-1} \left[\int \mathbf{G}'_{\mu} d\mathbf{U}' + \begin{pmatrix} \mathbf{0} \\ \bar{\xi}' \Sigma_{\beta 0} \end{pmatrix} \right]. \end{aligned}$$

For all values of ψ_{β} ,

$$\begin{aligned} \text{LR}_{\mu}^* & \Rightarrow \text{tr} \left(\int_0^1 d\mathbf{J}_{(\Gamma\psi)_{\alpha \perp}} \mathbf{G}^{*'} \left[\int_0^1 \mathbf{G}^* \mathbf{G}^{*'} \right]^{-1} \int_0^1 \mathbf{G}^* d\mathbf{J}'_{(\Gamma\psi)_{\alpha \perp}} \right), \\ \text{LR}_{\tau}^* & \Rightarrow \text{tr} \left(\int_0^1 d\mathbf{V} \mathbf{V}' \left[\int_0^1 \mathbf{V}^* \mathbf{V}' \right]^{-1} \int_0^1 \mathbf{V}^* d\mathbf{V}' \right), \\ \text{LR}_{\tau} & \Rightarrow \text{tr} \left\{ \int_0^1 d\mathbf{V} \mathbf{V}'_{\tau} \left[\int_0^1 \mathbf{V}_{\tau} \mathbf{V}'_{\tau} \right]^{-1} \int_0^1 \mathbf{V}_{\tau} d\mathbf{V}' \right\}. \end{aligned}$$

Corollary 2 *Under the assumptions of Proposition 1, then whether $\psi_{\beta} = \mathbf{0}$ or not, as $\|\psi\| \rightarrow 0$ the asymptotic distributions of LR, LR_{μ} and LR_{μ}^* tend to their null distributions under correct specification of the deterministic components.*

Proposition 1 shows that the asymptotic distributions of LR_{τ} and LR_{τ}^* do not depend on ψ (as in Nielsen and Rahbek, 2000). By contrast LR, LR_{μ} and LR_{μ}^* lack robustness to the presence of a local trend, i.e. $\psi \neq \mathbf{0}$: their distributions no longer depend only on the number of common trends $p - q$ irrespective of the parameters of deterministic components. When $\psi_{\beta} = \mathbf{0}$, so $\psi_{\beta \perp} \neq 0$, misspecification impacts through the definition of the diffusion \mathbf{G} . For instance, in the case of the DGP considered in the simulations below (Section 4, experiment A) with $\psi_{\beta} = \mathbf{0}$ and $\psi_{\beta \perp} = \psi \mathbf{i}$ (\mathbf{i} is a vector consisting of zeros in $sp(\beta)$ and ones in $sp(\beta_{\perp})$), the coefficient of the deterministic trend component in \mathbf{G} is $\left(\bar{\psi}'_{\beta \perp} \mathbf{C} \Omega \mathbf{C}' \bar{\psi}_{\beta \perp} \right)^{-1/2} = \sqrt{\psi^2 (p - q)}$. Deterministic misspecification (in the form of a local trend) impacts the probability of rejection under the null.

Under models M, M_{μ} and M_{μ}^* , and when $\psi_{\beta} = \mathbf{0}$ in the DGP, i.e. what Perron and Campbell (1993) labeled “deterministic” cointegration and which can be treated using Johansen (1991), the test statistics have asymptotic distributions which are expressed as (mathematical) traces that are comparable to those under correct specification – although they now depend on the nuisance parameters $\psi_{\beta \perp}$ and $\psi_{\alpha \perp}$. In the presence of “stochastic” cointegration, when the linear trend is not entirely lying in $sp(\beta_{\perp})$: the systems under M and M_{μ} present $p - q + 1$ asymptotically zero eigenvalues (Proposition 3.1 of Demetrescu et al. (2009) states this result for model M_{μ} in the presence of a non-local trend); under the null of at most q cointegration relations, the statistic for a $q - 1$ rank has a distribution that can be expressed through a trace statistic.³ Under the null of

³It is possible to modify the LR test statistic by *regularization* (a form of shrinkage, or ridge regression) where

cointegration rank equal to q , the asymptotic distribution depends on the magnitude of ψ in the directions of β and α_{\perp} . Due to the complicated distributions which result, we observe the rejection rates via a Monte Carlo experiment in Section 4. Tests that account for the non-similarity of the test statistics with respect to ψ can in principle be obtained either (i) controlling the error rate while simulating the asymptotic distributions for a range of ψ (using e.g. a Bonferroni correction for the familywise error rate or the false discovery rate, see Benjamini and Hochberg, 1995); or (ii) using a median estimator for ψ by minimizing the distance between the observed statistic and the median of the asymptotic distribution (see Stock, 1991, and Andrews, 1993). We do not consider these options in the paper for the combination approach of Demetrescu et al. (2009) where the tests involving LR_{μ} use critical values computed under the null that $\psi = \mathbf{0}$.

The proposition also shows that short run dynamics (when $\Gamma \neq \mathbf{I}_p$) may affect the asymptotic distributions through two channels: one is through the definition of the matrix \mathbf{C} which appears in the definition of \mathbf{G} ; this concerns models \mathbf{M} , \mathbf{M}_{μ} and \mathbf{M}_{μ}^* when $\psi_{\beta_{\perp}} \neq 0$. Interestingly Γ also plays a role in the asymptotic distribution of LR_{μ}^* unless $\Gamma\psi$ lies in the space spanned by α – which is the case if the common stochastic trends exhibit no short run dynamics, i.e. $\alpha'_{\perp}\Delta\mathbf{y}_t = \epsilon_t$.

Corollary 2 shows that, no matter whether cointegration is deterministic or stochastic, if we let the trend coefficient tend to zero, the asymptotic distributions of the test statistics tend to those under the null for a correctly specified model.

Moreover, because of the difficulty associated with choosing the correct deterministic specification for the Likelihood Ratio test, H. Lütkepohl and P. Saikkonen have proposed in a series of papers (Lütkepohl and Saikkonen, 2000, L&S henceforth, and Saikkonen and Lütkepohl, 1999, 2000a, 2000b) an alternative Lagrange-Multiplier test. This LM test consists in estimating Ψ under the null hypothesis of q cointegrating relations, then detrending \mathbf{y}_t and testing for cointegration. We show in the Appendix, Section 7.2, that the procedure is robust to local trends.

Finally, the analytical results presented above also apply to the weak broken trends in processes \mathbf{z}_t and \mathbf{w}_t : the cointegration tests LR_{μ} , LR_{τ} and LR_{τ}^* applied to $\mathbf{z}_t, \mathbf{w}_t$ follow the same distributions as when applied to \mathbf{y}_t in Proposition 1, as the following corollary shows.

Corollary 3 *Define the processes \mathbf{z}_t and \mathbf{w}_t as in expression (5) under Assumption 1, where \mathbf{x}_t satisfies the assumptions of Proposition 1. Then the asymptotic distributions of the test statistics LR_{μ} , LR_{τ}^* and LR_{τ} are given by Proposition 1.*

the eigenvalue problem is modified to

$$\left| \lambda (S_{11} + R'R) - S_{10}S_{00}^{-1}S_{01} \right| = 0 \tag{9}$$

with appropriate choice of R . The matrix R has to be specified under the null for the rank of β and the space it spans; it involves estimating ψ_{β} and rotating the solutions to (9) away from the space spanned by ξ . For a discussion of Regularized Reduced Rank Regression, see Dobrev and Schamburg (2012). We do not consider this approach here as it turns out that the LR statistics exhibit robustness to $\psi_{\beta} \neq 0$.

The proof of the corollary follows the same lines as that of Proposition 1 and is omitted in the paper. It relies, for \mathbf{z}_t , on (6) and the fact that $\delta_t = \psi t/\sqrt{T} + o_p(\sqrt{T})$ and $(T^{-3/2} \sum_{t=1}^T \delta_t : T^{-5/2} \sum_{t=1}^T t\delta_t) \xrightarrow{m.s.} \psi(1/2 : 1/3)$.⁴ The proof for \mathbf{w}_t is similar. The procedure recommended by Demetrescu et al. hence holds also for a wide class of trending processes.

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 20,000 replications (computations were performed using Ox version 6.00, see Doornik, 2007; the code is available from the author's website). We consider three sets of experiments: first, a simple one based on a VAR(1) to assess the individual tests in the presence of a local trend which may or not be orthogonal to the cointegrating vector. We then consider the Demetrescu et al. joint testing procedure and compare it with the Lagrange-Multiplier technique of Lütkepohl and Saikkonen. Second, we extend the analysis to consider the more realistic VAR(2) DGP of Toda (1994, see Section 7.4 of the Appendix) considered by Demetrescu et al. and many authors. Finally, we also evaluate how the techniques fare in the presence of a weak Bernoulli trend.

4.1 LR tests without short run dynamics

4.1.1 Simulation setting

In the first set of simulations, we do not consider the effect of short-run dynamics and set $\mathbf{\Omega} = \mathbf{I}_p$, $k = 1$ in (1). We let, for a cointegration rank q , the vectors $\beta = -\alpha = (\mathbf{I}_q : \mathbf{0}_{q \times (p-q)})'$ where \mathbf{I}_q is the q dimensional unit matrix and $\mathbf{0}_{m \times n}$ a $(m \times n)$ -matrix of zeros. Then $\mathbf{\Pi} = \text{diag}(-\mathbf{I}_q, \mathbf{0}_{(p-q) \times (p-q)})$ and we choose $\alpha_{\perp} = \beta_{\perp} = (\mathbf{0}_{(p-q) \times q} : \mathbf{I}_{(p-q)})'$ and $\mathbf{C} = \beta_{\perp} (\alpha'_{\perp} \beta_{\perp})^{-1} \alpha'_{\perp} = \text{diag}(\mathbf{0}_{q \times q}, \mathbf{I}_{(p-q)})$. 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

⁴The proof follows from (when $p = 1$ so $\mathbf{\Omega}_\nu$ reduces to σ_ν^2)

$$T^{-1/2} \left(\mathbb{E} \left[T^{-1} \sum_{t=1}^T \delta_t \right], \text{Var} \left[T^{-1} \sum_{t=1}^T \delta_t \right] \right) \rightarrow \psi \left(\frac{1}{2}, \frac{\sigma_\nu^2}{3} \right),$$

and

$$T^{-1/2} \left(\mathbb{E} \left[T^{-2} \sum_{t=1}^T t\delta_t \right], \text{Var} \left[T^{-2} \sum_{t=1}^T t\delta_t \right] \right) \rightarrow \psi \left(\frac{1}{3}, \frac{2\sigma_\nu^2}{15} \right).$$

(non-deterministic) cointegration and a hybrid, as in

$$A : \psi = \psi (\mathbf{0}'_q : \mathbf{1}'_{p-q})' \quad (\text{deterministic cointegration})$$

$$B : \psi = \psi (\mathbf{1}'_q : \mathbf{0}'_{p-q})' \quad (\text{pure stochastic cointegration})$$

$$C : \psi = \psi \mathbf{1}'_p$$

for $\psi \in [0, 10]$, with $\mathbf{1}_{(p-q)}$ a vector $(1 : 1 : \dots)'$ of dimension $p - q$. Hence, in experiment A , $\psi_\beta = 0$ and $\psi_{\beta_\perp} = \psi_{\alpha_\perp} \neq 0$; in B , $\psi_\beta \neq 0$ and $\psi_{\beta_\perp} = \psi_{\alpha_\perp} = 0$ and in C , ψ_β and ψ_{β_\perp} are both nonzero. In this setting, the coefficient of the trend in \mathbf{G} in (8) is $(\overline{\psi}'_{\beta_\perp} \mathbf{C} \boldsymbol{\Omega} \mathbf{C}' \overline{\psi}_{\beta_\perp})^{-1/2} = \sqrt{\psi^2 (p - q)}$ in experiments A and C , and 0 in B . We show in Section 7.1 of the Appendix that a univariate test for the presence of a deterministic trend has 50% power when $\tau/\sigma \approx 2T^{-1/2}$, so the values of interest are here for $\psi\sqrt{p - q} \in (0, 1)$.

As Nielsen (1997) and Doornik (1998) point out, there exist several ways to compute the LR statistics that are equivalent asymptotically but differ numerically in finite samples; this issue is relevant here since localizing parameters depend on the sample size. We therefore compare two computational techniques: the first involves solving numerically the eigen problem $|S(\lambda)|$ and computing the statistics as $-T \sum_{i=q+1}^p \log(1 - \lambda_i)$ where the eigenvalues are ordered decreasingly $\lambda_1 > \lambda_2 > \dots > \lambda_p$; the second consists in computing a Cholesky decomposition of S_{11}^{-1} into PP' and obtaining the trace of $PS_{10}S_{00}^{-1}S_{10}P'$. Nielsen (1997) shows that first technique converges faster to the asymptotic distribution, unreported simulations show that it also the case in the presence of weak trends, hence we use it to compute the results.

4.1.2 Individual LR tests

We report in Table 1 the rejection probabilities of the null of q cointegrating relations out of $p = 6$ variables at a 5% nominal size (results for other sizes are comparable). In order to reduce Monte Carlo and finite sample variability, we compute the critical values under the corresponding null using the distribution of \mathbf{x}_t . This allows to focus specifically on robustness with respect to the local trend although it does not appear to affect our results.⁵ In the simulations, we let the sample vary from 100 to 1000 but with very little impact on the conclusions. As discussed previously and under the simulation specifications, the t -test for the null of no drift in a differenced random walk model has an approximate 50% rejection probability at the 5% nominal level under the alternative that $\psi = 2$. Hence values of ψ as large as 5 should be detected by the modeler.

As expected from Proposition 1, LR_τ^* and LR_τ are robust to local trends in finite samples. The table shows that the effect of trends orthogonal to the cointegrating space starts affecting LR_τ^* for

⁵A feasible alternative method would consist in using the bootstrap proposed by Cavaliere et al. (2012). We find in the experiment that the simulated critical values are very close to the asymptotic values reported by MacKinnon, Haug and Michelis (1999).

values of $\psi\sqrt{p-q}$ reaching the value of 5.

Since robustness is understood here as rejection frequencies that remain close to the nominal size, non robustness can arise with the test being conservative or liberal. If the null is under (over) rejected, the modeler is led to under (over) estimate the cointegration rank. The table shows that LR, LR_μ are inconsistent in all experiments even for low values of ψ . By contrast, LR_μ^* remains conservative in the presence of a local trend as in Demetrescu et al. 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_μ is more robust and LR and LR_μ^* reject even more. Clearly *A* and *C* coincide in this case. Also, experiment *B* then implies that $\psi = 0$ since $\beta_\perp = \mathbf{I}_p$.

The simulation experiment argues for the use of LR_μ in the combination of tests proposed by Demetrescu et al., and against the use of the either LR_μ^* or LR.

4.2 Combined LR and LM tests

We now assess the implication of the analysis above for the combination of LR tests. Table 2 reports the success rate in correctly choosing the rank of cointegration via either the joint $\text{LR}_\mu \cap \text{LR}_\tau^*$ test or the LM test. We use the standard sequential testing procedure starting with zero cointegration rank. Critical values are computed for the individual tests for \mathbf{x}_t using the same sample size: they are close to the asymptotic values reported in MacKinnon et al. (1999) and Lütkepohl and Saikkonen (2000). The conclusions are similar when using the asymptotic critical values. For comparability with the Toda (1994) DGP below, we set $p = 5$.

The table shows that whereas the LM more accurately selects the correct rank of cointegration in large samples (here $T = 1,000$), it suffers in small samples: the sequential testing procedure terminates too soon and the rank of cointegration is underestimated. By contrast, the combined LR methodology determines the rank, with very reasonable success rate irrespective of the position of the linear trend and the sample size.

We now consider extending the DGP into two directions (but retain the critical values previously computed). First, we consider the VAR(2) simulation setting of Toda (1994) used by Demetrescu et al. (see Section 7.4 of the Appendix). They consider linear trends whose “nonlocal” parameters are 0.1 (so $\psi = 1$ or 1.6 for $T = 100$ or 250) and 0.5 ($\psi = 5$ or 7.9 for $T = 100$ or 250). Here we consider smaller drifts (ψ less than 2). Table 3 records the simulated rejection frequencies of the combined tests when $\psi = \psi \mathbf{1}'_p$ (as in Experiment *C*). The simulations show that serial correlation impacts finite sample inference. When $\psi = .5$, joint testing of $\text{LR}_\mu \cap \text{LR}_\tau^*$ leads to under selection when $T = 100$, but not when $T = 1,000$. Comparing the table with those reported in Demetrescu et al. (their Table 3), we see that the presence of a trend that is local leads to similar accuracy in

selecting the rank of cointegration in finite samples.

By contrast, the presence of short-run dynamics has significant impact on the selected cointegration rank using the LM test. The procedure then terminates too early, even in significantly large samples. Unreported simulations show that the LM test performs much better when k is known: the impact of the Akaike Information Criterion for selecting the lag order is non negligible in this context.

Finally, the lower rows consider apply the tests to the weakly trending process \mathbf{z}_t (as in Experiment C again). The conclusions are similar except that the procedure performs less well in finite samples when only one stochastic trend is present: the rank of cointegration is only correctly selected in 80% of the experiments. To understand how differently a weak $I(1)$ Bernoulli trend impacts the finite sample statistics, Table 4 records the simulated quantiles at probability .95 of the distributions of LR_μ and LR_τ^* applied to \mathbf{x}_t , \mathbf{y}_t and \mathbf{z}_t for samples of size 100 and 1000. Here $\psi = \psi \mathbf{1}'_p$ (Experiment C) with $\psi = 1$. The table shows that the LR_τ^* quantiles are virtually identical for all processes. Those for LR_μ are larger for \mathbf{y}_t and \mathbf{z}_t than for \mathbf{x}_t , yet they are much less so for \mathbf{z}_t than for \mathbf{x}_t .⁶ Hence if the trend is under-specified, there is a small asymptotic property that the sequential testing strategy terminates too early under model M_μ .

5 Application: the Human Origin of Global Warming

We now apply the theoretical results to the vexed issue of the human origin of global warming. There exists a relatively large literature that debates the time series properties of global surface temperatures and human factors hypothesized to cause changes in temperatures (measured as radiative forcing of greenhouse gases, aerosols, and solar insolation, see e.g. Kaufmann, Kauppi and Stock, 2010). A key question is whether these series are better described as difference or trend-stationary, so cointegration and/or Granger causality tests can be performed. This literature originates in Bloomfield and Nychka (1992) but various authors find mixed results depending on which test is used. Kaufmann, Kauppi and Stock (2006) and Kaufmann and Stern (2002) conclude that surface temperatures cointegrate with the radiative forcing of human emissions. This finding was subsequently challenged by Gay, Estrada and Sanchez (2009) who conclude that temperatures are stationary around a broken deterministic trend. Kaufmann, Kauppi and Stock (2010) and Stern and Kaufmann (2014) disagree and confirm their evidence of cointegration and Granger causality from radiative forcing of human origin to surface temperatures. In a related study, Beenstock, Reingewertz and Paldor (2012) conclude that cointegration is impossible as evidence shows anthropogenic radiative forcing to be integrated of order 2 whereas temperature only contain one unit root. This evidence is challenged by Pretis and Hendry (2012) who show that the series for

⁶Note that the critical values for LR_μ reported in the literature vary to a large extent for low values of $p - q$.

radiative forcing exhibits breaks in recent periods due to changes in measurement methodologies (the dates differ across gases). Stern and Kaufmann (2014) also invoke a similar argument.

Here, we revisit the issue of cointegration between temperatures and anthropogenic radiative forcing in the light of our results on tests in the presence of weak trends. Figure 2 presents the time series⁷ of Temperature and Radiative Forcing of human origin that are used in Stern and Kaufmann (2014). The data span 1850-2011. In the figure, the levels of Temperature and Radiative Forcing were adjusted so their means and ranges match. The series in level present some long term comovements as well as an overall upward pattern (a linear trend is added to the graph, the data are scaled so the linear trend in either coincide).⁸ In light of the stochastic versus piecewise deterministic trends controversy, a question of debated interest is whether the processes may cointegrate and present a deterministic trend of small magnitude (whose position with respect to the potential cointegrating vector is unclear); the latter being subject to potential shifts occurring, say, over the 1960-1980 period. The formulation of our question hence differs from approaches that aim to detect breaks and perform inference in deterministic (possibly nonlinear) trends in temperatures (see e.g. Gao and Hawthorne, 2006, or Wu and Zhao, 2007), in such a context, the question could be whether broken trends with similar shapes drive both temperatures and radiative forcing. Here, by contrast, our interest lies in establishing whether a common stochastic trend drives the variables, and we aim to use a statistical technique that is robust to deterministic trends (possibly broken, as we showed in section 2.2 and Corollary 3) of small magnitude.

To assess the relevance of the weak trend framework, we fit autoregressive models of order 3, $AR(3)$, to the differences in *Temperature* and *RF*.⁹ The drifts are not significant, with t -values of 1.06 for $\Delta Temperature$ and 1.80 for ΔRF , so their associated p -values are 0.29 and 0.07. Hence the trend is not significant at the 5% level. The p -values are compatible with the local trend framework and possibly, the presence of a weak broken trend.

We perform a VAR(4) analysis for the levels of *Temperature* and *RF* (this corresponds to the minimum lag order for which residuals are not tested serially correlated at the 5% level). We follow Stern and Kaufmann and consider in addition their data for radiative forcing of natural origin (mostly volcanic and organic), denoted *RFnat*. The latter variable is considered stationary, and possibly heteroskedastic. To reduce the dimensionality of the system, we let *RFnat* enter

⁷The series reported here are those of Stern and Kaufmann (2014), HADCRUT4 (*Temperature*) and RFANTH (*RF*).

⁸For an explanation of the apparent recent disconnect, see Kaufmann, Kauppi, Mann and Stock (2011).

⁹We started from AR(4) processes, for which F -tests do not reject at the 5% level that residuals do not serially correlate with their first two lags. Fourth lags were tested insignificant so models were reduced to AR(3), still with insignificant residual autocorrelation. AR(3) models also minimize the Schwarz Criterion within the class of AR models up to lag order 7 for both variables.

All computations were performed using Oxmetrics 7.0 and the Ox programming language, see Doornik (2006) and Doornik and Hendry (2006).

unrestrictedly. We therefore estimate 28 parameters (30 when a trend is included) over a sample of 158 observations. Table 5 records the trace statistics for the model with (un)restricted intercepts and trends. Models with trends conclude to the presence of a unique stochastic trend, whereas M_μ and M_μ^* provide weaker evidence, rejecting the presence of a stochastic trend at the 5% level, but not at the 1%. Joint testing using $LR_\mu \cap LR_\tau^*$ hence concludes only at the 1% level to the presence of one cointegration relation and one stochastic trend.

In this system, under model M_μ , the selected cointegration vector is (with standard deviations in parentheses):

$$c_t = Temperature_t - \underset{(0.14)}{0.71} RF_t,$$

which is recorded in Figure 3. RF_t is tested weakly exogenous for the coefficients in the *Temperature* equation (at the 1%). A parsimonious multiple equation model for $\Delta Temperature_t$ and ΔRF_t is derived¹⁰ as

$$\begin{aligned} \Delta Temperature_t &= -\underset{(0.017)}{0.02} - \underset{(0.06)}{0.19} c_{t-1} - \underset{(0.014)}{0.05} RF_{nat_t} - \underset{(0.06)}{0.27} \Delta_3 Temperature_{t-1}, \\ \Delta RF_t &= \underset{(0.0017)}{0.003} + \underset{(0.013)}{0.03} \Delta Temperature_{t-1} + \underset{(0.07)}{0.43} \Delta RF_{t-1} + \underset{(0.07)}{0.30} \Delta RF_{t-3}, \end{aligned}$$

where $\Delta_3 x_t = x_t - x_{t-3}$. Figure 3 presents graphs for the fit of the model. The correlation between observed and fitted values is 0.54 and 0.63. The null that anthropogenic radiative forcing does not Granger cause global temperature is tested using a Likelihood Ratio test of the exclusion of c_t in the equation for $\Delta Temperature_t$. The test statistic is 9.8 with associated p -value of 0.002, thus rejecting long-run Granger non-causality.

This confirms the analysis of Stern and Kaufmann (2014) that human radiative forcing is a potential source of global warming. Yet, to establish this result formally, further study using various measures of global temperatures and radiative forcing must be carried out. Also, now that we have established that a weak trend may be present, statistical analyses of its specific form are warranted (possibly using the semi- or non-parametric techniques of Gao and Hawthorne, 2006, or Wu and Zhao, 2007). However, our previous theoretical and simulation results show that the cointegration testing procedure of Demetrescu et al. (2009) is reasonably robust to the presence of various forms of weak trends.

6 Conclusion

In this paper, we have studied the robustness of tests for the rank of cointegration in the VAR process with respect to misspecified local linear trends. This situation may correspond to data that

¹⁰Standard specification tests for joint serial correlation and heteroscedasticity do not reject; the p -value of the test for joint normality is 0.012. The p -value of the LR test of over-identifying restrictions is 0.94.

exhibit both stochastic and deterministic trends but the latter have a low magnitude that render them hardly noticeable and possibly insignificant. In this setting we have considered five versions of the likelihood ratio (LR) test and the Lagrange multiplier (LM) test. The LR with trend (LR_{τ}^* and LR_{τ}) 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). By contrast the LR statistic that do not allow for a trend (LR , LR_{μ} and LR_{μ}^*) only exhibit varying degrees of robustness, the latter also depending on the position of the trend with respect to the space spanned by cointegrating vectors. If the modeler aims to achieve robustness through a combination of tests, then as in Demetrescu et al. (2009) where overall rejection occurs if either test rejects, it is preferable to use one statistic with correct size under in the presence of a trend (LR_{τ}^*) and another with is more efficient in the absence of trend but that is conservative when the trend is not noticeable: LR_{μ} satisfies this requirement. Our results and simulations show that the recommendation of Demetrescu et al. (2009) also holds in the presence of trends with low impact. By contrast the LM test tends to underestimate the rank of cointegration in finite samples.

We show also that the weak trend framework applies to more general patterns than the pure linear trend. Indeed the trend driven by Bernoulli increments yields asymptotically identical distributions. This type of trend having locally nonzero expected change may represent the broken trends that are encountered empirically more accurately than the linear deterministic trend. In an application to the possible human origin of global warming, we find new results reinstating existing evidence that global temperature cointegrate with and may be Granger caused by anthropogenic emissions of greenhouse gases. We show in particular that this evidence is robust to the presence of weak trends, and possibly to breaks caused by changes in measurement.

References

- Ahlgren, N. and M. Juselius (2012). Tests for cointegration rank and the initial condition. *Empirical Economics* 42, 667691.
- Andrews, D. W. K. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. *Econometrica* 61(1), 139–65.
- Balke, N. S. and T. B. Fomby (1991). Shifting trends, segmented trends, and infrequent permanent shocks. *Journal of Monetary Economics* 28(1), 61 – 85.
- Beenstock, M., Y. Reingewertz, and N. Paldor (2012). Polynomial cointegration tests of anthropogenic impact on global warming. *Earth System Dynamics* 3, 173–188.
- Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and

- powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B* 57, 289–300.
- Bloomfield, P. and D. Nychka (1992). Climate spectra and detecting climate change. *Climatic Change* 21, 275–287.
- Cavaliere, G., A. Rahbek, and A. M. R. Taylor (2012). Bootstrap determination of the cointegration rank in vector autoregressive models. *Econometrica* 80(4), 1721–1740.
- Cavaliere, G., A. M. R. Taylor, and C. Trenkler (2013). Bootstrap cointegration rank testing: The role of deterministic variables and initial values in the bootstrap recursion. *Econometric Reviews* 32(7), 814–847.
- Demetrescu, M., H. Lütkepohl, and P. Saikkonen (2009). Testing for the cointegrating rank of a vector autoregressive process with an uncertain deterministic trend term. *Econometrics Journal* 12, 414–35.
- Diebold, F. X. and A. Inoue (2001). Long memory and regime switching. *Journal of Econometrics* 105(1), 131–159.
- Diebold, F. X. and A. S. Senhadji (1996). The uncertain unit root in real GNP: Comment. *The American Economic Review* 86(5), 1291–1298.
- Dobrev, D. and E. Schaumburg (2012). Robust forecasting by regularization. Working paper, Federal Reserve Board of Governors.
- Doornik, J. (1998). Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys* 12, 573–93.
- Doornik, J. A. (2006). *An Object-Oriented Matrix Language - Ox 4*. London: Timberlake Consultants Press.
- Doornik, J. A. (2007). *Object-Oriented Matrix Programming Using Ox, 3rd ed.* London: Timberlake Consultants Press and Oxford: www.doornik.com.
- Doornik, J. A. and D. F. Hendry (2006). *Empirical Econometric Modelling Using PcGive 11*. London: Timberlake Consultants Press.
- Elliott, G. and U. Müller (2006). Efficient tests for general persistent time variation in regression coefficients. *Review of Economic Studies* 73, 907–940.
- Gao, J. and K. Hawthorne (2006). Semiparametric estimation and testing of the trend of temperature series. *Econometrics Journal* 9, 332–355.

- Gay, C., F. Estrada, and A. Sanchez (2009). Global and hemispheric temperature revisited. *Climatic Change* 94, 333–349.
- Haldrup, N. and S. Hylleberg (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.
- Harris, R. and G. Judge (1998). Small sample testing for cointegration using the bootstrap approach. *Economics Letters* 58(1), 31–37.
- Harvey, A. C. and P. H. J. Todd (1983). Forecasting economic time series with structural and box-jenkins models: A case study. *Journal of Business Economic Statistics* 1(4), 299–307.
- Harvey, D. I., S. J. Leybourne, and A. M. R. Taylor (2009). Unit root testing in practice: dealing with uncertainty over the trend and initial condition (with commentaries and rejoinder). *Econometric Theory* 25, 587–636.
- Hubrich, K., H. Lütkepohl, and P. Saikkonen (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.
- Kaufmann, R. K., H. Kauppi, M. L. Mann, and J. H. Stock (2011). Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proceedings of the National Academy of Sciences* 108(29), 11790–11793.

- Kaufmann, R. K., H. Kauppi, and J. H. Stock (2006). Emissions, concentrations, temperature: a time series analysis. *Climatic Change* 77, 249–278.
- Kaufmann, R. K., H. Kauppi, and J. H. Stock (2010). Does temperature contain a stochastic trend? Evaluating conflicting statistical results. *Climatic Change* 101, 395–405.
- Kaufmann, R. K. and D. I. Stern (2002). Cointegration analysis of hemispheric temperature relations. *Journal of Geophysical Research* 107(D2), ACL 8–1 – ACL 8–10.
- Lütkepohl, H. and P. Saikkonen (2000). Testing for the cointegrating rank of a VAR process with a time trend. *Journal of Econometrics* 95, 177–98.
- MacKinnon, J. G., A. A. Haug, and L. Michelis (1999). Numerical distribution functions of likelihood ratio tests for cointegration. *Journal of Applied Econometrics* 14, 563–577.
- Mantalos, P. and G. Shukur (2001). Booststrapped johansen tests for cointegration relationships: a graphical analysis. *Journal of Statistical Computation and Simulation* 68(4), 351–371.
- Müller, U. K. and G. Elliott (2003). Tests for unit roots and the initial condition. *Econometrica* 71, 1269–86.
- Murray, C. J. and C. R. Nelson (2000). The uncertain trend in U.S. GDP. *Journal of Monetary Economics* 46, 79–95.
- Nielsen, B. (1997). Bartlett correction of the unit root test in autoregressive models. *Biometrika* 84, 500–04.
- Nielsen, B. and A. Rahbek (2000). Similarity issues in cointegration models. *Oxford Bulletin of Economics and Statistics* 62, 5–22.
- Perron, P. and J. Y. Campbell (1993). A note on Johansen’s cointegration procedure when trends are present. *Empirical Economics* 18, 777–89.
- Perron, P. and Z. Qu (2007). An analytical evaluation of the log-periodogram estimate in the presence of level shifts. working paper, Boston University.
- Perron, P. and Z. Qu (2010). Long-memory and level shifts in the volatility of stock market return indices. *Journal of Business and Economic Statistics* 28, 275–290.
- Pretis, F. and D. F. Hendry (2013). Comment on “Polynomial cointegration tests of anthropogenic impact on global warming” by Beenstock et al. (2012) some hazards in econometric modelling of climate change. *Earth System Dynamics* 4, 375–384.

- Rahbek, A. (1994). The power of some cointegration tests. Department of theoretical statistics, University of Copenhagen.
- Saikkonen, P. and H. Lütkepohl (1999). Local power of likelihood ratio tests for the cointegrating rank of a VAR process. *Econometric Theory* 15, 50–78.
- Saikkonen, P. and H. Lütkepohl (2000a). Testing for the cointegrating rank of a VAR process with an intercept. *Econometric Theory* 16, 373–406.
- Saikkonen, P. and H. Lütkepohl (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.
- Stern, D. I. and R. K. Kaufmann (2014). Anthropogenic and natural causes of climate change. *Climatic Change* 122, 277–269.
- Stock, J. H. (1991). Confidence intervals for the largest autoregressive root in u.s. macroeconomic time series. *Journal of Monetary Economics* 28(3), 435–459.
- Stock, J. H. and M. W. Watson (1996). Confidence sets in regressions with highly serially correlated regressors. Mimeo, Princeton University.
- Swensen, A. R. (2006). Bootstrap algorithms for testing and determining the cointegration rank in VAR models. *Econometrica* 74(6), 1699–1714.
- 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.
- Trenkler, C. (2009). Bootstrapping systems cointegration tests with a prior adjustment for deterministic terms. *Econometric Theory* 25(01), 243–269.
- van Giersbergen, N. (1996). Bootstrapping the trace statistic in VAR models: Monte Carlo results and applications. *Oxford Bulletin of Economics and Statistics* 58(2), 391–408.
- Wu, W. and Z. Zhao (2007). Inference of trends in time series. *Journal of the Royal Statistical Society, Series B* 69, 391–410.

7 Appendix

7.1 Power

Consider the random walk with drift process

$$\Delta y_t = \tau + \epsilon_t$$

where $\epsilon_i \stackrel{i.i.d}{\sim} N(0, \sigma^2)$, the t -statistic of $\hat{\tau}$ computed over a sample of $t = 1, \dots, T$ observations is

$$t_{\hat{\tau}} = \frac{\tau + \bar{\epsilon}_T}{\hat{\sigma}_{\hat{\tau}}}$$

where $\bar{\epsilon}_T = T^{-1} \sum_{t=1}^T \epsilon_t$ and $\hat{\sigma}_{\hat{\tau}} = \frac{T^{-1/2}}{T-1} \sum_{t=1}^T (\epsilon_t - \bar{\epsilon}_T)^2$. Let $\eta \sim N(0, 1)$ so

$$t_{\hat{\tau}} = T^{1/2} \frac{\tau}{\sigma} + T^{1/2} \eta + o_p(\sqrt{T})$$

The power function for a two-sided test with 0.05 size given by

$$\begin{aligned} \Pr(|t_{\hat{\tau}}| > 1.96) &= \Pr\left(\left|T^{1/2} \frac{\tau}{\sigma} + T^{1/2} \eta\right| > 1.96\right) \\ &= \Pr\left(T^{1/2} \frac{\tau}{\sigma} + T^{1/2} \eta > 1.96\right) + \Pr\left(T^{1/2} \frac{\tau}{\sigma} + T^{1/2} \eta < -1.96\right) \\ &= \Pr\left(\eta > 1.96T^{-1/2} - \frac{\tau}{\sigma}\right) + \Pr\left(\eta < -1.96T^{-1/2} - \frac{\tau}{\sigma}\right) \\ &= 1 - \Phi\left(1.96T^{-1/2} - \frac{\tau}{\sigma}\right) + \Phi\left(-1.96T^{-1/2} - \frac{\tau}{\sigma}\right) \end{aligned}$$

so the power is approximately 50% if $1.96T^{-1/2} - \frac{\tau}{\sigma} \approx 0$ since $\Phi\left(-1.96T^{-1/2} - \frac{\tau}{\sigma}\right)$ is then negligible.

Using a one-sided test yields similar results.

7.2 Lagrange Multiplier test

The LM test consists in estimating Ψ under the null hypothesis of q cointegrating relations, then detrending \mathbf{y}_t into $\tilde{\mathbf{x}}_t = \mathbf{y}_t - (\tilde{\mu} : \tilde{\psi})(1 : t)'$ and testing for $\rho_* = 0$ in a feasible version of

$$\alpha'_{\perp} \Delta \tilde{\mathbf{x}}_t = \rho_* \beta'_{\perp} \tilde{\mathbf{x}}_{t-1} + \sum_{i=1}^{k-1} \Gamma_{*,i} \Delta \tilde{\mathbf{x}}_{t-i} + \epsilon_{x,t}$$

In the presence of a local trend, $\tilde{\psi}$ is not consistent for ψ since it is an estimator of $T^{-1/2}\psi$. In the presence 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 $\tilde{\psi}$ in lemma (A.3) of Lütkepohl and Saikkonen (2000) (it is $\tilde{\mu}_1$ in their notation). $\sqrt{T} \beta'_{\perp} (\tilde{\psi} - T^{-1/2}\psi) \xrightarrow{L} \beta'_{\perp} \mathbf{C} \Sigma \mathbf{B}(1)$. $\tilde{\psi}$ is in turn used to derive the distribution of the sample moments of

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

hence the procedure provide the asymptotic distribution as in Lütkepohl and Saikkonen (2000, Section A.2).

7.3 Proof of Proposition 1.

7.3.1 Sample Moments

We follow the lines of the proofs Johansen (1991). We first let $\psi = \psi_\beta + \psi_{\beta_\perp}$ and choose γ orthogonal to β and ψ_{β_\perp} such that $(\beta : \gamma : \psi_{\beta_\perp})$ has full rank. When $\psi_\beta \neq \mathbf{0}$, we also decompose β into $\beta\xi$ and $\beta\xi_\perp$, with $\xi = \beta'\psi_\beta$ and ξ_\perp orthogonal to ξ within $sp(\beta)$.

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

and let also

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

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

$$\text{Var} \begin{bmatrix} \Delta \mathbf{x}_t \\ \beta' \mathbf{x}_{t-1} \end{bmatrix} = \begin{bmatrix} \Sigma_{00} & \Sigma_{0\beta} \\ \Sigma_{\beta 0} & \Sigma_{\beta\beta} \end{bmatrix} \left| \Delta \mathbf{x}_{t-1}, \dots, \Delta \mathbf{x}_{t-k+1} \right|$$

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 assume that the process is corrected for lagged values but, for notational simplicity, do not write it explicitly (this clearly affects the definition of \mathbf{W}). 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 (i.e. none under M, a constant under M_μ and both constant and trend under M_τ). We consider the models M, M_μ and M_τ 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 $T^{-2} \sum_{t=1}^T \left[\frac{t-(T+1)/2}{\sqrt{T}} \right]^2 = \frac{1}{12} - \frac{1}{12T^2}$. Under the hypotheses, the DGP rewrites as

$$\begin{aligned} \mathbf{M} : R_{0t} &= T^{-1/2} (\mathbf{\Gamma}\psi - \alpha\beta'\psi_\beta(t-1)) + \alpha\beta'R_{1t} + \hat{\epsilon}_t \\ &= T^{-1/2} \mathbf{\Gamma}\psi + \alpha\beta'\hat{\mathbf{x}}_t + \hat{\epsilon}_t \end{aligned}$$

$$\begin{aligned} \mathbf{M}_\mu : R_{0t} &= -T^{-1/2} \alpha\beta'\psi_\beta(t-1-T/2) + \alpha\beta'R_{1t} + \hat{\epsilon}_{\mu t} \\ &= \alpha\beta'\hat{\mathbf{x}}_{\mu t} + \hat{\epsilon}_{\mu t} \end{aligned}$$

$$\begin{aligned} \mathbf{M}_\tau : R_{0t} &= \alpha\beta'R_{1t} + \hat{\epsilon}_{\tau t} \\ &= \alpha\beta'\hat{\mathbf{x}}_{\tau t} + \hat{\epsilon}_{\tau t} \end{aligned}$$

where $\hat{\mathbf{x}}_{jt}, \hat{\epsilon}_{jt}$, for $j = \emptyset, \mu, \tau$, are residuals from the corrections of \mathbf{x}_t, ϵ_t for the lags $\Delta \mathbf{x}_{t-1}, \dots, \Delta \mathbf{x}_{t-k+1}$, and a constant ($j = \mu$) and possibly a trend also ($j = \tau$). Replacing $R_{1t} - \psi_\beta T^{-1/2}(t-1)$ with the same, corrected for a constant and possibly a trend leads to the same limit. Also, $T^{-1/2} \sum_{t=1}^T \hat{\epsilon}_{jt} \Rightarrow \Sigma W(1)$ for $j = \emptyset, \mu, \tau$.

Hence the following results.

First, under M, the residuals R_{it} are not corrected, hence, different limits result, depending on whether $\psi_\beta = \mathbf{0}$:

If $\psi_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \Sigma_{00}$, $\beta' S_{11} \beta \xrightarrow{p} \Sigma_{\beta\beta}$, $\beta' S_{10} \xrightarrow{p} \Sigma_{\beta 0}$ and

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

and for $\psi_\beta \neq \mathbf{0}$, then in the direction $\beta : S_{00} \xrightarrow{p} \Sigma_{00}$, and,

$$\beta' S_{11} \beta = T^{-1} \sum_{t=1}^T \left(\beta' \hat{\mathbf{x}}_{t-1} + \xi \frac{t-1}{\sqrt{T}} \right) \left(\beta' \hat{\mathbf{x}}_{t-1} + \xi \frac{t-1}{\sqrt{T}} \right)'$$

so

$$\begin{aligned} T^{-1} (\beta \bar{\xi})' S_{11} \beta \bar{\xi} &\xrightarrow{p} \frac{1}{3} \\ (\beta \xi_\perp)' S_{11} \beta \xi_\perp &\xrightarrow{p} \xi'_\perp \Sigma_{\beta\beta} \xi_\perp \\ (\beta \xi_\perp)' S_{11} \beta \bar{\xi} &= O_p(1) \end{aligned}$$

and the scaling by T^{-1} leads to an asymptotically singular matrix:

$$T^{-1} (\beta \bar{\xi} : \beta \xi_\perp)' S_{11} \beta (\bar{\xi} : \xi_\perp) \rightarrow \begin{bmatrix} \frac{1}{3} & \mathbf{0}_{1 \times (q-1)} \\ \mathbf{0}_{(q-1) \times 1} & \mathbf{0}_{(q-1) \times (q-1)} \end{bmatrix}$$

Recalling that $\Sigma_{\beta 0} = \Sigma_{\beta\beta}\alpha'$, $\beta' S_{10} \Rightarrow \Sigma_{\beta 0} + \xi \int_0^1 r d\mathbf{K}'_{\Gamma\psi, \mathbf{C}\Sigma}$. Now in the direction β_\perp :

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

also $T^{-1/2} (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{11} \beta \xi_\perp = o_p(1)$. We note that $\mathbf{K}_{\psi, \mathbf{C}\Sigma}$ satisfies the following stochastic differential equation:

$$d\mathbf{K}_{\psi, \mathbf{C}\Sigma}(r) = \alpha \beta' (\mathbf{K}_{\psi, \mathbf{C}\Sigma}(r) - \psi_\beta r) dr + \Sigma d\mathbf{W}(r)$$

so $\left[\int_0^1 \mathbf{H}(r) d\mathbf{K}'_{\psi, \mathbf{C}\Sigma}(r) \right] \alpha_\perp = \left[\int_0^1 \mathbf{H}(r) d\mathbf{K}'_{\psi, \Sigma}(r) \right] \alpha_\perp$ and $(\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{10} \alpha_\perp \Rightarrow \int_0^1 \mathbf{H}(r) d\mathbf{K}'_{\psi, \Sigma}(r) \alpha_\perp$.

This also applies in all cases below.

Under \mathbf{M}_μ , first if $\psi_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \Sigma_{00}$, $\beta' S_{11} \beta \xrightarrow{p} \Sigma_{\beta\beta}$, $\beta' S_{10} \xrightarrow{p} \Sigma_{\beta 0}$

and also

$$\begin{aligned} (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' (S_{10} - S_{11}\beta\alpha') &\Rightarrow \int_0^1 \mathbf{H}_\mu d\mathbf{K}'_{\mathbf{0}, \Sigma}(r) \\ (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{10} &\Rightarrow \int_0^1 \mathbf{H}_\mu d\mathbf{K}'_{\mathbf{0}, \mathbf{C}\Sigma}(r) \\ T^{-1} (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{11} (\bar{\gamma} : \bar{\psi}_{\beta_\perp}) &\Rightarrow \int_0^1 \mathbf{H}_\mu \mathbf{H}'_\mu dr \\ (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{11}\beta &= O_p(1) \end{aligned}$$

When $\psi_\beta \neq \mathbf{0}$, $S_{00} \xrightarrow{p} \Sigma_{00}$,

$$T^{-1} (\beta \bar{\xi}, \beta \xi_\perp)' S_{11} \beta (\bar{\xi}, \xi_\perp) \rightarrow \begin{bmatrix} \frac{1}{12} & 0_{1 \times q-1} \\ 0_{q-1 \times 1} & 0_{q-1 \times q-1} \end{bmatrix}$$

$\beta' S_{10} \Rightarrow \Sigma_{\beta 0} + \xi \int_0^1 (r - \frac{1}{2}) d\mathbf{K}'_{\mathbf{0}, \mathbf{C}\Sigma}(r)$ and

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

And finally, under \mathbf{M}_τ , whether or not $\psi_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \Sigma_{00}$, $\beta' S_{11} \beta \xrightarrow{p} \Sigma_{\beta\beta}$, $\beta' S_{10} \xrightarrow{p} \Sigma_{\beta 0}$ and

$$\begin{aligned} (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' (S_{10} - S_{11} \beta \alpha') &\Rightarrow \int_0^1 \mathbf{H}_\tau d\mathbf{K}'_{\mathbf{0}, \Sigma}(r) \\ (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{10} &\Rightarrow \int_0^1 \mathbf{H}_\tau d\mathbf{K}'_{\mathbf{0}, \mathbf{C}\Sigma}(r) \\ T^{-1} (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{11} (\bar{\gamma} : \bar{\psi}_{\beta_\perp}) &\Rightarrow \int_0^1 \mathbf{H}_\tau \mathbf{H}'_\tau dr \\ (\bar{\gamma} : \bar{\psi}_{\beta_\perp})' S_{11} \beta &= O_p(1) \end{aligned}$$

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

We complete the analysis above with the two restricted models \mathbf{M}_μ^* and \mathbf{M}_τ^* . Let, under \mathbf{M}_μ^* , $R_{2t} = 1$ and under \mathbf{M}_τ^* , $R_{2t} = t - 1 - T/2$. Now, let as before $\Gamma\psi = (\Gamma\psi)_\alpha + (\Gamma\psi)_{\alpha_\perp}$ with $(\Gamma\psi)_\alpha = \alpha (\alpha' \alpha)^{-1} \alpha' \Gamma\psi$ then

$$\begin{aligned} \mathbf{M}_\mu^* : R_{0t} &= T^{-1/2} \left((\Gamma\psi)_{\alpha_\perp} - \alpha \beta' \psi_\beta (t-1) \right) + \alpha \beta' R_{1t} + T^{-1/2} (\Gamma\psi)_\alpha R_{2t} + \epsilon_t \\ &= T^{-1/2} (\Gamma\psi)_{\alpha_\perp} + \alpha \left[\beta' \hat{\mathbf{x}}_t + (\alpha' \alpha)^{-1} \alpha' T^{-1/2} \Gamma\psi R_{2t} \right] + \hat{\epsilon}_t \\ \mathbf{M}_\tau^* : R_{0t} &= \alpha \beta' R_{1t} - T^{-1/2} \alpha \xi R_{2t} + \epsilon_t \\ &= \alpha \beta' \hat{\mathbf{x}}_{\mu t} + \hat{\epsilon}_{\mu t} \end{aligned}$$

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

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

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

if $\psi_\beta = \mathbf{0}$, then $S_{00} \xrightarrow{p} \Sigma_{00}$, $\beta^{*'} S_{11}^* \beta^* \xrightarrow{p} \Sigma_{\beta\beta} + \beta_2 \beta_2'$, $\beta^{*'} S_{10}^* \xrightarrow{p} \Sigma_{\beta 0}$ and

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

and for $\psi_\beta \neq \mathbf{0}$, then $S_{00} \xrightarrow{p} \Sigma_{00}$,

$$T^{-1} \left((\beta \bar{\xi} : \beta \xi_\perp)' : \beta_2 \right) S_{11}^* \left((\beta \bar{\xi} : \beta \xi_\perp)' : \beta_2 \right)' \xrightarrow{p} \begin{bmatrix} \frac{1}{3} & \mathbf{0}_{1 \times (q-1)} \\ \mathbf{0}_{(q-1) \times 1} & \mathbf{0}_{(q-1) \times (q-1)} \end{bmatrix}$$

and

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

with also

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

And finally under \mathbf{M}_τ^* , letting $R_{1t}^* = (R'_{1t} : R'_{2t})'$ and $\beta^{*'} = (\beta' : \xi)$. We also define the matrices $\gamma^{*'} = ([\gamma : \psi_{\beta \perp}]' : 0)$ and $\tau^{*'} = (\mathbf{0} : 1)$ such that $(\beta^* : \gamma^* : \tau^*)$ is of full rank $p+1$, then

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

Hence, if $\psi_\beta = \mathbf{0}$, $S_{00} \xrightarrow{p} \Sigma_{00}$, $\beta^{*'} S_{11}^* \beta^* \xrightarrow{p} \Sigma_{\beta\beta}$, $\beta^{*'} S_{10}^* \xrightarrow{p} \Sigma_{\beta 0}$ and, letting $\mathbf{H}_\mu^*(r) = (\mathbf{H}'_\mu(r) : r - 1/2)'$

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

and for $\psi_\beta \neq \mathbf{0}$,

$$(T^{-1} \bar{\xi} : \xi_\perp)' \beta^{*'} S_{11}^* \beta^* (T^{-1} \bar{\xi} : \xi_\perp) \xrightarrow{p} \begin{bmatrix} 1/12 & \mathbf{0} \\ \mathbf{0} & \xi'_\perp \Sigma_{\beta\beta} \xi_\perp \end{bmatrix}$$

with also

$$\begin{aligned} (\beta' : T^{-1/2} \xi) S_{10}^* &\Rightarrow \left[\Sigma_{\beta 0} + \xi \int_0^1 \left(r - \frac{1}{2} \right) d\mathbf{K}'_{\mathbf{0}, \Sigma}(r) \right] + \xi \int_0^1 \left(r - \frac{1}{2} \right) d\mathbf{K}'_{\mathbf{0}, \mathbf{C}\Sigma}(r) \\ T^{-1/2} (\beta' : \xi) S_{10}^* &\Rightarrow \xi \int_0^1 \left(r - \frac{1}{2} \right) d\mathbf{K}'_{\mathbf{0}, \mathbf{C}\Sigma}(r) \end{aligned}$$

hence

$$\begin{aligned} T^{-1} \left(\bar{\gamma}^* : T^{-1/2} \bar{\tau}^* \right)' (S_{10}^* - S_{11}^* \beta^* \alpha') &\Rightarrow - \int_0^1 \left(r - \frac{1}{2} \right) \mathbf{H}_\mu^*(r) \xi' \alpha' dr \\ \left(\bar{\gamma}^* : T^{-1/2} \bar{\tau}^* \right)' S_{10}^* &\Rightarrow \int_0^1 \mathbf{H}_\mu^*(r) d\mathbf{K}'_{\mathbf{0}, \mathbf{C}\Sigma}(r) \\ T^{-1} \left(\bar{\gamma}^* : T^{-1/2} \bar{\tau}^* \right)' S_{11}^* \left(\bar{\gamma}^* : T^{-1/2} \bar{\tau}^* \right) &\Rightarrow \int_0^1 \mathbf{H}_\mu^*(r) \mathbf{H}_\mu^{*'}(r) dr \\ T^{-3/2} \left(\bar{\gamma}^* : T^{-1/2} \bar{\tau}^* \right)' S_{11}^* \beta^* &\Rightarrow \left(\int_0^1 \left(r - \frac{1}{2} \right) \mathbf{H}_\mu^*(r) dr \right) \xi' \end{aligned}$$

7.3.2 LR statistic

We turn next to the asymptotic distribution of

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

where the trace statistic is the sum of the $p - q$ smallest solutions to the equation $|S(\lambda)| = 0$ (except for \mathbf{M}_μ^* and \mathbf{M}_τ^* which we treat later). The matrix $(\beta(\bar{\xi}: \xi_\perp) : \psi_{\beta_\perp} : \gamma)$ has full rank (when $\psi_\beta = 0$, $\bar{\xi} = 0$ and $\xi_\perp = \mathbf{I}_q$). Let $\mathbf{A}_T = (\beta(T^{-1/2}\bar{\xi}: \xi_\perp) : T^{-1/2}\psi_{\beta_\perp} : T^{-1/2}\gamma)$ such that we obtain distributions under the various hypotheses. We provide standardizing matrices at the end of the Section.

Under model \mathbf{M} , first if $\psi_\beta = 0$

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

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

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

where the first factor has no roots:

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

and, letting $\rho = T\lambda$

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

hence

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

where $\Sigma_{00}^{-1} - \Sigma_{00}^{-1} \Sigma_{0\beta} [\Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta}]^{-1} \Sigma_{\beta 0} \Sigma_{00}^{-1} = \alpha_{\perp} (\alpha'_{\perp} \Omega \alpha_{\perp})^{-1} \alpha'_{\perp}$. We note that $S_{10} \alpha_{\perp} = (S_{10} - S_{11} \beta \alpha') \alpha_{\perp}$.

The expression above is therefore equal to

$$\rho \int_0^1 \mathbf{H} \mathbf{H}' dr - \left(\int_0^1 \mathbf{H}(r) d\mathbf{K}'_{\psi, \Sigma}(r) \right) \alpha_{\perp} (\alpha'_{\perp} \Omega \alpha_{\perp})^{-1} \alpha'_{\perp} \int_0^1 d\mathbf{K}_{\psi, \Sigma}(r) \mathbf{H}'(r). \quad (10)$$

Now for $\psi_{\beta} \neq 0$,

$$\begin{aligned} \left| \mathbf{A}'_T S(\lambda) \mathbf{A}_T \right| &\Rightarrow \left| \lambda \begin{bmatrix} 1/3 & 0 & \int_0^1 r \mathbf{H}' dr \\ 0 & \xi'_{\perp} \Sigma_{\beta\beta} \xi_{\perp} & \mathbf{0} \\ \int_0^1 r \mathbf{H} dr & 0 & \int_0^1 \mathbf{H} \mathbf{H}' dr \end{bmatrix} - \begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma'_{\beta 0} \xi_{\perp} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \right| \\ &= \left| \lambda \begin{bmatrix} 1/3 & \int_0^1 r \mathbf{H}' dr \\ \int_0^1 r \mathbf{H} dr & \lambda \int_0^1 \mathbf{H} \mathbf{H}' dr \end{bmatrix} \right| \left| \xi'_{\perp} (\lambda \Sigma_{\beta\beta} - \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma'_{\beta 0}) \xi_{\perp} \right| \end{aligned}$$

The equation $|\mathbf{A}'_T S(\lambda) \mathbf{A}_T| = 0$ has, asymptotically, $p - q + 1$ zero solution and $q - 1$ positive roots given by the solutions to

$$\left| \xi'_{\perp} (\lambda \Sigma_{\beta\beta} - \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma'_{\beta 0}) \xi_{\perp} \right| = 0.$$

We now let $\lambda = \rho T^{-1}$, and

$$\mathbf{N} = \begin{bmatrix} \frac{1}{3} & \mathbf{0}_{1 \times (q-1)} & \int_0^1 r \mathbf{H}'(r) dr \\ \mathbf{0}_{(q-1) \times 1} & \mathbf{0}_{(q-1) \times (q-1)} & \mathbf{0}_{(q-1) \times (p-q)} \\ \int_0^1 r \mathbf{H}(r) dr & \mathbf{0}_{(p-q) \times (q-1)} & \int_0^1 \mathbf{H} \mathbf{H}' dr \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \xi'_{\perp} \Sigma_{\beta 0} \\ \int_0^1 \mathbf{H} d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix} \Sigma_{00}^{-1} \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \xi'_{\perp} \Sigma_{\beta 0} \\ \int_0^1 \mathbf{H} d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix}'$$

then

$$(\beta(\bar{\xi}; \xi_{\perp}) : \bar{\gamma} : \bar{\psi}_{\beta \perp})' S(\lambda) (\beta(\bar{\xi}; \xi_{\perp}) : \bar{\gamma} : \bar{\psi}_{\beta \perp}) = \rho \mathbf{N} - \mathbf{M} + o_p(1)$$

\mathbf{M} is nonsingular with probability one, and so is \mathbf{N} if $q = 1$. In this case, the test statistic has the same distribution as

$$T \text{tr} \{ S_{11}^{-1} S_{10} S_{00}^{-1} S_{01} \} \Rightarrow \text{tr} \{ \mathbf{N}^{-1/2} \mathbf{M} \mathbf{N}^{-1/2} \}.$$

If $q > 1$, \mathbf{N} is singular and of rank $p - q + 1$ almost surely. We let

$$\mathbf{N}^+ = \begin{bmatrix} \frac{1}{3} & \int_0^1 r \mathbf{H}'(r) dr \\ \int_0^1 r \mathbf{H}(r) dr & \int_0^1 \mathbf{H} \mathbf{H}' dr \end{bmatrix} = \int \begin{bmatrix} r \\ \mathbf{H}(r) \end{bmatrix} \begin{bmatrix} r \\ \mathbf{H}(r) \end{bmatrix}' dr$$

$$\mathbf{M}^+ = \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \int_0^1 \mathbf{H} d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix} \Sigma_{00}^{-1} \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \int_0^1 \mathbf{H} d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix}'$$

then $|\rho \mathbf{N} - \mathbf{M}| = 0$ is equivalent to

$$\left| \begin{array}{cc} \rho \mathbf{N}^+ - \mathbf{M}^+ & - \left(\Sigma_{\beta 0} \bar{\xi} + \int_0^1 r d\mathbf{K}_{\Gamma \psi, \mathbf{C}\Sigma} : \int_0^1 d\mathbf{K}_{\psi, \mathbf{C}\Sigma} \mathbf{H}' \right)' \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} \\ - \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \left(\Sigma_{0\beta} \bar{\xi} + \int_0^1 r d\mathbf{K}_{\Gamma \psi, \mathbf{C}\Sigma} : \int_0^1 d\mathbf{K}_{\psi, \mathbf{C}\Sigma} \mathbf{H}' \right) & - \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} \end{array} \right| = 0$$

which, using the formula for the determinant of partitioned matrices, is equal to

$$|-\xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp}| |\rho \mathbf{N}^+ - \mathbf{M}^{++}| = 0$$

where

$$\mathbf{M}^{++} = \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 rd\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \int_0^1 \mathbf{H}d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix} \Omega_{\xi} \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 rd\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \int_0^1 \mathbf{H}d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix}'$$

with

$$\Omega_{\xi} = \Sigma_{00}^{-1} - \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} (\xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp})^{-1} \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1}$$

multiplying

$$\begin{aligned} & \Omega_{\xi} (\Sigma_{0\beta} \xi_{\perp}, \Sigma_{00} (\Sigma_{0\beta} \xi_{\perp})_{\perp}) \\ &= \begin{pmatrix} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} - \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} (\xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp})^{-1} \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp}, \\ \Sigma_{00}^{-1} \Sigma_{00} (\Sigma_{0\beta} \xi_{\perp})_{\perp} - \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} (\xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp})^{-1} \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{00} (\Sigma_{0\beta} \xi_{\perp})_{\perp} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{0}, \\ (\Sigma_{0\beta} \xi_{\perp})_{\perp} \end{pmatrix} \end{aligned}$$

so

$$\Omega_{\xi} = (\Sigma_{0\beta} \xi_{\perp})_{\perp} ((\Sigma_{0\beta} \xi_{\perp})'_{\perp} \Sigma_{00} (\Sigma_{0\beta} \xi_{\perp})_{\perp})^{-1} (\Sigma_{0\beta} \xi_{\perp})'_{\perp}$$

where

$$\Sigma_{0\beta} \xi_{\perp} = \alpha \Sigma_{\beta\beta} \xi_{\perp}$$

which is of dimension $p \times (q-1)$, so there exists a matrix Σ_{ξ} such that the orthogonal complement is $p \times (p-q+1)$

$$(\Sigma_{0\beta} \xi_{\perp})_{\perp} = \alpha_{\perp} \Sigma_{\xi}$$

and hence

$$\Omega_{\xi} = \alpha_{\perp} \Sigma_{\xi} (\Sigma'_{\xi} \alpha'_{\perp} \Sigma_{00} \alpha_{\perp} \Sigma_{\xi})^{-1} \Sigma'_{\xi} \alpha'_{\perp}$$

and recall that $\alpha'_{\perp} \int_0^1 d\mathbf{K}_{\psi, \mathbf{C}\Sigma} \mathbf{H}' = \alpha'_{\perp} \int_0^1 d\mathbf{K}_{\psi, \Sigma} \mathbf{H}'$ so \mathbf{M}^{++} rewrites

$$\mathbf{M}^{++} = \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 rd\mathbf{K}'_{\psi, \Sigma} \\ \int_0^1 \mathbf{H}d\mathbf{K}'_{\psi, \Sigma} \end{bmatrix} \Omega_{\xi} \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 rd\mathbf{K}'_{\psi, \Sigma} \\ \int_0^1 \mathbf{H}d\mathbf{K}'_{\psi, \Sigma} \end{bmatrix}'$$

Both \mathbf{N}^+ and \mathbf{M}^{++} are nonsingular with probability one. This shows that the equation admits $p-q+1$ solutions that decrease to zero at rate T . The distribution of the likelihood ratio test statistic hence converges to the sum of the $p-q$ smallest eigenvalues of $(\mathbf{N}^+)^{-1} \mathbf{M}^{++}$.

Under model M_μ , when $\psi_\beta \neq 0$

$$\begin{aligned} \mathbf{N}_\mu^+ &= \begin{bmatrix} \frac{1}{12} & \int_0^1 (r - \frac{1}{2}) \mathbf{H}'_\mu(r) dr \\ \int_0^1 (r - \frac{1}{2}) \mathbf{H}_\mu(r) dr & \int_0^1 \mathbf{H}_\mu \mathbf{H}'_\mu dr \end{bmatrix} \\ &= \int \begin{bmatrix} r - 1/2 \\ \mathbf{H}_\mu(r) \end{bmatrix} \begin{bmatrix} r - 1/2 \\ \mathbf{H}_\mu(r) \end{bmatrix}' dr \\ \mathbf{M}_\mu^{++} &= \begin{bmatrix} (\bar{\xi}' \Sigma_{\beta 0} + \int_0^1 (r - \frac{1}{2}) d\mathbf{K}'_{0,\Sigma}(r)) \\ \int_0^1 \mathbf{H}_\mu(r) d\mathbf{K}'_{0,\Sigma}(r) \end{bmatrix} \boldsymbol{\Omega}_\xi \begin{bmatrix} (\bar{\xi}' \Sigma_{\beta 0} + \int_0^1 (r - \frac{1}{2}) d\mathbf{K}'_{0,\Sigma}(r)) \\ \int_0^1 \mathbf{H}_\mu(r) d\mathbf{K}'_{0,\Sigma}(r) \end{bmatrix}' \end{aligned}$$

The distribution of the likelihood ratio test statistic hence converges to the sum of the $p - q$ smallest eigenvalues of $(\mathbf{N}_\mu^+)^{-1} \mathbf{M}_\mu^{++}$.

Under M_τ , if $\psi_\beta = 0$, or $\psi_\beta \neq 0$

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

which has q positive roots given by $|\lambda \Sigma_{\beta\beta} - \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta}|$. Now consider

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

where the first factor has no roots as $\lambda \rightarrow 0$ since $\beta' S(\lambda) \beta \rightarrow -\Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta}$, and,

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

hence

$$\begin{aligned} & (\bar{\gamma} : \bar{\psi}_{\beta\perp})' \left\{ S(\lambda) - S(\lambda) \beta [\beta' S(\lambda) \beta]^{-1} \beta' S(\lambda) \right\} (\bar{\gamma} : \bar{\psi}_{\beta\perp}) \\ & \Rightarrow \rho \int_0^1 \mathbf{H}_\tau \mathbf{H}'_\tau dr - \left(\int_0^1 \mathbf{H}_\tau(r) d\mathbf{K}'_{0,\Sigma}(r) \right) \alpha_\perp (\alpha'_\perp \boldsymbol{\Omega} \alpha_\perp)^{-1} \alpha'_\perp \int_0^1 d\mathbf{K}_{0,\Sigma}(r) \mathbf{H}'_\tau(r) \end{aligned}$$

Under M_μ^* , with $\mathbf{A}_T^* = \left(\left((T^{-1/2} \bar{\xi} : \xi_\perp) \right)' \beta' : \beta_2 (\bar{\xi} : \xi_\perp) \right)' : T^{-1/2} (\bar{\gamma}^* : T^{1/2} \bar{\tau}^*)$, first if $\psi_\beta = 0$, then

$$\begin{aligned} & |\mathbf{A}_T^{*'} S^*(\lambda) \mathbf{A}_T^*| \\ & \Rightarrow \left| \lambda \begin{bmatrix} \Sigma_{\beta\beta} + \beta_2 \beta'_2 & \mathbf{0} \\ \mathbf{0} & \int_0^1 \mathbf{H}^* \mathbf{H}^{*'} dr \end{bmatrix} - \begin{bmatrix} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma'_{\beta 0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right| \end{aligned}$$

which has q positive roots given by $|\lambda (\Sigma_{\beta\beta} + \beta_2 \beta'_2) - \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta}| = 0$. Now, consider

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

where the first factor has no roots as $\lambda \rightarrow 0$ since $\beta^{*'} S^* (\lambda) \beta^* \rightarrow -\Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta}$ and, letting $\rho = T\lambda$

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

hence

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

Now, for $\psi_{\beta} \neq 0$, $S_{00} \xrightarrow{p} \Sigma_{00}$, $\mathbf{A}_T^* = \left(\left(T^{-1/2} \bar{\xi} : \xi_{\perp} \right)' (\beta' : \beta_2)' : T^{-1/2} \left(\bar{\gamma}^* : T^{1/2} \bar{\tau}^* \right) \right)$ and

$$\begin{aligned} & |\mathbf{A}_T^{*'} S^* (\lambda) \mathbf{A}_T^*| \\ & \Rightarrow \left| \lambda \begin{bmatrix} \begin{bmatrix} 1/3 & \mathbf{0} \\ \mathbf{0} & \xi'_{\perp} (\Sigma_{\beta\beta} + \beta_2 \beta_2') \xi_{\perp} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \mathbf{0} \\ \mathbf{0} & \int_0^1 \mathbf{H}^* \mathbf{H}^{*'} dr \end{bmatrix} - \begin{bmatrix} \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma'_{\beta 0} \xi_{\perp} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right| \end{aligned}$$

which, as under \mathbf{M} admits only $q-1$ non zero roots. The test statistic has then a distribution which can be derived in a way similar to that under \mathbf{M} . Notice that if the lag order $k=1$ so $\Gamma = \mathbf{I}_p$, then $\beta_2 = (\beta' \alpha)^{-1} \xi$, the contribution of β_2 only appears in the direction ξ .

Now letting $\lambda = \rho T$, and $\mathbf{D}_T^* = \left(\left(\bar{\xi} : \xi_{\perp} \right)' \beta' : \beta_2 (\bar{\xi} : \xi_{\perp}) \right)' : \left(\bar{\gamma}^* : T^{1/2} \bar{\tau}^* \right)$ then

$$|\mathbf{D}_T^{*'} S^* (\lambda) \mathbf{D}_T^*| \Rightarrow 0$$

is equivalent to

$$\begin{aligned} 0 & = \left| \begin{array}{cc} \rho \mathbf{N}^* - \mathbf{M}^* & - \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \int_0^1 \mathbf{H}^* d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} \\ - \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \\ \int_0^1 \mathbf{H}^* d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} \end{bmatrix}' & - \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} \end{array} \right| \\ & \Leftrightarrow 0 = | - \xi'_{\perp} \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta} \xi_{\perp} | |\rho \mathbf{N}^* - \mathbf{M}^{**}| \end{aligned}$$

with

$$\begin{aligned} \mathbf{N}^* & = \begin{bmatrix} \frac{1}{3} & \int_0^1 r \mathbf{H}^{*'} dr \\ \int_0^1 r \mathbf{H}^* dr & \int_0^1 \mathbf{H}^* \mathbf{H}^{*'} dr \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \int_0^1 r \mathbf{H}^* dr & \frac{1}{3} \\ \int_0^1 r \mathbf{H}^* dr & \int_0^1 \mathbf{H} \mathbf{H}' dr & \int_0^1 r \mathbf{H} dr \\ \frac{1}{3} & \int_0^1 r \mathbf{H}^* dr & \frac{1}{3} \end{bmatrix} \\ \mathbf{M}^* & = \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} (r) \\ \int_0^1 \mathbf{H}^* (r) d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} (r) \end{bmatrix} \Sigma_{00}^{-1} \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} (r) \\ \int_0^1 \mathbf{H}^* (r) d\mathbf{K}'_{\psi, \mathbf{C}\Sigma} (r) \end{bmatrix}' \\ \mathbf{M}^{**} & = \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \Sigma} (r) \\ \int_0^1 \mathbf{H}^* (r) d\mathbf{K}'_{\psi, \Sigma} (r) \end{bmatrix} \Omega_{\xi} \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \Sigma} (r) \\ \int_0^1 \mathbf{H}^* (r) d\mathbf{K}'_{\psi, \Sigma} (r) \end{bmatrix}' \end{aligned}$$

The rank of the $(p - q + 2)$ -square matrix \mathbf{N}^* is $p - q + 1$ so 0 is eigenvalue of \mathbf{N}^* . As \mathbf{N}^* is symmetric, there exists a basis \mathbf{v} , such that

$$\mathbf{N}^* = \mathbf{v} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{\Lambda}^* \end{bmatrix} \mathbf{v}^{-1}$$

where $\mathbf{\Lambda}^*$ is the diagonal matrix of eigenvalues of $\int_0^1 \mathbf{H}^* \mathbf{H}^{*'} dr$. Hence the problem $|\rho \mathbf{N}^* - \mathbf{M}^{**}| = 0$ becomes

$$\left| \rho \mathbf{v} \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{\Lambda}^* \end{bmatrix} \mathbf{v}^{-1} - \mathbf{M}^{**} \right| = |\mathbf{v} \mathbf{v}^{-1}| \left| \rho \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{\Lambda}^* \end{bmatrix} - \mathbf{v}^{-1} \mathbf{M}^{**} \mathbf{v} \right| = 0$$

this only admits $p - q + 1$ solutions that coincide with the solutions to

$$\left| \rho \int_0^1 \mathbf{H}^* \mathbf{H}^{*'} dr - \int_0^1 \mathbf{H}^* (r) d\mathbf{K}'_{\psi, \Sigma} (r) \Omega_{\xi} \int_0^1 d\mathbf{K}_{\psi, \Sigma} \mathbf{H}^* (r)' (r) \right|$$

Under \mathbf{M}_{τ}^* , with $\mathbf{A}_{1T}^* = \left(\left((T^{-1/2} \bar{\xi} : \xi_{\perp})' \beta' : T^{-1/2} \xi \right)' : T^{-1/2} \bar{\gamma}^* : T^{-1} \bar{\tau}^* \right)$, first if $\psi_{\beta} = 0$, then

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

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

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

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

$$\begin{aligned} & \left(\bar{\gamma}^* : T^{-1/2} \bar{\tau}^* \right)' S^* (\lambda) \left(\bar{\gamma}^* : T^{-1/2} \bar{\tau}^* \right) \\ & \Rightarrow \rho \int_0^1 \mathbf{H}_{\mu}^* \mathbf{H}_{\mu}^{*'} dr - \int_0^1 \mathbf{H}_{\mu}^* d\mathbf{K}'_{0, \mathbf{C}\Sigma} (r) \mathbf{\Sigma}_{00}^{-1} \int_0^1 d\mathbf{K}_{0, \mathbf{C}\Sigma} (r) \mathbf{H}_{\mu}^{*'} \\ & \text{and } \left(\bar{\gamma}^* : \sqrt{T} \bar{\tau}^* \right)' S^* (\lambda) \beta^* \\ & \Rightarrow \int_0^1 \mathbf{H}_{\mu}^* d\mathbf{K}'_{0, \mathbf{C}\Sigma} (r) \mathbf{\Sigma}_{00}^{-1} \mathbf{\Sigma}_{\beta 0} \end{aligned}$$

hence

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

we show below that we can use a rotation similar to that of lemma 2 in Johansen (1994) so that the test is similar with respect to $\psi_{\beta_{\perp}}$.

Now, for $\psi_\beta \neq 0$,

$$\begin{aligned} & \left(\begin{array}{c} \left(T^{-1/2} \bar{\xi}' \right) \\ \xi'_\perp \\ \left(\bar{\gamma}^{*'} \right) \\ T^{-1/2} \bar{\tau}^{*'} \end{array} \right) \beta^{*'} \right) S(\rho/T) \left(\begin{array}{c} \left(T^{-1/2} \bar{\xi}' \right) \\ \xi'_\perp \\ \left(\bar{\gamma}^{*'} \right) \\ T^{-1/2} \bar{\tau}^{*'} \end{array} \right) \beta^{*'} \right)' \\ \Rightarrow & \rho \begin{bmatrix} 1/12 & 0 & \int_0^1 (r - \frac{1}{2}) \mathbf{H}_\mu^{*'} dr \\ 0 & 0 & 0 \\ \int_0^1 (r - \frac{1}{2}) \mathbf{H}_\mu^* dr & 0 & \int_0^1 \mathbf{H}_\mu^* \mathbf{H}_\mu^{*'} dr \end{bmatrix} \\ & - \begin{bmatrix} \int_0^1 (r - \frac{1}{2}) d\mathbf{K}'_{0, \mathbf{C}\Sigma}(r) \\ \xi'_\perp \Sigma_{\beta 0} \\ \int_0^1 \mathbf{H}_\mu^*(r) d\mathbf{K}'_{0, \mathbf{C}\Sigma}(r) \end{bmatrix} \Sigma_{00}^{-1} \begin{bmatrix} \int_0^1 (r - \frac{1}{2}) d\mathbf{K}'_{0, \mathbf{C}\Sigma}(r) \\ \xi'_\perp \Sigma_{\beta 0} \\ \int_0^1 \mathbf{H}_\mu^*(r) d\mathbf{K}'_{0, \mathbf{C}\Sigma}(r) \end{bmatrix}' \end{aligned}$$

so let

$$\begin{aligned} \mathbf{N}_\mu^* &= \begin{bmatrix} \frac{1}{12} & \int_0^1 (r - \frac{1}{2}) \mathbf{H}_\mu^{*'} dr \\ \int_0^1 (r - \frac{1}{2}) \mathbf{H}_\mu^* dr & \int_0^1 \mathbf{H}_\mu^* \mathbf{H}_\mu^{*'} dr \end{bmatrix} \\ \mathbf{M}_\mu^{**} &= \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 (r - \frac{1}{2}) d\mathbf{K}'_{0, \Sigma}(r) \\ \int_0^1 \mathbf{H}_\mu^*(r) d\mathbf{K}'_{0, \Sigma}(r) \end{bmatrix} \Omega_\xi \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 (r - \frac{1}{2}) d\mathbf{K}'_{0, \Sigma}(r) \\ \int_0^1 \mathbf{H}_\mu^*(r) d\mathbf{K}'_{0, \Sigma}(r) \end{bmatrix}' \end{aligned}$$

As with model \mathbf{M}_μ^* , \mathbf{N}_μ^* is not invertible and the problem $|\rho \mathbf{N}_\mu^* - \mathbf{M}_\mu^{**}|$ coincides with the solution to

$$\left| \rho \int_0^1 \mathbf{H}_\mu^* \mathbf{H}_\mu^{*'} dr - \int_0^1 \mathbf{H}_\mu^*(r) d\mathbf{K}'_{0, \Sigma}(r) \Omega_\xi \int_0^1 d\mathbf{K}_{0, \Sigma}(r) \mathbf{H}_\mu^{*'}(r) \right|$$

Define $\mathbf{V} = (\mathbf{V}'_\gamma : \mathbf{V}'_\psi)'$ with $\mathbf{V}_\gamma = [\bar{\gamma}' \mathbf{C} \Omega \mathbf{C}' \bar{\gamma}]^{-1/2} \bar{\gamma}' \mathbf{C} \Sigma \mathbf{W}$

and $\mathbf{V}_\psi = [\bar{\psi}'_{\beta \perp} \mathbf{C} \Omega \mathbf{C}' \bar{\psi}_{\beta \perp}]^{-1/2} \bar{\psi}'_{\beta \perp} \mathbf{C} \Sigma \mathbf{W}$. \mathbf{V} is a standard Brownian Motion. Also, let

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

The trace statistic therefore admits the following distributions

$$\text{LR} \Rightarrow \text{tr} \left(\int_0^1 d\mathbf{J}_{\psi \alpha_\perp} \mathbf{G}' \left[\int_0^1 \mathbf{G} \mathbf{G}' \right]^{-1} \int_0^1 \mathbf{G} d\mathbf{J}'_{\psi \alpha_\perp} \right) \quad \text{if } \psi_\beta = 0 \quad (11)$$

Now, let $\mathbf{G}_\mu(r) = \mathbf{G}(r) - \int_0^1 \mathbf{G}(u) du$, and $\mathbf{G}_\tau(r) = \mathbf{V}(r) - a_v - b_v r$ where a_v, b_v are coefficients correcting \mathbf{V} for a constant and a trend. Then

$$\begin{aligned} \text{LR}_\mu &\Rightarrow \text{tr} \left(\int_0^1 d\mathbf{J}_0 \mathbf{G}'_\mu \left[\int_0^1 \mathbf{G}_\mu \mathbf{G}'_\mu \right]^{-1} \int_0^1 \mathbf{G}_\mu d\mathbf{J}'_0 \right) \quad \text{if } \psi_\beta = 0 \\ \text{LR}_\tau &\Rightarrow \text{tr} \left(\int_0^1 d\mathbf{J}_0 \mathbf{G}'_\tau \left[\int_0^1 \mathbf{G}_\tau \mathbf{G}'_\tau \right]^{-1} \int_0^1 \mathbf{G}_\tau d\mathbf{J}'_0 \right) \end{aligned}$$

where $\mathbf{J}_0 = \mathbf{V}$ and for the two restricted hypotheses, whether or not $\psi_\beta = 0$ does not matter. Consider first \mathbf{M}_μ^* , let $\mathbf{G}^* = (\mathbf{G}' : r)'$,

$$\text{LR}_\mu^* \Rightarrow \text{tr} \left(\int_0^1 d\mathbf{J}_{(\Gamma\psi)\alpha_\perp} \mathbf{G}^{*'} \left[\int_0^1 \mathbf{G}^* \mathbf{G}^{*'} \right]^{-1} \int_0^1 \mathbf{G}^* d\mathbf{J}'_{(\Gamma\psi)\alpha_\perp} \right).$$

In the case of \mathbf{M}_τ^* , let

$$\mathbf{\Upsilon} = \begin{bmatrix} \mathbf{I}_{p-q-1} & \mathbf{0}_{(p-q-1) \times 1} & 0 \\ \mathbf{0}_{1 \times (p-q-1)} & 1 & 0 \\ 0 & -[\bar{\psi}'_{\beta_\perp} \mathbf{C} \Omega \mathbf{C}' \bar{\psi}_{\beta_\perp}]^{-1/2} & 1 \end{bmatrix}$$

so $\Upsilon' (\mathbf{G}'_\mu(r) : r - 1/2)' = (\mathbf{V}'(r) : r - 1/2)'$ with Υ invertible and define $\mathbf{G}_\mu^* = (\mathbf{V}'(r) : r - 1/2)'$ such that

$$\begin{aligned} \text{LR}_\tau^* &\Rightarrow \text{tr} \left(\int_0^1 d\mathbf{J}_0 \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix}' \left[\int_0^1 \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix} \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix}' \right]^{-1} \int_0^1 \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix} d\mathbf{J}'_0 \right) \\ &= \text{tr} \left(\int_0^1 d\mathbf{J}_0 \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix}' \mathbf{\Upsilon} \left[\int_0^1 \mathbf{\Upsilon}' \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix} \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix}' \mathbf{\Upsilon} \right]^{-1} \int_0^1 \mathbf{\Upsilon}' \begin{pmatrix} \mathbf{G}_\mu \\ r - 1/2 \end{pmatrix} d\mathbf{J}'_0 \right) \end{aligned}$$

i.e. whatever the value of ψ_β

$$\text{LR}_\tau^* \Rightarrow \text{tr} \left(\int_0^1 d\mathbf{J}_0 \mathbf{G}_\mu^{*'} \left[\int_0^1 \mathbf{G}_\mu^* \mathbf{G}_\mu^{*'} \right]^{-1} \int_0^1 \mathbf{G}_\mu^* d\mathbf{J}'_0 \right)$$

Now for $\psi_\beta \neq 0$, first model M:

$$\begin{aligned} &\begin{bmatrix} 1 & 0 & 0 \\ 0 & [\bar{\gamma}' \mathbf{C} \Omega \mathbf{C}' \bar{\gamma}]^{-1/2} & \\ 0 & & [\bar{\psi}'_{\beta_\perp} \mathbf{C} \Omega \mathbf{C}' \bar{\psi}_{\beta_\perp}]^{-1/2} \end{bmatrix} \mathbf{M}^{++} \begin{bmatrix} 1 & 0 & 0 \\ 0 & [\bar{\gamma}' \mathbf{C} \Omega \mathbf{C}' \bar{\gamma}]^{-1/2} & \\ 0 & & [\bar{\psi}'_{\beta_\perp} \mathbf{C} \Omega \mathbf{C}' \bar{\psi}_{\beta_\perp}]^{-1/2} \end{bmatrix}' \\ &= \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \Sigma} \\ \int_0^1 \mathbf{G} d\mathbf{K}'_{\psi, \Sigma} \end{bmatrix} \Omega_\xi \begin{bmatrix} \bar{\xi}' \Sigma_{\beta 0} + \int_0^1 r d\mathbf{K}'_{\psi, \Sigma} \\ \int_0^1 \mathbf{G} d\mathbf{K}'_{\psi, \Sigma} \end{bmatrix}' \end{aligned}$$

where $\mathbf{U}(r) = (\Sigma'_\xi \alpha'_\perp \Sigma_{00} \alpha_\perp \Sigma_\xi)^{-1/2} \Sigma'_\xi \alpha'_\perp \Sigma_{00}^{1/2} \mathbf{W}(r)$ is a standard Brownian motion. Let also $\mathbf{I}_{\alpha'_\perp \psi}(r) = \mathbf{U}(r) + (\Sigma'_\xi \alpha'_\perp \Sigma_{00} \alpha_\perp \Sigma_\xi)^{-1/2} \Sigma'_\xi \alpha'_\perp \psi r$. Denote then by \mathbf{T} (resp. \mathbf{T}_μ) the sum of the $p - q + 1$ smallest eigenvalues of $S(\lambda)$ under \mathbf{M} (resp. \mathbf{M}_μ). It follows that

$$\begin{aligned} \mathbf{T} &\Rightarrow \text{tr} \left\{ \left[\begin{pmatrix} \bar{\xi}' \Sigma_{\beta 0} \\ \mathbf{0} \end{pmatrix} + \int d\mathbf{I}_{\alpha'_\perp \psi} \begin{pmatrix} r \\ \mathbf{G} \end{pmatrix} \right]' \left(\int \begin{pmatrix} r \\ \mathbf{G} \end{pmatrix} \begin{pmatrix} r \\ \mathbf{G} \end{pmatrix}' \right)^{-1} \right. \\ &\quad \left. \times \left(\int \begin{pmatrix} r \\ \mathbf{G} \end{pmatrix} d\mathbf{I}'_{\alpha'_\perp \psi} + \begin{pmatrix} \Sigma'_{\beta 0} \bar{\xi} \\ \mathbf{0} \end{pmatrix} \right) \right\} \\ \mathbf{T}_1 &\Rightarrow \text{tr} \left\{ \left[\begin{pmatrix} \bar{\xi}' \Sigma_{\beta 0} \\ \mathbf{0} \end{pmatrix} + \int d\mathbf{I}_0 \begin{pmatrix} r - \frac{1}{2} \\ \mathbf{G}_\mu \end{pmatrix} \right]' \left(\int \begin{pmatrix} r - \frac{1}{2} \\ \mathbf{G}_\mu \end{pmatrix} \begin{pmatrix} r - \frac{1}{2} \\ \mathbf{G}_\mu \end{pmatrix}' \right)^{-1} \right. \\ &\quad \left. \times \left(\int \begin{pmatrix} r - \frac{1}{2} \\ \mathbf{G}_\mu \end{pmatrix} d\mathbf{I}'_0 + \begin{pmatrix} \Sigma'_{\beta 0} \bar{\xi} \\ \mathbf{0} \end{pmatrix} \right) \right\} \end{aligned}$$

and LR (resp. LR_μ) coincides with the sum of the $p - q$ smallest eigenvalues of the corresponding random matrices entering the traces.

7.4 Simulation setting of Toda (1994)

In the Toda (1994) experiment used by Demetrescu et al., the p -variate \mathbf{x}_t satisfies $\mathbf{x}_0 = 0$ and

$$\mathbf{x}_t = \begin{bmatrix} b\mathbf{I}_q & 0 \\ 0 & \mathbf{I}_{p-q} \end{bmatrix} \mathbf{x}_{t-1} + \mathbf{u}_t, \quad \mathbf{u}_t = \phi \mathbf{u}_{t-1} + \varepsilon_t, \quad \varepsilon_t \stackrel{i.i.d.}{\sim} \mathbf{N} \left(0, \begin{bmatrix} \mathbf{I}_q & \Theta \\ \Theta' & \mathbf{I}_{p-q} \end{bmatrix} \right).$$

We set $p = 5$, $\phi = -.8$, $b = .5$, and

$$\Theta = \begin{cases} (.4 : .2 : .4 : .2) & \text{if } q = 1 \\ \begin{bmatrix} .4 & .2 & .4 \\ .2 & .4 & .2 \end{bmatrix} & \text{if } q = 2 \\ \begin{bmatrix} .4 & .2 & .4 \\ .2 & .4 & .2 \end{bmatrix} & \text{if } q = 3 \\ (.4 : .2 : .4 : .2)' & \text{if } q = 4 \end{cases},$$

In the simulations, we let the deterministic trend span the whole space:

$$\mathbf{y}_t = \psi \frac{t}{\sqrt{T}} (1 : \dots : 1)' + \mathbf{x}_t.$$

$p - q =$	LR			LR_{μ}^*			LR_{μ}			LR_{τ}^*			LR_{τ}			
	1	3	5	1	3	5	1	3	5	1	3	5	1	3	5	
Experiment A																
$\psi =$	0	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	
	0.1	5.82	17.9	45.9	6.47	18.0	46.5	5.42	4.23	4.68	4.86	4.90	5.10	4.75	4.97	5.15
	0.2	7.60	52.7	97.4	9.39	52.8	97.7	5.37	3.20	4.19	4.85	5.01	5.07	4.75	5.02	5.08
	0.5	19.2	99.5	100.0	21.0	99.8	100.0	0.78	2.84	4.15	5.08	4.76	4.80	4.97	4.83	4.94
	1	66.4	99.7	100.0	70.6	99.9	100.0	0.30	2.96	3.77	5.22	4.77	4.81	5.29	4.84	4.90
	1.5	87.8	99.7	100.0	96.0	99.8	100.0	0.17	3.19	3.97	5.37	5.22	5.28	5.40	5.05	5.24
	2	90.2	99.6	100.0	97.1	99.9	100.0	0.18	3.06	3.86	4.90	5.05	5.28	4.85	4.87	5.03
	3	91.7	99.7	100.0	97.3	99.8	100.0	0.21	3.08	3.60	4.85	5.10	6.19	4.82	4.89	5.21
	5	92.6	99.7	100.0	97.6	99.9	100.0	0.19	3.13	3.50	4.94	5.97	10.1	5.08	5.19	5.85
Experiment B																
$\psi =$	0	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	
	0.1	19.1	17.0	5.98	20.8	17.3	6.05	0.67	4.40	5.00	4.46	4.67	4.81	4.50	4.84	4.78
	0.2	64.4	52.1	10.6	70.9	51.9	10.3	0.36	3.26	4.93	4.44	4.56	5.00	4.45	4.83	4.91
	0.5	90.9	99.4	45.8	97.5	99.9	45.6	0.20	3.17	4.32	4.54	4.80	5.10	4.65	5.01	4.97
	1	91.6	99.5	97.3	97.7	99.9	97.5	0.13	3.14	4.04	5.28	4.78	4.98	5.12	4.95	4.91
	1.5	92.0	99.7	100.0	97.7	99.9	100.0	0.23	2.92	4.04	4.92	4.88	5.18	4.82	4.90	5.19
	2	91.8	99.7	100.0	97.5	99.9	100.0	0.16	3.14	3.79	5.25	5.21	4.96	5.03	5.01	4.81
	3	91.5	99.6	100.0	97.6	99.9	100.0	0.11	3.12	3.75	5.72	5.59	4.93	5.14	5.04	4.86
	5	92.0	99.5	100.0	97.7	99.9	100.0	0.14	2.80	3.71	6.76	5.99	5.18	5.25	4.97	4.96
Experiment C																
$\psi =$	0	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	
	0.1	26.3	52.0	61.0	27.3	51.5	62.0	0.40	3.12	4.53	5.21	5.09	5.24	5.21	4.95	5.25
	0.2	78.0	98.3	99.7	86.5	99.0	99.7	0.20	2.73	4.21	4.95	5.08	4.91	5.00	4.97	4.98
	0.5	90.9	99.6	100.0	97.4	99.9	100.0	0.18	3.25	4.21	4.96	4.97	5.23	5.09	5.03	5.08
	1	91.6	99.6	100.0	97.6	99.9	100.0	0.14	2.75	4.26	5.33	5.21	5.39	5.22	4.99	5.28
	1.5	92.0	99.6	100.0	97.5	99.8	100.0	0.17	2.63	4.19	5.14	5.41	5.31	5.06	5.06	5.04
	2	91.7	99.7	100.0	97.6	99.9	100.0	0.15	3.14	4.08	5.33	5.52	6.02	5.21	5.16	5.23
	3	91.8	99.7	100.0	97.5	99.9	100.0	0.19	2.80	4.12	5.97	6.61	7.14	5.46	5.32	5.38
	5	92.4	99.6	100.0	97.8	99.9	100.0	0.14	2.78	3.69	7.78	10.7	13.3	5.57	5.93	6.32

Table 1: Simulated finite sample rejection probabilities at the 5% nominal size of the null of q cointegration relations in the $p = 6$ dimensional VAR(1). The sample size is set to 100 observations and the number of Monte Carlo replications is 20,000. Experiments A , B , and C refer to whether the linear trend lies in the space spanned by β_{\perp} , β or (β, β_{\perp}) .

ψ	$p - q$	$\text{LR}_\mu \cap \text{LR}_\tau^*$		LM	
		$T = 100$	$T = 1000$	$T = 100$	$T = 1000$
Experiment A					
0	1	0.92	0.92	0.76	0.95
	3	0.92	0.92	0.32	0.95
	5	0.93	0.93	0.95	0.95
0.5	1	0.95	0.95	0.76	0.95
	3	0.93	0.93	0.32	0.95
	5	0.93	0.93	0.95	0.95
1	1	0.95	0.94	0.76	0.95
	3	0.94	0.93	0.32	0.95
	5	0.94	0.94	0.95	0.95
2	1	0.95	0.95	0.76	0.95
	3	0.94	0.94	0.32	0.95
	5	0.94	0.93	0.95	0.95
Experiment B					
0	1	0.92	0.92	0.76	0.95
	3	0.92	0.92	0.32	0.95
	5	0.93	0.93	0.95	0.95
0.5	1	0.95	0.95	0.76	0.95
	3	0.93	0.93	0.32	0.95
	5	0.93	0.93	0.95	0.95
1	1	0.95	0.94	0.76	0.95
	3	0.94	0.93	0.32	0.95
	5	0.93	0.93	0.95	0.95
2	1	0.95	0.95	0.76	0.95
	3	0.94	0.94	0.32	0.95
	5	0.93	0.93	0.95	0.95
Experiment C					
0	1	0.92	0.92	0.76	0.95
	3	0.92	0.92	0.30	0.95
	5	0.93	0.93	0.95	0.95
0.5	1	0.95	0.95	0.74	0.95
	3	0.93	0.93	0.31	0.97
	5	0.93	0.93	0.95	0.94
1	1	0.95	0.94	0.78	0.95
	3	0.94	0.93	0.30	0.96
	5	0.94	0.94	0.95	0.95
2	1	0.95	0.95	0.77	0.96
	3	0.94	0.94	0.32	0.96
	5	0.94	0.93	0.95	0.95

Table 2: Proportion of experiments where the rank of cointegration is correctly determined using the usual sequential testing procedure (starting with zero rank). The number of Monte Carlo replications is 20,000 and T denotes the sample size. The critical values for the likelihood ratio tests with unrestricted intercept and restricted trend, LR_μ , LR_τ^* , and the Lagrange-Multiplier test, LM, are computed using the finite sample simulation with $\psi = 0$. The experiments are labeled as in the text, with trend orthogonal to the cointegration space (A) restricted to lie within the cointegration space (B) or spanning both (C).

ψ	$p - q$	$\text{LR}_\mu \cap \text{LR}_\tau^*$		LM	
		$T = 100$	$T = 1000$	$T = 100$	$T = 1000$
Toda (1994) Experiment					
0	1	0.90	0.92	0.38	0.17
	3	0.60	0.92	0.11	0.00
	5	0.93	0.93	0.00	0.00
0.5	1	0.80	0.95	0.38	0.18
	3	0.45	0.94	0.11	0.00
	5	0.93	0.93	0.00	0.00
1	1	0.80	0.94	0.39	0.18
	3	0.46	0.94	0.11	0.00
	5	0.92	0.93	0.00	0.00
2	1	0.80	0.95	0.38	0.18
	3	0.51	0.93	0.11	0.00
	5	0.81	0.93	0.00	0.01
Weak Broken Trend					
0	1	0.92	0.92	0.76	0.95
	3	0.92	0.92	0.31	0.95
	5	0.93	0.93	0.95	0.95
0.5	1	0.79	0.95	0.34	0.96
	3	0.92	0.93	0.26	0.92
	5	0.92	0.93	0.95	0.95
1	1	0.79	0.95	0.32	0.96
	3	0.93	0.93	0.25	0.92
	5	0.93	0.93	0.95	0.95
2	1	0.80	0.95	0.32	0.95
	3	0.93	0.93	0.25	0.92
	5	0.93	0.93	0.95	0.95

Table 3: Proportion of experiments where the rank of cointegration is correctly determined using the usual sequential testing procedure (starting with zero rank). The number of Monte Carlo replications is 20,000 and T denotes the sample size. The critical values for the likelihood ratio tests with unrestricted intercept and restricted trend, LR_μ , LR_τ^* , and the Lagrange-Multiplier test, LM, are computed using the finite sample simulation of pure random walks with $\psi = 0$. The Toda (1994) experiment allows for short-run dynamics (see section 7.5 of the Appendix) and the weak broken trend experiment corresponds to the process with cumulated local Bernoulli increments (\mathbf{z}_t in Section 2).

T	$p - q$	LR_μ			LR_τ^*		
		\mathbf{x}_t	\mathbf{y}_t	\mathbf{z}_t	\mathbf{x}_t	\mathbf{y}_t	\mathbf{z}_t
100	1	9.20	45.22	11.91	12.40	12.40	11.71
	2	20.36	67.27	23.80	25.88	25.90	25.79
	3	35.68	92.14	39.43	43.33	43.35	43.26
	4	55.21	126.05	58.60	65.27	65.32	64.97
	5	79.14	431.80	82.98	91.26	91.31	91.28
1000	1	9.06	423.42	25.28	12.42	12.25	12.19
	2	20.23	575.23	36.22	25.87	25.86	26.16
	3	35.16	662.01	50.75	43.03	42.93	42.93
	4	53.85	739.91	69.61	64.02	64.04	64.46
	5	77.13	3697.70	91.93	89.29	89.10	88.73
∞	1	8.19			12.52		
	2	18.11			25.86		
	3	31.88			42.92		
	4	49.64			63.87		
	5	71.44			88.79		

Table 4: Monte Carlo quantiles at probability 0.95 for the p -vector with cointegrating rank q . \mathbf{x}_t denotes the process without deterministic terms; \mathbf{y}_t with local trend spanning all dimensions (Experiment C) and parameter $\psi = 1$; \mathbf{z}_t denotes a process with Bernoulli trend spanning all dimensions and parameter $\psi = 1$. The number of Monte Carlo replications is 20,000 and $p = 5$. The asymptotic critical values are from MacKinnon et al. (1999).

$H_0 :$	LR_μ^*	LR_μ	LR_τ^*	LR_τ
$q = 0$	31.4 [0.001]**	27.8 [0.000]**	34.7 [0.002]**	28.2 [0.001]**
$q \leq 1$	9.4 [0.044]*	5.9 [0.015]*	7.4 [0.316]	1.4 [0.233]

Table 5: Trace Statistics for the null of q cointegration relations between *Temperature* and *RF*. The data consists of annual observations spanning 1850-2011. Values within square brackets denote the p -value reported by Oxmetrics 6.30.

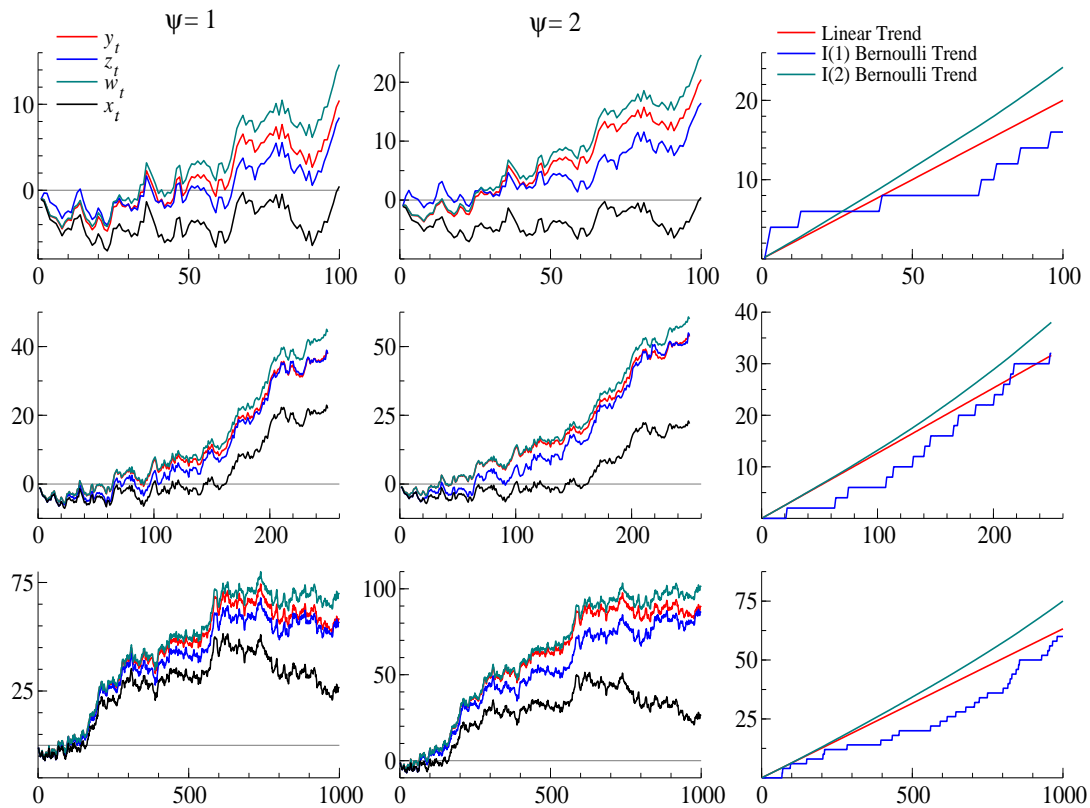


Figure 1: Simulations of processes over $T = 250$ observations. x_t denotes a random walk. y_t is the same random walk with drift ψ/\sqrt{T} . z_t is the same random walk with a random drift $q_t v_t$ where $q_t \stackrel{i.i.d.}{\sim} \text{Bernoulli}\left(\frac{\psi}{\sqrt{T}}\right)$ and $v_t \stackrel{i.i.d.}{\sim} \mathbf{N}(0,1)$. In the third column, ‘Linear Trend’ denotes the difference $y_t - x_t$, and ‘Bernoulli Trend’ the difference $z_t - x_t$.

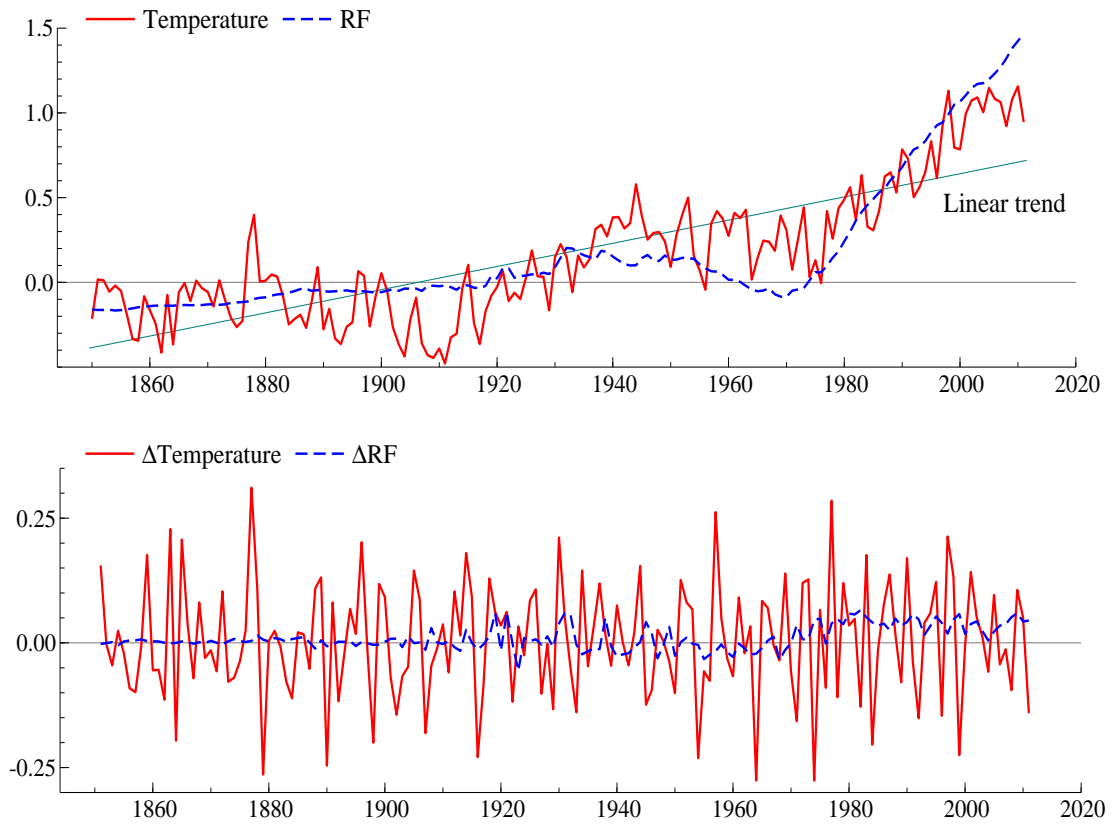


Figure 2: The figure reports the data on global surface temperature (*Temperature*) and radiative forcing of human origin (*RF*). The observations are annual. Top panel: a linear trend has been fitted to either variable and both variables have been adjusted so their means and ranges match. The fitted linear trends therefore match. Bottom panel: the data have not been adjusted.

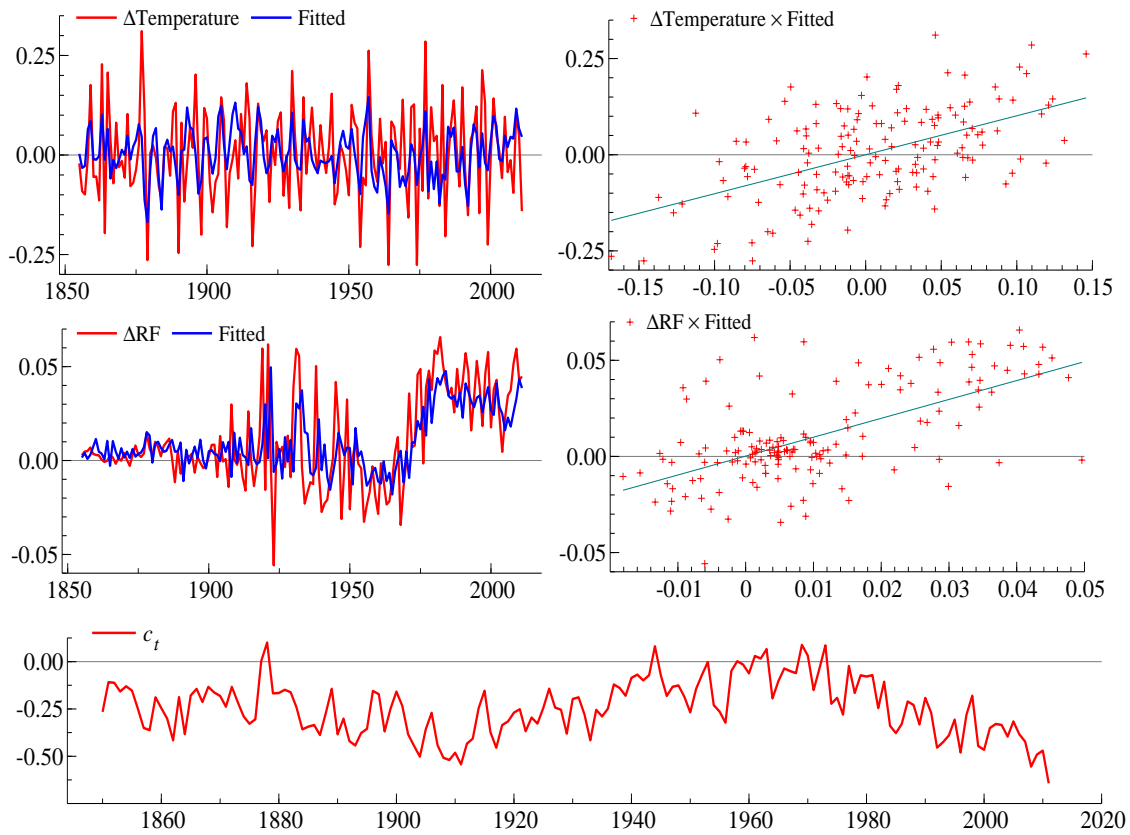


Figure 3: The figure reports output from estimation of a cointegrated VAR for *Temperature* and *RF*. Top panels report the time series and cross plot of observed and fitted series. The bottom panel records the time series of the estimated cointegrating vector.