CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Testing for Instability in Covariance Structures

Chihwa Kao, Lorenzo Trapani, and Giovanni Urga

CEA@Cass Working Paper Series

WP–CEA–07-2014

Testing for Instability in Covariance Structures*

Chihwa Kao† Syracuse University

Lorenzo Trapani‡ Cass Business School

Giovanni Urga§ Cass Business School and Università di Bergamo

June 10, 2014

Abstract

We propose a test for the stability over time of the covariance matrix of multivariate time series. The analysis is extended to the eigensystem to ascertain changes due to instability in the eigenvalues and/or eigenvectors. Using strong Invariance Principles and Law of Large Numbers, we normalise the CUSUM-type statistics to calculate their supremum over the whole sample. The power properties of the test versus local alternatives and alternatives close to the beginning/end of sample are investigated theoretically and via simulation. We extend our theory to test for the stability of the covariance matrix of a multivariate regression model, and we develop tests for the stability of the loadings in a panel factor model context. The testing procedures are illustrated through two applications: we study the stability of the principal components of the term structure of 18 US interest rates; and we investigate the stability of the covariance matrix of the error term of a Vector AutoRegression applied to exchange rates.

JEL codes: C1, C22, C5.

Keywords: Covariance Matrix, Eigensystem, Factor Models, Changepoint, CUSUM Statistic.

---

*We wish to thank participants in the conference on "High-Dimensional Econometric Modelling" (London, 3-4 December, 2010), in particular Elena Andreou, Jianqing Fan, Domenico Giannone, David Hendry and Soren Johansen, the New York Camp Econometrics VI (Syracuse University, 9-10 April, 2011), the Rimini Time Series Workshop (Rimini, 2-3 June, 2011), the Faculty of Finance Workshops at Cass Business School, the Glasgow Business School Seminar Series, the Lancaster University Management School Seminar Series, the LSE Department of Statistics Seminar Series, in particular Piotr Fryzlewicz and Myung Hwan Seo. Special thanks to Lajos Horvath for providing us with inspiring comments. The usual disclaimer applies.

†Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020 (USA). Tel. +.1. 315.4433233, Fax. +1-315-443-1081, e-mail: dkao@maxwell.syr.edu

‡Centre for Econometric Analysis, Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ (U.K.). Tel. +.44.20.70405260 Fax. +.44.20.70408881, e-mail: L.Trapani@city.ac.uk

§Centre for Econometric Analysis, Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ (U.K.) and Bergamo University (Italy). Tel. +.44.20.70408698 Fax. +.44.20.70408881, e-mail: g.urga@city.ac.uk

1 INTRODUCTION

In this paper, we propose a testing procedure to evaluate the structural stability of the covariance matrix (and its eigensystem) of multivariate time series. A large amount of empirical evidence shows that the issue of changepoint detection in a covariance matrix is of great importance. A classical example is the application of Principal Component Analysis (PCA) to the term structure of interest rates, with the three main principal components interpreted as "slope", "level" and "curvature" (Litterman and Scheinkman, 1991). Bliss (1997), Bliss and Smith (1997) and Perignon and Villa (2006) show that the principal components of the term structure change substantially over time. Similar findings, using a different methodology, are in Audrino et al. (2005). PCA is also widely used in macroeconometrics, for instance to forecast inflation (Stock and Watson, 1999, 2002, 2005). The importance of verifying the stability of a covariance matrix is also evident in the context of Vector AutoRegressive (VAR) models. In the context of forecasting, Castle et al. (2010) show that changes in the smallest eigenvalue of the covariance matrix of the error term have a large impact on predictive ability. Furthermore, the Choleski decomposition is routinely employed in the context of variance decomposition analysis, when examining how much of the variance of the forecast error of each variable in a VAR is due to exogenous shocks to the other variables (see e.g. Pesaran and Shin, 1998).

Despite the relevance of the topic, most studies either assume stability as a working assumption without testing for it, or the testing is carried out by splitting the sample, thus assuming knowledge of the break date *a priori*. This calls for a rigorous testing procedure to estimate the location of the changepoint when breaks are detected. Further, a typical requirement of "classical" PCA is that the data are *i.i.d.* and Gaussian (Flury, 1984, 1988; Perignon and Villa, 2006), which is unsuitable for financial data (see also Audrino et al., 2005).

The theoretical framework developed in this paper builds on a plethora of results for the changepoint problem available in statistics and in econometrics. Existing testing procedures (see e.g. the reviews by Csorgo and Horvath, 1997, and Perron, 2006) are typically based on taking the supremum (or some other metric - see Andrews and Ploberger, 1994) of a sequence of CUSUM-type statistics, thus not requiring prior knowledge of the breakdate. In particular, Aue et al. (2009) develop a test for the structural stability of a covariance matrix, based on minimal assumptions. However, a feature of this test is that, by construction, it has power versus breaks occurring at least (respectively, at most) $O\left(\sqrt{T}\right)$ time periods from the beginning (respectively to the end) of the sample. Lack of power versus alternatives close to either end of the sample is a typical feature in this literature (see also Andrews, 1993), which somewhat limits the applicability of the test. Situations where breaks are due to recent events, like e.g. the 2008 recession, are left out of the analysis. Our contribution complements that of Aue et al. (2009) by proposing a test that has power versus breaks occurring close to the beginning/end of the sample.

The main contribution of this paper is twofold. First, testing for changepoints is extended to PCA. In addition, the extension to testing for the stability of principal components is useful for the purpose of dimension reduction. Our simulations show that tests for the stability of the whole covariance matrix have severe size distortions in finite samples. Contrary to this, testing for the stability of eigenvalues is found to have the correct size and good power even for relatively small samples. As a second contribution, our testing procedure is able to detect breaks occurring up to $O(\ln \ln T)$ periods to the end of the sample. This is achieved by using a Strong Invariance Principle (SIP) and a Strong Law of Large Numbers (SLLN) for the partial sample estimators of the covariance matrix, and by using these results to normalize the CUSUM-type test statistic, using a Darling-Erdos limit theory (see Csorgo and Horvath, 1997; Horvath, 1993). We also extend our results to the case of testing for the stability of the covariance matrix of the error term in a multivariate regression setting.

The theory derived in our paper is illustrated through two applications, one to the US term structure of interest rates, and one to verifying the stability of the covariance matrix of the error term in a VAR model for exchange rates. In both cases, the span of the datasets is from the late nineties to the current date. As far as the former exercise is concerned, we find (as expected) evidence of changes in the volatility and in the loading of the principal components of the term structure around the end of 2007/beginning of 2008. As regards the latter, although the sample covers roughly the same period, we find very little evidence of any changes, suggesting that the covariance matrix of the error term has been stable even during the recession.

The paper is organized as follows. Section 2 contains the SIP and its extension to the eigensystem. The test statistic and its distribution under the null (as well as its behaviour under local-to-null alternatives) is in Section 3. Monte Carlo evidence is in Section 4, while the applications to the term structure of interest rates and to exchange rates are in Section 5. Section 6 concludes. An Internet Appendix reports Lemmas (Appendix A), and Proofs of Theorems and Propositions (Appendix B) in the paper.

A word on notation. Limits are denoted as "$\to$" (the ordinary limit); "$\xrightarrow{p}$" (convergence in probability); "$\xrightarrow{d}$" and (convergence in distribution). Orders of magnitude for an almost surely convergent sequence (say s_T) are denoted as $O_{a.s.}(T^\varsigma)$ and $o_{a.s.}(T^\varsigma)$ when, for some $\varepsilon > 0$ and $\tilde{T} < \infty$, $P\left[|T^{-\varsigma} s_T| < \varepsilon \text{ for all } T \geq \tilde{T}\right] = 1$ and $T^{-\varsigma} s_T \to 0$ almost surely respectively. Orders of magnitude for a sequence converging in probability (say s'_T) are denoted as $O_p(T^\varsigma)$ and $o_p(T^\varsigma)$ when, for some $\varepsilon > 0$, $\Delta_\varepsilon > 0$ and $\tilde{T}_\varepsilon < \infty$, $P[|T^{-\varsigma} s'_T| > \Delta_\varepsilon] < \varepsilon$ for all $T > \tilde{T}_\varepsilon$ and $T^{-\varsigma} s'_T \to 0$ in probability respectively. Standard Wiener processes and Brownian bridges of dimension q are denoted as $W_q(\cdot)$ and $B_q(\cdot)$ respectively; $\|v\|$ denotes the Euclidean norm of a vector v in $\mathbb{R}^n$; similarly, $\|A\|$ denotes the Euclidean norm of a matrix A in $\mathbb{R}^n$, and $|\cdot|_p$ the L_p-norm; the integer part of a real number x is denoted as $\lfloor x \rfloor$. Constants that do not depend on the sample size are denoted as M, M', M'', etc.

2 THEORETICAL FRAMEWORK

Let $\{y_t\}_{t=1}^{T}$ be a time series of dimension n. We assume, without loss of generality, that y_t has zero mean and covariance matrix $\Sigma \equiv E\left(y_t y_t'\right)$ - see also Section 3.1 on this. This section contains the asymptotics of the partial sample estimates of Σ; the results are used in Section 3 in order to construct the CUSUM-type test statistic to test for breaks in Σ and its eigensystem. Specifically, we report a SIP for the partial sample estimators of Σ and an estimator of the long run covariance matrix of the estimated Σ, say V_Σ; and we extend the asymptotics to PCA.

Strong Invariance Principle and estimation of V_Σ

Let $\hat{\Sigma}$ be the sample covariance matrix, i.e. $\hat{\Sigma} = T^{-1}\sum_{t=1}^{T} y_t y_t'$. For a given $\tau \in [0,1]$, we define a point in time $\lfloor T\tau \rfloor$, and we use the subscripts τ and $1-\tau$ to denote quantities calculated using the subsamples $t = 1, ..., \lfloor T\tau \rfloor$ and $t = \lfloor T\tau \rfloor + 1, ..., T$ respectively. In particular, we consider the sequence of partial sample estimators $\hat{\Sigma}_\tau = (T\tau)^{-1}\sum_{t=1}^{\lfloor T\tau \rfloor} y_t y_t'$, and similarly $\hat{\Sigma}_{1-\tau} = \left[T(1-\tau)\right]^{-1}\sum_{t=\lfloor T\tau \rfloor+1}^{T} y_t y_t'$. Finally, henceforth we denote $w_t = vec\left(y_t y_t'\right)$ and $\bar{w}_t = vec\left(y_t y_t' - \Sigma\right)$.

In the sequel, we need the following assumption.

Assumption 1 *(i)* $\sup_t E\left\|y_t\right\|^{2r} < \infty$ for some $r > 2$; *(ii)* y_t is $L_{2+\epsilon}$-NED (Near Epoch Dependent) for some $\epsilon > 0$, of size $\alpha \in (1, +\infty)$ on a strong mixing base $\{v_t\}_{t=-\infty}^{+\infty}$ of size $-r/(r-2)$ and $r > \frac{2\alpha-1}{\alpha-1}$; *(iii)* letting $V_{\Sigma,T} = T^{-1}E\left[\left(\sum_{t=1}^{T}\bar{w}_t\right)\left(\sum_{t=1}^{T}\bar{w}_t\right)'\right]$, $V_{\Sigma,T}$ is positive definite uniformly in T, and as $T \to \infty$, $V_{\Sigma,T} \to V_\Sigma$ with $\left\|V_\Sigma\right\| < \infty$; *(iv)* letting $\bar{w}_{it}$ be the i-th element of $\bar{w}_t$ and defining $S_{iT,m} \equiv \sum_{t=m+1}^{m+T}\bar{w}_{it}$, there exists a positive definite matrix $\bar{\Omega} = \{\varpi_{ij}\}$ such that $T^{-1}\left|E\left[S_{iT,m}S_{jT,m}\right] - \varpi_{ij}\right| \leq MT^{-\psi}$, for all i and j and uniformly in m, with $\psi > 0$.

Assumption 1 specifies the moment conditions and the memory allowed in y_t; no distributional assumptions are required. According to part *(i)*, at least the 4-th moment of y_t is required to be finite, similarly to Aue et al. (2009). As far as serial dependence is concerned, the requirement that y_t be NED is typical in nonlinear time series analysis (see Gallant and White, 1988) and it implies that y_t is a mixingale (Davidson, 2002a). Many of the DGPs considered in the literature generate NED series - examples include GARCH, bilinear and threshold models (see Davidson, 2002b). Part *(ii)* illustrates the trade-off between the memory of y_t (i.e. its NED size α), and its largest existing moment: as α (the memory of y_t) approaches 1, r has to increase. Note that in our context, the data (y_t) undergo a non-Lipschitz transformation (viz., they are squared), and therefore the relationship between moment conditions and memory is not the "standard" one (see e.g. the IP in Theorem 29.6 in Davidson, 2002a). In principle, moment conditions such as the one in part *(ii)* could be tested for, e.g. using a test based on some tail-index estimator - J. B. Hill (2010, 2011) extends the well-known Hill's estimator to the context of dependent data. Other types of dependence could be considered, e.g. assuming a linear process for y_t - an IP for the sample variance is in Phillips and Solo (1992, Theorem 3.8). Part *(iv)* is a bound on the growth rate of the variance of partial sums of $\bar{w}_t$, and it is the same as equation (1.5) in Eberlein (1986); see also Assumption A.3 in Corradi (1999). Although it is not needed to prove the IP for the partial sum process of $\bar{w}_t$, it is a sufficient condition for the SIP.

Theorem 1 contains the IP and the SIP for the partial sums of $\bar{w}_t$.

Theorem 1 *Under Assumptions 1(i)-(iii), as $T \to \infty$*

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor T\tau \rfloor} \bar{w}_t \xrightarrow{d} [V_\Sigma]^{1/2} W_{n^2}(\tau), \tag{1}$$

uniformly in τ. Redefining $\bar{w}_t$ in a richer probability space, under Assumptions 1(i)-(iv)

$$\sum_{t=1}^{\lfloor T\tau \rfloor} \bar{w}_t = \sum_{t=1}^{\lfloor T\tau \rfloor} X_t + O_{a.s.}\left(\lfloor T\tau \rfloor^{\frac{1}{2}-\delta} \right), \tag{2}$$

uniformly in τ, where X_t is a zero mean, i.i.d. Gaussian sequence with $E(X_t X_t') = V_\Sigma$ and $\delta > 0$.

Remarks

T1.1 Equation (1) is an IP for $\bar{w}_t$ (i.e. a weak convergence result), which is sufficient to use the test statistics discussed e.g. in Andrews (1993) and Andrews and Ploberger (1994).

T1.2 Equation (2) is an almost sure result, which also provides a rate of convergence. The practical consequence of (2) is that the dependent, heteroskedastic series $\bar{w}_t$ can be replaced with a sequence of *i.i.d.* normally distributed random variables, with the same long run variance as $\bar{w}_t$.

We now turn to the estimation of V_Σ. If no serial dependence is present, a possible choice is the full sample estimator $\hat{V}_\Sigma = \frac{1}{T}\sum_{t=1}^{T} w_t w_t' - \left[vec\left(\hat{\Sigma}\right)\right]\left[vec\left(\hat{\Sigma}\right)\right]'$. Alternatively, one could use the sequence of partial sample estimators

$$\hat{V}_{\Sigma,\tau} = \frac{1}{T}\sum_{t=1}^{T} w_t w_t' - \left\{ \tau \left[vec\left(\hat{\Sigma}_\tau\right)\right]\left[vec\left(\hat{\Sigma}_\tau\right)\right]' + (1-\tau)\left[vec\left(\hat{\Sigma}_{1-\tau}\right)\right]\left[vec\left(\hat{\Sigma}_{1-\tau}\right)\right]' \right\}.$$

To accommodate for the case $\Psi_l \equiv E\left(\bar{w}_t \bar{w}_{t-l}'\right) \neq 0$ for some l, we propose a weighted sum-of-covariance estimator with bandwidth m:

$$\tilde{V}_\Sigma = \hat{\Psi}_0 + \sum_{l=1}^{m}\left(1 - \frac{l}{m}\right)\left[\hat{\Psi}_l + \hat{\Psi}_l'\right], \tag{3}$$

where $\hat{\Psi}_l = \frac{1}{T-(l+1)} \sum_{t=l+1}^{T} \left[w_t - vec\left(\hat{\Sigma}\right)\right] \left[w_{t-l} - vec\left(\hat{\Sigma}\right)\right]'$; or $\tilde{V}_{\Sigma,\tau} = \left(\hat{\Psi}_{0,\lfloor T\tau \rfloor} + \hat{\Psi}_{0,1-\lfloor T\tau \rfloor}\right)$ $+ \sum_{l=1}^{m} \left(1 - \frac{l}{m}\right) \left[\left(\hat{\Psi}_{l,\lfloor T\tau \rfloor} + \hat{\Psi}'_{l,\lfloor T\tau \rfloor}\right) + \left(\hat{\Psi}_{l,1-\lfloor T\tau \rfloor} + \hat{\Psi}'_{l,1-\lfloor T\tau \rfloor}\right)\right]$, where $\hat{\Psi}_{l,\lfloor T\tau \rfloor} = \frac{1}{T-(l+1)} \sum_{t=l+1}^{\lfloor T\tau \rfloor}$ $\left[w_t - vec\left(\hat{\Sigma}_\tau\right)\right] \left[w_{t-l} - vec\left(\hat{\Sigma}_\tau\right)\right]'$, and similarly for $\hat{\Psi}_{l,1-\lfloor T\tau \rfloor}$.

In order to derive the asymptotics of $\hat{V}_{\Sigma,\tau}$ and $\tilde{V}_{\Sigma,\tau}$, consider the following assumption:

Assumption 2. *(i)* either (a) $\Psi_l = 0$ for all $l \neq 0$ or (b) $\sum_{l=0}^{\infty} l^s \|\Psi_l\| < \infty$ for some $s \geq 1$; *(ii)* $\sup_t E \|y_t\|^{4r} < \infty$ for some $r > 2$; *(iii)* letting $\Omega_T = T^{-1} E \left\{\sum_{t=1}^{T} vec\left[\bar{w}_t \bar{w}'_t - E\left(\bar{w}_t \bar{w}'_t\right)\right] vec\left[\bar{w}_t \bar{w}'_t - E\left(\bar{w}_t \bar{w}'_t\right)\right]'\right\}$, Ω_T is positive definite uniformly in T, and $\Omega_T \to \Omega$ with $\|\Omega\| < \infty$.

Assumption 2 encompasses various possible cases. Part *(i)*(a) considers the basic, non autocorrelated case, for which both $\hat{V}_\Sigma$ and $\hat{V}_{\Sigma,\tau}$ are valid choices. Part *(i)*(b) considers the possibility of non-zero autocorrelations. Intuitively, the assumption that the 4-th moment of y_t exists, as in Assumption 1*(i)*, entails, through a Law of Large Numbers (LLN), the consistency of $\hat{V}_{\Sigma,\tau}$. Part *(ii)* supersedes Assumption 1*(i)*, by requiring the existence of moments up to the 8-th. Intuitively, this implies that an IP holds for the partial sums of $vec\left[\bar{w}_t \bar{w}'_t - E\left(\bar{w}_t \bar{w}'_t\right)\right]$.

The consistency of $\hat{V}_{\Sigma,\tau}$ and of $\tilde{V}_{\Sigma,\tau}$ is in Theorem 2:

Theorem 2 *Under H_0, as $T \to \infty$:*

if Assumptions 1(i)-(iii) and 2(i)(a) hold:

$$\sup_{1 \leq \lfloor T\tau \rfloor \leq T} \left\|\hat{V}_{\Sigma,\tau} - V_\Sigma\right\| = o_p\left(\frac{1}{T^{\delta'}}\right), \tag{4}$$

if Assumptions 1(i)-(iii) and 2(i)(b) hold:

$$\sup_{1 \leq \lfloor T\tau \rfloor \leq T} \left\|\tilde{V}_{\Sigma,\tau} - V_\Sigma\right\| = O_p\left(\frac{1}{m}\right) + O_p\left(\frac{m}{T^{\delta'}}\right), \tag{5}$$

if Assumptions 1(i)-(iii) and 2(i)(b)-(ii)-(iii) hold:

$$\sup_{1 \leq \lfloor T\tau \rfloor \leq T} \left\|\tilde{V}_{\Sigma,\tau} - V_\Sigma\right\| = O_p\left(\frac{1}{m}\right) + O_p\left(\frac{m}{\sqrt{T}}\right), \tag{6}$$

where $\delta' > 0$. The same rates hold for $\hat{V}_\Sigma$ or $\tilde{V}_\Sigma$.

Remarks

T2.1 Equation (4) is based on a SLLN for the case of no autocorrelation in w_t - see also Ling (2007). Theorem 2 provides a uniform rate of convergence for $\hat{V}_{\Sigma,\tau}$ and $\tilde{V}_{\Sigma,\tau}$, as it is usually required in this literature (e.g. Lemma 2.1.2 in Csorgo and Horvath, 1997, p. 76; see also the proof of Theorem 3 below).

T2.2 In case of serial dependence, (5) states that it is possible to construct an estimator of V_Σ with a rate of convergence. This can be refined as in (6). Indeed, Assumptions 2*(ii)-(iii)* allow for an IP to hold for partial sums of $vec\left[\bar{w}_t\bar{w}'_{t-l} - E\left(\bar{w}_t\bar{w}'_{t-l}\right)\right]$, whence the $O_p\left(T^{-1/2}\right)$ convergence rate, uniformly in τ. Equation (6) also provides a selection rule for the bandwidth m: in addition to the conditions $m \to \infty$ and $\frac{m}{\sqrt{T}} \to 0$, the speed of convergence of $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\left\|\tilde{V}_{\Sigma,\tau} - V_\Sigma\right\|$ can be maximised by setting $m = O\left(T^{1/4}\right)$.

Estimation of the eigensystem

In this section, we extend the asymptotics for the partial sample estimates of Σ to its eigensystem.

Let the i-th eigenvalue/eigenvector couple be defined as (λ_i, x_i); the eigenvectors are defined as an orthonormal basis, i.e. $x'_i x_j = \delta_{ij}$, where δ_{ij} is Kronecker's delta. Since $\Sigma x_i = \lambda_i x_i$, a natural estimator for (λ_i, x_i) is the solution to the system

$$\left\{\begin{array}{c} \hat{\Sigma}\hat{X} = \hat{X}\hat{\Lambda} \\ \hat{X}'\hat{X} = I \end{array}\right. , \tag{7}$$

where $\hat{X} = [\hat{x}_1, ..., \hat{x}_n]$, $\hat{x}_i$ denotes the estimate of x_i, and $\hat{\Lambda}$ is a diagonal matrix containing the estimated eigenvalues $\hat{\lambda}_i$ in decreasing order. Estimation of $\{(\lambda_i, x_i)\}_{i=1}^n$ based on (7) is known as Anderson's Principal Component (PC) estimator. Similarly, the partial sample estimators of the eigenvalues and eigenvectors are the solutions to $\hat{\Sigma}_\tau\hat{x}_{i,\tau} = \hat{\lambda}_{i,\tau}\hat{x}_{i,\tau}$.

As we mention below (see Remark P1.2), one disadvantage of Anderson's PC estimator is that the estimated eigenvectors have a singular asymptotic covariance matrix (see Kollo and Neudecker, 1997). In order to avoid this issue, an estimator based on a different normalisation can be proposed, known as the Pearson-Hotelling's PC estimator; in this case, the estimated eigenvalues are the same as from (7), but the eigenvectors γ_i are defined (and estimated) as an eigenvalue-normed basis, viz. $\gamma'_i\gamma_j = \lambda_i\delta_{ij}$. Thus, $\gamma_i \equiv \lambda_i^{1/2}x_i$. A typical interpretation of the γ_is in the context of the term structure of interest rates (Litterman and Scheinkman, 1991; Perignon and Villa, 2006) is that λ_i is the "volatility" of γ_i, and x_i represents its "loading". The estimates of the eigensystem according to the Pearson-Hotelling approach are the solution to the system

$$\left\{\begin{array}{c} \hat{\Sigma}\hat{X} = \hat{X}\hat{\Lambda} \\ \hat{X}'\hat{X} = \hat{\Lambda} \end{array}\right. . \tag{8}$$

Upon calculating the solutions of (8), it turns out that the eigenvectors are estimated by $\hat{\gamma}_i = \hat{\lambda}_i^{1/2}\hat{x}_i$, i.e. by the same estimator for the eigenvector as in (7) multiplied by the square root of the corresponding estimate of the eigenvalue. Similarly, we define the partial sample estimator of γ_i as $\hat{\gamma}_{i,\tau} = \hat{\lambda}_{i,\tau}^{1/2}\hat{x}_{i,\tau}$.

Consider the following assumption.

Assumption 3. The matrix Σ has distinct eigenvalues.

Assumption 3 is typical of PCA and it allows to use Matrix Perturbation Theory (MPT); the assumption could be relaxed at the price of a more complicated analysis, still based on MPT. In essence, the asymptotics of $\left(\hat{\lambda}_{i,\tau}, \hat{x}_{i,\tau}\right)$ is derived by treating $\hat{\Sigma}_\tau$ as a perturbation of Σ, thus deriving the expressions for the estimation errors of $\hat{\lambda}_{i,\tau}$ and $\hat{x}_{i,\tau}$.

The extension of the IP and the SIP to the eigensystem of Σ is reported in Proposition 1:

Proposition 1 *Under Assumptions 1 and 3, as $T \to \infty$, uniformly in τ*

$$\hat{\lambda}_{i,\tau} - \lambda_i = \left(x_i' \otimes x_i'\right) vec\left(\hat{\Sigma}_\tau - \Sigma\right) + O_p\left(T^{-1}\right), \tag{9}$$

$$\hat{x}_{i,\tau} - x_i = v_{x,i} vec\left(\hat{\Sigma}_\tau - \Sigma\right) + O_p\left(T^{-1}\right), \tag{10}$$

$$\hat{\gamma}_{i,\tau} - \gamma_i = v_{\gamma,i} vec\left(\hat{\Sigma}_\tau - \Sigma\right) + O_p\left(T^{-1}\right), \tag{11}$$

where $v_{x,i} = \left[\sum_{k \neq i} \frac{x_k}{\lambda_i - \lambda_k}\left(x_k' \otimes x_i'\right)\right]$ *and* $v_{\gamma,i} = \frac{1}{2}\frac{x_i}{\lambda_i^{1/2}}\left(x_i' \otimes x_i'\right) + \sum_{k \neq i} \frac{\lambda_i^{1/2} x_k}{\lambda_i - \lambda_k}\left(x_i' \otimes x_k'\right)$.

Remarks

P1.1 Proposition 1 states that the estimation errors $\hat{\lambda}_{i,\tau} - \lambda_i$, $\hat{x}_{i,\tau} - x_i$ and $\hat{\gamma}_{i,\tau} - \gamma_i$ are, asymptotically, linear functions of $\hat{\Sigma}_\tau - \Sigma$; thus, the IP and the SIP in Theorem 1 carry through to the estimated eigensystem. The results in Proposition 1, and the method of proof, can be compared to related results in Kollo and Neudecker (1997).

P1.2 By (10), the asymptotic covariance matrix of $\sqrt{T}\left(\hat{x}_{i,\tau} - x_i\right)$ is $v_{x,i} V_\Sigma v_{x,i}'$. It can be shown (see e.g. Kollo and Neudecker, 1997, p. 66) that $v_{x,i} V_\Sigma v_{x,i}'$ is singular. Thus, in order to carry out tests on the $\hat{x}_i$s, one would have to use a generalised inverse of the estimated $v_{x,i} V_\Sigma v_{x,i}'$. However, there is no obvious way to calculate the rank of $v_{x,i} V_\Sigma v_{x,i}'$. This makes it difficult to prove the consistency of the Moore-Penrose inverse for $v_{x,i} V_\Sigma v_{x,i}'$ (see Andrews, 1987). Thus, we recommend to carry out tests on the eigenvectors using the γ_is.

P1.3 We show in the appendix that

$$E\left[T\left(\hat{\lambda}_{i,\tau} - \lambda_i\right)\right] = \sum_{k \neq i} \frac{\left(x_i' \otimes x_k'\right) V_\Sigma \left(x_k \otimes x_i\right)}{\lambda_i - \lambda_k}, \tag{12}$$

as $T \to \infty$. As far as the impact of n is concerned, V_Σ is an n^2-dimensional matrix; thus, in general the quadratic form $\left(x_i' \otimes x_k'\right) V_\Sigma \left(x_k \otimes x_i\right)$ has magnitude of order $O\left(n^2\right)$. Also, due to the summation on the right hand side of (12) involving $n-1$ elements, $E\left[T\left(\hat{\lambda}_{i,\tau} - \lambda_i\right)\right]$

$= O\left(n^3\right)$. Thus, the asymptotic bias is of order $O\left(\frac{n^3}{T}\right)$, and it is always positive for the largest eigenvalue; a bias-corrected version is $\tilde{\lambda}_{i,\tau} = \hat{\lambda}_{i,\tau} - T^{-1}\sum_{k\neq i}\left[\hat{x}_i' \otimes \hat{x}_k'\right] \frac{\tilde{V}_\Sigma}{\hat{\lambda}_i - \hat{\lambda}_k}$ $\left[\hat{x}_k \otimes \hat{x}_i\right]$. This result is of independent interest; it could be useful e.g. when measuring the percentage of the total variance of y_t explained by each of its principal components. Similarly, we show that

$$E\left[T\left(\hat{x}_{i,\tau} - x_i\right)\right] = \sum_{k\neq i}\sum_{j\neq i} \frac{\left(x_k' \otimes x_j'\right) V_\Sigma \left(x_j \otimes x_i\right)}{\left(\lambda_i - \lambda_k\right)\left(\lambda_i - \lambda_j\right)} x_k, \tag{13}$$

which provides an expression to correct the bias of the $\hat{x}_{i,\tau}$; combining (12) and (13), a bias-corrected version of $\hat{\gamma}_{i,\tau}$s the can also be computed.

P1.4 In Appendix B, we show an ancillary result concerning the estimation of the eigenvalues of a matrix which undergoes a change over time. Suppose that the covariance matrix Σ has a break at time $k_{0,T}$, such that $\Sigma = \Sigma_0$ for $t = 1, ..., k_{0,T}$ and $\Sigma = \Sigma_0 + \Delta_\Sigma$ for $t = k_{0,T} + 1, ..., T$. Assume further that the break affects only a subspace of the matrix, i.e. if $\Sigma_0 = \sum_{i=1}^n \lambda_i x_i x_i'$, it holds that $\Delta_\Sigma = \sum_{i\in C} \tilde{\lambda}_i \tilde{x}_i \tilde{x}_i' - \sum_{i\in C} \lambda_i x_i x_i'$, where C is a (nontrivial) family of indices, and $\tilde{\lambda}_i$ and $\tilde{x}_i$ are perturbations of λ_i and x_i; we assume that $\tilde{x}_i'\tilde{x}_j = \delta_{ij}$ and that $\tilde{x}_i' x_j = 0$ for all $i \in C$ and $j \notin C$. By virtue of this, clearly $\Sigma_0 x_i = \lambda_i x_i$ for all $t = 1, ..., T$ for $i \notin C$. In the proof of Proposition 1, we show that $\hat{\lambda}_{i,\tau} - \lambda_i = \left(x_i' \otimes x_i'\right) vec\left(\hat{\Sigma}_\tau - \Sigma_0\right) + O_p\left(T^{-1}\right)$ and $\hat{x}_{i,\tau} - x_i = v_{x,i}\, vec\left(\hat{\Sigma}_\tau - \Sigma_0\right) + O_p\left(T^{-1}\right)$ for all $i \notin C$ and uniformly in τ. This entails that as long as one eigenvalue/eigenvector couple is found to be unchanged over time, then it can be estimated consistently, with no need to test for the time stability of the rest of the eigensystem. Similar results, with a different method of proof and a focus on the eigenvalues only, have also been derived in the context of the so-called "Google matrices" (see Zhou, 2011). This result can also be compared with the findings in Bates et al. (2013), who, in the context of panel factor models, show that factors can be estimated consistently even in presence of breaks in the loadings.

Define $\lambda \equiv [\lambda_1, ..., \lambda_n]'$ as the n-dimensional vector containing the eigenvalues sorted in descending order, and $\Gamma \equiv [\gamma_1, ..., \gamma_n]$; $\hat{z} \equiv \left[\hat{\lambda}', vec\left(\hat{\Gamma}\right)'\right]'$ with $\hat{z}_\tau - z = D_{\lambda\gamma} vec\left(\hat{\Sigma}_\tau - \Sigma\right) + O_p\left(T^{-1}\right)$ and $D_{\lambda\gamma} \equiv \left[x_1 \otimes x_1, ..., x_n \otimes x_n, v_{\gamma,1}', ..., v_{\gamma,n}'\right]'$. The matrix $D_{\lambda\gamma}$ can be estimated as $\hat{D}_{\lambda\gamma} = \left[\hat{x}_1 \otimes \hat{x}_1, ..., \hat{x}_n \otimes \hat{x}_n, \hat{v}_{\gamma,1}', ..., \hat{v}_{\gamma,n}'\right]'$, with $\hat{v}_{\gamma,i} = \frac{1}{2}\frac{\hat{x}_i}{\hat{\lambda}_i^{1/2}}\left(\hat{x}_i' \otimes \hat{x}_i'\right) + \sum_{k\neq i}\frac{\hat{\lambda}_i^{1/2}\hat{x}_k}{\hat{\lambda}_i - \hat{\lambda}_k}\left(\hat{x}_i' \otimes \hat{x}_k'\right)$.

The asymptotics of $\hat{z}_\tau$ follows from Theorem 1 and Proposition 1, and we summarize it below.

Corollary 1 *Under Assumptions 1 and 3, as $T \to \infty$, it holds that $\sqrt{T}\left(\hat{z}_\tau - z\right) \xrightarrow{d} \left[V_z\right]^{1/2} W_{n(2n+1)}(\tau)$. Also, $T\left(\hat{z}_\tau - z\right) = \sum_{t=1}^{\lfloor T\tau \rfloor} \tilde{X}_t + O_{a.s.}\left(\lfloor T\tau \rfloor^{\frac{1}{2}-\delta}\right)$, uniformly in τ, where $V_z = D_{\lambda\gamma} V_\Sigma D_{\lambda\gamma}'$ and $\tilde{X}_t$ is a zero mean, i.i.d. Gaussian sequence with $E\left(\tilde{X}_t \tilde{X}_t'\right) = V_z$ and $\delta > 0$.*

Corollary 1 entails that

$$\sqrt{T}\left(\hat{\lambda}_\tau-\lambda\right)\xrightarrow{d}\left[V_\lambda\right]^{1/2}W_n(\tau),$$
$$\sqrt{T}vec\left(\hat{\Gamma}_\tau-\Gamma\right)\xrightarrow{d}\left[V_\Gamma\right]^{1/2}W_{n^2}(\tau),$$

with: V_λ a matrix with (i,j)-th element given by $V_{ij}^\lambda=(x_i'\otimes x_i')V_\Sigma(x_j\otimes x_j)$, and V_Γ is an $(n^2\times n^2)$-dimensional matrix whose (i,j)-th $n\times n$ block is defined as $V_{ij}^\Gamma=v_{\gamma,i}V_\Sigma v_{\gamma,j}'$.

3 TESTING

This section studies the null distribution and the consistency of tests based on CUSUM-type statistics.

Henceforth, we define the CUSUM process $S(\tau)=\sum_{t=1}^{\lfloor T\tau\rfloor}vec(y_ty_t')$. In light of Corollary 1, test statistics for Σ and its eigensystem can be based on

$$\tilde{S}(\tau)=R\times D_{\lambda x\gamma}\times\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right],\tag{14}$$

with $\tilde{S}(\tau)=0$ for $\tau\leq\frac{1}{T}$ or $\geq 1-\frac{1}{T}$, and R a $p\times n(n+1)$ matrix. For example, when testing for the null of no changes in the first eigenvalue, R is the matrix that extracts the first element of $D_{\lambda\gamma}\times\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]$. Thence, testing is carried out by using

$$\Lambda_T(\tau)=\sqrt{\frac{T}{\lfloor T\tau\rfloor\times\lfloor T(1-\tau)\rfloor}}\times\left[\tilde{S}(\tau)'\tilde{V}_{z,\tau}^{-1}\tilde{S}(\tau)\right]^{1/2},\tag{15}$$

with $\tilde{V}_{z,\tau}=RD_{\lambda\gamma}\tilde{V}_{\Sigma,\tau}D_{\lambda\gamma}'R'$.

Theorem 3 contains the asymptotics of $\sup_{\lfloor T\tau\rfloor}\Lambda_T(\tau)$ under the null.

Theorem 3 *Under Assumptions 1-3, as* $(m,T)\to\infty$ *with* $\frac{1}{m}+\frac{m}{\sqrt{T}}\to 0$,

$$\sup_{\lfloor T\tau_1\rfloor\leq\lfloor T\tau\rfloor\leq\lfloor T\tau_2\rfloor}\Lambda_T(\tau)\xrightarrow{d}\sup_{\tau_1\leq\tau\leq\tau_2}\frac{\|B_p(\tau)\|}{\sqrt{\tau(1-\tau)}},\tag{16}$$

where $B_p(\tau)$ *is a* p*-dimensional standard Brownian bridge and* $[\tau_1,\tau_2]\in(0,1)$. *Also, as* $(m,T)\to\infty$ *with* $\frac{\sqrt{\ln\ln T}}{m}+m\sqrt{\frac{\ln\ln T}{T}}\to 0$,

$$P\left\{a_T\left[\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\Lambda_T(\tau)\right]\leq x+b_T\right\}\to e^{-2e^{-x}},\tag{17}$$

where $a_T=\sqrt{2\ln\ln T}$ *and* $b_T=2\ln\ln T+\frac{p}{2}\ln\ln\ln T-\ln\Gamma\left(\frac{p}{2}\right)$, *with* $\Gamma(\cdot)$ *the Gamma function.*

Remarks

T3.1 According to (16), the maximum is taken in a subset of $[0,1]$, namely $[\tau_1, \tau_2]$. This approach requires an IP for $S(\tau)$, and the Continuous Mapping Theorem (CMT). As noted in Corollary 1 in Andrews (1993, p. 838), $\Lambda_T(\tau)$ is not continuous at $\{0,1\}$ and $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\Lambda_T(\tau)\xrightarrow{p}\infty$ under H_0. Thus, trimming is necessary in this case. Further, in this case it sfuccies to have a consistent estimator of the long-run covariance matrix V_Σ which, in light of equation (6) in Theorem 2, entails that both $m\to\infty$ and $\frac{m}{\sqrt{T}}\to 0$. The considerations in Remark T2.2 apply here.

T3.2 As an alternative approach, the SIP can be used: sums of $\bar{w}_t$ can be replaced by sums of *i.i.d.* Gaussian variables, with an approximation error. Upon normalising $\Lambda_T(\tau)$ with the appropriate norming constants, say a_T and b_T, an Extreme Value (EV henceforth) theorem can be employed. Tests based on $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}[a_T\Lambda_T(\tau)-b_T]$ are designed to be able to detect breaks close to the end of the sample. Results like (17) have been derived by Horvath (1993), for *i.i.d.* Gaussian data, and extended to the case of dependence by Ling (2007), *inter alia*. As far as the long-run covariamce matrix estimator is concerned, in this case the theory requires a consistent estimator at a rate (at least) $o_p\left[\left(\sqrt{\ln\ln T}\right)^{-1}\right]$: therefore, from (6), we need the restrictions $\frac{\sqrt{\ln\ln T}}{m}\to 0$ and $m\sqrt{\frac{\ln\ln T}{T}}\to 0$. Note that the "optimal" choice of the bandwidth m that maximizes the speed of convergence of $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\left\|\tilde{V}_{\Sigma,\tau}-V_\Sigma\right\|$ is still $m=O\left(T^{1/4}\right)$.

T3.3 Theorem 3 allows to test for breaks in Σ when n is finite. As n passes to infinity, Aue et al. (2009) study the behaviour of $\hat{\Lambda}_T^2\equiv\sup_{1\leq\lfloor T\tau\rfloor\leq T}T^{-1}\tilde{S}(\tau)'\left[R\tilde{V}_{z,\tau}R'\right]^{-1}\tilde{S}(\tau)$, showing that, as $T\to\infty$ followed by $n\to\infty$, $\frac{4\hat{\Lambda}_T^2-n^2}{\sqrt{2}n}\xrightarrow{d}N(0,1)$. Based on Goetze and Zaitsev (2012; see also the references therein), it is possible to extend (17) to the case of $n\to\infty$. Although this goes beyond the scope of this paper, we point out that, based on the results available in the literature, the theory developed here can only be extended to the case of n passing to infinity at a very slow rate compared with T. This is also due to the fact that the maximum of the sequence of Wald-type statistics ought to be searched between n and $T-n$, so as not to have a singular matrix. Therefore, in order to detect breaks close to either end of the sample, n cannot be too large, or too many initial observations would have to be trimmed away. In this respect, dimension reduction techniques such as PCA are bound to be advantageous.

Consistency of the test

We now turn to studying the behaviour of $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\Lambda_T(\tau)$ under alternatives. As a leading example, we consider the case of testing for no change in Σ in presence of one abrupt change

$$H_a^{(T)} : vech(\Sigma_t) = \begin{cases} vech\left(\Sigma\right) \text{ for } t = 1, ..., k_{0,T} \\ vech\left(\Sigma\right) + \Delta_T \text{ for } t = k_{0,T} + 1, ..., T \end{cases}, \tag{18}$$

where both the changepoint ($k_{0,T}$) and the size of the break (Δ_T) could depend on T. More general alternatives could be considered (see e.g. Andrews, 1993; Csorgo and Horvath, 1997): these include epidemic alternatives, and also breaks that occur as a smooth transition over time as opposed to abruptly as in (18). Further, note that (18) does not rule out the possibility that only some series (i.e. only some of the coordinates of y_t) actually have a break. This entails that tests based on $\Lambda_T(\tau)$ are capable of detecting breaks that only affect some of the series, and possibly at different points in time.

Theorem 4 illustrates the dependence of the power on Δ_T and $k_{0,T}$.

Theorem 4 *Let Assumptions 1-3 hold, and define $c_{\alpha,T}$ such that, under H_0, $P\left[\sup_{n\leq\lfloor T\tau\rfloor\leq T-n} \Lambda_T(\tau) \leq c_{\alpha,T}\right] = 1-\alpha$ for some $\alpha \in [0,1]$. If, under $H_a^{(T)}$, as $T \to \infty$*

$$\frac{1}{\ln\ln T}\left[\frac{(T-k_{0,T})\,k_{0,T}}{T}\left\|\Delta_T\right\|^2\right] \to \infty, \tag{19}$$

it holds that

$$P\left[\sup_{n\leq\lfloor T\tau\rfloor\leq T-n} \Lambda_T(\tau) > c_{\alpha,T}\right] = 1. \tag{20}$$

Remarks

T4.1 Theorem 4 illustrates the impact of $k_{0,T}$ and Δ_T on the power of tests based on $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n} \Lambda_T(\tau)$. Particularly, consider the two extreme cases:

T4.1.a $\|\Delta_T\| = O(1)$, i.e. finite break size. In this case, the test has power as long as $k_{0,T}$ is strictly bigger than $O(\ln\ln T)$. This can be compared with tests based on $\sup_{1\leq\lfloor T\tau\rfloor\leq T} T^{-1}\tilde{S}(\tau)'\,\tilde{V}_{z,\tau}^{-1}\,\tilde{S}(\tau)$. Using similar algebra as in the proof of Theorem 4, it can be shown that the noncentrality parameter of $\sup_{1\leq\lfloor T\tau\rfloor\leq T} T^{-1}\tilde{S}(\tau)'\,\tilde{V}_{z,\tau}^{-1}\,\tilde{S}(\tau)$ is proportional to $\|\Delta_T\|^2 \frac{k_{0,T}^2}{T}$. Under $\|\Delta_T\| = O(1)$, this entails that nontrivial power is attained as long as $k_{0,T} = O\left(\sqrt{T}\right)$.

T4.1.b $k_{0,T} = O(T)$ - i.e. the break occurs in the middle of the sample. The test is powerful as long as the size of the break is strictly bigger than $O\left(\sqrt{\frac{\ln\ln T}{T}}\right)$. When using trimmed statistics such as in (16), the test is powerful versus mid-sample alternatives of size $O\left(\frac{1}{\sqrt{T}}\right)$: when no trimming is used, there is some, limited loss of power versus mid-sample alternatives.

T4.2 The presence of "large" breaks in Σ is bound to affect inference on the eigensystem - see e.g. Stock and Watson (2002). Consider, as a leading example, testing for the stability of

λ_i. From Proposition 1, the long-run variance of the estimated eigenvalues is estimated by $(\hat{x}_i' \otimes \hat{x}_i') \tilde{V}_\Sigma (\hat{x}_i \otimes \hat{x}_i)$, thus depending on $\hat{x}_i$. In presence of a mid-sample break of magnitude $\|\Delta_T\| = O(1)$, x_i is estimated by $\hat{x}_i$ with an error of magnitude $O_p(1)$, as a consequence of Proposition 1; similarly, $\left\|\tilde{V}_\Sigma - V_\Sigma\right\| = O_p(1)$ also. This entails that the estimated variance of $\hat{\lambda}_i - \lambda_i$, $(\hat{x}_i' \otimes \hat{x}_i') V_\Sigma (\hat{x}_i \otimes \hat{x}_i)$ is $O_p(1)$. On the other hand, the numerator of the test statistic is proportional to $T\left\|\hat{\Sigma} - \Sigma\right\| = O_p(T)$. This entails that the test based on $\sup_{n \leq \lfloor T\tau \rfloor \leq T-n} \Lambda_T(\tau)$ is consistent as $T \to \infty$.

3.1 Applications to the residuals of a multivariate regression

In this section, we extend the results developed above by proposing a test for the time stability of the covariance matrix (and its eigensystem) of the error term in a multivariate regression, such as a VAR.

Consider the following multivariate regression model

$$y_t = \beta x_t + \varepsilon_t, \tag{21}$$

where $t = 1, ..., T$ and y_t and ε_t are $n \times 1$ vectors, x_t is of dimension $q \times 1$ and the matrix of regressors β has dimensions $n \times q$. Of course, n can be equal to 1, thereby recovering the usual univariate regression; alternatively, the regressors x_t can be the lags of y_t, thus having a VAR. Bai (2000) considers a test for structural changes in a VAR, extending it to test for changes in the covariance matrix of the error term. The theory developed in this section also extends naturally to the case of linear or polynomial trends among the regressors - see also Aue et al. (2012).

Let $\Sigma_\varepsilon = E(\varepsilon_t \varepsilon_t')$ be the covariance matrix of the error term ε_t; using the same notation as above, we denote its eigenvalues and eigenvectors as λ_i^ε and x_i^ε respectively, satisfying the relationship $\Sigma_\varepsilon x_i^\varepsilon = \lambda_i^\varepsilon x_i^\varepsilon$ for $i = 1, ..., n$.

Tests for the time stability of Σ_ε and of its eigensystem are based on the residuals $\hat{\varepsilon}_t = y_t - \hat{\beta} x_t$, where $\hat{\beta} = \left[\sum_{t=1}^T y_t x_t'\right] \left[\sum_{t=1}^T x_t x_t'\right]^{-1}$. Hence, the full sample estimator of Σ_ε is defined as $\hat{\Sigma}_\varepsilon = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_t'$. As above, for a given point in time $\lfloor T\tau \rfloor$ with $\tau \in [0, 1]$, we use the subscript τ to denote quantities calculated using the subsamples $t = 1, ..., \lfloor T\tau \rfloor$. In particular, we extensively use the sequence of partial sample estimators $\hat{\Sigma}_{\varepsilon,\tau} = (T\tau)^{-1} \sum_{t=1}^{\lfloor T\tau \rfloor} \hat{\varepsilon}_t \hat{\varepsilon}_t'$. We also employ the following definitions: $w_t^\varepsilon = vec(\varepsilon_t \varepsilon_t')$ and $\bar{w}_t^\varepsilon = vec(\varepsilon_t \varepsilon_t' - \Sigma_\varepsilon)$.

We need the following Assumptions, which are closely related to Assumptions 1-3 above.

Assumption R1. *(i)* Assumptions 1*(i)*-1*(ii)* hold for x_t; *(ii)* Assumptions 1*(i)*-1*(ii)* hold for ε_t; *(iii)* let $w_t^{x\varepsilon} = vec(\varepsilon_t x_t')$; it holds that: (a) $E(w_t^{x\varepsilon}) = 0$ for all t; (b) defining $V_T^{x\varepsilon} = T^{-1} E\left[\left(\sum_{t=1}^T w_t^{x\varepsilon}\right) \left(\sum_{t=1}^T w_t^{x\varepsilon})'\right]$, $V_T^{x\varepsilon}$ is positive definite uniformly in T, and as $T \to$

∞, $V_T^{x\varepsilon} \rightarrow V^{x\varepsilon}$ with $\|V^{x\varepsilon}\| < \infty$; (c) letting $w_{it}^{x\varepsilon}$ be the i-th element of $w_t^{x\varepsilon}$ and defining $S_{iT,m}^{x\varepsilon} \equiv \sum_{t=m+1}^{m+T} w_{it}^{x\varepsilon}$, there exists a positive definite matrix $\bar{\Omega}^{x\varepsilon} = \left\{\varpi_{ij}^{x\varepsilon}\right\}$ such that $T^{-1}\left|E\left[S_{iT,m}^{x\varepsilon}S_{jT,m}^{x\varepsilon}\right] - \varpi_{ij}^{x\varepsilon}\right| \leq MT^{-\psi}$, for all i and j and uniformly in m, with M a constant and $\psi > 0$; *(iv)* (a) letting $V_T^{\varepsilon} = T^{-1}E\left[\left(\sum_{t=1}^{T} \bar{w}_t^{\varepsilon}\right)\left(\sum_{t=1}^{T} \bar{w}_t^{\varepsilon}\right)'\right]$, V_T^{ε} is positive definite uniformly in T, and as $T \rightarrow \infty$, $V_T^{\varepsilon} \rightarrow V^{\varepsilon}$ with $\|V^{\varepsilon}\| < \infty$; (b) letting $\bar{w}_{it}^{\varepsilon}$ be the i-th element of $\bar{w}_t^{\varepsilon}$ and defining $S_{iT,m}^{\varepsilon} \equiv \sum_{t=m+1}^{m+T} \bar{w}_{it}^{\varepsilon}$, there exists a positive definite matrix $\bar{\Omega}^{\varepsilon} = \left\{\varpi_{ij}^{\varepsilon}\right\}$ such that $T^{-1}\left|E\left[S_{iT,m}^{\varepsilon}S_{jT,m}^{\varepsilon}\right] - \varpi_{ij}^{\varepsilon}\right| \leq M'T^{-\psi'}$, for all i and j and uniformly in m, with $\psi' > 0$.

Assumption R2. Let $\Psi_l^{\varepsilon} = E\left(\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime}\right)$; it holds that *(i)* $\sum_{l=0}^{\infty} l^s \|\Psi_l^{\varepsilon}\| < \infty$ for some $s \geq 1$; *(ii)* $\sup_t E\|\varepsilon_t\|^{4r} < \infty$ for some $r > 2$; *(iii)* letting $\Omega_T^{l\varepsilon} = T^{-1}E\left\{\sum_{t=1}^{T} vec\left[\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime} - E\left(\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime}\right)\right]\right.$ $\left. vec\left[\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime} - E\left(\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime}\right)\right]'\right\}$, $\Omega_T^{l\varepsilon}$ is positive definite uniformly in T, and $\Omega_T^{l\varepsilon} \rightarrow \Omega^{l\varepsilon}$ with $\left\|\Omega^{l\varepsilon}\right\| < \infty$ for every l.

Assumption R3. The matrix Σ_{ε} has distinct eigenvalues.

Let $\hat{w}_t^{\varepsilon} = vec\left(\hat{\varepsilon}_t\hat{\varepsilon}_t'\right)$. In order to estimate the long run covariance matrix V^{ε} defined in Assumption R1*(iv)*(b), we propose the weighted sum-of-covariance estimator $\tilde{V}_{\tau}^{\varepsilon} = \left(\hat{\Psi}_{0,\lfloor T\tau\rfloor}^{\varepsilon} + \hat{\Psi}_{0,1-\lfloor T\tau\rfloor}^{\varepsilon}\right)$ $+\sum_{l=1}^{m}\left(1 - \frac{l}{m}\right)\left[\left(\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon} + \hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon\prime}\right) + \left(\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon} + \hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon\prime}\right)\right]$, where $\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon} = \frac{1}{T-(l+1)}\sum_{t=l+1}^{\lfloor T\tau\rfloor}$ $\left[\hat{w}_t^{\varepsilon} - vec\left(\hat{\Sigma}_{\varepsilon}\right)\right]\left[\hat{w}_{t-l}^{\varepsilon} - vec\left(\hat{\Sigma}_{\varepsilon}\right)\right]'$, and similarly $\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon}$.

In order to set up the test statistic, consider the following notation. When estimating the eigensystem of Σ_{ε}, we define the estimators of $(\lambda_i^{\varepsilon}, x_i^{\varepsilon})$ as the solution to $\hat{\Sigma}_{\varepsilon}\hat{x}_i^{\varepsilon} = \hat{\lambda}_i^{\varepsilon}\hat{x}_i^{\varepsilon}$, where $\hat{\lambda}_i^{\varepsilon}$ and $\hat{x}_i^{\varepsilon}$ denote the estimates of λ_i^{ε} and x_i^{ε} respectively; the Pearson-Hotelling estimator of the eigenvectors is defined as $\hat{\gamma}_i^{\varepsilon} = \sqrt{\hat{\lambda}_i^{\varepsilon}}\hat{x}_i^{\varepsilon}$. Similarly, the partial sample estimators of the eigenvalues and eigenvectors are the solutions to $\hat{\Sigma}_{\varepsilon,\tau}\hat{x}_{i,\tau}^{\varepsilon} = \hat{\lambda}_{i,\tau}^{\varepsilon}\hat{x}_{i,\tau}^{\varepsilon}$, and we define $\hat{\gamma}_{i,\tau}^{\varepsilon} = \sqrt{\hat{\lambda}_{i,\tau}^{\varepsilon}}\hat{x}_{i,\tau}^{\varepsilon}$. Thence, let $\lambda^{\varepsilon} \equiv [\lambda_1^{\varepsilon}, ..., \lambda_n^{\varepsilon}]'$ be the n-dimensional vector containing the eigenvalues sorted in descending order; $\Gamma^{\varepsilon} \equiv [\gamma_1^{\varepsilon}, ..., \gamma_n^{\varepsilon}]$; and let $D_{\lambda\gamma}^{\varepsilon} \equiv \left[x_1^{\varepsilon} \otimes x_1^{\varepsilon}, ..., \ x_n^{\varepsilon} \otimes x_n^{\varepsilon}, v_{\gamma,1}^{\varepsilon\prime}, \ ..., v_{\gamma,n}^{\varepsilon\prime}\right]'$, with $v_{\gamma,i}^{\varepsilon} = \frac{1}{2}\frac{x_i^{\varepsilon}}{\sqrt{\lambda_i^{\varepsilon}}}$ $\left(x_i^{\varepsilon\prime} \otimes x_i^{\varepsilon\prime}\right) + \sum_{k\neq i}\frac{\sqrt{\lambda_i^{\varepsilon}}x_k^{\varepsilon}}{\lambda_i^{\varepsilon}-\lambda_k^{\varepsilon}}\left(x_i^{\varepsilon\prime} \otimes x_k^{\varepsilon\prime}\right)$. Estimation of $D_{\lambda\gamma}^{\varepsilon}$ is based on $\hat{\lambda}_i^{\varepsilon}$ and $\hat{x}_i^{\varepsilon}$.

Based on the notation spelt out above, we define the test statistic for the time stability of Σ_{ε} and its eigensystem. Let

$$\tilde{S}^{\varepsilon}(\tau) = R \times D_{\lambda\gamma}^{\varepsilon} \times \left[S^{\varepsilon}(\tau) - \frac{\lfloor T\tau\rfloor}{T}S^{\varepsilon}(T)\right],$$

where $S^{\varepsilon}(\tau)$ is the CUSUM process defined as $\sum_{t=1}^{\lfloor T\tau\rfloor} vec\left(\hat{\varepsilon}_t\hat{\varepsilon}_t'\right)$, $S^{\varepsilon}(\tau) = 0$ for $\tau \leq \frac{1}{T}$ or $\geq 1-\frac{1}{T}$, and R is defined in (14). The test statistic is defined as

$$\Lambda_T^{\varepsilon}(\tau) = \sqrt{\frac{T}{\lfloor T\tau\rfloor \times \lfloor T(1-\tau)\rfloor}} \times \left[\tilde{S}^{\varepsilon}(\tau)'\left(\tilde{V}_{\varepsilon,\tau}\right)^{-1}\tilde{S}^{\varepsilon}(\tau)\right]^{1/2},$$

with $\tilde{V}_{\varepsilon,\tau} = RD_{\lambda\gamma}^{\varepsilon}\tilde{V}_{\tau}^{\varepsilon}D_{\lambda\gamma}^{\varepsilon\prime}R'$.

Let $\Lambda_T^{\varepsilon *} \equiv \sup_{nq \leq \lfloor T\tau \rfloor \leq T-nq} \Lambda_T^{\varepsilon}(\tau)$. The following theorem summarizes the asymptotics of $\Lambda_T^{\varepsilon *}$ under the null and under the alternative. Similarly to Theorem 4, we define the critical value $c_{\alpha,T}$ to be such that, under H_0, $P\left[\Lambda_T^{\varepsilon *} \leq c_{\alpha,T}\right] = 1-\alpha$ for some $\alpha \in [0,1]$.

Theorem 5 *Let Assumptions R1-R3 hold, and let p be defined as in Theorem 3. Under the null, as $(m,T) \to \infty$ with $\frac{\sqrt{\ln\ln T}}{m} + \frac{m}{\sqrt{T}} \ln\ln T \to 0$, it holds that*

$$P\left\{a_T \Lambda_T^{\varepsilon *} \leq x + b_T\right\} \to e^{-2e^{-x}},$$

where $a_T = \sqrt{2\ln\ln T}$ and $b_T = 2\ln\ln T + \frac{p}{2}\ln\ln\ln T - \ln\Gamma\left(\frac{p}{2}\right)$, with $\Gamma(\cdot)$ the Gamma function. Under the alternative (18) for Σ_ε, if (19) holds, $P\left[\Lambda_T^{\varepsilon *} > c_{\alpha,T}\right] = 1$.

Remarks

T5.1 The main feature of Theorem 5 is that the same theory applies to residual-based tests as to observable data. This is due to the fact that, under our assumptions, an almost sure rate of convergence can be derived for the OLS estimator $\hat{\beta}$, with $\hat{\beta} - \beta = O_{a.s.}\left(\sqrt{\frac{\ln\ln T}{T}}\right)$. By virtue of this result, a SIP and a SLLN can be shown for the partial sums of the squared residuals, so that the asymptotics of tests based on $\Lambda_T^{\varepsilon}(\tau)$ can be derived in the same way as for tests based on $\Lambda_T(\tau)$.

T5.2 Theorem 5 can be readily extended to accommodate for the presence of linear or polynomial trends among the x_ts; the only thing that changes is the rate of convergence of $\hat{\beta}$, which in presence of linear trends is actually faster than $O_{a.s.}\left(\sqrt{\frac{\ln\ln T}{T}}\right)$. A consequence of Theorem 5 is that the test developed in Section 3 can be applied to data with nonzero mean and/or trends: it suffices to filter out such deterministics, and then apply the test to the resulting residuals.

T5.3 On a similar note to Remark T5.2, the test in Section 3 is developed for the case of $E(y_t)$ being constant over time. As a consequence of Theorem 5, this does not need to be the case: in case of temporal shifts in the mean of y_t, it suffices to demean y_t by running (21) with time dummies corresponding to the dates at which $E(y_t)$ changes, and then applying the test to the residuals.

4 MONTE CARLO SIMULATIONS

In this section, we discuss: *(a)* the calculation of critical values, and *(b)* size and power of the test. Our experiments refer to the small sample properties of the tests developed in Section 2, i.e. to the context of testing for changes in the eigensystem of a matrix of fixed dimension.

There are two possible approaches to the computation of critical values: either using the EV distribution in (17) or using an approximation similar to that proposed in Csorgo and Horvath

(1997, Section 1.3.2); other approaches, such as bootstrap, are also possible (see Aue et al., 2012).

Direct computation of critical values $c_{\alpha,T}$ for a test of level α is based on $c_{\alpha,T} = a_T^{-1} \left\{ b_T - \ln\left[-\frac{1}{2} \ln(1-\alpha)\right]\right\}$. Thus, critical values only depend on p and T. It is well known that convergence to the EV distribution is usually very slow, which hampers the quality of $c_{\alpha,T}$. Alternatively, critical values can be simulated from

$$P\left\{ \sup_{h_{nT} \leq \tau \leq 1-h_{nT}} \left[\sum_{i=1}^{p} \frac{B_{1,i}^2(\tau)}{\tau(1-\tau)} \right]^{1/2} \leq c'_{\alpha,T} \right\} = 1-\alpha, \tag{22}$$

where the $B_{1,i}(\tau)$s are independent, univariate, squared Brownian bridges, generated over a grid of dimension T. We set $T \times h_{nT} = \max\left\{n, \ln^{3/2} T\right\}$. The "time series" part of this bound (i.e. the $\ln^{3/2} T$ part) is based on Csorgo and Horvath (1997, p. 25), who show that computing the maxima over restricted intervals (specifically, by truncating at $T \times h_{nT} = \ln^{3/2} T$) yields tests with good size properties; in our simulations, we have tried other solutions to restrict the interval over which the maximum is taken, but truncating at $\ln^{3/2} T$ yielded the best size properties. In addition to this, due to the multivariate nature of the problem, we also need to truncate at n; this is in order to have full rank estimated covariance matrices. In view of this, critical values $c'_{\alpha,T}$ are to be simulated for a given combination of p, n and T.

4.1 Finite sample properties of the test

We evaluate size and power through a Monte Carlo exercise. As a benchmark, we consider a test for a change in the first eigenvalue of the covariance matrix. Unreported simulations show that the finite sample performance of tests for changes in the other eigenvalues are very similar. In order to evaluate the impact of large n on finite sample properties, we also report a smaller Monte Carlo exercise applied to testing for changes in the whole covariance matrix Σ. We also assess, by way of comparison, the size and power of Aue et al.'s (2009) test, which is based on

$$\sup_{1 \leq \lfloor T\tau \rfloor \leq T} \Lambda_T(\tau) = \sup_{1 \leq \lfloor T\tau \rfloor \leq T} \frac{1}{\sqrt{T}} \times \left[\tilde{S}(\tau)' \tilde{V}_{z,\tau}^{-1} \tilde{S}(\tau) \right]^{1/2}. \tag{23}$$

Data are generated as follows. In order to avoid dependence on initial conditions, $T + 1000$ data are generated, discarding the first 1000 observations. We carry out our simulations for $T = \{50, 100, 200, 500\}$ and $n = \{3, 5, 7, 10, 15, 20\}$. Serial dependence in y_t is introduced through an ARMA(1,1) process:

$$y_t = \rho y_{t-1} + e_t + \theta e_{t-1}, \tag{24}$$

where $e_t \sim N(0, I_n)$ (see below for more details about simulations under the alternative). We conduct our experiments for the cases $(\rho, \theta) = \{(0,0), (0.5,0), (0,0.5), (0,-0.5), (0.5,0.5)\}$. Evidence from other experiments shows that little changes when the covariance matrix of e_t is non-diagonal, or when it has different elements on the main diagonal. All experiments have been conducted using the long run variance estimator in (3), based on full sample estimation of the autocovariance matrices with $m = T^{2/5}$. Other simulations show a heavy dependence of the results on m; in general, the larger m, the more conservative the test.

Finally, the number of replications is 2000; all routines are written using Gauss 10.

Size

We calculate the empirical rejection frequencies for tests of level 5%. We base the test on

$$\sup_{Th_{nT} \leq \lfloor T\tau \rfloor \leq T - Th_{nT}} \Lambda_T(\tau). \tag{25}$$

This is in accordance with the approximation of the critical values discussed above. Whilst this procedure no longer yields power versus breaks occurring $O(\ln \ln T)$ periods from the beginning or the end of the sample, however the test retains power versus breaks occurring $O(Th_{nT})$ periods from the beginning/end of sample.

[Insert Table 1 somewhere here]

The test (see Table 1) is, in general, undersized in small samples; this tends to disappear as T increases, with empirical rejection frequencies belonging, in general, to the interval $[0.04, 0.06]$ with few exceptions. The test has the correct size for large samples. Considering the *i.i.d.* case, the nominal size level is attained for $T = 200$ or larger. In general, as far as the presence of time dependence is concerned, this does not seem to affect the size properties of the test in a strong way, as it only makes the size slightly worse (with a tendency towards oversizement). The table also shows that higher values of n have a slight tendency to reduce the size. Similar results are found with Aue et al.'s (2009) test based on (23), which, in small samples, tends to be slightly more conservative.

Power

We conduct our simulations under alternatives defined as

$$\begin{cases} \Sigma \text{ for } t = 1, ..., k \\ \Sigma + \Delta \text{ for } t = k+1, ..., T \end{cases} \tag{26}$$

with $k = \lfloor \frac{T}{2} \rfloor$ and

$$\Delta = \sqrt{\frac{\ln \ln T}{T^{\nu}}} \times I_n. \tag{27}$$

We set $\nu = \left\{ \frac{2}{3}, \frac{1}{2} \right\}$ in Tables 2a and 2b respectively. In Table 2c, we also report power versus alternatives close to the beginning of the sample, with $k = Th_{nT} + 1$ and $\Delta = I_n$.

[Insert Tables 2a-2b somewhere here]

Considering mid-sample alternatives (Tables 2a and 2b), the test has nontrivial power versus "local" alternatives (represented here by the case $\nu = \frac{2}{3}$), and good power when $\nu = \frac{1}{2}$; the power becomes higher than 50%, in general, when T is larger than 200. As n increases, the test becomes increasingly powerful for all the cases considered (sample size T and dynamics in the error term); in general, the power of the test is not affected by the presence of MA disturbances, although a reduction in power is seen in presence of autoregressive roots. As predicted by the theory, tests based on (23) have better power properties, at least for $T = 200$ or higher. Indeed, tests based on (25) have very similar (in fact, slightly higher) power for $T = 100$; in that case, however, the factor $\ln \ln T$, by which the power of tests based on (25) is reduced with respect to tests based on (23), is negligible.

[Insert Tables 2c-2d somewhere here]

Tables 2c and 2d consider the power of the test versus finite alternatives that are close to the beginning of the sample. In particular, Table 2c reports the power under the alternative that the breakdate is close to the boundary. As predicted by the theory, there is power versus such alternatives. Referring the *i.i.d.* case as a benchmark, the power becomes higher than 50% when $T = 500$. As also observed in Tables 2a and 2b, as n increases the power slightly increases. As is natural, tests based on (23) have very little power versus beginning of sample alternatives; by construction, such tests do not have power versus changes that occur closer than $O\left(\sqrt{T}\right)$ periods to the beginning (or the end) of the sample. Indeed, the presence of some, limited power is to be interpreted as $\sqrt{T}$ and $k = Th_{nT}+$ being of comparable magnitude. As far as the impact of serial correlation is concerned, the power deteriorates in presence of AR roots, which is even more evident in the ARMA case. A less dramatic power reduction is also observed in presence of MA roots. This is due to the use of the full sample estimator of the long run covariance matrix, $\tilde{V}_\Sigma$, which is consistent under the alternative but biased. Unreported simulations show that tests based on (23) and (25) exhibit a substantial power gain when using the partial sample estimates, $\tilde{V}_{\Sigma,\tau}$. However, both tests become grossly oversized. Therefore, as a guideline for empirical applications, we suggest that if an AR structure is found in the data, pre-whitening should be applied.

As far as Table 2d is concerned, by way of comparison we report the evolution of the power of both tests, based on (25) and (23), as the changepoint moves from the beginning of the sample towards the middle. Results are reported for *i.i.d.* data only; when considering serial correlation, no changes were noted in the results, save for a general decrease in power (see also the comments to Table 2c), and we therefore omit them to save space. The alternatives considered have the following times of change: $k = Th_{nT} + 1$ (this is the same as in Table 2c;

indeed, the first column of Table 2d is the same as the first column of Table 2c); $k = \frac{1}{2}(\ln T)^2$; $k = \frac{1}{2}(\ln T)^{5/2}$ and $k = 3\sqrt{T}$. We note little variation in the power of both tests between the experiments when comparing the first and the second alternative. The third column, containing the case $k = (\ln T)^{5/2}$, shows a different pattern. Tests based on (23) are more powerful for small samples; in this case, however, $\frac{1}{2}(\ln T)^{5/2}$ is usually larger than $\sqrt{T}$, which is the threshold after which tests based on (23) become powerful. When T increases (e.g. when $T = 200$ and, more noticeably, 500), tests based on (25) have more power than tests based on (23). The power of tests based on (25) becomes very similar, and in fact worse in some cases, to that of tests based on (23) when considering alternatives closer to the mid-sample, i.e. when $k = 3\sqrt{T}$. In this case, it can be expected that tests based on (23) have nontrivial power, which is indeed the case. However, it can be noted that tests based on (25) perform very well, indeed (marginally) better, than tests based on (23) when T is large. Finally, as noted for Table 2c, the power of tests based on (25) increases monotonically with T across all experiments, and it also has a tendency to increase as n increases, albeit not monotonically.

Testing for the constancy of Σ: the role of n

We report a simulation exercise for the null of no change in Σ. Data are generated as Gaussian, with the same scheme as in (24). We report results for $n = \{3, 4, 5, 6, 7\}$ and $T = \{50, 100, 200, 500\}$. When generating data under the alternative, this is defined as in (26) and (27). In the last column, $k = Th_{nT} + 1$ and $\Delta = I_n$.

[Insert Tables 3a-3d somewhere here]

Table 3 illustrates the role played by n. As n increases, the test becomes increasingly conservative in finite samples; however, as $T \to \infty$, the empirical rejection frequencies approach their nominal values. As far as the power is concerned, when considering mid-sample alternatives, the power increases monotonically with T as expected, whereas it decreases with n. Such dependence on n is expected, since as n increases a larger trimming occurs, which entails a loss of information. However, for large samples ($T = 500$), we note that as n increases, the power of the test also increases. The results in Table 3b and 3c could be contrasted with those in Tables 2a-2b: under mid-sample alternatives, tests for the stability of the whole matrix are more powerful than tests based on the individual eigenvalues, at least in small samples; when $T = 500$, the figures become comparable. A heuristic explanation of this is in Remark T4.2: under the alternative, the long run variance of the estimated eigenvalues is affected by the eigenvectors, and thus it can be expected to be bounded in probability but biased.

Different considerations apply to the case of breaks closer to the beginning of the sample, i.e. Table 3d. In this case, as n increases, the power decreases. Again, this is likely to be due to the extent of the trimming, which always entails a loss of power. In addition to this, tests based on (25) are decidedly more powerful than tests based on (23) for large samples ($T = 500$). Comparing Table 3d with Table 2c, it can be noted that the power of tests for the whole matrix is lower (substantially, in the *i.i.d.* case) than that of tests for changes in the eigenvalues, which is a reversal of the power performance under mid-sample alternatives. Moreover, the power of tests for changes in the eigenvalues is not negatively affected by n; indeed, it tends to increase with n.

5 EMPIRICAL ILLUSTRATIONS USING INTEREST RATES AND EXCHANGE RATES

In this section, we report two empirical applications that illustrate the theory developed in Sections 2 and 3.1. Firstly, we consider a "classical" application of PCA to finance - namely, we investigate the time stability of the covariance matrix of the term structure of interest rates - Section 5.1. Secondly, we apply the theory developed in Section 3.1 to a VAR applied to exchange rates.

5.1 The time stability of the covariance matrix of interest rates

In this section, we apply the theory developed above to test for the stability of the covariance matrix of the term structure of interest rates, and to infer the sources of instability if present. Our analysis is motivated by the study in Perignon and Villa (2006), and follows similar steps.

As a first step, we investigate whether the "volatility curve" (i.e. the term structure of the volatility of interest rates) changes over time; this corresponds to testing for the stability of the main diagonal of the covariance matrix. Further, we verify whether the whole covariance matrix changes. This could be done by directly testing for the constancy of the matrix. Alternatively, in order to reduce the dimensionality of the problem, one could check whether the main three principal components (level, slope and curvature) are stable through time. We choose the latter approach, verifying separately, for each principal component, whether sources of time variation are in the loadings (i.e. the eigenvectors) or in the volatility (i.e. the eigenvalues), or both.

Previous studies have found evidence of changes in the yield curve. Using a descriptive approach based on splitting the sample at some predetermined points in time, indicated by stylised facts, Bliss (1997) finds that the eigenvectors of the covariance matrix of interest rates are quite stable, although the eigenvalues differ across subsamples. Perignon and Villa (2006), under the assumption that data are *i.i.d.* Gaussian, find evidence of changes in the volatilities (eigenvalues) of the principal components across four different subperiods (chosen *a priori*) in the time interval January 1960 - December 1999.

We conduct a similar exercise to Perignon and Villa (2006), relaxing the assumptions of *i.i.d.* Gaussian data, and avoiding the *a priori* selection of breakdates which could be rather arbitrary. We apply our test to US data, considering monthly and weekly frequencies, spanning from April 1997 to November 2010 (monthly - the sample size is $T_m = 164$) and from the second week of February 1997 to the last week of February 2011 (weekly - the sample size is $T_w = 733$); the number of maturities which we consider is $n = 18$, corresponding to (1m, 3m, 6m, 9m, 12m, 15m, 18m, 21m, 24m, 30m, 3y, 4y, 5y, 6y, 7y, 8y, 9y, 10y). Figure 1 reports the term structure in the period considered.

[Insert Figure 1 somewhere here]

Preliminary analysis shows that yields are highly persistent. Therefore, we use returns, which are found to be much less autocorrelated (particularly, they have no autocorrelation pattern for higher maturities). Table 4 shows some descriptive statistics for both monthly and weekly data; it is interesting to note how the large values of skewness and kurtosis of each maturity lead to reject the assumption of normality.

[Insert Table 4 somewhere here]

As far as the notation is concerned, y_t denotes, henceforth, the demeaned 18-dimensional vector of first-differenced maturities. Preliminary evidence based on the autocorrelation function of the squared returns shows that there is very little serial correlation, and, with higher maturities, no correlation at all. In light of this, we set the bandwidth, for the estimation of the long-run variance, as $m = \sqrt{\ln T}$ (see equation (3)).

The first step of our analysis is an evaluation of the stability of the variances of the first differenced maturities, i.e. of the elements on the main diagonal of $\Sigma = E\left(y_t y_t'\right)$. Instead of checking for the stability of the whole main diagonal, we test the volatilities one by one; this approach should be more constructive if the null of no changes were to be rejected, in that it would indicate which maturity changes and when. In order to control for the size of this multiple comparison, we propose a Bonferroni correction. We calculate the critical values for each test as $\alpha_I = \frac{\alpha_P}{n}$, where α_P is the size of the whole procedure. Using these critical values yields, approximately, a level α_P not greater than 1%, 5% and 10% corresponds to conducting each test at levels $\alpha_I = 0.056\%$, 0.28% and 0.56% respectively.

Critical values for individual tests of levels $\alpha_I = 1\%$, 5% and 10%, and for procedure level $\alpha_P = 1\%$, 5% and 10%, are in Table 5.

[Insert Table 5 somewhere here]

As a second step, we verify whether slope, level and curvature are constant over time. Particularly, we carry out separately the detection of changes in the volatility of the principal components (verifying the time stability of the three largest eigenvalues, say λ_1, λ_2 and λ_3), and in their loading (verifying the stability of the eigenvalue-normed eigenvectors corresponding to the three largest eigenvalues, denoted as γ_1, γ_2 and γ_3). As far as eigenvectors are concerned, (10) and (11) ensure that, when running the test, the CUSUM transformation of the estimated γ_is has the same sign for all values of τ, thus overcoming the issue of the eigenvectors being defined up to a sign.

Results for both experiments, at both frequencies, are reported in Table 6.

[Insert Table 6 somewhere here]

It is well known that controlling the procedure-wise error by a Bonferroni correction can be rather conservative. In our case, the values of test statistics (Table 6) can be contrasted with the critical values to be used for single hypothesis testing (reported in Table 5 as cv_1), which is the least conservative approach. When using a 5% level, results are exactly the same. The only exception is the test for the stability of the second eigenvector, γ_2, when using weekly data, where the null of no change is now rejected at 5%. A marginal discrepancy can be observed in the first panel of Table 6, when testing for the constancy of the diagonal elements of Σ with weekly data. When using cv_1 as critical values, two maturities (the 30 months and the 3 years ones) now appear to have a break. The rest of the results (especially the absence of breaks in monthly data) is the same as when using a Bonferroni correction.

Table 6 shows an interesting discrepancy between monthly and weekly data. Monthly data, as a whole, have a stable covariance structure over time: no changes are present either in the volatilities of the maturities, or in any of the principal components. Such result can be ascribed also to the test being less powerful in small samples, but Table 4 shows that the two datasets (monthly and weekly) do have different features, at least as far as descriptive statistics are concerned. As far as covariance instability is concerned, when monthly data are considered the only change is recorded in λ_3, the volatility of the curvature, which has a break significant at 10%. The second and third panel of the table show that the principal component structure has a change in the size of the curvature, significant at 5%. The corresponding estimated breakdate, selected as the maximizer of the CUSUM statistic, is January 2008. Table 7 reports the proportion of the total variance explained by each principal component before and after this date.

[Insert Table 7 somewhere here]

As far as weekly data are concerned, there is evidence of instability in the covariance structure. At a "macro" level, the variances of longer maturities (from 4 years onwards) change, whilst the variances of shorter maturities are constant. For most maturities, the breakdate is around the first week of December 2007. This is expected, since December 2007 is generally associated with the deepening of the recent recession. It is interesting to note that the longest maturity, the 10-year one, has a break at around the last week of August 2008. As far as principal components are concerned, the second panel of Table 6 shows that whilst the volatility of slope and curvature does not change over time, the loading of the level changes at the first week of December 2007, consistently with the findings for the variances. As the third panel of the table shows, the loadings of principal components are subject to change: the level and the curvature change significantly around the middle/end of March 2008 (possibly due to an "attraction" effect of the variance of the 10-year maturity); the slope has a significant break also, a few weeks later. The presence of significant changes in the loadings of each principal component as a result of the 2007-2009 recession is a different feature to what Perignon and Villa (2006) found in the time period they consider, when eigenvectors were not subject to changes over time.

5.2 A VAR model for exchange rates: the stability of the covariance matrix of the error term

In this section, we apply the test developed in Section 3.1 to a VAR, checking whether the covariance matrix of the error term is stable over time. Following Carriero et al. (2009), we consider a VAR consisting of 4 exchange rates vis-a-vis the US Dollar. Specifically, we consider monthly averages of the following currencies: Euro, British Pound, Yen, Canadian Dollar; data are collected from January 1999 to April 2013, which entails a sample size (modulo the first entry) of $T = 172$. All data are taken from Datastream.

Two tests are carried out in parallel. We verify whether the covariance matrix of the error term as a whole is subject to any changes, in a similar spirit to the simulations in Section 4 (Tables 3a-3d). Furthermore, we also test for the stability of each of the eigenvalues of the covariance matrix. Since we find very weak evidence of breaks (if any) in the whole matrix and in its spectrum, we do not carry out any analysis on the eigenvectors: the key finding in this section is that, despite a major event such as the 2007-2009 recession, the covariance matrix of the VAR model of exchange rates does not change over time.

Carriero et al. (2009) propose a Bayesian approach to model the exchange rates directly, imposing a prior on the VAR coefficients in order to shrink them towards a unit root representation. Preliminary analysis shows that, as expected, our series have a unit root. We therefore fit a VAR to the returns of the exchange rates. We report descriptive statistics for the returns, and a plot of the original series, in Table 8 and Figure 2 respectively.

[Insert Table 8 and Figure 2 somewhere here]

As the evidence in Carriero et al. (2009) shows, applying VAR models to exchange rates could yield a good level of forecasting accuracy, also in light of the fact that exchange rates have a tendency to co-move. However, as pointed out in the Introduction, Castle et al. (2010) find that when the covariance matrix of the error term in a regression changes, this could have pernicious effects on the forecasting ability of the model; thus, applying our test as discussed in Section 3.1 to the covariance matrix of the error term (and to its eigenvalues) could shed some light as to the forecasting ability of the VAR. Estimation results and mis-specification tests are reported in Table 9: the best specification was found to be a VAR of order 5, using lags 1, 2 and 5. As can be seen from the Table, the model is correctly specified, save for the lack of normality in the error term; in view of Assumption R1, this is not an issue when applying our test. We also note that, as can be expected as regards the residuals of a correctly specified model, the error term is found to be serially uncorrelated, and we therefore estimate the long run variance without adjusting for serial correlation.

[Insert Table 9 somewhere here]

The test is applied, preliminarily, to the (lower triangular part of the) whole covariance matrix: the results in Section 4 suggest that the dimension of the matrix ($n = 4$ in our context) affords the test to have the correct size and power versus the alternative of one or more breaks. Given that we consider only the lower triangular part of the matrix, this entails applying our test with $p = 10$ constraints under the null, as opposed to $p = 16$ constraints which is the case in the Monte Carlo exercise in Section 4.

We also verify the stability of each of the four eigenvalues of the matrix. As in Section 5.1, this procedure is implemented by controlling for the family-wise rejection rate though a Bonferroni correction. Critical values for individual tests of levels $\alpha_I = 1\%$, 5% and 10%, and for procedure level $\alpha_P = 1\%$, 5% and 10%, with the Bonferroni correction $\alpha_I = \frac{\alpha_P}{4}$ are in Table 5.

Results are reported in Table 10.

[Insert Table 10 somewhere here]

As can be seen from the first panel of the Table, the matrix, as a whole, is stable over time: as pointed out above, despite the presence of a major event such as the 2007-2009 recession, the covariance structure of the error term of the VAR does not change. This could be interpreted in light of the stylised fact that exchange rates tend to co-move (see the discussion in Carriero et al., 2009). This finding is also reinforced by the second panel of Table 10, where we find that the first three largest eigenvalues (denoted as $\lambda_1 - \lambda_3$) do not change over time. As far as the smallest one is concerned (λ_4), some evidence of a change is indeed found. However, the test rejects the null of a stable eigenvalue only when applied to each eigenvalue separately.

Conversely, when adjusting for the family-wise error, no break is found at 5% level (the null of no break is rejected at 10% level for λ_4); as mentioned in Section 5.1, this is a consequence of the Bonferroni correction yielding relatively conservative results. However, this suggests very weak evidence of a break in the smallest eigenvalue; the estimated breakdate is found to be around June 2008, which has the natural interpretation of being a consequence of the latest recession.

6 CONCLUSIONS

In this paper, we propose a test for the null of no breaks in the eigensystem of a covariance matrix. The assumptions under which we derive our results are sufficiently general to accommodate for a wide variety of datasets. We show that our test is powerful versus alternatives as close to the boundaries of the sample as $O(\ln \ln T)$. Results are extended to testing for the stability of the eigensystem. We also derive a correction for the finite sample bias when estimating eigenvalues and eigenvectors, which can be relatively severe for large n or small T. The theory is also extended to develop tests for the null of no change in the covariance matrix of the error term in a multivariate regression (including the case of VARs). As shown in Section 4, the properties of the test are satisfactory: the correct size is attained under various degrees of serial dependence, and the test exhibits good power.

The results in this paper suggest several avenues for research. The test discussed here is a stability test for the null of no change in a covariance matrix. Although we only consider one alternative, the test could be extended to be applied sequentially, i.e. by splitting the sample around an estimated breakdate and test for breaks in each subsample. Also, results are derived under the minimal assumption that the 4-th moment exists. Aue et al. (2009) provide a discussion as to how to proceed if this is not the case, which involves fractional transformations of the series, viz. y_{it}^{Δ} for some $\Delta \in (0,1)$, although the optimal choice of Δ is not straightforward. These issues are currently under investigation by the authors.

REFERENCES

Andrews, D.W.K., 1987. Asymptotic results for generalised Wald tests. Econometric Theory 3, 348-358.

Andrews, D.W.K., 1993. Tests for parameter instability and structural changes with unknown change point. Econometrica 61, 821-856.

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383-1414.

Audrino, F., Barone-Adesi, G., Mira, A., 2005. The stability of factor models of interest rates. Journal of Financial Econometrics 3, 422-441.

Aue, A., Hormann, S., Horvath, L., Reimherr. M., 2009. Break detection in the covariance structure of multivariate time series models. Annals of Statistics 37, 4046-4087.

Aue, A., Horvath, L., Huskova, M., 2012. Segmenting mean-nonstationary time series via trending regressions. Journal of Econometrics 168, 367-381.

Bai, J., 2000. Vector autoregressive models with structural changes in regression coefficients and in variance-covariance matrices. Annals of Economics and Finance 1, 303-339.

Bates, B.J., Plagborg-Muller, M., Stock, J.H., Watson, M.W., 2013. Consistent factor estimation in dynamic factor models with structural instability. Forthcoming, Journal of Econometrics.

Bliss, R.R., 1997. Movements in the term structure of interest rates. Economic Review, Federal Reserve Bank of Atlanta, Q4, 16-33.

Bliss, R.R., Smith, D.C., 1997. The stability of interest rate processes. Federal Reserve Bank of Atlanta, Working Paper 97-13.

Carriero, A., Kapetanios, G., Marcellino, M., 2009. Forecasting exchange rates with a large Bayesian VAR. International Journal of Forecasting 25, 400-417.

Castle, J.R., Fawcett, N., Hendry, D.F., 2010. Forecasting with equilibrium-correction models during structural breaks. Journal of Econometrics 158, 25-36.

Corradi, V., 1999. Deciding between $I(0)$ and $I(1)$ via FLIL-based bounds. Econometric Theory 15, 643-663.

Csorgo, M., Horvath, L., 1997. Limit Theorems in Change-point Analysis. Chichester: Wiley.

Davidson, J., 2002a. Stochastic Limit Theory. Oxford: Oxford University Press.

Davidson, J., 2002b. Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. Journal of Econometrics 106, 243-269.

Davis, C., Kahan, W., 1970. The rotation of eigenvectors by a perturbation. III. SIAM Journal of Numerical Analysis 7, 1-46.

Eberlein, E., 1986. On strong invariance principles under dependence assumptions. Annals of Probability 14, 260-270.

Flury, B.H., 1984. Common principal components in k groups. Journal of the American Statistical Association 79, 892-898.

Flury, B.H., 1988. Common principal components and related multivariate models. New York: Wiley.

Gallant, A.R., White, H., 1988. A unified theory of estimation and inference for nonlinear dynamic models. Oxford: Basil Blackwell.

Götze, F., Zaitsev, A., 2012. Estimates for the rate of strong approximation in Hilbert space. Mimeo.

Hill, J.B., 2010. On tail index estimation for dependent, heterogeneous data. Econometric Theory, 26, 1398-1436.

Hill, J.B., 2011. Tail and non-tail memory with applications to extreme value and robust statistics. Econometric Theory, 27, 844-884.

Horn, R.A., Johnson, C.R., 1999. Matrix Analysis. Cambridge: Cambridge University Press.

Horvath, L., 1993. The maximum likelihood method for testing changes in the parameters of normal observations. Annals of Statistics 21, 672-680.

Kollo, T., Neudecker, H., 1997. Asymptotics of Pearson-Hotelling principal-component vectors of sample variance and correlation matrices. Behaviormetrika 24, 51-59.

Ling, S., 2007. Testing for change points in time series models and limiting theorems for NED sequences. Annals of Statistics 35, 1213-1237.

Litterman, R., Scheinkman, J.A., 1991. Common factors affecting bond returns. Journal of Fixed Income 1, 54-61.

Perignon, C., Villa, C., 2006. Sources of time variation in the covariance matrix of interest rates. Journal of Business 79, 1535-1549.

Perron, P., 2006. Dealing with breaks. In Mills, T.C. and Patterson, K. (eds.). Palgrave Handbook of Econometrics: Vol. 1 Econometric Theory. Basingstoke: Palgrave Macmillan.

Pesaran, M.H., Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models. Economics Letters 58, 17-29.

Phillips, P.C.B., Solo, V., 1992. Asymptotics for linear processes. Annals of Statistics 20, 971-1001.

Stock, J. H., Watson, M.W., 1999. Forecasting inflation. Journal of Monetary Economics 44, 293-335.

Stock, J. H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167-1179.

Stock, J. H., Watson, M.W., 2005. Implications of dynamic factor models for VAR analysis. Manuscript.

Trapani, L., 2013. On bootstrapping panel factor series. Journal of Econometrics 172, 127-141.

Zhou, Y., 2011. On the eigenvalues of specially low-rank perturbed matrices. Applied Mathematics and Computation, 217, 10267-10270.

n	T		(ρ, ϑ)	$(0, 0)$	$(0.5, 0)$	$(0, 0.5)$	$(0, -0.5)$	$(0.5, 0.5)$
3	50	Test based on (25)		0.013	0.015	0.009	0.010	0.013
		Test based on (23)		0.006	0.002	0.003	0.005	0.004
	100	Test based on (25)		0.025	0.033	0.026	0.025	0.034
		Test based on (23)		0.012	0.017	0.014	0.014	0.021
	200	Test based on (25)		0.041	0.048	0.041	0.040	0.058
		Test based on (23)		0.029	0.023	0.025	0.029	0.020
	500	Test based on (25)		0.044	0.058	0.043	0.045	0.067
		Test based on (23)		0.034	0.027	0.029	0.033	0.030
5	50	Test based on (25)		0.010	0.011	0.011	0.010	0.012
		Test based on (23)		0.003	0.007	0.007	0.004	0.007
	100	Test based on (25)		0.024	0.030	0.029	0.030	0.036
		Test based on (23)		0.011	0.013	0.018	0.019	0.016
	200	Test based on (25)		0.032	0.041	0.032	0.042	0.047
		Test based on (23)		0.025	0.030	0.024	0.025	0.023
	500	Test based on (25)		0.044	0.063	0.049	0.050	0.065
		Test based on (23)		0.033	0.043	0.038	0.039	0.037
7	50	Test based on (25)		0.007	0.009	0.012	0.007	0.008
		Test based on (23)		0.003	0.005	0.005	0.004	0.003
	100	Test based on (25)		0.022	0.023	0.022	0.031	0.025
		Test based on (23)		0.015	0.009	0.010	0.011	0.011
	200	Test based on (25)		0.032	0.049	0.043	0.038	0.047
		Test based on (23)		0.023	0.029	0.029	0.021	0.030
	500	Test based on (25)		0.041	0.064	0.052	0.039	0.065
		Test based on (23)		0.046	0.043	0.040	0.034	0.042
10	50	Test based on (25)		0.004	0.006	0.004	0.002	0.003
		Test based on (23)		0.005	0.003	0.005	0.004	0.003
	100	Test based on (25)		0.018	0.034	0.029	0.019	0.038
		Test based on (23)		0.016	0.017	0.016	0.007	0.016
	200	Test based on (25)		0.030	0.053	0.040	0.029	0.053
		Test based on (23)		0.023	0.033	0.028	0.026	0.024
	500	Test based on (25)		0.044	0.063	0.057	0.051	0.065
		Test based on (23)		0.036	0.033	0.035	0.040	0.036
15	50	Test based on (25)		0.001	0.000	0.001	0.001	0.000
		Test based on (23)		0.004	0.003	0.001	0.004	0.003
	100	Test based on (25)		0.013	0.015	0.017	0.011	0.019
		Test based on (23)		0.014	0.011	0.010	0.013	0.013
	200	Test based on (25)		0.027	0.053	0.040	0.031	0.055
		Test based on (23)		0.024	0.032	0.025	0.026	0.031
	500	Test based on (25)		0.047	0.061	0.052	0.047	0.063
		Test based on (23)		0.028	0.035	0.038	0.038	0.038
20	50	Test based on (25)		0.000	0.001	0.001	0.000	0.001
		Test based on (23)		0.003	0.004	0.003	0.002	0.003
	100	Test based on (25)		0.007	0.009	0.010	0.008	0.011
		Test based on (23)		0.013	0.016	0.016	0.017	0.013
	200	Test based on (25)		0.017	0.033	0.020	0.023	0.039
		Test based on (23)		0.018	0.026	0.022	0.022	0.028
	500	Test based on (25)		0.040	0.058	0.047	0.057	0.059
		Test based on (23)		0.035	0.042	0.037	0.037	0.042

Table 1. Empirical rejection frequencies for the null of no changes in the largest eigenvalue of Σ. Data are generated according to (24). The empirical sizes have confidence interval $[0.04, 0.06]$.

n	T		(ρ, ϑ)	$\Delta = \sqrt{\frac{\ln\ln(T)}{T^{2/3}}}$ (0,0)	(0.5,0)	(0,0.5)	(0,−0.5)	(0.5,0.5)
3	50	Test based on (25)		0.035	0.021	0.027	0.034	0.016
		Test based on (23)		0.014	0.001	0.010	0.006	0.005
	100	Test based on (25)		0.133	0.121	0.128	0.122	0.112
		Test based on (23)		0.128	0.073	0.084	0.101	0.052
	200	Test based on (25)		0.235	0.180	0.211	0.191	0.167
		Test based on (23)		0.293	0.185	0.228	0.232	0.140
	500	Test based on (25)		0.427	0.302	0.371	0.335	0.259
		Test based on (23)		0.533	0.350	0.424	0.410	0.272
5	50	Test based on (25)		0.039	0.026	0.028	0.026	0.025
		Test based on (23)		0.016	0.016	0.020	0.015	0.012
	100	Test based on (25)		0.171	0.117	0.138	0.138	0.103
		Test based on (23)		0.153	0.099	0.121	0.118	0.075
	200	Test based on (25)		0.318	0.245	0.246	0.241	0.197
		Test based on (23)		0.374	0.244	0.291	0.265	0.186
	500	Test based on (25)		0.460	0.337	0.377	0.370	0.281
		Test based on (23)		0.584	0.395	0.469	0.449	0.309
7	50	Test based on (25)		0.040	0.034	0.037	0.031	0.033
		Test based on (23)		0.018	0.012	0.016	0.012	0.010
	100	Test based on (25)		0.210	0.154	0.149	0.168	0.132
		Test based on (23)		0.184	0.112	0.145	0.126	0.084
	200	Test based on (25)		0.344	0.247	0.257	0.290	0.205
		Test based on (23)		0.409	0.247	0.305	0.313	0.193
	500	Test based on (25)		0.516	0.339	0.387	0.407	0.288
		Test based on (23)		0.627	0.397	0.482	0.506	0.312
10	50	Test based on (25)		0.032	0.026	0.033	0.021	0.025
		Test based on (23)		0.020	0.012	0.021	0.015	0.008
	100	Test based on (25)		0.217	0.157	0.183	0.175	0.140
		Test based on (23)		0.203	0.116	0.150	0.151	0.080
	200	Test based on (25)		0.356	0.247	0.273	0.309	0.209
		Test based on (23)		0.440	0.248	0.328	0.336	0.192
	500	Test based on (25)		0.528	0.364	0.449	0.434	0.298
		Test based on (23)		0.661	0.430	0.536	0.537	0.347
15	50	Test based on (25)		0.006	0.001	0.004	0.008	0.003
		Test based on (23)		0.029	0.008	0.016	0.020	0.009
	100	Test based on (25)		0.211	0.136	0.165	0.178	0.121
		Test based on (23)		0.241	0.145	0.181	0.181	0.118
	200	Test based on (25)		0.440	0.315	0.365	0.341	0.281
		Test based on (23)		0.514	0.315	0.411	0.426	0.265
	500	Test based on (25)		0.623	0.390	0.483	0.493	0.350
		Test based on (23)		0.727	0.471	0.570	0.574	0.386
20	50	Test based on (25)		0.000	0.000	0.000	0.000	0.000
		Test based on (23)		0.035	0.014	0.021	0.015	0.014
	100	Test based on (25)		0.211	0.126	0.164	0.149	0.099
		Test based on (23)		0.280	0.152	0.204	0.208	0.117
	200	Test based on (25)		0.486	0.296	0.377	0.383	0.278
		Test based on (23)		0.585	0.311	0.447	0.466	0.270
	500	Test based on (25)		0.644	0.401	0.506	0.512	0.403
		Test based on (23)		0.764	0.493	0.594	0.603	0.426

Table 2a. Power of the test under the null of no changes in the largest eigenvalue of Σ. Data are generated according to (24) and under the alternative hypothesis as in (27).

n	T		$\Delta = \sqrt{\frac{\ln\ln(T)}{T^{1/2}}}$				
		(ρ, ϑ)	$(0,0)$	$(0.5,0)$	$(0,0.5)$	$(0,-0.5)$	$(0.5,0.5)$
3	50	Test based on (25)	0.055	0.030	0.045	0.050	0.023
		Test based on (23)	0.020	0.004	0.021	0.013	0.008
	100	Test based on (25)	0.272	0.188	0.213	0.213	0.164
		Test based on (23)	0.258	0.136	0.182	0.198	0.100
	200	Test based on (25)	0.514	0.380	0.450	0.415	0.298
		Test based on (23)	0.593	0.413	0.510	0.486	0.306
	500	Test based on (25)	0.874	0.652	0.753	0.760	0.524
		Test based on (23)	0.934	0.762	0.850	0.836	0.625
5	50	Test based on (25)	0.070	0.041	0.056	0.045	0.042
		Test based on (23)	0.028	0.022	0.029	0.020	0.018
	100	Test based on (25)	0.314	0.206	0.256	0.259	0.181
		Test based on (23)	0.314	0.168	0.255	0.220	0.119
	200	Test based on (25)	0.616	0.420	0.505	0.508	0.356
		Test based on (23)	0.717	0.447	0.584	0.579	0.349
	500	Test based on (25)	0.900	0.704	0.810	0.785	0.568
		Test based on (23)	0.955	0.795	0.892	0.879	0.658
7	50	Test based on (25)	0.073	0.043	0.062	0.063	0.048
		Test based on (23)	0.038	0.019	0.039	0.028	0.013
	100	Test based on (25)	0.362	0.253	0.286	0.295	0.219
		Test based on (23)	0.385	0.200	0.292	0.274	0.161
	200	Test based on (25)	0.679	0.467	0.526	0.585	0.401
		Test based on (23)	0.764	0.513	0.615	0.631	0.406
	500	Test based on (25)	0.916	0.702	0.815	0.846	0.592
		Test based on (23)	0.949	0.795	0.871	0.915	0.687
10	50	Test based on (25)	0.056	0.040	0.053	0.039	0.043
		Test based on (23)	0.037	0.017	0.030	0.027	0.009
	100	Test based on (25)	0.406	0.252	0.308	0.299	0.222
		Test based on (23)	0.405	0.215	0.300	0.308	0.161
	200	Test based on (25)	0.711	0.465	0.566	0.598	0.403
		Test based on (23)	0.796	0.524	0.648	0.674	0.419
	500	Test based on (25)	0.954	0.768	0.889	0.866	0.665
		Test based on (23)	0.979	0.856	0.943	0.920	0.748
15	50	Test based on (25)	0.016	0.008	0.010	0.014	0.004
		Test based on (23)	0.053	0.018	0.027	0.037	0.013
	100	Test based on (25)	0.385	0.234	0.308	0.332	0.207
		Test based on (23)	0.466	0.241	0.355	0.357	0.200
	200	Test based on (25)	0.800	0.571	0.666	0.681	0.502
		Test based on (23)	0.867	0.602	0.733	0.764	0.496
	500	Test based on (25)	0.964	0.811	0.907	0.903	0.707
		Test based on (23)	0.987	0.883	0.948	0.954	0.801
20	50	Test based on (25)	0.000	0.000	0.000	0.000	0.000
		Test based on (23)	0.049	0.024	0.032	0.037	0.019
	100	Test based on (25)	0.431	0.224	0.311	0.290	0.184
		Test based on (23)	0.553	0.282	0.389	0.371	0.227
	200	Test based on (25)	0.835	0.572	0.720	0.711	0.508
		Test based on (23)	0.906	0.610	0.787	0.787	0.499
	500	Test based on (25)	0.980	0.825	0.931	0.932	0.741
		Test based on (23)	0.992	0.902	0.966	0.971	0.827

Table 2b. Power of the test for the null of no changes in the largest eigenvalue of Σ. Data are generated according to (24) and under the alternative hypothesis as in (27).

n	T		$k = 2 \times [\ln(T)]^{3/2}$				
		(ρ, ϑ)	$(0,0)$	$(0.5,0)$	$(0,0.5)$	$(0,-0.5)$	$(0.5,0.5)$
3	50	Test based on (25)	0.071	0.023	0.025	0.027	0.030
		Test based on (23)	0.054	0.040	0.049	0.039	0.034
	100	Test based on (25)	0.177	0.147	0.143	0.156	0.086
		Test based on (23)	0.162	0.100	0.108	0.106	0.111
	200	Test based on (25)	0.485	0.169	0.234	0.222	0.140
		Test based on (23)	0.224	0.151	0.183	0.185	0.101
	500	Test based on (25)	0.886	0.226	0.501	0.467	0.223
		Test based on (23)	0.156	0.127	0.129	0.125	0.112
5	50	Test based on (25)	0.018	0.024	0.027	0.019	0.022
		Test based on (23)	0.067	0.044	0.059	0.036	0.038
	100	Test based on (25)	0.211	0.120	0.134	0.141	0.090
		Test based on (23)	0.156	0.067	0.092	0.107	0.060
	200	Test based on (25)	0.549	0.169	0.242	0.254	0.164
		Test based on (23)	0.261	0.157	0.193	0.181	0.136
	500	Test based on (25)	0.888	0.212	0.497	0.466	0.221
		Test based on (23)	0.158	0.144	0.156	0.107	0.124
7	50	Test based on (25)	0.014	0.028	0.026	0.019	0.031
		Test based on (23)	0.057	0.048	0.051	0.043	0.037
	100	Test based on (25)	0.202	0.125	0.154	0.148	0.090
		Test based on (23)	0.186	0.078	0.111	0.102	0.069
	200	Test based on (25)	0.536	0.152	0.253	0.247	0.117
		Test based on (23)	0.247	0.145	0.177	0.185	0.103
	500	Test based on (25)	0.909	0.231	0.488	0.477	0.226
		Test based on (23)	0.194	0.151	0.175	0.138	0.131
10	50	Test based on (25)	0.012	0.020	0.017	0.019	0.016
		Test based on (23)	0.071	0.033	0.053	0.056	0.029
	100	Test based on (25)	0.187	0.114	0.135	0.121	0.099
		Test based on (23)	0.158	0.089	0.103	0.089	0.085
	200	Test based on (25)	0.516	0.158	0.225	0.227	0.155
		Test based on (23)	0.262	0.129	0.180	0.180	0.092
	500	Test based on (25)	0.915	0.262	0.495	0.488	0.251
		Test based on (23)	0.179	0.114	0.126	0.153	0.099
15	50	Test based on (25)	0.011	0.006	0.009	0.006	0.005
		Test based on (23)	0.067	0.031	0.054	0.043	0.030
	100	Test based on (25)	0.182	0.101	0.122	0.142	0.084
		Test based on (23)	0.174	0.055	0.072	0.086	0.050
	200	Test based on (25)	0.558	0.160	0.231	0.244	0.156
		Test based on (23)	0.250	0.135	0.182	0.179	0.114
	500	Test based on (25)	0.919	0.246	0.464	0.449	0.253
		Test based on (23)	0.155	0.106	0.126	0.133	0.107
20	50	Test based on (25)	0.017	0.001	0.003	0.002	0.001
		Test based on (23)	0.074	0.034	0.039	0.037	0.003
	100	Test based on (25)	0.194	0.100	0.142	0.142	0.091
		Test based on (23)	0.148	0.048	0.071	0.072	0.041
	200	Test based on (25)	0.588	0.167	0.225	0.225	0.162
		Test based on (23)	0.249	0.121	0.171	0.183	0.079
	500	Test based on (25)	0.931	0.259	0.475	0.444	0.261
		Test based on (23)	0.161	0.124	0.146	0.124	0.107

Table 2c. Power of the test for the null of no changes in the largest eigenvalue of Σ. Data are generated according to (24), under the alternative specified as in (27).

n	T					
		k	$Th_{nT}+1$	$\frac{1}{2}\left[\ln(T)\right]^2$	$\frac{1}{2}\left[\ln(T)\right]^{5/2}$	$3\sqrt{T}$
3	50	Test based on (25)	0.071	0.017	0.017	0.059
		Test based on (23)	0.054	0.037	0.037	0.162
	100	Test based on (25)	0.177	0.182	0.088	0.198
		Test based on (23)	0.162	0.142	0.164	0.413
	200	Test based on (25)	0.485	0.488	0.304	0.609
		Test based on (23)	0.224	0.171	0.249	0.614
	500	Test based on (25)	0.886	0.834	0.904	0.997
		Test based on (23)	0.156	0.157	0.306	0.847
5	50	Test based on (25)	0.018	0.022	0.022	0.049
		Test based on (23)	0.067	0.052	0.052	0.178
	100	Test based on (25)	0.211	0.177	0.091	0.230
		Test based on (23)	0.156	0.152	0.172	0.442
	200	Test based on (25)	0.549	0.534	0.362	0.702
		Test based on (23)	0.261	0.195	0.291	0.704
	500	Test based on (25)	0.888	0.851	0.915	0.998
		Test based on (23)	0.158	0.167	0.312	0.877
7	50	Test based on (25)	0.014	0.015	0.015	0.055
		Test based on (23)	0.057	0.048	0.048	0.185
	100	Test based on (25)	0.202	0.187	0.110	0.254
		Test based on (23)	0.186	0.135	0.169	0.474
	200	Test based on (25)	0.536	0.546	0.358	0.683
		Test based on (23)	0.247	0.184	0.273	0.689
	500	Test based on (25)	0.909	0.914	0.930	0.997
		Test based on (23)	0.194	0.189	0.353	0.900
10	50	Test based on (25)	0.012	0.010	0.010	0.061
		Test based on (23)	0.071	0.057	0.057	0.234
	100	Test based on (25)	0.187	0.174	0.083	0.237
		Test based on (23)	0.158	0.137	0.159	0.474
	200	Test based on (25)	0.516	0.561	0.375	0.709
		Test based on (23)	0.262	0.202	0.292	0.718
	500	Test based on (25)	0.915	0.936	0.942	1.000
		Test based on (23)	0.179	0.195	0.356	0.905
15	50	Test based on (25)	0.011	0.009	0.009	0.047
		Test based on (23)	0.067	0.064	0.640	0.219
	100	Test based on (25)	0.182	0.184	0.094	0.246
		Test based on (23)	0.174	0.130	0.151	0.485
	200	Test based on (25)	0.558	0.596	0.422	0.762
		Test based on (23)	0.250	0.197	0.296	0.757
	500	Test based on (25)	0.919	0.936	0.957	1.000
		Test based on (23)	0.155	0.176	0.337	0.921
20	50	Test based on (25)	0.017	0.004	0.004	0.033
		Test based on (23)	0.074	0.630	0.630	0.251
	100	Test based on (25)	0.194	0.187	0.107	0.276
		Test based on (23)	0.148	0.146	0.184	0.535
	200	Test based on (25)	0.588	0.504	0.471	0.782
		Test based on (23)	0.249	0.191	0.336	0.787
	500	Test based on (25)	0.931	0.899	0.972	0.999
		Test based on (23)	0.161	0.175	0.369	0.935

Table 2d. Power of the test for the null of no changes in the largest eigenvalue of Σ. Data are generated as i.i.d., under the alternative specified in (27); the values of the time of change k are tabulated in the second row of the table.

n	T		(ρ, ϑ)	$(0, 0)$	$(0.5, 0)$	$(0, 0.5)$	$(0, -0.5)$	$(0.5, 0.5)$
3	50	Test based on (25)		0.047	0.000	0.000	0.000	0.000
		Test based on (23)		0.052	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.045	0.000	0.000	0.000	0.000
		Test based on (23)		0.043	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.056	0.036	0.038	0.031	0.035
		Test based on (23)		0.044	0.024	0.023	0.028	0.029
	500	Test based on (25)		0.062	0.044	0.041	0.040	0.041
		Test based on (23)		0.042	0.040	0.038	0.040	0.040
4	50	Test based on (25)		0.005	0.000	0.000	0.000	0.000
		Test based on (23)		0.009	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.032	0.000	0.000	0.000	0.000
		Test based on (23)		0.021	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.050	0.030	0.032	0.031	0.028
		Test based on (23)		0.023	0.013	0.016	0.020	0.014
	500	Test based on (25)		0.065	0.042	0.041	0.039	0.040
		Test based on (23)		0.037	0.032	0.030	0.031	0.032
5	50	Test based on (25)		0.002	0.000	0.000	0.000	0.000
		Test based on (23)		0.001	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.028	0.000	0.000	0.000	0.000
		Test based on (23)		0.019	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.047	0.031	0.024	0.025	0.026
		Test based on (23)		0.020	0.010	0.009	0.010	0.010
	500	Test based on (25)		0.059	0.038	0.038	0.038	0.068
		Test based on (23)		0.029	0.032	0.032	0.028	0.023
6	50	Test based on (25)		0.001	0.000	0.000	0.000	0.000
		Test based on (23)		0.002	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.016	0.000	0.000	0.000	0.000
		Test based on (23)		0.009	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.037	0.034	0.030	0.029	0.029
		Test based on (23)		0.021	0.018	0.012	0.013	0.016
	500	Test based on (25)		0.065	0.041	0.041	0.040	0.042
		Test based on (23)		0.040	0.039	0.036	0.035	0.036
7	50	Test based on (25)		0.000	0.000	0.000	0.000	0.000
		Test based on (23)		0.000	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.008	0.000	0.000	0.000	0.000
		Test based on (23)		0.003	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.040	0.037	0.033	0.031	0.030
		Test based on (23)		0.022	0.024	0.029	0.032	0.030
	500	Test based on (25)		0.071	0.056	0.040	0.041	0.044
		Test based on (23)		0.039	0.032	0.030	0.040	0.040

Table 3a. Empirical rejection frequencies for the null of no change in Σ. Data are generated according to (24). The empirical sizes reported here have confidence interval $[0.04, 0.06]$.

n	T		(ρ, ϑ)	$(0, 0)$	$(0.5, 0)$	$(0, 0.5)$	$(0, -0.5)$	$(0.5, 0.5)$
3	50	Test based on (25)		0.490	0.224	0.201	0.276	0.184
		Test based on (23)		0.486	0.221	0.183	0.200	0.211
	100	Test based on (25)		0.573	0.320	0.356	0.312	0.344
		Test based on (23)		0.551	0.290	0.288	0.311	0.316
	200	Test based on (25)		0.628	0.502	0.551	0.476	0.521
		Test based on (23)		0.605	0.471	0.503	0.524	0.478
	500	Test based on (25)		0.724	0.618	0.576	0.555	0.598
		Test based on (23)		0.737	0.580	0.554	0.583	0.523
4	50	Test based on (25)		0.133	0.000	0.000	0.002	0.000
		Test based on (23)		0.112	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.401	0.165	0.197	0.201	0.188
		Test based on (23)		0.313	0.111	0.123	0.154	0.165
	200	Test based on (25)		0.567	0.444	0.462	0.504	0.406
		Test based on (23)		0.547	0.391	0.412	0.411	0.427
	500	Test based on (25)		0.824	0.778	0.794	0.698	0.752
		Test based on (23)		0.845	0.821	0.786	0.802	0.739
5	50	Test based on (25)		0.092	0.000	0.000	0.000	0.000
		Test based on (23)		0.061	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.407	0.203	0.222	0.241	0.210
		Test based on (23)		0.313	0.216	0.193	0.200	0.214
	200	Test based on (25)		0.624	0.509	0.512	0.570	0.552
		Test based on (23)		0.579	0.476	0.500	0.504	0.488
	500	Test based on (25)		0.851	0.898	0.723	0.748	0.774
		Test based on (23)		0.867	0.806	0.816	0.799	0.783
6	50	Test based on (25)		0.031	0.000	0.000	0.000	0.000
		Test based on (23)		0.017	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.381	0.271	0.134	0.120	0.136
		Test based on (23)		0.235	0.162	0.156	0.194	0.207
	200	Test based on (25)		0.668	0.701	0.504	0.536	0.499
		Test based on (23)		0.593	0.561	0.512	0.498	0.507
	500	Test based on (25)		0.899	0.734	0.805	0.888	0.855
		Test based on (23)		0.892	0.862	0.797	0.913	0.828
7	50	Test based on (25)		0.000	0.000	0.000	0.000	0.000
		Test based on (23)		0.000	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.293	0.098	0.103	0.061	0.084
		Test based on (23)		0.186	0.150	0.076	0.077	0.100
	200	Test based on (25)		0.671	0.680	0.606	0.631	0.577
		Test based on (23)		0.534	0.533	0.612	0.598	0.562
	500	Test based on (25)		0.921	1.000	0.903	0.926	0.947
		Test based on (23)		0.903	1.000	0.945	0.888	0.912

Table 3b. Power of the test for the null of no changes in Σ. Data are generated according to (24) and under the alternative hypothesis specified in (27) with $\nu = \frac{2}{3}$.

n	T		(ρ, ϑ)	$(0,0)$	$(0.5,0)$	$(0,0.5)$	$(0,-0.5)$	$(0.5,0.5)$
3	50	Test based on (25)		0.736	0.600	0.598	0.567	0.616
		Test based on (23)		0.721	0.555	0.599	0.602	0.589
	100	Test based on (25)		0.855	0.778	0.756	0.812	0.734
		Test based on (23)		0.833	0.803	0.799	0.787	0.831
	200	Test based on (25)		0.944	0.906	0.912	0.836	0.899
		Test based on (23)		0.939	0.888	0.847	0.935	0.921
	500	Test based on (25)		0.995	1.000	1.000	1.000	1.000
		Test based on (23)		0.996	1.000	1.000	1.000	1.000
4	50	Test based on (25)		0.270	0.136	0.122	0.146	0.151
		Test based on (23)		0.227	0.204	0.231	0.218	0.199
	100	Test based on (25)		0.712	0.741	0.632	0.681	0.658
		Test based on (23)		0.675	0.608	0.661	0.639	0.655
	200	Test based on (25)		0.932	0.899	0.873	0.881	0.876
		Test based on (23)		0.937	0.936	0.931	0.952	0.908
	500	Test based on (25)		1.000	1.000	1.000	1.000	1.000
		Test based on (23)		1.000	1.000	1.000	1.000	1.000
5	50	Test based on (25)		0.204	0.099	0.072	0.070	0.101
		Test based on (23)		0.135	0.100	0.101	0.122	0.162
	100	Test based on (25)		0.729	0.701	0.698	0.732	0.677
		Test based on (23)		0.686	0.678	0.696	0.648	0.699
	200	Test based on (25)		0.973	1.000	0.998	0.833	0.898
		Test based on (23)		0.964	1.000	0.972	0.959	0.932
	500	Test based on (25)		1.000	1.000	1.000	1.000	1.000
		Test based on (23)		1.000	1.000	1.000	1.000	1.000
6	50	Test based on (25)		0.077	0.000	0.000	0.000	0.000
		Test based on (23)		0.032	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.721	0.654	0.671	0.688	0.700
		Test based on (23)		0.591	0.623	0.684	0.673	0.618
	200	Test based on (25)		0.977	1.000	0.998	0.973	1.000
		Test based on (23)		0.971	1.000	1.000	1.000	1.000
	500	Test based on (25)		1.000	1.000	1.000	1.000	1.000
		Test based on (23)		1.000	1.000	1.000	1.000	1.000
7	50	Test based on (25)		0.000	0.000	0.000	0.000	0.000
		Test based on (23)		0.000	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.652	0.583	0.603	0.652	0.669
		Test based on (23)		0.525	0.782	0.816	0.624	0.713
	200	Test based on (25)		0.987	1.000	1.000	0.997	1.000
		Test based on (23)		0.972	1.000	1.000	1.000	1.000
	500	Test based on (25)		1.000	1.000	1.000	1.000	1.000
		Test based on (23)		1.000	1.000	1.000	1.000	1.000

Table 3c. Power of the test for the null of no changes in Σ. Data are generated according to (24) and under the alternative hypothesis as in (27) with $\nu = \frac{1}{2}$.

n	T		(ρ,ϑ)	$(0,0)$	$(0.5,0)$	$(0,0.5)$	$(0,-0.5)$	$(0.5,0.5)$
3	50	Test based on (25)		0.077	0.000	0.000	0.000	0.000
		Test based on (23)		0.111	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.116	0.065	0.083	0.099	0.068
		Test based on (23)		0.125	0.072	0.064	0.074	0.066
	200	Test based on (25)		0.198	0.201	0.197	0.231	0.199
		Test based on (23)		0.125	0.103	0.112	0.111	0.098
	500	Test based on (25)		0.539	0.592	0.608	0.589	0.345
		Test based on (23)		0.063	0.073	0.076	0.079	0.052
4	50	Test based on (25)		0.052	0.000	0.000	0.000	0.000
		Test based on (23)		0.095	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.103	0.001	0.003	0.000	0.002
		Test based on (23)		0.090	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.157	0.191	0.301	0.277	0.132
		Test based on (23)		0.089	0.063	0.053	0.040	0.069
	500	Test based on (25)		0.507	0.599	0.600	0.522	0.351
		Test based on (23)		0.082	0.098	0.083	0.079	0.067
5	50	Test based on (25)		0.023	0.000	0.000	0.000	0.000
		Test based on (23)		0.074	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.005	0.000	0.000	0.000	0.000
		Test based on (23)		0.076	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.106	0.200	0.279	0.275	0.174
		Test based on (23)		0.082	0.089	0.067	0.079	0.092
	500	Test based on (25)		0.373	0.398	0.401	0.402	0.138
		Test based on (23)		0.066	0.089	0.079	0.097	0.082
6	50	Test based on (25)		0.007	0.000	0.000	0.000	0.000
		Test based on (23)		0.028	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.047	0.000	0.000	0.000	0.000
		Test based on (23)		0.059	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.089	0.024	0.036	0.012	0.000
		Test based on (23)		0.084	0.000	0.000	0.000	0.000
	500	Test based on (25)		0.243	0.333	0.394	0.390	0.146
		Test based on (23)		0.067	0.038	0.027	0.025	0.044
7	50	Test based on (25)		0.000	0.000	0.000	0.000	0.000
		Test based on (23)		0.002	0.000	0.000	0.000	0.000
	100	Test based on (25)		0.029	0.000	0.000	0.000	0.000
		Test based on (23)		0.060	0.000	0.000	0.000	0.000
	200	Test based on (25)		0.087	0.000	0.032	0.021	0.000
		Test based on (23)		0.065	0.000	0.000	0.000	0.000
	500	Test based on (25)		0.147	0.390	0.334	0.338	0.163
		Test based on (23)		0.079	0.012	0.008	0.002	0.001

Table 3d. Power of the test for the null of no changes in Σ. Data are generated according to (24) and under the alternative hypothesis specified in (26), with $k = Th_{nT} + 1$ and $\Delta = I_n$.

	monthly data						weekly data					
	mean	std. dev	skew	kurt	AR(1)	ARCH(7)	mean	std. dev	skew	kurt	AR(1)	ARCH(7)
$1m$	-0.035	0.438	-0.481	25.266	-0.129	0.000^{***}	-0.007	0.159	0.119	66.052	0.151	0.000^{***}
$3m$	-0.035	0.366	-0.030	24.291	0.097	0.000^{***}	-0.008	0.121	-0.664	46.650	0.324	0.000^{***}
$6m$	-0.036	0.295	-1.591	15.137	0.209	0.011^{**}	-0.008	0.108	-1.525	21.640	0.215	0.000^{***}
$9m$	-0.037	0.274	-1.624	10.921	0.257	0.668	-0.008	0.121	-1.301	14.454	0.035	0.000^{***}
$12m$	-0.037	0.262	-1.188	8.251	0.286	0.989	-0.008	0.136	-1.252	16.110	-0.098	0.000^{***}
$15m$	-0.038	0.267	-0.851	6.835	0.268	0.972	-0.008	0.130	-0.883	9.984	-0.049	0.000^{***}
$18m$	-0.038	0.273	-0.549	5.589	0.242	0.957	-0.008	0.127	-0.474	6.827	-0.014	0.000^{***}
$21m$	-0.038	0.282	-0.323	4.695	0.210	0.962	-0.008	0.129	-0.138	5.693	-0.003	0.000^{***}
$24m$	-0.038	0.294	-0.160	4.086	0.172	0.974	-0.008	0.134	0.026	5.298	-0.012	0.000^{***}
$30m$	-0.038	0.303	-0.043	3.937	0.146	0.984	-0.008	0.139	0.080	4.930	-0.022	0.000^{***}
$3y$	-0.038	0.314	0.052	3.982	0.118	0.980	-0.007	0.144	0.101	4.741	-0.029	0.000^{***}
$4y$	-0.036	0.319	0.060	4.198	0.073	0.970	-0.007	0.149	0.041	4.560	-0.045	0.000^{***}
$5y$	-0.035	0.323	0.116	4.706	0.037	0.988	-0.006	0.152	0.007	4.486	-0.051	0.000^{***}
$6y$	-0.033	0.321	0.140	5.151	0.033	0.994	-0.006	0.152	-0.029	4.555	-0.052	0.000^{***}
$7y$	-0.032	0.320	0.144	5.560	0.023	0.995	-0.005	0.152	-0.028	4.596	-0.049	0.000^{***}
$8y$	-0.031	0.318	0.138	5.868	0.017	0.994	-0.005	0.152	-0.033	4.776	-0.044	0.000^{***}
$9y$	-0.030	0.318	0.112	6.251	0.009	0.992	-0.005	0.151	-0.063	4.968	-0.047	0.000^{***}
$10y$	-0.030	0.318	0.048	6.653	-0.001	0.990	-0.005	0.151	-0.059	4.946	-0.046	0.000^{***}

Table 4. Descriptive statistics for monthly and weekly data. We report the mean, standard deviation, skewness and kurtosis, the AR(1) coefficient and the p-value of the ARCH(7) test for the returns - *, **, and *** denote rejection of the null of no ARCH effects at 10%, 5% and 1% levels, respectively.

(T,p)	*level*	10%	5%	1%
(163, 1)	cv_1	2.8044	3.0716	3.5758
	cv_3	3.1987	3.4222	3.9344
	cv_{18}	3.7843	3.9690	4.1273
(732, 1)	cv_1	2.9471	3.1891	3.6379
	cv_3	3.3266	3.5153	3.8789
	cv_{18}	3.7779	4.0051	4.5603
(163, 18)	cv_1	6.1897	6.4075	6.7747
	cv_3	6.5068	6.6486	7.0053
(732, 18)	cv_1	6.3066	6.5202	6.9209
	cv_3	6.6288	6.8040	7.1498
(166, 1)	cv_1	2.7952	3.0336	3.5951
	cv_4	3.2688	3.5213	3.9874
(166, 10)	cv_{10}	5.0741	5.3172	5.7916

Table 5. Critical values for the applications in Sections 5.1 and 5.2. cv_N refers to the critical value to be used when N hypotheses are being tested for, in order to have a procedure-wise level of 10%, 5% and 1%, respectively. Panels with $(T,p)=(163,1)$ and $(732,1)$ contain critical values for unidimensional tests (monthly and weekly frequencies, respectively) to test for changes in eigenvalues or to verify the stability of the diagonal elements of Σ one at a time. Panels with $(T,p)=(163,1)$ and $(732,1)$ contain critical values for tests with 18 hypotheses under the null designed for tests for the stability of one eigenvector. Finally, panels with $(T,p)=(166,1)$ and $(166,10)$ contain critical values for the application in Section 5.2: the former are for changes in the eigenvalues, both individual and family-wise tests unidimensional tests (monthly and weekly frequencies respectively), the latter are for testing the stability of the whole matrix.

i	$H_0 : \Sigma_{ii}$ constant	
	monthly	*weekly*
$1m$	2.6989	2.8136
$3m$	2.7656	3.7004
$6m$	2.7394	3.1770
$9m$	2.3924	2.3132
$12m$	1.5350	3.1266
$15m$	1.4991	2.8294
$18m$	1.6467	2.9063
$21m$	1.8065	3.0928
$24m$	1.9827	3.1274
$30m$	2.0718	3.3169
$3y$	2.0815	3.5926
$4y$	1.9314	4.0180^{**} [3rd week, 09/2007]
$5y$	1.8964	4.1170^{**} [1st week, 12/2007]
$6y$	1.8369	4.2779^{**} [1st week, 12/2007]
$7y$	1.7677	4.2595^{**} [1st week, 12/2007]
$8y$	1.9601	4.3342^{**} [1st week, 12/2007]
$9y$	2.1046	4.3549^{**} [1st week, 12/2007]
$10y$	2.1967	4.4386^{**} [last week, 08/2008]

	$H_0 : \lambda_i$ constant	
	monthly	*weekly*
λ_1	1.6921	3.7156^{**} [1st week, 12/2007]
λ_2	2.5513	2.8518
λ_3	3.4328^{**} [01/2008]	2.7495

	$H_0 : \gamma_i$ constant	
	monthly	*weekly*
x_1	3.9142	7.1807^{**} [3rd week, 03/2008]
x_2	4.3898	7.2960^{**} [3rd week, 04/2008]
x_3	4.2340	7.2526^{**} [2nd week, 03/2008]

Table 6. Tests for changes in the variances of the term structure; in the volatilities of each principal component; and in the eigenvalue-normed eigenvectors. Rejection at 10%, 5% and 1% levels are denoted with *, ** and *** respectively. Where present, numbers in square brackets are the estimated breakdates, defined as $T \times \arg\max \Lambda_T(\tau)$.

	monthly data			*weekly data*	
	1st subsample	*2nd subsample*		*1st subsample*	*2nd subsample*
λ_1	0.790	0.729	λ_1	0.737	0.780
λ_2	0.163	0.214	λ_2	0.164	0.142
λ_3	0.029	0.047	λ_3	0.056	0.056

Table 7. Proportion of the total variance explained by principal components (λ_1, λ_2 and λ_3 refer to the level, slope and curvature respectively) for each subsample. The samples are split based on the results in Table 6. When considering monthly data, the sample was split at January 2008; when using weekly data, at the first week of December 2007.

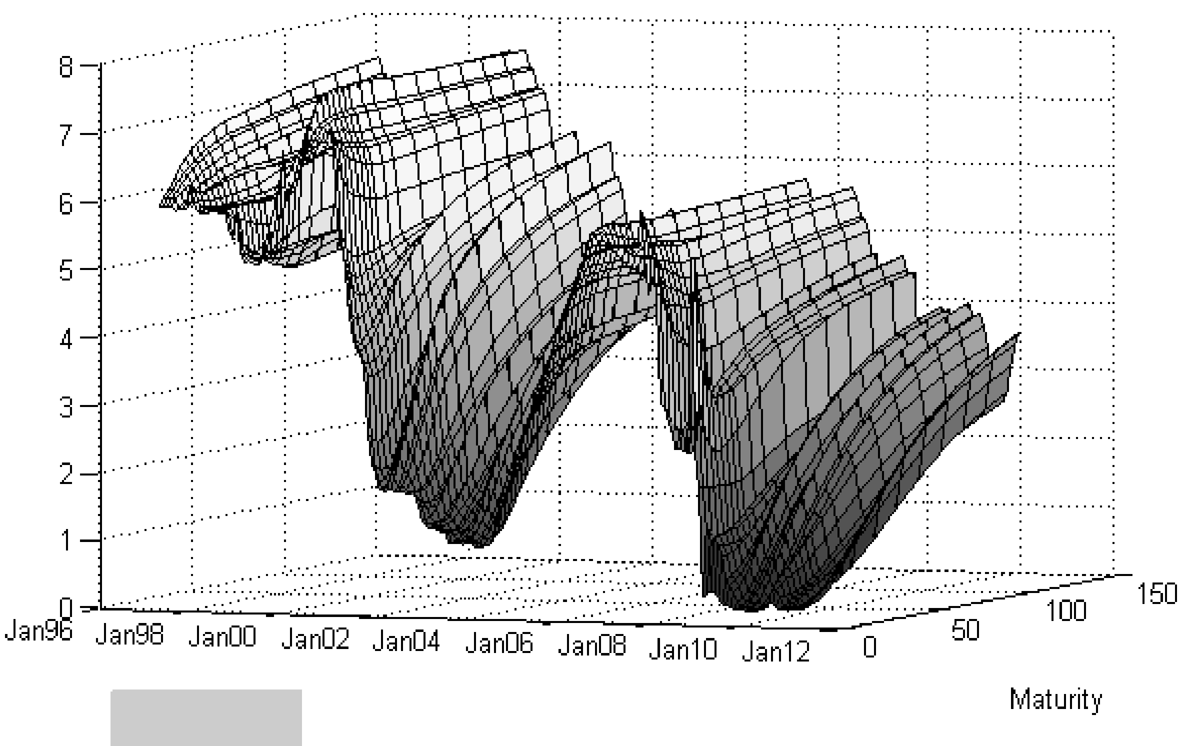

Figure 1. Term structure of the US interest rates. Maturities correspond to 1m, 3m, 6m, 9m, 12m, 15m, 18m, 21m, 24m, 30m, 3y, 4y, 5y, 6y, 7y, 8y, 9y, 10y over the period April 1997-November 2010.

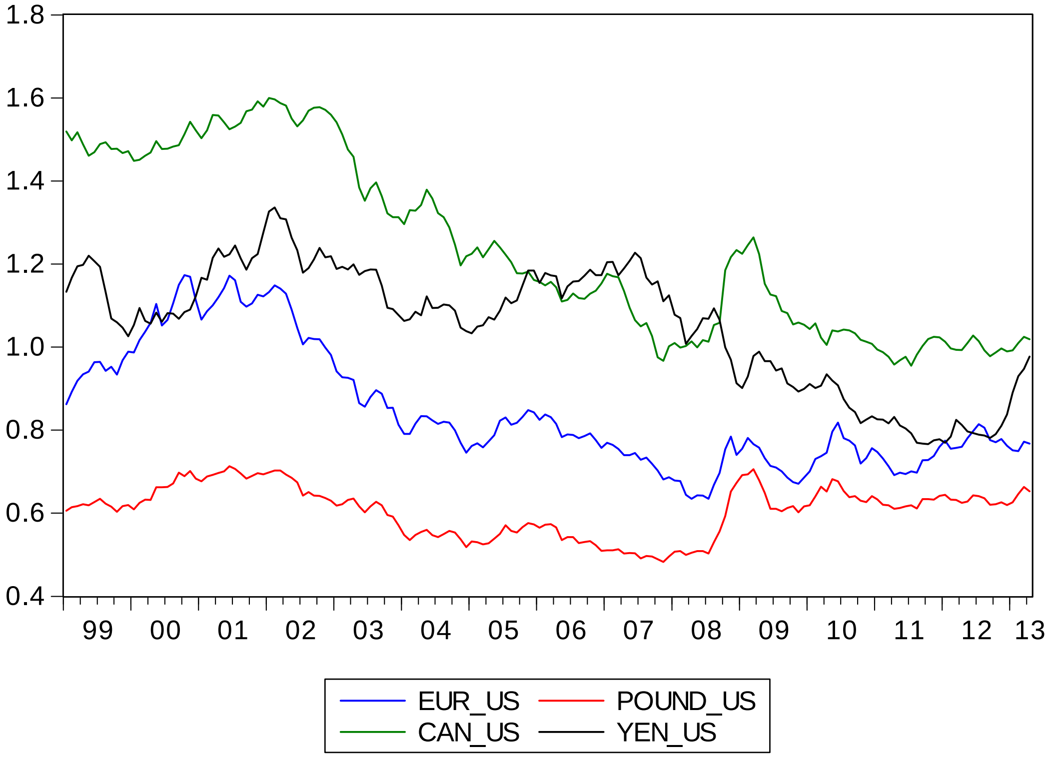

Figure 2. Exchange rates vis-a-vis the US Dollar. Monthly data recorded between January 1999 and April 2013; for the Yen/US Dollar exchange rate is concerned, the original series have been divided by 100 to make figures comparable with the other exchange rates.

	mean	std. dev	skew	kurt	AR(1)	ARCH(7)
EUR	−0.0007	0.0252	−0.0026	2.990	0.297	0.113
GBP	0.0004	0.0219	0.4460	4.841	0.294	0.000***
CAN	−0.0023	0.0192	0.9417	9.774	0.272	0.505
YEN	−0.0009	0.0238	−0.1859	2.983	0.263	0.385

Table 8. Descriptive statistics for the returns on (vis-a-vis the US Dollar): the Euro (EUR), the British Pound (GBP), the Canadian Dollar (CAN) and the Yen (YEN). See Table 4 for an explanation of the output.

Estimation Output

		EUR	GBP	CAN	YEN
Constant		−0.0015	0.0004	−0.0020	−0.0004
Lag 1	EUR	0.3760*	0.2080*	0.1604	−0.0408*
	GBP	−0.1642	0.0442	−0.1053	−0.0650
	CAN	0.2966*	0.4057*	0.2786*	0.0667
	YEN	−0.0389	−0.0139	−0.0396	0.2045
Lag 2	EUR	−0.2868*	−0.2044*	−0.1499	−0.0674
	GBP	0.2425	0.2680*	0.1300	−0.0614
	CAN	−0.1387	−0.1891	0.0099	0.0169
	YEN	0.1462	−0.0405	0.0394	0.1195
Lag 5	EUR	0.0710	0.0165	0.2031*	0.0779
	GBP	0.0500	−0.1051	−0.0851	−0.1171
	CAN	−0.1385	0.0119	−0.1180	0.2350*
	YEN	−0.1603	−0.0684	−0.1311*	−0.2036*

Mis-specification tests

Serial Correlation LM test	134.378 [0.904]
Normality	100.308 [0.000]

Table 9. Estimation results and misspecification analysis for the VAR; the components of the vector are the following (vis-a-vis the US Dollar): the Euro (EUR), the British Pound (GBP), the Canadian Dollar (CAN) and the Yen (YEN). In the top half of the Table we report the estimation results: the symbol "*" denotes rejection of the null that the corresponding coefficient is zero at 5% level. The second half of the Table contains misspecification tests: we report (in order of appearance) the portmanteu serial correlation test (up to lag 12) and the Jarque-Bera test.

$H_0 : \Sigma$ constant	4.8049	$H_0 : \lambda_1$ constant	1.5457
		$H_0 : \lambda_2$ constant	2.2527
		$H_0 : \lambda_3$ constant	1.7077
		$H_0 : \lambda_4$ constant	3.4154^{*} [June 2008]

Table 10. Tests for changes in the covariance matrix of the error term of the VAR estimated in Table 9. Rejection at 10%, 5% and 1% levels are denoted with *, ** and *** respectively. Numbers in square brackets refer to the estimated breakdates, defined as $T \times \arg\max \Lambda_T^{\varepsilon}(\tau)$.

INTERNET APPENDIX
Appendix A: Preliminary Lemmas

Lemma 1 *Let B and ρ be non-negative random variables, with $|B|_{q^*/q-1} < \infty$ and $|\rho|_{q^*} < \infty$, where $q^* = q\left(1+\frac{\delta}{2}\right)$, $q \geq 1$ and $\delta > 0$. Assume $|B\rho|_r < \infty$ for $r > 2+\delta$. Then*

$$|B\rho|_{2+\delta} \leq \left[|\rho|_{q^*}^{r-2} |B|_{q^*/q-1}^{r-2}\right]^{1/2(r-1)} \times \left[|B\rho|_r^{\left[2r(1+\delta)-\delta r^2\right]/[2+\delta]} + |B\rho|_r^r\right]^{1/2(r-1)}. \tag{28}$$

Proof. The proof is fairly similar to Gallant and White (1988; see Davidson, 2002a, p. 271). Let $C = \left[|\rho|_{q^*} |B|_{q^*/q-1} |B\rho|_r^{-r}\right]^{\frac{1}{1-r}}$, and define $B_1 = I_{\{B\rho \leq C\}} B$. By construction, $|B\rho|_{2+\delta} \leq |B_1\rho|_{2+\delta} + |(B-B_1)\rho|_{2+\delta}$. We have

$$\begin{aligned} |B_1\rho|_{2+\delta} &= \left[\int_{B\rho \leq C} (B\rho)^{2+\delta} dP\right]^{\frac{1}{2+\delta}} \\ &\leq C^{1/2}\left[\int_{B\rho \leq C} (B\rho)^{\frac{2+\delta}{2}} dP\right]^{\frac{2}{2+\delta}\frac{1}{2}} \\ &\leq C^{1/2}\left[|\rho|_{q^*} |B|_{q^*/q-1}\right]^{1/2}, \end{aligned}$$

where the last passage follows from Holder's inequality. Also, since $r > 2$ and $(B\rho)^r > C^r$,

$$\begin{aligned} |(B-B_1)\rho|_{2+\delta} &= \left[\int_{B\rho > C} (B\rho)^{2+\delta} dP\right]^{\frac{1}{2+\delta}} \\ &\leq C^{1-\frac{r}{2}}\left[\int_{B\rho > C} (B\rho)^r dP\right]^{\frac{r}{2+\delta}\frac{1}{r}} \\ &\leq C^{1-\frac{r}{2}} |B\rho|_r^{\frac{r}{2+\delta}}. \end{aligned}$$

Substituting for C gives (28). ■

Remarks

1. The Lemma is an extension of Lemma 4.1 in Gallant and White (1988, p. 47). Their result is derived for the L_2-norm, and the method of proof here is exactly the same.
2. Equation (28) is very similar to Lemma 17.15 in Davidson (2002a, p. 271). Particularly, $|\rho|_{q^*}^{r-2}$, for some $q^* < 2$, is raised to the power of $r-2$: this is exactly the same as in Lemma 17.15 in Davidson (2002a). This is an important consequence of the Lemma. Given a non-Lipschitz transformation of a NED sequence u_t, say $\phi(u_t)$, setting $u_t^m = E[u_t | u_{t-m}, ..., u_{t+m}]$, one would look for a bound to $|\phi(u_t) - \phi(u_t^m)|_{2+\delta}$. This would be majorized by some suitably chosen $|B(u_t, u_t^m)\rho(u_t, u_t^m)|_{2+\delta}$; since normally one would choose $\rho(u_t, u_t^m)$ as the taxicab distance, it is $|\rho|_{q^*}^{r-2}$ that gives the size of $\phi(u_t)$.

Lemma 2 *Under Assumption 1, w_t is $L_{2+\epsilon}$-NED of size $\alpha' > \frac{1}{2}$ on $\{v_t\}_{t=-\infty}^{+\infty}$.*

Proof. The proof follows similar passages as in Example 17.17 in Davidson (2002a, p. 273); for simplicity, assume $n = 1$. Let $x_t^a = w_t$ and $x_t^b = E[w_t | w_{t-m}, ..., w_{t+m}]$; and define, in a similar fashion, $y_t^a = y_t$ and $y_t^b = E[y_t | y_{t-m}, ..., y_{t+m}]$. Then

$$\begin{aligned} \left|x_t^a - x_t^b\right|_{2+\epsilon} &\leq \left|\left(|y_t^a| + \left|y_t^b\right|\right)\left(\left|y_t^a - y_t^b\right|\right)\right|_{2+\epsilon} \\ &= \left|B\left(y_t^a, y_t^b\right)\rho\left(y_t^a, y_t^b\right)\right|_{2+\epsilon} = |B\rho|_{2+\epsilon} \end{aligned}$$

Lemma 1 entails that $\left|x_t^a - x_t^b\right|_{2+\epsilon}$ is bounded by $\left[|\rho|_{q^*}^{r-2} |B|_{q^*/q-1}^{r-2} \left(|B\rho|_r^{\left[2r(1+\epsilon)-\epsilon r^2\right]/[2+\epsilon]} + |B\rho|_r^r\right)\right]^{1/2(r-1)}$. It holds that $|\rho|_{q^*}^{r-2} < \infty$ for $q^* \leq 2r$, and thus for $q < 2r$; also, $|B|_{q^*/q-1}^{r-2} < \infty$ if $q^* \geq \frac{4}{3}$, i.e. $q > \frac{4}{3}$. Since, for $q^* \leq 2$, $|\rho|_{q^*} \leq |\rho|_2 \leq M v_m$, where M is a constant and v_m a non-negative sequence of size $v_m = O\left(m^{-\alpha}\right)$, we have $\left|x_t^a - x_t^b\right|_{2+\epsilon} = \left|x_t^a - E\left[w_t | w_{t-m}, ..., w_{t+m}\right]\right|_{2+\epsilon} \leq M' \left[v_m^{r-2}\right]^{1/2(r-1)} = M' v_m'$. The size of v_m' is $v_m' = O\left(m^{-\alpha \frac{r-2}{2(r-1)}}\right) = O\left(m^{-\alpha'}\right)$. Assumption 1*(ii)* entails that $\alpha' > \frac{1}{2}$. ■

Lemma 3 *Under Assumption 1, vec* $\left[\bar{w}_t \bar{w}_{t-l}' - E\left(\bar{w}_t \bar{w}_{t-l}'\right)\right]$ *is* $L_{1+\epsilon/2}$*-NED of size* α' *on* $\{v_t\}_{t=-\infty}^{+\infty}$*, for every* l*.*

Proof. The Lemma is an application of Theorem 17.9 in Davidson (2002a, p. 268), where L_1- and L_2-norms are replaced, respectively, by $L_{1+\epsilon/2}$- and $L_{2+\epsilon}$-norms. ■

Lemma 4 *Under Assumption 1 and 2(i)(b)-(ii), vec* $\left[\bar{w}_t \bar{w}_t' - E\left(\bar{w}_t \bar{w}_t'\right)\right]$ *is* $L_{2+\epsilon}$*-NED of size* α' *on* $\{v_t\}_{t=-\infty}^{+\infty}$*.*

Proof. The proof is very similar to that of Lemma 2. Assuming $n = 1$ and letting $\omega_t = w_t^2$

$$\begin{aligned} \left|\omega_t^a - \omega_t^b\right|_{2+\epsilon} &\leq \left|\left(x_t^a + x_t^b\right)\left(x_t^a - x_t^b\right)\right|_{2+\epsilon} \\ &= \left|\left(|y_t^a|^3 + \left|y_t^b\right|^3 + |y_t^a|^2 \left|y_t^b\right| + |y_t^a| \left|y_t^b\right|^2\right)\left(\left|y_t^a - y_t^b\right|\right)\right|_{2+\epsilon} \\ &= \left|B\left(y_t^a, y_t^b\right) \rho\left(y_t^a, y_t^b\right)\right|_{2+\epsilon} = |B\rho|_{2+\epsilon}, \end{aligned}$$

so that the Lemma follows from Assumption 2*(ii)* and Lemma 1. ■

Lemma 5 *Under Assumption R1, it holds that:*

(i) $\bar{w}_t^\varepsilon$ *is* $L_{2+\epsilon}$*-NED of size* $\alpha' > \frac{1}{2}$ *on* $\{v_t\}_{t=-\infty}^{+\infty}$*;*

(ii) *redefining* $\bar{w}_t^\varepsilon$ *in a richer probability space,* $\sum_{t=1}^{\lfloor T\tau \rfloor} \bar{w}_t^\varepsilon = \sum_{t=1}^{\lfloor T\tau \rfloor} X_t^\varepsilon + O_{a.s.}\left(\lfloor T\tau \rfloor^{\frac{1}{2}-\delta}\right)$*, uniformly in* τ*, where* X_t^ε *is a zero mean, i.i.d. Gaussian sequence with* $E\left(X_t^\varepsilon X_t^{\varepsilon\prime}\right) = V^\varepsilon$ *and* $\delta > 0$*;*

(iii) $\hat{\beta} - \beta = O_{a.s.}\left(\sqrt{\frac{\ln \ln T}{T}}\right)$*;*

(iv) *defining* $\overline{\overline{w}}_t^\varepsilon = \hat{w}_t^\varepsilon - vec\left(\Sigma_\varepsilon\right)$*, it holds that* $\sum_{t=1}^{\lfloor T\tau \rfloor} \overline{\overline{w}}_t^\varepsilon = \sum_{t=1}^{\lfloor T\tau \rfloor} X_t^{\varepsilon *} + O_{a.s.}\left(\lfloor T\tau \rfloor^{\frac{1}{2}-\delta}\right)$*, uniformly in* τ*, where* $X_t^{\varepsilon *}$ *is a zero mean, i.i.d. Gaussian sequence with* $E\left(X_t^{\varepsilon *} X_t^{\varepsilon * \prime}\right) = V^\varepsilon$ *and* $\delta > 0$*;*

(v) *vec* $\left[\bar{w}_t^\varepsilon \bar{w}_{t-l}^{\varepsilon\prime} - E\left(\bar{w}_t^\varepsilon \bar{w}_{t-l}^{\varepsilon\prime}\right)\right]$ *is* $L_{1+\epsilon/2}$*-NED of size* α' *on* $\{v_t\}_{t=-\infty}^{+\infty}$*, for every* l*.*

Further, under Assumptions R1(ii) and R2(i)-(ii), it holds that:

(vi) *vec* $\left[\bar{w}_t^\varepsilon \bar{w}_t^{\varepsilon\prime} - E\left(\bar{w}_t^\varepsilon \bar{w}_t^{\varepsilon\prime}\right)\right]$ *is* $L_{2+\epsilon}$*-NED of size* α' *on* $\{v_t\}_{t=-\infty}^{+\infty}$*.*

Proof. Parts (i) and (v) of the Lemma are the same as Lemmas 2 and 3 respectively. Part (ii) is the same as equation (2), and we refer to the proof of Theorem 1 for details. Similarly, part (vi) of the Lemma is the same as Lemma 4.

As far as part (iii) is concerned, let $n = q = 1$ for the sake of notational simplicity, and consider $\hat{\beta} - \beta = \left(\sum_{t=1}^T x_t^2\right)^{-1} \left(\sum_{t=1}^T x_t \varepsilon_t\right)$. We start from the denominator. By virtue of Assumption R1*(i)*, Lemma 2 entails that x_t^2 is $L_{2+\epsilon}$-NED of size $\alpha' > \frac{1}{2}$ on the strong mixing basis $\{v_t\}_{t=-\infty}^{+\infty}$. This result allows to use a SLLN, since the assumptions of Theorem 2.1 in Ling (2007) are satisfied; thus, $\sum_{t=1}^T x_t^2 = O_{a.s.}(T)$. Consider now the numerator; Assumptions R1*(i)*-*(ii)* entail that $x_t \varepsilon_t$ is $L_{2+\epsilon}$-NED of size $\alpha' > \frac{1}{2}$ on $\{v_t\}_{t=-\infty}^{+\infty}$; this is a direct application of Theorem 17.17 in Davidson (2002a, p. 273), and it follows the same lines as the proof of Lemma 2. Using this result and Assumption R1*(iii)*, a SIP holds, based on the same passages as in the proof of Theorem 1, so that $\sum_{t=1}^T x_t \varepsilon_t = \sum_{t=1}^T N_t^{x\varepsilon} + O_{a.s.}\left(T^{\frac{1}{2}-\delta}\right)$ for some $\delta > 0$, where $N_t^{x\varepsilon}$ is a sequence of *i.i.d.* zero mean, Gaussian random variables with covariance matrix $V^{x\varepsilon}$. Therefore, by the Law of the Iterated Logarithm (LIL), $\sum_{t=1}^{T} x_t\varepsilon_t = O_{a.s.}\left(\sqrt{T\ln\ln T}\right)$. Part (iii) follows by combining the two results.

Finally, we show part (iv). We can write $\sum_{t=1}^{\lfloor T\tau\rfloor} \overline{\hat{w}}_t^{\varepsilon} = \sum_{t=1}^{\lfloor T\tau\rfloor} \bar{w}_t^{\varepsilon} - 2vec\left[\left(\hat{\beta}-\beta\right)\left(\sum_{t=1}^{\lfloor T\tau\rfloor} x_t\varepsilon_t'\right)\right] + vec\left[\left(\hat{\beta}-\beta\right)\left(\sum_{t=1}^{\lfloor T\tau\rfloor} x_t x_t'\right)\left(\hat{\beta}-\beta\right)'\right] = \sum_{t=1}^{\lfloor T\tau\rfloor} \bar{w}_t^{\varepsilon} + I + II$. By virtue of the passages above, I and II are both $O_{a.s.}(\ln\ln T)$. Part (iv) then follows from part (ii). ■

Lemma 6 *Under Assumptions R1(ii)-(iv)(a) and R2, it holds that* $\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\tilde{V}_\tau^{\varepsilon}-V^{\varepsilon}\right\| = O_p\left(\frac{1}{m}\right)+O_p\left(m\sqrt{\frac{\ln\ln T}{T}}\right)$.

Proof. Let, preliminarily, $\tilde{V}_\tau^{\varepsilon\varepsilon} = \left(\hat{\Psi}_{0,\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}+\hat{\Psi}_{0,1-\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}\right)+\sum_{l=1}^{m}\left(1-\frac{l}{m}\right)\left[\left(\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}+\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon\varepsilon\prime}\right)+\left(\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}+\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon\varepsilon\prime}\right)\right]$, where $\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon\varepsilon} = \frac{1}{T-(l+1)}\sum_{t=l+1}^{\lfloor T\tau\rfloor}\left[w_t^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon}\right)\right]\left[w_{t-l}^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon}\right)\right]'$, and similarly for $\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}$. We start by showing that $\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\tilde{V}_\tau^{\varepsilon\varepsilon}-V^{\varepsilon}\right\| = O_p\left(\frac{1}{m}\right)+O_p\left(\frac{m}{\sqrt{T}}\right)$; the proof is very similar to that of Theorem 2, and therefore only the most relevant passages are reported. Let $\hat{\Sigma}_{\varepsilon\varepsilon} = T^{-1}\sum_{t=1}^{T}\varepsilon_t\varepsilon_t'$, and note that $\hat{\Sigma}_{\varepsilon}-\hat{\Sigma}_{\varepsilon\varepsilon} = \left(\hat{\beta}-\beta\right)\left(T^{-1}\sum_{t=1}^{T}x_t x_t'\right)\left(\hat{\beta}-\beta\right)'-\left(T^{-1}\sum_{t=1}^{T}\varepsilon_t x_t'\right)\left(\hat{\beta}-\beta\right)'-\left(\hat{\beta}-\beta\right)\left(T^{-1}\sum_{t=1}^{T}x_t\varepsilon_t'\right) = I+II+II'$. By Lemma 5, we have $I = O_{a.s.}\left(\frac{\ln\ln T}{T}\right)$, and similarly $II = O_{a.s.}\left(\frac{\ln\ln T}{T}\right)$. Thus, for every l, $\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon\varepsilon} = [T-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}\left[w_t^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon\varepsilon}\right)\right]\left[w_{t-l}^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon\varepsilon}\right)\right]'+O_{a.s.}\left(\frac{\ln\ln T}{T}\right)$. By using part (ii) of Lemma 5, it follows that $\hat{\Sigma}_{\varepsilon\varepsilon}-\Sigma_{\varepsilon} = O_{a.s.}\left(T^{-\frac{1}{2}-\delta}\right)$, so that $\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon\varepsilon} = [T-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime}+O_{a.s.}\left(T^{-\frac{1}{2}-\delta}\right)$; the same holds for $\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}$. Therefore, $\hat{\Psi}_{l,\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}+\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}^{\varepsilon\varepsilon}-\Psi_l^{\varepsilon\varepsilon} = [T-(l+1)]^{-1}\sum_{t=l+1}^{T}\left[\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime}-E\left(\bar{w}_t^{\varepsilon}\bar{w}_{t-l}^{\varepsilon\prime}\right)\right]+O_{a.s.}\left(T^{-\frac{1}{2}-\delta}\right)$. Henceforth, the proof is the same as for equation (6), by virtue of part (vi) of Lemma 5 and since the assumptions on the autocovariances Ψ_l^{ε} are the same as in Assumption 2. Finally, we show that $\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\tilde{V}_\tau^{\varepsilon\varepsilon}-\tilde{V}_\tau^{\varepsilon}\right\| = O_p\left(m\sqrt{\frac{\ln\ln T}{T}}\right)$. Consider the quantities $I = [T-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}2vec\left[\left(\hat{\beta}-\beta\right)'x_t\varepsilon_t'\right]\left[w_{t-l}^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon}\right)\right]'$ and $II = [T-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}vec\left[\left(\hat{\beta}-\beta\right)\left(x_t x_t'\right)\left(\hat{\beta}-\beta\right)'\right]\left[w_{t-l}^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon}\right)\right]'$. We have $I\le\left\|\hat{\beta}-\beta\right\|\left\{[T-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}\left\|x_t\varepsilon_t'\right\|^2\right\}^{1/2}\left\{[T-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}\left\|w_{t-l}^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon}\right)\right\|^2\right\}^{1/2}$. By applying Theorem 17.9 in Davidson (2002a, p. 268) following similar lines as in the proof of Lemma 3, $\left\|x_t\varepsilon_t'\right\|^2$ can be shown to be $L_{1+\epsilon/2}$-NED of size $\alpha'>\frac{1}{2}$ on $\{v_t\}_{t=-\infty}^{+\infty}$; thus, a SLLN (see e.g. Theorem 2.1 in Ling, 2007) can be applied with $[\lfloor T\tau\rfloor-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}\left\|x_t\varepsilon_t'\right\|^2 = O_{a.s.}(1)$; similarly, part (v) of Lemma 5 entails a SLLN whereby $[\lfloor T\tau\rfloor-(l+1)]^{-1}\sum_{t=l+1}^{\lfloor T\tau\rfloor}\left\|w_{t-l}^{\varepsilon}-vec\left(\hat{\Sigma}_{\varepsilon}\right)\right\|^2 = O_{a.s.}(1)$; hence, $I = O_{a.s.}\left(\sqrt{\frac{\ln\ln T}{T}}\right)$. By the same logic, $II = O_{a.s.}\left(\frac{\ln\ln T}{T}\right)$. After some algebra, $\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\tilde{V}_\tau^{\varepsilon\varepsilon}-\tilde{V}_\tau^{\varepsilon}\right\| = O_p\left(m\sqrt{\frac{\ln\ln T}{T}}\right)$ follows. ■

Lemma 7 *Under Assumptions R1 and R3, it holds that as* $T\to\infty$, *uniformly in* τ, $\hat{\lambda}_{i,\tau}^{\varepsilon}-\lambda_i^{\varepsilon} = \left(x_i^{\varepsilon\prime}\otimes x_i^{\varepsilon\prime}\right)vec\left(\hat{\Sigma}_{\varepsilon,\tau}-\Sigma_{\varepsilon}\right)+O_p\left(T^{-1}\right)$, $\hat{x}_{i,\tau}^{\varepsilon}-x\varepsilon_i = v_{x,i}^{\varepsilon}\,vec\left(\hat{\Sigma}_{\varepsilon,\tau}-\Sigma_{\varepsilon}\right)+O_p\left(T^{-1}\right)$ *and* $\hat{\gamma}_{i,\tau}^{\varepsilon}-\gamma_i^{\varepsilon} = v_{\gamma,i}^{\varepsilon}\,vec\left(\hat{\Sigma}_{\varepsilon,\tau}-\Sigma_{\varepsilon}\right)+O_p\left(T^{-1}\right)$.

Proof. By part (i) of Lemma 5, an FCLT holds for $vec\left(\hat{\Sigma}_{\varepsilon,\tau}-\Sigma_{\varepsilon}\right)$, so that $\sup_{\lfloor T\tau\rfloor}\left\|\hat{\Sigma}_{\varepsilon,\tau}-\Sigma_{\varepsilon}\right\| = O_p\left(T^{-1/2}\right)$ - see the proof of equation (1) in Theorem 1. Henceforth, the proof is the same as for Proposition 1. ■

Appendix B: Proofs

Proof of Theorem 1. The proof of (1) is essentially based on checking the validity of the assumptions in Theorem 29.6 in Davidson (2002a, p. 481) for the normalized sequence $\bar{w}_{T,t} = V_{\Sigma,T}^{-1/2}\bar{w}_t$. In light of Lemma 2, $\bar{w}_{T,t}$, for given values of α and r in Assumption 2, is L_2-NED on the strong mixing base $\{v_t\}_{t=-\infty}^{+\infty}$ with size $\alpha' > \frac{1}{2}$, which entails the validity of Assumption (c) in Davidson (2002a; Theorem 29.6). Assumption 1*(ii)* implies that $E(\bar{w}_{T,t}) = V_{\Sigma,T}^{-1/2}E(\bar{w}_t) = 0$. Assumption (b) in Theorem 29.6 in Davidson (2002a, p. 482) follows from Assumption 1*(ii)* and from noting that, in light of Assumption 1*(i)*, $\sup_t E\left(\|\bar{w}_t\|^{r/2}\right) < \infty$. Assumptions (d) and (f) in Theorem 29.6 in Davidson (2002a) are implied by Assumption 1*(iii)*. Finally, Assumption (e) follows from the LLN entailed by Assumptions 1*(iii)*. Thus, (1) holds.

As far as (2) is concerned, its proof is based on Theorem 1 in Eberlein (1986, p. 263). Lemma 2 entails that $\bar{w}_t$ is a zero-mean $L_{2+\epsilon}$-mixingale of size $\alpha'' > \frac{1}{2}$. Letting $\Im_m = \{\bar{w}_{1,...}, \bar{w}_m\}$ and $S_{Tm} \equiv \sum_{t=m+1}^{m+T}\bar{w}_t$, (2) follows if $|E[S_{Tm}|\Im_m]|_2 < \infty$ and $|E[S_{iTm}S_{jTm}|\Im_m] - E[S_{iTm}S_{jTm}]| = O(T^{1-\theta})$ for $\theta > 0$ and all i, j. Both conditions can be proved following the same passages as in Corradi (1999, pp. 651-652). ■

Proof of Theorem 2. The proof is similar to the proof of Lemma 2.1.1 in Csorgo and Horvath (1997, p. 74-75). In view of Lemma 3, a SLLN holds (see Ling, 2007, Theorem 2.1), whereby for all l

$$\frac{1}{\lfloor T\tau\rfloor}\sum_{t=1}^{\lfloor T\tau\rfloor} vec\left[\bar{w}_t\bar{w}_{t-l}' - E\left(\bar{w}_t\bar{w}_{t-l}'\right)\right] = o_{a.s.}\left(\frac{1}{\lfloor T\tau\rfloor^{\delta'}}\right);$$

similarly, $\hat{\Sigma}_\tau - \Sigma = o_{a.s.}\left(\lfloor T\tau\rfloor^{-\delta'}\right)$, since w_t also satisfies the assumptions needed for Theorem 2.1 in Ling (2007). This entails that, for any $\varepsilon > 0$ and $\varepsilon' > 0$, there is an integer $g_T = g_T(\varepsilon, \varepsilon')$ such that

$$\begin{aligned} P\left[\sup_{g_T\le\lfloor T\tau\rfloor\le T}\lfloor T\tau\rfloor^{\delta'}\left\|\hat{\Psi}_{l,\lfloor T\tau\rfloor} - \Psi_l\right\| > \varepsilon\right] &\le \varepsilon', \\ P\left[\sup_{1\le\lfloor T\tau\rfloor\le T-g_T}\lfloor T\tau\rfloor^{\delta'}\left\|\hat{\Psi}_{l,\lfloor T\tau\rfloor} - \Psi_l\right\| > \varepsilon\right] &\le \varepsilon'. \end{aligned}$$

These yield $\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\hat{\Psi}_{l,\lfloor T\tau\rfloor} - \Psi_l\right\| = o_p\left(\frac{1}{T^{\delta'}}\right)$. This proves (4).

In order to prove (5), write

$$\begin{aligned} \sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\tilde{V}_{\Sigma,\tau} - V_\Sigma\right\| &\le \sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\hat{V}_{\Sigma,\tau} - V_\Sigma\right\| \\ &\quad +2\sup_{1\le\lfloor T\tau\rfloor\le T}\sum_{l=0}^{m}\left(1-\frac{l}{m}\right)\left\|\hat{\Psi}_{l,\lfloor T\tau\rfloor} - \Psi_l + \hat{\Psi}_{l,1-\lfloor T\tau\rfloor} - \Psi_l\right\| \\ &\quad +2\sum_{l=1}^{m}\frac{l}{m}\|\Psi_l\| + 2\sum_{l=m+1}^{\infty}\|\Psi_l\| \\ &= I + II + III + IV. \end{aligned} \tag{29}$$

Assumption 2*(i)*(b) entails $IV = o(m^{-s})$. Equation (4) ensures that $I = o_p\left(T^{-\delta'}\right)$; as far as II is concerned, this is of the same order as

$$\max_{1\le l\le m} E\left[\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\hat{\Psi}_{l,\lfloor T\tau\rfloor} + \hat{\Psi}_{l,1-\lfloor T\tau\rfloor} - \Psi_l\right\|\right]\sum_{l=1}^{m}\left(1-\frac{l}{m}\right) = O\left(\frac{1}{T^{\delta'}}\right)O(m); \tag{30}$$

finally, in light of Assumption 2*(i)*(b), $III = 2m^{-1}O(1) = O(m^{-1})$. Thus, (5) follows.

Consider (6). We still use (29) in our proof. The orders of magnitude of I, III and IV are the same as above. Turning to II, similar passages as in the proof of Theorem 1 yield an IP for each l, so that $\sup_{1\le\lfloor T\tau\rfloor\le T}$

$\left\|\hat{\Psi}_{l,\lfloor T\tau\rfloor}-\Psi_l\right\|=O_p\left(T^{-1/2}\right)$. Thus (30) becomes

$$\max_{1\leq l\leq m} E\left[\sup_{1\leq\lfloor T\tau\rfloor\leq T}\left\|\hat{\Psi}_{l,\lfloor T\tau\rfloor}+\hat{\Psi}_{l,1-\lfloor T\tau\rfloor}-\Psi_l\right\|\right]\sum_{l=1}^{m}\left(1-\frac{l}{m}\right)=O\left(\frac{1}{\sqrt{T}}\right)O(m).$$

Putting all together, (5) follows. ■

Proof of Proposition 1. The estimation error in $\hat{\Sigma}$ can be represented as a perturbation of Σ, with $\hat{\Sigma}_\tau=\Sigma+\left(\hat{\Sigma}_\tau-\Sigma\right)$. Recall that in light of Theorem 1, $\sup_{\lfloor T\tau\rfloor}\left\|\hat{\Sigma}_\tau-\Sigma\right\|=O_p\left(T^{-1/2}\right)$. The eigenvalue problem for the perturbed matrix is

$$\left[\Sigma+\left(\hat{\Sigma}_\tau-\Sigma\right)\right]\left[x_i+\left(\hat{x}_{i,\tau}-x_i\right)\right]=\left[\lambda_i+\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)\right]\left[x_i+\left(\hat{x}_{i,\tau}-x_i\right)\right]. \tag{31}$$

After expanding the product, consider the terms $\left(\hat{\Sigma}_\tau-\Sigma\right)\left(\hat{x}_{i,\tau}-x_i\right)$ and $\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)\left(\hat{x}_{i,\tau}-x_i\right)$. It holds that $\hat{\lambda}_{i,\tau}-\lambda_i=O_p\left(T^{-1/2}\right)$ uniformly in τ. This is because Σ is symmetric, and therefore Corollary 6.3.4 in Horn and Johnson (1999, p. 367) entails that $\left|\hat{\lambda}_{i,\tau}-\lambda_i\right|\leq\left\|\hat{\Sigma}_\tau-\Sigma\right\|$. Equation (1) yields the result. Also, it holds that $\hat{x}_{i,\tau}-x_i=O_p\left(T^{-1/2}\right)$ uniformly in τ. This follows from the $\sin\Theta$ Theorem in Davis and Kahan (1970, p. 10), whereby $\left\|\hat{X}_\tau-X\right\|\leq\|X\|\left\|\hat{\Sigma}_\tau-\Sigma\right\|$. Thus, $\left(\hat{\Sigma}_\tau-\Sigma\right)\left(\hat{x}_{i,\tau}-x_i\right)$ and $\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)\left(\hat{x}_{i,\tau}-x_i\right)$ are $O_p\left(T^{-1}\right)$ uniformly in τ; omitting them, (31) can be written as

$$\Sigma\left(\hat{x}_{i,\tau}-x_i\right)+\left(\hat{\Sigma}_\tau-\Sigma\right)x_i=\lambda_i\left(\hat{x}_{i,\tau}-x_i\right)+\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)x_i. \tag{32}$$

Consider (9). Premultiplying (32) by x_i', we obtain $x_i'\Sigma\left(\hat{x}_{i,\tau}-x_i\right)+x_i'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i=\lambda_i x_i'\left(\hat{x}_{i,\tau}-x_i\right)+\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)x_i'x_i$. Recalling that $x_i'\Sigma=\lambda_i x_i'$, and that $x_i'x_i=1$, we have $x_i'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i=\lambda_i x_i'\left(\hat{x}_{i,\tau}-x_i\right)$, which entails (9). In order to prove (10), note that the x_is are a complete (and orthonormal) basis. This entails that, for an arbitrary set of constants $\phi_{j,\tau}$, it holds that $\hat{x}_{i,\tau}-x_i=\sum_{j=1}^{n}\phi_{j,\tau}x_j$. Premultiplying (32) by any x_k' for $i\neq k$, and using the identity $x_i'\Sigma=\lambda_i x_i'$, we obtain $\lambda_k\phi_{k,\tau}+x_k'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i=\lambda_i\phi_{k,\tau}$. This yields $\phi_{k,\tau}=\frac{x_k'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i}{\lambda_i-\lambda_k}$, under Assumption 2 which stipulates that $\lambda_i\neq\lambda_k$ for all $i\neq k$. From $\hat{x}_{i,\tau}-x_i=\sum_{j=1}^{n}\phi_{j,\tau}x_j$ we obtain $\hat{x}_{i,\tau}-x_i=\sum_{k\neq i}\frac{x_k'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i}{\lambda_i-\lambda_k}x_k+\phi_{i,\tau}x_i$; (10) follows from setting $\phi_{i,\tau}=0$. Finally, consider (11). Using the results above, it holds that

$$\begin{aligned}\hat{\gamma}_{i,\tau} &= \hat{\lambda}_{i,\tau}^{1/2}\hat{x}_{i,\tau}=\lambda_i^{1/2}\left[1+\frac{\hat{\lambda}_{i,\tau}-\lambda_i}{2\lambda_i}+O_p\left(\left\|\hat{\lambda}_{i,\tau}-\lambda_i\right\|^2\right)\right]\left[x_i+\left(\hat{x}_{i,\tau}-x_i\right)\right]\\ &= \lambda_i^{1/2}x_i+\lambda_i^{1/2}\left(\hat{x}_{i,\tau}-x_i\right)+\frac{\hat{\lambda}_{i,\tau}-\lambda_i}{2\lambda_i^{1/2}}x_i+O_p\left(T^{-1}\right),\end{aligned}$$

which, combining (9) and (10), yields (11).

We now turn to deriving the bias for $\hat{\lambda}_{i,\tau}-\lambda_i$, presented in (12). Expanding (31) and premultiplying by x_i' we obtain

$$\begin{aligned}\hat{\lambda}_{i,\tau}-\lambda_i &= x_i'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i-\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)x_i'\left(\hat{x}_{i,\tau}-x_i\right)+x_i'\left(\hat{\Sigma}_\tau-\Sigma\right)\left(\hat{x}_{i,\tau}-x_i\right)\\ &= x_i'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i-I+II.\end{aligned}$$

From (10), $I = x_i' \sum_{k\neq i} \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i}{\lambda_i - \lambda_k} x_k = 0$. Also (focusing on first order terms only):

$$
\begin{aligned}
II &= \left[x_i' \otimes (\hat{x}_{i,\tau} - x_i)'\right] vec\left(\hat{\Sigma}_\tau - \Sigma\right) \\
&= \left[x_i' \otimes \sum_{k\neq i} \frac{x_k'}{\lambda_i - \lambda_k} x_k' \left(\hat{\Sigma}_\tau - \Sigma\right) x_i\right] vec\left(\hat{\Sigma}_\tau - \Sigma\right) \\
&= \sum_{k\neq i} \left[x_i' \otimes \frac{x_k'}{\lambda_i - \lambda_k}\right] \left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right] \left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right]' \left[x_k \otimes x_i\right].
\end{aligned}
$$

The bias of $\hat{\lambda}_{i,\tau} - \lambda_i$ is given by II, with

$$
\begin{aligned}
& E\left[T x_i' \left(\hat{\Sigma}_\tau - \Sigma\right)(\hat{x}_{i,\tau} - x_i)\right] \qquad (33) \\
=& \sum_{k\neq i} \left[\frac{x_i' \otimes x_k'}{\lambda_i - \lambda_k}\right] E\left\{\left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right]\left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right]'\right\}\left[x_k \otimes x_i\right] \\
=& \sum_{k\neq i} \frac{(x_i' \otimes x_k') V_\Sigma (x_k \otimes x_i)}{\lambda_i - \lambda_k}.
\end{aligned}
$$

The bias for $\hat{x}_{i,\tau} - x_i$ can be derived from (10) following similar passages. Using (31), $\lambda_k \phi_{k,\tau} + x_k' \left(\hat{\Sigma}_\tau - \Sigma\right) x_i$ $+ x_k' \left(\hat{\Sigma}_\tau - \Sigma\right)(\hat{x}_{i,\tau} - x_i) = \lambda_i \phi_{k,\tau} + \left(\hat{\lambda}_{i,\tau} - \lambda_i\right)\phi_{k,\tau}$, whence

$$
\begin{aligned}
\phi_{k,\tau} &= \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i + x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)(\hat{x}_{i,\tau} - x_i)}{\lambda_i - \lambda_k + \left(\hat{\lambda}_{i,\tau} - \lambda_i\right)} \\
&= \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i + x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)(\hat{x}_{i,\tau} - x_i)}{(\lambda_i - \lambda_k)} + o_p(1).
\end{aligned}
$$

Thus, since $\hat{x}_{i,\tau} - x_i = \sum_{k\neq i} \phi_{k,\tau} x_k$, it holds that $\hat{x}_{i,\tau} - x_i = \sum_{k\neq i} \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i}{\lambda_i - \lambda_k} x_k + \sum_{k\neq i} \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)(\hat{x}_{i,\tau} - x_i)}{\lambda_i - \lambda_k} x_k$ $+ o_p(1)$. The bias is given by the second term, with

$$
\begin{aligned}
& E\left[T \sum_{k\neq i} \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)(\hat{x}_{i,\tau} - x_i)}{\lambda_i - \lambda_k} x_k\right] \\
=& \sum_{k\neq i}\sum_{j\neq i} \frac{(x_k' \otimes x_j') E\left\{T\left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right]\left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right]'\right\}(x_j \otimes x_i)}{(\lambda_i - \lambda_k)(\lambda_i - \lambda_j)} x_k \\
=& \sum_{k\neq i}\sum_{j\neq i} \frac{(x_k' \otimes x_j') V_\Sigma (x_j \otimes x_i)}{(\lambda_i - \lambda_k)(\lambda_i - \lambda_j)} x_k.
\end{aligned}
$$

Finally, we consider the case of a "low-rank perturbation", i.e. of a break that affects only a subspace of the matrix Σ. Formally, this entails that

$$
\Sigma = \begin{cases} \Sigma_0 = \sum_{i=1}^n \lambda_i x_i x_i' \text{ for } t = 1, ..., k \\ \Sigma_0 + \Delta_\Sigma \text{ for } t = k+1, ..., T \end{cases},
$$

where, as mentioned in Remark P1.4, we define $\Delta_\Sigma = \sum_{i\in C} \tilde{\lambda}_i \tilde{x}_i \tilde{x}_i' - \sum_{i\in C} \lambda_i x_i x_i'$. Here, C is a family of indices, $\tilde{\lambda}_i$ and $\tilde{x}_i$ are the "new" eigenvalues and eigenvectors. Recall that $\tilde{x}_i' \tilde{x}_j = \delta_{ij}$ and that $\tilde{x}_i' x_j = 0$ for all $i \in C$ and $j \notin C$. Given that there is no change in the couple (λ_i, x_i) when $i \notin C$, throughout $t = 1, ..., T$ it holds that $\Sigma_0 x_i = \lambda_i x_i$. We start by showing that the estimate of λ_i for $i \notin C$ is not affected by the break in Σ. Consider

(31), and multiply both sides by x_i': we get

$$x_i'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i+x_i'\left(\hat{x}_{i,\tau}-x_i\right)=\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)+\left(\hat{\lambda}_{i,\tau}-\lambda_i\right)x_i'\left(\hat{x}_{i,\tau}-x_i\right).$$

Note however that

$$\hat{x}_{i,\tau}-x_i=\sum_{k\neq i}\frac{x_k'\left(\hat{\Sigma}-\Sigma\right)\hat{x}_i}{\hat{\lambda}_i-\lambda_k}x_k, \tag{34}$$

so that $x_i'\left(\hat{x}_{i,\tau}-x_i\right)=0$. Thus, we have $\hat{\lambda}_{i,\tau}-\lambda_i=x_i'\left(\hat{\Sigma}_\tau-\Sigma\right)x_i$. Consider the full sample estimator $\hat{\Sigma}$ for the sake of notational simplicity; note that

$$\begin{aligned}\hat{\Sigma} &= \frac{1}{T}\sum_{t=1}^{T}\left[y_ty_t'-E\left(y_ty_t'\right)\right]+\frac{1}{T}\sum_{t=1}^{T}E\left(y_ty_t'\right) \\ &= \frac{1}{T}\sum_{t=1}^{T}\left[y_ty_t'-E\left(y_ty_t'\right)\right]+\Sigma_0+\frac{T-(k+1)}{T}\Delta_\Sigma,\end{aligned} \tag{35}$$

so that $\hat{\Sigma}-\Sigma_0=O_p\left(T^{-1/2}\right)+\frac{T-(k+1)}{T}\Delta_\Sigma$. However, $x_i'\Delta_\Sigma=0$ for all $i\notin C$, by definition of Δ_Σ. Therefore, $\hat{\lambda}_{i,\tau}-\lambda_i=x_i'\left(\hat{\Sigma}_\tau-\Sigma_0\right)x_i$. The same holds for the estimated eigenvectors; indeed, by (34), and again considering the full sample estimator we have:

$$\begin{aligned}\hat{x}_i-x_i &= \sum_{k\neq i}\frac{x_k'\left(\hat{\Sigma}-\Sigma\right)x_i}{\hat{\lambda}_i-\lambda_k}x_k+\sum_{k\neq i}\frac{x_k'\left(\hat{\Sigma}-\Sigma\right)\left(\hat{x}_i-x_i\right)}{\hat{\lambda}_i-\lambda_k}x_k \\ &= \sum_{k\neq i}\frac{x_k'\left(\hat{\Sigma}-\Sigma_0\right)x_i}{\hat{\lambda}_i-\lambda_k}x_k+\sum_{k\neq i}\frac{x_kx_k'\left(\hat{\Sigma}-\Sigma\right)}{\hat{\lambda}_i-\lambda_k}\left(\hat{x}_i-x_i\right),\end{aligned}$$

whence

$$\left(\hat{x}_i-x_i\right)=\left[I_n-\sum_{k\neq i}\frac{x_kx_k'}{\hat{\lambda}_i-\lambda_k}\left(\hat{\Sigma}-\Sigma\right)\right]^{-1}\left[\sum_{k\neq i}\frac{x_k'\left(\hat{\Sigma}-\Sigma_0\right)x_i}{\hat{\lambda}_i-\lambda_k}x_k\right];$$

given that $\left\|\hat{\Sigma}-\Sigma\right\|=O_p\left(T^{-1/2}\right)$, we recover (10). ■

Proof of Theorem 3. The proof of (16) follows from (1), Theorem 2 and the CMT. As far as (17) is concerned, the proof is based on the proof of Theorem A.4.1 in Csorgo and Horvath (1997, p. 368-370). Here we summarize the main steps, using, as a leading example, $\dot{\Lambda}(\tau)=\frac{1}{\sqrt{T\tau(1-\tau)}}\left[\bar{S}(\tau)'\tilde{V}_{\Sigma,\tau}^{-1}\bar{S}(\tau)\right]^{1/2}$, where $\bar{S}(\tau)=S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)$. We also define $\ddot{\Lambda}(\tau)=\frac{1}{\sqrt{T\tau(1-\tau)}}\left[\bar{S}(\tau)'V_\Sigma^{-1}\bar{S}(\tau)\right]^{1/2}$; further, letting $B_{1i}(\tau)$ be a sequence of standard, independent Brownian bridges for $i=1,...,n^2$, we define $M(\tau)=\left[\sum_{i=1}^{n^2}\frac{B_{1i}^2(\tau)}{\tau(1-\tau)}\right]^{1/2}$. The Darling-Erdos Theorem (see e.g. Corollary A.3.1 in Csorgo and Horvath, 1997, p. 366) states that $P\left[a_T\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}M(\tau)\leq x+b_T\right]=e^{-2e^{-x}}$, where the norming constants a_T and b_T are defined in the Theorem. In order to prove (17), it is enough to show that $\left|\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}\dot{\Lambda}(\tau)-\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}M(\tau)\right|=o_p\left[(\ln\ln T)^{-1/2}\right]$. By virtue of Theorem 2, this entails that, as far as the estimated long-run covariance matrix is concerned, we need to have $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\left\|\tilde{V}_{\Sigma,\tau}-V_\Sigma\right\|=o_p\left[(\ln\ln T)^{-1/2}\right]$. This holds, by virtue of equation (6), if both $\frac{\sqrt{\ln\ln T}}{m}\to0$ and $\frac{m}{\sqrt{T}}\sqrt{\ln\ln T}\to0$, whence the restrictions on m in the statement of the Theorem. Under such restrictions, it suffices to prove that

$$\left|\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}\ddot{\Lambda}(\tau)-\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}M(\tau)\right|=o_p\left(\frac{1}{\sqrt{\ln\ln T}}\right). \tag{36}$$

In order to show (36), note first that (2) yields the (weak) result

$$\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\left|\ddot{\Lambda}(\tau)-M(\tau)\right|=o_p\left(\sqrt{\ln\ln T}\right). \tag{37}$$

Indeed, (2) entails

$$\sup_{u(T,\varepsilon)\leq\tau\leq\frac{1}{2}}\left[\lfloor T\tau\rfloor\right]^{\delta}\left|\ddot{\Lambda}(\tau)-M(\tau)\right| = o_p(1), \tag{38}$$

$$\sup_{\frac{1}{2}\leq\tau\leq 1-u(T,\varepsilon)}\left[\lfloor T(1-\tau)\rfloor\right]^{\delta}\left|\ddot{\Lambda}(\tau)-M(\tau)\right| = o_p(1), \tag{39}$$

for all sequences $u(T,\varepsilon)$ such that $u(T,\varepsilon)\to 0$ and $Tu(T,\varepsilon)\to\infty$ as $T\to\infty$; here, ε is a number between 0 and 1. Choosing $Tu(T,\varepsilon)=e^{(\ln T)^{\varepsilon}}$, and applying Theorem A.3.1 in Csorgo and Horvath (1997, p. 363) it holds that

$$\frac{1}{\sqrt{2\ln\ln T}}\sup_{\frac{1}{T}\leq\tau\leq u(T,\varepsilon)}M(\tau)\xrightarrow{p}\sqrt{\varepsilon}, \tag{40}$$

$$\frac{1}{\sqrt{2\ln\ln T}}\sup_{1-u(T,\varepsilon)\leq\tau\leq 1-\frac{1}{T}}M(\tau)\xrightarrow{p}\sqrt{\varepsilon}.$$

Hence, from (37)

$$\frac{1}{\sqrt{2\ln\ln T}}\sup_{\frac{1}{T}\leq\tau\leq u(T,\varepsilon)}\ddot{\Lambda}(\tau)\xrightarrow{p}\sqrt{\varepsilon},$$

$$\frac{1}{\sqrt{2\ln\ln T}}\sup_{1-u(T,\varepsilon)\leq\tau\leq 1-\frac{1}{T}}\ddot{\Lambda}(\tau)\xrightarrow{p}\sqrt{\varepsilon}.$$

Defining $\xi(T)$ and $\eta(T)$ as $\sup_{1\leq\lfloor T\tau\rfloor\leq T}M(\tau)=M[\xi(T)]$ and $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\ddot{\Lambda}(\tau)=\ddot{\Lambda}[\eta(T)]$, the relationships above entail $P[u(T,\varepsilon)\leq\xi(T),\eta(T)\leq 1-u(T,\varepsilon)]=1$ as $T\to\infty$. Indeed, using (40) as an illustrative example, as $T\to\infty$ and $\varepsilon\to 0$

$$P\left[a_T\sup_{\frac{1}{T}\leq\tau\leq u(T,\varepsilon)}M(\tau)-b_T\geq -K\right]=P\left[\left(\sqrt{\varepsilon}-1\right)\ln\ln T\geq -K\right]=0,$$

for some $K>0$. Hence, (38) and (39) entail

$$\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\left|\ddot{\Lambda}(\tau)-M(\tau)\right|=o_p\left(e^{-\delta\ln^{\varepsilon}T}\right),$$

and since $\left|\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\ddot{\Lambda}(\tau)-\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}M(\tau)\right|\leq\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\left|\ddot{\Lambda}(\tau)-M(\tau)\right|$, (36) follows in view of $\sqrt{\ln\ln T}\,e^{-\delta\ln^{\varepsilon}T}\to 0$. ■

Proof of Theorem 4. In order to prove (20), we show that, under $H_a^{(T)}$

$$P\left[\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\Lambda_T(\tau)>c_{\alpha,T}\right]=P\left[\Lambda_0>c_{\alpha,T}-NC_T\right],$$

where Λ_0 is the distribution of $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\Lambda_T(\tau)$ under the null of no change and NC_T is a non-centrality parameter. Tests based on $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\Lambda_T(\tau)$ are consistent as long as $c_{\alpha,T}-NC_T\to-\infty$ as $T\to\infty$.

Note that

$$\Lambda_T(\tau)=\frac{T}{\lfloor T\tau\rfloor\lfloor T(1-\tau)\rfloor}\tilde{S}(\tau)'\hat{V}_{\Sigma}^{-1}\tilde{S}(\tau),$$

where we consider Assumption 2*(i)* and the full sample estimator of V_{Σ}^{-1} only, for simplicity. Consider $\tilde{S}(\tau)$.

Under $H_a^{(T)}$

$$\begin{aligned}
\sqrt{\frac{T}{\lfloor T\tau\rfloor \lfloor T(1-\tau)\rfloor}}\tilde{S}(\tau) &= \sqrt{\frac{T}{\lfloor T\tau\rfloor \lfloor T(1-\tau)\rfloor}}\left[\sum_{t=1}^{\lfloor T\tau\rfloor}\bar{w}_t - \frac{\lfloor T\tau\rfloor}{T}\sum_{t=1}^{T}\bar{w}_t\right] \\
&\quad +\Delta_T\sqrt{\frac{T}{\lfloor T\tau\rfloor \lfloor T(1-\tau)\rfloor}}\left[\sum_{t=1}^{\lfloor T\tau\rfloor}I(t\leq k_{0,T}) - \frac{\lfloor T\tau\rfloor}{T}\sum_{t=1}^{T}I(t\leq k_{0,T})\right] \\
&= \tilde{S}_1(\tau)+\tilde{S}_2(\tau),
\end{aligned}$$

where $I(\cdot)$ is the indicator function. The sequence $\bar{w}_t$ is zero mean, and it satisfies the assumptions of Theorem 1; thus, $\tilde{S}_1(\tau)$ follows the null distribution as $T\to\infty$. As far as $\tilde{S}_2(\tau)$ is concerned, we have

$$\tilde{S}_2(\tau)=\Delta_T\sqrt{\frac{T}{\lfloor T\tau\rfloor \lfloor T(1-\tau)\rfloor}}\left[\left(\frac{\lfloor T(1-\tau)\rfloor}{T}k_{0,T}\right)I(k_{0,T}<\lfloor T\tau\rfloor)+\left(\frac{T-k_{0,T}}{T}\lfloor T\tau\rfloor\right)I(k_{0,T}\geq\lfloor T\tau\rfloor)\right],$$

with

$$\begin{aligned}
&\sup_{1\leq\lfloor T\tau\rfloor\leq T}\Delta_T\sqrt{\frac{T}{\lfloor T\tau\rfloor \lfloor T(1-\tau)\rfloor}}\left[\left(\frac{\lfloor T(1-\tau)\rfloor}{T}k_{0,T}\right)I(k_{0,T}<\lfloor T\tau\rfloor)\right. \\
&\left.+\left(\frac{T-k_{0,T}}{T}\lfloor T\tau\rfloor\right)I(k_{0,T}\geq\lfloor T\tau\rfloor)\right] \\
&= \Delta_T\sqrt{k_{0,T}\left(\frac{T-k_{0,T}}{T}\right)}. \qquad (41)
\end{aligned}$$

Turning to $\hat{V}_{\Sigma,\tau}^{-1}$, we show that, under $H_a^{(T)}$, $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\left\|\hat{V}_{\Sigma,\tau}-V_\Sigma\right\|$ is bounded in probability. Consider $\hat{\Sigma}$; it holds that $vec\left(\hat{\Sigma}\right)=vec(\Sigma_t)+\left[\frac{T-k_{0,T}}{T}-I(t\geq k_{0,T})\right]\Delta_T+o_p(1)$, where the $o_p(1)$ term comes from a LLN. Therefore

$$\begin{aligned}
\hat{V}_\Sigma &= \frac{1}{T}\sum_{t=1}^{T}\bar{w}_t\bar{w}_t' - \frac{1}{T}\sum_{t=1}^{T}\bar{w}_t\left[\frac{T-k_{0,T}}{T}-I(t\geq k_{0,T})\right]\Delta_T' \\
&\quad -\frac{1}{T}\sum_{t=1}^{T}\left[\frac{T-k_{0,T}}{T}-I(t\geq k_{0,T})\right]\Delta_T\bar{w}_t' \\
&\quad +\frac{1}{T}\sum_{t=1}^{T}\left[\frac{T-k_{0,T}}{T}-I(t\geq k_{0,T})\right]^2\Delta_T\Delta_T' \\
&= I+II+III+IV.
\end{aligned}$$

The LLN entails that $I\stackrel{p}{\to}V_\Sigma$; II and III have the same order of magnitude as each other. Particularly, since $\sum_{t=1}^{T}\bar{w}_t\left[\frac{T-k_{0,T}}{T}-I(t\geq k_{0,T})\right]=O_p\left(\sqrt{T}\right)$, $II=O_p\left(\frac{\|\Delta_T\|}{\sqrt{T}}\right)$. Finally

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\left[\frac{T-k_{0,T}}{T}-I(t\geq k_{0,T})\right]^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(\frac{T-k_{0,T}}{T}\right)^2-2\left(\frac{T-k_{0,T}}{T}\right)^2+\frac{1}{T}\sum_{t=1}^{T}I(t\geq k_{0,T}) \\
&= \frac{k_{0,T}}{T}\frac{T-k_{0,T}}{T},
\end{aligned}$$

thus, $IV=O_p\left(\frac{k_{0,T}}{T}\frac{T-k_{0,T}}{T}\|\Delta_T\|^2\right)$, which is $O_p(1)$ under $H_a^{(T)}$. This entails that $\left\|\hat{V}_\Sigma-V_\Sigma\right\|=O_p(1)$ under $H_a^{(T)}$. Applying Taylor's expansion, we can write $\hat{V}_\Sigma^{-1}=V_\Sigma^{-1}+\mathring{V}_\Sigma^{-1}\left(\hat{V}_\Sigma-V_\Sigma\right)\hat{V}_\Sigma^{-1}$, for some invertible matrix

$\mathring{V}_\Sigma$.

Putting all together, we have

$$\begin{aligned}\Lambda_T(\tau) &= \tilde{S}_1(\tau)' V_\Sigma^{-1}\tilde{S}_1(\tau) + \tilde{S}_2(\tau)'\hat{V}_\Sigma^{-1}\tilde{S}_2(\tau) \\ &\quad +2\tilde{S}_1(\tau)'\hat{V}_\Sigma^{-1}\tilde{S}_2(\tau) + \tilde{S}_1(\tau)'\mathring{V}_\Sigma^{-1}\left(\hat{V}_\Sigma - V_\Sigma\right)\mathring{V}_\Sigma^{-1}\tilde{S}_1(\tau) \\ &= I + II + III + IV.\end{aligned}$$

Term I follows the null distribution under $H_a^{(T)}$, i.e. $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\tilde{S}_1(\tau)' V_\Sigma^{-1}\tilde{S}_1(\tau) \xrightarrow{d} \Lambda_0$ as $T\to\infty$. Given that $\hat{V}_\Sigma$ is $O_p(1)$ under $H_a^{(T)}$, term II has the same order as $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\left\|\tilde{S}_2(\tau)\right\|^2$, which is $O\left(k_{0,T}\frac{T-k_{0,T}}{T}\|\Delta_T\|^2\right)$ in view of (41). Terms III and IV are of smaller order of magnitude than II: e.g. as far as III is concerned, it holds that $E\left[\tilde{S}_1(\tau)'\hat{V}_\Sigma^{-1}\tilde{S}_2(\tau)\right] \leq \left(E\left\|\tilde{S}_1(\tau)\right\|^2\right)^{1/2}\left(E\left\|\tilde{S}_2(\tau)\right\|^2\right)^{1/2}$, since $\hat{V}_\Sigma^{-1}$ is $O_p(1)$; thus, $\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\tilde{S}_1(\tau)'\hat{V}_\Sigma^{-1}\tilde{S}_2(\tau) = O\left(\sqrt{\ln\ln T}\sqrt{k_{0,T}\frac{T-k_{0,T}}{T}}\|\Delta_T\|\right)$, which is smaller than II, as $T\to\infty$, due to (19). Therefore, under $H_a^{(T)}$, $P\left[\sup_{n\leq\lfloor T\tau\rfloor\leq T-n}\Lambda_T(\tau) > c_{\alpha,T}\right] = P\left[\Lambda_0 > c_{\alpha,T} - NC_T\right]$, with

$$NC_T = \|\Delta_T\|\sqrt{k_{0,T}\left(\frac{T-k_{0,T}}{T}\right)} + o\left[\|\Delta_T\|\sqrt{k_{0,T}\left(\frac{T-k_{0,T}}{T}\right)}\right].$$

In view of $c_{\alpha,T}$ being $O\left(\sqrt{\ln\ln T}\right)$ and of (19), it holds that $c_{\alpha,T} - NC_T \to -\infty$ as $T\to\infty$, whence (20) follows. ■

Proof of Theorem 5. By Lemma 6, it suffices to show that $\left|\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\mathring{\Lambda}_{iT}^{\varepsilon}(\tau) - \sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}} M^{\varepsilon}(\tau)\right| = o_p\left[(\ln\ln T)^{-1/2}\right]$, where $M^{\varepsilon}(\tau) = \left[\sum_{i=1}^{p}\frac{B_{1i}^2(\tau)}{\tau(1-\tau)}\right]^{1/2}$, $B_{1i}(\tau)$ is a sequence of standard, independent Brownian bridges for $i=1,...,p$, and

$$\mathring{\Lambda}_{iT}^{\varepsilon}(\tau) = \sqrt{\frac{T}{\lfloor T\tau\rfloor\times\lfloor T(1-\tau)\rfloor}}\times\left[\tilde{S}^{\varepsilon}(\tau)'(V_{\varepsilon,\tau})^{-1}\tilde{S}^{\varepsilon}(\tau)\right]^{1/2}.$$

Indeed, we can show that $\left|\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\ddot{\Lambda}_{iT}^{\varepsilon}(\tau) - \sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\mathring{\Lambda}_{iT}^{\varepsilon}(\tau)\right| = O_p\left(\frac{\ln\ln T}{T}\right)$, where

$$\ddot{\Lambda}_{iT}^{\varepsilon}(\tau) = \sqrt{\frac{T}{\lfloor T\tau\rfloor\times\lfloor T(1-\tau)\rfloor}}\times\left[\ddot{S}^{\varepsilon}(\tau)'(V_{\varepsilon,\tau})^{-1}\ddot{S}^{\varepsilon}(\tau)\right]^{1/2},$$

and $\ddot{S}^{\varepsilon}(\tau) = \sum_{t=1}^{\lfloor T\tau\rfloor} vec(\varepsilon_t\varepsilon_t')$ with $\ddot{S}^{\varepsilon}(\tau) = 0$ for $\tau\leq\frac{1}{T}$ or $\geq 1-\frac{1}{T}$. Since $\left|\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\ddot{\Lambda}_{iT}^{\varepsilon}(\tau) - \sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\mathring{\Lambda}_{iT}^{\varepsilon}(\tau)\right| \leq \sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\left|\ddot{\Lambda}_{iT}^{\varepsilon}(\tau) - \mathring{\Lambda}_{iT}^{\varepsilon}(\tau)\right|$, the passages in the proof of Lemma 6 and the same logic as in the proof of Theorem 2 yield $\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\left|\ddot{\Lambda}_{iT}^{\varepsilon}(\tau) - \mathring{\Lambda}_{iT}^{\varepsilon}(\tau)\right| = O_p\left(\frac{\ln\ln T}{T}\right)$. Thus, it suffices to show that $\left|\sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}}\ddot{\Lambda}_{iT}^{\varepsilon}(\tau) - \sup_{\frac{1}{T}\leq\tau\leq 1-\frac{1}{T}} M^{\varepsilon}(\tau)\right| = o_p\left[(\ln\ln T)^{-1/2}\right]$. This can be shown following the passages in the proof of Theorem 3, using the SIP in part (ii) of Lemma 5 and Lemma 7. The behaviour of the test statistic under the alternative can be shown following the same lines as the proof of Theorem 4. ■