TESTING FOR INSTABILITY IN COVARIANCE STRUCTURES

Chihwa Kao, Lorenzo Trapani, and Giovanni Urga

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

April 2012
[Revision of August 2011]

$5.00

Up-to-date information about CPR’s research projects and other activities is available from our World Wide Web site at www-cpr.maxwell.syr.edu. All recent working papers and Policy Briefs can be read and/or printed from there as well.
CENTER FOR POLICY RESEARCH – Spring 2012

Christine L. Himes, Director
Maxwell Professor of Sociology

Associate Directors
Margaret Austin
Associate Director
Budget and Administration

Douglas Wolf
Gerald B. Cramer Professor of Aging Studies
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi Baltagi ................................................. Economics
Robert Bifulco.................................Public Administration
Leonard Burman ..Public Administration/Economics
Thomas Dennison .................Public Administration
William Duncombe ..........Public Administration
Gary Engelhardt .........................Economics
Madonna Harrington Meyer ........... Sociology
William C. Horrace .............Economics
Duke Kao..........................Economics
Eric Kingson .......Social Work
Sharon Kioko.........Public Administration
Thomas Kniesner .............Economics
Jeffrey Kubik .....................Economics
Andrew London ..........Sociology
Len Lopoo ..........Public Administration

Amy Lutz............................................. Sociology
Jerry Miner...................................... Economics
Jan Ondrich.................................Economics
John Palmer .........................Public Administration
David Popp ..........Public Administration
Gretchen Purser .................Sociology
Christopher Rohlf.....................Economics
Stuart Rosenthal ..........Economics
Ross Rubenstein ..........Public Administration
Perry Singleton ..........Economics
Margaret Usdansky ..........Sociology
Michael Wasylenko ..........Economics
Jeffrey Weinstein ..........Economics
Peter Wilcoxon ....Public Administration/Economics
Janet Wilmoth ..........Sociology

GRADUATE ASSOCIATES

Kanika Arora ......................Public Administration
Christopher Bianchi ....................Public Administration
Christian Buerger .................Public Administration
Il Hwan Chung ..............Public Administration
Alissa Dubnicki ............Economics
Alexander Falevich ..........Economics
Andrew Friedson ..........Economics
Pallab Ghosh .................Economics
Lincoln Groves ...............Public Administration
Clorise Harvey ..........Public Administration
Hee Seung Lee ..........Public Administration
Jae Yoon Lee ..........Economics
Jing Li ....................Economics

Shimeng Liu ...................Economics
Yan Liu ......................Sociology
Allison Marier ..........Economics
Qing Miao ..............Public Administration
Gabrielle Mirsaidi ..........Public Administration
Nuno Abreu Faro E Mota ..........Economics
Wael Moussa ..........Economics
Kerri Raissian ..........Public Administration
Heather Ruby ..........Public Administration
Natalee Simpson ..........Sociology
Tian Tang ..........Public Administration
Liu Tian ..........Economics
Pengju Zhang ..........Public Administration

STAFF

Kelly Bogart....................Administrative Specialist
Karen Cimilluca...............Office Coordinator
Alison Kirsche...............Administrative Secretary
Kitty Nasto....................Administrative Secretary
Candi Patterson...............Computer Consultant
Mary Santy...............Administrative Secretary
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 Principle and Law of Large Numbers, we normalize 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. The testing procedure is validated through an application to 18 US interest rates over 1997-2011, finding instability at the end-2007/beginning-2008.

JEL No. C1, C22, C5

Key Words: Covariance Matrix, Eigensystem, Changepoint, Term Structure of Interest Rates, CUSUM statistic

Authors: Chihwa Kao, Syracuse University; Lorenzo Trapani, Cass Business School; Giovanni Urga, Cass Business School and Università di Bergamo

We wish to thank the participants to the Faculty of Finance Workshops at Cass Business School; to 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; to the Glasgow Business School Seminar Series; to the New York Camp Econometrics VI (Syracuse University, 9-10 April, 2011); to the Rimini Time Series Workshop (Rimini, 2-3 June, 2011). The usual disclaimer applies.

Corresponding author: Centre for Econometric Analysis, Faculty of Finance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ (U.K.). Tel. 00.44.(0)20.70405260 Fax. 00.44.(0)20.70408881, e-mail: L.Trapani@city.ac.uk
1 Introduction

In this paper, we propose a testing procedure to evaluate the structural stability of the covariance matrix of multivariate time series. A large amount of empirical evidence shows that the issue of changepoint detection in a covariance matrix and in its eigensystem 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 change substantially over time. Similar findings, using a different methodology, are in Audrino et al. (2005). Another popular field of application of PCA is the prediction of mortality rates based on the Lee-Carter model (Lee and Carter, 1992; Hyndman and Ullah, 2007). Yang et al. (2010) show that the second principal component of the log mortality rates is subject to changes over time. PCA is also widely used in macroeconometrics, for instance to forecast inflation (Stock and Watson, 1999, 2002, 2005). Finally, the importance of verifying the stability of a covariance matrix is also evident in the context of VAR 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.

Despite the relevance of the topic, these 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). This assumption is unsuitable for financial data, which, in general, are serially dependent, heterogeneous, and for which it is difficult to make distributional assumptions. Thus, testing procedures cannot rely on assuming i.i.d. normal data. Audrino et al. (2005) accommodate for serial dependence through filtering, but this is done at the price of losing the classical interpretation of principal components.

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 break date. 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 (resp. 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 whereby 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, we extend testing for changepoints to PCA; this is useful e.g. when studying the stability of the term structure. 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 proving 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).

The theory derived in our paper is validated through an application to the US term structure of interest rates, in a similar spirit to Perignon and Villa (2006). As expected, we find 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.

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, and the application to the term structure of interest rates is in Section 5. Section 6 concludes.

A word on notation. Limits are denoted as “$\rightarrow$” (the ordinary limit); “$\mathbb{P}$” (convergence in probability); “$\mathbb{L}$” and (convergence in distribution). Orders of magnitude for an almost surely convergent sequence (say $s_T$) are denoted as $O_{a.s.}(T^n)$ and $o_{a.s.}(T^n)$ when, for some $\varepsilon > 0$ and $\tilde{T} < \infty$, $P\left[|T^{-\varepsilon}s_T| < \varepsilon \text{ for all } T \geq \tilde{T}\right] = 1$ and $T^{-\varepsilon}s_T \to 0$ almost surely respectively. Orders of magnitude for a sequence converging in probability (say $s_T'$) are denoted as $O_p(T^n)$ and $o_p(T^n)$ when, for some $\varepsilon > 0$, $\Delta_\varepsilon > 0$ and $\tilde{T}_\varepsilon < \infty$, $P\left[|T^{-\varepsilon}s_T'| > \Delta_\varepsilon \text{ for all } T > \tilde{T}_\varepsilon \text{ and } T^{-\varepsilon}s_T' \to 0\right] < \varepsilon$ 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 $|x|_p$ the $L_p$-norm; the integer part of a real number $x$ is denoted as $\lfloor x \rfloor$.

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 $\Sigma = E(y_t y_t')$. This section contains the asymptotics of the partial sample estimates of $\Sigma$; the results are used in Section 3 in order to construct the CUSUM-type test statistic to test for breaks in $\Sigma$ and its eigensystem. Specifically, we derive a SIP for the partial sample estimators of $\Sigma$ and an estimator of the long run covariance matrix.
of the estimated $\Sigma$, say $V_\Sigma$; and we extend the asymptotics to PCA. All results are derived for $n < \infty$.

**Strong Invariance Principle and estimation of $V_\Sigma$**

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 $[T\tau]$, and we use the subscripts $\tau$ and $1 - \tau$ to denote quantities calculated using the subsamples $t = 1, ..., [T\tau]$ and $t = [T\tau] + 1, ..., T$ respectively. In particular, we consider the sequence of partial sample estimators $\hat{\Sigma}_\tau = (T\tau)^{-1} \sum_{t=[T\tau]+1}^T y_t y_t'$, and similarly $\hat{\Sigma}_{1-\tau} = [T(1-\tau)]^{-1} \sum_{t=[T\tau]+1}^T y_t y_t'$. Finally, henceforth we denote $w_t = \text{vec} (y_t y_t')$ and $\bar{w}_t = \text{vec} (y_t y_t' - \Sigma)$.

In the sequel, we need the following assumption.

**Assumption 1** (i) $\sup_1 E \|y_t\|^{2r} < \infty$ 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 $r > \frac{2\alpha-1}{\alpha}$; (iii) letting $V_{\Sigma,T} = T^{-1} E \left[ \left( \sum_{t=1}^T w_t \right) \left( \sum_{t=1}^T \bar{w}_t \right) \right]$, $V_{\Sigma,T}$ is positive definite uniformly in $T$, and as $T \to \infty$, $V_{\Sigma,T} \to V_{\Sigma}$ with $\|V_{\Sigma}\| < \infty$; (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 $\Omega = \{\omega_{ij}\}$ such that $T^{-1} E [S_{iT,m} S_{jT,m} - \omega_{ij}] \leq MT^{-\psi}$, for all $i$ and $j$ and uniformly in $m$, with $M$ a constant and $\psi > 0$.

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 $\alpha$), and its largest existing moment: as $\alpha$ (the memory of $y_t$) approaches 1, $r$ has to increase. 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 Assumption 1(i)-(iii), as $T \to \infty$

\[
\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\tau]} \bar{w}_t \overset{d}{\to} \frac{[V_{\Sigma}]^{1/2} W_{n2}(\tau)}{},
\] (1)
uniformly in $\tau$. Redefining $\tilde{w}_t$ in a richer probability space, under Assumptions 1(i)-(iv)

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

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

**Remarks**

T1.1 Equation (1) is an IP for $\tilde{w}_t$ (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 $\tilde{w}_t$ can be replaced with a sequence of i.i.d. normally distributed random variables, with the same long run variance as $\tilde{w}_t$. The rate $\delta$ could, in principle, be derived under different assumptions on the dependence of $y_t$. A typical finding is $\delta = \frac{1}{2} \left(1 - \frac{1}{T}\right)$ - see Shorack and Wellner (1986).

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

$$\hat{V}_{\Sigma,r} = \frac{1}{T} \sum_{t=1}^{T} \tilde{w}_t \tilde{w}'_t - \left\{ r \left[\text{vec}\left(\tilde{\Sigma}_r\right)\right] \left[\text{vec}\left(\tilde{\Sigma}_r\right)'\right] + (1 - r) \left[\text{vec}\left(\tilde{\Sigma}_{1-r}\right)\right] \left[\text{vec}\left(\tilde{\Sigma}_{1-r}\right)'\right]\right\}.$$

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

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

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

In order to derive the asymptotics of $\hat{V}_{\Sigma,r}$ and $\hat{V}_{\Sigma,r}$, 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) $\sup_t E\|y_t\|^{2r} < \infty$ for some $r > 2$; (iii) letting $\Omega_T = T^{-1} E\left\{\sum_{t=1}^{T} \text{vec}[\tilde{w}_t\tilde{w}'_t - E(\tilde{w}_t\tilde{w}'_t)]\right\}$, $\Omega_T$ is positive definite uniformly in $T$, and $\Omega_T \to \Omega$ with $\|\Omega\| < \infty$. 

5
Assumption 2 encompasses various possible cases. Part (i)(a) considers the basic, non-autocorrelated case, for which both $\hat{V}_\Sigma$ and $\tilde{V}_{\Sigma,\tau}$ 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 $\tilde{V}_{\Sigma,\tau}$.

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 $\text{vec} \left[ w_t \bar{w}_t' - E (w_t \bar{w}_t') \right]$.

The consistency of $\tilde{V}_{\Sigma,\tau}$ 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:

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

(4)

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

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

(5)

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

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

(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 $\tilde{V}_{\Sigma,\tau}$ and $\tilde{V}_{\Sigma,\tau}$, as it is usually required in this literature (see e.g. Lemma 2.1.2 in Csorgo and Horvath, 1997, p. 76; see also the proof of Theorem 3 below). The value of $\delta'$, in principle, could be determined. Typically, this would be done by using some maximal inequality for mixingales, and $\delta'$ would be expressed in terms of the $L_p$-norm used to define the NED process (in our notation, from Assumption 1, this would be $2 + \epsilon$). For example, using Theorem 16.11 in Davidson (2005, p. 255) and Lemma 3 in Appendix, it could be shown that, if $\alpha > 3$ in Assumption 1(ii), then $\delta' = 1 - \frac{1}{1+\epsilon/2}$. A similar result is also in Wu (2007).

T2.2 In case of serial dependence, (5) states that it is possible to construct an estimator of $V_\Sigma$ with a rate of convergence. A possible choice could be $m = O \left( T^{\delta'-\kappa} \right)$ for some $\kappa > 0$, although there is no “natural” way of determining $\delta'$, and therefore $\kappa$. The convergence rate in (5) can be refined as in (6). Indeed, Assumptions 2(ii)-(iii) allow for an IP to hold
for partial sums of $\text{vec}[\bar{w}_t w^t_{t-1} - E(\bar{w}_t w^t_{t-1})]$, whence the $O_p(T^{-1/2})$ convergence rate, uniformly in $\tau$.

Estimation of the eigensystem

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

Let the $i$-th eigenvalue-eigenvector couple be defined as $(\lambda_i, x_i)$; the eigenvectors are defined as an orthonormal basis, i.e. $x'_i x_j = \delta_{ij}$, where $\delta_{ij}$ is Kronecker’s delta. Since $\Sigma x_i = \lambda_i x_i$, a natural estimator for $(\lambda_i, x_i)$ is the solution to $\hat{\Sigma} x_i = \hat{\lambda}_i \hat{x}_i$, where $\hat{\lambda}_i$ and $\hat{x}_i$ denote the estimates of $\lambda_i$ and $x_i$ respectively. Similarly, the partial sample estimators of the eigenvalues and eigenvectors are the solutions to $\hat{\Sigma}_\tau \hat{x}_{i,\tau} = \hat{\lambda}_{i,\tau} \hat{x}_{i,\tau}$.

Consider the following assumption.

**Assumption 3.** The matrix $\Sigma$ 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 $(\hat{\lambda}_{i,\tau}, \hat{x}_{i,\tau})$ is derived by treating $\hat{\Sigma}_\tau$ as a perturbation of $\Sigma$, 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 $\Sigma$ is reported in Proposition 1:

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

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

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

where $v_{x,i} = \left[ \sum_{k \neq i} \frac{x_k}{\lambda_i - \lambda_k} \left( x'_k \otimes x'_i \right) \right]$.

**Remarks**

P1.1 Proposition 1 states that the estimation errors $\hat{\lambda}_{i,\tau} - \lambda_i$ and $\hat{x}_{i,\tau} - x_i$ 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 can be compared to related results in Wateraux (1976) and Tyler (1981).

P1.2 The asymptotic covariance matrix of $\sqrt{T} (\hat{x}_{i,\tau} - x_i)$ is $v_{x,i} \Sigma V_{x,i}$; in view of (8), it is singular and has rank $n - 1$. In order to invert $\hat{v}_{x,i} \Sigma \hat{v}^t_{x,i}$, a Moore-Penrose generalised inverse can be used. The validity of this approach can be shown based on Andrews
(1987). Since \( \hat{v}_{x,i} \tilde{V}_t v_{x,i} \), it holds that \( P \left[ r ( \hat{v}_{x,i} \tilde{V}_t v_{x,i} ) \geq r ( v_{x,i} V \Sigma v_{x,i} ) \right] = 1 \) as \( T \to \infty \), where \( r (A) \) denotes the rank of \( A \). Also, by construction and for any \( T \), \( r ( \hat{v}_{x,i} \tilde{V}_t v_{x,i} ) = n - 1 \). Thus, as \( T \to \infty \), \( P \left[ r ( \hat{v}_{x,i} \tilde{V}_t v_{x,i} ) = r ( v_{x,i} V \Sigma v_{x,i} ) \right] = 1 \). This is a sufficient condition that allows to use the Moore-Penrose inverse for \( v_{x,i} V \Sigma v_{x,i} \), e.g., when computing quadratic forms.

P1.3 We show in the appendix that

\[
E \left[ T \left( \hat{\lambda}_{i,t} - \lambda_i \right) \right] = \sum_{k \neq i} \frac{(x'_k \otimes x'_k) V \Sigma (x_k \otimes x_i)}{\lambda_i - \lambda_k}, \tag{9}
\]

as \( T \to \infty \). As far as the impact of \( n \) is concerned, \( V \Sigma \) is an \( n^2 \)-dimensional matrix; thus, in general the quadratic form \( (x'_i \otimes x'_k) V \Sigma (x_k \otimes x_i) \) has magnitude of order \( O \left( n^2 \right) \). Also, due to the summation on the right hand side of (9) involving \( n - 1 \) elements, \( E \left[ T \left( \hat{\lambda}_{i,t} - \lambda_i \right) \right] = O \left( n^{3} \right) \). Thus, the asymptotic bias is of order \( O \left( \frac{n^2}{T} \right) \), and it is always positive for the largest eigenvalue; a bias-corrected version is \( \hat{\lambda}_{i,t} = \hat{\lambda}_{i,t} - T^{-1} \sum_{k \neq i} \frac{(x'_i \otimes x'_k) \tilde{V}_t}{\lambda_i - \lambda_k} \). 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

\[
E \left[ T ( \hat{x}_{i,t} - x_i ) \right] = \sum_{k \neq i} \sum_{j \neq i} \frac{(x'_k \otimes x'_j) V \Sigma (x_j \otimes x_i)}{(\lambda_i - \lambda_k)(\lambda_i - \lambda_j)} x_k, \tag{10}
\]

which provides an expression to correct the bias of the estimated eigenvectors.

P1.4 One may also be interested in the principal components \( \gamma_i = \lambda_i^{1/2} x_i \). A typical interpretation in the context of the term structure of interest rates (Litterman and Scheinkman, 1991; Perignon and Villa, 2006) is that \( \lambda_i \) is the “volatility” of \( \gamma_i \), and \( x_i \) represents its “loading”. It holds that

\[
\hat{\gamma}_{i,t} = \lambda_i^{1/2} \hat{x}_{i,t} = \lambda_i^{1/2} \left[ 1 + \frac{\hat{\lambda}_{i,t} - \lambda_i}{2\lambda_i} + o_p \left( \left( \hat{\lambda}_{i,t} - \lambda_i \right) \right) \right] x_i + (\hat{x}_{i,t} - x_i).
\]

Thus, \( \hat{\gamma}_{i,t} - \gamma_i = v_{\gamma,i} vec (\tilde{\Sigma}_t - \Sigma) + o_p (1) \), with \( v_{\gamma,i} = \frac{1}{\lambda_i^{1/2}} (x'_i \otimes x'_i) + \sum_{k \neq i} \frac{\lambda_k^{1/2}}{\lambda_i - \lambda_k} (x'_i \otimes x'_k) \).

Define \( \lambda = [\lambda_1, ..., \lambda_n]' \) as the \( n \)-dimensional vector containing the eigenvalues sorted in descending order; \( X = [x_1, ..., x_n] \), and \( \Gamma = [\gamma_1, ..., \gamma_n] \); \( z = [\hat{\lambda}', vec \hat{X}', vec \hat{\Gamma}]' \) with \( \hat{\lambda}_t - z = D_{\lambda x' \gamma} vec (\hat{\Sigma}_t - \Sigma) + o_p (1) \) and \( D_{\lambda x' \gamma} = [x_1 \otimes x_1, ..., x_n \otimes x_n, v'_{x,1}, ..., v'_{x,n}, v'_{x,1}, ..., v'_{x,n}, v'_{x,1}, ..., v'_{x,n}]' \).
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 
\[
\left( \frac{\hat{z}_\tau - z}{p(T)} \right) \xrightarrow{d} \left[ V_{\hat{z}} \right]^{1/2} W_n(2n+1) \quad (11)
\]

Also, 
\[
T(\hat{z}_\tau - z) = \sum_{t=1}^{[TR]} X_t + O_{a.s.} \left( (T^{-1/2}) \right), \quad \text{uniformly in } \tau, \quad \text{where } V_{\hat{z}} = D_{\lambda x} V \Sigma D'_{\lambda x} \quad \text{and } \hat{X}_t \text{ is a zero mean, i.i.d. Gaussian sequence with } E(\hat{X}_t) = V_{\hat{z}}.
\]

Corollary 1 entails that 
\[
\left( \frac{\hat{z}_\tau - z}{p(T)} \right) \xrightarrow{d} \left[ V_{\hat{z}} \right]^{1/2} W_n(2n+1) \quad (11)
\]

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

with: 
- \( V_{\lambda} \) a matrix with \((i, j)\)-th element given by 
  \[ V_{ij}^{\lambda} = (x_i' x_j) V \Sigma (x_j' x_j) \]
- \( V_X \) an \((n^2 \times n^2)\)-dimensional matrix whose \((i, j)\)-th \( n \times n \) block is 
  \[ V_{ij}^X = \sum_{k=1}^{n^2} \sum_{h=1}^{n^2} \frac{x_k x_h}{x_k x_h} \] \( V \Sigma x_j x_h \); 
- \( V_T \) is an \((n^2 \times n^2)\)-dimensional matrices whose \((i, j)\)-th \( n \times n \) block is 
  \[ V_{ij}^T = v_{ij} v_{ij}' \].

### 3 Testing

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

Henceforth, we define the CUSUM process 
\[
S(T) = \sum_{t=1}^{[TR]} vec(y_t y_t').
\]

In light of Corollary 1, test statistics for \( \Sigma \) and its eigensystem can be based on

\[
\tilde{S}(\tau) = R \times D_{\lambda x} \times \left[ S(\tau) - \frac{[TR]}{T} \frac{S(T)}{T} \right], \quad (11)
\]

with \( \tilde{S}(\tau) = 0 \) for \( \tau \leq \frac{1}{T} \) or \( \tau \geq 1 - \frac{1}{T} \), and \( R \) a \( p \times (2n+1) \) 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 
\[ D_{\lambda x} \times \left[ S(\tau) - \frac{[TR]}{T} \frac{S(T)}{T} \right]. \]

Thence, testing is carried out by using

\[
\Lambda_T(\tau) = \sqrt{\frac{T}{[TR] \times [T(1-\tau)]}} \times \left[ \tilde{S}(\tau)' \tilde{V}_{\tilde{z},\tau}^{-1} \tilde{S}(\tau) \right]^{1/2}, \quad (12)
\]

with 
\[ \tilde{V}_{\tilde{z},\tau} = RD_{\lambda x} \tilde{V}_{\Sigma,\tau} D'_{\lambda x} R'. \]

Theorem 3 contains the asymptotics of \( \sup_{[TR]} \frac{\Lambda_T(\tau)}{|\tilde{z}|} \) under the null.
Theorem 3 Under Assumptions 1-3, as \( T \to \infty \)
\[
\sup_{[T\tau_1] \leq [T\tau] \leq [T\tau_2]} \Lambda_T (\tau) \to \sup_{\tau_1 \leq \tau \leq \tau_2} \frac{||B_p (\tau)||}{\sqrt{\tau} (1 - \tau)},
\]
where \( B_p (\tau) \) is a \( p \)-dimensional standard Brownian bridge and \([\tau_1, \tau_2] \in (0, 1)\). Also
\[
P \left\{ a_T \left( \sup_{n \leq [T\tau] \leq T-n} \Lambda_T (\tau) \right) \leq x + b_T \right\} \to e^{-2e^{-x}},
\]
where \( a_T = \sqrt{2\ln \ln T} \) and \( b_T = 2\ln \ln T + \frac{\pi}{3} \ln \ln T - \ln \Gamma \left( \frac{p}{2} \right) \), with \( \Gamma (\cdot) \) the Gamma function.

Remarks
Theorem 3 states that (12) can be used in two different ways:

T3.1 According to (13), 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 \( \sup_{1 \leq [T\tau] \leq T} \Lambda_T (\tau) \to 0 \) under \( H_0 \). Thus, trimming is necessary in this case; alternatively, a weighted norm can be employed, defining \( \Lambda_T (\tau) = \frac{S(\tau)[R\tilde{V}_{\tau+R}T\tilde{V}_{\tau}T]}{1-\tau} \tilde{S}(\tau) \) for \( \alpha \in [0, 1]\) (see Csorgo and Horvath, 1997; Chen et al., 2005) and taking \( \sup_{1 \leq [T\tau] \leq T} \Lambda_T (\tau) \to 0 \).

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 \( \sup_{n \leq [T\tau] \leq T-n} [a_T \Lambda_T (\tau) - b_T] \) are designed to be able to detect breaks close to the end of the sample.

Theorem 3 allows to test for breaks in \( \Sigma \) when \( n \) is finite. As \( n \) passes to infinity, Aue et al. (2009) study the behaviour of \( \hat{\Lambda}^2_T \equiv \sup_{1 \leq [T\tau] \leq T} T^{-1} \tilde{S}(\tau)' \left[ R\tilde{V}_{\tau+R}T\tilde{V}_{\tau}T \right]^{-1} \tilde{S}(\tau) \), showing that, as \( T \to \infty \) followed by \( n \to \infty \),
\[
\frac{4\hat{\Lambda}^2_T - n^2}{\sqrt{2n}} \to N(0, 1).
\]

Corollary 2 Let Assumptions 1-3 hold, with \( \sup_{i,t} E [y_{it}] \) < \( \infty \) for \( r > 2 \), and define \( \hat{\Lambda}^2_{n,T} \equiv \sup_{[T\tau_1] \leq [T\tau] \leq [T\tau_2]} \hat{\Lambda}^2_{n,T} (\tau) \), where \( \hat{\Lambda}^2_{n,T} (\cdot) = \frac{T}{[T\tau](T(1-\tau))] \left[ S(\tau) - \frac{T}{T(1-\tau)} S(T) \right]' \tilde{V}_{\tau}\Sigma \tilde{V}_{\tau}' \left[ S(\tau) - \frac{T}{T(1-\tau)} S(T) \right] \). As \( (n, T) \to \infty \) with \( \frac{n}{T^2} \to 0 \), \( \frac{nm}{T^2} \to 0 \) and \( \frac{n}{m} \to 0 \), it holds that
\[
\frac{1}{\sqrt{2n}} \left( \hat{\Lambda}^2_{n,T} - n^2 \right) \to N(0, 1).
\]

The Corollary states that normality holds for large \( T \) and relatively small \( n \). The restriction \( \frac{n}{T^2} \to 0 \) arises from the approximation error in the SIP. As far as both \( \frac{nm}{T^2} \to 0 \) and \( \frac{n}{m} \to 0 \) are concerned, they come from the estimated long run covariance matrix, \( \tilde{V}_{\Sigma} \). Heuristically, each
element in $\hat{V}_2$ has an error of magnitude $O_p\left(\frac{m}{\sqrt{T}}\right) + O_p\left(\frac{1}{n}\right)$; each column contains $n^2$ elements, so that the contribution of the matrix estimation error is $O_p\left(\frac{n^2m}{T}\right) + O_p\left(\frac{n^2}{m}\right)$. This is then normalised by $n$, whence the restrictions.

**Consistency of the test**

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

$$H_a^{(T)}: vech(\Sigma_t) = \begin{cases} vech(\Sigma) & \text{for } t = 1, \ldots, k_0, T \\ vech(\Sigma) + \Delta_T & \text{for } t = k_0, T + 1, \ldots, T \end{cases},$$

(15)

where both the changepoint $(k_0, T)$ and the size of the break $(\Delta_T)$ could depend on $T$. More general alternatives could be considered (see e.g. Andrews, 1993; Csorgo and Horvath, 1997).

Theorem 4 illustrates the dependence of the power on $\Delta_T$ and $k_0, T$.

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

$$\frac{1}{\ln \ln T} \left[ \frac{(T - k_0, T)k_0, T}{T} \|\Delta_T\|^2 \right] \to \infty,$$

(16)

it holds that

$$P\left[\sup_{n \leq |T\tau| \leq T-n} \Lambda_T(\tau) > c_{a,T}\right] = 1.$$

(17)

**Remarks**

T4.1 Theorem 4 illustrates the impact of $k_0, T$ and $\Delta_T$ on the power of tests based on $\sup_{n \leq |T\tau| \leq T-n} \Lambda_T(\tau)$. 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 $\sup_{1 \leq |T\tau| \leq T} T^{-1} \hat{S}(\tau)' \hat{V}_{\tau, r}^{-1} \hat{S}(\tau)$. Using similar algebra as in the proof of Theorem 4, it can be shown that the noncentrality parameter of $\sup_{1 \leq |T\tau| \leq T} T^{-1} \hat{S}(\tau)' \hat{V}_{\tau, r}^{-1} \hat{S}(\tau)$ is proportional to $\|\Delta_T\|^2 \frac{k_0^2}{T}$. 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 $O\left(\sqrt{\ln \ln T}\right)$. When using trimmed statistics such as in (13), 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.
The test can be used in presence of multiple breaks also. Although this goes beyond the scope of this paper, the sequential procedure discussed by Bai (1997) can be used. Upon finding evidence of one break, its date can be estimated as $\hat{T_\tau} = T \times \arg \sup \Lambda_T(\tau)$; this estimator is such that $|\hat{T_\tau} - T_\tau| = O_p(1)$. The sample can be then split around $\hat{T_\tau}$, and testing can be applied to each subsample.

The presence of “large” breaks in $\Sigma$ 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 $\lambda_i$. From Proposition 1, the long-run variance of the estimated eigenvalues is estimated by $(\hat{x}_i^\prime \otimes \hat{x}_i^\prime) \hat{V}_\Sigma (\hat{x}_i \otimes \hat{x}_i)$, 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, $\|\hat{V}_\Sigma - V_\Sigma\| = O_p(1)$ also. This entails that the estimated variance of $\hat{\lambda}_i - \lambda_i$, $(\hat{x}_i^\prime \otimes \hat{x}_i^\prime) \hat{V}_\Sigma (\hat{x}_i \otimes \hat{x}_i)$ is $O_p(1)$. On the other hand, the numerator of the test statistic is proportional to $T \|\hat{\Sigma} - \Sigma\| = O_p(T)$. This entails that the test based on $\sup_{n \leq |T_\tau| \leq T-n} \Lambda_T(\tau)$ is consistent as $T \to \infty$.

4 Monte Carlo evidence

In this section, we discuss: (a) the calculation of critical values, and (b) size and power of the test.

There are two possible approaches to the computation of critical values: either using the EV distribution in (14) or using an approximation similar to that proposed in Csorgo and Horvath (1997, ch. 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 $\alpha$ is based on $c_{\alpha,T} = a_T^{-1} \{b_T - \ln \left[ -\frac{1}{2} \ln (1 - \alpha) \right] \}$. 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

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

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 truncating at $T \times h_{nT} = \ln^{3/2} T$ yields tests with good size properties; our simulations reinforce this finding. 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 $T$. 

12
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 (only for the i.i.d. case) applied to testing for changes in the whole covariance matrix \( \Sigma \). We also assess, by way of comparison, the size and power of Aue et al.’s (2009) test, which is based on

\[
\sup_{1 \leq |T \tau| \leq T} \Lambda_T (\tau) = \sup_{1 \leq |T \tau| \leq T} \frac{1}{\sqrt{T}} \times \left[ \hat{S}(\tau)' \hat{V}_{z,\tau}^{-1} \hat{S}(\tau) \right]^{1/2}.
\]  

(19)

In view of Theorem 4 and the comments thereafter, it can be expected that our test has less power versus mid-sample alternatives than tests based on (19), but it is powerful versus changes close to the beginning/end of sample.

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:

\[
y_t = \rho y_{t-1} + e_t + \theta e_{t-1},
\]  

(20)

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. Finally, 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

\[
\sup_{Th_n T \leq |T \tau| \leq T - Th_n T} \Lambda_T (\tau).
\]  

(21)

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 \( O(Th_n T) \) periods from the beginning/end of sample.
The test 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 (19), which, in small samples, tends to be slightly more conservative.

**Power**

We conduct our simulations under alternatives defined as

$$\left\{ \begin{array}{ll} \Sigma & \text{for } t = 1, \ldots, k \\ \Sigma + \Delta & \text{for } t = k + 1, \ldots, T \end{array} \right.$$  \hspace{1cm} (22)

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

$$\Delta = \sqrt{\frac{\ln\ln T}{T^\beta}} \times I_n.$$  \hspace{1cm} (23)

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

Considering mid-sample alternatives (Tables 2a and 2b), the test has nontrivial power versus “local” alternatives (represented here by the case $\beta = \frac{2}{3}$), and good power when $\beta = \frac{1}{2}$; 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 (19) have better power properties, at least for $T = 200$ or higher. Indeed, tests based on (21) 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 (21) is reduced with respect to tests based on (19), is negligible.

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 it is natural, tests based on (19) have very little power versus beginning of sample alternatives; by construction, such tests do not have power versus changes that occur closer than $O(\sqrt{T})$ 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 $k = Th_{nT} + 1$ being of comparable magnitude.

We note that, in Table 2c, 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, $\hat{\Sigma}$, which is consistent under the alternative but biased. Unreported simulations show that tests based on (19) and (21) exhibit a substantial power gain when using the partial sample estimates, $\tilde{\Sigma}$. However, both tests becomes 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.

Testing for the constancy of $\Sigma$: the role of $n$

We report a simulation exercise for the null of no change in $\Sigma$. The alternative hypothesis is the same as in (22); we refer to Table 3 for the values of $\beta$ employed. Data are generated as i.i.d. Gaussian with $\Sigma = I_n$. 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 (22) and (23). In the last column, $k^* = Th_{nT} + 1$ and $\Delta = I_n$.

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. By way of comparison, the results in Table 3 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 T4.3: 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. the last column of Table 3. 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 (21) are decidedly more powerful than tests based on (19) for large samples ($T = 500$). Comparing the last column of Table 3 with Table 2c, it can be noted that the power of tests for the whole matrix is substantially lower 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 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 $\Sigma = E(y_ty_t')$. 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 $\alpha_P$ is the size of the whole procedure. Using these critical values yields, approximately, a level $\alpha_P$ 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 $\lambda_1$, $\lambda_2$ and $\lambda_3$), and in their loading (verifying the stability of the eigenvectors corresponding to the three largest eigenvalues, denoted $x_1$, $x_2$ and $x_3$). As far as eigenvectors are concerned, (8) ensures that, when running the test, the CUSUM transformation of the estimated $x_i$s has the same sign for all values of $\tau$, thus overcoming the issue of the eigenvectors being defined up to a sign. The singularity of the covariance matrix of the estimated eigenvectors is addressed by using a Moore-Penrose inverse - see also Remark P1.2.

Results for both experiments, at both frequencies, are reported in Table 6.
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, $x_2$, 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 $\Sigma$ 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. The only exception is $\lambda_3$, 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, according to Remark T4.2, as the maximizer of the CUSUM statistic, is January 2008. Table 7 and Figure 2 report the proportion of the total variance explained by each principal component before and after this date, and the eigenvectors of the first and third principal component, before and after mid-April 2008, respectively.

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 eigenvectors corresponding to level and curvature change significantly around the third and the last weeks of March 2008 respectively (possibly due to an “attraction” effect of the variance of the 10-year maturity). The loading of the second principal component has a change, significant at 10%, at
around the last week of June 1999. A closer look at the target FED fund rate reveals that June 1999 was the first time since 1995 (with an exception in 1997) that the FED increased the rate, starting a trend that would continue until late 2000. This does not rule out the possibility that other, less impactful breaks exist. The presence of significant changes in the loadings of each principal component as a result of the 2008 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.

6 Conclusions

This paper proposes 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 $T$. 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. Finally, our theory is applied to the US term structure of interest rates, using 18 maturities and monthly and weekly data. We find evidence of instability in the volatility of the level factor, and in the loadings of all factors, during the first half of 2008.

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. However, we can extend our framework to the case of multiple breaks, along the same lines as in Bai (1997). The test can 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. $y^\Delta_t$ for some $\Delta \in (0,1)$, although the optimal choice of $\Delta$ is not straightforward. An open issue, moreover, is the impact of the dimensionality, $n$, on the properties of the test. Corollary 2 is a first step in this direction. Finally, the issue of controlling the size is important, since our procedure involves $n$ tests for eigenvalues and for eigenvectors. We address this in Section 5 by proposing a Bonferroni-type correction. Alternative methodologies could be explored (see e.g. Lehmann and Romano, 2005). These issues are currently under investigation by the authors.
References


Appendix A: Preliminary Lemmas

Lemma 1 Let $B$ and $\rho$ 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

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

$$\left[|B\rho|_r^{2r(1+\delta)-\delta^2}/[2+\delta] + |B\rho|_r^r \right]^{1/2(r-1)}.$$

Proof. The proof is fairly similar to Gallant and White (1988; see Davidson, 2002a, p. 271). Let $C = \left[|\rho|_{q^*} |B|_{q^*/q-1} |B\rho|_r^{-1}\right]^{1/2}$, 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

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

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

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

Substituting for $C$ gives (24).

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 (24) is very similar to Lemma 17.15 in Davidson (2002a, p. 271). Particularly, $|\rho|_{q^*}^{-2}$, 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 $u_t^n = E \left[ u_t \mid u_{t-m}, ..., u_{t+m} \right]$, one would look for a bound to $|\phi(u_t) - \phi(u_t^n)|_{2+\delta}$. This would be majorized by some suitably chosen $|B(u_t, u_t^n)\rho(u_t, u_t^n)|_{2+\delta}$: since normally one would choose $\rho(u_t, u_t^n)$ as the taxicab distance, it is $|\rho|_{q^*}^{-2}$ that gives the size of $\phi(u_t)$. 

23
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 \( x_t^0 = w_t \) and \( x_t^0 = E[w_t | w_{t-m}, \ldots, w_{t+m}] \); and define, in a similar fashion, \( y_t^0 = y_t \) and \( y_t^0 = E[y_t | y_{t-m}, \ldots, y_{t+m}] \). Then

\[
\left| x_t^0 - x_t^b \right|_{2+\epsilon} \leq \left| \left( |y_t^0| + |y_t^b| \right) \left( |y_t^0 - y_t^b| \right) \right|_{2+\epsilon} = |B(y_t^0, y_t^b)\rho(y_t^0, y_t^b)|_{2+\epsilon} = |B\rho|_{2+\epsilon}.
\]

Lemma 1 entails that \( |x_t^0 - x_t^b|_{2+\epsilon} \) is bounded by

\[
\left[ |\rho|_{q^*} - 2 |B|_{q^*/q-1} \left( |B\rho|_{[2r(1+\epsilon)-\epsilon^2]/[2+\epsilon]} + |B\rho|_{r} \right) \right]^{1/2(r-1)}.
\]

It holds that \( |\rho|_{q^*} < \infty \) for \( q^* \leq 2r \), and thus for \( q < 2r \); also, \( |B|_{q^*/q-1} < \infty \) if \( q^* \geq \frac{4}{3} \), i.e. \( q > \frac{3}{2} \). Since, for \( q^* \leq 2r \), \( |\rho|_{q^*} \leq |\rho|_{2} \leq Mv_{m} \), where \( M \) is a constant and \( v_m \) a non-negative sequence of size \( v_m = O\left(m^{-\alpha}\right) \), we have \( |x_t^0 - x_t^b|_{2+\epsilon} = |x_t^0 - E[w_t | w_{t-m}, \ldots, w_{t+m}]|_{2+\epsilon} \leq M'\left[v_m^{-2}\right]^{1/2(r-1)} = M'v_m' \). The size of \( v_m' \) is \( v_m' = O\left(m^{-\alpha/2}\right) = O\left(m^{-\alpha'}\right) \). Assumption 1(ii) entails that \( \alpha' > \frac{1}{2} \). 

Lemma 3  Under Assumption 1, \( \text{vec} \left[ \tilde{\omega}_t \tilde{\omega}_t' \right] - E \left( \tilde{\omega}_t \tilde{\omega}_t' \right) \) is \( L_{1+\epsilon/2} \)-NED of size \( \alpha' \) 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), \( \text{vec} \left[ \tilde{\omega}_t \tilde{\omega}_t' - E \left( \tilde{\omega}_t \tilde{\omega}_t' \right) \right] \) is \( L_{2+\epsilon} \)-NED of size \( \alpha' \) on \( \{v_t\}_{t=-\infty}^{+\infty} \).

**Proof.** The proof is very similar to that of Lemma 2. Assuming \( n = 1 \) and letting \( \omega_t = w_t^0 \)

\[
\left| \omega_t^0 - \omega_t^b \right|_{2+\epsilon} \leq \left| \left( x_t^0 + x_t^b \right) \left( x_t^0 - x_t^b \right) \right|_{2+\epsilon} = \left| \left( |y_t^0|^3 + |y_t^b|^3 + |y_t^0|^2 |y_t^b| + |y_t^0| |y_t^b|^2 \right) \left( |y_t^0 - y_t^b| \right) \right|_{2+\epsilon} = |B(y_t^0, y_t^b)\rho(y_t^0, y_t^b)|_{2+\epsilon} = |B\rho|_{2+\epsilon},
\]

so that the Lemma follows from Assumption 2(ii) and Lemma 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 $\tilde{w}_{T,t} = V_{\Sigma,T}^{-1/2} \tilde{w}_t$. In light of Lemma 2, $\tilde{w}_{T,t}$, for given values of $\alpha$ 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 $E (\tilde{w}_{T,t}) = V_{\Sigma,T}^{-1/2} E (\tilde{w}_t) = 0$. 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), $\sup_t E \left( \|\tilde{w}_t\|^{r/2} \right) < \infty$. 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 $w_t$ is a zero-mean $L_2$-mixingale of size $\alpha'' > \frac{1}{2}$. Letting $\mathcal{F}_t = \{w_{1-}, w_m\}$ and $S_{T_m} \equiv \sum_{t=m+1}^{m+T} \tilde{w}_t$, (2) follows if $\sup_1 E \left( \|S_{T_m}\|^{r/2} \right) < \infty$ and $\sup_1 E \left( \|S_{T_m} S_{J_{T_m}} \| \right) = O(T^{1-\theta})$ 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 $l$

$$\hat{\Psi}_{T,T_l} = \left[ T_l \right] \sum_{t=1}^{T_l} vec \left[ \tilde{w}_t \tilde{w}_{t-l} - E \left( \tilde{w}_t \tilde{w}_{t-l} \right) \right] = o_{a.s.} \left( \frac{1}{[T_l]^{\delta/2}} \right);$$

similarly, $\hat{\Sigma}_{T,T_l} = \Sigma = o_{a.s.} \left( \frac{1}{[T_l]^{\delta/2}} \right)$, 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

$$P \left( \sup_{g_T \le [T_T] \le T} [T_T]^{\delta/2} \left\| \hat{\Psi}_{T,T_T} - \Psi_l \right\| > \varepsilon \right) \le \varepsilon',$$

$$P \left( \sup_{1 \le [T_T] \le T - g_T} [T_T]^{\delta/2} \left\| \hat{\Psi}_{T,T_T} - \Psi_l \right\| > \varepsilon \right) \le \varepsilon'. $$

These yield $\sup_{1 \le [T_T] \le T} \left\| \hat{\Psi}_{T,T_T} - \Psi_l \right\| = o_p \left( \frac{1}{T_T} \right)$. This proves (4).
In order to prove (5), write

\[
\sup_{1 \leq |T\tau| \leq T} \left\| \hat{V}_{\Sigma,\tau} - V_\Sigma \right\| \leq \sup_{1 \leq |T\tau| \leq T} \left\| \hat{V}_{\Sigma,\tau} - V_\Sigma \right\| \\
+ 2 \sup_{1 \leq |T\tau| \leq T} \sum_{l=0}^{m} \left( 1 - \frac{l}{m} \right) \left\| \hat{\Psi}_{l,[T\tau]} - \Psi_l + \hat{\Psi}_{l,1-[T\tau]} - \Psi_l \right\| \\
+ 2 \sum_{l=1}^{m} \frac{l}{m} \|\Psi_l\| + 2 \sum_{l=m+1}^{\infty} \|\Psi_l\| \\
= I + II + III + IV. \tag{25}
\]

Assumption 2(i)/(b) entails \( IV = o(m^{-8}) \). Equation (4) ensures that \( I = o_p\left(T^{-\delta}\right) \); as far as \( II \) is concerned, this is of the same order as

\[
\max_{1 \leq l \leq m} E \left[ \sup_{1 \leq |T\tau| \leq T} \left\| \hat{\Psi}_{l,[T\tau]} + \hat{\Psi}_{l,1-[T\tau]} - \Psi_l \right\| \right] \sum_{l=1}^{m} \left( 1 - \frac{l}{m} \right) = O \left( \frac{1}{T^\delta} \right) O(m); \tag{26}
\]

finally, in light of Assumption 2(ii)/(b), \( III = 2m^{-1}O(1) = O(m^{-1}) \). Thus, (5) follows.

Consider (6). We still use (25) 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 \( \sup_{1 \leq |T\tau| \leq T} \left\| \hat{\Psi}_{l,[T\tau]} - \Psi_l \right\| = O_p\left(T^{-1/2}\right) \). Thus (26) becomes

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

Putting all together, (5) follows. \( \blacksquare \)

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

\[
\left[ \Sigma + (\hat{\Sigma}_\tau - \Sigma) \right] [x_i + (\hat{x}_{i,\tau} - x_i)] = \lambda_i + (\hat{\lambda}_{i,\tau} - \lambda_i) [x_i + (\hat{x}_{i,\tau} - x_i)]. \tag{27}
\]

After expanding the product, consider the terms \( (\hat{\Sigma}_\tau - \Sigma) (\hat{x}_{i,\tau} - x_i) \) and \( (\hat{\lambda}_{i,\tau} - \lambda_i) (\hat{x}_{i,\tau} - x_i) \). It holds that \( \hat{\lambda}_{i,\tau} - \lambda_i = O_p\left(T^{-1/2}\right) \) uniformly in \( \tau \). This is because \( \Sigma \) is symmetric, and therefore Corollary 6.3.4 in Horn and Johnson (1995, p. 367) entails that \( |\hat{\lambda}_{i,\tau} - \lambda_i| \leq \left\| \hat{\Sigma}_\tau - \Sigma \right\| \). Equation (1) yields the result. Also, it holds that \( \hat{x}_{i,\tau} - x_i = O_p\left(T^{-1/2}\right) \) uniformly in \( \tau \). This follows from the sin \( \Theta \) Theorem in Davis and Kahan (1970, p. 10), whereby \( \left\| \hat{X} - X \right\| \leq \left\| X \right\| \left\| \hat{\Sigma}_\tau - \Sigma \right\| \). Thus, \( (\hat{\Sigma}_\tau - \Sigma) (\hat{x}_{i,\tau} - x_i) \) and \( (\hat{\lambda}_{i,\tau} - \lambda_i) (\hat{x}_{i,\tau} - x_i) \) are \( O_p\left(T^{-1}\right) \) uniformly in \( \tau \); omitting them, (27) can be written as

\[
\Sigma (\hat{x}_{i,\tau} - x_i) + (\hat{\Sigma}_\tau - \Sigma) x_i = \lambda_i (\hat{x}_{i,\tau} - x_i) + (\hat{\lambda}_{i,\tau} - \lambda_i) x_i. \tag{28}
\]
Consider (7). Premultiplying (28) by \( x'_i \), we obtain 
\[
 x'_i \Sigma (\hat{x}_{i,\tau} - x_i) + x'_i \left( \Sigma_{\tau} - \Sigma \right) x_i = \lambda_i x'_i (\hat{x}_{i,\tau} - x_i)
\]
+ \( (\hat{\lambda}_{i,\tau} - \lambda_i) x'_i x_i \). Recalling that \( x'_i \Sigma = \lambda_i x'_i \), and that \( x'_i x_i = 1 \), we have \( x'_i \left( \Sigma_{\tau} - \Sigma \right) x_i = \lambda_i x'_i (\hat{x}_{i,\tau} - x_i) \), which entails (7). In order to prove (8), note that the \( x_i \)'s 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 (28) by any \( x'_k \) for \( i \neq k \), and using the identity 
\[
 x'_i \Sigma = \lambda_i x'_i \), we obtain \( \lambda_k \phi_{k,\tau} + x'_k \left( \Sigma_{\tau} - \Sigma \right) x_i = \lambda_i \phi_{k,\tau} \). This yields \( \phi_{k,\tau} = \frac{x'_i \left( \Sigma_{\tau} - \Sigma \right) x_i}{\lambda_i - \lambda_k} \), 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'_i \left( \Sigma_{\tau} - \Sigma \right) x_i}{\lambda_i - \lambda_k} x_k + \phi_{i,\tau} x_i; (8) follows from setting \( \phi_{i,\tau} = 0 \).

We now turn to deriving the bias for \( \hat{x}_{i,\tau} - \lambda_i \), presented in (9). Expanding (27) and premultiplying by \( x'_i \) we obtain

\[
 \hat{\lambda}_{i,\tau} - \lambda_i = x'_i \left( \Sigma_{\tau} - \Sigma \right) x_i - (\hat{\lambda}_{i,\tau} - \lambda_i) x'_i (\hat{x}_{i,\tau} - x_i) + x'_i \left( \Sigma_{\tau} - \Sigma \right) (\hat{x}_{i,\tau} - x_i)
\]

\[
 = x'_i \left( \Sigma_{\tau} - \Sigma \right) x_i - I + II. 
\]

From (8), \( I = x'_i \sum_{k \neq i} \frac{x'_i \left( \Sigma_{\tau} - \Sigma \right) x_i}{\lambda_i - \lambda_k} x_k = 0 \). Also (focusing on first order terms only):

\[
 II = \left[ x'_i \otimes (\hat{x}_{i,\tau} - x_i) \right] vec \left( \Sigma_{\tau} - \Sigma \right)
\]

\[
 = \left[ x'_i \otimes \sum_{k \neq i} \frac{x'_i}{\lambda_i - \lambda_k} x'_k \left( \Sigma_{\tau} - \Sigma \right) x_i \right] vec \left( \Sigma_{\tau} - \Sigma \right)
\]

\[
 = \sum_{k \neq i} \left[ x'_i \otimes \frac{x'_k}{\lambda_i - \lambda_k} \right] \left[ vec \left( \Sigma_{\tau} - \Sigma \right) \right] \left[ vec \left( \Sigma_{\tau} - \Sigma \right) \right] ^{\prime} [x_k \otimes x_i].
\]

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

\[
 E \left[ T x'_i \left( \Sigma_{\tau} - \Sigma \right) (\hat{x}_{i,\tau} - x_i) \right] = \sum_{k \neq i} \left[ x'_i \otimes x'_k \right] = \sum_{k \neq i} \frac{x'_i \otimes x'_k}{\lambda_i - \lambda_k} E \left[ vec \left( \Sigma_{\tau} - \Sigma \right) \right] \left[ vec \left( \Sigma_{\tau} - \Sigma \right) \right] ^{\prime} [x_k \otimes x_i]
\]

\[
 = \sum_{k \neq i} \frac{x'_i \otimes x'_k}{\lambda_i - \lambda_k} V_{\Sigma} (x_k \otimes x_i).
\]

The bias for \( \hat{x}_{i,\tau} - x_i \) can be derived from (8) following similar passages. Using (27), \( \lambda_k \phi_{k,\tau} + \)
\[ x_k' \left( \hat{\Sigma} - \Sigma \right) x_i + x_k' \left( \hat{\Sigma} - \Sigma \right) (\hat{x}_{i,\tau} - x_i) = \lambda_i \phi_{k,\tau} + \left( \lambda_{i,\tau} - \lambda_i \right) \phi_{k,\tau}, \]

whence
\[ \phi_{k,\tau} = \frac{x_k' \left( \hat{\Sigma} - \Sigma \right) x_i + x_k' \left( \hat{\Sigma} - \Sigma \right) (\hat{x}_{i,\tau} - x_i)}{\lambda_i - \lambda_k + \left( \lambda_{i,\tau} - \lambda_i \right)} + o_p(1). \]

Thus, since \( \hat{x}_{i,\tau} - x_i = \sum_{k \neq i} \phi_{k,\tau} x_k \), it holds that
\[ \hat{x}_{i,\tau} - x_i = \sum_{k \neq i} x_k' \left( \hat{\Sigma} - \Sigma \right) x_i x_k + \sum_{k \neq i} x_k' \left( \hat{\Sigma} - \Sigma \right) (\hat{x}_{i,\tau} - x_i) x_k + o_p(1). \]

The bias is given by the second term, with
\[
E \left[ T \sum_{k \neq i} x_k' \left( \hat{\Sigma} - \Sigma \right) (\hat{x}_{i,\tau} - x_i) \right] \\
= \sum_{k \neq i} \sum_{j \neq i} (x_k' \otimes x_j') E \left[ T \left[ \text{vec} \left( \hat{\Sigma} - \Sigma \right) \right] \left[ \text{vec} \left( \hat{\Sigma} - \Sigma \right) \right]' \right] (x_j \otimes x_i) \\
= \sum_{k \neq i} \sum_{j \neq i} x_k' \otimes x_j' \text{vec} (x_j \otimes x_i) (\lambda_i - \lambda_k) (\lambda_i - \lambda_j) x_k.
\]

**Proof of Theorem 3.** The proof of (13) follows from (1), Theorem 2 and the CMT. As far as (14) 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, \( \hat{\Lambda}(\tau) = \frac{1}{\sqrt{T(1-\tau)}} \left[ \hat{S}(\tau)' \hat{V}_{\Sigma_{\tau}}^{-1} S(\tau) \right]^{1/2} \), where \( S(\tau) = S(\tau) - \frac{T}{T} S(T) \). We also define \( \bar{\Lambda}(\tau) = \frac{1}{\sqrt{T(1-\tau)}} \left[ \bar{S}(\tau)' \bar{V}_{\Sigma_{\tau}}^{-1} S(\tau) \right]^{1/2} \); further, letting \( B_{1,i}(\tau) \) be a sequence of standard, independent Brownian bridges for \( i = 1, \ldots, n^2 \), we define \( M(\tau) = \left[ \sum_{i=1}^{n^2} B_{1,i}(\tau) \right]^{1/2} \). The Darling-Erdos Theorem (see e.g. Corollary A.3.1 in Csorgo and Horvath, 1997, p. 366) states that
\[ P \left[ a_T \sup_{\frac{1}{T} \leq \tau \leq 1 - \frac{1}{T}} M(\tau) \leq x + b_T \right] = e^{-2e^{-x}}, \]
where the norming constants \( a_T \) and \( b_T \) are defined in the Theorem. In order to prove (14), it is enough to show that
\[ \sup_{\frac{1}{T} \leq \tau \leq 1 - \frac{1}{T}} \bar{\Lambda}(\tau) - \sup_{\frac{1}{T} \leq \tau \leq 1 - \frac{1}{T}} M(\tau) = o_p \left[ (\ln \ln T)^{-1/2} \right]. \]

By virtue of Theorem 2, it suffices to prove that
\[
\sup_{\frac{1}{T} \leq \tau \leq 1 - \frac{1}{T}} |\bar{\Lambda}(\tau) - M(\tau)| = o_p \left( \frac{1}{\sqrt{\ln \ln T}} \right). \tag{30}
\]

In order to show (30), note first that (2) yields the (weak) result
\[
\sup_{\frac{1}{T} \leq \tau \leq 1 - \frac{1}{T}} |\bar{\Lambda}(\tau) - M(\tau)| = o_p \left( \sqrt{\ln \ln T} \right). \tag{31}
\]
Indeed, (2) entails

\[
\sup_{u(T, \varepsilon) \leq \tau \leq \frac{1}{2}} \left[ \left\lceil T \tau \right\rceil \delta \right] \left| \tilde{A}(\tau) - M(\tau) \right| = o_p(1), \tag{32}
\]

\[
\sup_{\frac{1}{2} \leq \tau \leq 1 - u(T, \varepsilon)} \left[ \left\lceil T(1 - \tau) \right\rceil \delta \right] \left| \tilde{A}(\tau) - M(\tau) \right| = o_p(1), \tag{33}
\]

for all sequences \( u(T, \varepsilon) \) such that \( u(T, \varepsilon) \to 0 \) and \( Tu(T, \varepsilon) \to \infty \) as \( T \to \infty \); here, \( \varepsilon \) 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

\[
\frac{1}{\sqrt{2 \ln \ln T}} \sup_{\frac{1}{2} \leq \tau \leq u(T, \varepsilon)} M(\tau) \overset{P}{\to} \sqrt{\varepsilon}, \tag{34}
\]

\[
\frac{1}{\sqrt{2 \ln \ln T}} \sup_{1 - u(T, \varepsilon) \leq \tau \leq 1 - \frac{1}{2}} M(\tau) \overset{P}{\to} \sqrt{\varepsilon}.
\]

Hence, from (31)

\[
\frac{1}{\sqrt{2 \ln \ln T}} \sup_{\frac{1}{2} \leq \tau \leq u(T, \varepsilon)} \tilde{A}(\tau) \overset{P}{\to} \sqrt{\varepsilon},
\]

\[
\frac{1}{\sqrt{2 \ln \ln T}} \sup_{1 - u(T, \varepsilon) \leq \tau \leq 1 - \frac{1}{2}} \tilde{A}(\tau) \overset{P}{\to} \sqrt{\varepsilon}.
\]

Defining \( \xi(T) \) and \( \eta(T) \) as \( \sup_{\frac{1}{2} \leq u(T, \varepsilon) \leq T} M(\tau) = M[\xi(T)] \) and \( \sup_{\frac{1}{2} \leq \tau \leq T} \tilde{A}(\tau) = \tilde{A}[\eta(T)] \), the relationships above entail \( P \left[ u(T, \varepsilon) \leq \xi(T), \eta(T) \leq 1 - u(T, \varepsilon) \right] = 1 \) as \( T \to \infty \). Using (34) as an illustrative example, this follows from the fact that \( \sup_{\frac{1}{2} \leq \tau \leq T} M(\tau) \) is essentially equal to \(-\infty\) since, as \( T \to \infty \) and \( \varepsilon \to 0 \)

\[
P \left[ a_T \sup_{\frac{1}{2} \leq \tau \leq u(T, \varepsilon)} \tilde{A}(\tau) - b_T \geq -K \right] = P \left[ (\sqrt{\varepsilon} - 1) \ln \ln T \geq -K \right] = 0,
\]

for some \( K > 0 \). Hence, (32) and (33) entail

\[
\sup_{\frac{1}{2} \leq \tau \leq 1 - \frac{1}{2}} \left| \tilde{A}(\tau) - M(\tau) \right| = o_p \left( e^{-\delta \ln T} \right),
\]

and since \( \sup_{\frac{1}{2} \leq \tau \leq 1 - \frac{1}{2}} \tilde{A}(\tau) - \sup_{\frac{1}{2} \leq \tau \leq 1 - \frac{1}{2}} M(\tau) \leq \sup_{\frac{1}{2} \leq \tau \leq 1 - \frac{1}{2}} \left| \tilde{A}(\tau) - M(\tau) \right| \), (30) follows in view of \( \sqrt{\ln \ln T} e^{-\delta \ln T} \to 0 \).

**Proof of Corollary 2.** Consider \( T^{-1} \left[ S(\tau) - \frac{|T\tau|}{T} S(T) \right] \) \( \tilde{V}_T^{-1} \left[ S(\tau) - \frac{|T\tau|}{T} S(T) \right] \) It
Theorem 2 states that
\[ a \]
The asymptotics of \( \sup \) holds whereby the 8-th moment condition on \( V_\Sigma^{-1/2} \left( \tilde{V}_{\Sigma,\tau} - V_{\Sigma} \right) V_{\Sigma}^{-1} + o \left( \left\| \tilde{V}_{\Sigma,\tau} - V_{\Sigma} \right\| \right) ; \) thence
\[
\begin{align*}
&\frac{1}{T} \left[ S(\tau) - \left\lfloor \frac{T\tau}{T} \right\rfloor S(T) \right] \tilde{V}_{\Sigma,\tau}^{-1} \left[ S(\tau) - \left\lfloor \frac{T\tau}{T} \right\rfloor S(T) \right] \\
= &\frac{1}{T} \left[ S(\tau) - \left\lfloor \frac{T\tau}{T} \right\rfloor S(T) \right] V_{\Sigma}^{-1} \left[ S(\tau) - \left\lfloor \frac{T\tau}{T} \right\rfloor S(T) \right] + \\
&\frac{1}{T} \left[ S(\tau) - \left\lfloor \frac{T\tau}{T} \right\rfloor S(T) \right] \tilde{V}_{\Sigma,\tau}^{-1} \left( \tilde{V}_{\Sigma,\tau} - V_{\Sigma} \right) V_{\Sigma}^{-1} \left[ S(\tau) - \left\lfloor \frac{T\tau}{T} \right\rfloor S(T) \right]
\end{align*}
\]

Write \( I_\tau \), with zero mean and unit variance, and approximation error of order \( O_{a.s.} (T^{-\delta}) \). Thus
\[
\frac{1}{\sqrt{2n}} \sum_{i=1}^{n^2} (w_{i,\tau}^2 - 1) = \frac{1}{\sqrt{2n}} \sum_{i=1}^{n^2} (w_{i,\tau}^2 - 1) + \frac{1}{\sqrt{2n}} \sum_{i=1}^{n^2} (w_{i,\tau}^2 - w_{i,\tau}^2) = I_{a,\tau} + I_{b,\tau}.
\]
The asymptotics of \( \sup_{[T\tau]} I_{a,\tau} \) is studied in Remark 2.1 in Aue et al. (2009), which stipulates that \( \sup_{[T\tau]} I_{a,\tau} = O_p (1) \). From the SIP, \( \sup_{[T\tau]} I_{b,\tau} = nO_p \left( T^{-\delta'} \right) \). We now turn to \( II_\tau \), where \( \sum_{i=1}^{n^2} \sum_{j=1}^{n^2} w_{i,\tau}w_{j,\tau} a_{i,j,\tau} \); \( a_{i,j,\tau} \) is element in position \( (i,j) \) of matrix \( V^{-1} \left( \tilde{V}_{\Sigma,\tau} - V_{\Sigma} \right) V_{\Sigma}^{-1} \). Theorem 2 states that \( a_{i,j,\tau} = O_p \left( \frac{m}{\sqrt{T}} \right) + O_p \left( \frac{1}{\sqrt{n}} \right) \), uniformly in \( \tau \). For some constant \( M, \)
\[
\sup_{[T\tau]} II_\tau \leq \sup_{[T\tau]} \sum_{i=1}^{n^2} \sum_{j=1}^{n^2} w_{i,\tau}w_{j,\tau} a_{i,j,\tau} \leq \max_{1 \leq i,j \leq n} \sup_{[T\tau]} |a_{i,j,\tau}| \sup_{[T\tau]} \sum_{i=1}^{n^2} \sum_{j=1}^{n^2} |w_{i,\tau}w_{j,\tau}|
\]
\[
\leq \max_{1 \leq i,j \leq n} \sup_{[T\tau]} |a_{i,j,\tau}| \sup_{[T\tau]} \left( \sum_{i=1}^{n^2} |w_{i,\tau}| \right)^2 \leq M \max_{1 \leq i,j \leq n} \sup_{[T\tau]} |a_{i,j,\tau}| \sup_{[T\tau]} \sum_{i=1}^{n^2} w_{i,\tau}^2
\]
\[
= \left[ O_p \left( \frac{m}{\sqrt{T}} \right) + O_p \left( \frac{1}{\sqrt{n}} \right) \right] O_p (n^2). \]

Thus, \( (\sqrt{2n})^{-1} \sup_{[T\tau]} II_\tau = O_p \left( \frac{m}{\sqrt{T}} \right) + O_p \left( \frac{1}{\sqrt{n}} \right) \). \( \blacksquare \)

**Proof of Theorem 4.** In order to prove (17), we show that, under \( H_a^{(T)} \)
\[
P \left[ \sup_{n \leq \lfloor T\tau \rfloor \leq T-n} \Lambda_{T} (\tau) > c_{a,T} \right] = P [\Lambda_0 > c_{a,T} - NC_T],
\]
where \( \Lambda_0 \) 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 \( c_{a,T} - NC_T \rightarrow -\infty \) as \( T \rightarrow \infty \).

Note that
\[
\Lambda_{T} (\tau) = \frac{T}{\lfloor T\tau \rfloor} \tilde{S} (\tau) ^{1/2} \tilde{S}_{\Sigma}^{-1} \tilde{S} (\tau),
\]
where we consider Assumption 2(i) and the full sample estimator of \( V_{\Sigma}^{-1} \) only, for simplicity.
Consider $\tilde{S}(\tau)$. Under $H_a^{(T)}$

$$
\sqrt{\frac{T}{[T \tau] [T (1 - \tau)]}} \tilde{S}(\tau) = \sqrt{\frac{T}{[T \tau] [T (1 - \tau)]}} \left[ \sum_{t=1}^{T} \hat{w}_t - \frac{|T \tau|}{T} \sum_{t=1}^{T} \hat{w}_t \right] + \Delta_T \sqrt{\frac{T}{[T \tau] [T (1 - \tau)]}} \left[ \sum_{t=1}^{T} I(t \leq k_{0,T}) - \frac{|T \tau|}{T} \sum_{t=1}^{T} I(t \leq k_{0,T}) \right]
$$

$$
= \tilde{S}_1(\tau) + \tilde{S}_2(\tau),
$$

where $I(\cdot)$ is the indicator function. The sequence $\hat{w}_t$ 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

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

with

$$
\sup_{1 \leq |T \tau| \leq T} \Delta_T \sqrt{\frac{T}{[T \tau] [T (1 - \tau)]}} \left[ \left( \frac{|T (1 - \tau)|}{T} \right) k_{0,T} I(k_{0,T} < |T \tau|) + \left( \frac{T - k_{0,T}}{T} \right) [T \tau] I(k_{0,T} \geq |T \tau|) \right]
$$

$$
= \Delta_T \sqrt{k_{0,T} \left( \frac{T - k_{0,T}}{T} \right)}. \tag{35}
$$

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

$$
\hat{V}_{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} \hat{w}_t \hat{w}_t' - \frac{1}{T} \sum_{t=1}^{T} \hat{w}_t \left[ \frac{T - k_{0,T}}{T} I(t \geq k_{0,T}) \right] \Delta_T' - \frac{1}{T} \sum_{t=1}^{T} \left[ \frac{T - k_{0,T}}{T} I(t \geq k_{0,T}) \right] \Delta_T \hat{w}_t' + \frac{1}{T} \sum_{t=1}^{T} \left[ \frac{T - k_{0,T}}{T} I(t \geq k_{0,T}) \right]^2 \Delta_T \Delta_T'.
$$

The LLN entails that $I \overset{D}{\rightarrow} V_{\Sigma}$; $II$ and $III$ have the same order of magnitude as each other.
Particularly, since \( \sum_{t=1}^{T} \bar{w}_t \left[ \frac{T-k_0,T}{T} - I(t \geq k_0,T) \right] = O_p \left( \sqrt{T} \right) , II = O_p \left( \frac{\| \Delta_T \|^2}{T} \right) . \) Finally

\[
\frac{1}{T} \sum_{t=1}^{T} \left[ \frac{T-k_0,T}{T} - I(t \geq k_0,T) \right]^2
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \left( \frac{T-k_0,T}{T} \right)^2 - 2 \left( \frac{T-k_0,T}{T} \right)^2 + \frac{1}{T} \sum_{t=1}^{T} I(t \geq k_0,T)
\]

\[
= k_0,T \frac{T - k_0,T}{T},
\]

thus, \( IV = O_p \left( \frac{k_0,T T-k_0,T}{T} \frac{\| \Delta_T \|^2}{T} \right) , \) which is \( O_p \left( 1 \right) \) under \( H_a^{(T)} . \) This entails that \( \| \hat{V}_\Sigma - V_\Sigma \| = O_p \left( 1 \right) \) under \( H_a^{(T)} . \) Applying Taylor’s expansion, we can write \( \hat{V}_\Sigma^{-1} = V_\Sigma^{-1} + \hat{V}_\Sigma^{-1} (\hat{V}_\Sigma - V_\Sigma) \hat{V}_\Sigma^{-1} , \) for some invertible matrix \( \hat{V}_\Sigma . \)

Putting all together, we have

\[
\Lambda_T(\tau) = \hat{S}_1(\tau)^T V_\Sigma^{-1} \hat{S}_1(\tau) + \hat{S}_2(\tau)^T \hat{V}_\Sigma^{-1} \hat{S}_2(\tau)
\]

\[
+ 2 \hat{S}_1(\tau)^T \hat{V}_\Sigma^{-1} \hat{S}_2(\tau) + \hat{S}_1(\tau)^T \hat{V}_\Sigma^{-1} (\hat{V}_\Sigma - V_\Sigma) \hat{V}_\Sigma^{-1} \hat{S}_1(\tau)
\]

\[
= I + II + III + IV.
\]

Term \( I \) follows the null distribution under \( H_n^{(T)} , \) i.e. \( \sup_{n \leq |T| \leq T-n} \hat{S}_1(\tau)^T V_\Sigma^{-1} \hat{S}_1(\tau) \overset{d}{\rightarrow} \Lambda_0 \) as \( T \rightarrow \infty . \) Given that \( \hat{V}_\Sigma \) is \( O_p \left( 1 \right) \) under \( H_a^{(T)} , \) term \( II \) has the same order as \( \sup_{n \leq |T| \leq T-n} \| \hat{S}_2(\tau) \|^2 \), which is \( O \left( k_0,T T-k_0,T \frac{\| \Delta_T \|^2}{T} \right) \) in view of (35). Terms \( III \) and \( IV \) are of smaller order of magnitude than \( II : \) e.g. as far as \( III \) is concerned, it holds that \( E \left( \hat{S}_1(\tau)^T V_\Sigma^{-1} \hat{S}_2(\tau) \right) \leq \left( E \left( \| \hat{S}_1(\tau) \|^2 \right)^{1/2} \left( E \left( \| \hat{S}_2(\tau) \|^2 \right)^{1/2} \right) , \) since \( \hat{V}_\Sigma^{-1} \) is \( O_p \left( 1 \right) ; \) thus, \( \sup_{n \leq |T| \leq T-n} \hat{S}_1(\tau)^T V_\Sigma^{-1} \hat{S}_2(\tau) \)

\[
= O \left( \sqrt{\ln \ln T} \sqrt{k_0,T T-k_0,T \frac{\| \Delta_T \|^2}{T}} \right) , \] which is smaller than \( II , \) as \( T \rightarrow \infty , \) due to (16). Therefore, under \( H_a^{(T)} , \)

\[
P \left[ \sup_{n \leq |T| \leq T-n} \Lambda_T(\tau) > c_{a,T} \right] = P \left[ \Lambda_0 > c_{a,T} - NC_T \right] , \]

\[
NC_T = \| \Delta_T \| \sqrt{k_0,T \left( \frac{T - k_0,T}{T} \right)} + o \left( \| \Delta_T \| \sqrt{k_0,T \left( \frac{T - k_0,T}{T} \right)} \right).
\]

In view of \( c_{a,T} \) being \( O \sqrt{\ln \ln T} \) and of (16), it holds that \( c_{a,T} - NC_T \rightarrow -\infty \) as \( T \rightarrow \infty , \) whence (17) follows.
Table 1. Empirical rejection frequencies for the null of no changes in the largest eigenvalue of $\Sigma$. Data are generated according to (20). The empirical sizes reported here have confidence interval $[0.04, 0.06]$.
<table>
<thead>
<tr>
<th>(n)</th>
<th>(T)</th>
<th>((p, \theta))</th>
<th>(0.0, 0.5)</th>
<th>(0.0, 0.5)</th>
<th>(0.0, -0.5)</th>
<th>(0.5, 0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.035</td>
<td>0.021</td>
<td>0.027</td>
<td>0.034</td>
<td>0.016</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.014</td>
<td>0.001</td>
<td>0.010</td>
<td>0.006</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Test based on (21)</td>
<td>0.133</td>
<td>0.121</td>
<td>0.128</td>
<td>0.122</td>
<td>0.112</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.128</td>
<td>0.073</td>
<td>0.084</td>
<td>0.101</td>
<td>0.052</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.235</td>
<td>0.180</td>
<td>0.211</td>
<td>0.191</td>
<td>0.167</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.293</td>
<td>0.185</td>
<td>0.228</td>
<td>0.232</td>
<td>0.140</td>
<td></td>
</tr>
<tr>
<td>Test based on (21)</td>
<td>0.427</td>
<td>0.302</td>
<td>0.371</td>
<td>0.335</td>
<td>0.259</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.533</td>
<td>0.350</td>
<td>0.424</td>
<td>0.410</td>
<td>0.272</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.039</td>
<td>0.026</td>
<td>0.028</td>
<td>0.026</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Test based on (21)</td>
<td>0.016</td>
<td>0.016</td>
<td>0.020</td>
<td>0.015</td>
<td>0.012</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.171</td>
<td>0.117</td>
<td>0.138</td>
<td>0.138</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.153</td>
<td>0.099</td>
<td>0.121</td>
<td>0.118</td>
<td>0.075</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.318</td>
<td>0.245</td>
<td>0.246</td>
<td>0.241</td>
<td>0.197</td>
<td></td>
</tr>
<tr>
<td>Test based on (21)</td>
<td>0.374</td>
<td>0.244</td>
<td>0.291</td>
<td>0.265</td>
<td>0.186</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.460</td>
<td>0.337</td>
<td>0.377</td>
<td>0.370</td>
<td>0.281</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.584</td>
<td>0.395</td>
<td>0.469</td>
<td>0.449</td>
<td>0.309</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Test based on (21)</td>
<td>0.040</td>
<td>0.034</td>
<td>0.037</td>
<td>0.031</td>
<td>0.033</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.018</td>
<td>0.012</td>
<td>0.016</td>
<td>0.012</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Test based on (21)</td>
<td>0.210</td>
<td>0.154</td>
<td>0.149</td>
<td>0.168</td>
<td>0.132</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.184</td>
<td>0.112</td>
<td>0.145</td>
<td>0.126</td>
<td>0.084</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.344</td>
<td>0.247</td>
<td>0.257</td>
<td>0.290</td>
<td>0.205</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.409</td>
<td>0.247</td>
<td>0.305</td>
<td>0.313</td>
<td>0.193</td>
<td></td>
</tr>
<tr>
<td>Test based on (21)</td>
<td>0.516</td>
<td>0.339</td>
<td>0.387</td>
<td>0.407</td>
<td>0.288</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.627</td>
<td>0.397</td>
<td>0.482</td>
<td>0.506</td>
<td>0.312</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.032</td>
<td>0.026</td>
<td>0.033</td>
<td>0.021</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Test based on (21)</td>
<td>0.020</td>
<td>0.012</td>
<td>0.021</td>
<td>0.015</td>
<td>0.008</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.217</td>
<td>0.157</td>
<td>0.183</td>
<td>0.175</td>
<td>0.140</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.203</td>
<td>0.116</td>
<td>0.150</td>
<td>0.151</td>
<td>0.080</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.356</td>
<td>0.247</td>
<td>0.273</td>
<td>0.309</td>
<td>0.209</td>
<td></td>
</tr>
<tr>
<td>Test based on (21)</td>
<td>0.440</td>
<td>0.248</td>
<td>0.328</td>
<td>0.336</td>
<td>0.192</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.528</td>
<td>0.364</td>
<td>0.449</td>
<td>0.434</td>
<td>0.298</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.661</td>
<td>0.430</td>
<td>0.536</td>
<td>0.537</td>
<td>0.347</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Test based on (21)</td>
<td>0.006</td>
<td>0.001</td>
<td>0.004</td>
<td>0.008</td>
<td>0.003</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.029</td>
<td>0.008</td>
<td>0.016</td>
<td>0.020</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.211</td>
<td>0.136</td>
<td>0.165</td>
<td>0.178</td>
<td>0.121</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.241</td>
<td>0.145</td>
<td>0.181</td>
<td>0.181</td>
<td>0.118</td>
<td></td>
</tr>
<tr>
<td>Test based on (21)</td>
<td>0.440</td>
<td>0.315</td>
<td>0.365</td>
<td>0.341</td>
<td>0.281</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.514</td>
<td>0.315</td>
<td>0.411</td>
<td>0.426</td>
<td>0.265</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.623</td>
<td>0.390</td>
<td>0.483</td>
<td>0.493</td>
<td>0.350</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.727</td>
<td>0.471</td>
<td>0.570</td>
<td>0.574</td>
<td>0.386</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Test based on (21)</td>
<td>0.035</td>
<td>0.014</td>
<td>0.021</td>
<td>0.015</td>
<td>0.014</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.211</td>
<td>0.126</td>
<td>0.164</td>
<td>0.149</td>
<td>0.099</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.280</td>
<td>0.152</td>
<td>0.204</td>
<td>0.208</td>
<td>0.117</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.486</td>
<td>0.296</td>
<td>0.377</td>
<td>0.383</td>
<td>0.278</td>
<td></td>
</tr>
<tr>
<td>Test based on (21)</td>
<td>0.585</td>
<td>0.311</td>
<td>0.447</td>
<td>0.466</td>
<td>0.270</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.644</td>
<td>0.401</td>
<td>0.506</td>
<td>0.512</td>
<td>0.403</td>
</tr>
<tr>
<td>Test based on (19)</td>
<td>0.764</td>
<td>0.493</td>
<td>0.594</td>
<td>0.603</td>
<td>0.426</td>
<td></td>
</tr>
</tbody>
</table>

Table 2a. Power of the test for the null of no changes in the largest eigenvalue of \( \Sigma \). Data are generated according to (20) and under the alternative hypothesis specified in (23).
<table>
<thead>
<tr>
<th>n</th>
<th>$T$</th>
<th>($p, q$)</th>
<th>(0, 0)</th>
<th>(0.5, 0)</th>
<th>(0, 0.5)</th>
<th>(0, -0.5)</th>
<th>(0.5, 0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.055</td>
<td>0.030</td>
<td>0.045</td>
<td>0.050</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.020</td>
<td>0.004</td>
<td>0.021</td>
<td>0.013</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Test based on (21)</td>
<td>0.272</td>
<td>0.188</td>
<td>0.213</td>
<td>0.213</td>
<td>0.164</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.258</td>
<td>0.136</td>
<td>0.182</td>
<td>0.198</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.514</td>
<td>0.380</td>
<td>0.450</td>
<td>0.415</td>
<td>0.298</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.593</td>
<td>0.413</td>
<td>0.510</td>
<td>0.486</td>
<td>0.306</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.874</td>
<td>0.652</td>
<td>0.753</td>
<td>0.760</td>
<td>0.524</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.934</td>
<td>0.762</td>
<td>0.850</td>
<td>0.836</td>
<td>0.625</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.070</td>
<td>0.041</td>
<td>0.056</td>
<td>0.045</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.028</td>
<td>0.022</td>
<td>0.029</td>
<td>0.020</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Test based on (21)</td>
<td>0.314</td>
<td>0.206</td>
<td>0.256</td>
<td>0.259</td>
<td>0.181</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.314</td>
<td>0.168</td>
<td>0.255</td>
<td>0.220</td>
<td>0.119</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.616</td>
<td>0.420</td>
<td>0.505</td>
<td>0.508</td>
<td>0.356</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.717</td>
<td>0.447</td>
<td>0.584</td>
<td>0.579</td>
<td>0.349</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.900</td>
<td>0.704</td>
<td>0.810</td>
<td>0.785</td>
<td>0.568</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.955</td>
<td>0.795</td>
<td>0.892</td>
<td>0.879</td>
<td>0.658</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.073</td>
<td>0.043</td>
<td>0.062</td>
<td>0.063</td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.038</td>
<td>0.019</td>
<td>0.039</td>
<td>0.028</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Test based on (21)</td>
<td>0.362</td>
<td>0.253</td>
<td>0.286</td>
<td>0.295</td>
<td>0.219</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.385</td>
<td>0.200</td>
<td>0.292</td>
<td>0.274</td>
<td>0.161</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.679</td>
<td>0.467</td>
<td>0.526</td>
<td>0.585</td>
<td>0.401</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.764</td>
<td>0.513</td>
<td>0.615</td>
<td>0.631</td>
<td>0.406</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.916</td>
<td>0.702</td>
<td>0.815</td>
<td>0.846</td>
<td>0.592</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.949</td>
<td>0.795</td>
<td>0.871</td>
<td>0.915</td>
<td>0.687</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.056</td>
<td>0.040</td>
<td>0.053</td>
<td>0.039</td>
<td>0.043</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.037</td>
<td>0.017</td>
<td>0.030</td>
<td>0.027</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Test based on (21)</td>
<td>0.406</td>
<td>0.252</td>
<td>0.308</td>
<td>0.299</td>
<td>0.222</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.405</td>
<td>0.215</td>
<td>0.300</td>
<td>0.308</td>
<td>0.161</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.711</td>
<td>0.465</td>
<td>0.566</td>
<td>0.598</td>
<td>0.403</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.796</td>
<td>0.524</td>
<td>0.648</td>
<td>0.674</td>
<td>0.419</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.954</td>
<td>0.768</td>
<td>0.889</td>
<td>0.866</td>
<td>0.665</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.979</td>
<td>0.856</td>
<td>0.943</td>
<td>0.920</td>
<td>0.748</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.016</td>
<td>0.008</td>
<td>0.010</td>
<td>0.014</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.053</td>
<td>0.018</td>
<td>0.027</td>
<td>0.037</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Test based on (21)</td>
<td>0.385</td>
<td>0.234</td>
<td>0.308</td>
<td>0.332</td>
<td>0.207</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.466</td>
<td>0.241</td>
<td>0.355</td>
<td>0.357</td>
<td>0.200</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.800</td>
<td>0.571</td>
<td>0.666</td>
<td>0.681</td>
<td>0.502</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.867</td>
<td>0.602</td>
<td>0.733</td>
<td>0.764</td>
<td>0.496</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.964</td>
<td>0.811</td>
<td>0.907</td>
<td>0.903</td>
<td>0.707</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.987</td>
<td>0.883</td>
<td>0.948</td>
<td>0.954</td>
<td>0.801</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Test based on (21)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.049</td>
<td>0.024</td>
<td>0.032</td>
<td>0.037</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.431</td>
<td>0.224</td>
<td>0.311</td>
<td>0.290</td>
<td>0.184</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.553</td>
<td>0.282</td>
<td>0.389</td>
<td>0.371</td>
<td>0.227</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.835</td>
<td>0.572</td>
<td>0.720</td>
<td>0.711</td>
<td>0.508</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.906</td>
<td>0.610</td>
<td>0.787</td>
<td>0.787</td>
<td>0.499</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.980</td>
<td>0.825</td>
<td>0.931</td>
<td>0.932</td>
<td>0.741</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.992</td>
<td>0.902</td>
<td>0.966</td>
<td>0.971</td>
<td>0.827</td>
<td></td>
</tr>
</tbody>
</table>

Table 2b. Power of the test for the null of no changes in the largest eigenvalue of $\Sigma$. Data are generated according to (20) and under the alternative hypothesis specified in (23).
<table>
<thead>
<tr>
<th>n</th>
<th>T</th>
<th>( (\rho, \delta) )</th>
<th>((0,0))</th>
<th>((0.5, 0))</th>
<th>((0.5, 0.5))</th>
<th>((0, -0.5))</th>
<th>((0.5, 0.5))</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.071</td>
<td>0.023</td>
<td>0.025</td>
<td>0.027</td>
<td>0.030</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.054</td>
<td>0.040</td>
<td>0.049</td>
<td>0.039</td>
<td>0.034</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Test based on (21)</td>
<td>0.177</td>
<td>0.147</td>
<td>0.143</td>
<td>0.156</td>
<td>0.086</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.162</td>
<td>0.100</td>
<td>0.108</td>
<td>0.106</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.485</td>
<td>0.169</td>
<td>0.234</td>
<td>0.222</td>
<td>0.140</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.224</td>
<td>0.151</td>
<td>0.183</td>
<td>0.185</td>
<td>0.101</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.886</td>
<td>0.226</td>
<td>0.501</td>
<td>0.467</td>
<td>0.223</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.156</td>
<td>0.127</td>
<td>0.129</td>
<td>0.125</td>
<td>0.112</td>
<td></td>
</tr>
</tbody>
</table>

| 50 | Test based on (21) | 0.018                | 0.024     | 0.027       | 0.019       | 0.022       |
|    | Test based on (19) | 0.067                | 0.044     | 0.059       | 0.036       | 0.038       |
| 100| Test based on (21) | 0.211                | 0.120     | 0.134       | 0.141       | 0.090       |
|    | Test based on (19) | 0.156                | 0.067     | 0.092       | 0.107       | 0.060       |
| 200| Test based on (21) | 0.549                | 0.169     | 0.242       | 0.254       | 0.164       |
|    | Test based on (19) | 0.261                | 0.157     | 0.193       | 0.181       | 0.136       |
| 500| Test based on (21) | 0.888                | 0.212     | 0.497       | 0.466       | 0.221       |
|    | Test based on (19) | 0.158                | 0.144     | 0.156       | 0.107       | 0.124       |

| 50 | Test based on (21) | 0.014                | 0.028     | 0.026       | 0.019       | 0.031       |
|    | Test based on (19) | 0.057                | 0.048     | 0.051       | 0.043       | 0.037       |
| 100| Test based on (21) | 0.202                | 0.125     | 0.154       | 0.148       | 0.090       |
|    | Test based on (19) | 0.186                | 0.078     | 0.111       | 0.102       | 0.069       |
| 200| Test based on (21) | 0.536                | 0.152     | 0.253       | 0.247       | 0.117       |
|    | Test based on (19) | 0.247                | 0.145     | 0.177       | 0.185       | 0.103       |
| 500| Test based on (21) | 0.909                | 0.231     | 0.488       | 0.477       | 0.226       |
|    | Test based on (19) | 0.194                | 0.151     | 0.175       | 0.138       | 0.131       |

| 50 | Test based on (21) | 0.012                | 0.020     | 0.017       | 0.019       | 0.016       |
|    | Test based on (19) | 0.071                | 0.033     | 0.053       | 0.056       | 0.029       |
| 100| Test based on (21) | 0.187                | 0.114     | 0.135       | 0.121       | 0.099       |
|    | Test based on (19) | 0.158                | 0.089     | 0.103       | 0.089       | 0.085       |
| 200| Test based on (21) | 0.516                | 0.158     | 0.225       | 0.227       | 0.155       |
|    | Test based on (19) | 0.262                | 0.129     | 0.180       | 0.180       | 0.092       |
| 500| Test based on (21) | 0.915                | 0.262     | 0.495       | 0.488       | 0.251       |
|    | Test based on (19) | 0.179                | 0.114     | 0.126       | 0.153       | 0.099       |

| 50 | Test based on (21) | 0.011                | 0.006     | 0.009       | 0.006       | 0.005       |
|    | Test based on (19) | 0.067                | 0.031     | 0.054       | 0.043       | 0.030       |
| 100| Test based on (21) | 0.182                | 0.101     | 0.122       | 0.142       | 0.084       |
|    | Test based on (19) | 0.174                | 0.055     | 0.072       | 0.086       | 0.050       |
| 200| Test based on (21) | 0.558                | 0.160     | 0.231       | 0.244       | 0.156       |
|    | Test based on (19) | 0.250                | 0.135     | 0.182       | 0.179       | 0.114       |
| 500| Test based on (21) | 0.919                | 0.246     | 0.464       | 0.449       | 0.253       |
|    | Test based on (19) | 0.155                | 0.106     | 0.126       | 0.133       | 0.107       |

| 50 | Test based on (21) | 0.017                | 0.001     | 0.003       | 0.002       | 0.001       |
|    | Test based on (19) | 0.074                | 0.034     | 0.039       | 0.037       | 0.003       |
| 100| Test based on (21) | 0.194                | 0.100     | 0.142       | 0.142       | 0.091       |
|    | Test based on (19) | 0.148                | 0.048     | 0.071       | 0.072       | 0.041       |
| 200| Test based on (21) | 0.588                | 0.167     | 0.225       | 0.225       | 0.162       |
|    | Test based on (19) | 0.249                | 0.121     | 0.171       | 0.183       | 0.079       |
| 500| Test based on (21) | 0.931                | 0.259     | 0.475       | 0.444       | 0.261       |
|    | Test based on (19) | 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 \( \Sigma \). Data are generated according to (20), under the alternative specified in (23).
<table>
<thead>
<tr>
<th>n</th>
<th>T</th>
<th>$\Delta = 0$</th>
<th>$\Delta = \sqrt{\ln \ln T}$</th>
<th>$\Delta = \sqrt{\ln \ln T}$</th>
<th>$\Delta = \sqrt{\ln \ln T}$</th>
<th>$k = k^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>Test based on (21)</td>
<td>0.047</td>
<td>0.087</td>
<td>0.490</td>
<td>0.736</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.052</td>
<td>0.086</td>
<td>0.486</td>
<td>0.721</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.045</td>
<td>0.214</td>
<td>0.573</td>
<td>0.855</td>
<td>0.116</td>
</tr>
<tr>
<td>100</td>
<td>Test based on (19)</td>
<td>0.043</td>
<td>0.171</td>
<td>0.551</td>
<td>0.833</td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.056</td>
<td>0.304</td>
<td>0.628</td>
<td>0.944</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.044</td>
<td>0.269</td>
<td>0.605</td>
<td>0.939</td>
<td>0.125</td>
</tr>
<tr>
<td>200</td>
<td>Test based on (21)</td>
<td>0.062</td>
<td>0.390</td>
<td>0.724</td>
<td>0.995</td>
<td>0.539</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.042</td>
<td>0.383</td>
<td>0.737</td>
<td>0.996</td>
<td>0.063</td>
</tr>
<tr>
<td>500</td>
<td>Test based on (21)</td>
<td>0.005</td>
<td>0.088</td>
<td>0.133</td>
<td>0.270</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.009</td>
<td>0.062</td>
<td>0.112</td>
<td>0.227</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.032</td>
<td>0.228</td>
<td>0.401</td>
<td>0.712</td>
<td>0.103</td>
</tr>
<tr>
<td>100</td>
<td>Test based on (19)</td>
<td>0.021</td>
<td>0.166</td>
<td>0.313</td>
<td>0.675</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.050</td>
<td>0.319</td>
<td>0.567</td>
<td>0.932</td>
<td>0.157</td>
</tr>
<tr>
<td>4</td>
<td>Test based on (19)</td>
<td>0.023</td>
<td>0.264</td>
<td>0.547</td>
<td>0.937</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.065</td>
<td>0.427</td>
<td>0.824</td>
<td>1.000</td>
<td>0.507</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.037</td>
<td>0.422</td>
<td>0.845</td>
<td>1.000</td>
<td>0.082</td>
</tr>
<tr>
<td>5</td>
<td>Test based on (21)</td>
<td>0.002</td>
<td>0.050</td>
<td>0.092</td>
<td>0.204</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.001</td>
<td>0.032</td>
<td>0.061</td>
<td>0.135</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.028</td>
<td>0.213</td>
<td>0.407</td>
<td>0.729</td>
<td>0.005</td>
</tr>
<tr>
<td>100</td>
<td>Test based on (19)</td>
<td>0.019</td>
<td>0.156</td>
<td>0.313</td>
<td>0.686</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.047</td>
<td>0.357</td>
<td>0.624</td>
<td>0.973</td>
<td>0.106</td>
</tr>
<tr>
<td>5</td>
<td>Test based on (19)</td>
<td>0.020</td>
<td>0.288</td>
<td>0.579</td>
<td>0.964</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.059</td>
<td>0.456</td>
<td>0.851</td>
<td>1.000</td>
<td>0.373</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.029</td>
<td>0.415</td>
<td>0.867</td>
<td>1.000</td>
<td>0.660</td>
</tr>
<tr>
<td>6</td>
<td>Test based on (21)</td>
<td>0.001</td>
<td>0.015</td>
<td>0.031</td>
<td>0.077</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.002</td>
<td>0.007</td>
<td>0.017</td>
<td>0.032</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.016</td>
<td>0.216</td>
<td>0.381</td>
<td>0.721</td>
<td>0.047</td>
</tr>
<tr>
<td>100</td>
<td>Test based on (19)</td>
<td>0.009</td>
<td>0.096</td>
<td>0.235</td>
<td>0.591</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.037</td>
<td>0.341</td>
<td>0.668</td>
<td>0.977</td>
<td>0.089</td>
</tr>
<tr>
<td>7</td>
<td>Test based on (19)</td>
<td>0.021</td>
<td>0.246</td>
<td>0.593</td>
<td>0.971</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.065</td>
<td>0.483</td>
<td>0.899</td>
<td>1.000</td>
<td>0.243</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.040</td>
<td>0.447</td>
<td>0.892</td>
<td>1.000</td>
<td>0.067</td>
</tr>
<tr>
<td>5</td>
<td>Test based on (21)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.008</td>
<td>0.160</td>
<td>0.293</td>
<td>0.652</td>
<td>0.029</td>
</tr>
<tr>
<td>100</td>
<td>Test based on (19)</td>
<td>0.003</td>
<td>0.071</td>
<td>0.186</td>
<td>0.525</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.040</td>
<td>0.351</td>
<td>0.671</td>
<td>0.987</td>
<td>0.087</td>
</tr>
<tr>
<td>7</td>
<td>Test based on (19)</td>
<td>0.022</td>
<td>0.210</td>
<td>0.534</td>
<td>0.972</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td>Test based on (21)</td>
<td>0.071</td>
<td>0.523</td>
<td>0.921</td>
<td>1.000</td>
<td>0.147</td>
</tr>
<tr>
<td></td>
<td>Test based on (19)</td>
<td>0.039</td>
<td>0.445</td>
<td>0.903</td>
<td>1.000</td>
<td>0.079</td>
</tr>
</tbody>
</table>

Table 3. Empirical rejection frequencies and power for the null of no change in $\Sigma$. Data are generated as i.i.d. normal with $E(y_1y_1') = I_n$ under $H_0$, and using equation (20) and (23) under $H_1$. 

37
Table 4. Descriptive statistics for monthly and weekly data. We report the mean of the returns, the standard deviation, the skewness and the 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.
Table 5. Critical values. In the Table, the notation $cv_N$ 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. The panels with $(T, p) = (163, 1)$ and $(732, 1)$ contain critical values for unidimensional tests (monthly and weekly frequencies respectively), and therefore are used to test for changes in eigenvalues or when verifying the stability of the diagonal elements of $\Sigma$ one at a time. Panels where $(T, p) = (163, 18)$ and $(732, 18)$ contain critical values for tests with 18 hypotheses under the null, and thus are designed for tests for the stability of one eigenvector.

<table>
<thead>
<tr>
<th>$(T, p)$</th>
<th>level</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$cv_1$</td>
<td></td>
<td>2.8044</td>
<td>3.0716</td>
<td>3.5758</td>
</tr>
<tr>
<td>$(163, 1)$ $cv_3$</td>
<td></td>
<td>3.1987</td>
<td>3.4222</td>
<td>3.9344</td>
</tr>
<tr>
<td>$cv_{18}$</td>
<td></td>
<td>3.7843</td>
<td>3.9690</td>
<td>4.1273</td>
</tr>
<tr>
<td>$cv_1$</td>
<td></td>
<td>2.9471</td>
<td>3.1891</td>
<td>3.6379</td>
</tr>
<tr>
<td>$(732, 1)$ $cv_3$</td>
<td></td>
<td>3.3266</td>
<td>3.5153</td>
<td>3.8789</td>
</tr>
<tr>
<td>$cv_{18}$</td>
<td></td>
<td>3.7779</td>
<td>4.0051</td>
<td>4.5603</td>
</tr>
<tr>
<td>$cv_1$</td>
<td></td>
<td>6.1897</td>
<td>6.4075</td>
<td>6.7747</td>
</tr>
<tr>
<td>$(163, 18)$ $cv_3$</td>
<td></td>
<td>6.5068</td>
<td>6.6486</td>
<td>7.0053</td>
</tr>
<tr>
<td>$cv_1$</td>
<td></td>
<td>6.3066</td>
<td>6.5202</td>
<td>6.9209</td>
</tr>
<tr>
<td>$(732, 18)$ $cv_3$</td>
<td></td>
<td>6.6288</td>
<td>6.8040</td>
<td>7.1498</td>
</tr>
</tbody>
</table>
Table 6. Tests for changes in the variances of the term structure; in the volatilities of each principal component; and in the loadings of each principal component. Rejection at 10%, 5% and 1% levels are denoted with *, ** and *** respectively. Where present, numbers in square brackets are the estimated breakdates, defined as in Remark T4.2.
Table 7. Proportion of the total variance explained by principal components ($\lambda_1$, $\lambda_2$ and $\lambda_3$ 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.

<table>
<thead>
<tr>
<th></th>
<th>monthly data</th>
<th>weekly data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st subsample</td>
<td>2nd subsample</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.790</td>
<td>0.729</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.163</td>
<td>0.214</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.029</td>
<td>0.047</td>
</tr>
</tbody>
</table>
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: Loadings (eigenvectors) of the first and third principal components, before and after mid-April 2008.