Fw: Journal of the American Statistical Association - Decision on Manuscript

PARUOLO Paolo <Paolo.PARUOLO@ec.europa.eu> Tue, Jun 16, 2026 at 7:37 AM
To: Massimo Franchi <mas.franchi@gmail.com>, Iliyan Georgiev <i.georgiev@unibo.it>

:-)

Sent from Nine Work

From: Journal of the American Statistical Association <onbehalfof@manuscriptcentral.com>
Sent: Tuesday, June 16, 2026 07:12
To: PARUOLO Paolo (JRC-ISPRA)
Subject: Journal of the American Statistical Association - Decision on Manuscript

16-Jun-2026

Dear Dr. Paruolo:

Ref: Canonical correlation analysis of stochastic trends via functional approximation

I am happy to report that your paper, "Canonical correlation analysis of stochastic trends via functional approximation", has been accepted to appear in the Journal of the American Statistical Association. Please accept my warmest congratulations.

If you have not already uploaded the following files, please send them to eric@amstat.org:

• One zip file containing the TeX or Word file of your final, unblinded paper, including all figure files as separate files, and any additional files needed to compile your article. For example, .bbl and .bib files if you used BiBTeX (and we hope you did!), and any special macros. These files should be uploaded using the Source Files designation.

• One zip file containing all files to be included as online supplements to your paper (appendices, datasets, code, etc). There is no need to include the TeX file(s) for online PDF files. Please use the Supplementary Materials for Review designation for this zip file. A README file containing a list of your online supplements is very helpful.

• PLEASE NOTE: If you use LaTeX cross-referencing between your main paper file and online supplement files, please be sure to use the {xr} LaTeX package and let us know of its use. Otherwise, please do not use cross-referencing between the main paper file and the supplement.

• In addition to funding acknowledgments, please be sure to include a disclosure statement. This is to acknowledge any financial or non-financial interest that has arisen from the direct applications of your research. If there are no relevant competing interests to declare please state this within the article, for example: The authors report there are no competing interests to declare. Further guidance on what is a conflict of interest and how to disclose it can be found here: https://urldefense.com/v3/__https://authorservices.taylorandfrancis.com/editorial-policies/competing-interest/__;!!DOxrgLBm!GsVAYBaxt5dw1r385iAFJNsoWHmUUjSup84xZgNgu3JdHdXOiv0jfkU7hGHXPQ06vDsTTePVpBry6XGib_jmcXn10Qbg0wdJ-lMx$

Please ensure that the Conflict of Interest statement is included at the end of your manuscript, after any acknowledgements and prior to the references.

• PLEASE NOTE: Your paper will be posted online shortly after it is placed into production. All funding acknowledgments should be noted and any minor corrections (typos, etc.) should be made at this time, as we are unable to correct files once posted on the Latest Articles site. Of course, all corrections can be made at proofs stage, prior to final publication.

IMPORTANT: Please also remember to keep the length of your paper 35 pages or less (which includes tables, figures, references). (You may choose to remove material from both copies and place the material into a separate supplement file (blinded, for review) in order to meet these requirements.)

If you have any questions, please feel free to contact Eric Sampson, ASA Journals Manager, eric@amstat.org.

Your paper in its current form will now be forwarded to the publisher for copy editing and typesetting. The reviewer comments are included at the bottom of this letter, along with those of the editor who coordinated the review of your paper.

You will receive proofs for checking.

The publisher also requests that proofs are checked through the publisher's tracking system and returned within 48 hours of receipt.

Thank you for your contribution to Journal of the American Statistical Association and we look forward to receiving further submissions from you.

Sincerely,
Dylan Small
Co-editor, Journal of the American Statistical Association, Theory and Methods Section

Reviewers' Comments to Author:

Reviewer: 3

Comments to the Author
This is the reproducibility assessment from the JASA Associate Editor for Reproducibility. The authors have done a commendable job at providing their code for the reproducibility review. I don't have other comments and recommend this pass the reproducibility review.

Canonical correlation analysis of stochastic trends via functional approximation

M. Franchi, I. Georgiev, P. Paruolo[1]

March 23, 2026

Abstract

This paper proposes a novel approach for semiparametric inference on the number s of common trends and their loading matrix ψ in $I(1)/I(0)$ systems. It combines functional approximation of limits of random walks and canonical correlations analysis, performed between the p observed time series of length T and the first K discretized elements of an L^2 basis. Tests and selection criteria on s, and estimators and tests on ψ are proposed; their properties are discussed as T and K diverge sequentially for fixed p and s. It is found that tests on s are asymptotically pivotal, selection criteria of s are consistent, estimators of ψ are T-consistent, mixed-Gaussian and efficient, so that Wald tests on ψ are asymptotically Normal or χ^2. The paper also discusses asymptotically pivotal misspecification tests for checking model assumptions. The approach can be coherently applied to subsets or aggregations of variables in a given panel. Monte Carlo simulations show that these tools have reasonable performance for $T \geq 10p$ and $p \leq 300$. An empirical analysis of 20 exchange rates illustrates the methods.

Keywords: Unit roots; cointegration; $I(1)$; large panels; semiparametric inference; canonical correlation analysis
JEL Classification: C32, C33

[1]The views expressed in the paper are those of the authors and do not necessarily reflect the opinion of the institutions of affiliation. The authors thankfully acknowledge useful comments from: the Associate Editor, two anonymous Referees, Søren Johansen, Ye Lu, Anders Rahbek, and participants to paper presentations at University of Copenhagen, University of Sydney, Workshop in memory of Stefano Fachin at Sapienza University of Rome, University of Oxford, University of Aarhus, Workshop on Methodological and Computational Issues in Large-Scale Time Series Models for Economics and Finance at University of Messina, Econometric Workshop in memory of Carlo Giannini at University of Pavia, Villa Mondragone Time Series Symposium in honor of Marco Lippi at Tor Vergata University of Rome, Eleventh Italian Congress of Econometrics and Empirical Economics at University of Palermo, Kobe University, Hitotsubashi University. M. Franchi gratefully acknowledges partial financial support from the project PRIN 20223725WE "Methodological and computational issues in large-scale time series models for economics and finance".

M. Franchi, Sapienza University of Rome, ORCID: 0000-0002-3745-2233, Email: massimo.franchi@uniroma1.it, Address: P.le A. Moro 5, 00185 Rome, Italy.

I. Georgiev, University of Bologna, ORCID: 0000-0003-3392-806X, Email: i.georgiev@unibo.it, Address: Piazza Scaravilli 2, 40126 Bologna, Italy.

P. Paruolo, European Commission, Joint Research Centre (JRC), ORCID: 0000-0002-3982-4889, Email: paolo.paruolo@ec.Europa.eu, Address: Via Enrico Fermi 2749, 21027 Ispra (VA), Italy.

1 Introduction

The analysis of multiple time series with nonstationarity and comovements has been steadily developing since the introduction of the notion of cointegration in Engle and Granger (1987), and the seminal contributions of Stock and Watson (1988), Johansen (1988, 1991), Ahn and Reinsel (1990), Phillips and Hansen (1990); see Watson (1994) for an early survey. These contributions discussed asymptotic results for diverging sample size T and fixed cross-sectional dimension p. Associated applications considered a cross-sectional dimension p in the low single digits.

A central parameter in this context is the $p \times r$ cointegration matrix β and its rank r, the cointegration rank; inference on these quantities can be approached either parametrically, as in the context of Vector Autoregressive (VAR) processes, see Johansen (1991, 1996), or semiparametrically, as in Phillips and Hansen (1990), and Bierens (1997).

Data with p in the double digits have recently become ubiquitous. The label 'moderate dimensional' has been used when such data are analyzed assuming a fixed p in the asymptoptics for diverging T, see Chen and Schienle (2024); the present paper falls into this framework. Contributions for the moderate dimensional case include modifications of parametric procedures, such as Bartlett corrections, see Johansen (2000, 2002a,b), bootstrap implementations, see Swensen (2006), Cavaliere et al. (2012) and Cavaliere et al. (2015), and Lasso-type procedures, see Chen and Schienle (2024).

One alternative stream of literature, called 'high dimensional', considers p and T diverging proportionally; leading contributions are Onatski and Wang (2018, 2019), Bykhovskaya and Gorin (2022, 2024), Liang and Schienle (2019). Finally, the cross sectional dimension p can also be infinite from the start, as in functional time series, see Petersen et al. (2022), Nielsen et al. (2022). Combinations of the moderate and high dimensional cases have been studied in Zhang et al. (2019), Barigozzi et al. (2022), Franchi et al. (2023).

The present paper provides novel inferential tools for semiparametric inference on $I(1)/I(0)$ p-dimensional vectors X_t that admit the Common Trends (CT) representation

$$X_t = \gamma + \psi\kappa' \sum_{i=1}^{t} \varepsilon_i + C_1(L)\varepsilon_t, \qquad t = 1, 2, \ldots,$$

where γ is a vector of initial values, $\kappa' \sum_{i=1}^{t} \varepsilon_i$ are $0 \leq s \leq p$ stochastic trends, ψ is a $p \times s$ full column rank loading matrix, L is the lag operator and $C_1(L)\varepsilon_t$ is an $I(0)$ linear process.

The current analysis can be coherently applied to subsets or aggregations of variables. This fact rests on the property, called *dimensional coherence*, that linear combinations of the observables X_t also admit a common trends representation with a number of stochastic trends that is at most equal to the one of the original system.

The proposed approach is based on the empirical canonical correlations between the p observed variables $\{X_t\}_{t=1}^T$ and K deterministic variables $\{d_t\}_{t=1}^T$, constructed as the first K elements of an orthonormal $L^2[0, 1]$ basis discretized over the equispaced grid $1/T, 2/T, \ldots, 1$. In the asymptotic analysis, the cross-sectional dimension p is kept fixed while T and K diverge sequentially, K after T, denoted by $(T, K)_{seq} \to \infty$.

The techniques in the present paper build on the results in Phillips (1998, 2005, 2014), Sun (2011) where functional approximations with an increasing number K of $L^2[0, 1]$ basis elements are used for understanding regressions between $I(1)$ and deterministic variables, for the heteroskedasticity and autocorrelation-consistent estimation of long-run variance (LRV) matrices and for efficient instrumental-variables (IV) estimation of cointegrating vectors.

Related approaches based on a fixed number K of cosine functions are presented in Bierens (1997), Hwang and Sun (2018) and in Müller and Watson (2008, 2017, 2018), who use an IV approach to estimate the number of stochastic trends, test hypotheses on cointegrating relations, or analyze long-run variability and covariability of filtered time series.

The present study combines canonical correlation analysis (CCA) and the functional approximation of limits of random walks. It is shown that the s largest squared canonical correlations (SCC) tend to one while the $p - s$ smallest SCC tend to zero for any cross-sectional dimension p, for any number of stochastic trends $0 \leq s \leq p$ and for any choice of $L^2[0,1]$ basis, all fixed, as $(T, K)_{seq} \to \infty$. This implies that a criterion based on the maximal gap between SCC selects the correct number of common trends consistently. Similar results apply to criteria designed following Bierens (1997) and Ahn and Horenstein (2013).

The cause of these results is that the stochastic trends in X_t converge to Brownian motions as T diverges while the deterministic variables d_t used in the CCA converge to the first K elements of an $L^2[0,1]$ basis; Brownian motions have an L^2-representation (unlike $I(0)$ components), and in the K-limit they are completely recovered. This implies that the largest SCC are associated with the nonstationary directions in X_t and the smallest ones with the stationary linear combinations.

This is the opposite association with respect to the one in Onatski and Wang (2018, 2019) and in Bykhovskaya and Gorin (2022, 2024), which involve a CCA between differences ΔX_t and lagged levels X_{t-1} (possibly correcting for lagged differences and deterministic terms) derived as the Gaussian Quasi Maximum Likelihood (QML) estimator in cointegrated VARs, see Johansen (1988, 1991) and Ahn and Reinsel (1990).

The present approach also differs from the one in Franchi et al. (2023), who discuss a CCA between levels X_t and their cumulations $\sum_{i=1}^{t} X_i$, and consider an asymptotic regime in which T and p diverge sequentially. The advantage of the proposal here is that it leads to efficient asymptotic inference on the loading matrix ψ, as well as to tests on s and misspecification tests, as explained next.

The canonical variates in the CCA are used to define estimators of the space spanned by the loading matrix ψ and of the cointegrating space spanned by the cointegrating matrix β; these estimators are shown to be T-consistent for any cross-sectional dimension p, for any number of stochastic trends $0 \leq s \leq p$ and for any basis of $L^2[0,1]$, all fixed. When the CCA is iterated once in a specific way, for $(T,K)_{seq} \to \infty$ the associated estimators of ψ and β are asymptotically mixed-Gaussian and efficient, as in Johansen (1996) and in Phillips (2005, 2014). The mixed-Gaussian asymptotic distribution is then used to construct asymptotically Normal and χ^2 Wald test statistics for hypotheses on ψ and β. Identification of ψ and β is discussed and an inferential criterion for the validity of the chosen identification condition is proposed.

Further asymptotic results are derived for the specific basis in the Karhunen-Loève (KL) representation of a Brownian Motion, see Loève (1978). Using the KL basis, canonical correlations are used (i) to estimate s by a top-down sequences of asymptotically pivotal tests for $H_0: s = j$ versus $H_1: s < j$, for $j = p, p-1, \ldots, 1$ and (ii) to perform misspecification tests for model assumptions. The sequential tests are shown to select the correct number of stochastic trends with limit probability equal to 1 minus the size of each test in the sequence, similarly to what happens for sequences of likelihood ratio (LR) tests in the VAR case, see Johansen (1996).

The remainder of the paper is organized as follows. Section 2 discusses assumptions on the Data Generating Process (DGP) and Section 3 defines the main estimators. Section 4 discusses the asymptotic properties of estimators of the number of stochastic trends and Section 5 those of the loading and the cointegrating matrices. Section 6 contains the Monte

Carlo study, Section 7 presents an empirical application to exchange rates and Section 8 concludes. The Supplementary material contains proofs and additional material.

Concerning notation, for any $a \in \mathbb{R}$, $\lfloor a \rfloor$ and $\lceil a \rceil$ denote the floor and ceiling functions; for any condition c, 1_c is the indicator function of c, taking value 1 if c is true and 0 otherwise; ι_n indicates the $n \times 1$ vectors of ones. The space of right continuous functions $f : [0,1] \mapsto \mathbb{R}^n$ having finite left limits, endowed with the Skorokhod topology, is denoted by $D_n[0,1]$, see Jacod and Shiryaev (2003, Ch. VI) for details, also abbreviated to $D[0,1]$ if $n = 1$. A similar notation is employed for the space of square integrable functions $L^2_n[0,1]$, abbreviated to $L^2[0,1]$ if $n = 1$. Riemann integrals $\int_0^1 a(u)b(u)'\mathrm{d}u$, Ito stochastic integrals $\int_0^1 \mathrm{d}a(u)b(u)'$ and Stratonovich stochastic integrals $\int_0^1 \partial a(u)b(u)'$ are abbreviated respectively as $\int ab'$, $\int \mathrm{d}ab'$ and $\int \partial ab'$.

2 Process, identification and L^2-representation

This section introduces the class of processes studied in the paper and their limiting L^2-representation; it also discusses identification of the parameters of interest.

2.1 Data Generating Process

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a $p \times 1$ linear process generated by

$$\Delta X_t = C(L)\varepsilon_t, \tag{2.1}$$

where L is the lag operator, $\Delta := 1 - L$, $C(z) = \sum_{n=0}^{\infty} C_n z^n$, $z \in \mathbb{C}$, satisfies Assumption 2.1 below, and the innovations $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are independently and identically distributed (i.i.d.) with expectation $\mathbb{E}(\varepsilon_t) = 0$, finite moments of order $2 + \epsilon$, $\epsilon > 0$, and positive definite variance-covariance matrix $\mathbb{V}(\varepsilon_t) = \Omega_\varepsilon$, indicated as ε_t i.i.d. $(0, \Omega_\varepsilon)$, $\Omega_\varepsilon > 0$.

Assumption 2.1 below concerns the first-order expansion of $C(z)$, $C(z) = C + C_1(z)(1-z)$, and the rank $s := \operatorname{rank} C(1) = \operatorname{rank} C$. For $s > 0$, let $\breve{\psi}$ denote a matrix whose columns form a basis of $\operatorname{col} C$; then $\breve{\kappa}$ is the unique matrix such that $C = \breve{\psi}\breve{\kappa}'$.[2] Moreover, for $s < p$, let $\breve{\beta} := \breve{\psi}_\perp$ and $\breve{\kappa}_\perp$ be matrices whose columns form bases of $(\operatorname{col} C)^\perp$ and $(\operatorname{col} C')^\perp$ respectively; thus, for $0 < s < p$, it holds that $(\operatorname{col} \breve{\psi})^\perp = \operatorname{col}(\breve{\psi}_\perp)$ and likewise for $\breve{\kappa}$. The space $\operatorname{col} C$ is called the attractor space and its orthogonal complement $(\operatorname{col} C)^\perp$ is the cointegrating space of X_t, of dimension $0 \leq r := p - s \leq p$ (the cointegrating rank).

Assumption 2.1 (Assumptions on $C(z)$)**.** *$C(z)$ in (2.1) is $p \times n_\varepsilon$, $n_\varepsilon \geq p$, and*

(i) $C(z) = \sum_{n=0}^{\infty} C_n z^n$ *converges for all* $|z| < 1 + \delta$, $\delta > 0$*;*

(ii) $\operatorname{rank} C(z) < p$ *may only hold at isolated points* z*, where either* $z = 1$ *or* $|z| > 1$*;*

(iii) if $s < p$*, then* $\breve{\psi}_\perp' C_1 \breve{\kappa}_\perp$ *has full row rank* $r := p - s$*, with* $C_1 := C_1(1)$*.*

Assumption 2.1 includes possibly nonstationary VARMA processes (for $n_\varepsilon = p$) in line with e.g. Stock and Watson (1988) and Dynamic Factor Models (for $n_\varepsilon > p$) as in e.g. Bai (2004). Cumulating (2.1), one finds the Common Trends (CT) representation

$$X_t = \gamma + \breve{\psi}\breve{\kappa}' \sum_{i=1}^{t} \varepsilon_i + C_1(L)\varepsilon_t, \qquad t = 1, 2, \ldots, \tag{2.2}$$

which gives a decomposition of X_t into initial conditions $\gamma := X_0 - C_1(L)\varepsilon_0$, the s-dimensional random walk component $\breve{\kappa}' \sum_{i=1}^{t} \varepsilon_i$ with loading matrix $\breve{\psi}$ and the stationary component $C_1(L)\varepsilon_t$. Observe that X_t is $I(0)$ if $s = 0$, i.e. $r = p$; it is $I(1)$ and cointegrated if $0 < r, s < p$ and it is $I(1)$ and non-cointegrated if $s = p$, i.e. $r = 0$.

The aim of the paper is to make inferences about s and $\breve{\psi}$ (and on their complements r and $\breve{\beta}$) from a sample $\{X_t\}_{t=1,\ldots,T}$ observed from (2.2).

[2] For any matrix a, $\operatorname{col} a$ indicates the linear space spanned by its columns; when $(\operatorname{col} a)^\perp \neq \{0\}$, $a_\perp$ is used to indicate any basis of $(\operatorname{col} a)^\perp$.

Remark 2.2 (Initial conditions). Definition 3.4 in Johansen (1996) stipulates that a vector $v \neq 0$ is a cointegrating vector if $v'X_t$ can be made stationary by a suitable choice of its initial distribution; in the present case, this corresponds to setting $\breve{\beta}'\gamma = 0$, i.e. $\breve{\beta}'X_0 = \breve{\beta}'C_1(L)\varepsilon_0$, so that $\gamma := X_0 - C_1(L)\varepsilon_0 = \breve{\psi}\gamma_0$ and hence

$$X_t = \breve{\psi}\left(\gamma_0 + \breve{\kappa}'\sum_{i=1}^{t}\varepsilon_i\right) + C_1(L)\varepsilon_t, \qquad t = 1, 2, \dots . \tag{2.3}$$

Alternatively, as in Elliott (1999), see also Müller and Watson (2008) and Onatski and Wang (2018), if one also observed X_0, one can consider

$$X_t - X_0 = \breve{\psi}\breve{\kappa}'\sum_{i=1}^{t}\varepsilon_i + C_1(L)(\varepsilon_t - \varepsilon_0), \qquad t = 1, 2, \dots . \tag{2.4}$$

The statistical analysis is performed on $x_t := X_t$ or $x_t := X_t - X_0$ according to whether specification (2.3) or (2.4) is chosen.

2.2 Identification

The parameters $\breve{\psi}$, $\breve{\kappa}$ and $\breve{\beta}$ are not identified; in fact $C = \breve{\psi}\breve{\kappa}' = (\breve{\psi}a)(a^{-1}\breve{\kappa}')$ for any nonsingular $a \in \mathbb{R}^{s\times s}$, so that $\breve{\psi}$ and $\breve{\psi}a$ are bases of the attractor space $\operatorname{col} C$, and $\breve{\kappa}$ and $\breve{\kappa}(a^{-1})'$ of $\operatorname{col}(C')$; similarly, $\breve{\beta}$ and $\breve{\beta}a$ with $a \in \mathbb{R}^{r\times r}$ nonsingular are bases of the cointegrating space $(\operatorname{col} C)^{\perp}$.

As done in Johansen (1996) for $\breve{\beta}$, just-identification of the loading matrix $\breve{\psi}$ can be achieved by linear restrictions, employing a known and full column rank matrix $b \in \mathbb{R}^{p\times s}$ such that $b'\breve{\psi} \in \mathbb{R}^{s\times s}$ is nonsingular. Similarly, just-identification of the cointegrating vectors is obtained using a known and full column rank matrix $c \in \mathbb{R}^{p\times r}$ such that $c'\breve{\beta} \in \mathbb{R}^{r\times r}$ is nonsingular, where c is a basis of $(\operatorname{col} b)^{\perp}$. The following Lemma discusses the relationship between the identified parameters obtained in this way; here $\bar{a} := a(a'a)^{-1}$ for a full column rank matrix a and $\breve{\psi}$, ψ (respectively $\breve{\beta}$, β) are bases of $\operatorname{col} C$ (respectively $\operatorname{col}(C)^{\perp}$).

Lemma 2.3 (Duality). *Let $C = \breve{\psi}\breve{\kappa}'$ be a rank factorization of C, let $b \in \mathbb{R}^{p\times s}$ be such that $b'\breve{\psi} \in \mathbb{R}^{s\times s}$ is nonsingular and let $c \in \mathbb{R}^{p\times r}$ be a basis of $(\operatorname{col} b)^{\perp}$; then*

(i) for $s > 0$, $\psi := \breve{\psi}(b'\breve{\psi})^{-1}$ and $\kappa := \breve{\kappa}(\breve{\psi}'b) = C'b$ are identified as the unique solution of the system of equations $C = \psi\kappa'$ and $b'\psi = I_s$;

(ii) for $s < p$, $c'\breve{\beta} \in \mathbb{R}^{r\times r}$ is nonsingular and $\beta := \breve{\beta}(c'\breve{\beta})^{-1}$ is identified as the unique solution of the system of equations $\beta'C = 0$ and $c'\beta = I_r$;

(iii) for $0 < s < p$, the unrestricted coefficients $\psi_* := \bar{c}'\psi$ and $\beta_* := \bar{b}'\beta$ satisfy $\psi_* = -\beta_*'$;

(iv) for any estimator $\widetilde{\psi}$ of ψ normalized as $b'\widetilde{\psi} = b'\psi = I_s$ and any estimator $\widetilde{\beta}$ of β normalized as $c'\widetilde{\beta} = c'\beta = I_r$, one has $\bar{c}'(\widetilde{\psi} - \psi) = \beta'(\widetilde{\psi} - \psi)$ and similarly $\bar{b}'(\widetilde{\beta} - \beta) = \psi'(\widetilde{\beta} - \beta)$.

Point (iii) shows that the unrestricted coefficients in the identified parameters are linearly dependent, $\psi_* = -\beta_*'$; this implies a duality between hypotheses on ψ and β, see (5.5) below.

Importantly, one does not know a priori which restrictions are just-identifying, i.e., which b delivers a nonsingular $b'\breve{\psi}$. Take for instance $b = (I_s, 0)'$: because $\breve{\psi}$ has rank s, there is always at least one permutation of the order of variables for which these restrictions are just-identifying; however, one does not know which permutation should be used. A consistent decision rule on the nonsingularity of $b'\breve{\psi}$ is proposed in Section 4.2 below.

The choice of a specific b implies a unique space $\operatorname{col}(b)^{\perp} = \operatorname{col} c$, but leaves the choice of the specific basis c open, so that c can be chosen to best fit the application. Observe also that the choice of b and then of c can be reversed, by interchanging the role of $(\breve{\psi}, b)$ and $(\breve{\beta}, c)$ in Lemma 2.3. In the remainder of the paper it is assumed that ψ, κ and β in (2.2) are identified by the choice of b and $c = b_{\perp}$ such that $b'\psi = I_s$, $c'\beta = I_r$ and $\kappa = C'b$.

2.3 Dimensional coherence

This section shows that whenever X_t has a CT representation (2.2), also linear combinations of X_t admit a CT representation, with a number of stochastic trends that is at most equal to that in X_t. This property, called dimensional coherence, is formally stated in the next theorem.

Theorem 2.4 (Dimensional coherence)**.** *Let H be a $p \times m$ full column rank matrix; if X_t satisfies (2.1) with $C(z)$ fulfilling Assumption 2.1, then the same holds for $H'X_t$, i.e. $\Delta H'X_t = G(L)\varepsilon_t$, where $G(z) := H'C(z)$ satisfies Assumption 2.1 with $G(1) = H'C$ of rank $q \leq s$.*

Special cases of H are the selection matrix $H = (I_m, 0)'$ of the first m variables, and the group-aggregation matrix $H = (I_q \otimes \iota_n)$, where $p = qn$, $m = q$ and ι_n is an $n \times 1$ vector of ones. One consequence of Theorem 2.4 is that the semiparametric approach described in this paper applies to subsets or aggregations of the original variables X_t; this property is exploited in the empirical application in Section 7.

2.4 Functional central limit theorem and long-run variance

By the functional central limit theorem, see Phillips and Solo (1992), the partial sums $T^{-\frac{1}{2}} \sum_{i=1}^{t} \varepsilon_i$ converge weakly to an n_ε-dimensional Brownian Motion $W_\varepsilon(u)$, $u \in [0,1]$, with variance $\Omega_\varepsilon > 0$; by the CT representation (2.2), one hence has that

$$T^{-\frac{1}{2}} \begin{pmatrix} \bar{\psi}' x_t \\ \beta' \sum_{i=1}^{t} x_i \end{pmatrix} = T^{-\frac{1}{2}} \begin{pmatrix} \kappa' \\ \beta' C_1 \end{pmatrix} \sum_{i=1}^{\lfloor Tu \rfloor} \varepsilon_i + o_p(1) \underset{T\to\infty}{\overset{w}{\to}} \begin{pmatrix} W_1(u) \\ W_2(u) \end{pmatrix} := \begin{pmatrix} \kappa' \\ \beta' C_1 \end{pmatrix} W_\varepsilon(u), \tag{2.5}$$

where $t = \lfloor Tu \rfloor \in \mathbb{N}$ and the $o_p(1)$-term is infinitesimal uniformly in $u \in [0,1]$. Here $\overset{w}{\to}$ indicates weak convergence of probability measures on $D_p[0,1]$ and $W(u) := (W_1(u)', W_2(u)')'$

is a $p \times 1$ Brownian motion with variance $\Omega := D\Omega_\varepsilon D'$, $D := (\kappa, C_1'\beta)'$ and $\beta = \psi_\perp$.

Remark that long-run variance (LRV) Ω is positive definite because D is nonsingular by Assumption 2.1.(iii); in fact,

$$\operatorname{rank} D = \operatorname{rank}\begin{pmatrix} \kappa' \\ \beta' C_1 \end{pmatrix}(\bar{\kappa}, \kappa_\perp) = \operatorname{rank}\begin{pmatrix} I_s & 0 \\ \beta' C_1 \bar{\kappa} & \beta' C_1 \kappa_\perp \end{pmatrix} = s + \operatorname{rank} \beta' C_1 \kappa_\perp.$$

Partition Ω conformably to $W(u) = (W_1(u)', W_2(u)')'$, namely

$$\Omega := \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} = \begin{pmatrix} \kappa'\Omega_\varepsilon\kappa & \kappa'\Omega_\varepsilon C_1'\beta \\ \beta' C_1 \Omega_\varepsilon \kappa & \beta' C_1 \Omega_\varepsilon C_1'\beta \end{pmatrix}, \tag{2.6}$$

and define the independent Brownian motions

$$W_{2.1}(u) := W_2(u) - \Omega_{21}\Omega_{11}^{-1} W_1(u), \qquad B_1(u) := \Omega_{11}^{-\frac{1}{2}} W_1(u) \tag{2.7}$$

with nonsingular variance matrices $\Omega_{22.1} := \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}$ and I_s respectively.

2.5 L^2-representation

Recall, see e.g. Loève (1978) section 37.5B, that the p-dimensional Brownian motion $W(u)$ admits the representation

$$W(u) \simeq \sum_{k=1}^{\infty} c_k \phi_k(u), \quad c_k := \int_0^1 W(u)\phi_k(u)\mathrm{d}u = \nu_k \xi_k, \quad \nu_k \in \mathbb{R}_+, \qquad \xi_k \sim N(0, \Omega), \tag{2.8}$$

where $\{\phi_k(u)\}_{k=1}^{\infty}$, $\int_0^1 \phi_j(u)\phi_k(u)\mathrm{d}u = 1_{j=k}$, is an orthonormal basis of $L^2[0,1]$ and $\simeq$ indicates that the series in (2.8) is a.s. convergent in the L^2 sense to the l.h.s..

In the special case where $(\nu_k^2, \phi_k(u))$, $k = 1, 2, \dots$, is an eigenvalue-eigenvector pair of the covariance kernel of the standard Brownian motion, i.e. $\nu_k^2 \phi_k(u) = \int_0^1 \min(u, v)\phi_k(v)\mathrm{d}v$, (2.8) is the Karhunen-Loève (KL) representation of $W(u)$, for which one has

$$\nu_k = \frac{1}{(k - \frac{1}{2})\pi}, \qquad \phi_k(u) = \sqrt{2}\sin(u/\nu_k), \qquad \xi_k \sim \text{i.i.d. } N(0, \Omega). \tag{2.9}$$

In this case the series in (2.8) is a.s. uniformly convergent in u and $\simeq$ is replaced by $=$. In the following, a basis $\{\phi_k(u)\}_{k=1}^{\infty}$ is indicated as the KL basis when $\phi_k(u)$ is chosen as in (2.9). Expansion (2.8) underlies the large-K asymptotic theory in this paper; see Appendix A for details and references.

3 Definition of the estimators

This section presents the proposed estimators of the number s of stochastic trends and the cointegrating rank $r = p - s$, as well as of ψ (the identified basis of the attractor space) and β (the identified basis of the cointegrating space).

Let $\varphi_K(u) := (\phi_1(u), \ldots, \phi_K(u))'$ be a $K \times 1$ vector function of $u \in [0, 1]$, where $\phi_1(u), \ldots, \phi_K(u)$ are the first K elements of some fixed orthonormal càdlàg basis of $L^2[0, 1]$, see (2.8). Let d_t be the $K \times 1$ vector constructed by evaluating $\varphi_K(\cdot)$ at the discrete sample points $1/T, \ldots, (T-1)/T, 1$, i.e.

$$d_t := \varphi_K(t/T) := (\phi_1(t/T), \ldots, \phi_K(t/T))', \qquad K \geq p, \qquad t = 1, \ldots, T. \tag{3.1}$$

For a generic p-dimensional variable f_t observed for $t = 1, \ldots, T$, the sample canonical correlation analysis of f_t and d_t in (3.1), denoted as $\textsc{cca}(f_t, d_t)$, consists in solving the following generalized eigenvalue problem, see e.g. Johansen (1996) and references therein,

$$\textsc{cca}(f_t, d_t): \qquad |\lambda M_{ff} - M_{fd} M_{dd}^{-1} M_{df}| = 0, \qquad M_{ij} := T^{-1} \sum_{t=1}^{T} i_t j_t'. \tag{3.2}$$

This delivers eigenvalues $1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, collected in $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$, and corresponding eigenvectors $V = (v_1, \ldots, v_p)$, organized as

$$\Lambda_1 := \operatorname{diag}(\lambda_1, \ldots, \lambda_s), \quad \Lambda_0 := \operatorname{diag}(\lambda_{s+1}, \ldots, \lambda_p), \quad V_1 := (v_1, \ldots, v_s), \quad V_0 := (v_{s+1}, \ldots, v_p), \tag{3.3}$$

with the convention that $(\Lambda_1, V_1) = (\Lambda, V)$ when $s = p$ and $(\Lambda_0, V_0) = (\Lambda, V)$ when $s = 0$.

Definition 3.1 (Max-gap estimators of s and r). *Define the maximal gap (max-gap) estimator of s as*

$$\widehat{s} := \underset{i\in\{0,\ldots,p\}}{\operatorname{argmax}}(\lambda_i - \lambda_{i+1}), \qquad \lambda_0 := 1, \qquad \lambda_{p+1} := 0, \tag{3.4}$$

where $1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ are the eigenvalues of $\text{CCA}(x_t, d_t)$, x_t is defined after (2.4) in Remark 2.2; further define the estimator of r as $\widehat{r} := p - \widehat{s}$.

The choice $\lambda_0 := 1$ and $\lambda_{p+1} := 0$ permits the calculation of $\lambda_i - \lambda_{i+1}$ also for $i = 0$ and $i = p$, and it is motivated by the results in Theorem 4.1 below.

First stage estimators of ψ and β are defined in terms of the eigenvectors in $\text{CCA}(x_t, d_t)$, and they are indicated as $\widehat{\psi}^{(1)}$ and $\widehat{\beta}^{(1)}$. The Iterated Canonical Correlation (ICC) estimators are denoted as $\widehat{\psi}$ and $\widehat{\beta}$, and they are defined in terms of the eigenvectors in $\text{CCA}(e_t, d_t)$, where e_t denotes the residual of x_t regressed on the fitted values from the regression of $\widehat{\psi}^{(1)\prime}\Delta x_t$ on d_t. The motivation for the ICC estimation is to obtain mixed-Gaussian limit distributions, see Theorem 5.2 below.

Definition 3.2 (Estimators of ψ and β). *Let $0 < s < p$.*[3] *Let $b \in \mathbb{R}^{p\times s}$ and $c = b_\perp \in \mathbb{R}^{p\times r}$ be a pair of matrices for the just-identification of $\breve{\psi}$ and $\breve{\beta}$, see Section 2.2, with $\psi := \breve{\psi}(b'\breve{\psi})^{-1}$, $\beta := \breve{\beta}(c'\breve{\beta})^{-1}$; for any V as in (3.3), the first stage estimators of ψ and β are defined as*

$$\widehat{\psi}^{(1)} := M_{xx}V_1(b'M_{xx}V_1)^{-1}, \qquad \widehat{\beta}^{(1)} := V_0(c'V_0)^{-1},$$

where the eigenvectors V are those of $\text{CCA}(x_t, d_t)$. The ICC estimators are defined as

$$\widehat{\psi} := M_{ee}V_1(b'M_{ee}V_1)^{-1}, \qquad \widehat{\beta} := V_0(c'V_0)^{-1},$$

where the eigenvectors V are those of $\text{CCA}(e_t, d_t)$, with

$$e_t := x_t - M_{xg}M_{gg}^{-1}g_t, \qquad g_t := \bar{a}'M_{\Delta xd}M_{dd}^{-1}d_t, \qquad a = \widehat{\psi}^{(1)}. \tag{3.5}$$

[3] For $s = 0$ (resp. $s = p$), ψ (resp. β) is undefined and $\beta = (c')^{-1}$ (resp. $\psi = (b')^{-1}$) is known.

The asymptotic analysis implies that the inverse matrices in Definition 3.2 are well-defined with probability approaching one. Note moreover that, (i), the estimators in Definition 3.2 are invariant to the normalization of the eigenvectors V and, (ii), the ICC estimators $\widehat{\psi}$ and $\widehat{\beta}$ depend on $\widehat{\psi}^{(1)}$ only through $\text{col}(\widehat{\psi}^{(1)})$.

Remark that the estimators in Definition 3.2 satisfy the identifying constraints of ψ and β, i.e.

$$b'\psi = b'\widehat{\psi}^{(1)} = b'\widehat{\psi} = I_s, \qquad c'\beta = c'\widehat{\beta}^{(1)} = c'\widehat{\beta} = I_r \tag{3.6}$$

and that the duality in Lemma 2.3(iii) equally applies to them, i.e.

$$\psi_* = -\beta_*', \quad \widehat{\psi}_*^{(1)} = -\widehat{\beta}_*^{(1)\prime}, \quad \widehat{\psi}_* = -\widehat{\beta}_*', \tag{3.7}$$

where $a_* := \bar{c}'a$ and $h_* := \bar{b}'h$, for $a = \psi, \widehat{\psi}^{(1)}, \widehat{\psi}$ and $h = \beta, \widehat{\beta}^{(1)}, \widehat{\beta}$.

4 Inference on the number of common trends

This section discusses the limits of the eigenvalues of $\textsc{cca}(x_t, d_t)$ and their implications for estimating the number s of stochastic trends (and the number r of cointegrating relations).

First, results are provided for the max-gap estimator and for alternative argmax estimators based on $\textsc{scc}$. Second, by employing the KL basis, several statistics are introduced (i) to perform misspecification tests for model assumptions and (ii) to define top-down sequences of asymptotically pivotal tests of $H_0 : s = j$ versus $H_1 : s < j$ for $j = p, p-1, \ldots, 1$, which provide alternative estimators of s.

4.1 Consistency

The asymptotic behavior of $\textsc{scc}$ is given in the following theorem.

Theorem 4.1 (Limits of the eigenvalues of $\text{CCA}(x_t, d_t)$)**.** *Let $W_2(u)$, $B_1(u)$ be defined as in (2.5), (2.7), let d_t in (3.1) be constructed using any orthonormal càdlàg basis of $L^2[0,1]$ and let $1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ be the eigenvalues of $\text{CCA}(x_t, d_t)$, see (3.2) and (3.3); then for fixed $K \geq p$ and $T \to \infty$,*

(i) the s largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s$ converge weakly to the ordered eigenvalues of

$$\Upsilon_K := \left(\int B_1 B_1'\right)^{-1} \int B_1 \varphi_K' \int \varphi_K B_1' \overset{a.s.}{>} 0; \tag{4.1}$$

(ii) the r smallest eigenvalues $\lambda_{s+1} \geq \lambda_{s+2} \geq \cdots \geq \lambda_p$ multiplied by T converge weakly to the ordered eigenvalues of

$$\Psi_K := Q_{00}^{-1} \int \mathrm{d}W_2 \varphi_K' Q_{\varphi\varphi} \int \varphi_K \mathrm{d}W_2' \overset{a.s.}{>} 0, \tag{4.2}$$

where $Q_{00} := \mathbb{E}(\beta' x_t x_t' \beta) > 0$ and $Q_{\varphi\varphi} := I_K - \int \varphi_K B_1' \left(\int B_1 \varphi_K' \int \varphi_K B_1'\right)^{-1} \int B_1 \varphi_K'$.

Moreover,

(iii) for $K \to \infty$, $\Upsilon_K \xrightarrow{p} I_s$ and $K^{-1}\Psi_K \xrightarrow{p} Q_{00}^{-1}\Omega_{22}$;

(iv) for $(T, K)_{seq} \to \infty$, all the s largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s$ converge in probability to 1, while all the r smallest eigenvalues $\lambda_{s+1} \geq \lambda_{s+2} \geq \cdots \geq \lambda_p$ converge in probability to 0.

Let $a \overset{p}{\asymp} g(T,K)$ for some function g indicate that $a/g(T,K)$ is bounded and bounded away from 0 in probability as $(T,K)_{seq} \to \infty$; points (i), (ii) and (iii) in Theorem 4.1 imply that

$$\lambda_i \overset{p}{\asymp} \begin{cases} 1 & i = 1, \ldots, s \\ K/T & i = s+1, \ldots, p \end{cases} \tag{4.3}$$

for $(T,K)_{seq} \to \infty$. Because $K/T \to 0$ as $(T,K)_{seq} \to \infty$, this implies $\lambda_i \xrightarrow{p} 1$ for $i = 1, \ldots, s$ while $\lambda_i \xrightarrow{p} 0$ for $i = s+1, \ldots, p$, see Theorem 4.1.(iv).

Corollary 4.2 (Consistency of the max-gap estimator $\widehat{s}$). *Let $\widehat{s}$ be the max-gap estimator in (3.4) computed on* $\textsc{cca}(x_t, d_t)$ *and let the assumptions of Theorem 4.1 hold; then* $\mathbb{P}(\widehat{s} = s) \to 1$ *as* $(T, K)_{seq} \to \infty$ *for any* $0 \leq s \leq p$.

This result holds because Theorem 4.1.(iv) implies that $f_0(i) := \lambda_i - \lambda_{i+1} \xrightarrow{p} 1_{i=s}$ as $(T, K)_{seq} \to \infty$; note that the function $1_{i=s}$ provides the best discrepancy that the $\textsc{scc}$ $\{\lambda_i\}_{i=1}^p$ can give between the $I(1)$ and $I(0)$ components.

In the context of eigenvalue problems different from the ones considered here, Bierens (1997) and Ahn and Horenstein (2013) propose to estimate s selecting the integer i that maximizes some other function $f_j(i)$ of the eigenvalues, possibly using the rates of convergence of λ_i in the definition of $f_j(i)$. These approaches, applied to the $\{\lambda_i\}_{i=1}^p$ eigenvalues from $\textsc{cca}(x_t, d_t)$, suggest the following additional estimators of s, called here *alternative argmax estimators*:

$$\widetilde{s}^{(j)} := \underset{i \in \mathcal{I}_j}{\operatorname{argmax}} \, f_j(i), \qquad j = 1, 2, 3, \tag{4.4}$$

$$f_1(i) := \frac{\prod_{h=1}^{i} \lambda_h}{\prod_{h=i+1}^{p} \left(\frac{T}{K} \lambda_h\right)}, \qquad f_2(i) := \frac{\lambda_i}{\lambda_{i+1}}, \qquad f_3(i) := \frac{\log\left(1 + \lambda_i / \sum_{h=i+1}^{p} \lambda_h\right)}{\log\left(1 + \lambda_{i+1} / \sum_{h=i+2}^{p} \lambda_h\right)},$$

where $\mathcal{I}_1 := \{0, 1, \ldots, p\}$, $\mathcal{I}_2 := \{1, \ldots, p-1\}$, $\mathcal{I}_3 := \{1, \ldots, p-2\}$, empty sums (products) are equal to 0 (1), and $f_1(i)$ uses the rates in (4.3) for $(T, K)_{seq} \to \infty$.

The estimators $\widetilde{s}^{(j)}$, $j = 1, 2, 3$, optimize over the set of integers $\mathcal{I}_j$; note that f_1 is defined for $0 \leq i \leq p$, f_2 for $1 \leq i \leq p - 1$ and f_3 for $1 \leq i \leq p - 2$. One may extend f_2 and f_3 to cover $i = 0$ by defining $\lambda_0 := 1$ as in (3.4); this leads to replacing $\mathcal{I}_j$ in (4.4) with the sets $\mathcal{I}_j^0 := \{0\} \cup \mathcal{I}_j$, $j = 2, 3$. Note however that adding $\lambda_{p+1} := 0$ is not an option, as λ_{p+1} would appear in the denominators of f_2 and f_3. This implies that f_2 (respectively f_3) cannot be defined for $i = p$ (respectively $i = p - 1, p$).

Corollary 4.3 (Consistency of the alternative argmax estimators of s). *Let* $\widetilde{s}^{(j)}$, $j = 1, 2, 3$, *be the alternative argmax estimators of s in* (4.4) *computed on* $\textsc{cca}(x_t, d_t)$ *and let the assumptions of Theorem* 4.1 *hold; then* $\mathbb{P}(\widetilde{s}^{(j)} = s) \to 1$ *for* $(T, K)_{seq} \to \infty$ *for any* $s \in \mathcal{I}_j$, $j = 1, 2, 3$. *The same holds replacing* $\mathcal{I}_j$ *with* $\mathcal{I}_j^0$ *both in the optimization and in the class of DGPs.*

These results are in line with the ones reported in Bierens (1997) and Ahn and Horenstein (2013) in their respective setups.

4.2 Testing identification

Recall that the joint identification of ψ and β via b and c given in Lemma 2.3 rests on the validity of the rank restriction $\text{rank}(b'\breve{\psi}) = s$. Here a decision rule on the correctness of $\text{rank}(b'\breve{\psi}) = s$ is proposed, using the property of dimensional coherence; let

$$H_0 : \text{rank}(b'\breve{\psi}) = s \qquad \text{versus} \qquad H_1 : \text{rank}(b'\breve{\psi}) < s. \tag{4.5}$$

The proposed decision rule consists of estimating the number of stochastic trends of $b'x_t$, and rejecting H_0 if the estimated number of stochastic trends is lower than the one estimated for x_t. The rationale for the criterion is the following: from (2.2) one has $b'x_t = b'\breve{\psi}\breve{\kappa}' \sum_{i=1}^{t} \varepsilon_i + b'C_1(L)\varepsilon_t$, so that the number of stochastic trends in $b'x_t$ coincides with the one in x_t if and only if $b'\breve{\psi}$ is nonsingular. Formally,

$$\text{reject } H_0 \quad \text{if} \quad \widehat{s}(b'x_t) < \widehat{s}(x_t), \tag{4.6}$$

where $\widehat{s}(x_t)$ (respectively $\widehat{s}(b'x_t)$) is the max-gap estimator calculated for x_t (respectively $b'x_t$), see (3.4).

Theorem 4.4 (Consistency of decision rule (4.6))**.** *If H_0 in* (4.5) *holds, one has* $\mathbb{P}(\widehat{s}(b'x_t) < \widehat{s}(x_t)) \to 0$ *and* $\mathbb{P}(\widehat{s}(b'x_t) = \widehat{s}(x_t)) \to 1$ *as* $(T, K)_{seq} \to \infty$. *If H_1 in* (4.5) *holds, one has*

$\mathbb{P}(\widehat{s}(b'x_t) < \widehat{s}(x_t)) \to 1$ and $\mathbb{P}(\widehat{s}(b'x_t) = \widehat{s}(x_t)) \to 0$ as $(T, K)_{seq} \to \infty$. The same statements are true substituting $\widehat{s}$ with $\widetilde{s}^{(j)}$ provided $s \in \mathcal{I}_j$ (respectively $s \in \mathcal{I}_j^0$) both in the DGP and in the optimization.

4.3 Asymptotic distributions

This subsection discusses misspecification tests and test sequences for the estimation of s; these are derived using the specific KL basis in (2.9) in the definition of d_t in (3.1).

Theorem 4.5 (Asymptotic distribution of the eigenvalues with the KL basis)**.** *Let $B_1(u)$ be defined as in (2.7), let d_t in (3.1) be constructed using the KL basis in (2.9) and let $1 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ be the eigenvalues of* CCA(x_t, d_t)*; then for $(T, K)_{seq} \to \infty$,*

(i) the s largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s$ satisfy

$$K\pi^2\tau^{(s)} \xrightarrow{w} \zeta^{(s)}, \tag{4.7}$$

where $\tau^{(s)} := (1 - \lambda_s, 1 - \lambda_{s-1}, \ldots, 1 - \lambda_1)'$, $\zeta^{(s)} := (\zeta_1, \zeta_2, \ldots, \zeta_s)'$, and $\zeta_1 \geq \zeta_2 \geq \cdots \geq \zeta_s$ are the eigenvalues of $\omega := \left(\int B_1 B_1'\right)^{-1}$;

(ii) the r smallest eigenvalues $\lambda_{s+1} \geq \lambda_{s+2} \geq \cdots \geq \lambda_p$ satisfy $K\pi^2(1 - \lambda_i) \xrightarrow{p} \infty$, $i = s + 1, \ldots, p$.

Observe that ω does not depend on nuisance parameters, and that tests based on $K\pi^2\tau^{(s)}$ are asymptotically pivotal. Further, the limit distribution in Theorem 4.5(i) depends on the tail $\sum_{k=K+1}^{\infty} \nu_k^2$, see (2.9), and is therefore not invariant to the choice of an $L_2[0, 1]$ basis.

The rest of this section discusses misspecification tests and sequences of tests on s based on the results of Theorem 4.5.

Remark 4.6 (Confidence stripe for misspecification analysis). The results in Theorem 4.5 can be used to construct confidence sets for misspecification analysis. For example, consider $f(a) = \log(a) - \mathbb{E}(\log \zeta^{(s)})$, $a \in \mathbb{R}^s$, where the log function is applied component-wise, and let $\|\cdot\|_n$ be the n-norm for vectors. Define the confidence misspecification stripe $\mathcal{B} := \{a \in \mathbb{R}^s : \|f(a)\|_\infty < \delta\}$ with δ, such that $\mathbb{P}(\zeta^{(s)} \in \mathcal{B}) = 1 - \eta$ for some pre-assigned confidence level $1-\eta$. The continuous mapping theorem and Theorem 4.5 then imply that $\mathbb{P}(K\pi^2\tau^{(s)} \in \mathcal{B}) \to 1 - \eta$ as $(T, K)_{seq} \to \infty$. The set $\mathcal{B}$ is a stripe around $\mathbb{E}(\log \zeta^{(s)})$ with equal deviations above or below $\mathbb{E}(\log(\zeta_i))$ across $i = 1, \ldots, s$, see Figure 4 below for an illustration.

In practice, one can compute $\mathcal{B}$ and check if $K\pi^2\tau^{(\widehat{s})} \in \mathcal{B}$. If so, the model assumptions appear valid, including the selection of s; otherwise it is necessary to question which assumptions need to be modified. The same approach can be used for continuous transformations other than $\log \zeta_i$ and any well-defined location indicator, such as the median.

Remark 4.7 (Sequential tests on s). The results in Theorem 4.5 can be used to test hypotheses of the following form

$$H_{0j} : s = j \qquad \text{versus} \qquad H_{1j} : s < j, \qquad j = p, p-1, \ldots, 1. \tag{4.8}$$

Define the test statistics $F_{j,n} := \|K\pi^2\tau^{(j)}\|_n$. Under H_{0s}, Theorem 4.5 shows that $F_{s,n} \xrightarrow{w} \|\zeta^{(s)}\|_n$, while under H_{1s}, $F_{s,n} \to \infty$. For $n = 1, \infty$, one finds

$$F_{j,1} = \|K\pi^2\tau^{(j)}\|_1 = K\pi^2 \sum_{i=1}^{j} (1 - \lambda_i), \qquad F_{j,\infty} = \|K\pi^2\tau^{(j)}\|_\infty = K\pi^2(1 - \lambda_j), \tag{4.9}$$

which are similar to the trace and λ-max test statistics in Johansen (1991), and satisfy $F_{s,1} \xrightarrow{w} \operatorname{tr}(\omega) = \sum_{i=1}^{s} \zeta_i^{(s)}$, and $F_{s,\infty} \xrightarrow{w} \max \operatorname{eig}(\omega) = \zeta_1^{(s)}$ as $(T, K)_{seq} \to \infty$.

The $F_{j,n}$ test rejects for large values of the statistics, and critical values $c_{n,\eta}$ are defined as $\Pr(\|\zeta^{(s)}\|_n \le c_{n,\eta}) = 1 - \eta$, see (4.7); they can be estimated by simulation of the limit distribution. The following test sequence is defined using these quantities: choose an asymptotic test size $\eta \in (0,1)$ and test H_{0j} versus H_{1j} in (4.8) for $j = p, p-1, \ldots, 1$ using test statistic $F_{j,n}$ with significance level η until a non-rejection is found. The resulting estimated value for s equals the first non-rejected value of j; this test sequence is indicated as $\{F_{j,n}\}_{j=p,\ldots,1}$.

Corollary 4.8 (Asymptotic properties of test sequences)**.** *Under the assumptions of Theorem 4.5, as $(T, K)_{seq} \to \infty$, the test sequence $\{F_{j,n}\}_{j=p,\ldots,1}$, has asymptotic probability to select the true s equal to $1-\eta$ for $s > 0$ and equal to 1 if $s = 0$, $n = 1, \infty$.*

Results for test sequences are similar to the ones for LR tests in a VAR context for cointegration rank determination, see Johansen (1996).

5 Inference on the attractor space

This section discusses the limit distributions of the first stage and ICC estimators of the identified parameters ψ and β as well as Wald tests of hypotheses on their unrestricted parameters ψ_* and β_*, see (3.7), valid for any choice of basis of $L^2[0,1]$. Because $b'(\widehat{\psi} - \psi) = 0$ and $c'(\widehat{\beta} - \beta) = 0$, the unrestricted parameters are in $\bar{c}'(\widehat{\psi} - \psi) = \widehat{\psi}_* - \psi_*$ and $\bar{b}'(\widehat{\beta} - \beta) = \widehat{\beta}_* - \beta_*$, which are linearly related by the duality in Lemma 2.3 and (3.7); see also Theorem 5.1 below.

Results can be summarized as follows. First stage and ICC estimators are T-consistent and the ICC estimators are asymptotically mixed-Gaussian, unlike the first stage estimators. Thus, asymptotically Normal or χ^2 Wald test statistics for hypotheses on ψ and β can be constructed using the ICC estimators.

Theorem 5.1 (Asymptotic distributions of the first stage and ICC estimators)**.** *Let d_t in*

(3.1) *be constructed using any orthonormal càdlàg basis of* $L^2[0,1]$ *and let the unrestricted coefficients be as in* (3.6) *and* (3.7)*; then for* $(T,K)_{seq}\to\infty$*, the first stage estimators satisfy*

$$T(\widehat{\psi}_*^{(1)}-\psi_*)=-T(\widehat{\beta}_*^{(1)}-\beta_*)'\xrightarrow{w}\int\partial W_2W_1'\left(\int W_1W_1'\right)^{-1}=:Z^{(1)}, \tag{5.1}$$

where $\int\partial W_2W_1'$ *denotes a Stratonovich stochastic integral and* $Z^{(1)}$ *is not mixed-Gaussian in general. On the contrary, for the ICC estimators one has*

$$T(\widehat{\psi}_*-\psi_*)=-T(\widehat{\beta}_*-\beta_*)'\xrightarrow{w}\int \mathrm{d}W_{2.1}W_1'\left(\int W_1W_1'\right)^{-1}=:Z, \tag{5.2}$$

where $\int \mathrm{d}W_{2.1}W_1'$ *denotes an Ito stochastic integral and* Z *is mixed-Gaussian because* $W_{2.1}$ *and* W_1 *are independent.*

Remark 5.2 (Mixed Gaussianity). Note that W_2 and W_1 in (5.1) are dependent because Ω_{21} in (2.6) is in general different from 0; this implies that $Z^{(1)}$ is not mixed-Gaussian in general. On the contrary, $W_{2.1}$ and W_1 in (5.2) are independent and thus Z is mixed-Gaussian; specifically, conditionally on W_1, one has $\mathrm{vec}(Z)\sim N(0,\left(\int W_1W_1'\right)^{-1}\otimes\Omega_{22.1})$, which is a variance mixture of Normals. This shows that the dependence between W_2 and W_1 is cleansed by correcting x_t for the fitted values of $\psi'\Delta x_t$ on d_t, similarly to the heteroskedasticity and autocorrelation-consistent and LRV estimators in Phillips (2005) and the IV estimator in Phillips (2014).

The limiting mixed-Gaussianity of the ICC estimators in Theorem 5.1 can be exploited for testing hypothesis on ψ_* and β_* with Wald-type statistics, using consistent estimators of the LRV $\Omega_{22.1}$, see (2.6). Among others, in line with Hwang and Sun (2018), Phillips (2005, 2014), the estimator $\widehat{\Omega}_{22.1}:=\widehat{\Omega}_{22}-\widehat{\Omega}_{21}\widehat{\Omega}_{11}^{-1}\widehat{\Omega}_{12}$ could be used, with

$$\widehat{\Omega}:=\begin{pmatrix}\widehat{\Omega}_{11} & \widehat{\Omega}_{12}\\ \widehat{\Omega}_{21} & \widehat{\Omega}_{22}\end{pmatrix}:=\frac{T}{K}\begin{pmatrix}\bar{a}'M_{\Delta xd}\\ h'M_{xd}\end{pmatrix}M_{dd}^{-1}\begin{pmatrix}\bar{a}'M_{\Delta xd}\\ h'M_{xd}\end{pmatrix}',\quad a=\widehat{\psi},\ \ h=\widehat{\beta}. \tag{5.3}$$

Theorem 5.3 (Consistency of $\widehat{\Omega}$)**.** *Let d_t in (3.1) be constructed using any orthonormal càdlàg basis of $L^2[0,1]$ and let $\widehat{\Omega}$ be as in (5.3); then $\widehat{\Omega} \xrightarrow{p} \Omega$ as $(T,K)_{seq} \to \infty$, and as a consequence also $\widehat{\Omega}_{22.1} \xrightarrow{p} \Omega_{22.1}$.*

Consider the following hypothesis on the unrestricted coefficients of ψ,

$$H_0 : R' \operatorname{vec}(\psi_*) = h, \tag{5.4}$$

where R is a given $sr \times m$ matrix of full column rank and h is a given $m \times 1$ vector.

Remark 5.4 (Hypotheses on ψ_* and ψ). Recall that $\psi_* := \bar{c}'\psi$ so that (5.4) corresponds to a linear hypotheses on ψ of the form $R_0' \operatorname{vec}(\psi) = h$ with $R_0' := R'(I_s \otimes \bar{c}')$; R_0 is also of full column rank m because R and c are of full column rank. Symmetrically, one can write $-h = R' \operatorname{vec}(\beta_*') = R_1' \operatorname{vec}(\beta')$, where $R_1' := R'(\bar{b}' \otimes I_r)$ is of full row rank m.

Remark 5.5 (Duality between hypothesis on ψ and β). Note that by Lemma 2.3.(iii) and (iv), or (3.7), one has $R' \operatorname{vec}(\psi_*) - h = R' \operatorname{vec}(\beta_*') + h$. This shows that

$$H_0 : R' \operatorname{vec}(\psi_*) = h \qquad \Leftrightarrow \qquad H_0 : R' \operatorname{vec}(\beta_*') = -h. \tag{5.5}$$

The next theorem discusses the Wald test of (5.5).

Theorem 5.6 (Wald tests based on the ICC estimators)**.** *Let the assumptions of Theorem 5.1 and the m linear restrictions on ψ_* and β_* in (5.5) hold. Let also $\widehat{\Omega}_{22.1}$ be a consistent estimator of $\Omega_{22.1}$; then as $(T,K)_{seq} \to \infty$ the Wald test statistic Q satisfies*

$$\begin{aligned} Q &:= T^2(R' \operatorname{vec}(\widehat{\psi}_*) - h)'(R'\widehat{U}R)^{-1}(R' \operatorname{vec}(\widehat{\psi}_*) - h) \\ &= T^2(R' \operatorname{vec}(\widehat{\beta}_*') + h)'(R'\widehat{U}R)^{-1}(R' \operatorname{vec}(\widehat{\beta}_*') + h) \xrightarrow{w} \chi^2_m, \end{aligned} \tag{5.6}$$

where $\widehat{U} := (T^{-1}\bar{a}'M_{xx}\bar{a})^{-1} \otimes \widehat{\Omega}_{22.1} \xrightarrow{w} \left(\int W_1 W_1'\right)^{-1} \otimes \Omega_{22.1}$ with $a = \widehat{\psi}$. Similarly when $m = 1$, the t-ratio statistic satisfies

$$\frac{T(R' \operatorname{vec}(\widehat{\psi}_*) - h)}{\sqrt{R'\widehat{U}R}} = \frac{T(R' \operatorname{vec}(\widehat{\beta}_*') + h)}{\sqrt{R'\widehat{U}R}} \xrightarrow{w} N(0,1). \tag{5.7}$$

The limit distribution in (5.2) is next compared with the distribution in Johansen (1996) of the ML estimator for VAR processes and with the one in Phillips (2014) of the optimal IV estimator in the semiparametric case; it is shown that Z coincides with both of them, thus proving that the ICC estimators are asymptotically efficient.

Theorem 5.7 (ML in Johansen (1996))**.** *Assume that the DGP is a VAR satisfying the $I(1)$ condition in* Johansen (1996)*; then the DGP has representation* (2.1) *with $C(z)$ satisfying Assumption* 2.1 *and the asymptotic distribution Z in* (5.2) *coincides with the asymptotic distribution of the corresponding ML estimator of $-\beta_*'$ in* Johansen (1996, Chapter 13) *and of ψ_* in* Paruolo (1997)*.*

Consider next the semiparametric triangular form in Phillips (2014)

$$y_t = Az_t + u_{0t}, \qquad \Delta z_t = u_{1t}, \tag{5.8}$$

where $u_t := (u_{0t}', u_{1t}')' = Q(L)\varepsilon_t \sim I(0)$ with $Q(1)$ nonsingular, y_t is $r \times 1$ and z_t is $s \times 1$. This corresponds to a partition of x_t into $(y_t', z_t')' := x_t$.

Theorem 5.8 (IV in Phillips (2014))**.** *Assume* (2.1) *holds with $C(z)$ satisfying Assumption* 2.1*; then $Q(z)$ satisfies the summability condition (**L**) in* Phillips (2014) *and the asymptotic distribution Z in* (5.2) *coincides with the asymptotic distribution of the corresponding IV estimator of $A := -\beta_*'$ in* Phillips (2014)*.*

6 Monte Carlo simulations

This section reports a set of Monte Carlo (MC) simulations. The DGP is taken from Onatski and Wang (2018) and Bykhovskaya and Gorin (2022) and data is generated from

$$\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \text{i.i.d.} N(0, I_p), \qquad t = 1, \ldots, T, \qquad \Delta X_0 = X_0 = 0, \tag{6.1}$$

with $\beta = (I_r, 0)'$ and $\alpha = -a\beta$. The DGP depends on the values of (p, T, s, a) and identification is achieved with $b = (0, I_s)'$ and $c = (I_r, 0)'$ except for $s = p$, where α and β are undefined and a plays no role.

When $s < p$, the first $p - s$ coordinate processes of X_t in (6.1) are univariate AR(1)'s with coefficient $1 - a$ and the remaining s are pure random walks. In terms of (2.2), this implies that $\psi = (0, I_s)'$, $\kappa' = (0, I_s)$, $C_1(L) = \mathrm{diag}(0, (1 - (1 - a)L)^{-1} I_r)$, so that the relevant LRV is $\Omega_{22.1} = a^{-2} I_r$. One expects small-sample inference on s and ψ to be more difficult when a is smaller, because one needs to distinguish unit roots from autoregressive roots equal to $1 - a$.

The properties of the max-gap and sequential estimators of s, see (3.4) and (4.9), are simulated with $K = \lceil T^{3/4} \rceil$ for the following values of (p, T, s, a): $p = 10, 20, 50, 100, 300$, $T = (10, 20, 30)p$, $s = \lceil jp/4 \rceil$, $j = 0, 1, 2, 3, 4$, $a = 0.25, 0.5, 0.75, 1$. Tables A.1, A.2 and A.3 in the supplement report full MC results; Figure 1 reports a representative subset for $p = 20$.

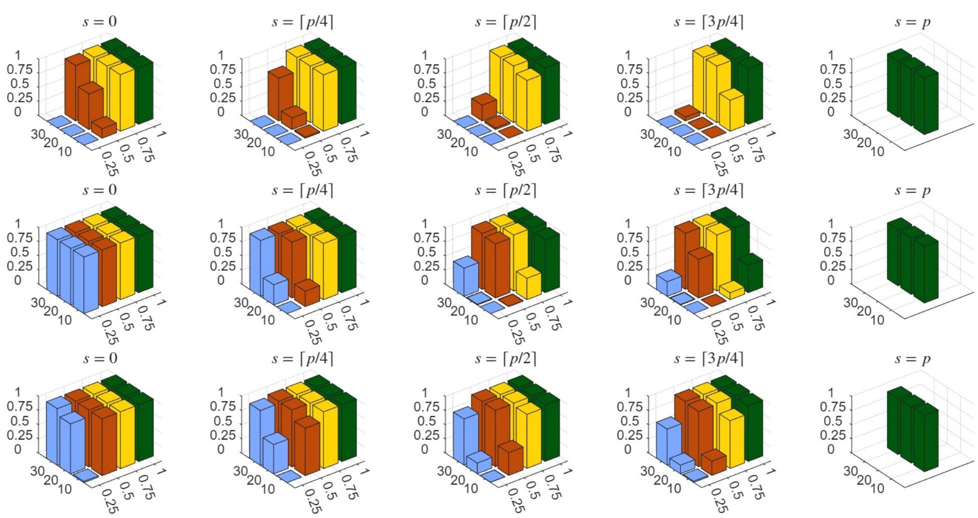

Figure 1: MC relative frequency of correct selection (vertical axis) of $s = \lceil jp/4 \rceil$, $j = 0, 1, 2, 3, 4$, $p = 20$, $T = (10, 20, 30)p$, $a = 0.25, 0.5, 0.75, 1$, and $N = 10^4$ replications. Top: max-gap estimator. Middle: test sequence $\{F_{j,1}\}_{j=p,\ldots,1}$. Bottom: hybrid with $F_{p,1}$. Significance level: 5%.

The top row in Figure 1 reports results for the max-gap estimator. In absence of stationary AR components, i.e. when $s = p$ or $a = 1$, the max-gap estimator selects the correct number of common trends s with frequency 1 from $T = 10p$ onward. When $a = 0.75$ and $a = 0.5$ the stationary AR components are increasingly autocorrelated, and the correct selection of s with frequency 1 is observed for higher values of T. Finally, when $a = 0.25$ and there are strong stationary AR components, the max-gap estimator never select the correct s up to $T = 30p$.

Overall, for fixed (p, T, s) the performance of the max-gap worsens for lower values of a because convergence to 0 of the smaller SCC λ_j, $s < j \leq p$, appears to be slower for lower values of a, see Fig. A.1.

The middle row in Fig. 1 shows that when the max-gap estimator does not perform well, the estimator based on the test sequence $\{F_{j,n}\}_{j=p,\ldots,1}$, $n = 1$, has more reliable behavior (similar results hold for $n = \infty$, see Table A.2). Table A.2 and A.3 in the Supplement further document that the tests $F_{s,1}$ and $F_{s,\infty}$ are undersized in finite samples.

One can combine inference based on the max-gap estimator and on the test $F_{p,n}$, defining a hybrid estimator $\widehat{s}_n$ that selects p when $F_{p,n}$ does not reject, and uses the restricted max-gap estimator over $\{0, \ldots, p-1\}$. The bottom row of Fig. 1 shows that the hybrid estimator with $n = 1$ performs well (similar results hold for $n = \infty$, see Tables A.4 and A.5).

Tables A.6 and A.7 show the corresponding results for the cointegration rank determination procedures based on the trace test of Johansen (1996) with or without Bartlett correction. For $s = p$, Table A.8 reports results for the trace test corrected as in Onatski and Wang (2019). The results of the hybrid estimator with $F_{p,1}$ in Table A.5 compare favourably with the ones of these benchmarks.

Next consider the Wald Q and t-ratio statistics for linear hypotheses on ψ, see Theorem

5.6. Two hypotheses of type (5.4) are considered, namely (i) the element 1,1 of ψ_* equals 0, tested with the t-statistic in (5.7), and (ii) the first column of ψ_* equals 0, tested using the Wald Q statistic in (5.6) with a χ^2_r weak limit; both hypotheses are true under the DGP.

The main challenge in the implementation of the tests lies in the precise estimation of the LRV $\Omega_{22.1}$, which is notoriously difficult by nonparametric means, see e.g. Haug (2002). Three different estimators of $\Omega_{22.1}$ are considered in the computations of the test statistics: the first one is the true value in the DGP for $\Omega_{22.1} = a^{-2} I_r$, called LRV-U (unfeasible); the second option employs $\widehat{\Omega}_{22.1}$ in (5.3) following Phillips (2014), called LRV-P, and the third one (LRV-A) computes $\widehat{\Omega}_{22.1}$ as in Andrews (1991) and Andrews and Monahan (1992) using Parzen's kernel.

The properties of the tests are simulated with $K = \lceil T^{3/4} \rceil$ for the following values of (p, T, s, a): $p = 10, 20$, $T = (30, 60, 90)p$, $s = \lceil jp/4 \rceil$, $j = 1, 2, 3$, $a = 0.25, 0.5, 0.75, 1$. Table A.9 in the supplement reports full Monte Carlo results, and Figure 2 reports the subset for $T = 30p$ and varying a. It can be seen that the tests are largely oversized in finite samples, in line with the literature (Haug, 2002).

Table A.10 shows the results for the likelihood-ratio test on β of Johansen (1996) of the hypothesis (ii), with or without Bartlett correction. The empirical size of these benchmarks appears closer to the nominal level than the one of the Q-test.

The unfeasible t-ratio and Q-test (LRV-U) perform considerably better than the feasible versions (LRV-P and LRV-A), thus suggesting that a major part of the size distortions are due to the imprecise LRV estimation. Finally, the $N(0, 1)$ approximation improves with increasing a (i.e. with decreasing level of autocorrelation of the stationary AR components).

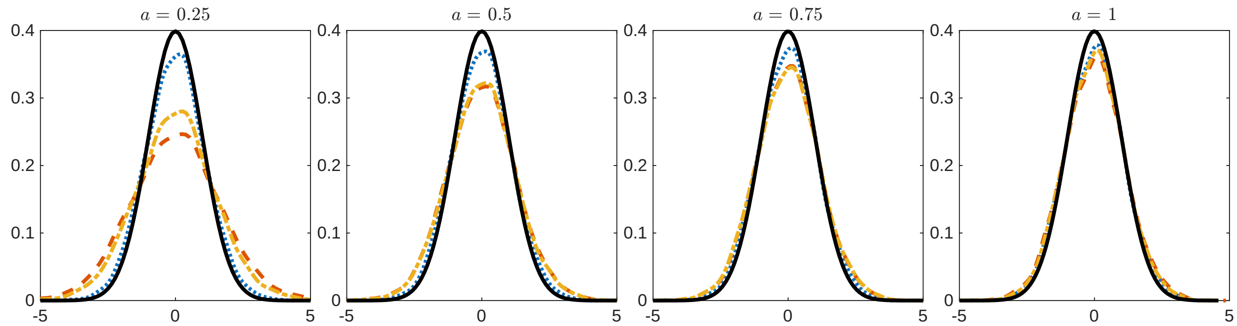

Figure 2: MC distributions of the t-ratio for $p = 20$, $T = 30p$, $s = \lceil p/4 \rceil$, $a = 0.25, 0.5, 0.75, 1$ and $N = 10^4$ replications. Lines: solid black $N(0,1)$, dotted blue t-ratio with LRV-U, dashed red t-ratio with LRV-P, dot-dashed yellow t-ratio with LRV-A.

7 Empirical application

This section provides an illustration on a panel of daily exchange rates between Jan 4, 2022 and Aug 30, 2024 of the US dollar against 20 World Markets (WM) currencies, see Onatski and Wang (2019) for a similar dataset and a review of the literature on the topic. The sample size is $T = 667$ and $K = \lceil T^{3/4} \rceil = 132$. The illustration of the present methods is compared with the high dimensional VAR analysis of Onatski and Wang (2018) and Bykhovskaya and Gorin (2024) for inference on s, and with the likelihood analysis of Johansen (1996) for inference on s and ψ. Details are reported in the supplement.

Data (in logs and normalized to start at 0) are plotted in the first four panels of Figure 3. According to the MSCI 2024 Market Classification, countries are grouped into Emerging Markets (EM: Brazil, China, India, Malaysia, Mexico, South Africa, South Korea, Taiwan, Thailand) and Developed Markets (DM). On the basis of economic and geographic criteria, DM are broken down into non European (Non-EU: Australia, Canada, Hong Kong, Japan, Singapore) and European (EU). EU is partitioned into UK-SZ (United Kingdom and Switzerland), and Nordic and Eurozone (Nc-EZ: Denmark, Eurozone, Norway, Sweden). The tools presented in this paper are employed to investigate the presence of stochastic trends in the whole system and in the country-groups described above.

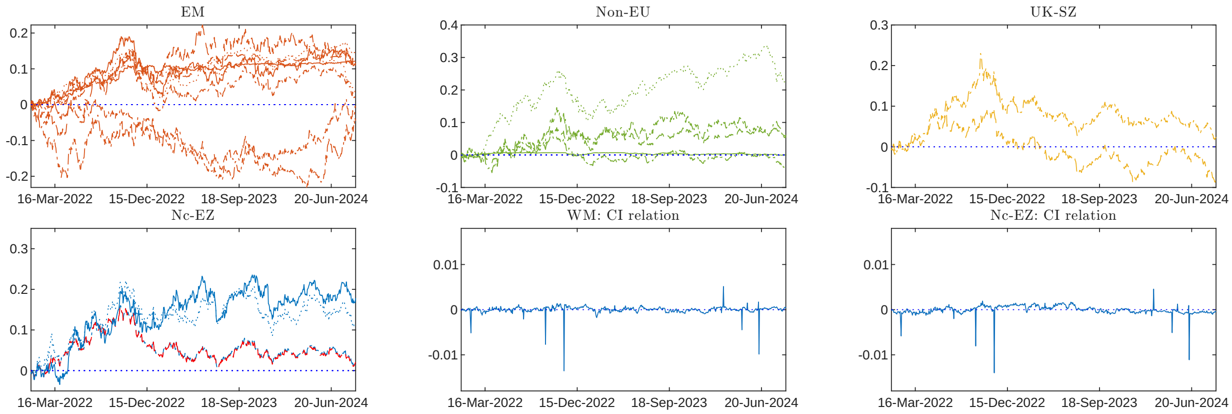

Figure 3: Top: EM, Non-EU, UK-SZ time series. Bottom: Nc-EZ time series, cointegrating relation in WM, cointegrating relation in Nc-EZ.

The presence of cointegration among the 20 WM currencies is first investigated using Wachter plots as in Onatski and Wang (2018) and Bykhovskaya and Gorin (2024), and the test of no cointegration ($s = p$) of Bykhovskaya and Gorin (2024) in a VAR(k), $k = 1, 2, 3, 4$. For all values of k, Figure *A*.2 in the supplement displays eigenvalues to the right of the support of the Wachter distribution, which is an indication of the presence of cointegration ($s < 20$). This is further confirmed by the results of their test at 1% significance level, see Table *A*.12 in the supplement.

The results of the present analysis are next considered. The CCA analysis applied to all the $p = 20$ WM time series is summarized in the first two subgraphs in Figure 4, which report the associated SCC profile and the misspecification stripe of Remark 4.6 at 95% confidence level.

The largest drop in the SCC profile is between eigenvalue 19 and 20 (red vertical line), so that the max-gap estimator selects $\widehat{s} = 19$, indicating $\widehat{r} = 1$ cointegrating relation (CI). This is confirmed by the sequential test procedure based on $K\pi^2\|\tau^{(i)}\|_\infty$, while the one based on $K\pi^2\|\tau^{(i)}\|_1$ indicates $s = 18$, see Table 1. Given these overall results, s is fixed at

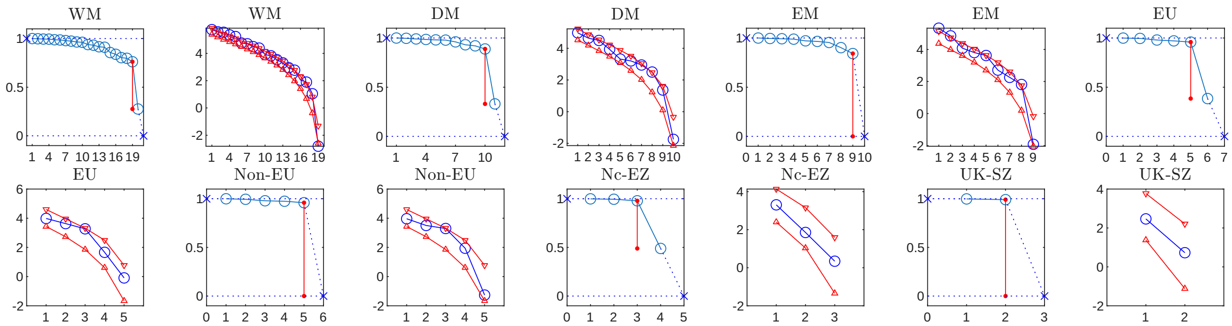

Figure 4: Pair of graphs for each CCA; SCC profile (first graph) with largest drop in red, misspecification stripe at 95% confidence level (second graph). Numbering in the text is left-to-right, top-to-bottom.

19 in the WM system.

$\widehat{s}\ \backslash$ group (p)	WM (20)	EM (9)	DM (11)	Non-EU (5)	EU (6)	UK-SZ (2)	Nc-EZ (4)
max-gap	19	9	10	5	5	2	3
$K\pi^2\|\tau^{(i)}\|_\infty$	19	8	10	5	5	2	3
$K\pi^2\|\tau^{(i)}\|_1$	18	8	10	5	5	2	3

Table 1: Estimates of s in the subgroups with the max-gap estimator and with the sequential tests at 1% significance level.

The plot of the cointegrating relation $\widehat{\beta}' X_t$ is reported in the center bottom panel of Figure 3; the plot includes some large infrequent shocks, whose analysis, however, goes beyond the scope of the present illustration. The estimate of the 20×1 cointegrating vector β in the WM system, normalized on the Euro, is reported in Table A.11 in the supplement.

Due to the difficulties related to the precise estimation of the LRV discussed in Section 6, in order to gain additional insight into the structure of the cointegrating relationship, the analysis is performed on the EM, DM, Non-EU, EU, UK-SZ, and Nc-EZ subgroups via the criteria for the determination of s. This allows to investigate whether the cointegrating relation in the WM system is due to countries belonging to a specific subgroup.

The max-gap estimator finds: no cointegration in EM and 1 CI relation in DM, no cointegration in Non-EU and 1 CI relation in EU, no cointegration in the UK-SZ group and 1 CI relation in the Nc-EZ group. The results for the sequential tests at 1% significance level agree, see Table 1, with the exception of the EM subgroup, where the max-gap selects $s = 9$ and the sequential tests indicate $s = 8$.

These overall results lead to concluding that the cointegrating relation in the WM system can be attributed to the Nc-EZ group.

The estimates of ψ and β in Nc-EZ subgroup are next reported, with identification restrictions $c = e_2$ and $b = (e_1, e_3, e_4)$, where e_i indicates the i-th unit vector; this identification restriction is tested using the criterion in (4.5) and it is not rejected. The resulting $\widehat{\beta}$ is

$$\widehat{\beta} = (\underset{(<0.01)}{-1.0147}, 1, \underset{(<0.01)}{0.0153}, \underset{(0.7256)}{0.0012}) \tag{7.1}$$

with p-values reported in parenthesis for the unrestricted coefficients $\psi_* = e_2'\widehat{\psi} = -\beta_*'$, which correspond to the second row of $\widehat{\psi}$. The hypothesis $H_0 : \psi_{2,1} = 0$ and $H_0 : \psi_{2,2} = 0$ are strongly rejected while $H_0 : \psi_{2,3} = 0$ is not, i.e. the Euro loads the first and second stochastic trends but not the third one, which only affects the Swedish Krone. The plot of the cointegrating relation $\widehat{\beta}' X_t$ in the Nc-EZ subgroup is reported in the bottom right panel of Figure 3; this CI relation essentially coincides with the one among the 20 WM, thus reinforcing the conclusion that cointegration in the 20 WM system is due to the Nc-EZ group. Finally, the estimates of s and β are compared with the ones derived from the likelihood ratio test on s and the QML estimate of β in Johansen (1996) for the Nc-EZ subgroup. Table A.13 in the supplement reports an estimate of s equal to 3, and an estimate of β essentially indistinguishable from the one in (7.1).

8 Conclusions

This paper proposes a unified semiparametric framework for inference on the number and the loadings of stochastic trends (and on their duals, the number and the coefficients of the cointegrating relations) for $I(1)/I(0)$ processes, which includes VARIMA and Dynamic Factor Model processes. The setup directly applies also to subsets of variables or to their aggregation, i.e. it is dimensionally coherent.

The approach employs a novel canonical correlation analysis between the observed multiple time series and a finite subset of an $L^2[0,1]$ basis. Canonical correlations deliver test sequences for the number of stochastic trends, as well as consistent estimators of it. Canonical variates provide estimators of the loading matrix of stochastic trends and of their duals, the matrix of cointegrating coefficients. The ICC estimators are T-consistent, mixed-Gaussian and efficient. Wald tests on the parameters are developed, as well as misspecification tests for checking model assumptions.

The finite sample properties of the estimators and tests investigated via a Monte Carlo simulation study display reasonable performance. An empirical analysis illustrates the range of available inferences in the proposed framework.

References

Ahn, S. and A. Horenstein (2013). Eigenvalue Ratio Test for the Number of Factors. *Econometrica 81*, 1203–1227.

Ahn, S. and C. Reinsel (1990). Estimation for partially nonstationary multivariate autoregressive models. *Journal of the American Statistical Association 85*, 813–23.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*(3), 817–858.

Andrews, D. W. K. and J. C. Monahan (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica 60*(4), 953–966.

Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics 122*, 137–183.

Barigozzi, M., G. Cavaliere, and L. Trapani (2022). Inference in Heavy-Tailed Nonstationary Multivariate Time Series. *Journal of the American Statistical Association*, 565–581.

Bierens, H. (1997). Nonparametric cointegration analysis. *Journal of Econometrics 77*, 379–404.

Bykhovskaya, A. and V. Gorin (2022). Cointegration in large VARs. *The Annals of Statistics 50*, 1593 – 1617.

Bykhovskaya, A. and V. Gorin (2024). Asymptotics of cointegration tests for high-dimensional VAR(k). *The Review of Economics and Statistics*, 1–38.

Cavaliere, G., H. B. Nielsen, and A. Rahbek (2015). Bootstrap testing of hypotheses on co-integration relations in vector autoregressive models. *Econometrica 83*, 813–831.

Cavaliere, G., A. Rahbek, and R. Taylor (2012). Bootstrap determination of the co-integration rank in vector autoregressive models. *Econometrica 80*, 1721–1740.

Chen, S. and M. Schienle (2024). Large Spillover Networks of Nonstationary Systems. *Journal of Business & Economic Statistics 42*, 422–436.

Elliott, G. (1999). Efficient Tests for a Unit Root When the Initial Observation is Drawn from Its Unconditional Distribution. *International Economic Review 40*, 767–783.

Engle, R. and C. Granger (1987). Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica 55*, 251–276.

Franchi, M., I. Georgiev, and P. Paruolo (2023). Estimating the number of common trends in large T and N factor models via canonical correlations analysis. *Econometrics and Statistics*, forthcoming.

Haug, A. A. (2002). Testing linear restrictions on cointegrating vectors: sizes and powers of Wald and likelihood ratio tests in finite samples. *Econometric Theory 18*(2), 505–524.

Hwang, J. and Y. Sun (2018). Simple, robust, and accurate F and t tests in cointegrated systems. *Econometric Theory 34*(5), 949–984.

Jacod, J. and A. Shiryaev (2003). *Limit Theorems for Stochastic Processes*. Springer.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control 12*, 231–254.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica 59*, 1551–1580.

Johansen, S. (1996). *Likelihood-based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford University Press.

Johansen, S. (2000). A Bartlett Correction Factor for Tests on the Cointegrating Relations. *Econometric Theory 16*, 740–778.

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

Johansen, S. (2002b). A Small Sample Correction for the Test of Cointegrating Rank in the Vector Autoregressive Model. *Econometrica 70*, 1929–1961.

Liang, C. and M. Schienle (2019). Determination of vector error correction models in high dimensions. *Journal of Econometrics 208*, 418–441.

Loève, M. (1978). *Probability theory II* (4th ed.). Graduate Texts in Mathematics. Springer.

Müller, U. and M. Watson (2008). Testing Models of Low-Frequency Variability. *Econometrica 76*, 979–1016.

Müller, U. and M. Watson (2017). Low-Frequency Econometrics. In e.-a. Honoré, B. (Ed.), *Advances in Economics and Econometrics: 11-th WC*, pp. 53–94. CUP.

Müller, U. and M. Watson (2018). Long-Run Covariability. *Econometrica 86*, 775–804.

Nielsen, M., W. Seo, and D. Seong (2022). Inference on the dimension of the nonstationary subspace in functional time series. *Econometric Theory 39*, 443–480.

Onatski, A. and C. Wang (2018). Alternative Asymptotics for Cointegration Tests in Large VARs. *Econometrica 86*, 1465–1478.

Onatski, A. and C. Wang (2019). Extreme canonical correlations and high-dimensional cointegration analysis. *Journal of Econometrics 212*, 307–322.

Paruolo, P. (1997). Asymptotic Inference on the Moving Average Impact Matrix in Cointegrated I(1) Var Systems. *Econometric Theory 13*, 79–118.

Petersen, A., C. Zhang, and P. Kokoszka (2022). Modeling probability density functions as data objects. *Econometrics and Statistics 21*, 159–178.

Phillips, P. and B. Hansen (1990). Statistical inference in Instrumental Variables regression with I(1) processes. *Review of Economic Studies 57*(1), 99–125.

Phillips, P. C. B. (1998). New Tools for Understanding Spurious Regressions. *Econometrica 66*, 1299–1326.

Phillips, P. C. B. (2005). HAC Estimation by Automated Regression. *Econometric Theory 21*, 116–142.

Phillips, P. C. B. (2014). Optimal estimation of cointegrated systems with irrelevant instruments. *Journal of Econometrics 178*, 210–224.

Phillips, P. C. B. and V. Solo (1992). Asymptotics for Linear Processes. *The Annals of Statistics 20*, 971 – 1001.

Stock, J. and M. Watson (1988). Testing for common trends. *Journal of the American Statistical Society 83*, 1097–1107.

Sun, Y. (2011). Robust trend inference with series variance estimator and testing-optimal smoothing parameter. *Journal of Econometrics 164*(2), 345–366.

Swensen, A. (2006). Bootstrap Algorithms for Testing and Determining the Cointegration Rank in VAR Models. *Econometrica 74*, 1699–1714.

Watson, M. W. (1994). Chapter 47: Vector autoregressions and cointegration. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, pp. 2843–2915. Elsevier.

Zhang, R., P. Robinson, and Q. Yao (2019). Identifying cointegration by eigenanalysis. *Journal of the American Statistical Association 114*, 916–927.

Supplementary material for "Canonical correlation analysis of stochastic trends via functional approximation"

M. Franchi, I. Georgiev, P. Paruolo[1]

March 23, 2026

Abstract

This is the supplementary material for Franchi, Georgiev and Paruolo (2026) *Canonical correlation analysis of stochastic trends via functional approximation*. It contains lemmas in Appendix A, proofs in Appendix B, Monte Carlo simulations in Appendix C, and the empirical application in Appendix D.

[1]M. Franchi, Sapienza University of Rome, ORCID: 0000-0002-3745-2233, Email: massimo.franchi@uniroma1.it, Address: P.le A. Moro 5, 00185 Rome, Italy.

I. Georgiev, University of Bologna, ORCID: 0000-0003-3392-806X, Email: i.georgiev@unibo.it, Address: Piazza Scaravilli 2, 40126 Bologna, Italy.

P. Paruolo, European Commission, Joint Research Centre (JRC), ORCID: 0000-0002-3982-4889, Email: paolo.paruolo@ec.europa.eu, Address: Via Enrico Fermi 2749, 21027 Ispra (VA), Italy.

A Lemmas

The following lemma collects well known facts on linear processes, see Phillips and Solo (1992) or Theorem B.13 in Johansen (1996) (Thm SJ.B.13 henceforth).

Lemma A.1 (Limits for fixed $K \geq p$ and $T \to \infty$). *Let $x_{1t} := \bar{\psi}' x_t$ and $x_{0t} := \beta' x_t$, and consider $\varphi_K(u) := (\phi_1(u), \ldots, \phi_K(u))'$ for any orthonormal càdlàg basis $\{\phi_i\}_{i=1}^{\infty}$ of $L^2[0,1]$; for fixed $K \geq p$ and $T \to \infty$, one has that, jointly:*

(i) $T^{-\frac{1}{2}} \sum_{t=1}^{\lfloor Tu \rfloor} x_{0t} \xrightarrow{w} W_2(u) := \beta' C_1 W_\varepsilon(u)$ *in* $D_r[0,1]$*;*

(ii) $T^{-\frac{1}{2}} x_{1\lfloor Tu \rfloor} \xrightarrow{w} W_1(u) := \kappa' W_\varepsilon(u)$ *in* $D_s[0,1]$*;*

(iii) $\varphi_K(\lfloor Tu \rfloor / T) \to \varphi_K(u)$ *in* $D_K[0,1]$ *and* $L_K^2[0,1]$*;*

(iv) $T^{-1} \bar{\psi}' M_{xx} \bar{\psi} = T^{-1} \sum_{t=1}^{T} (T^{-\frac{1}{2}} x_{1t})(T^{-\frac{1}{2}} x_{1t})' \xrightarrow{w} \ell_{11} := \int W_1 W_1'$*;*

(v) $\bar{\psi}' M_{xx} \beta = T^{-\frac{1}{2}} \sum_{t=1}^{T} (T^{-\frac{1}{2}} x_{1t}) x_{0t}' = O_p(1)$*;*

(vi) $\beta' M_{xx} \beta = T^{-1} \sum_{t=1}^{T} x_{0t} x_{0t}' \xrightarrow{p} Q_{00} := \mathbb{E}(x_{0t} x_{0t}') > 0$*;*

(vii) $T^{-\frac{1}{2}} \bar{\psi}' M_{xd} = T^{-1} \sum_{t=1}^{T} (T^{-\frac{1}{2}} x_{1t}) \varphi_K(t/T)' \xrightarrow{w} \ell_{1\varphi} := \int W_1 \varphi_K'$*;*

(viii) $T^{1/2} \beta' M_{xd} = T^{-\frac{1}{2}} \sum_{t=1}^{T} x_{0t} \varphi_K(t/T)' \xrightarrow{w} \ell_{\mathrm{d}2\varphi}^{\circ} := \int \mathrm{d}W_2 \varphi_K'$*;*

(ix) $T^{1/2} \bar{\psi}' M_{\Delta xd} = T^{-\frac{1}{2}} \sum_{t=1}^{T} \Delta x_{1t} \varphi_K(t/T)' \xrightarrow{w} \ell_{\mathrm{d}1\varphi}^{\circ} := \int \mathrm{d}W_1 \varphi_K'$*;*

(x) $T^{1/2} \beta' M_{\Delta xd} = T^{-\frac{1}{2}} \sum_{t=1}^{T} \Delta x_{0t} \varphi_K(t/T)' = o_p(1)$*.*

Proof. (i) From (2.5) and $T^{-\frac{1}{2}} \sum_{t=1}^{\lfloor Tu \rfloor} x_{0t} = T^{-\frac{1}{2}} \sum_{t=1}^{\lfloor Tu \rfloor} \beta' C_1 \varepsilon_t + o_p(1)$ uniformly in $u \in [0,1]$. (ii) From (2.5) and $T^{-\frac{1}{2}} x_{1\lfloor Tu \rfloor} = T^{-\frac{1}{2}} \kappa' \sum_{t=1}^{\lfloor Tu \rfloor} \varepsilon_t + o_p(1)$ uniformly in $u \in [0,1]$. (iii) By the dominated convergence theorem for $L_K^2[0,1]$. (iv) By (B.19) in Thm SJ.B.13.

(v) By (B.20) in Thm SJ.B.13. (vi) By the law of large numbers for linear processes. (vii) By (B.22) in Thm SJ.B.13. (viii)-(x) By (B.23) in Thm SJ.B.13. ■

Let

$$g_t(a) := a' M_{\Delta x d} M_{dd}^{-1} d_t, \qquad e_t(a) := x_t - M_{xg} M_{gg}^{-1} g_t(a). \tag{A.1}$$

Note that substituting $a = \widehat{\psi}^{(1)}$ in (A.1) one finds (3.5). The next lemma collects the limits of the sample moments M_{ij} that appear in (3.2) for $\text{CCA}(f_t, d_t)$, $f = x, e$.

Lemma A.2 ($T \to \infty$ limits for fixed $K \geq p$)**.** *The following $T \to \infty$ limits hold jointly and jointly with those in Lemma A.1, for fixed $K \geq p$. Consider d_t in (3.1) constructed using any orthonormal càdlàg basis of $L^2[0,1]$ and define $A_T = (T^{-\frac{1}{2}}\bar{\psi}, \beta)$; one has that:*

(i) $A_T' M_{xx} A_T \xrightarrow{w} \text{diag}(\ell_{11}, Q_{00})$;

(ii) $M_{dd} = T^{-1} \sum_{t=1}^{T} \varphi_K(t/T) \varphi_K(t/T)' \to \int \varphi_K \varphi_K' = I_K$;

(iii) $A_T' M_{xd} \xrightarrow{w} (\ell_{1\varphi}', 0')'$.

Let g_t, e_t be $g_t(\psi)$ and $e_t(\psi)$ in (A.1) using the true parameter value of ψ;[2] *then one has that:*

(iv) $a_4(\psi) := T\beta' M_{xg} \xrightarrow{w} \ell^{\circ}_{\mathrm{d}2\varphi} \ell^{\circ}_{\varphi\mathrm{d}1}$;

(v) $a_5(\psi) := T M_{gg} \xrightarrow{w} \ell^{\circ}_{\mathrm{d}1\varphi} \ell^{\circ}_{\varphi\mathrm{d}1}$;

(vi) $a_6(\psi) := T^{1/2} M_{gd} \xrightarrow{w} \ell^{\circ}_{\mathrm{d}1\varphi}$;

(vii) $a_7(\psi) := A_T' M_{xg} M_{gg}^{-1} M_{gx} A_T \xrightarrow{w} \text{diag}(\ell_{1\varphi} U_\circ \ell_{\varphi 1}, 0)$, *where* $U_\circ := \ell^{\circ}_{\varphi\mathrm{d}1} (\ell^{\circ}_{\mathrm{d}1\varphi} \ell^{\circ}_{\varphi\mathrm{d}1})^{-1} \ell^{\circ}_{\mathrm{d}1\varphi}$;

(viii) $a_8(\psi) := A_T' M_{xg} M_{gg}^{-1} M_{gd} \xrightarrow{w} ((\ell_{1\varphi} U_\circ)', 0')'$;

[2]Points (iv) and (xii) involve β, which is a function of ψ, $\beta = (I - b\psi')\bar{c}$ in the present notation, see Paruolo (1997) eq. (B.1).

(ix) $a_9(\psi) := M_{dg}M_{gg}^{-1}M_{gd} \xrightarrow{w} U_\circ$;

(x) $a_{10}(\psi) := A_T' M_{ee} A_T \xrightarrow{w} \mathrm{diag}(\ell_{11} - \ell_{1\varphi}U_\circ \ell_{\varphi 1}, Q_{00})$;

(xi) $a_{11}(\psi) := A_T' M_{ed} \xrightarrow{w} ((\ell_{1\varphi}P_\circ)', 0')'$, *where* $P_\circ := I_K - U_\circ$;

(xii) $a_{12}(\psi) := T^{1/2}\beta' M_{ed} \xrightarrow{w} \ell^\circ_{\mathrm{d}2\varphi}P_\circ$.

Finally, local tightness results hold for the sample moments in items (vii)-(xii)*, in the following sense. Let* $\widetilde{\psi}$ *be any estimator of* ψ *such that for an arbitrary* $L > 0$

$$\widetilde{\psi} = \psi(h_1 + h_2 o_p(1)) + \beta h_3 o_p(T^{-\frac{1}{2}}), \quad \mathrm{eig}_{\min}(h_1' h_1) \overset{a.s.}{\geq} L^{-1}, \quad \|h_i\| \overset{a.s.}{\leq} L, \quad i = 1, 2, 3, \tag{A.2}$$

with $h_1, h_2 \in \mathbb{R}^{s\times s}$ *and* $h_3 \in \mathbb{R}^{r\times r}$*, and where* $\mathrm{eig}_{\min}$ *is the minimal eigenvalue of the argument matrix. Let* $\tilde{g}_t := g_t(\widetilde{\psi})$*,* $\tilde{e}_t := e_t(\widetilde{\psi})$ *be the quantities* g_t, e_t *in* (A.1) *calculated for this estimator* $\widetilde{\psi}$*; then the replacement of* e, g *by* $\tilde{e}, \tilde{g}$ *in items* (vii)-(xii) *produces* $o_p(1)$ *perturbations uniformly over* h_1, h_2, h_3*, i.e.* $\|a_i(\widetilde{\psi}) - a_i(\psi)\| = o_p(1)$*,* $i = 4, \ldots, 12$.

Proof. (i) By Lemma *A.1*(iv),(v),(vi), one finds

$$A_T' M_{xx} A_T = T^{-1}\sum_{t=1}^{T} \begin{pmatrix} (T^{-\frac{1}{2}}x_{1t})(T^{-\frac{1}{2}}x_{1t})' & (T^{-\frac{1}{2}}x_{1t})x_{0t}' \\ x_{0t}(T^{-\frac{1}{2}}x_{1t})' & x_{0t}x_{0t}' \end{pmatrix} \xrightarrow{w} \begin{pmatrix} \ell_{11} & 0 \\ 0 & Q_{00} \end{pmatrix}.$$

(ii) By the continuous mapping theorem and orthonormality of the basis in (2.8). (iii) By Lemma *A.1*(vii),(viii), one finds

$$A_T' M_{xd} = T^{-1}\sum_{t=1}^{T} \begin{pmatrix} (T^{-\frac{1}{2}}x_{1t})d_t' \\ x_{0t}d_t' \end{pmatrix} \xrightarrow{w} \begin{pmatrix} \ell_{1\varphi} \\ 0 \end{pmatrix}.$$

(iv) By (ii) and Lemma *A.1*(viii), (ix), one finds

$$T\beta' M_{xg} = \left(T^{-\frac{1}{2}}\sum_{t=1}^{T} x_{0t}d_t'\right) M_{dd}^{-1} T^{1/2} M_{d\Delta x_1} \xrightarrow{w} \ell^\circ_{\mathrm{d}2\varphi}\ell^\circ_{\varphi\mathrm{d}1}.$$

(v) By (ii) and Lemma A.1(ix), one finds $TM_{gg} = (T^{1/2}M_{\Delta x_1 d})M_{dd}^{-1}(T^{1/2}M_{d\Delta x_1}) \xrightarrow{w} \ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1}$.

(vi) By Lemma A.1(ix), one finds $T^{1/2}M_{gd} = T^{-1}\sum_{t=1}^{T}(T^{1/2}g_t)d_t' = T^{1/2}M_{\Delta x_1 d} \xrightarrow{w} \ell^{\circ}_{\mathrm{d}1\varphi}$.

(vii) Consider $A_T'M_{xg}M_{gg}^{-1}M_{gx}A_T$, where

$$
\begin{aligned}
T^{1/2}A_T'M_{xg} = T^{-1}\sum_{t=1}^{T}\begin{pmatrix} (T^{-\frac{1}{2}}x_{1t})(T^{1/2}g_t)' \\ x_{0t}(T^{1/2}g_t)' \end{pmatrix} &= T^{-1}\sum_{t=1}^{T}\begin{pmatrix} (T^{-\frac{1}{2}}x_{1t})d_t' \\ x_{0t}d_t' \end{pmatrix} M_{dd}^{-1}(T^{1/2}M_{d\Delta x_1}), \\
&\xrightarrow{w} \begin{pmatrix} \ell_{1\varphi} \\ 0 \end{pmatrix} \ell^{\circ}_{\varphi\mathrm{d}1} \qquad \text{(A.3)}
\end{aligned}
$$

by (ii) and Lemma A.1(vii),(viii),(ix); hence, by (v) one has

$$
A_T'M_{xg}M_{gg}^{-1}M_{gx}A_T = (T^{1/2}A_T'M_{xg})(TM_{gg})^{-1}(T^{1/2}M_{gx}A_T) \xrightarrow{w} \begin{pmatrix} \ell_{1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1}(\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1})^{-1}\ell^{\circ}_{\mathrm{d}1\varphi}\ell_{\varphi 1} & 0 \\ 0 & 0 \end{pmatrix}.
$$

(viii) By (A.3), (v) and (vi), one has

$$
A_T'M_{xg}M_{gg}^{-1}M_{gd} = (T^{1/2}A_T'M_{xg})(TM_{gg})^{-1}(T^{1/2}M_{gd}) \xrightarrow{w} \begin{pmatrix} \ell_{1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1}(\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1})^{-1}\ell^{\circ}_{\mathrm{d}1\varphi} \\ 0 \end{pmatrix}.
$$

(ix) By (v) and (vi), one has $M_{dg}M_{gg}^{-1}M_{gd} = (T^{1/2}M_{dg})(TM_{gg})^{-1}(T^{1/2}M_{gd}) \xrightarrow{w} \ell^{\circ}_{\varphi\mathrm{d}1}(\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1})^{-1}\ell^{\circ}_{\mathrm{d}1\varphi}$.

(x) By (ii) and (ix), using $M_{ee} = M_{xx} - M_{xg}M_{gg}^{-1}M_{gx}$. (xi) By (iii) and (viii), using $M_{ed} = M_{xd} - M_{xg}M_{gg}^{-1}M_{gd}$. (xii) By Lemma A.1(vii),(iv),(v),(vi), one has

$$
\begin{aligned}
T^{1/2}\beta'M_{ed} &= T^{1/2}\beta'M_{xd} - (T\beta'M_{xg})(TM_{gg})^{-1}(T^{1/2}M_{gd}) \\
&\xrightarrow{w} \ell^{\circ}_{\mathrm{d}2\varphi} - \ell^{\circ}_{\mathrm{d}2\varphi}\ell^{\circ}_{\varphi\mathrm{d}1}(\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1})^{-1}\ell^{\circ}_{\mathrm{d}1\varphi}.
\end{aligned}
$$

The tight versions of items (vii)-(xii) follow from the tight versions of items (iv)-(vi). As to the latter,

$$
T\|\beta'(M_{xg}h_1 - M_{x\tilde{g}})\| \leq o_p(1)\|T\beta'M_{xg}\|\|h_2\| + o_p(T^{-\frac{1}{2}})\|\beta'M_{xd}\|\|M_{dd}^{-1}\|\|M_{d\Delta x}\beta\|\|h_3\| = o_p(1)
$$

by Lemma A.1(viii),(x) and Lemma A.2(ii),(iv); similarly, uniformly over h_1, h_2, h_3 satisfying (A.2), one has

$$T\|(h_1' M_{gg} h_1) - M_{\tilde{g}\tilde{g}}\| = o_p(1),$$

$$T^{1/2}\|h_1' M_{gd} - M_{\tilde{g}d}\| = o_p(1),$$

$$T^{-1}\|(h_1' M_{gg} h_1)^{-1} - M_{\tilde{g}\tilde{g}}^{-1}\| \leq \|(Th_1' M_{gg} h_1)^{-1}\| \|(Th_1' M_{gg} h_1) - TM_{\tilde{g}\tilde{g}}\| \|(TM_{\tilde{g}\tilde{g}})^{-1}\| = o_p(1).$$

Finally note that a_7, a_8, a_9 are invariant with respect to the transformation from g_t to $h_1' g_t$ with h_1 square and nonsingular. This completes the proof. ■

The tightness results in Lemma A.2 imply that the limits in items (x)-(xii) therein are invariant to the replacement of $e_t(\psi)$ with $e_t(\widetilde{\psi})$ for any estimator $\widetilde{\psi} \in \mathbb{R}^{p\times s}$ of ψ such that $\bar{\psi}'\widetilde{\psi}$ converges in distribution to an a.s. nonsingular matrix h_1 and $\beta'\widetilde{\psi} = o_p(T^{-\frac{1}{2}})$, as $T \to \infty$. This is the case of $\widetilde{\psi} = \widehat{\psi}^{(1)}$ with $h_1 = I_s$, see Theorem *B.1*.

The following lemma discusses the limit of $S(\lambda) := \lambda M_{ff} - M_{fd} M_{dd}^{-1} M_{df}$, $f = x, e$, that appears in (3.2) for $\text{CCA}(f_t, d_t)$, $f = x, e$, where $e_t := e_t(\psi)$ (or asymptotically equivalently, $e_t(\widetilde{\psi})$ with $\widetilde{\psi}$ as in the previous paragraph).

Lemma A.3 (Limits for fixed $K \geq p$ and $T \to \infty$). *Consider $A_T = (T^{-\frac{1}{2}}\bar{\psi}, \beta)$ and let $S(\lambda) := \lambda M_{ff} - M_{fd} M_{dd}^{-1} M_{df}$, $f = x, e$; for fixed $K \geq p$ and $T \to \infty$, one has that*

$$A_T' S(\lambda) A_T \xrightarrow{w} \lambda \begin{pmatrix} \varsigma_{11} & 0 \\ 0 & Q_{00} \end{pmatrix} - \begin{pmatrix} \varsigma_{1\varphi}\varsigma_{\varphi 1} & 0 \\ 0 & 0 \end{pmatrix},$$

where

$$\varsigma_{11} = \begin{cases} \ell_{11} & \text{for } f = x \\ \ell_{11} - \ell_{1\varphi} U_\circ \ell_{\varphi 1} & \text{for } f = e \end{cases}, \qquad \varsigma_{1\varphi} = \begin{cases} \ell_{1\varphi} & \text{for } f = x \\ \ell_{1\varphi} P_\circ & \text{for } f = e \end{cases}, \tag{A.4}$$

and $U_\circ$ and $P_\circ$ are as in Lemma A.2.

Proof. For $f = x$, by Lemma A.2(i),(ii),(iii), using $S(\lambda) = \lambda M_{xx} - M_{xd}M_{dd}^{-1}M_{dx}$; for $f = e$, by Lemma A.2(ii),(x),(xi), using $S(\lambda) = \lambda M_{ee} - M_{ed}M_{dd}^{-1}M_{de}$. ■

Lemma A.4 (Convergence to $\int B\partial W$)**.** *Let B and W be two scalar standard Brownian motions, with $c_k = \int B\phi_k$ and $g_k^\circ = \int (\mathrm{d}W)\phi_k$; then as $K \to \infty$ one has $\sum_{k=1}^K c_k g_k^\circ \xrightarrow{p} \int B\partial W$, a Stratonovich integral, for any orthonormal basis $\{\phi_k\}_{k=1}^\infty$ of $L^2[0,1]$.*

Proof. By linearity considerations, the claim follows from two special cases: (i) when $B = W$ is the same Brownian Motion, proved in Ogawa (2017, Example 3.1), and (ii) when B and W are two independent standard Brownian Motions. Consider two steps to prove case (ii).

First, $\sum_{k=1}^K c_k g_k^\circ \xrightarrow{p} \int B\partial W$ for the KL basis by Wong-Zakai's theorem, see Twardowska (1996). Second, it can be shown that the probability limit of $\sum_{k=1}^K c_k g_k^\circ$ is independent of the choice of an orthonormal basis. Indeed, consider an orthonormal basis $\{\psi_k\}$ with $\gamma_k = \int B\psi_k$ and $\delta_k^\circ = \int \mathrm{d}W\psi_k$. Because of the independence of B and W, the a.s. equality of $\sum_{k=1}^\infty c_k g_k^\circ$ and $\sum_{k=1}^\infty \gamma_k \delta_k^\circ$ can be established by conditioning on B (effectively on c_k, γ_k).

Observe in fact that both $\{g_k^\circ\}$ and $\{\delta_k^\circ\}$ are i.i.d. $N(0,1)$, whereas $\sum_{k=1}^K c_k^2$ and $\sum_{k=1}^K \gamma_k^2$ converge a.s.; hence it follows that $\sum_{k=1}^K c_k g_k^\circ$ and $\sum_{k=1}^K \gamma_k \delta_k^\circ$ converge in L^2-norm a.s. conditionally on B. Further,

$$
\begin{aligned}
\mathbb{E}_B\left(\sum_{k=1}^\infty c_k g_k^\circ - \sum_{k=1}^\infty \gamma_k \delta_k^\circ\right)^2 &= \operatorname*{plim}_{K\to\infty} \mathbb{E}_B\left(\sum_{k=1}^K c_k g_k^\circ - \sum_{k=1}^K \gamma_k \delta_k^\circ\right)^2 \\
&= \operatorname*{plim}_{K\to\infty}\left\{\sum_{k=1}^K (c_k^2 + \gamma_k^2) - 2\sum_{k,l=1}^K c_k\gamma_l \int \phi_k\psi_l\right\} \\
&= 2\int B^2 - 2\operatorname*{plim}_{K\to\infty}\int\left(\sum_{k=1}^K c_k\phi_k\right)\left(\sum_{k=1}^K \gamma_k\psi_k\right) \\
&= 2\int B^2 - 2\int B^2 = 0,
\end{aligned}
$$

where $\mathbb{E}_B$ indicates the expectation operator conditional on B. This shows that $\sum_{k=1}^\infty c_k g_k^\circ =$

$\sum_{k=1}^{\infty} \gamma_k \delta_k^\circ$ a.s. in conditional probability, and hence, also unconditionally. ∎

Lemma *A.5* below collects facts on large-K asymptotic theory related to Parseval's theorem, the ergodic theorem and Ogawa's stochastic integration theory, see e.g. Pedersen (2012) and Ogawa (2017). Recall the L^2-expansion (2.8) of $W(u)$; by Parseval's theorem applied to almost all sample paths, it holds that

$$\sum_{k=1}^{\infty} c_k c_k' = \int W W', \qquad c_k := \int W \phi_k,$$

where the series is a.s. convergent. If W was continuously differentiable with gradient $\dot{W}$, again by Parseval's theorem it would hold that

$$\sum_{k=1}^{\infty} c_k^\circ c_k' = \int \dot{W} W', \qquad \sum_{k=1}^{\infty} c_k^\circ c_k^{\circ\prime} = \int \dot{W} \dot{W}' \tag{A.5}$$

with $c_k^\circ := \int \mathrm{d}W \phi_k$. Although W is not differentiable, c_k° are well-defined as stochastic integrals, and properties formally similar to the relations in (A.5) do hold, though with no connection to Parseval's theorem. On the one hand,

$$\sum_{k=1}^{\infty} c_k^\circ c_k' = \int (\delta_\phi W) W',$$

a noncausal stochastic integral with respect to the basis $\{\phi_k\}_{k=1}^{\infty}$ in the sense of Ogawa (2017, Chapter 3), where the series on the l.h.s. is required to converge in probability. By Lemma *A.4* applied componentwise, it is seen that $\int (\delta_\phi W) W'$ exists and equals the Stratonovich integral $\int \partial W W'$ independently of the choice of $\{\phi_k\}_{k=1}^{\infty}$; this is the formal analogue of the first relation in (A.5). On the other hand, the orthonormality of $\{\phi_k\}_{k=1}^{\infty}$ is sufficient for $\{c_k^\circ\}_{k=1}^{\infty}$ to be an i.i.d. $N(0, \Omega)$ sequence; thus by the ergodic theorem,

$$K^{-1} \sum_{k=1}^{K} c_k^\circ c_k^{\circ\prime} \xrightarrow[K\to\infty]{a.s.} \int \mathrm{d}W \mathrm{d}W' = \Omega,$$

where the integrand is understood as the quadratic variation $\mathrm{d}W(u)\mathrm{d}W(u)' = \Omega \mathrm{d}u$; this is the formal analogue of the second relation in (A.5). It is by virtue of this latter property that some interactions of stationary processes and the basis functions are not asymptotically negligible and can be used to cleanse limit distributions from dependence.

Lemma A.5 (Limits for $K \to \infty$)**.** *For any given orthonormal càdlàg basis $\{\phi_k\}_{k=1}^{\infty}$ of $L^2[0,1]$, consider*

$$\ell_{00} := \int WW', \qquad \ell_{0\varphi} := \int W\varphi_K', \qquad \ell^{\circ}_{\mathrm{d}0\varphi} := \int \mathrm{d}W\varphi_K',$$

where $\varphi_K := (\phi_1, \ldots, \phi_K)'$ and $W := (W_1', W_2')'$, see (2.5), with L^2-representation in (2.8); for $K \to \infty$, one has that:

(i) $\ell_{0\varphi}\ell_{\varphi 0} \overset{a.s.}{\to} \ell_{00}$*;*

(ii) $K^{-1}\ell^{\circ}_{\mathrm{d}0\varphi}\ell^{\circ}_{\varphi\mathrm{d}0} \overset{a.s.}{\to} \Omega$.

(iii) $\ell^{\circ}_{\mathrm{d}0\varphi}\ell_{\varphi 0} \overset{p}{\to} \int \partial WW'$*, where $\int \partial WW'$ indicates a Stratonovich integral;*

(iv) $\ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1} \overset{p}{\to} \int \partial W_2 W_1' \left(\int W_1W_1'\right)^{-1}$*;*

(v) $\ell^{\circ}_{\mathrm{d}2\varphi}P_{\circ}\ell_{\varphi 1}\left(\ell_{1\varphi}P_{\circ}\ell_{\varphi 1}\right)^{-1} \overset{p}{\to} \int \mathrm{d}W_{2.1}W_1' \left(\int W_1W_1'\right)^{-1}$*, where $W_{2.1}(u)$ is as in (2.7).*

Proof. (i) By definition, $\ell_{0\varphi} = (c_1, c_2, \ldots, c_K)$, so that $\ell_{0\varphi}\ell_{\varphi 0} = \sum_{k=1}^{K} \nu_k^2\xi_k\xi_k'$; by Parseval's theorem, $\ell_{00} = \sum_{k=1}^{\infty} \nu_k^2\xi_k\xi_k'$ a.s. and the statement follows from the a.s. convergence of the latter series. (ii) By definition, $\ell^{\circ}_{\mathrm{d}0\varphi} = (c_1^{\circ}, c_2^{\circ}, \ldots, c_K^{\circ})$, where $c_k^{\circ} := \int \mathrm{d}W\phi_k \sim N(0,\Omega)$ and $\{c_k^{\circ}\}_{k=1}^{\infty}$ is i.i.d. because the basis $\{\phi_k\}_{k=1}^{\infty}$ is orthonormal. Hence $\ell^{\circ}_{\mathrm{d}0\varphi}\ell^{\circ}_{\varphi\mathrm{d}0} = \sum_{k=1}^{K} c_k^{\circ}c_k^{\circ\prime}$ and the statement follows from the ergodic theorem. (iii) This follows by Lemma *A.4* applied componentwise. (iv) By points (iii) and (i). (v) It is first shown that $\ell^{\circ}_{\mathrm{d}2\varphi}P_{\circ}\ell_{\varphi 1} \overset{p}{\to} \int \mathrm{d}W_{2.1}W_1'$. Observe that $\ell^{\circ}_{\mathrm{d}2\varphi}P_{\circ} = \ell^{\circ}_{\mathrm{d}2\varphi} - H_K\ell^{\circ}_{\mathrm{d}1\varphi}$, where $H_K := \ell^{\circ}_{\mathrm{d}2\varphi}\ell^{\circ}_{\varphi\mathrm{d}1}(\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1})^{-1}$. By point (ii), $H_K \xrightarrow{p} \Omega_{21}\Omega_{11}^{-1}$ and hence, again by Lemma A.4 applied componentwise,

$$\begin{aligned}\ell^{\circ}_{\mathrm{d}2\varphi} P_{\circ} \ell_{\varphi 1} &= \ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1} - H_K \ell^{\circ}_{\mathrm{d}1\varphi}\ell_{\varphi 1} \\ &\xrightarrow{p} \int \partial W_2 W_1' - \Omega_{21}\Omega_{11}^{-1} \int \partial W_1 W_1' = \int \partial [W_2 - \Omega_{21}\Omega_{11}^{-1} W_1] W_1' \\ &= \int \partial W_{2.1} W_1' = \int \mathrm{d} W_{2.1} W_1',\end{aligned}$$

where the last equality follows by the independence of the processes $W_{2.1}$ and W_1. Furthermore,

$$\ell_{1\varphi} P_{\circ} \ell_{\varphi 1} = \ell_{1\varphi}\ell_{\varphi 1} - K^{-1}\ell_{1\varphi}\ell^{\circ}_{\varphi \mathrm{d}1}(K^{-1}\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi \mathrm{d}1})^{-1}\ell^{\circ}_{\mathrm{d}1\varphi}\ell_{\varphi 1} = \sum_{k=1}^{K} \nu_k^2 \xi_{k1}\xi_{k1}' + O_p(K^{-1}),$$

because $\ell_{1\varphi}\ell^{\circ}_{\varphi \mathrm{d}1} = (\ell^{\circ}_{\mathrm{d}1\varphi}\ell_{\varphi 1})' \xrightarrow{p} \int W_1 \partial W_1'$ and $K^{-1}\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi \mathrm{d}1} \xrightarrow{p} \Omega_{11}$ by points (iii) and (ii). Thus, $\ell_{1\varphi} P_{\circ} \ell_{\varphi 1} \xrightarrow{p} \sum_{k=1}^{\infty} \nu_k^2 \xi_{k1}\xi_{k1}' = \int W_1 W_1'$ as $K \to \infty$. ∎

B Proofs

Proof of Lemma 2.3. Results (i) and (ii) are well known, see e.g. Exercise 3.7 and Section 13.2 in Johansen (1996), while (iii) and (iv) are proved in eq. (B.1) and (B.2) in Paruolo (1997), see also Paruolo (2006). ∎

Proof of Theorem 2.4. Consider H of dimension $p \times m$ and full column rank. Assumption 2.1.(i) is satisfied by $G(z) := H'C(z)$, of dimension $m \times n_\varepsilon$, because its disc of convergence is at least equal to the one of $C(z)$. Assumption 2.1.(ii) is also satisfied by $G(z)$ because $m \le p \le n_\varepsilon$, so that $\operatorname{rank} H'C(z) < m$ only when $C(z)$ is of reduced rank; write $G(z) = G + G_1(z)(1-z)$, where $G_1 := G_1(1) = H'C_1$ and $G = H'C = H'\psi\kappa'$ of rank $q \le s$. Here superscripts $\breve{}$ are omitted for brevity and because the results are invariant to the choice of bases of various spaces. Let ϱ be a basis of $\operatorname{col} G = \operatorname{col}(H'\psi)$ and write $H'\psi = \varrho\vartheta'$; then $G = \varrho\tau'$, where $\tau := \kappa\vartheta$, is a rank decomposition of G with ϱ and τ full column rank matrices with q columns.

Next it is shown that Assumption 2.1.(iii) also holds for $G(z)$, i.e. that $\varrho_\perp' G_1 \tau_\perp = \varrho_\perp' H' C_1 \tau_\perp$ has full row rank $m-q$. First note that $H\varrho_\perp = \psi_\perp g$ where $g := \bar{\psi}_\perp' H \varrho_\perp$; in fact, $H\varrho_\perp = (\bar{\psi}\psi' H + \psi_\perp \bar{\psi}_\perp' H)\varrho_\perp = (\bar{\psi}\vartheta\varrho' + \psi_\perp\bar{\psi}_\perp' H)\varrho_\perp = \psi_\perp g$. Next note that $\tau_\perp = (\kappa_\perp, \bar{\kappa}\vartheta_\perp)$, so that

$$\varrho_\perp' G_1 \tau_\perp = \varrho_\perp' H' C_1 \tau_\perp = g'\psi_\perp' C_1(\kappa_\perp, \bar{\kappa}\vartheta_\perp) = g'(\psi_\perp' C_1 \kappa_\perp, \psi_\perp' C_1 \bar{\kappa}\vartheta_\perp),$$

where $\operatorname{rank}(\psi_\perp' C_1 \kappa_\perp) = p-s$ by Assumption 2.1(iii) on $C(z)$ and $g := \bar{\psi}_\perp' H\varrho_\perp$ is $(p-s)\times(m-q)$ of full column rank $m-q$. This completes the proof. ■

Proof of Theorem 4.1. (i) Consider $\textsc{cca}(x_t, d_t)$ and normalize $S(\lambda) := \lambda M_{xx} - M_{xd}M_{dd}^{-1}M_{dx}$ in (3.2) as $A_T' S(\lambda) A_T = \lambda A_T' M_{xx} A_T - A_T' M_{xd} M_{dd}^{-1} M_{dx} A_T$, where $A_T = (T^{-\frac{1}{2}}\bar{\psi}, \beta)$; by Lemma *A*.3 one finds

$$A_T' S(\lambda) A_T \xrightarrow{w} \lambda \begin{pmatrix} \ell_{11} & 0 \\ 0 & Q_{00} \end{pmatrix} - \begin{pmatrix} \ell_{1\varphi}\ell_{\varphi 1} & 0 \\ 0 & 0 \end{pmatrix}, \tag{B.1}$$

where $\ell_{\varphi 1} = \ell_{1\varphi}'$. From (B.1) one finds that the s largest eigenvalues in (3.2) converge weakly to the ones of $\ell_{11}^{-1}\ell_{1\varphi}\ell_{\varphi 1}$, which coincide with those of the symmetric and a.s. positive definite $s \times s$ matrix $\ell_{11}^{-1/2}\ell_{1\varphi}\ell_{\varphi 1}\ell_{11}^{-1/2}$, and the $r = p - s$ smallest eigenvalues converge to 0 in probability. Observe that the eigenvalues of $\ell_{11}^{-1}\ell_{1\varphi}\ell_{\varphi 1}$ are invariant with respect to the transformation of $W_1(u)$ into $AW_1(u)$ for any nonsingular $A \in \mathbb{R}^{s\times s}$; choosing $A = \Omega_{11}^{-1/2}$, where Ω_{11} is the variance of $W_1(1)$, see (2.6), one obtains (4.1). In order to prove (ii), set $\lambda = \rho/T$ and consider

$$B_T' S(\rho/T) B_T = \begin{pmatrix} T^{-1}\bar{\psi}' S(\rho/T)\bar{\psi} & \bar{\psi}' S(\rho/T)\beta \\ \beta' S(\rho/T)\bar{\psi} & T\beta' S(\rho/T)\beta \end{pmatrix}, \qquad B_T := (T^{-\frac{1}{2}}\bar{\psi}, T^{1/2}\beta);$$

using Lemma A.1, one has

$$T^{-1}\bar{\psi}' S(\rho/T)\bar{\psi} = \rho T^{-2}\bar{\psi}' M_{xx}\bar{\psi} - T^{-\frac{1}{2}}\bar{\psi}' M_{xd} M_{dd}^{-1} M_{dx}\bar{\psi} T^{-\frac{1}{2}} \xrightarrow{w} -\ell_{1\varphi}\ell_{\varphi 1},$$

$$\bar{\psi}' S(\rho/T)\beta = \rho T^{-1}\bar{\psi}' M_{xx}\beta - T^{-\frac{1}{2}}\bar{\psi}' M_{xd} M_{dd}^{-1} M_{dx}\beta T^{1/2} \xrightarrow{w} -\ell_{1\varphi}\ell^{\circ}_{\varphi \mathrm{d}2},$$

$$T\beta' S(\rho/T)\beta = \rho\beta' M_{xx}\beta - T^{1/2}\beta' M_{xd} M_{dd}^{-1} M_{dx}\beta T^{1/2} \xrightarrow{w} \rho Q_{00} - \ell^{\circ}_{\mathrm{d}2\varphi}\ell^{\circ}_{\varphi \mathrm{d}2}$$

jointly, where $\ell^{\circ}_{\varphi \mathrm{d}2} = (\ell^{\circ}_{\mathrm{d}2\varphi})'$, and hence

$$|B_T' S(\rho/T) B_T| = \left|T^{-1}\bar{\psi}' S(\rho/T)\bar{\psi}\right| \left|T\beta' S(\rho/T)\beta - \beta' S(\rho/T)\bar{\psi}(T^{-1}\bar{\psi}' S(\rho/T)\bar{\psi})^{-1}\bar{\psi}' S(\rho/T)\beta\right|$$
$$\xrightarrow{w} \left|-\ell_{1\varphi}\ell_{\varphi 1}\right| \left|\rho Q_{00} - \ell^{\circ}_{\mathrm{d}2\varphi}\ell^{\circ}_{\varphi \mathrm{d}2} + \ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1}\ell_{1\varphi}\ell^{\circ}_{\varphi \mathrm{d}2}\right|.$$

Because $\ell_{1\varphi}\ell_{\varphi 1}$ is a.s. nonsingular, this shows that the r smallest eigenvalues in (3.2) multiplied by T converge weakly to the ones of $Q_{00}^{-1}\ell^{\circ}_{\mathrm{d}2\varphi}(I_K - \ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1}\ell_{1\varphi})\ell^{\circ}_{\varphi \mathrm{d}2}$, which coincide with those of the symmetric and a.s. positive definite $r \times r$ matrix $Q_{00}^{-1/2}\ell^{\circ}_{\mathrm{d}2\varphi}(I_K - \ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1}\ell_{1\varphi})\ell^{\circ}_{\varphi \mathrm{d}2} Q_{00}^{-1/2}$. Observe that the eigenvalues of $\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1}\ell_{1\varphi}$ are invariant with respect to the transformation of $W_1(u)$ into $AW_1(u)$ for any nonsingular $A \in \mathbb{R}^{s\times s}$; choosing $A = \Omega_{11}^{-1/2}$, one obtains (4.2).

(iii) It is next shown that $\Upsilon_K := \left(\int B_1 B_1'\right)^{-1} \int B_1 \varphi_K' \int \varphi_K B_1' \xrightarrow{a.s.} I_s$ as $K \to \infty$. Indeed, for the standard Brownian motion B_1 one has, see (2.8), $\int B_1 B_1' = \sum_{k=1}^{\infty} \nu_k^2 \xi_k \xi_k'$ a.s. by Parseval Theorem and $\int B_1 \varphi_K' = (\nu_1\xi_1, \nu_2\xi_2, \ldots, \nu_K\xi_K)$ by definition; this implies

$$\Upsilon_K = \left(\sum_{k=1}^{\infty} \nu_k^2 \xi_k \xi_k'\right)^{-1} \sum_{k=1}^{K} \nu_k^2 \xi_k \xi_k' \xrightarrow{a.s.} I_s \tag{B.2}$$

as $K \to \infty$. This further implies that the eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_s$ of Υ_K all converge to one a.s. as $K \to \infty$. Consider next (4.2) and observe that

$$K^{-1}\Psi_K = K^{-1} Q_{00}^{-1}(\ell^{\circ}_{\mathrm{d}2\varphi}\ell^{\circ}_{\varphi \mathrm{d}2} - \ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1}\ell_{1\varphi}\ell^{\circ}_{\varphi \mathrm{d}2});$$

from Lemma A.5.(ii) one has $K^{-1}\ell^{\circ}_{\mathrm{d}2\varphi}\ell^{\circ}_{\varphi \mathrm{d}2} \xrightarrow{a.s.} \Omega_{22}$ and $K^{-1}\ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1}\ell_{1\varphi}\ell^{\circ}_{\varphi \mathrm{d}2} = O_p(K^{-1})$ by Lemma A.5.(i),(iii). This shows that $K^{-1}\Psi_K \xrightarrow{p} Q_{00}^{-1}\Omega_{22} \overset{a.s.}{>} 0$. (iv) By (iii), the s largest eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s$ all converge weakly to one (and hence also in probability) as $(T,K)_{seq} \to \infty$, while $(\lambda_{s+1},\ldots,\lambda_p)$ converge to 0 in probability with $K^{-1}T(\lambda_{s+1},\ldots,\lambda_p) = O_p(1)$ as $(T,K)_{seq} \to \infty$. ■

Proof of Corollary 4.2. Direct consequence of Theorem 4.1.(iv). ■

Proof of Corollary 4.3. Note that $K/T = o(1)$ as $(T,K)_{seq} \to \infty$. Use Theorem 4.1.(i) and (4.3) to find $f_1(i) \overset{p}{\asymp} (K/T)^{|i-s|}$ for $i \in \mathcal{I}_1$, $f_2(i) \overset{p}{\asymp} 1_{i\neq s} + (T/K)1_{i=s}$ for $i \in \mathcal{I}_2$, $f_3(i) \overset{p}{\asymp} 1_{i\notin\{s-1,s\}} + 1_{i=s-1,s>1}/\log(T/K) + \log(T/K)1_{i=s}$ for $i \in \mathcal{I}_3$. Hence asymptotically: $f_1(i)$ converges to 0 for all $i \neq s$ while it converges to $1/\det(Q_{00}^{-1}\Omega_{22}) > 0$ for $i = s$; hence it has a unique maximum at $i = s$; $f_2(i)$ and $f_3(i)$ converge to finite values for $i \neq s$ and diverge for $i = s$, and hence have a unique maximum at $i = s$. This proves $\operatorname{argmax}_{i\in\mathcal{I}_j} f_j(i) \xrightarrow{p} s$ for $j = 1,2,3$ when $s \in \mathcal{I}_j$. When considering maximization over $\mathcal{I}_j^0$ for $j = 2,3$, one still has $f_2(i) \overset{p}{\asymp} 1_{i\neq s}+(T/K)1_{i=s}$ for $i \in \mathcal{I}_2^0$, $f_3(i) \overset{p}{\asymp} 1_{i\notin\{s-1,s\}}+1_{i=s-1,s>0}/\log(T/K)+\log(T/K)1_{i=s}$ for $i \in \mathcal{I}_3^0$, and, as above, this implies $\operatorname{argmax}_{i\in\mathcal{I}_j^0} f_j(i) \xrightarrow{p} s$ for $j = 2,3$ when $s \in \mathcal{I}_j^0$. All of the above is still valid when K is kept constant. ■

Proof of Theorem 4.4. By dimensional coherence, see Theorem 2.4, Corollary 4.2 applies to both $\widehat{s}(x_t)$ and $\widehat{s}(b'x_t)$, the max-gap estimator based on x_t and on $b'x_t$; hence $\mathbb{P}(\widehat{s}(b'x_t) = s = \widehat{s}(x_t)) \to 1$ as $(T,K)_{seq} \to \infty$. The case when H_0 is invalid is similar, with $\mathbb{P}(\widehat{s}(x_t) = s) \to 1$ and $\mathbb{P}(\widehat{s}(b'x_t) = \operatorname{rank}(b'\psi) < s) \to 1$, and hence $\mathbb{P}(\widehat{s}(b'x_t) < \widehat{s}(x_t)) \to 1$. This completes the proof. ■

Proof of Theorem 4.5. Consider Υ_K in (4.1); from (B.2) one has

$$K\left(I_s - \Upsilon_K\right) = \left(\int B_1 B_1'\right)^{-1} K \sum_{k=K+1}^{\infty} \nu_k^2 \xi_k \xi_k'.$$

It is next shown that $K\sum_{k=K+1}^{\infty} \nu_k^2 \xi_k \xi_k' \xrightarrow{p} \frac{1}{\pi^2} I_s$ as $K \to \infty$. Indeed,

$$K \sum_{k=K+1}^{\infty} \nu_k^2 \xi_k \xi_k' = I_s K \sum_{k=K+1}^{\infty} \nu_k^2 + K \sum_{k=K+1}^{\infty} \nu_k^2 \left(\xi_k \xi_k' - I_s\right)$$

and for the Karhunen-Loève representation, see (2.9), $\nu_k = \frac{1}{(k-\frac{1}{2})\pi}$; hence

$$\pi^2 K \sum_{k=K+1}^{\infty} \nu_k^2 = 4K \sum_{k=K+1}^{\infty} \frac{1}{(2k-1)^2} = \frac{1}{K} \sum_{k=1}^{\infty} \frac{1}{(1+\frac{k}{K}-\frac{1}{2K})^2} = \frac{1}{K} \sum_{k=1}^{\infty} \frac{1}{(1+\frac{k}{K})^2} + o(1) \tag{B.3}$$

$$\to \int_0^\infty \frac{dx}{(1+x)^2} = 1$$

by dominated convergence, whereas $\mathbb{E}\left(\sum_{k=K+1}^{\infty} \nu_k^2 (\xi_k \xi_k' - I_s)\right) = 0$ and, for every fixed $a, b \in \mathbb{R}^s$,

$$\mathbb{E}\left(\sum_{k=K+1}^{\infty} \nu_k^2 (a'\xi_k\xi_k' b - a'b)\right)^2 = \sum_{k=K+1}^{\infty} \nu_k^4 \mathbb{E}(a'\xi_k\xi_k' b - a'b)^2 = [a'ab'b + 2(a'b)^2] \sum_{k=K+1}^{\infty} \nu_k^4 = O(K^{-3}). \tag{B.4}$$

Thus $\pi^2 K \sum_{k=K+1}^{\infty} \nu_k^2 \xi_k \xi_k' \xrightarrow{p} I_s$, which implies that $\pi^2 K\left(I_s - \Upsilon_K\right) \xrightarrow{p} \left(\int B_1 B_1'\right)^{-1}$. Because $\lambda_i \xrightarrow{w} \mu_i$, $i = 1, \ldots, s$, by Theorem 4.1.(i), where $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_s$ are the eigenvalues of Υ_K, and the eigenvalues of $I_s - \Upsilon_K$ are $1 - \mu_i$, $i = 1, \ldots, s$, this proves (i). In order to prove (ii), recall that by Theorem 4.1.(iv), one has $\lambda_i \xrightarrow{p} 0$, $i = s+1, \ldots, p$, as $T \to \infty$. ∎

Proof of Corollary 4.8. Same as in the proof of Theorem 12.3 in Johansen (1996). ∎

Lemma B.1 (Super-consistency). *Let d_t in (3.1) be constructed using any orthonormal càdlàg basis of $L^2[0,1]$; then the one-step estimators $\widehat{\psi}^{(1)}, \widehat{\beta}^{(1)}$ and the ICC estimators $\widehat{\psi}^{(2)} := \widehat{\psi}, \widehat{\beta}^{(2)} := \widehat{\beta}$ of ψ and β are T-consistent, i.e.*

$$\widehat{\psi}^{(j)} - \psi = o_p(T^{-\frac{1}{2}}), \qquad \widehat{\beta}^{(j)} - \beta = o_p(T^{-\frac{1}{2}}), \qquad j = 1, 2, \tag{B.5}$$

for any fixed $K \geq p$.

Thanks to (3.7), the results in (B.5) can be equivalently expressed as

$$\widehat{\psi}_*^{(j)} - \psi_* = -(\widehat{\beta}_*^{(j)} - \beta_*)' = o_p(T^{-\frac{1}{2}}), \qquad j = 1, 2.$$

Proof of Lemma B.1. By the duality in (3.7), it is sufficient to prove the lemma for the estimators of ψ. Let $\Lambda = \operatorname{diag}(\Lambda_1, \Lambda_0)$ and $V = (V_1, V_0)$ in (3.3) satisfy

$$M_{ff} V \Lambda = M_{fd} M_{dd}^{-1} M_{df} V, \qquad V' M_{ff} V = I_p, \qquad V' M_{fd} M_{dd}^{-1} M_{df} V = \Lambda, \tag{B.6}$$

where $f = x$ for $j = 1$ and $f = e$ for $j = 2$. Consider the first equation in (B.6) for the s largest eigenvalues,

$$M_{ff} V_1 \Lambda_1 = M_{fd} M_{dd}^{-1} M_{df} V_1, \qquad f = x, e, \tag{B.7}$$

with associated normalization $V_1' M_{ff} V_1 = I_s$. Then one has

$$A_T' M_{ff} A_T (A_T^{-1} V_1) \Lambda_1 = A_T' M_{fd} M_{dd}^{-1} M_{df} A_T (A_T^{-1} V_1),$$

where $A_T := (T^{-\frac{1}{2}} \bar{\psi}, \beta)$ and $A_T^{-1} V_1 = (T^{1/2} \psi, \bar{\beta})' V_1 =: (v_{1T}', v_{0T}')'$; hence, by Lemma A.3 the limit eigenvalue problem is

$$\begin{pmatrix} \varsigma_{11} & 0 \\ 0 & Q_{00} \end{pmatrix} \begin{pmatrix} v_1^* \\ v_0^* \end{pmatrix} \Upsilon_K = \begin{pmatrix} \varsigma_{1\varphi} \varsigma_{\varphi 1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1^* \\ v_0^* \end{pmatrix}, \tag{B.8}$$

where $\Lambda_1 \xrightarrow{w} \Upsilon_K$ and $\varsigma_{11}, \varsigma_{1\varphi} \varsigma_{\varphi y}$ are a.s. nonsingular, with associated normalization of the eigenvectors $(v_1^{*\prime}, v_0^{*\prime}) \operatorname{diag}(\varsigma_{11}, Q_{00}) (v_1^{*\prime}, v_0^{*\prime})' = I_s$.

The lower block of (B.8) reads $Q_{00} v_0^* \Upsilon_K = 0$, which implies $v_{0T} \xrightarrow{p} v_0^* = 0$ because Q_{00} and Υ_K are a.s. nonsingular. The upper block of (B.8) reads $\varsigma_{11} v_1^* \Upsilon_K = \varsigma_{1\varphi} \varsigma_{\varphi 1} v_1^*$; hence it follows by the continuous mapping theorem that $v_{1T} \xrightarrow{w} v_1^*$, where v_1^* are eigenvectors of the limit eigenvalue problem $\det(\mu \varsigma_{11} - \varsigma_{1\varphi} \varsigma_{\varphi 1}) = 0$ such that $v_1^{*\prime} \varsigma_{11} v_1^* = I_s$. In particular, v_1^* is a.s. nonsingular. Summarizing, $A_T^{-1} V_1$ converges in distribution to $(v_1^{*\prime}, 0)'$, and

$$\begin{pmatrix} T^{-\frac{1}{2}} \bar{\psi}' M_{ff} V_1 \\ \beta' M_{ff} V_1 \end{pmatrix} = (A_T' M_{ff} A_T)(A_T^{-1} V_1) \xrightarrow{w} \begin{pmatrix} \varsigma_{11} & 0 \\ 0 & Q_{00} \end{pmatrix} \begin{pmatrix} v_1^* \\ 0 \end{pmatrix} = \begin{pmatrix} \vartheta_1 \\ 0 \end{pmatrix}, \tag{B.9}$$

where $\vartheta_1 := \varsigma_{11} v_1^*$ is a.s. nonsingular, and thus

$$T^{-\frac{1}{2}}\bar{\psi}' M_{ff} V_1 = O_p(1), \qquad (T^{-\frac{1}{2}}\bar{\psi}' M_{ff} V_1)^{-1} = O_p(1), \qquad \beta' M_{ff} V_1 = o_p(1), \tag{B.10}$$

$$T^{-\frac{1}{2}} b' M_{ff} V_1 = T^{-\frac{1}{2}} b' \psi \bar{\psi}' M_{ff} V_1 + b'\beta o_p(T^{-\frac{1}{2}}) = T^{-\frac{1}{2}}\bar{\psi}' M_{ff} V_1 + o_p(T^{-\frac{1}{2}}). \tag{B.11}$$

Hence using (B.10) and (B.11)

$$\bar{\psi}'\widehat{\psi}^{(j)} = T^{-\frac{1}{2}}\bar{\psi}' M_{ff} V_1 (T^{-\frac{1}{2}} b' M_{ff} V_1)^{-1} = I_s + o_p(T^{-\frac{1}{2}}), \tag{B.12}$$

$$\beta'\widehat{\psi}^{(j)} = T^{-\frac{1}{2}}\beta' M_{ff} V_1 (T^{-\frac{1}{2}} b' M_{ff} V_1)^{-1} = o_p(T^{-\frac{1}{2}}) O_p(1) = o_p(T^{-\frac{1}{2}}), \tag{B.13}$$

and one finds $\widehat{\psi}^{(j)} = \psi\bar{\psi}'\widehat{\psi}^{(j)} + \beta\bar{\beta}'\widehat{\psi}^{(j)} = \psi + o_p(T^{-\frac{1}{2}})$. This completes the proof. ∎

The following proofs employ a simplifying convention, without loss of generality. By Lemma *B.1*, $\bar{\psi}'\widehat{\psi}^{(1)}$ converges in probability to a positive-definite matrix and $\beta'\widehat{\psi}^{(1)} = o_p(T^{-\frac{1}{2}})$ as $T \to \infty$, such that the statements in Lemmas A.2(x)-(xii) and A.3 remain valid with $e_t(\widehat{\psi}^{(1)})$ in place of $e_t(\psi)$. Thus, the ICC estimators $\widehat{\psi}, \widehat{\beta}$ are discussed as if computed using $e_t(\psi)$, abbreviated to e_t.

Lemma B.2 (Fixed K asymptotic distribution of $\widehat{\psi}^{(j)}$ and $\widehat{\beta}^{(j)}$)**.** *Let $\widehat{\psi}^{(1)}$, $\widehat{\beta}^{(1)}$, $\widehat{\psi}^{(2)} := \widehat{\psi}$ and $\widehat{\beta}^{(2)} := \widehat{\beta}$ be as in Definition 3.2; then for fixed $K \geq p$ and $T \to \infty$, one has*

$$T\bar{c}'(\widehat{\psi}^{(j)} - \psi) = -T(\widehat{\beta}^{(j)} - \beta)'\bar{b} \xrightarrow{w} \begin{cases} \ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1} & \text{for } j = 1 \\ \ell^{\circ}_{\mathrm{d}2\varphi}P_{\circ}\ell_{\varphi 1}(\ell_{1\varphi}P_{\circ}\ell_{\varphi 1})^{-1} & \text{for } j = 2 \end{cases}, \tag{B.14}$$

where $P_{\circ} := I_K - \ell^{\circ}_{\varphi\mathrm{d}1}(\ell^{\circ}_{\mathrm{d}1\varphi}\ell^{\circ}_{\varphi\mathrm{d}1})^{-1}\ell^{\circ}_{\mathrm{d}1\varphi}$.

Proof of Lemma B.2. By the duality in (3.7), it is enough to prove the lemma for the estimators of ψ. Consider the estimator $\widehat{\psi}^{(j)} = M_{ff} V_1 (b' M_{ff} V_1)^{-1}$, where $f = x$ for $j = 1$ and $f = e$ for $j = 2$, and drop the superscript $^{(j)}$ in the rest of the proof. Write $\widehat{\psi} = \widetilde{\psi}\theta^{-1}$ where $\widetilde{\psi} := T^{-\frac{1}{2}} M_{ff} V_1$ and $\theta = T^{-\frac{1}{2}} b' M_{ff} V_1 = O_p(1)$, see (B.11). From (B.7) one has

$\widehat{\psi}(\theta\Lambda_1\theta^{-1}) = M_{fd}M_{dd}^{-1}M_{df}M_{ff}^{-1}\widehat{\psi}$; pre-multiplying by $(b,\beta)'$ and using $A_T = T^{-\frac{1}{2}}(\bar{\psi}, T^{1/2}\beta)$ in the r.h.s., one has

$$(b,\beta)'\widehat{\psi}(\theta\Lambda_1\theta^{-1}) = T^{-\frac{1}{2}}(b,\beta)'M_{fd}M_{dd}^{-1}(M_{df}A_T)(A_T'M_{ff}A_T)^{-1}(\bar{\psi}, T^{1/2}\beta)'\widehat{\psi}$$

and one finds

$$\begin{pmatrix} I_s \\ \beta'\widehat{\psi} \end{pmatrix} \theta\Lambda_1\theta^{-1} = T^{-\frac{1}{2}}(b,\beta)'M_{fd}Q_T, \tag{B.15}$$

where by Lemma A.3 and (B.12), (B.13) one has

$$Q_T := M_{dd}^{-1}(M_{df}A_T)(A_T'M_{ff}A_T)^{-1}(\bar{\psi}, T^{1/2}\beta)'\widehat{\psi} \xrightarrow{w} (\varsigma_{\varphi 1}, 0) \begin{pmatrix} \varsigma_{11} & 0 \\ 0 & Q_{00} \end{pmatrix}^{-1} \begin{pmatrix} I_s \\ 0 \end{pmatrix} = \varsigma_{\varphi 1}\varsigma_{11}^{-1}.$$

Because $b = \bar{\psi} + \beta h$, $h := \bar{\beta}'b$, and

$$T^{-\frac{1}{2}}b'M_{fd} = T^{-\frac{1}{2}}\bar{\psi}'M_{fd} + h'T^{-\frac{1}{2}}\beta'M_{fd} \xrightarrow{w} \varsigma_{1\varphi},$$

from the first block of s equalities in (B.15) one finds

$$\theta\Lambda_1\theta^{-1} = T^{-\frac{1}{2}}b'M_{fd}Q_T \xrightarrow{w} (\varsigma_{1\varphi}\varsigma_{\varphi 1})\varsigma_{11}^{-1},$$

which is a.s. nonsingular, and hence

$$G_T := Q_T(\theta\Lambda_1\theta^{-1})^{-1} \xrightarrow{w} \varsigma_{\varphi 1}(\varsigma_{1\varphi}\varsigma_{\varphi 1})^{-1} = \begin{cases} \ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1} & \text{for } j = 1 (f = x) \\ P_\circ\ell_{\varphi 1}(\ell_{1\varphi}P_\circ\ell_{\varphi 1})^{-1} & \text{for } j = 2 (f = e) \end{cases},$$

see (A.4). From the second block of r equalities in (B.15) one has

$$T\beta'\widehat{\psi} = (T^{\frac{1}{2}}\beta'M_{fd})G_T,$$

where $\beta'\widehat{\psi} = \beta'(\widehat{\psi} - \psi)$, and by Lemma A.1(vii), and Lemma A.2(xii),

$$T^{\frac{1}{2}}\beta'M_{fd} \xrightarrow{w} \varsigma_{\mathrm{d}2\varphi}^\circ := \begin{cases} \ell_{\mathrm{d}2\varphi}^\circ & \text{for } j = 1 (f = x) \\ \ell_{\mathrm{d}2\varphi}^\circ P_\circ & \text{for } j = 2 (f = e) \end{cases}.$$

This shows that

$$T\beta'(\widehat{\psi}^{(j)}-\psi) \xrightarrow{w} \varsigma^{\circ}_{\mathrm{d}2\varphi}\varsigma_{\varphi 1}(\varsigma_{1\varphi}\varsigma_{\varphi 1})^{-1} = \begin{cases} \ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1} & \text{for } j=1 \\ \ell^{\circ}_{\mathrm{d}2\varphi}P_{\circ}\ell_{\varphi 1}(\ell_{1\varphi}P_{\circ}\ell_{\varphi 1})^{-1} & \text{for } j=2 \end{cases};$$

the statement follows from $\bar{c}'(\widehat{\psi}^{(j)}-\psi) = \beta'(\widehat{\psi}^{(j)}-\psi)$ and the duality in (3.7). ■

Proof of Theorem 5.1. Consider (B.14) in Lemma *B*.2; by Lemma *A*.5(iv),(v)

$$\ell^{\circ}_{\mathrm{d}2\varphi}\ell_{\varphi 1}(\ell_{1\varphi}\ell_{\varphi 1})^{-1} \xrightarrow{p} \int \partial W_2 W_1' \left(\int W_1 W_1'\right)^{-1}$$
$$\ell^{\circ}_{\mathrm{d}2\varphi}P_{\circ}\ell_{\varphi 1}(\ell_{1\varphi}P_{\circ}\ell_{\varphi 1})^{-1} \xrightarrow{p} \int \mathrm{d}W_{2.1} W_1' \left(\int W_1 W_1'\right)^{-1}$$

and thus the statement. ■

Proof of Theorem 5.3. By Lemma *A*.1(viii),(ix), for fixed $K \geq p$ and $T \to \infty$, one has that

$$T^{1/2}\begin{pmatrix} \bar{\psi}' M_{\Delta x d} \\ \beta' M_{xd} \end{pmatrix} \xrightarrow{w} \begin{pmatrix} \int \mathrm{d}W_1 \varphi_K' \\ \int \mathrm{d}W_2 \varphi_K' \end{pmatrix} = \int \mathrm{d}W \varphi_K',$$

by Lemma *A*.2(ii), for fixed $K \geq p$ and $T \to \infty$, one has that $M_{dd} \to I_K$ and by Lemma *A*.5(ii), for $K \to \infty$, $K^{-1}\int \mathrm{d}W\varphi_K' \int \varphi_K \mathrm{d}W' \xrightarrow{a.s.} \Omega$. Combining these results, one finds that

$$\frac{T}{K}\begin{pmatrix} \bar{\psi}' M_{\Delta x d} \\ \beta' M_{xd} \end{pmatrix} M_{dd}^{-1} \begin{pmatrix} \bar{\psi}' M_{\Delta x d} \\ \beta' M_{xd} \end{pmatrix}' \underset{(T,K)_{seq}\to\infty}{\xrightarrow{w}} \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}.$$

Because Ω is non-stochastic, the last convergence is also in probability. By the super-consistency result in Theorem *B*.1, the same is achieved if $\widehat{\psi}^{(j)}$ and $\widehat{\beta}^{(j)}$ are used instead of ψ and β, $j = 1, 2$. ■

Proof of Theorem 5.6. Direct consequence of Theorem 5.1, Remark 5.2 and Theorem 5.3. ■

Proof of Theorem 5.7. Consider a VAR $A(L)X_t = \varepsilon_t$ under the $I(1)$ condition $|\alpha_\perp' A_1 \beta_\perp| \neq 0$, where $A(z) = A + A_1(z)(1-z)$, $A = -\alpha\beta'$ and $A_1 := A_1(1)$; in this case, Granger's Representation Theorem, see Johansen (1996, Theorem 4.2), $(1-z)A(z)^{-1} = C + C_1(z)(1-z)$

and hence $C(z) = C + C_1(z)(1-z)$ in (2.1) satisfies

$$C = \beta_\perp \Psi \alpha'_\perp, \qquad \beta' C_1 = \bar{\alpha}'(A_1 C - I_p), \qquad \begin{array}{c} C_1 := C_1(1) \\ \Psi := (\alpha'_\perp A_1 \beta_\perp)^{-1} \end{array}. \tag{B.16}$$

Note that $\psi = \beta_\perp$ and $\kappa' = \Psi\alpha'_\perp$.

Using the present notation for the parameters, Johansen (1996, Theorem 13.3) states $T\bar{b}'(\widehat{\beta}-\beta) \overset{w}{\to} (\int GG')^{-1} \int G \mathrm{d}V'_\alpha =: U$ where $G(u) := \bar{\beta}'_\perp C W_\varepsilon(u)$ and $V_\alpha(u) := (\alpha'\Omega_\varepsilon^{-1}\alpha)^{-1}\alpha'\Omega_\varepsilon^{-1} W_\varepsilon(u)$. It is next shown that $U = -Z'$, where Z is as in (5.2). Note that $G(u) = W_1(u)$; in fact $\bar{\beta}'_\perp C W_\varepsilon(u) = \kappa' W_\varepsilon(u) = W_1(u)$, so that $U = (\int W_1 W_1')^{-1} \int W_1 \mathrm{d}V'_\alpha$. Next it is shown that $V_\alpha = -W_{2.1}$; substituting $\psi = \beta_\perp$ and $\kappa' = \Psi\alpha'_\perp$ in (2.5) one finds that

$$W_1(u) := \kappa' W_\varepsilon(u) = \Psi\alpha'_\perp W_\varepsilon(u), \qquad W_2(u) := \beta' C_1 W_\varepsilon(u) = \bar{\alpha}'(A_1 C - I_p) W_\varepsilon(u)$$

whose variance matrix at $u = 1$ is

$$\Omega := \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} = \begin{pmatrix} \Psi\alpha'_\perp \Omega_\varepsilon \alpha_\perp \Psi' & \Psi\alpha'_\perp \Omega_\varepsilon (A_1 C - I_p)'\bar{\alpha} \\ \bar{\alpha}'(A_1 C - I_p)\Omega_\varepsilon \alpha_\perp \Psi' & \bar{\alpha}'(A_1 C - I_p)\Omega_\varepsilon (A_1 C - I_p)'\bar{\alpha} \end{pmatrix}.$$

It is useful to recall the following non-orthogonal projection identity

$$I_p = \Omega_\varepsilon \alpha_\perp (\alpha'_\perp \Omega_\varepsilon \alpha_\perp)^{-1} \alpha'_\perp + \alpha(\alpha'\Omega_\varepsilon^{-1}\alpha)^{-1}\alpha'\Omega_\varepsilon^{-1} =: P_1 + P_2. \tag{B.17}$$

One finds

$$\begin{aligned} \Omega_{21}\Omega_{11}^{-1} &= \bar{\alpha}'(A_1 C - I_p)\Omega_\varepsilon \alpha_\perp \Psi'(\Psi\alpha'_\perp \Omega_\varepsilon \alpha_\perp \Psi')^{-1} = \bar{\alpha}'(I_p - \Omega_\varepsilon \alpha_\perp (\alpha'_\perp \Omega_\varepsilon \alpha_\perp)^{-1}\alpha'_\perp) A_1 \beta_\perp \\ &= (\alpha'\Omega_\varepsilon^{-1}\alpha)^{-1}\alpha'\Omega_\varepsilon^{-1} A_1 \beta_\perp = \bar{\alpha}' P_2 A_1 \beta_\perp, \end{aligned}$$

where the second equality follows by substituting $C = \beta_\perp \Psi \alpha'_\perp$ and Ψ from (B.16) and the third one by the non-orthogonal projection identity (B.17). Substituting the above expressions into the definition of $W_{2.1}(u)$, see (2.7), one finds, using (B.17) again,

$$\begin{aligned} W_{2.1}(u) &:= W_2(u) - \Omega_{21}\Omega_{11}^{-1} W_1(u) = \bar{\alpha}'(A_1 C - I_p) W_\varepsilon(u) - \bar{\alpha}' P_2 A_1 C W_\varepsilon(u) \\ &= \bar{\alpha}'(-I_p + (I_p - P_2) A_1 C) W_\varepsilon(u) = \bar{\alpha}'(-P_2 + P_1(A_1 C - I_p)) W_\varepsilon(u), \end{aligned}$$

where $P_1(A_1C - I_p) = 0$ follows from $\alpha'_\perp(A_1C - I_p) = \alpha'_\perp A_1\beta_\perp(\alpha'_\perp A_1\beta_\perp)^{-1}\alpha'_\perp - \alpha'_\perp = 0$; hence

$$W_{2.1}(u) = -\bar{\alpha}' P_2 W_\varepsilon(u) = -(\alpha'\Omega_\varepsilon^{-1}\alpha)^{-1}\alpha'\Omega_\varepsilon^{-1}W_\varepsilon(u) = -V_\alpha(u).$$

This completes the proof that $U = -Z'$. ■

Proof of Theorem 5.8. If $C(z)$ in (2.1) satisfies Assumption 2.1, then $Q(z)$ satisfies the summability condition (**L**) in Phillips (2014), as e.g. in the discussion in Franchi and Paruolo (2020, page 809). $Q(1)$ nonsingular and it is normalized to I_p without loss of generality. From (5.8) one has

$$\begin{pmatrix} \Delta y_t \\ \Delta z_t \end{pmatrix} = \begin{pmatrix} Au_{1t} + \Delta u_{0t} \\ u_{1t} \end{pmatrix} = \begin{pmatrix} 0 & A \\ 0 & I_s \end{pmatrix}\begin{pmatrix} u_{rt} \\ u_{st} \end{pmatrix} + \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \Delta u_{0t} \\ \Delta u_{1t} \end{pmatrix}$$
$$= \left(\begin{pmatrix} 0 & A \\ 0 & I_s \end{pmatrix} + \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\Delta\right) Q(L)\varepsilon_t,$$

where $Q(z) = I_p + Q_1(z)(1-z)$ and $Q_1 := Q_1(1)$; hence $C(z) = C + C_1(z)(1-z)$ in (2.1) satisfies

$$C = \begin{pmatrix} 0 & A \\ 0 & I_s \end{pmatrix}, \qquad C_1 = \begin{pmatrix} 0 & A \\ 0 & I_s \end{pmatrix} Q_1 + \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

This shows that one can choose $\psi = (A', I_s)'$ with $\kappa' = (0, I_s)$ and $\beta' = (I_r, -A)$. Note also that $\beta' C_1 \kappa_\perp = (I_r, -A)C_1(I_r, 0)' = I_r$ of full row rank so that Assumption 2.1 holds. Substituting $\psi = (A', I_s)'$ with $\kappa' = (0, I_s)$ and $\beta' = (I_r, -A)$ in (2.2), from (2.5) one finds that

$$W_1(u) := \kappa' W_\varepsilon(u) = (0, I_s)W_\varepsilon(u), \qquad W_2(u) := \beta' C_1 W_\varepsilon(u) = (I_r, 0)W_\varepsilon(u).$$

Using the present notation for the parameters, Phillips (2014, point (a) in Section 3) reads $T(\widehat{\beta} - \beta)'\bar{b} \xrightarrow{w} -\int \mathrm{d}B_{0.x}B_x'(\int B_x B_x')^{-1} =: U$, where $B_x(u) := (0, I_s)W_\varepsilon(u) = W_1(u)$, $B_0(u) := (I_r, 0)W_\varepsilon(u) = W_2(u)$ and $B_{0.x}(u) := B_0(u) - \Omega_{0x}\Omega_{xx}^{-1}B_x(u) = W_{2.1}(u)$.

It is then shown that $U = -Z$, where Z is as in (5.2). Note that the variance matrix of $B_x(u) = W_1(u)$ and $B_0(u) = W_2(u)$ is

$$\begin{pmatrix} \Omega_{xx} & \Omega_{x0} \\ \Omega_{0x} & \Omega_{00} \end{pmatrix} = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix};$$

because $W_{2.1}(u) = B_{0.x}(u)$ and $W_1(u) = B_x(u)$, one can write Z in (5.2) as $Z = \int \mathrm{d}B_{0.x}B'_x \left(\int B_x B'_x\right)^{-1}$, which shows that $U = -Z$. ■

C Simulations

Results for the Monte Carlo simulations are split between inference on s and ψ.

C.1 Estimation and inference on s

First, the SCC profile $\{\lambda_i\}_{i=1,\dots,p}$ was simulated, to illustrate Theorem 4.1.(iv). Fig. *A*.1 reports quantiles 0.05 and 0.95 ((l, u), lower-upper) of the SCC profile λ_i for $a = 0.25, 1$, $p = 20, 100$ and $T = 30p, 100p$. Lines with the same colour refer to (l, u), colors (blue, red, orange, purple, green) correspond to $s = 0, \lceil p/4\rceil, \lceil p/2\rceil, \lceil 3p/4\rceil, p$, the first and second (third and fourth) row of panels correspond to $a = 0.25$ ($a = 1$), the first and third (second and fourth) row of panels correspond to $p = 20$ ($p = 100$), the first (second) column of panels correspond to $T = 30p$ ($T = 100p$). The eight panels illustrate how the curve $(\lambda_i)_{i=1,\dots,p}$ converges to $1_{i\leq s}$, see Theorem 4.1. Note that (l, u) lines get closer passing from the first column to the second one.

For fixed (p, T, s), the MC behavior of the max-gap estimator worsens as a decreases because of slower convergence to 0 of the last part of the SCC profile $\{\lambda_i\}_{i=s+1}^{p}$ for lower values of a. This can be seen by comparing the top 2 rows with the bottom 2 rows in Fig. *A*.1.

Next consider the finite sample properties of the max-gap estimator and of the sequential testing procedures $\{F_{j,n}\}_{j=p,\ldots,1}$, $n = 1, \infty$. Table A.1 reports MC statistics on the max-gap $\widehat{s}$. The top panel reports the MC frequency of correct selection of s, $\frac{1}{N}\sum_{j=1}^{N} 1_{\widehat{s}_j = s}$, where $\widehat{s}_j$ indicates the value of $\widehat{s}$ in MC replication j. The bottom panel reports the Mean Absolute Error (MAE), $\frac{1}{N}\sum_{j=1}^{N} |\widehat{s}_j - s|$. The number of replications of DGP (6.1) is $N = 10^4$.

When $a = 1$ or $s = p$, it can be seen that the max-gap estimator selects s with MC frequency 1 except for $T = 10p$ at $s = \lceil 3p/4 \rceil$, $p = 10, 20, 50$, and $T = 10p$, $p = 10$ at $s = \lceil p/2 \rceil$. For $a = 0.25, 0.5, 0.75$, the cases of less-than-perfect selection of the max-gap estimator increase, with worse performance for low values of a, and high values of s (excluding $s = p$). The MAE indicates that estimation errors in s are moderate except for $a = 0.25$ and large values of p.

Table A.2 reports MC statistics for the estimator $\widetilde{s}$ of s based on the test sequence $\{F_{j,\infty}\}_{j=p,\ldots,1}$. The top two panels are of the same type as the ones reported in Table A.1 for the max-gap estimator. The lower panel reports the MC size of test $F_{s,\infty}$ for the true value of s.

It can be seen that the test $F_{s,\infty}$ appears to be undersized in finite samples. The MC frequency of reaching the test $F_{s,\infty}$ in the test sequence is 1 for $T = 30p$ except for $a = 0.25$. At $a = 0.25$ this frequency is also 1 for $s = \lceil p/4 \rceil$ but deteriorates for higher values of s. Unlike for the max-gap, the predicted asymptotic behavior for $\{F_{j,\infty}\}_{j=p,\ldots,1}$ is in line with MC frequencies at $T = 30p$ for all values of a except for $a = 0.25$ and $s = \lceil jp/4 \rceil$, $j = 2, 3$.

Table A.3 reports MC statistics for estimator of s employing the test sequence $\{F_{j,1}\}_{j=p,\ldots,1}$, also indicated as $\widetilde{s}$ in the caption. The structure of the table is the same as for Table A.2. It can be seen that also the test $F_{s,1}$ is undersized in finite samples. Similarly to $\{F_{j,\infty}\}_{j=p,\ldots,1}$, the predicted asymptotic behavior for $\{F_{j,1}\}_{j=p,\ldots,1}$ is in line with MC frequencies at $T = 30p$

for all value of a except for $a = 0.25$ and $s = \lceil jp/4 \rceil$, $j = 2, 3$.

The max-gap criterion and the tests $F_{p,n}$ can be combined into a hybrid estimator of s, that first tests the hypothesis $H_0 : s = p$ using the $F_{p,n}$ statistic, and, in case of rejection, employs the restricted max-gap estimator that replaces the domain $\{0, 1, \dots p\}$ with $\{0, 1, \dots p - 1\}$; formally the hybrid estimator is defined as

$$\widehat{s}_n := 1_{\{F_{p,n} \leq c_{n,\eta}\}} \cdot p + 1_{\{F_{p,n} > c_{n,\eta}\}} \cdot \text{argmax}_{j \in \{0,1,\dots p-1\}} (\lambda_j - \lambda_{j+1}), \qquad n = 1, \infty. \tag{C.1}$$

Note that $\widehat{s}_n$ does not involve $\lambda_{p+1} := 0$. Under the conditions of Theorem 4.8 one finds that the hybrid estimator $\widehat{s}_n$ satisfies $\Pr(\widehat{s}_n = s) \to 1 - \eta$ if $s = p$ and $\Pr(\widehat{s}_n = s) \to 1$ if $s < p$, where η is the asymptotic significance level of the test $F_{p,n}$.

Monte Carlo results for the hybrid estimators $\widehat{s}_n$, $n = 1, \infty$, are reported in Tables A.4 and A.5. The performance of the hybrid estimator based on $F_{p,n}$ compares favorably with the one of the max-gap estimator for all values of a.

Tables A.6 and A.7 show the corresponding results for the cointegration rank determination procedures based on the trace test of Johansen (1996) with or without the Bartlett correction discussed in Johansen (2002) and implemented as in Johansen et al. (2005). For $p = 300$, the Bartlett correction could not be computed because a matrix exceeds the maximal dimension allowed in MATLAB, The MathWorks Inc. (2025); for $p = 50, 100$, the Bartlett correction was calculated for the first MC replication and it was kept fixed for the remaining ones. Critical values for the trace test were constructed as in Doornik (1998).

For $s = p$, Table A.8 reports results for the trace test corrected as in Onatski and Wang (2019). The results of the hybrid estimator with $F_{p,1}$ in Table A.5 compare favourably with the ones of these benchmarks.

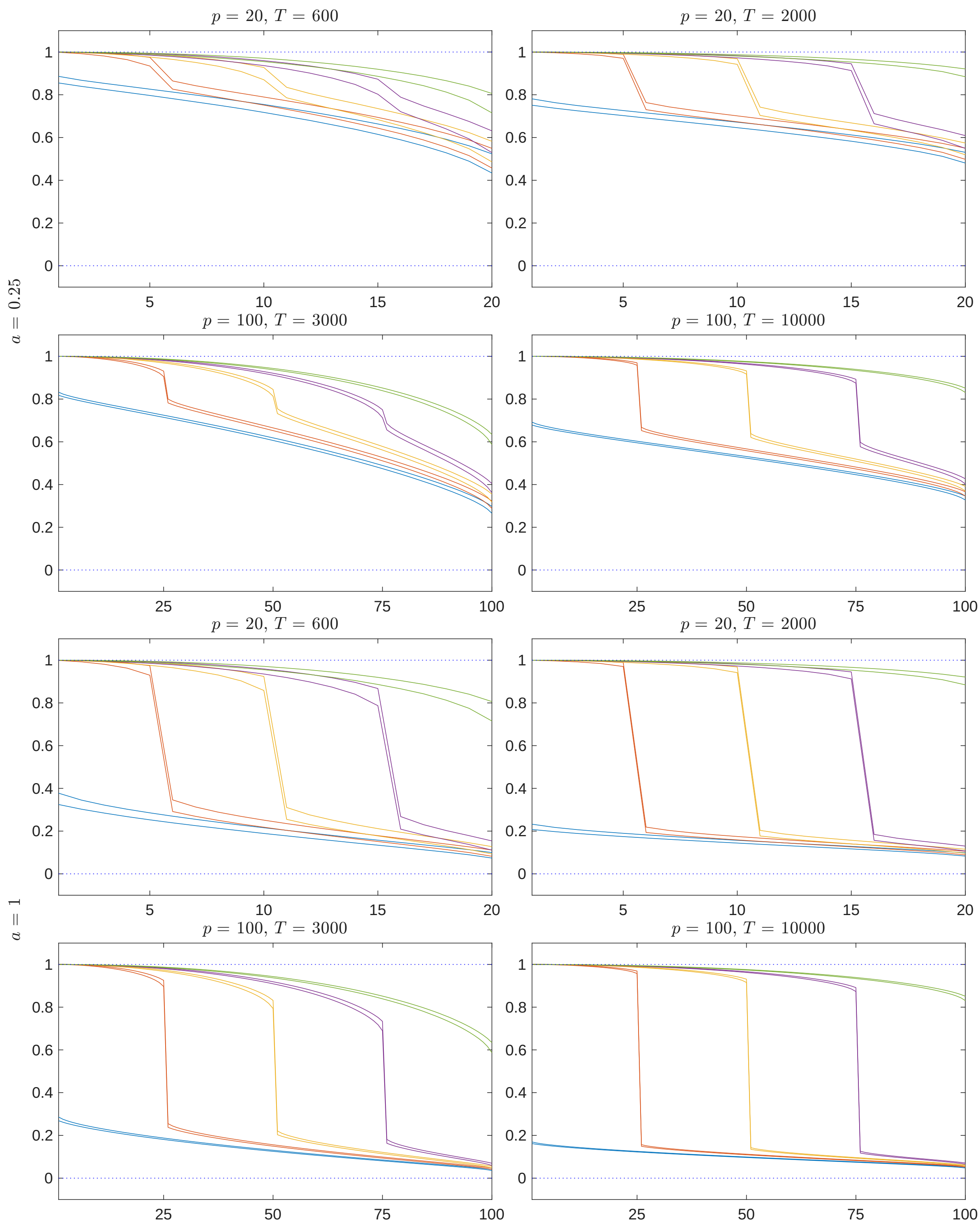

Figure A.1: Quantiles 0.05 and 0.95 of λ_i, $i = 1, \ldots, p$. $p = 20, 100$, $T/p = 30, 100$ and $a = 0.25$ (first and second panels from top) and $a = 1$ (bottom two panels) in DGP (6.1). Color codes for $s = \lceil jp/4 \rceil$, $j = 0, 1, 2, 3, 4$: blue $j = 0$, orange $j = 1$, yellow $j = 2$, purple $j = 3$, green $j = 4$. $N = 10^4$ replications of DGP (6.1).

Max-gap

Frequency of correct selection

		$s=0,\ a=$				$s=\lceil p/4\rceil,\ a=$				$s=\lceil p/2\rceil,\ a=$				$s=\lceil 3p/4\rceil,\ a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	0	0	0.96	1	0	0	0.80	1	0	0	0.59	0.99	0	0.01	0.34	0.93	1
	20	0	0.01	1	1	0	0.01	1	1	0	0	0.98	1	0	0.01	0.85	1	1
	30	0	0.06	1	1	0	0.03	1	1	0	0.02	1	1	0	0.02	0.99	1	1
20	10	0	0.17	1	1	0	0.02	1	1	0	0	0.90	1	0	0	0.55	0.98	1
	20	0	0.63	1	1	0	0.21	1	1	0	0.03	1	1	0	0	1	1	1
	30	0	0.98	1	1	0	0.75	1	1	0	0.28	1	1	0	0.07	1	1	1
50	10	0	1	1	1	0	0.58	1	1	0	0.01	0.99	1	0	0	0.78	0.99	1
	20	0	1	1	1	0	1	1	1	0	0.72	1	1	0	0.04	1	1	1
	30	0	1	1	1	0	1	1	1	0	1	1	1	0	0.87	1	1	1
100	10	0	1	1	1	0	0.99	1	1	0	0.04	1	1	0	0	0.86	1	1
	20	0	1	1	1	0	1	1	1	0	1	1	1	0	0.39	1	1	1
	30	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1
300	10	1	1	1	1	0	1	1	1	0	0.27	1	1	0	0	0.59	1	1
	20	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1
	30	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	1

MAE

		$s=0,\ a=$				$s=\lceil p/4\rceil,\ a=$				$s=\lceil p/2\rceil,\ a=$				$s=\lceil 3p/4\rceil,\ a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	9.99	9.93	0.35	0	6.99	6.96	1.30	0	4.99	4.96	1.92	0.01	2	1.96	1.26	0.10	0
	20	10	9.88	0	0	7	6.95	0.01	0	5	4.98	0.09	0	2	1.99	0.30	0	0
	30	10	9.42	0	0	7	6.78	0	0	5	4.92	0	0	2	1.96	0.01	0	0
20	10	20	16.64	0	0	15	14.69	0.03	0	10	9.98	0.81	0	5	4.99	2.13	0.02	0
	20	20	7.35	0	0	15	11.91	0	0	10	9.69	0	0	5	4.98	0.01	0	0
	30	20	0.41	0	0	15	3.70	0	0	10	7.23	0	0	5	4.66	0	0	0
50	10	50	0	0	0	37	15.51	0	0	25	24.75	0.01	0	12	12	2.05	0.01	0
	20	50	0	0	0	37	0.01	0	0	25	7.11	0	0	12	11.51	0	0	0
	30	50	0	0	0	37	0	0	0	25	0	0	0	12	1.56	0	0	0
100	10	99.99	0	0	0	75	0.18	0	0	50	47.63	0	0	25	25	1.39	0	0
	20	100	0	0	0	75	0	0	0	50	0.03	0	0	25	15.14	0	0	0
	30	100	0	0	0	75	0	0	0	50	0	0	0	25	0	0	0	0
300	10	0	0	0	0	225	0	0	0	150	45.35	0	0	75	74.98	1.70	0	0
	20	0	0	0	0	225	0	0	0	150	0	0	0	75	0.14	0	0	0
	30	0	0	0	0	225	0	0	0	150	0	0	0	75	0	0	0	0

Table A.1: Max-gap. Top panel: frequency of correct selection of s, $\frac{1}{N}\sum_{n=1}^{N} 1_{\widehat{s}_n=s}$; bottom panel: Mean Absolute Error (MAE), $\frac{1}{N}\sum_{n=1}^{N}|\widehat{s}_n - s|$. $N=10^4$ replications of DGP (6.1).

$F_{j,\infty}$

Frequency of correct selection

p	T/p	$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
		0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	0	0.81	1	1	0	0.05	0.98	0.98	0	0.01	0.88	0.98	0	0.07	0.76	0.98	0.98
	20	0.72	1	1	1	0.04	0.98	0.97	0.97	0.01	0.98	0.97	0.97	0.08	0.97	0.97	0.97	0.96
	30	1	1	1	1	0.97	0.97	0.96	0.96	0.77	0.97	0.97	0.97	0.68	0.97	0.97	0.97	0.96
20	10	0	1	1	1	0	0.34	0.99	0.98	0	0	0.98	0.98	0	0	0.77	0.99	0.99
	20	1	1	1	1	0.27	0.98	0.98	0.97	0	0.98	0.98	0.97	0	0.98	0.97	0.97	0.97
	30	1	1	1	1	0.99	0.98	0.97	0.97	0.96	0.97	0.97	0.97	0.72	0.97	0.97	0.97	0.96
50	10	0	1	1	1	0	0.84	1	0.99	0	0	1	1	0	0	0.22	1	1
	20	1	1	1	1	0.91	0.99	0.98	0.98	0	0.99	0.99	0.98	0	0.99	0.99	0.98	0.99
	30	1	1	1	1	0.99	0.98	0.98	0.97	1	0.99	0.98	0.98	0.76	0.98	0.98	0.97	0.98
100	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	20	1	1	1	1	1	1	0.99	0.99	0	1	1	0.99	0	1	1	1	1
	30	1	1	1	1	1	0.99	0.98	0.98	1	1	0.99	0.99	0.35	1	1	1	1
300	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	20	0	0.55	1	1	0	0	1	1	0	0	1	1	0	0	0.12	1	1
	30	1	1	1	1	1	1	1	1	0.98	1	1	1	0	1	1	1	1

MAE

p	T/p	$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
		0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	9.61	1.20	0	0	6.76	3.68	0.02	0.02	4.84	3.18	0.14	0.02	1.93	1.46	0.26	0.02	0.02
	20	1.87	0	0	0	3.91	0.02	0.03	0.03	3.20	0.02	0.03	0.03	1.41	0.03	0.03	0.03	0.04
	30	0	0	0	0	0.03	0.03	0.04	0.04	0.29	0.03	0.03	0.03	0.35	0.03	0.03	0.03	0.04
20	10	19.39	0.01	0	0	14.64	3.31	0.01	0.02	9.82	5.86	0.02	0.02	4.93	3.59	0.25	0.01	0.01
	20	0.01	0	0	0	4.08	0.02	0.02	0.03	5.62	0.02	0.02	0.03	3.20	0.02	0.03	0.03	0.03
	30	0	0	0	0	0.01	0.02	0.03	0.03	0.04	0.03	0.03	0.04	0.31	0.03	0.03	0.03	0.04
50	10	49.92	0	0	0	36.98	1.14	0	0.01	25	19.84	0	0	12	11.47	1.41	0	0
	20	0	0	0	0	0.39	0.01	0.02	0.02	12.98	0.01	0.01	0.02	7.80	0.01	0.01	0.02	0.01
	30	0	0	0	0	0.01	0.02	0.02	0.03	0	0.01	0.02	0.02	0.25	0.02	0.02	0.03	0.02
100	10	100	100	100	100	75	75	75	75	50	50	50	50	25	25	25	25	0
	20	0	0	0	0	0	0	0.01	0.01	31.38	0	0	0.01	19.63	0	0	0	0
	30	0	0	0	0	0	0.01	0.02	0.02	0	0	0.01	0.01	0.85	0	0	0	0
300	10	300	300	300	300	225	225	225	225	150	150	150	150	75	75	75	75	0
	20	300	134.84	0	0	225	225	0	0	150	150	0	0	75	75	65.67	0	0
	30	0	0	0	0	0	0	0	0	0.03	0	0	0	59.35	0	0	0	0

Empirical size

p	T/p	$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
		0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10					0	0	0.01	0.02	0	0	0.01	0.02	0	0	0.01	0.02	0.02
	20					0.01	0.02	0.03	0.03	0	0.02	0.03	0.03	0	0.02	0.03	0.03	0.04
	30					0.02	0.03	0.04	0.04	0.01	0.03	0.03	0.03	0.01	0.03	0.03	0.03	0.04
20	10					0	0	0.01	0.02	0	0	0.01	0.02	0	0	0	0.01	0.01
	20					0	0.02	0.02	0.03	0	0.02	0.02	0.03	0	0.02	0.03	0.03	0.03
	30					0.01	0.02	0.03	0.03	0.01	0.03	0.03	0.03	0.01	0.03	0.03	0.03	0.04
50	10					0	0	0	0.01	0	0	0	0	0	0	0	0	0
	20					0	0.01	0.02	0.02	0	0.01	0.01	0.02	0	0.01	0.01	0.02	0.01
	30					0.01	0.02	0.02	0.03	0	0.01	0.02	0.02	0	0.02	0.02	0.03	0.02
100	10					0	0	0	0	0	0	0	0	0	0	0	0	0
	20					0	0	0.01	0.01	0	0	0	0.01	0	0	0	0	0
	30					0	0.01	0.02	0.02	0	0	0.01	0.01	0	0	0	0	0
300	10					0	0	0	0	0	0	0	0	0	0	0	0	0
	20					0	0	0	0	0	0	0	0	0	0	0	0	0
	30					0	0	0	0	0	0	0	0	0	0	0	0	0

Table A.2: $F_{j,\infty}$. Top panel: frequency of correct selection of s, $\frac{1}{N}\sum_{j=1}^{N} 1_{\widetilde{s}_j=s}$, with test sequence at 5% significance level; second from top panel: Mean Absolute Error (MAE), $\frac{1}{N}\sum_{j=1}^{N}|\widetilde{s}_j - s|$; third from top panel: MC size. Note that $F_{0,\infty}$ is undefined. $N = 10^4$ replications of DGP (6.1).

$F_{j,1}$																		
Frequency of correct selection																		
		$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	0.31	1	1	1	0	0.19	0.95	0.98	0	0.05	0.58	0.97	0	0.07	0.36	0.76	0.99
	20	1	1	1	1	0.22	0.98	0.97	0.97	0.06	0.96	0.98	0.98	0.09	0.82	0.98	0.98	0.98
	30	1	1	1	1	0.94	0.97	0.97	0.97	0.58	0.98	0.98	0.97	0.45	0.98	0.98	0.98	0.98
20	10	1	1	1	1	0	0.27	0.99	0.99	0	0.01	0.38	0.94	0	0	0.12	0.48	1
	20	1	1	1	1	0.36	0.99	0.98	0.98	0.01	0.96	0.99	0.99	0.01	0.68	0.99	0.99	0.99
	30	1	1	1	1	0.99	0.98	0.98	0.98	0.50	0.99	0.98	0.98	0.26	0.99	0.99	0.99	0.99
50	10	1	1	1	1	0	0.02	0.94	1	0	0	0.01	0.34	0	0	0	0.01	1
	20	1	1	1	1	0.08	1	0.99	0.99	0	0.62	1	1	0	0.09	0.72	0.97	1
	30	1	1	1	1	0.99	0.99	0.99	0.99	0.10	1	1	1	0.01	0.90	1	1	1
100	10	1	1	1	1	0	0	0.40	1	0	0	0	0	0	0	0	0	1
	20	1	1	1	1	0	1	1	1	0	0.01	0.75	1	0	0	0	0.07	1
	30	1	1	1	1	0.88	1	1	1	0	0.94	1	1	0	0.04	0.66	0.96	1
300	10	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1
	20	1	1	1	1	0	0	0.96	1	0	0	0	0	0	0	0	0	1
	30	1	1	1	1	0	1	1	1	0	0	0	0	0	0	0	0	1
MAE																		
		$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	2.57	0	0	0	5.06	0.97	0.05	0.02	4.18	1.48	0.42	0.03	1.91	1.28	0.65	0.24	0.01
	20	0	0	0	0	0.91	0.02	0.03	0.03	1.43	0.04	0.02	0.02	1.19	0.18	0.02	0.02	0.02
	30	0	0	0	0	0.06	0.03	0.03	0.03	0.42	0.02	0.02	0.03	0.55	0.02	0.02	0.02	0.02
20	10	0	0	0	0	7.90	0.74	0.01	0.01	7.28	1.97	0.62	0.06	4.54	2.16	0.98	0.52	0
	20	0	0	0	0	0.65	0.01	0.02	0.02	1.85	0.04	0.01	0.01	1.96	0.32	0.01	0.01	0.01
	30	0	0	0	0	0.01	0.02	0.02	0.02	0.50	0.01	0.01	0.02	0.78	0.01	0.01	0.01	0.01
50	10	0	0	0	0	14.58	1.18	0.06	0	16.08	3.72	1.27	0.66	10.48	4.63	2.31	1.45	0
	20	0	0	0	0	0.99	0	0.01	0.01	3.32	0.38	0	0	3.95	0.94	0.28	0.03	0
	30	0	0	0	0	0.01	0.01	0.01	0.01	0.96	0	0	0	1.56	0.10	0	0	0
100	10	0	0	0	0	25.23	2.01	0.60	0	32.03	7.28	2.83	1.58	22.49	10.11	5.40	3.66	0
	20	0	0	0	0	1.62	0	0	0	6.29	1.04	0.25	0	8.13	2.22	1.11	0.93	0
	30	0	0	0	0	0.12	0	0	0	1.98	0.06	0	0	3.32	0.96	0.34	0.04	0
300	10	0	0	0	0	90.99	6.92	2.64	1.21	126.50	31.28	13.61	8.54	75	61.70	38.74	29.26	0
	20	0	0	0	0	5.48	1	0.04	0	24.94	5.90	3.09	2.12	39.06	14.63	9.70	7.88	0
	30	0	0	0	0	1.79	0	0	0	8.72	2.67	1.57	1	18.11	7.20	5.05	4.17	0
Empirical size																		
		$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10					0	0	0.01	0.02	0	0	0.01	0.01	0	0	0.01	0.01	0.01
	20					0.01	0.02	0.03	0.03	0	0.01	0.02	0.02	0	0.01	0.02	0.02	0.02
	30					0.02	0.03	0.03	0.03	0.01	0.02	0.02	0.03	0.01	0.02	0.02	0.02	0.02
20	10					0	0	0.01	0.01	0	0	0	0	0	0	0	0	0
	20					0	0.01	0.02	0.02	0	0.01	0.01	0.01	0	0	0.01	0.01	0.01
	30					0.01	0.02	0.02	0.02	0	0.01	0.01	0.02	0	0.01	0.01	0.01	0.01
50	10					0	0	0	0	0	0	0	0	0	0	0	0	0
	20					0	0	0.01	0.01	0	0	0	0	0	0	0	0	0
	30					0	0.01	0.01	0.01	0	0	0	0	0	0	0	0	0
100	10					0	0	0	0	0	0	0	0	0	0	0	0	0
	20					0	0	0	0	0	0	0	0	0	0	0	0	0
	30					0	0	0	0	0	0	0	0	0	0	0	0	0
300	10					0	0	0	0	0	0	0	0	0	0	0	0	0
	20					0	0	0	0	0	0	0	0	0	0	0	0	0
	30					0	0	0	0	0	0	0	0	0	0	0	0	0

Table A.3: $F_{j,1}$. Top panel: frequency of correct selection of s, $\frac{1}{N}\sum_{n=1}^{N} 1_{\widetilde{s}_n=s}$, with test sequence at 5% significance level; second from top panel: Mean Absolute Error (MAE), $\frac{1}{N}\sum_{j=1}^{N} |\widetilde{s}_j - s|$; third from top panel: MC size. Note that $F_{0,1}$ is undefined. $N = 10^4$ replications of DGP (6.1).

Hybrid with $F_{p,\infty}$																		
Frequency of correct selection																		
		$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	0	0.65	1	1	0	0.40	0.96	1	0	0.26	0.87	0.99	0.01	0.18	0.75	0.96	0.98
	20	0.46	1	1	1	0.26	0.99	1	1	0.18	0.95	1	1	0.17	0.89	1	1	0.96
	30	0.88	1	1	1	0.76	1	1	1	0.63	1	1	1	0.62	0.99	1	1	0.96
20	10	0	1	1	1	0	0.83	1	1	0	0.40	0.95	1	0	0.20	0.84	0.98	0.99
	20	0.88	1	1	1	0.52	1	1	1	0.19	1	1	1	0.14	0.98	1	1	0.97
	30	1	1	1	1	0.96	1	1	1	0.82	1	1	1	0.64	1	1	1	0.96
50	10	0.03	1	1	1	0	0.98	1	1	0	0.55	0.99	1	0	0.06	0.88	0.99	1
	20	1	1	1	1	0.89	1	1	1	0.24	1	1	1	0.07	1	1	1	0.99
	30	1	1	1	1	1	1	1	1	0.98	1	1	1	0.83	1	1	1	0.98
100	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	20	1	1	1	1	1	1	1	1	0.33	1	1	1	0.03	1	1	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	0.94	1	1	1	1
300	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	20	0	0.62	1	1	0	0	1	1	0	0	1	1	0	0	0.25	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	0.99	1	1	1	1
MAE																		
		$s=0$, $a=$				$s=\lceil p/4\rceil$, $a=$				$s=\lceil p/2\rceil$, $a=$				$s=\lceil 3p/4\rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	9.52	2.74	0.01	0	6.71	2.82	0.09	0	4.81	2.39	0.22	0.01	1.92	1.36	0.27	0.05	0.03
	20	4.55	0	0	0	3.80	0.05	0	0	2.80	0.11	0	0	1.33	0.11	0	0	0.04
	30	1.03	0	0	0	1.27	0	0	0	1.15	0	0	0	0.41	0.01	0	0	0.04
20	10	19.07	0.04	0	0	14.46	1.54	0	0	9.74	3.19	0.05	0	4.92	2.40	0.21	0.02	0.01
	20	2.23	0	0	0	6.12	0	0	0	5.92	0.01	0	0	2.63	0.04	0	0	0.04
	30	0.07	0	0	0	0.49	0	0	0	1.27	0	0	0	0.97	0	0	0	0.04
50	10	48.22	0	0	0	36.96	0.04	0	0	25	2.48	0.01	0	12	8.96	0.15	0.01	0
	20	0	0	0	0	3.31	0	0	0	14.95	0	0	0	7.59	0	0	0	0.01
	30	0	0	0	0	0	0	0	0	0.29	0	0	0	0.99	0	0	0	0.02
100	10	100	100	100	100	75	75	75	75	50	50	50	50	25	25	25	25	0
	20	0	0	0	0	0.16	0	0	0	24.55	0	0	0	17.94	0	0	0	0
	30	0	0	0	0	0	0	0	0	0.01	0	0	0	0.44	0	0	0	0
300	10	300	300	300	300	225	225	225	225	150	150	150	150	75	75	75	75	0
	20	300	113.73	0	0	225	224.98	0	0	150	150	0	0	75	75	55.94	0	0
	30	0	0	0	0	0	0	0	0	0	0	0	0	0.02	0	0	0	0

Table A.4: Hybrid with $F_{p,\infty}$. Top panel: frequency of correct selection of s, $\frac{1}{N}\sum_{n=1}^{N} 1_{\widehat{s}_n=s}$; bottom panel: Mean Absolute Error (MAE), $\frac{1}{N}\sum_{n=1}^{N} |\widehat{s}_n - s|$. $N = 10^4$ replications of DGP (6.1).

Hybrid with $F_{p,1}$

Frequency of correct selection

		$s = 0,\ a =$				$s = \lceil p/4 \rceil,\ a =$				$s = \lceil p/2 \rceil,\ a =$				$s = \lceil 3p/4 \rceil,\ a =$				$s = p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	0	0.65	1	1	0.01	0.42	0.96	1	0.02	0.32	0.87	0.99	0.03	0.27	0.77	0.96	0.99
	20	0.46	1	1	1	0.28	0.99	1	1	0.22	0.95	1	1	0.27	0.89	1	1	0.98
	30	0.88	1	1	1	0.76	1	1	1	0.63	1	1	1	0.64	0.99	1	1	0.98
20	10	0.02	1	1	1	0	0.83	1	1	0	0.40	0.95	1	0.02	0.25	0.84	0.98	1
	20	0.88	1	1	1	0.52	1	1	1	0.19	1	1	1	0.15	0.98	1	1	0.99
	30	1	1	1	1	0.96	1	1	1	0.82	1	1	1	0.64	1	1	1	0.99
50	10	0.83	1	1	1	0	0.98	1	1	0	0.55	0.99	1	0.01	0.17	0.88	0.99	1
	20	1	1	1	1	0.89	1	1	1	0.24	1	1	1	0.07	1	1	1	1
	30	1	1	1	1	1	1	1	1	0.98	1	1	1	0.83	1	1	1	1
100	10	1	1	1	1	0	1	1	1	0	0.55	1	1	0.01	0.11	0.89	1	1
	20	1	1	1	1	1	1	1	1	0.33	1	1	1	0.03	1	1	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	0.94	1	1	1	1
300	10	1	1	1	1	0.01	1	1	1	0	0.33	1	1	0	0.02	0.59	1	1
	20	1	1	1	1	1	1	1	1	0.47	1	1	1	0.01	1	1	1	1
	30	1	1	1	1	1	1	1	1	1	1	1	1	0.99	1	1	1	1

MAE

		$s = 0,\ a =$				$s = \lceil p/4 \rceil,\ a =$				$s = \lceil p/2 \rceil,\ a =$				$s = \lceil 3p/4 \rceil,\ a =$				$s = p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	8.15	2.73	0.01	0	5.27	2.65	0.09	0	3.85	1.92	0.22	0.01	1.90	1.11	0.25	0.05	0.01
	20	4.52	0	0	0	3.65	0.05	0	0	2.42	0.11	0	0	1.03	0.11	0	0	0.03
	30	1.03	0	0	0	1.27	0	0	0	1.15	0	0	0	0.37	0.01	0	0	0.03
20	10	17.27	0.04	0	0	12.65	1.54	0	0	7.69	3.16	0.05	0	4.12	1.83	0.21	0.02	0
	20	2.23	0	0	0	6.12	0	0	0	5.91	0.01	0	0	2.42	0.04	0	0	0.01
	30	0.07	0	0	0	0.49	0	0	0	1.27	0	0	0	0.97	0	0	0	0.02
50	10	7.94	0	0	0	33	0.04	0	0	21.31	2.14	0.01	0	8.89	3.75	0.15	0.01	0
	20	0	0	0	0	3.31	0	0	0	14.95	0	0	0	7.59	0	0	0	0
	30	0	0	0	0	0	0	0	0	0.29	0	0	0	0.99	0	0	0	0
100	10	0	0	0	0	66.74	0	0	0	42.71	1	0	0	18.86	5.36	0.16	0	0
	20	0	0	0	0	0.16	0	0	0	24.55	0	0	0	17.94	0	0	0	0
	30	0	0	0	0	0	0	0	0	0.01	0	0	0	0.44	0	0	0	0
300	10	0	0	0	0	139.04	0	0	0	85.73	2.27	0	0	75	16.71	1.55	0	0
	20	0	0	0	0	0	0	0	0	30.16	0	0	0	56.60	0	0	0	0
	30	0	0	0	0	0	0	0	0	0	0	0	0	0.02	0	0	0	0

Table A.5: Hybrid with $F_{p,1}$. Top panel: frequency of correct selection of s, $\frac{1}{N}\sum_{n=1}^{N} 1_{\widehat{s}_n = s}$; bottom panel: Mean Absolute Error (MAE), $\frac{1}{N}\sum_{n=1}^{N} |\widehat{s}_n - s|$. $N = 10^4$ replications of DGP (6.1).

Johansen trace test on rank, asymptotic																			
Frequency of correct selection																			
		$s=0,\ a=$				$s=\lceil p/4\rceil,\ a=$				$s=\lceil p/2\rceil,\ a=$				$s=\lceil 3p/4\rceil,\ a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	0.51	1	1	1	0	0.26	0.90	0.95	0	0.10	0.60	0.93	0.03	0.18	0.52	0.84	0.85
	20	1	1	1	1	0.21	0.95	0.94	0.94	0.06	0.88	0.95	0.94	0.10	0.70	0.93	0.93	0.91
	30	1	1	1	1	0.84	0.95	0.95	0.95	0.40	0.95	0.95	0.94	0.31	0.93	0.94	0.94	0.92
20	10	0.99	1	1	1	0	0.46	0.95	0.95	0	0.07	0.63	0.92	0	0.11	0.53	0.81	0.64
	20	1	1	1	1	0.39	0.95	0.95	0.95	0.03	0.89	0.94	0.94	0.03	0.71	0.91	0.91	0.85
	30	1	1	1	1	0.95	0.95	0.95	0.95	0.39	0.95	0.94	0.94	0.23	0.93	0.93	0.93	0.89
50	10	1	1	1	1	0	0.42	0.95	0.93	0	0.03	0.67	0.86	0	0.09	0.54	0.52	0.06
	20	1	1	1	1	0.33	0.95	0.94	0.94	0	0.89	0.92	0.90	0	0.72	0.82	0.80	0.55
	30	1	1	1	1	0.95	0.95	0.95	0.94	0.30	0.94	0.93	0.92	0.16	0.90	0.88	0.87	0.75
100	10	1	1	1	1	0	0.40	0.94	0.92	0	0	0.71	0.69	0	0.07	0.24	0.07	0
	20	1	1	1	1	0.29	0.95	0.94	0.93	0	0.89	0.87	0.84	0	0.72	0.60	0.52	0.08
	30	1	1	1	1	0.96	0.95	0.94	0.94	0.15	0.93	0.89	0.88	0.08	0.81	0.74	0.71	0.36
MAE																		
		$s=0,\ a=$				$s=\lceil p/4\rceil,\ a=$				$s=\lceil p/2\rceil,\ a=$				$s=\lceil 3p/4\rceil,\ a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	3.60	0	0	0	5.92	0.99	0.10	0.06	4.44	1.65	0.41	0.08	1.75	1.13	0.49	0.17	0.18
	20	0	0	0	0	1.17	0.06	0.06	0.06	2.01	0.12	0.06	0.06	1.37	0.31	0.08	0.08	0.11
	30	0	0	0	0	0.17	0.05	0.06	0.06	0.65	0.06	0.06	0.06	0.79	0.07	0.07	0.07	0.08
20	10	0.09	0	0	0	10.87	0.56	0.05	0.05	8.30	1.88	0.38	0.09	4.17	1.65	0.51	0.22	0.47
	20	0	0	0	0	0.65	0.05	0.06	0.06	2.47	0.12	0.06	0.07	2.31	0.30	0.10	0.11	0.17
	30	0	0	0	0	0.06	0.06	0.06	0.06	0.65	0.06	0.07	0.07	0.96	0.08	0.08	0.09	0.12
50	10	0	0	0	0	22.30	0.61	0.05	0.07	18.50	2.41	0.34	0.17	8.52	1.91	0.56	0.67	2.20
	20	0	0	0	0	0.72	0.05	0.06	0.07	3.55	0.12	0.10	0.11	3.67	0.30	0.21	0.25	0.61
	30	0	0	0	0	0.05	0.06	0.06	0.06	0.78	0.07	0.08	0.09	1.20	0.12	0.14	0.15	0.31
100	10	0	0	0	0	38.26	0.62	0.06	0.09	34.89	3.31	0.31	0.39	16.26	2.25	1.31	2.08	6.01
	20	0	0	0	0	0.77	0.05	0.07	0.08	5.34	0.12	0.15	0.19	5.89	0.32	0.53	0.65	1.97
	30	0	0	0	0	0.04	0.06	0.07	0.07	1.05	0.08	0.12	0.14	1.54	0.22	0.32	0.36	0.96
Empirical size																		
		$s=0,\ a=$				$s=\lceil p/4\rceil,\ a=$				$s=\lceil p/2\rceil,\ a=$				$s=\lceil 3p/4\rceil,\ a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10					0	0.03	0.05	0.05	0	0.02	0.05	0.06	0	0.03	0.08	0.10	0.15
	20					0.03	0.05	0.06	0.06	0.01	0.05	0.05	0.06	0.02	0.06	0.07	0.07	0.09
	30					0.04	0.05	0.05	0.05	0.03	0.05	0.05	0.06	0.03	0.06	0.06	0.06	0.08
20	10					0	0.03	0.04	0.05	0	0.01	0.06	0.08	0	0.03	0.11	0.15	0.36
	20					0.02	0.05	0.05	0.05	0	0.05	0.06	0.06	0.01	0.07	0.09	0.09	0.15
	30					0.04	0.05	0.05	0.05	0.03	0.05	0.06	0.06	0.03	0.06	0.07	0.07	0.11
50	10					0	0.02	0.05	0.07	0	0	0.08	0.14	0	0.03	0.32	0.48	0.94
	20					0.01	0.05	0.06	0.06	0	0.05	0.08	0.10	0	0.11	0.18	0.20	0.45
	30					0.03	0.05	0.05	0.06	0.01	0.06	0.07	0.08	0.02	0.10	0.12	0.13	0.25
100	10					0	0.01	0.05	0.08	0	0	0.12	0.31	0	0.02	0.75	0.93	1
	20					0	0.05	0.06	0.07	0	0.05	0.13	0.16	0	0.20	0.40	0.48	0.92
	30					0.03	0.05	0.06	0.06	0	0.07	0.11	0.12	0.01	0.19	0.26	0.29	0.64

Table A.6: Johansen trace test on rank, asymptotic. Top panel: frequency of correct selection of s, $\frac{1}{N}\sum_{n=1}^{N} 1_{\widetilde{s}_n=s}$, with test sequence at 5% significance level; second from top panel: Mean Absolute Error (MAE), $\frac{1}{N}\sum_{j=1}^{N}|\widetilde{s}_j - s|$; third from top panel: MC size. $N = 10^4$ replications of DGP (6.1).

Johansen trace test on rank, Bartlett corrected																		
Frequency of correct selection																		
		$s=0$, $a=$				$s=\lceil p/4 \rceil$, $a=$				$s=\lceil p/2 \rceil$, $a=$				$s=\lceil 3p/4 \rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	0.27	1	1	1	0	0.18	0.89	0.95	0	0.04	0.48	0.93	0	0.06	0.33	0.77	0.95
	20	1	1	1	1	0.14	0.96	0.95	0.95	0.03	0.85	0.96	0.95	0.05	0.62	0.95	0.95	0.95
	30	1	1	1	1	0.80	0.96	0.95	0.95	0.33	0.96	0.95	0.95	0.24	0.94	0.95	0.95	0.95
20	10	0.88	1	1	1	0	0.32	0.96	0.96	0	0.01	0.42	0.93	0	0.01	0.22	0.74	0.95
	20	1	1	1	1	0.27	0.96	0.95	0.95	0.01	0.85	0.96	0.95	0	0.55	0.96	0.95	0.96
	30	1	1	1	1	0.95	0.95	0.95	0.95	0.27	0.96	0.95	0.95	0.12	0.95	0.96	0.95	0.95
50	10	1	1	1	1	0	0.18	0.97	0.96	0	0	0.26	0.93	0	0	0.07	0.64	0.97
	20	1	1	1	1	0.14	0.97	0.96	0.95	0	0.78	0.97	0.96	0	0.37	0.97	0.96	0.96
	30	1	1	1	1	0.95	0.96	0.96	0.95	0.10	0.97	0.96	0.95	0.02	0.95	0.96	0.96	0.96
100	10	1	1	1	1	0	0.08	0.98	0.97	0	0	0.05	0.88	0	0	0	0.44	0.98
	20	1	1	1	1	0.06	0.98	0.96	0.95	0	0.59	0.98	0.97	0	0.14	0.97	0.98	0.97
	30	1	1	1	1	0.95	0.97	0.96	0.95	0.01	0.98	0.97	0.96	0	0.94	0.98	0.97	0.96
MAE																		
		$s=0$, $a=$				$s=\lceil p/4 \rceil$, $a=$				$s=\lceil p/2 \rceil$, $a=$				$s=\lceil 3p/4 \rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10	5.96	0	0	0	6.49	1.29	0.12	0.05	4.79	2.20	0.54	0.07	1.92	1.51	0.73	0.23	0.05
	20	0	0	0	0	1.45	0.05	0.05	0.06	2.38	0.15	0.04	0.05	1.55	0.39	0.06	0.05	0.06
	30	0	0	0	0	0.20	0.05	0.05	0.05	0.76	0.05	0.05	0.05	0.90	0.06	0.05	0.05	0.05
20	10	1.92	0	0	0	13.13	0.75	0.04	0.04	9.51	2.91	0.60	0.07	4.86	2.83	0.95	0.26	0.05
	20	0	0	0	0	0.84	0.04	0.05	0.05	3.25	0.15	0.04	0.05	2.98	0.45	0.05	0.05	0.05
	30	0	0	0	0	0.05	0.05	0.05	0.05	0.83	0.04	0.05	0.05	1.25	0.05	0.05	0.05	0.05
50	10	0	0	0	0	30.35	0.95	0.03	0.04	23.01	4.66	0.81	0.07	11.71	5.35	1.44	0.37	0.03
	20	0	0	0	0	1.07	0.03	0.04	0.05	5.37	0.22	0.03	0.05	5.71	0.65	0.04	0.04	0.04
	30	0	0	0	0	0.05	0.04	0.05	0.05	1.19	0.03	0.04	0.05	2	0.05	0.04	0.04	0.04
100	10	0	0	0	0	57.93	1.19	0.02	0.03	45.08	7.79	1.26	0.12	24.17	9.47	2.27	0.57	0.02
	20	0	0	0	0	1.28	0.02	0.04	0.05	8.76	0.41	0.02	0.03	10.22	0.97	0.03	0.02	0.03
	30	0	0	0	0	0.05	0.03	0.04	0.05	1.86	0.02	0.03	0.04	3.19	0.06	0.03	0.03	0.04
Empirical size																		
		$s=0$, $a=$				$s=\lceil p/4 \rceil$, $a=$				$s=\lceil p/2 \rceil$, $a=$				$s=\lceil 3p/4 \rceil$, $a=$				$s=p$
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1	
10	10					0	0.02	0.04	0.04	0	0	0.03	0.04	0	0	0.02	0.04	0.05
	20					0.01	0.04	0.05	0.05	0	0.03	0.04	0.05	0.01	0.04	0.05	0.05	0.05
	30					0.03	0.04	0.05	0.05	0.02	0.04	0.05	0.05	0.02	0.04	0.05	0.05	0.05
20	10					0	0.01	0.03	0.04	0	0	0.02	0.03	0	0	0.02	0.03	0.05
	20					0.01	0.04	0.05	0.05	0	0.03	0.04	0.05	0	0.02	0.04	0.05	0.04
	30					0.03	0.04	0.05	0.05	0.01	0.04	0.05	0.05	0.01	0.03	0.04	0.05	0.05
50	10					0	0	0.02	0.04	0	0	0	0.02	0	0	0	0.02	0.03
	20					0	0.03	0.04	0.05	0	0.01	0.03	0.04	0	0.01	0.03	0.04	0.04
	30					0.02	0.04	0.04	0.05	0	0.03	0.04	0.05	0	0.02	0.04	0.04	0.04
100	10					0	0	0.01	0.03	0	0	0	0.01	0	0	0	0	0.02
	20					0	0.02	0.04	0.05	0	0	0.02	0.03	0	0	0.01	0.02	0.03
	30					0.01	0.03	0.04	0.05	0	0.02	0.03	0.04	0	0.01	0.02	0.03	0.04

Table A.7: Johansen trace test on rank, Bartlett corrected. Top panel: frequency of correct selection of s, $\frac{1}{N}\sum_{n=1}^{N} 1_{\widetilde{s}_n = s}$, with test sequence at 5% significance level; second from top panel: Mean Absolute Error (MAE), $\frac{1}{N}\sum_{j=1}^{N} |\widetilde{s}_j - s|$; third from top panel: MC size. $N = 10^4$ replications of DGP (6.1).

$s = p$, Onatski and Wang (2019) correction					
$T/p \setminus p$	10	20	50	100	300
10	0.06	0.05	0.05	0.05	0.05
20	0.05	0.05	0.05	0.05	0.05
30	0.05	0.05	0.05	0.05	0.05

Table A.8: Empirical size of Johansen trace test on rank $s = p$, corrected as in Onatski and Wang (2019). $N = 10^4$ replications of DGP (6.1).

C.2 Inference on ψ

For hypotheses tests on ψ, the estimation of the long-run variance matrix $\Omega_{22.1}$ was performed using (5.3) (LRV-P), as well as the estimator in Andrews (1991) and Andrews and Monahan (1992) with Parzen's kernel (LRV-A). The latter calculations were performed using the `hac.m` function in MATLAB, The MathWorks Inc. (2025); the bandwidth was calculated for the first MC replication and it was kept fixed for the remaining ones. The unfeasible solution LRV-U was computed substituting $\Omega_{22.1}$ with $a^{-2}I_r$, the value in the DGP.

Table *A*.9 reports the MC frequency of rejection under the two null hypotheses on ψ introduced in the main text of the paper, a scalar hypothesis tested by means of a t-test and an r-vector hypothesis assessed using a Wald Q test, over $N = 10^4$ replications of DGP (6.1) at the 5% nominal significance level. The top panel reports results for the t-test of the scalar hypothesis and the bottom panel for the Q test of the vector hypothesis.

The t-test is oversized in finite samples for all LRV estimators considered. The best performance is observed for the unfeasible LRV-U variant. The MC test size approaches the nominal level as a and T increase. Results for $p = 10$ and $p = 20$ are very similar. Tests employing the LRV-P and LRV-A variants are more oversized; their size also becomes closer to the nominal level for large T, especially so when a is large.

The Wald Q test is more pronouncedly oversized in finite samples than the corresponding t-test for all LRV choices. The MC sizes tend to be closer to the nominal level as a and T increase, though convergence is slower than for the t-test. Comparing values for $p = 10, 20$, the MC sizes deteriorate when passing from $p = 10$ to 20, corresponding to an increase in the number of scalar restrictions under test (i.e., the degrees of freedom of the test).

Table A.10 shows the results for the likelihood-ratio test on β of Johansen (1996) of the hypothesis (ii) that the first column of ψ_* equals 0, with or without the Bartlett correction discussed in Johansen (2000). The empirical size of these benchmarks appears closer to the nominal level than the one of the Q-test.

The need to estimate the LRV non-parametrically appears to be a critical factor in the relatively worse performance of the Q-test with LRV-A, LRV-P with respect to the Q-test with LRV-U and the Bartlett corrected LR tests in the VAR case. See Sun (2011, 2014), Lazarus et al. (2021) and references therein for recent contributions on LRV estimation.

Empirical size of t-test													
LRV-U		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.13	0.11	0.09	0.09	0.19	0.17	0.13	0.12	0.30	0.31	0.24	0.21
	20	0.09	0.07	0.07	0.06	0.15	0.10	0.09	0.09	0.26	0.14	0.12	0.11
	30	0.08	0.06	0.06	0.06	0.10	0.08	0.07	0.07	0.17	0.11	0.10	0.09
	60	0.06	0.06	0.06	0.05	0.07	0.07	0.07	0.07	0.09	0.08	0.07	0.07
	90	0.06	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.08	0.07	0.07	0.07
20	10	0.13	0.10	0.09	0.08	0.22	0.19	0.14	0.13	0.32	0.35	0.26	0.22
	20	0.08	0.07	0.06	0.06	0.16	0.11	0.10	0.09	0.27	0.14	0.13	0.12
	30	0.07	0.07	0.06	0.06	0.10	0.08	0.08	0.08	0.16	0.11	0.10	0.09
	60	0.06	0.06	0.06	0.06	0.08	0.07	0.07	0.07	0.09	0.07	0.07	0.07
	90	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.06	0.08	0.07	0.06	0.06
LRV-P		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.38	0.24	0.16	0.13	0.40	0.34	0.25	0.21	0.49	0.51	0.44	0.38
	20	0.31	0.15	0.10	0.08	0.39	0.21	0.15	0.12	0.50	0.30	0.22	0.18
	30	0.26	0.12	0.09	0.07	0.32	0.16	0.11	0.10	0.43	0.23	0.16	0.14
	60	0.20	0.10	0.07	0.06	0.23	0.12	0.09	0.08	0.28	0.14	0.11	0.10
	90	0.17	0.08	0.06	0.06	0.18	0.09	0.07	0.06	0.22	0.11	0.09	0.08
20	10	0.35	0.21	0.15	0.13	0.34	0.35	0.27	0.23	0.37	0.51	0.46	0.42
	20	0.27	0.13	0.10	0.09	0.39	0.21	0.16	0.14	0.48	0.30	0.23	0.21
	30	0.23	0.12	0.09	0.08	0.30	0.16	0.12	0.10	0.40	0.22	0.17	0.15
	60	0.17	0.09	0.07	0.06	0.21	0.11	0.09	0.08	0.24	0.13	0.10	0.09
	90	0.15	0.08	0.07	0.06	0.17	0.09	0.08	0.07	0.20	0.11	0.09	0.08
LRV-A		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.40	0.26	0.17	0.12	0.51	0.39	0.26	0.17	0.60	0.53	0.37	0.28
	20	0.27	0.15	0.11	0.08	0.39	0.22	0.15	0.11	0.62	0.35	0.23	0.15
	30	0.20	0.12	0.09	0.07	0.29	0.16	0.12	0.08	0.46	0.24	0.16	0.11
	60	0.13	0.09	0.08	0.06	0.18	0.11	0.09	0.07	0.25	0.14	0.11	0.08
	90	0.10	0.08	0.06	0.05	0.12	0.08	0.07	0.06	0.18	0.11	0.09	0.07
20	10	0.39	0.23	0.15	0.10	0.47	0.42	0.27	0.18	0.53	0.62	0.47	0.30
	20	0.24	0.13	0.10	0.07	0.43	0.24	0.17	0.11	0.58	0.35	0.23	0.15
	30	0.17	0.11	0.09	0.07	0.31	0.17	0.12	0.09	0.48	0.25	0.17	0.12
	60	0.11	0.08	0.07	0.06	0.18	0.12	0.09	0.07	0.24	0.14	0.10	0.08
	90	0.09	0.07	0.07	0.06	0.14	0.09	0.08	0.06	0.18	0.11	0.09	0.07

Empirical size of Q-test													
LRV-U		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.32	0.20	0.16	0.15	0.45	0.34	0.26	0.23	0.48	0.45	0.34	0.29
	20	0.17	0.11	0.10	0.09	0.29	0.16	0.13	0.12	0.37	0.19	0.15	0.14
	30	0.12	0.08	0.08	0.07	0.18	0.12	0.11	0.10	0.22	0.13	0.12	0.11
	60	0.08	0.07	0.06	0.06	0.10	0.08	0.07	0.07	0.12	0.09	0.09	0.08
	90	0.06	0.06	0.06	0.06	0.08	0.07	0.07	0.07	0.09	0.08	0.07	0.07
20	10	0.46	0.26	0.21	0.19	0.74	0.55	0.43	0.39	0.79	0.72	0.58	0.52
	20	0.21	0.13	0.12	0.11	0.45	0.24	0.20	0.19	0.60	0.30	0.24	0.22
	30	0.15	0.10	0.10	0.09	0.25	0.16	0.14	0.13	0.33	0.19	0.16	0.15
	60	0.08	0.07	0.07	0.07	0.13	0.10	0.09	0.09	0.14	0.11	0.10	0.10
	90	0.07	0.06	0.06	0.06	0.10	0.08	0.08	0.07	0.11	0.09	0.08	0.08
LRV-P		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.85	0.74	0.56	0.47	0.73	0.76	0.64	0.55	0.67	0.70	0.62	0.55
	20	0.84	0.48	0.30	0.24	0.84	0.51	0.34	0.28	0.70	0.44	0.32	0.27
	30	0.74	0.35	0.21	0.17	0.74	0.37	0.25	0.20	0.60	0.32	0.22	0.19
	60	0.56	0.22	0.14	0.11	0.53	0.22	0.15	0.12	0.40	0.19	0.14	0.12
	90	0.45	0.17	0.11	0.09	0.43	0.17	0.12	0.10	0.31	0.15	0.11	0.10
20	10	0.91	0.91	0.79	0.72	0.71	0.94	0.90	0.86	0.70	0.89	0.92	0.88
	20	0.96	0.65	0.47	0.39	0.95	0.74	0.58	0.51	0.88	0.70	0.56	0.50
	30	0.89	0.49	0.32	0.27	0.90	0.54	0.39	0.33	0.83	0.51	0.38	0.33
	60	0.71	0.29	0.18	0.15	0.70	0.32	0.21	0.18	0.57	0.27	0.20	0.17
	90	0.59	0.22	0.14	0.12	0.56	0.23	0.15	0.13	0.45	0.20	0.15	0.13
LRV-A		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.95	0.85	0.59	0.31	0.88	0.86	0.64	0.39	0.78	0.73	0.54	0.40
	20	0.90	0.56	0.32	0.17	0.89	0.58	0.35	0.21	0.82	0.52	0.33	0.21
	30	0.77	0.40	0.23	0.11	0.76	0.41	0.25	0.14	0.66	0.35	0.22	0.14
	60	0.45	0.22	0.14	0.08	0.47	0.23	0.15	0.10	0.37	0.20	0.14	0.10
	90	0.30	0.16	0.11	0.07	0.32	0.17	0.12	0.08	0.26	0.15	0.11	0.08
20	10	0.98	0.99	0.81	0.42	0.88	0.98	0.90	0.61	0.83	0.94	0.92	0.69
	20	0.99	0.79	0.49	0.21	0.98	0.84	0.57	0.31	0.95	0.80	0.54	0.33
	30	0.94	0.60	0.33	0.15	0.96	0.66	0.40	0.19	0.93	0.61	0.37	0.21
	60	0.65	0.32	0.18	0.09	0.70	0.37	0.21	0.12	0.62	0.32	0.20	0.12
	90	0.46	0.22	0.14	0.08	0.50	0.25	0.16	0.09	0.43	0.22	0.15	0.09

Table A.9: Frequency of rejection under the null of t and Q test on ψ over $N = 10^4$ replications of DGP (6.1) at 5% significance level. LRV-U: $\widehat{\Omega}$ equal to its true value, LRV-P: $\widehat{\Omega}$ as in (5.3), LRV-A: $\widehat{\Omega}$ as in Andrews (1991) and Andrews and Monahan (1992).

Empirical size of Johansen test on β													
Asymptotic		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.57	0.29	0.19	0.15	0.61	0.34	0.21	0.16	0.48	0.33	0.21	0.15
	20	0.28	0.14	0.10	0.08	0.34	0.16	0.11	0.09	0.33	0.15	0.11	0.09
	30	0.18	0.10	0.08	0.07	0.22	0.12	0.09	0.08	0.22	0.11	0.09	0.07
	60	0.11	0.07	0.07	0.06	0.11	0.08	0.07	0.06	0.12	0.08	0.07	0.07
	90	0.08	0.06	0.06	0.05	0.09	0.07	0.06	0.06	0.09	0.07	0.06	0.06
20	10	0.79	0.42	0.26	0.19	0.83	0.52	0.32	0.23	0.73	0.51	0.31	0.21
	20	0.40	0.19	0.13	0.10	0.51	0.22	0.14	0.11	0.51	0.21	0.13	0.11
	30	0.26	0.13	0.10	0.08	0.31	0.15	0.10	0.08	0.32	0.15	0.10	0.08
	60	0.13	0.08	0.07	0.06	0.15	0.09	0.08	0.07	0.15	0.09	0.07	0.06
	90	0.10	0.07	0.06	0.06	0.11	0.07	0.06	0.06	0.11	0.07	0.06	0.06
Bartlett corrected		$s = \lceil p/4 \rceil$, $a =$				$s = \lceil p/2 \rceil$, $a =$				$s = \lceil 3p/4 \rceil$, $a =$			
p	T/p	0.25	0.5	0.75	1	0.25	0.5	0.75	1	0.25	0.5	0.75	1
10	10	0.30	0.15	0.11	0.09	0.31	0.16	0.10	0.08	0.29	0.18	0.11	0.08
	20	0.14	0.09	0.07	0.06	0.16	0.08	0.06	0.06	0.18	0.08	0.06	0.06
	30	0.10	0.07	0.06	0.06	0.10	0.07	0.06	0.06	0.11	0.06	0.05	0.05
	60	0.07	0.06	0.06	0.05	0.07	0.06	0.05	0.05	0.07	0.06	0.06	0.05
	90	0.06	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.06	0.05	0.05	0.05
20	10	0.48	0.21	0.14	0.11	0.49	0.23	0.14	0.10	0.43	0.25	0.14	0.10
	20	0.20	0.11	0.09	0.07	0.23	0.10	0.07	0.06	0.25	0.09	0.07	0.06
	30	0.14	0.08	0.07	0.06	0.13	0.08	0.06	0.06	0.14	0.07	0.06	0.06
	60	0.08	0.06	0.05	0.05	0.08	0.06	0.05	0.05	0.07	0.06	0.05	0.05
	90	0.07	0.06	0.06	0.05	0.06	0.06	0.05	0.05	0.06	0.05	0.05	0.05

Table A.10: Frequency of rejection under the null of Johansen test on β over $N = 10^4$ replications of DGP (6.1) at 5% significance level. Top panel: asymptotic. Bottom panel: Bartlett corrected.

D Empirical application

Data were downloaded from the Federal Reserve Economic Data (FRED) website:

https://fred.stlouisfed.org/.

Developed Markets (DM) and Emerging Markets (EM) are defined as in the MSCI 2024 Market Classification, available at:

https://www.msci.com/our-solutions/indexes/market-classification.

EM, Emerging Markets: Brazil (BZ), China (CH), India (IN), Malaysia (MA), Mexico (MX), South Africa (SA), South Korea (SK), Taiwan (TA), Thailand (TH); DM, Developed Markets are composed by EU, European countries: UK-SZ: United Kingdom (UK) and Switzerland (SZ); Nc-EZ, Nordic and Eurozone: Denmark (DK), Eurozone (EZ), Norway (NO), Sweden (SW) and Non-EU, non-European countries: Australia (AU), Canada (CA), Hong Kong (HK), Japan (JP), Singapore (SG).

	AU	EZ	UK	BZ	CA	CH	DK	HK	IN	JP
$\widehat{\beta}$	0.0194	1	-0.0091	0.0011	-0.0120	-0.0197	-1.0021	0.0165	-0.0124	0.0054
p-value	< 0.01		0.029	0.503	0.039	< 0.01	< 0.01	0.555	0.080	0.058
	SK	MA	MX	NO	SW	SA	SG	SZ	TA	TH
$\widehat{\beta}$	0.0019	0.0159	0.0045	0.0062	0.0017	-0.0040	-0.0405	-0.0008	0.0184	0.0004
p-value	0.568	< 0.01	0.054	0.014	0.629	0.105	< 0.01	0.793	< 0.01	0.9430

Table A.11: Estimate of the cointegrating vector in the WM system, normalized on the Eurozone.

The semiparametric results are compared with the high dimensional VAR analysis of Onatski and Wang (2018) and Bykhovskaya and Gorin (2024) for inference on s, and with the likelihood analysis of Johansen (1996) for inference on s and ψ.

The result on the presence of cointegration ($s < p$) among the of the 20 WM currencies is further investigated using Wachter plots as in Onatski and Wang (2018) and Bykhovskaya and Gorin (2024), and the test of no cointegration of Bykhovskaya and Gorin (2024). For a VAR(k), $k = 1, 2, 3, 4$, Figure A.2 displays eigenvalues and the Wachter distribution under $s = p$ and Table A.12 reports p-values of the Bykhovskaya and Gorin (2024) test of no cointegration. Computations were performed in R with the package `Largevars` documented in Bykhovskaya et al. (2025). The plots in Figure A.2 display eigenvalues to the right of the support of the Wachter distribution, which is an indication of the presence of cointegration, see Onatski and Wang (2018) and Bykhovskaya and Gorin (2024). This is in line with the results of the test of no cointegration in Table A.12, which reports p-values below 0.01.

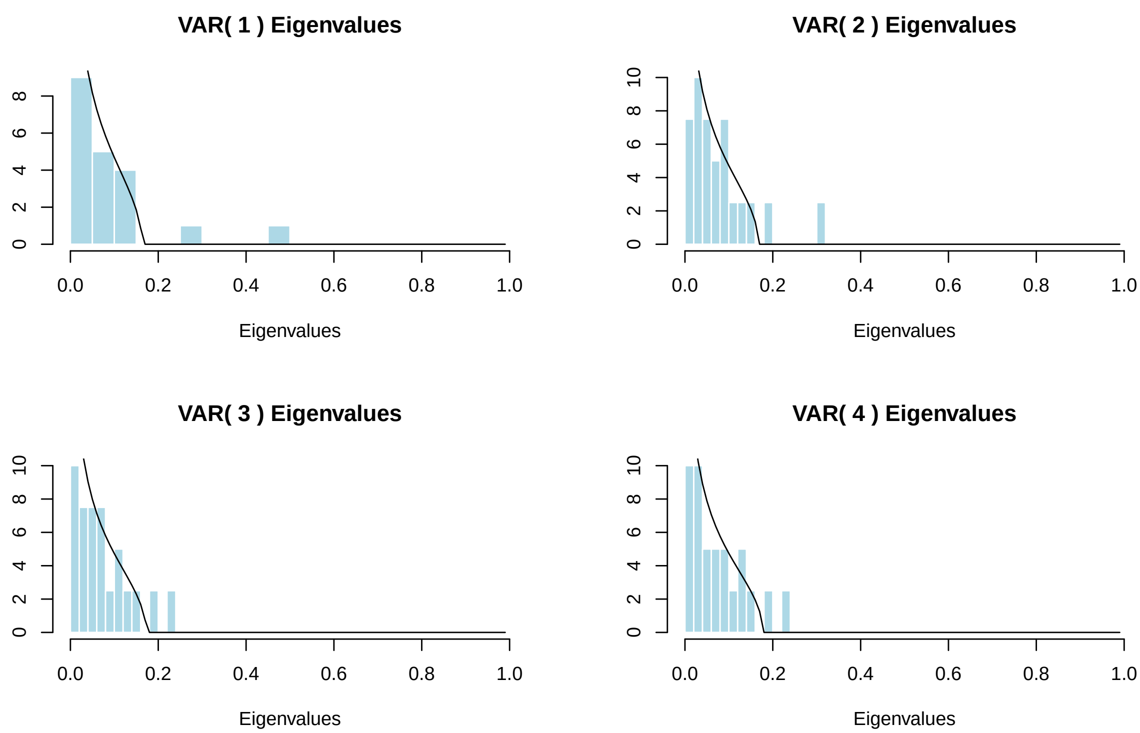

Figure A.2: Bykhovskaya and Gorin (2024) analysis of the 20 WM currencies. Eigenvalues (histogram) and Wachter distribution (line) based on VAR(k), $k = 1, 2, 3, 4$.

For the group of Nc-EZ currencies, the semiparametric results are compared with those based on the likelihood analysis of the Error Correction Model $\Delta X_t = \alpha\beta' X_{t-1} + \Gamma\Delta X_{t-1} +$

k	1	2	3	4
p-value	< 0.01	< 0.01	< 0.01	< 0.01

Table A.12: Bykhovskaya and Gorin (2024) test for the null of no cointegration among the 20 WM currencies based on VAR(k), $k = 1, 2, 3, 4$.

ε_t, with $\alpha, \beta \in \mathbb{R}^{p\times r}$, see Johansen (1996). Table A.13 reports the VAR likelihood ratio test on r and the QML estimate of β in Johansen (1996) (with no deterministic components).

The trace test strongly rejects $r = 0$, i.e. $s = 4$, and does not reject $r \leq 1$, i.e. $s \geq 3$, hence implying the same conclusion as the semiparametric analysis of $s = 3$ stochastic trends and $r = 1$ cointegrating relation in the Nc-EZ system. Moreover, the QML estimate of β reported in the table is very similar to the semiparametric estimate in (7.1).

r	s	Trace test	5% critical value	currency	$\widehat{\beta}$	p-values
0	4	128.63	39.71	DK	−1.015	< 0.01
≤ 1	≥ 3	12.75	24.08	Euro	1	
≤ 2	≥ 2	3.63	12.21	NO	0.015	< 0.01
≤ 3	≥ 1	0.22	4.14	SW	0.001	0.741

Table A.13: Likelihood based analysis of the 4 Nc-EZ currencies. Trace test and corresponding 5% critical values and ML estimate of β with p-values, see Johansen (1996).

References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*(3), 817–858.

Andrews, D. W. K. and J. C. Monahan (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica 60*(4), 953–966.

Bykhovskaya, A. and V. Gorin (2024). Asymptotics of cointegration tests for high-dimensional VAR(k). *The Review of Economics and Statistics*, 1–38.

Bykhovskaya, A., V. Gorin, and E. Kiss (2025). Largevars: An R package for testing large VARs for the presence of cointegration. arXiv:2509.06295.

Doornik, J. A. (1998). Approximations to the asymptotic distributions of cointegration tests. *Journal of Economic Surveys 12*(5), 573–593.

Franchi, M. and P. Paruolo (2020). Cointegration in functional autoregressive processes. *Econometric Theory 36*, 803–839.

Johansen, S. (1996). *Likelihood-based Inference in Cointegrated Vector Auto-Regressive Models*. Oxford University Press.

Johansen, S. (2000). A Bartlett Correction Factor for Tests on the Cointegrating Relations. *Econometric Theory 16*, 740–778.

Johansen, S. (2002). A Small Sample Correction for the Test of Cointegrating Rank in the Vector Autoregressive Model. *Econometrica 70*, 1929–1961.

Johansen, S., H. Hansen, and S. Fachin (2005). A simulation study of some functionals of random walk. Working paper.

Lazarus, E., D. J. Lewis, and J. H. Stock (2021). The size-power tradeoff in HAR inference. *Econometrica 89*(5), 2497–2516.

Ogawa, S. (2017). *Noncausal stochastic calculus*. Tokyo, Japan: Springer.

Onatski, A. and C. Wang (2018). Alternative Asymptotics for Cointegration Tests in Large VARs. *Econometrica 86*, 1465–1478.

Onatski, A. and C. Wang (2019). Extreme canonical correlations and high-dimensional cointegration analysis. *Journal of Econometrics 212*, 307–322.

Paruolo, P. (1997). Asymptotic Inference on the Moving Average Impact Matrix in Cointegrated I(1) Var Systems. *Econometric Theory 13*, 79–118.

Paruolo, P. (2006). The likelihood ratio test for the rank of a cointegration submatrix. *Oxford Bulletin of Economics and Statistics 68*(s1), 921–948.

Pedersen, G. (2012). *Analysis Now*. Graduate Texts in Mathematics. Springer New York, NY.

Phillips, P. C. B. (2014). Optimal estimation of cointegrated systems with irrelevant instruments. *Journal of Econometrics 178*, 210–224.

Phillips, P. C. B. and V. Solo (1992). Asymptotics for Linear Processes. *The Annals of Statistics 20*, 971 – 1001.

Sun, Y. (2011). Robust trend inference with series variance estimator and testing-optimal smoothing parameter. *Journal of Econometrics 164*(2), 345–366.

Sun, Y. (2014). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. *Journal of Econometrics 178*, 659–677.

The MathWorks Inc. (2025). Econometrics Toolbox Version 25.1 (R2025a).

Twardowska, K. (1996). Wong-Zakai Approximations for Stochastic Differential Equations. *Acta Applicandae Mathematicae 43*, 317–359.