Syracuse University

SURFACE

Testing for Instability in Covariance Sturctures

Chihwa Kao
Syracuse University

Lorenzo Trapani
Cass Business School

Giovanni Urga
Cass Business School and Università di Bergamo

Recommended Citation

Center for Policy Research
Working Paper No. 131

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

August 2011

$5.00

CENTER FOR POLICY RESEARCH – Summer 2011

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
Kalena Cortes………………………………Education
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 Rohlfs................................. Economics
Stuart Rosenthal..................................... Economics
Ross Rubenstein ....Public Administration
Perry Singleton……………………………Economics
Margaret Usdansky .... Sociology
Michael Wasylenko .... Economics
Jeffrey Weinstein…………………………Economics
Janet Wilmoth.. Sociology

GRADUATE ASSOCIATES

Kanika Arora ....Public Administration
Christian Buerger ....Public Administration
Il Hwan Chung.........................Public Administration
Alissa Dubnicki....................................... Economics
Andrew Friedson................................... Economics
Clorise Harvey........................Public Administration
Hee Seung Lee ....Public Administration
Jing Li.. Economics

Allison Marier..Economics
Qing Miao ....Public Administration
Wael Moussa..Economics
Kerri Raissian ....Public Administration
Natalee Simpson ....Sociology
Liu Tian....................................Public Administration
Ryan Yeung.............................Public Administration

STAFF

Kelly Bogart..…...…………Administrative Specialist
Karen Cimilluca…………………...Office Coordinator
Kitty Nasto...…...…………Administrative Secretary

Candi Patterson......................Computer Consultant
Roseann Presutti...….…...Administrative Secretary
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.

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

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 apparatus developed in this paper builds on a plethora of results for the changepoint problem available in statistics and in econometrics. Existing testing procedures

(see e.g. the reviews by Csorgo and Horvath, 1997 and Perron, 2006) are typically based on taking the supremum (or some other metric - see Andrews and Ploberger, 1994) of a sequence of CUSUM-type statistics, thus not requiring prior knowledge of the breakdate. In particular, Aue et al. (2009) develop a test for the structural stability of a covariance matrix, based on minimal assumptions. However, a feature of this test is that, by construction, it has power versus breaks occurring at least (respectively, at most) $O\left(\sqrt{T}\right)$ time periods from the beginning (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 overcomes this limitation.

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 $O\left(\ln\ln T\right)$ 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); "$\overset{p}{\rightarrow}$" (convergence in probability); "$\overset{d}{\rightarrow}$" and (convergence in distribution). Orders of magnitude for an almost surely convergent sequence (say s_T) are denoted as $O_{a.s.}(T^{\varsigma})$ and $o_{a.s.}(T^{\varsigma})$ when, for some $\varepsilon > 0$ and $\tilde{T} < \infty$, $P\left[|T^{-\varsigma}s_T| < \varepsilon \text{ for all } T \geq \tilde{T}\right] = 1$ and $T^{-\varsigma}s_T \rightarrow 0$ almost surely respectively. Orders of magnitude for a sequence converging in probability (say s'_T) are denoted as $O_p(T^{\varsigma})$ and $o_p(T^{\varsigma})$ when, for some $\varepsilon > 0$, $\Delta_\varepsilon > 0$ and $\tilde{T}_\varepsilon < \infty$, $P\left[|T^{-\varsigma}s'_T| > \Delta_\varepsilon\right] < \varepsilon$ for all $T > \tilde{T}_\varepsilon$ and $T^{-\varsigma}s'_T \rightarrow 0$ in probability respectively. Standard Wiener processes and Brownian bridges of dimension q are denoted as $W_q(\cdot)$ and $B_q(\cdot)$ respectively; $\|A\|$ denotes the Euclidean norm of a matrix A in $\mathbb{R}^n$, and $|\cdot|_p$ the L_p-norm; the integer part of a real number x is denoted as $\lfloor x \rfloor$.

2 Theoretical framework

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

Strong Invariance Principle and estimation of V_Σ

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

In the sequel, we need the following assumption.

Assumption 1 *(i)* $\sup_t E\|y_t\|^{2r} < \infty$ 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 $r > \frac{2\alpha-1}{\alpha-1}$; *(iii)* letting $V_{\Sigma,T} = T^{-1}E\left[\left(\sum_{t=1}^{T}\bar{w}_t\right)\left(\sum_{t=1}^{T}\bar{w}_t\right)'\right]$, $V_{\Sigma,T}$ is positive definite uniformly in T, and as $T \to \infty$, $V_{\Sigma,T} \to V_{\Sigma}$ with $\|V_{\Sigma}\| < \infty$; *(iv)* letting $\bar{w}_{it}$ be the i-th element of $\bar{w}_i$ and defining $S_{iT,m} \equiv \sum_{t=m+1}^{m+T}\bar{w}_{it}$, it holds that $T^{-1}|E[S_{iT,m}S_{jT,m}] - \varpi_{ij}| \le MT^{-\psi}$, 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, in essence, it implies that y_t is a mixingale (Davidson, 2002a). Many of the DGPs considered in the literature generate NED series - examples include GARCH, bilinear and threshold models (see Davidson, 2002b). Part *(ii)* illustrates the trade-off between the memory of y_t (i.e. its NED size α), and its largest existing moment: as α (the memory of y_t) approaches 1, r has to increase. 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 $T \to \infty$*

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

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

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

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

Remarks

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

T1.2 Equation (2) is an almost sure result, which also provides a rate of convergence. The practical consequence of (2) is that the dependent, heteroskedastic series $\bar{w}_t$ can be replaced with a sequence of *i.i.d.* normally distributed random variables, with the same long run variance as $\bar{w}_t$. The rate δ could, in principle, be derived under different assumptions on the dependence of y_t. A typical finding is $\delta = \frac{1}{2}\left(1 - \frac{1}{r}\right)$ - see Shorack and Wellner (1986).

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

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

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

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

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

In order to apply equation (14) in Theorem 3 below, when using $\hat{V}_\Sigma$ (or $\tilde{V}_\Sigma$), it must hold that $\left\|\hat{V}_\Sigma - V_\Sigma\right\| = o_p\left(\frac{1}{\sqrt{\ln\ln T}}\right)$. Similarly, when using the partial sample estimator $\hat{V}_{\Sigma,\tau}$ (or $\tilde{V}_{\Sigma,\tau}$), it must hold that $\sup_{\lfloor T\tau \rfloor}\left\|\hat{V}_{\Sigma,\tau} - V_\Sigma\right\| = o_p\left(\frac{1}{\sqrt{\ln\ln T}}\right)$.

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

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

6

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

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

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

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

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

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

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

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

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

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

Remarks

T2.1 Equation (4) is based on a SLLN for the case of no autocorrelation in w_t - see also Ling (2007). In principle, δ' can be determined. For example, upon strengthening certain parts of Assumption 1 (chiefly, the $L_{2+\epsilon}$-NED, assuming L_4-NED), a Law of the Iterated Logarithm for mixingales (Hall and Heyde, 1980, Th. 2.21) could be proved.

T2.2 In case of serial dependence, (5) states that it is possible to construct an estimator of V_Σ which has the required rate of convergence as long as both $\ln\ln T/m \to 0$ and $m\ln\ln T/T^{\delta'} \to 0$; a possible choice would be e.g. $m = O\left(\ln T\right)$.

T2.3 The convergence rate in (5) can be refined as in (6). Assumptions 2*(ii)-(iii)* allow for an IP to hold for partial sums of $vec\left[\bar{w}_t\bar{w}'_{t-l} - E\left(\bar{w}_t\bar{w}'_{t-l}\right)\right]$, whence the $O_p\left(T^{-1/2}\right)$ convergence rate, uniformly in τ. Thus, $\sup_{\lfloor T\tau\rfloor}\left\|\tilde{V}_{\Sigma,\tau} - V_\Sigma\right\| = o_p\left[\left(\ln\ln T\right)^{-1/2}\right]$ as long as $\sqrt{\ln\ln T}/m \to 0$ and $m\sqrt{\ln\ln T/T} \to \infty$.

Estimation of the eigensystem

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

Let the i-th eigenvalue-eigenvector couple be defined as (λ_i, x_i); the eigenvectors are defined as an orthonormal basis, i.e. $x'_i x_j = \delta_{ij}$, where δ_{ij} is Kronecker's delta. Since $\Sigma x_i = \lambda_i x_i$, a natural estimator for (λ_i, x_i) is the solution to $\hat{\Sigma}\hat{x}_i = \hat{\lambda}_i\hat{x}_i$, where $\hat{\lambda}_i$ and $\hat{x}_i$ denote the estimates of λ_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 Σ has distinct eigenvalues.

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

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

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

$$\hat{\lambda}_{i,\tau} - \lambda_i = \left(x'_i \otimes x'_i\right) vec\left(\hat{\Sigma}_\tau - \Sigma\right) + O_p\left(T^{-1}\right), \tag{7}$$

$$\begin{aligned} \hat{x}_{i,\tau} - x_i &= \left[\sum_{k\neq i}\frac{x_k}{\lambda_i - \lambda_k}\left(x'_k \otimes x'_i\right)\right] vec\left(\hat{\Sigma}_\tau - \Sigma\right) + O_p\left(T^{-1}\right) \\ &= v_{x,i} vec\left(\hat{\Sigma}_\tau - \Sigma\right) + O_p\left(T^{-1}\right). \end{aligned} \tag{8}$$

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 Waternaux (1976) and Tyler (1981).

P1.2 The asymptotic covariance matrix of $\sqrt{T}\left(\hat{x}_{i,\tau} - x_i\right)$ is $v_{x,i} V_\Sigma v'_{x,i}$; in view of (8), it is singular and has rank $n - 1$. In order to invert $\hat{v}_{x,i} \tilde{V}_\Sigma \hat{v}'_{x,i}$, a Moore-Penrose generalised inverse can be used. The validity of this approach can be shown based on Andrews (1987). Since $\hat{v}_{x,i} \tilde{V}_\Sigma \hat{v}'_{x,i} \xrightarrow{p} v_{x,i} V_\Sigma v'_{x,i}$, it holds that $P\left[r\left(\hat{v}_{x,i} \tilde{V}_\Sigma \hat{v}'_{x,i}\right) \geq r\left(v_{x,i} V_\Sigma v'_{x,i}\right)\right] = 1$ as $T \to \infty$, where $r\left(A\right)$ denotes the rank of A. Also, by construction and for any T, $r\left(\hat{v}_{x,i} \tilde{V}_\Sigma \hat{v}'_{x,i}\right) \leq n - 1$. Thus, as $T \to \infty$, $P\left[r\left(\hat{v}_{x,i} \tilde{V}_\Sigma \hat{v}'_{x,i}\right) = r\left(v_{x,i} V_\Sigma v'_{x,i}\right)\right] = 1$. 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 appendix that

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

as $T \to \infty$. As far as the impact of n is concerned, V_Σ is an n^2-dimensional matrix; thus, in general the quadratic form $\left(x'_i \otimes x'_k\right) V_\Sigma \left(x_k \otimes x_i\right)$ has magnitude of order $O\left(n^2\right)$; also, due to the summation on the right hand side of (9) involving $n - 1$ elements, $E\left[T\left(\hat{\lambda}_{i,\tau} - \lambda_i\right)\right] = O\left(n^3\right)$. Thus, the asymptotic bias is of order $O\left(\frac{n^3}{T}\right)$; a bias-corrected version is $\tilde{\lambda}_{i,\tau} = \hat{\lambda}_{i,\tau} - T^{-1} \sum_{k \neq i} \left[\hat{x}'_i \otimes \hat{x}'_k\right] \frac{\tilde{V}_\Sigma}{\hat{\lambda}_i - \hat{\lambda}_k} \left[\hat{x}_k \otimes \hat{x}_i\right]$. The bias is always positive for the largest eigenvalue. This result is of independent interest. It could be useful e.g. when measuring the percentage of the total variance of y_t explained by each of its principal components. Similarly, we show that

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

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

9

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

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

Thus, $\hat{\gamma}_{i,\tau} - \gamma_i = v_{\gamma,i} vec \left(\hat{\Sigma}_\tau - \Sigma \right) + o_p(1)$, with $v_{\gamma,i} = \frac{1}{2} \frac{x_i}{\lambda_i^{1/2}} (x_i' \otimes x_i') + \sum_{k \neq i} \frac{\lambda_i^{1/2} x_k}{\lambda_i - \lambda_k} (x_i' \otimes x_k')$.

Consider the following notation. Define $\lambda \equiv [\lambda_1, ..., \lambda_n]'$ as the n-dimensional vector containing the eigenvalues sorted in descending order; $X \equiv [x_1|...|x_n]$, and $\Gamma \equiv [\gamma_1|...|\gamma_n]$; $\hat{z} \equiv \left[\hat{\lambda}', vec\hat{X}', vec\hat{\Gamma}' \right]'$ with $\hat{z}_\tau - z = D_{\lambda x \gamma} vec \left(\hat{\Sigma}_\tau - \Sigma \right) + o_p(1)$ and $D_{\lambda x \gamma} \equiv [x_1 \otimes x_1, ..., x_n \otimes x_n, v_{x,1}', ..., v_{x,n}', v_{\gamma,1}', ..., v_{\gamma,n}']'$.

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

Corollary 1 *Under Assumptions 1 and 3, as $T \to \infty$*

$$
\sqrt{T} (\hat{z}_\tau - z) \xrightarrow{d} [V_z]^{1/2} W_{n(2n+1)} (\tau),
$$

$$
T (\hat{z}_\tau - z) = \sum_{t=1}^{\lfloor T\tau \rfloor} \tilde{X}_t + O_{a.s.} \left(\lfloor T\tau \rfloor^{\frac{1}{2} - \delta} \right),
$$

uniformly in τ, where $V_z = D_{\lambda x \gamma} V_\Sigma D_{\lambda x \gamma}'$ and $\tilde{X}_t$ is a zero mean, i.i.d. Gaussian sequence with $E \left(\tilde{X}_t \tilde{X}_t' \right) = V_z$ and $\delta > 0$.

Corollary 1 entails that

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

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

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

with: V_λ a matrix with (i,j)-th element given by $V_{ij}^\lambda = (x_i' \otimes x_i') V_\Sigma (x_j \otimes x_j)$; V_X an $(n^2 \times n^2)$-dimensional matrix whose (i,j)-th $n \times n$ block (say V_{ij}^X) is defined as $V_{ij}^X = \sum_{k \neq i} \sum_{h \neq j} \frac{x_k x_h'}{(\lambda_i - \lambda_k)(\lambda_i - \lambda_k)} [x_i' \otimes x_k'] V_\Sigma [x_j \otimes x_h]$; V_Γ is an $(n^2 \times n^2)$-dimensional matrices whose (i,j)-th $n \times n$ block is defined as $V_{ij}^\Gamma = v_{\gamma,i} V_\Sigma v_{\gamma,j}'$.

3   Testing

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

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

$$\tilde{S}(\tau) = R \times D_{\lambda x\gamma} \times \left[S(\tau) - \frac{\lfloor T\tau \rfloor}{T} S(T) \right], \tag{11}$$

with $\tilde{S}(\tau) = 0$ for $\tau \leq \frac{1}{T}$ or $\geq 1 - \frac{1}{T}$, and R a $p \times n(2n+1)$ 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 $D_{\lambda x\gamma} \times \left[S(\tau) - \frac{\lfloor T\tau \rfloor}{T} S(T) \right]$. Thence, testing is carried out by using

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

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

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

Theorem 3 *Under Assumptions 1-3, as $T \to \infty$*

$$\sup_{\lfloor T\tau_1 \rfloor \leq \lfloor T\tau \rfloor \leq \lfloor T\tau_2 \rfloor} \Lambda_T(\tau) \xrightarrow{d} \sup_{\tau_1 \leq \tau \leq \tau_2} \frac{\|B_p(\tau)\|}{\sqrt{\tau(1-\tau)}}, \tag{13}$$

where $B_p(\tau)$ is a p-dimensional standard Brownian bridge and $[\tau_1, \tau_2] \in (0,1)$. Also

$$P \left\{ a_T \left[\sup_{p < \lfloor T\tau \rfloor < T-p} \Lambda_T(\tau) \right] \leq x + b_T \right\} \to e^{-2e^{-x}}, \tag{14}$$

11

where $a_T = \sqrt{2 \ln \ln T}$ and $b_T = 2 \ln \ln T + \frac{p}{2} \ln \ln \ln T - \ln \Gamma \left(\frac{p}{2} \right)$, with $\Gamma \left(\cdot \right)$ 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 showing an IP for $S \left(\tau \right)$, and applying the Continuous Mapping Theorem (CMT). As noted in Corollary 1 in Andrews (1993, p. 838), $\Lambda_T \left(\tau \right)$ is not continuous at $\{0,1\}$ and $\sup_{1 \leq \lfloor T\tau \rfloor \leq T} \Lambda_T \left(\tau \right) \overset{p}{\rightarrow} \infty$ under H_0. Thus, trimming is necessary in this case; alternatively, a weighted norm can be employed, defining $\Lambda_T^{*2} \left(\tau \right) = \frac{\tilde{S}(\tau)' \left[R\tilde{V}_{z,\tau} R' \right]^{-1} \tilde{S}(\tau)}{T[\tau(1-\tau)]^\alpha}$ for $\alpha \in [0,1)$ (see Csorgo and Horvath, 1997, and Chen et al., 2005) and taking $\sup_{1 \leq \lfloor T\tau \rfloor \leq T} \Lambda_T^{*2} \left(\tau \right)$.

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 $\Lambda_T \left(\tau \right)$ with the appropriate norming constants, say a_T and b_T, an Extreme Value (EV henceforth) theorem can be employed. Tests based on $\sup_{p < \lfloor T\tau \rfloor < T-p} \left[a_T \Lambda_T \left(\tau \right) - b_T \right]$ are designed to be able to detect breaks close to the end of the sample.

Theorem 3 allows to test for breaks in Σ when n is finite. As n passes to infinity, Aue et al. (2009) derive a "sequential limit" (see Phillips and Moon, 1999, for the definition of sequential and joint convergence) result for $\hat{\Lambda}_T^2 \equiv \sup_{1 \leq \lfloor T\tau \rfloor \leq T} T^{-1} \tilde{S} \left(\tau \right)' \left[R\tilde{V}_{z,\tau} R' \right]^{-1} \tilde{S} \left(\tau \right)$, showing that, as $T \rightarrow \infty$ followed by $n \rightarrow \infty$, $\frac{4\hat{\Lambda}_T^2 - n^2}{\sqrt{2n}} \overset{d}{\rightarrow} N \left(0,1 \right)$.

The following corollary reports the result for $(n,T) \rightarrow \infty$ jointly.

Corollary 2 *Let Assumptions 1-3 hold, with $\sup_{i,t} E |y_{it}|^{4r} < \infty$ for $r > 2$, and define $\check{\Lambda}_{nT}^2 \equiv \sup_\tau \tilde{\Lambda}_{nT}^2 \left(\tau \right)$, where $\tilde{\Lambda}_{nT}^2 \left(\cdot \right) = \frac{T}{\lfloor T\tau \rfloor \lfloor T(1-\tau) \rfloor} \times \left[S \left(\tau \right) - \frac{\lfloor T\tau \rfloor}{T} S \left(T \right) \right]' \tilde{V}_{\Sigma,\tau}^{-1} \left[S \left(\tau \right) - \frac{\lfloor T\tau \rfloor}{T} S \left(T \right) \right]$. As $(n,T) \rightarrow \infty$ with $\frac{n}{T^\delta} \rightarrow 0$, $\frac{nm}{\sqrt{T}} \rightarrow 0$ and $\frac{n}{m} \rightarrow 0$, it holds that*

$$\frac{1}{\sqrt{2n}} \left(\check{\Lambda}_{nT}^2 - n^2 \right) \overset{d}{\rightarrow} N \left(0,1 \right).$$

The Corollary states, in essence, that normality holds for large T and relatively small n. The restriction $\frac{n}{T^\delta} \rightarrow 0$ arises from the approximation error in the SIP. As far as both $\frac{nm}{\sqrt{T}} \rightarrow 0$

and $\frac{n}{m} \to 0$ are concerned, they come from the estimated long run covariance matrix, $\tilde{V}_\Sigma$. Heuristically, each element in $\tilde{V}_\Sigma$ has an error of magnitude $O_p\left(\frac{m}{\sqrt{T}}\right) + O_p\left(\frac{1}{m}\right)$; each column contains n^2 elements, so that the contribution of the matrix estimation error is $O_p\left(\frac{n^2 m}{\sqrt{T}}\right) + O_p\left(\frac{n^2}{m}\right)$. This is then normalised by n, whence the restrictions.

Power/consistency of the test

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

$$H_a^{(T)} : vech(\Sigma_t) = \begin{cases} vech\left(\Sigma\right) \text{ for } t = 1, ..., k_{0,T} \\ vech\left(\Sigma\right) + \Delta_T \text{ for } t = k_{0,T} + 1, ..., T \end{cases}, \tag{15}$$

where both the changepoint $(k_{0,T})$ and the size of the break (Δ_T) could depend on T. More general alternatives could be considered (see e.g. Andrews, 1993; Csorgo and Horvath, 1997).

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

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

$$\frac{1}{\ln \ln T}\left[\frac{(T - k_{0,T})\, k_{0,T}}{T} \|\Delta_T\|^2\right] \to \infty, \tag{16}$$

it holds that

$$P\left[\sup_{p<\lfloor T\tau \rfloor < T-p} \Lambda_T(\tau) > c_\alpha\right] = 1. \tag{17}$$

Remarks

T4.1 Theorem 4 illustrates the impact of $k_{0,T}$ and Δ_T on the power of the test. Particularly, consider the two extreme cases:

T4.1.a $k_{0,T} = O(T)$ - i.e. the break occurs in the middle of the sample. The test is powerful as long as the size of the break is at least as big as $O\left(\sqrt{\frac{\ln \ln T}{T}}\right)$. When using trimmed statistics such as in (13), the test is powerful versus mid-sample alternatives of size

$O\left(\sqrt{\frac{1}{T}}\right)$: when no trimming is used, there is some, limited loss of power versus mid-sample alternatives.

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

T4.2 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 $\widehat{\lfloor T\tau\rfloor} = T\times \arg\sup_{\lfloor T\tau\rfloor} \Lambda_T(\tau)$; this estimator is such that $\widehat{\lfloor T\tau\rfloor} - \lfloor T\tau\rfloor = O_p(1)$. The sample can be then split around $\widehat{\lfloor T\tau\rfloor}$, and testing can be applied to each subsample.

T4.3 The presence of "large" breaks in Σ is bound to affect inference on the eigensystem - see e.g. Stock and Watson (2002). Consider, as a leading example, $\hat{\lambda}_i$. From Proposition 1, the long-run variance of the estimated eigenvalues is estimated by $(\hat{x}_i'\otimes\hat{x}_i')\,\tilde{V}_\Sigma\,(\hat{x}_i\otimes\hat{x}_i)$, thus depending on $\hat{x}_i$. In presence of a 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. Thus, $(x_i'\otimes x_i')\,V_\Sigma\,(x_i\otimes x_i)$ is estimated with an error of the same order of magnitude as $\|\hat{x}_i - x_i\|^4 \left\|\tilde{V}_\Sigma - V_\Sigma\right\|$. Since $\hat{\Sigma} - \Sigma = O_p(1)$, both $\|\hat{x}_i - x_i\|$ and $\left\|\tilde{V}_\Sigma - V_\Sigma\right\|$ are $O_p(1)$, which ensures that $(\hat{x}_i'\otimes\hat{x}_i')\,\tilde{V}_\Sigma\,(\hat{x}_i\otimes\hat{x}_i)$ is bounded in probability. This entails that the power of tests based on $\sup_{\lfloor T\tau\rfloor} \Lambda_T(\tau)$ does not vanish asymptotically.

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: using the EV

distribution in (14), or using the approximation proposed in Csorgo and Horvath (1997, ch. 1.3.2).

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

$$P \left\{ \sup_{h_T + \frac{p}{T} \leq \tau \leq 1 - \left(h_T + \frac{p}{T} \right)} \left[\sum_{i=1}^{p} \frac{B_{1,i}^2 (\tau)}{\tau (1 - \tau)} \right]^{1/2} \leq c_\alpha' \right\} = 1 - \alpha, \tag{18}$$

where the $B_{1,i} (\tau)$s are independent, univariate, squared Brownian bridges, generated over a grid of dimension T. We set $T \times h_T = [\ln (T)]^{1 + \ln \ln \ln T}$. We report here a table containing critical values c_α' for several combinations of p and T.

[Insert Table 1 somewhere here]

4.1 Finite sample properties of the test

We evaluate size and power through a Monte Carlo exercise. As a leading example, we consider a test for a change in the first eigenvalue of the covariance matrix: thus, the number of constraints is $p = 1$. 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 p 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 Σ.

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 $n = \{3, 4, 5, 6, 7, 10, 15, 20\}$. Serial dependence in y_t is introduced through an ARMA(1,1) process:

$$y_t = \rho y_{t-1} + e_t + \theta e_{t-1}, \tag{19}$$

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%. Unreported results based on $\sup_{p < \lfloor T\tau \rfloor < T-p} \Lambda_T(\tau)$ show that the empirical size overstates the nominal size level in small samples. To attenuate this, we propose to increase the trimming at each end of the sample as

$$\sup_{t^* \leq \lfloor T\tau \rfloor \leq T-t^*} \Lambda_T(\tau), \tag{20}$$

where $t^* = \left\lfloor (\ln T)^{1+\ln\ln\ln T} \right\rfloor$ - this is a slowly varying function of T, which we have found to work well for all the cases considered. Whilst this 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 $O(\ln T)$ periods from the beginning/end of sample.

[Insert Table 2 somewhere here]

The test is oversized 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 (always with a tendency towards oversizement). The table also

shows that n does not seem to play a role in affecting the behaviour of the size - this can be compared with Table 4 below, which illustrates the impact of p.

Power

We conduct our simulations under alternatives defined as

$$\begin{cases} \Sigma \text{ for } t = 1, ..., k \\ \Sigma + \Delta \text{ for } t = k+1, ..., T \end{cases} \quad (21)$$

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

$$\Delta = \sqrt{\frac{\ln \ln T}{T^\beta}} \times I_n. \quad (22)$$

We set $\beta = \left\{ \frac{2}{3}, \frac{1}{2} \right\}$ in Table 3a. In Table 3b, we also report power versus alternatives close to the beginning of the sample, with $k = 2 \times [\ln(T)]^{1 + \ln \ln \ln(T)}$ and $\Delta = I_n$.

[Insert Tables 3a-3b somewhere here]

Considering mid-sample alternatives, 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 AR or MA disturbances, although a reduction in power is seen in the ARMA(1,1) case.

Table 3b 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 Table 3a, as n increases the power slightly increases. As opposed to the results in Table 3a, 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. As a guideline for empirical applications, this entails that if an AR structure is found in the data, pre-whitening should be applied.

Testing for the constancy of Σ: the role of p

We report a simulation exercise for the null of no change in Σ. The alternative is the same as in (21). Data are generated as *i.i.d.* Gaussian with $\Sigma = I_n$. Thus, the number of hypotheses is $p = \frac{1}{2}n(n+1)$; in this setup, we report results for $n = \{3, 4, 5, 6, 7\}$ and $T = \{50, 100, 200, 500\}$.

Tests are found to be grossly oversized, in small samples, when trimming is done according to (20). Thus, we propose to trim away a larger portion of the data, namely

$$t^* = p \times [\ln \ln (T) - 1] + [\ln (T)]^{1 + \ln \ln \ln(T)},$$

which is the same as in (20), plus the extra term $p \times [\ln \ln (T) - 1]$. This yields a more conservative test. When generating data under the alternative, this is defined as in (21) and (22). In the last column, $k^* = p \times [\ln \ln (T) - 1] + [\ln (T)]^{1 + \ln \ln \ln(T)}$ and $\Delta = I_n$.

[Insert Table 4 somewhere here]

Table 4 illustrates the role played by p. As p increases, the test becomes increasingly conservative in finite samples. As $T \to \infty$, the empirical rejection frequencies approach their nominal values. Tables 2 and 4 show that size distortion arises from p rather than n itself, and this can be resolved with appropriate trimming. Although this hinders the ability of the test to detect changes closer to either end of the sample, however, as T increases, the amount of trimming is lower than when using the 15%-from-each-end rule (Andrews, 1993) - in some cases, decidedly lower (e.g. when $T = 1000$ and $n = 3$, trimming at t^* would entail eliminating 60 datapoints, whilst 300 would be eliminated with the 15% rule).

The power of the test is hampered by the trimming - see the cases $n = 6$ and $n = 7$ when $T = 50$ or 100. When considering mid-sample alternatives, the power increases monotonically with T as expected, and it does not seem to be affected by p in a significant way for large samples ($T = 500$). For finite samples, where the effect of the trimming is more severe, p does have a significant impact and the power decreases as p increases. Slightly different considerations seem to apply to the case of breaks closer to the beginning of the sample, i.e. the last column of Table

4. In this case, as p increases, so does the power, which could also reflect the fact that, as p increases, the location of the simulated changepoint approaches the middle of the sample.

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

19

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 5 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 5 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 dependence, and, with higher maturities, no dependence 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 $\Sigma = E(y_t y_t')$. 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 α_P is the size of the whole procedure. Using these critical values yields, approximately, a level α_P not greater than 1%, 5% and 10% corresponds to conducting each test at levels $\alpha_I = 0.056\%$, 0.28% and 0.56% respectively.

Critical values for individual tests of levels $\alpha_I = 1\%$, 5% and 10%, and for procedure level $\alpha_P = 1\%$, 5% and 10%, are in Table 6.

[Insert Table 6 somewhere here]

As a second step, we verify whether slope, level and curvature are constant over time. Particularly, we carry out separately the detection of changes in the volatility of the principal components (verifying the time stability of the three largest eigenvalues, say λ_1, λ_2 and λ_3), and in their loading (verifying the stability of the 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_is has the same sign for all values of τ, 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 7.

[Insert Table 7 somewhere here]

It is well known that controlling the procedure-wise error by a Bonferroni correction can be rather conservative. In our case, the values of test statistics (Table 7) can be contrasted with the critical values to be used for single hypothesis testing (reported in Table 6 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 7, when testing for the constancy of the diagonal elements of Σ with weekly data. When using cv_1 as critical values, two maturities (the 30 months and the 3 years ones) now appear to have a break. The rest of the results (especially the absence of breaks in monthly data) is the same as when using a Bonferroni correction.

Table 7 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 λ_3, the volatility of the curvature, which has a break significant at 10%. The second and third panel of the table show that the principal component structure has a change in the size of the curvature,

significant at 5%. The corresponding estimated breakdate, selected, according to Remark T4.2, as the maximizer of the CUSUM statistic, is January 2008. Table 8 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.

[Insert Table 8 and Figure 2 somewhere here]

As far as weekly data are concerned, there is evidence of instability in the covariance structure. At a "macro" level, the variances of longer maturities (from 4 years onwards) change, whilst the variances of shorter maturities are constant. For most maturities, the breakdate is around the first week of December 2007. This is expected, since December 2007 is generally associated with the deepening of the recent recession. It is interesting to note that the longest maturity, the 10-year one, has a break at around the last week of August 2008. As far as principal components are concerned, the second panel of Table 7 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 a covariance matrix and its eigensystem. 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 large sample properties for the test are satisfactory: the correct size is attained under various degrees of serial dependence, and the test exhibits good power. As far as the small sample performance is concerned, some trimming at the beginning/end of sample is required in order for the test to have the correct size, at least in finite samples. 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_{it}^{Δ} for some $\Delta \in (0, 1)$, although the optimal choice of Δ is not straightforward. An open issue, moreover, is the impact of the dimensionality, p/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

[1] Andrews, D.W.K. (1987), "Asymptotic Results for Generalised Wald Tests," *Econometric Theory* 3, 348-358.

[2] Andrews, D.W.K. (1993), "Tests for Parameter Instability and Structural Changes with Unknown Change Point," *Econometrica* 61, 821-856.

[3] Andrews, D.W.K. and W. Ploberger (1994), "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica* 62, 1383-1414.

[4] Audrino, F., Barone-Adesi, G., and A. Mira (2005), "The stability of factor models of interest rates," *Journal of Financial Econometrics* 3, 422-441.

[5] Aue, A., Hormann, S. Horvath, L., and M. Reimherr (2009), "Break Detection in the Covariance Structure of Multivariate Time Series Models," *Annals of Statistics* 37, 4046-4087.

[6] Bai, J. (1997), "Estimating Multiple Breaks One at a Time," *Econometric Theory* 13, 315-352.

[7] Bliss, R.R. (1997), "Movements in the Term Structure of Interest Rates," *Economic Review*, Federal Reserve Bank of Atlanta, Q4, 16-33.

[8] Bliss, R.R., and D.C. Smith (1997), "The Stability of Interest Rate Processes," Federal Reserve Bank of Atlanta, Working Paper 97-13.

[9] Castle, J.R., Fawcett, N., and D.F. Hendry (2010), "Forecasting with Equilibrium-Correction Models During Structural Breaks," *Journal of Econometrics* 158, 25-36.

[10] Chen, G., Choi, Y.K, and Y. Zhou (2005), "Nonparametric Estimation of Structural Change Points in Volatility Models for Time Series," *Journal of Econometrics* 126, 79-114.

[11] Corradi, V. (1999), "Deciding Between $I(0)$ and $I(1)$ via FLIL-Based Bounds," *Econometric Theory* 15, 643-663.

[12] Csorgo, M., and L. Horvath (1997), *Limit Theorems in Change-Point Analysis*, Wiley, Chichester.

[13] Davidson, J. (2002a), *Stochastic Limit Theory*. Oxford University Press, Oxford.

[14] Davidson, J. (2002b), "Establishing Conditions for the Functional Central Limit Theorem in Nonlinear and Semiparametric Time Series Processes," *Journal of Econometrics* 106, 243-269.

[15] Davis, C., and W. Kahan (1970), "The Rotation of Eigenvectors by a Perturbation. III," *SIAM Journal of Numerical Analysis* 7, 1-46.

[16] Eberlein, E. (1986), "On Strong Invariance Principles under Dependence Assumptions," *Annals of Probability* 14, 260-270.

[17] Flury, B.H. (1984), "Common Principal Components in k Groups," *Journal of the American Statistical Association* 79, 892-898.

[18] Flury, B.H. (1988), *Common Principal Components and Related Multivariate Models*, Wiley, New York.

[19] Gallant, A.R., and H. White (1988), *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*, Oxford: Basil Blackwell.

[20] Hall, P., and C.C. Heyde (1980), *Martingale Limit Theory and its Applications*, New York: New York University Press.

[21] Horn, R.A., and C.R. Johnson (1999), *Matrix Analysis*, Cambridge: Cambridge University Press.

[22] Horvath, L. (1993), "The Maximum Likelihood Method for Testing Changes in the Parameters of Normal Observations," *Annals of Statistics* 21, 672-680.

[23] Hyndman, R.J., and M.S. Ullah (2007), "Robust Forecasting of Mortality and Fertility Rates: a Functional Data Approach," *Computational Statistics and Data Analysis* 51, 4942–4956.

[24] Lee, R.D., and L. Carter (1992), "Modeling and Forecasting the Time Series of U.S. Mortality," *Journal of the American Statistical Association* 87, 659-71.

[25] Lehmann, E.L., and J.P. Romano (2005), "Generalizations of the Familywise Error Rate," *Annals of Statistics* 33, 1138-1154.

[26] Ling, S. (2007), "Testing for Change Points in Time Series Models and Limiting Theorems for NED Sequences," *Annals of Statistics* 35, 1213-1237.

[27] Litterman, R., and J.A. Scheinkman (1991), "Common Factors Affecting Bond Returns," *Journal of Fixed Income* 1, 54-61.

[28] Perignon, C., and C. Villa (2006), "Sources of Time Variation in the Covariance Matrix of Interest Rates," *Journal of Business* 79, 1535-1549.

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

[30] Phillips, P.C.B., and H.R. Moon (1999), "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica* 67, 1057-1112.

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

[32] Shorack, G.R., Wellner, J.A. (1986), *Empirical Processes with Applications to Statistics*, Wiley, New York.

[33] Stock, J. H., and M.W. Watson (1999), "Forecasting Inflation," *Journal of Monetary Economics* 44, 293-335.

[34] Stock, J. H., and M.W. Watson (2002), "Forecasting Using Principal Components From a Large Number of Predictors," *Journal of the American Statistical Association* 97, 1167-1179.

[35] Stock, J. H., and M.W. Watson (2005), "Implications of Dynamic Factor Models for Var Analysis," manuscript.

[36] Tyler, D.E. (1981), "Asymptotic Inference for Eigenvectors," *Annals of Statistics* 9, 725-736.

[37] Waternaux, C. (1976), "Asymptotic Distribution of the Sample Roots for a Non-Normal Population," *Biometrika* 63, 639-45.

[38] Yang, S.S., Yue, J.C., and H.C. Huang (2010), "Modeling Longevity Risks Using a Principal Component Approach: a Comparison with Existing Stochastic Mortality Models," *Insurance: Mathematics and Economics* 46, 254-270

Appendix A: Preliminary Lemmas

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

$$|B\rho|_{2+\delta} \leq \left[|\rho|_{q^*}^{r-2} |B|_{q^*/q-1}^{r-2}\right]^{1/2(r-1)} \times \tag{23}$$
$$\left[|B\rho|_r^{[2r(1+\delta)-\delta r^2]/[2+\delta]} + |B\rho|_r^r\right]^{1/2(r-1)}.$$

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

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

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

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

Substituting for C gives (23). ∎

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 (23) is very similar to Lemma 17.15 in Davidson (2002a, p. 271). Particularly, $|\rho|_{q^*}^{r-2}$, for some $q^* < 2$, is raised to the power of $r - 2$: this is exactly the same as in

Lemma 17.15 in Davidson (2002a). This is an important consequence of the Lemma. Given a non-Lipschitz transformation of a NED sequence u_t, say $\phi(u_t)$, setting $u_t^m = E[u_t| u_{t-m}, ..., u_{t+m}]$, one would look for a bound to $|\phi(u_t) - \phi(u_t^m)|_{2+\delta}$. This would be majorized by some suitably chosen $|B(u_t, u_t^m) \rho(u_t, u_t^m)|_{2+\delta}$; since normally one would choose $\rho(u_t, u_t^m)$ as the taxicab distance, it is $|\rho|_{q^*}^{r-2}$ that gives the size of $\phi(u_t)$.

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

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

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

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

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

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

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

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

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

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

Appendix B: Proofs

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

As far as (2) is concerned, its proof is based on Eberlein (1986). Lemma 2 entails that $\bar{w}_t$ is a zero-mean $L_{2+\epsilon}$-mixingale of size $\alpha'' > \frac{1}{2}$. Letting $\Im_m = \{\bar{w}_{1,\ldots,}, \bar{w}_m\}$ and $S_{Tm} \equiv \sum_{t=m+1}^{m+T} \bar{w}_t$, (2) follows if $\|E[S_{Tm}|\Im_m]\|_2 < \infty$ and $\|E[S_{iTm}S_{jTm}|\Im_m] - E[S_{iTm}S_{jTm}]\|_1 = O\left(T^{1-\theta}\right)$ 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

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

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

$$P\left[\sup_{g_T \leq \lfloor T\tau \rfloor \leq T} \frac{1}{\lfloor T\tau \rfloor^{\delta'}} \left\|\hat{\Psi}_{l,\lfloor T\tau \rfloor} - \Psi_l\right\| > \varepsilon\right] \leq \varepsilon',$$

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

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

In order to prove (5), write

$$
\begin{aligned}
\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\tilde V_{\Sigma,\tau}-V_\Sigma\right\| &\le \sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\hat V_{\Sigma,\tau}-V_\Sigma\right\|\\
&+2\sup_{1\le\lfloor T\tau\rfloor\le T}\sum_{l=0}^{m}\left(1-\frac{l}{m}\right)\left\|\hat\Psi_{l,\lfloor T\tau\rfloor}-\Psi_l+\hat\Psi_{l,1-\lfloor T\tau\rfloor}-\Psi_l\right\|\\
&+2\sum_{l=1}^{m}\frac{l}{m}\left\|\Psi_l\right\|+2\sum_{l=m+1}^{\infty}\left\|\Psi_l\right\|\\
&= I+II+III+IV.
\end{aligned}
\tag{24}
$$

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

$$
\max_{1\le l\le m}E\left[\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\hat\Psi_{l,\lfloor T\tau\rfloor}+\hat\Psi_{l,1-\lfloor T\tau\rfloor}-\Psi_l\right\|\right]\sum_{l=1}^{m}\left(1-\frac{l}{m}\right)=O\left(\frac{1}{T^{\delta'}}\right)O\left(m\right);
\tag{25}
$$

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

Consider (6). We still use (24) 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 proofs of Theorem 1 yield an IP for each l, so that $\sup_{1\le\lfloor T\tau\rfloor\le T}\left\|\hat\Psi_{l,\lfloor T\tau\rfloor}-\Psi_l\right\|=O_p\left(T^{-1/2}\right)$. Thus (25) becomes

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

Putting all together, (5) follows. ∎

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

$$
\left[\Sigma+\left(\hat\Sigma_\tau-\Sigma\right)\right]\left[x_i+\left(\hat x_{i,\tau}-x_i\right)\right]=\left[\lambda_i+\left(\hat\lambda_{i,\tau}-\lambda_i\right)\right]\left[x_i+\left(\hat x_{i,\tau}-x_i\right)\right].
\tag{26}
$$

After expanding the product, consider the terms $\left(\hat\Sigma_\tau-\Sigma\right)\left(\hat x_{i,\tau}-x_i\right)$ and $\left(\hat\lambda_{i,\tau}-\lambda_i\right)\left(\hat x_{i,\tau}-x_i\right)$. It holds that $\hat\lambda_{i,\tau}-\lambda_i=O_p\left(T^{-1/2}\right)$ uniformly in τ. This is because Σ is symmetric, and there-

fore Corollary 6.3.4 in Horn and Johnson (1995, p. 367) entails that $\left|\hat{\lambda}_{i,\tau} - \lambda_i\right| \leq \left\|\hat{\Sigma}_\tau - \Sigma\right\|$. Equation (1) yields the result. Also, it holds that $\hat{x}_{i,\tau} - x_i = O_p\left(T^{-1/2}\right)$ uniformly in τ. This follows from the $\sin\Theta$ Theorem in Davis and Kahan (1970), whereby $\left\|\hat{X}_\tau - X\right\| \leq \|X\| \left\|\hat{\Sigma}_\tau - \Sigma\right\|$. Thus, $\left(\hat{\Sigma}_\tau - \Sigma\right)(\hat{x}_{i,\tau} - x_i)$ and $\left(\hat{\lambda}_{i,\tau} - \lambda_i\right)(\hat{x}_{i,\tau} - x_i)$ are $O_p\left(T^{-1}\right)$ uniformly in τ; omitting them, (26) can be written as

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

The x_is are a complete (and orthonormal) basis. Thus, given an arbitrary set of constants $\phi_{j,\tau}$, it holds that $\hat{x}_{i,\tau} - x_i = \sum_{j=1}^n \phi_{j,\tau}x_j$. Recalling that $\Sigma x_i = \lambda_i x_i$, and premultiplying (27) by x_i' we obtain $\lambda_i\phi_{i,\tau} + x_i'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i = \lambda_i\phi_{i,\tau} + \left(\hat{\lambda}_{i,\tau} - \lambda_i\right)$, which entails (7). To prove (8), one can multiply (27) by any x_k', whence $\lambda_k\phi_{k,\tau} + x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i = \lambda_i\phi_{k,\tau}$. This yields $\phi_{k,\tau} = \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i}{\lambda_i - \lambda_k}$, under Assumption 2 which stipulates that $\lambda_i \neq \lambda_k$ for all $i \neq k$. From $\hat{x}_{i,\tau} - x_i = \sum_{j=1}^n \phi_{j,\tau}x_j$ we obtain $\hat{x}_{i,\tau} - x_i = \sum_{k \neq i}\frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i}{\lambda_i - \lambda_k}x_k + \phi_{i,\tau}x_i$; (8) follows from setting $\phi_{i,\tau} = 0$.

We now turn to deriving the bias for $\hat{\lambda}_{i,\tau} - \lambda_i$, presented in (9). Expanding (26) and premultiplying by x_i' we obtain

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

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

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

33

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

$$E\left[Tx_i'\left(\hat{\Sigma}_\tau - \Sigma\right)\left(\hat{x}_{i,\tau} - x_i\right)\right] \tag{28}$$

$$= \sum_{k\neq i}\left[\frac{x_i'\otimes x_k'}{\lambda_i - \lambda_k}\right]E\left\{\left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right]\left[vec\left(\hat{\Sigma}_\tau - \Sigma\right)\right]'\right\}[x_k\otimes x_i]$$

$$= \sum_{k\neq i}\frac{\left(x_i'\otimes x_k'\right)V_\Sigma\left(x_k\otimes x_i\right)}{\lambda_i - \lambda_k}.$$

The bias for $\hat{x}_{i,\tau} - x_i$ can be derived from (8) following similar passages. Using (26), $\lambda_k\phi_{k,\tau} + x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i + x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)\left(\hat{x}_{i,\tau} - x_i\right) = \lambda_i\phi_{k,\tau} + \left(\hat{\lambda}_{i,\tau} - \lambda_i\right)\phi_{k,\tau}$, whence

$$\phi_{k,\tau} = \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i^+ + x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)\left(\hat{x}_{i,\tau} - x_i\right)}{\lambda_i - \lambda_k + \left(\hat{\lambda}_{i,\tau} - \lambda_i\right)}$$

$$= \frac{x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)x_i + x_k'\left(\hat{\Sigma}_\tau - \Sigma\right)\left(\hat{x}_{i,\tau} - x_i\right)}{\left(\lambda_i - \lambda_k\right)} + o_p\left(1\right).$$

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

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

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

$$= \sum_{k\neq i}\sum_{j\neq i}\frac{\left(x_k'\otimes x_j'\right)V_\Sigma\left(x_j\otimes x_i\right)}{\left(\lambda_i - \lambda_k\right)\left(\lambda_i - \lambda_j\right)}x_k.$$

∎

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, $\dot{\Lambda}\left(\tau\right) = \frac{1}{\sqrt{T\tau(1-\tau)}}\left[\bar{S}\left(\tau\right)'\tilde{V}_{\Sigma,\tau}^{-1}\bar{S}\left(\tau\right)\right]^{1/2}$, where $\bar{S}\left(\tau\right) = S\left(\tau\right) - \frac{\lfloor T\tau\rfloor}{T}S\left(T\right)$. We also define $\ddot{\Lambda}\left(\tau\right) = \frac{1}{\sqrt{T\tau(1-\tau)}}\left[\bar{S}\left(\tau\right)'V_\Sigma^{-1}\bar{S}\left(\tau\right)\right]^{1/2}$; further, letting $B_{1i}\left(\tau\right)$ be a sequence of standard, inde-

pendent Brownian bridges for $i = 1, ..., n^2$, we define $M(\tau) = \left[\sum_{i=1}^{n^2} \frac{B_{1i}^2(\tau)}{\tau(1-\tau)}\right]^{1/2}$. The Darling-Erdos Theorem (see e.g. Corollary A.3.1 in Csorgo and Horvath, 1997, p. 366) states that $P\left[a_T \sup_{\frac{1}{T} \leq \tau \leq 1-\frac{1}{T}} M(\tau) \leq x + b_T\right] = e^{-2e^{-x}}$, where the norming constants a_T and b_T are defined in the Theorem. The proof of (14) is based on showing that

$$\left|\sup_{\frac{1}{T} \leq \tau \leq 1-\frac{1}{T}} \dot{\Lambda}(\tau) - \sup_{\frac{1}{T} \leq \tau \leq 1-\frac{1}{T}} M(\tau)\right| = o_p\left(\frac{1}{\sqrt{\ln \ln T}}\right). \tag{29}$$

We first note that, since, in view of Theorem 2, $\sup_{\lfloor T\tau \rfloor} \left\|\tilde{V}_{\Sigma,\tau} - \tilde{V}_{\Sigma}\right\| = o_p\left(\frac{1}{\sqrt{\ln \ln T}}\right)$, (29) can be rewritten as

$$\left|\sup_{\frac{1}{T} \leq \tau \leq 1-\frac{1}{T}} \ddot{\Lambda}(\tau) - \sup_{\frac{1}{T} \leq \tau \leq 1-\frac{1}{T}} M(\tau)\right| = o_p\left(\frac{1}{\sqrt{\ln \ln T}}\right). \tag{30}$$

In order to show (30), note first that (2) yields the (weak) result

$$\sup_{\frac{1}{T} \leq \tau \leq 1-\frac{1}{T}} \left|\ddot{\Lambda}(\tau) - M(\tau)\right| = o_p\left(\sqrt{\ln \ln T}\right). \tag{31}$$

Indeed, (2) entails

$$\sup_{u(T,\varepsilon) \leq \tau \leq \frac{1}{2}} \left[\lfloor T\tau \rfloor\right]^\delta \left|\ddot{\Lambda}(\tau) - M(\tau)\right| = o_p(1), \tag{32}$$

$$\sup_{\frac{1}{2} \leq \tau \leq 1-u(T,\varepsilon)} \left[\lfloor T(1-\tau) \rfloor\right]^\delta \left|\ddot{\Lambda}(\tau) - M(\tau)\right| = o_p(1), \tag{33}$$

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

$$\frac{1}{\sqrt{2 \ln \ln T}} \sup_{\frac{1}{T} \leq \tau \leq u(T,\varepsilon)} M(\tau) \xrightarrow{p} \sqrt{\varepsilon}, \tag{34}$$

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

Hence, from (31)

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

$$\frac{1}{\sqrt{2\ln\ln T}}\sup_{1-u(T,\varepsilon)\leq\tau\leq1-\frac{1}{T}}\ddot{\Lambda}(\tau)\xrightarrow{p}\sqrt{\varepsilon}.$$

Defining $\xi(T)$ and $\eta(T)$ as $\sup_{1\leq\lfloor T\tau\rfloor\leq T}M(\tau)=M[\xi(T)]$ and $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\ddot{\Lambda}(\tau)=\ddot{\Lambda}[\eta(T)]$, the relationships above entail $P[Tu(T,\varepsilon)\leq\xi(T),\eta(T)\leq1-u(T,\varepsilon)]=1$ as $T\to\infty$. Using (34) as an illustrative example, this follows from the fact that $\sup_{1\leq\lfloor T\tau\rfloor\leq Tu(T,\varepsilon)}M(\tau)$ is essentially $-\infty$ since, as $T\to\infty$ and $\varepsilon\to0$

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

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

$$\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}\left|\ddot{\Lambda}(\tau)-M(\tau)\right|=o_p\left(e^{-\delta\ln^{\varepsilon}T}\right),$$

and since $\left|\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}\ddot{\Lambda}(\tau)\right|-\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}M(\tau)\right|\leq\sup_{\frac{1}{T}\leq\tau\leq1-\frac{1}{T}}\left|\ddot{\Lambda}(\tau)-M(\tau)\right|$, (30) follows in view of $\sqrt{\ln\ln T}\,e^{-\delta\ln^{\varepsilon}T}\to0$. ∎

Proof of Corollary 2. Consider $T^{-1}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]'\tilde{V}_{\Sigma,\tau}^{-1}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]$. It holds that $\tilde{V}_{\Sigma,\tau}^{-1}=V_{\Sigma}^{-1}-V_{\Sigma}^{-1}\left(\tilde{V}_{\Sigma,\tau}-V_{\Sigma}\right)V_{\Sigma}^{-1}+o\left(\left\|\tilde{V}_{\Sigma,\tau}-V_{\Sigma}\right\|\right)$; thence

$$\frac{1}{T}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]'\tilde{V}_{\Sigma,\tau}^{-1}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]$$

$$=\frac{1}{T}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]'V_{\Sigma}^{-1}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]+$$

$$\frac{1}{T}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]'V_{\Sigma}^{-1}\left(\hat{V}_{\Sigma,\tau}-V_{\Sigma}\right)V_{\Sigma}^{-1}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]$$

$$=I_{\tau}+II_{\tau}.$$

Write $I_{\tau}=\sum_{i=1}^{n^2}w_{i,\tau}^2$, where $w_{i,\tau}$ is defined as the i-th element of $V_{\Sigma}^{-1/2}\left[S(\tau)-\frac{\lfloor T\tau\rfloor}{T}S(T)\right]$. The 8-th moment condition on y_t in Assumption 2(ii) entails, through Lemma 4, that a SIP holds whereby the $w_{i,\tau}$s can be approximated, uniformly in τ, by an i.i.d. normal sequence, say

$w_{i,\tau}^*$, with zero mean and unit variance, and approximation error of order $O_{a.s.}\left(T^{-\delta}\right)$. Thus

$$\frac{1}{\sqrt{2n}}\sum_{i=1}^{n^2}\left(w_{i,\tau}^2-1\right)=\frac{1}{\sqrt{2n}}\sum_{i=1}^{n^2}\left(w_{i,\tau}^{*2}-1\right)+\frac{1}{\sqrt{2n}}\sum_{i=1}^{n^2}\left(w_{i,\tau}^2-w_{i,\tau}^{*2}\right)=I_{a,\tau}+I_{b,\tau}.$$

Remark 2.1 in Aue et al. (2009) states that $\sup_{\lfloor T\tau\rfloor}I_{a,\tau}=O_p\left(1\right)$. From the SIP, $\sup_{\lfloor T\tau\rfloor}I_{b,\tau}=nO_p\left(T^{-\delta'}\right)$. We now turn to $II_\tau=\sum_{i=1}^{n^2}\sum_{j=1}^{n^2}w_{i,\tau}w_{j,\tau}a_{ij,\tau}$; $a_{ij,\tau}$ is element in position (i,j) of matrix $V_\Sigma^{-1/2}\left(\tilde{V}_{\Sigma,\tau}-V_\Sigma\right)V_\Sigma^{-1/2}$. Theorem 2 states that $a_{ij,\tau}=O_p\left(\frac{m}{\sqrt{T}}\right)+O_p\left(\frac{1}{m}\right)$, uniformly in τ. For some constant M, $\sup_{\lfloor T\tau\rfloor}II_\tau\leq\sup_{\lfloor T\tau\rfloor}\sum_{i=1}^{n^2}\sum_{j=1}^{n^2}|w_{i,\tau}w_{j,\tau}|\,|a_{ij,\tau}|\leq\max_{1\leq i,j\leq n}\sup_{\lfloor T\tau\rfloor}|a_{ij,\tau}|\sup_{\lfloor T\tau\rfloor}\sum_{i=1}^{n^2}\sum_{j=1}^{n^2}|w_{i,\tau}w_{j,\tau}|\leq\max_{1\leq i,j\leq n}\sup_{\lfloor T\tau\rfloor}|a_{ij,\tau}|\sup_{\lfloor T\tau\rfloor}\left(\sum_{i=1}^{n^2}|w_{i,\tau}|\right)^2\leq M\max_{1\leq i,j\leq n}\sup_{\lfloor T\tau\rfloor}|a_{ij,\tau}|\sup_{\lfloor T\tau\rfloor}\sum_{i=1}^{n^2}w_{i,\tau}^2=\left[O_p\left(\frac{m}{\sqrt{T}}\right)+O_p\left(\frac{1}{m}\right)\right]O_p\left(n^2\right)$. Thus, $\left(\sqrt{2}n\right)^{-1}\sup_{\lfloor T\tau\rfloor}II_\tau=O_p\left(\frac{nm}{\sqrt{T}}\right)+O_p\left(\frac{n}{m}\right)$. ∎

Proof of Theorem 4. In order to prove (17), we first show that, under $H_a^{(T)}$, $\sup_{1\leq\lfloor T\tau\rfloor\leq T}\left\|\hat{V}_{\Sigma,\tau}-V_\Sigma\right\|=o_p\left(1\right)$. For simplicity, we focus on Assumption 2(i) and on the full sample estimator $\hat{V}_\Sigma$. Consider $\hat{\Sigma}$; it holds that $vec\left(\hat{\Sigma}\right)=vec\left(\Sigma_t\right)+\left[\frac{T-k_{0,T}}{T}-I\left(t\geq k_{0,T}\right)\right]\Delta_T+o_p\left(1\right)$, where the $o_p\left(1\right)$ term comes from a LLN. Therefore

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

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

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

thus, $IV = O_p \left(\frac{k_{0,T}}{T} \| \Delta_T \|^2 \right)$, which is $o_p (1)$ under $H_a^{(T)}$.

After showing the consistency of $\hat{V}_\Sigma$, we turn to $\Lambda_T^2 (\tau)$. Set, for simplicity, $V_\Sigma = I_{n^2}$, so that $\Lambda_T^2 (\tau) = \frac{T}{\lfloor T\tau \rfloor \lfloor T(1-\tau) \rfloor} \tilde{S} (\tau)' \tilde{S} (\tau) + o_p (1)$. Under $H_a^{(T)}$

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

where $I (\cdot)$ is the indicator function. The sequence $\bar{w}_t$ is zero mean, and it satisfies the assumptions relevant for Theorem 1; thus, I follows the null distribution as $T \to \infty$. As far as the non-centrality parameter II is concerned,

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

$$(35)$$

with

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

Let Λ_0 denote a random variable with the same distribution as $\sup_{p \leq \lfloor T\tau \rfloor \leq T-p} \Lambda_T (\tau)$ under H_0. Under $H_a^{(T)}$

$$
P \left[\sup_{p \leq \lfloor T\tau \rfloor \leq T-p} \Lambda_T (\tau) > c_\alpha \right] = P \left[\Lambda_0 > c_\alpha - \| \Delta_T \| \sqrt{k_{0,T} \left(\frac{T - k_{0,T}}{T} \right)} \right];
$$

by definition, as $T \to \infty$, $c_\alpha = \sqrt{\ln \ln T} + o \left(\sqrt{\ln \ln T} \right)$. If (16) holds, $c_\alpha - \Delta_T \sqrt{k_{0,T} \left(\frac{T - k_{0,T}}{T} \right)} \to -\infty$ as $T \to \infty$, whence (17) follows. ■

T \ p	1	2	3	5	10	20
50	2.656	3.156	3.459	3.980	4.935	6.230
	2.919	3.398	3.705	4.211	5.191	6.444
	3.423	3.898	4.212	4.649	5.610	6.940
100	2.759	3.234	3.590	4.094	5.032	6.361
	3.035	3.476	3.832	4.332	5.271	6.575
	3.524	3.945	4.295	4.779	5.690	6.956
200	2.852	3.287	3.655	4.152	5.098	6.439
	3.128	3.499	3.901	4.376	5.311	6.642
	3.700	4.023	4.347	4.849	5.761	7.006
500	2.917	3.399	3.736	4.256	5.163	6.518
	3.187	3.664	3.958	4.495	5.374	6.718
	3.682	4.119	4.395	4.908	5.728	7.153
1000	2.969	3.441	3.795	4.274	5.246	6.529
	3.229	3.688	4.010	4.510	5.462	6.751
	3.711	4.119	4.444	4.965	5.906	7.157
2000	3.306	3.483	3.826	4.364	5.266	6.604
	3.287	3.714	4.057	4.588	5.460	6.823
	3.762	4.213	4.502	5.025	5.935	7.203

T \ p	30	40	50	75	100	200
50	7.273	8.084	8.846	10.440	11.797	15.941
	7.471	8.281	9.046	10.676	12.001	16.167
	7.861	8.698	9.481	11.094	12.391	16.554
100	7.369	8.198	8.950	10.542	11.898	16.019
	7.576	8.417	9.154	10.738	12.111	16.205
	7.961	8.853	9.566	11.082	12.483	16.611
200	7.445	8.272	9.028	10.616	11.934	16.099
	7.641	8.480	9.267	10.818	12.155	16.270
	8.071	8.920	9.594	11.241	12.532	16.644
500	7.500	8.367	9.089	10.700	12.045	16.180
	7.707	8.581	9.283	10.915	12.238	16.385
	8.078	8.944	9.637	11.282	12.618	16.751
1000	7.544	8.421	9.165	10.729	12.095	16.197
	7.746	8.604	9.353	10.956	12.277	16.383
	8.170	9.027	9.771	11.352	12.711	16.786
2000	7.595	8.436	9.186	10.783	12.117	16.251
	7.770	8.636	9.377	10.968	12.307	16.456
	8.188	9.009	9.750	11.374	12.690	16.830

T \ p	300	400	500	750	1000
50	19.103	21.787	24.127	29.167	33.420
	19.296	21.971	24.344	29.373	33.628
	19.702	22.401	24.701	29.735	34.009
100	19.199	21.893	24.231	29.280	33.507
	19.416	22.070	24.446	29.466	33.703
	19.828	22.458	24.856	29.845	34.119
200	19.288	21.972	24.292	29.351	33.575
	19.486	22.152	24.498	29.539	33.779
	19.855	22.581	24.870	29.915	34.144
500	19.357	22.026	24.377	29.400	33.636
	19.565	22.232	24.538	29.583	33.832
	19.945	22.589	24.922	29.899	34.225
1000	19.382	22.081	24.415	29.461	33.685
	19.587	22.273	24.604	29.636	33.868
	19.918	22.617	24.954	30.018	34.233
2000	19.407	22.125	24.463	29.500	33.724
	19.607	22.291	24.643	29.666	33.906
	20.021	22.701	25.021	30.008	34.289

Table 1. Approximated critical values c'_α. We report 90%, 95% and 99% quantiles, computed using (18). In the computations, 5000 replications are used; critical values could be calculated for all combinations of p (number of hypotheses under the null), and T (sample size). The code is available upon request.

n	T	(ρ,ϑ)	$(0,0)$	$(0.5,0)$	$(0,0.5)$	$(0,-0.5)$	$(0.5,0.5)$
3	50		0.094	0.125	0.118	0.114	0.118
	100		0.060	0.075	0.065	0.082	0.075
	200		0.055	0.063	0.059	0.064	0.058
	500		0.055	0.062	0.047	0.045	0.064
4	50		0.090	0.122	0.108	0.108	0.121
	100		0.057	0.076	0.070	0.073	0.077
	200		0.048	0.070	0.060	0.049	0.105
	500		0.046	0.055	0.048	0.059	0.059
5	50		0.087	0.109	0.103	0.098	0.105
	100		0.055	0.086	0.069	0.085	0.078
	200		0.051	0.063	0.058	0.055	0.064
	500		0.057	0.060	0.050	0.054	0.066
6	50		0.085	0.103	0.098	0.110	0.099
	100		0.062	0.078	0.075	0.075	0.077
	200		0.052	0.067	0.069	0.050	0.069
	500		0.062	0.065	0.053	0.050	0.069
7	50		0.082	0.109	0.108	0.105	0.101
	100		0.061	0.070	0.061	0.069	0.069
	200		0.056	0.067	0.060	0.055	0.065
	500		0.057	0.067	0.057	0.047	0.065
10	50		0.104	0.121	0.112	0.103	0.115
	100		0.070	0.076	0.066	0.059	0.082
	200		0.046	0.059	0.047	0.041	0.066
	500		0.043	0.052	0.050	0.046	0.051
15	50		0.090	0.113	0.093	0.099	0.101
	100		0.067	0.071	0.064	0.067	0.075
	200		0.045	0.066	0.064	0.046	0.076
	500		0.045	0.066	0.052	0.053	0.061
20	50		0.083	0.104	0.087	0.100	0.108
	100		0.066	0.068	0.060	0.071	0.076
	200		0.044	0.070	0.060	0.046	0.072
	500		0.042	0.060	0.046	0.057	0.056

Table 2. Empirical rejection frequencies for the null of no changes in the largest eigenvalue of Σ. Data are generated according to (19). The empirical sizes reported here have confidence interval $[0.04, 0.06]$.

Table showing power results. The table is rotated; reconstructed below.

n	T	(ρ, ϑ)	$\Delta = \sqrt{\frac{\ln\ln(T)}{T^{2/3}}}$					$\Delta = \sqrt{\frac{\ln\ln(T)}{T^{1/2}}}$				
			(0, 0)	(0.5, 0)	(0, 0.5)	(0, -0.5)	(0.5, 0.5)	(0, 0)	(0.5, 0)	(0, 0.5)	(0, -0.5)	(0.5, 0.5)
3	50		0.058	0.072	0.065	0.070	0.089	0.075	0.086	0.073	0.081	0.089
	100		0.107	0.093	0.094	0.095	0.090	0.191	0.132	0.149	0.150	0.118
	200		0.200	0.149	0.159	0.154	0.120	0.428	0.275	0.338	0.327	0.217
	500		0.363	0.224	0.291	0.280	0.175	0.847	0.574	0.709	0.704	0.421
	1000		0.578	0.360	0.438	0.449	0.280	0.978	0.843	0.923	0.926	0.695
4	50		0.056	0.082	0.067	0.068	0.081	0.077	0.087	0.086	0.080	0.087
	100		0.116	0.094	0.103	0.100	0.082	0.202	0.140	0.167	0.168	0.120
	200		0.213	0.141	0.161	0.182	0.127	0.473	0.262	0.344	0.367	0.206
	500		0.407	0.262	0.314	0.306	0.203	0.875	0.609	0.725	0.746	0.463
	1000		0.576	0.372	0.460	0.444	0.277	0.986	0.848	0.928	0.943	0.699
5	50		0.062	0.082	0.068	0.052	0.081	0.078	0.087	0.082	0.071	0.091
	100		0.119	0.094	0.106	0.102	0.119	0.225	0.149	0.166	0.169	0.123
	200		0.227	0.168	0.180	0.170	0.140	0.538	0.329	0.398	0.367	0.244
	500		0.394	0.260	0.316	0.302	0.212	0.884	0.611	0.748	0.751	0.478
	1000		0.619	0.394	0.485	0.483	0.302	0.992	0.883	0.957	0.951	0.753
6	50		0.066	0.087	0.071	0.069	0.087	0.081	0.097	0.087	0.082	0.100
	100		0.127	0.117	0.118	0.107	0.114	0.231	0.167	0.182	0.187	0.154
	200		0.231	0.157	0.188	0.191	0.133	0.512	0.297	0.384	0.399	0.237
	500		0.432	0.277	0.331	0.329	0.222	0.895	0.636	0.775	0.779	0.500
	1000		0.642	0.406	0.497	0.499	0.315	0.993	0.878	0.949	0.954	0.745
7	50		0.057	0.080	0.073	0.070	0.081	0.078	0.089	0.087	0.083	0.091
	100		0.141	0.114	0.108	0.127	0.108	0.243	0.172	0.180	0.205	0.151
	200		0.249	0.157	0.179	0.193	0.136	0.548	0.316	0.396	0.419	0.240
	500		0.457	0.280	0.335	0.352	0.219	0.898	0.633	0.775	0.792	0.493
	1000		0.641	0.410	0.507	0.511	0.320	0.993	0.890	0.960	0.959	0.772
10	50		0.072	0.088	0.083	0.068	0.087	0.084	0.102	0.098	0.084	0.105
	100		0.144	0.117	0.129	0.123	0.106	0.264	0.162	0.203	0.209	0.149
	200		0.283	0.176	0.206	0.215	0.146	0.604	0.338	0.450	0.452	0.269
	500		0.469	0.270	0.359	0.371	0.215	0.929	0.678	0.832	0.821	0.540
	1000		0.675	0.423	0.522	0.555	0.330	0.997	0.909	0.970	0.979	0.789
15	50		0.088	0.089	0.077	0.077	0.101	0.111	0.109	0.106	0.101	0.115
	100		0.169	0.121	0.130	0.137	0.119	0.308	0.177	0.222	0.220	0.164
	200		0.314	0.199	0.241	0.234	0.190	0.679	0.390	0.500	0.518	0.328
	500		0.559	0.331	0.415	0.428	0.257	0.955	0.737	0.870	0.873	0.584
	1000		0.699	0.442	0.559	0.559	0.344	0.997	0.929	0.973	0.976	0.838
20	50		0.095	0.086	0.087	0.094	0.091	0.122	0.105	0.111	0.118	0.114
	100		0.181	0.139	0.156	0.165	0.137	0.337	0.204	0.256	0.261	0.191
	200		0.374	0.211	0.278	0.272	0.184	0.745	0.387	0.557	0.551	0.319
	500		0.569	0.335	0.419	0.452	0.278	0.971	0.748	0.897	0.897	0.611
	1000		0.745	0.485	0.605	0.603	0.387	0.999	0.947	0.988	0.988	0.866

Table 3a. Power of the test for the null of no changes in the largest eigenvalue of Σ. Data are generated according to (19) and under the alternative hypothesis specified in (22).

n	T		$k = 2 \times [\ln(T)]^{1+\ln\ln(T)}$				
		(ρ, ϑ)	$(0,0)$	$(0.5, 0)$	$(0, 0.5)$	$(0, -0.5)$	$(0.5, 0.5)$
3	50		0.018	0.027	0.022	0.029	0.042
	100		0.041	0.040	0.033	0.039	0.044
	200		0.150	0.057	0.081	0.081	0.049
	500		0.671	0.160	0.345	0.338	0.079
	1000		0.881	0.233	0.527	0.530	0.091
4	50		0.020	0.028	0.025	0.030	0.040
	100		0.047	0.033	0.042	0.044	0.037
	200		0.177	0.059	0.080	0.102	0.054
	500		0.716	0.179	0.350	0.367	0.082
	1000		0.904	0.248	0.551	0.563	0.080
5	50		0.016	0.031	0.025	0.020	0.039
	100		0.041	0.039	0.039	0.041	0.041
	200		0.169	0.061	0.087	0.074	0.054
	500		0.722	0.160	0.358	0.343	0.069
	1000		0.910	0.258	0.574	0.560	0.078
6	50		0.019	0.040	0.026	0.028	0.049
	100		0.047	0.045	0.045	0.044	0.048
	200		0.169	0.061	0.074	0.081	0.054
	500		0.711	0.174	0.382	0.366	0.085
	1000		0.913	0.211	0.540	0.552	0.089
7	50		0.015	0.036	0.020	0.024	0.039
	100		0.045	0.037	0.041	0.037	0.042
	200		0.187	0.058	0.091	0.083	0.050
	500		0.745	0.159	0.386	0.374	0.079
	1000		0.919	0.249	0.588	0.555	0.089
10	50		0.020	0.035	0.031	0.025	0.038
	100		0.042	0.049	0.039	0.030	0.050
	200		0.205	0.056	0.094	0.092	0.049
	500		0.762	0.140	0.376	0.380	0.075
	1000		0.929	0.236	0.581	0.578	0.080
15	50		0.019	0.034	0.024	0.023	0.038
	100		0.050	0.047	0.036	0.045	0.042
	200		0.235	0.067	0.094	0.090	0.062
	500		0.792	0.139	0.392	0.389	0.067
	1000		0.932	0.203	0.588	0.582	0.088

Table 3b. Power of the test for the null of no changes in the largest eigenvalue of Σ. Data are generated according to (19), under the alternative specified in (22).

n	T	$\Delta = 0$	$\Delta = \sqrt{\frac{\ln\ln(T)}{T^{4/5}}}$	$\Delta = \sqrt{\frac{\ln\ln(T)}{T^{2/3}}}$	$\Delta = \sqrt{\frac{\ln\ln(T)}{T^{1/2}}}$	$k = k^*$
3	50	0.075	0.073	0.106	0.175	0.052
	100	0.070	0.126	0.207	0.434	0.066
	200	0.060	0.189	0.365	0.825	0.094
	500	0.057	0.271	0.666	0.996	0.467
4	50	0.051	0.027	0.041	0.083	0.030
	100	0.046	0.096	0.171	0.408	0.051
	200	0.066	0.187	0.404	0.864	0.145
	500	0.066	0.150	0.755	1.000	0.778
5	50	0.018	0.010	0.019	0.046	0.023
	100	0.030	0.072	0.156	0.388	0.077
	200	0.051	0.162	0.401	0.896	0.251
	500	0.057	0.326	0.789	1.000	0.965
6	50	0.005	0.000	0.000	0.002	0.005
	100	0.016	0.040	0.090	0.290	0.108
	200	0.027	0.159	0.381	0.887	0.453
	500	0.051	0.343	0.818	1.000	1.000
7	50	0.000	0.000	0.000	0.000	0.000
	100	0.004	0.022	0.046	0.186	0.149
	200	0.018	0.108	0.318	0.873	0.721
	500	0.050	0.319	0.832	1.000	1.000

Table 4. Empirical rejection frequencies and power for the null of no change in Σ. Data are generated as *i.i.d.* normal with $E\left(y_t y_t'\right) = I_n$ under H_0, and using equation (19) and (22) under H_a.

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

Table 5. 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 effectes at 10%, 5% and 1% levels respectively.

(T, p)	level	10%	5%	1%
	cv_1	2.8044	3.0716	3.5758
$(163, 1)$	cv_3	3.1987	3.4222	3.9344
	cv_{18}	3.7843	3.9690	4.1273
	cv_1	2.9471	3.1891	3.6379
$(732, 1)$	cv_3	3.3266	3.5153	3.8789
	cv_{18}	3.7779	4.0051	4.5603
	cv_1	6.1897	6.4075	6.7747
$(163, 18)$	cv_3	6.5068	6.6486	7.0053
	cv_1	6.3066	6.5202	6.9209
$(732, 18)$	cv_3	6.6288	6.8040	7.1498

Table 6. 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 Σ one at a time. Panels where $(T, p) = (163, 1)$ and $(732, 1)$ contain critical values for tests with 18 hypotheses under the null, and thus are designed for tests for the stability of one eigenvector.

i	$H_0: \Sigma_{ii}$ constant			$H_0: \lambda_i$ constant			$H_0: x_i$ constant	
	monthly	weekly		monthly	weekly		monthly	weekly
$1m$	2.6989	2.8136						
$3m$	2.7656	3.7004	λ_1	1.6921	3.7156** [1st week, 12/2007]	x_1	4.2950	7.1268** [3rd week, 03/2008]
$6m$	2.7394	3.1770						
$9m$	2.3924	2.3132	λ_2	2.5513	2.8518	x_2	4.6617	6.7893* [last week, 06/1999]
$12m$	1.5350	3.1266						
$15m$	1.4991	2.8294	λ_3	3.4328** [01/2008]	2.7495	x_3	5.0185	7.0000** [last week, 03/2008]
$18m$	1.6467	2.9063						
$21m$	1.8065	3.0928						
$24m$	1.9827	3.1274						
$30m$	2.0718	3.3169						
$3y$	2.0815	3.5926						
$4y$	1.9314	4.0180** [3rd week, 09/2007]						
$5y$	1.8964	4.1170** [1st week, 12/2007]						
$6y$	1.8369	4.2779** [1st week, 12/2007]						
$7y$	1.7677	4.2595** [1st week, 12/2007]						
$8y$	1.9601	4.3342** [1st week, 12/2007]						
$9y$	2.1046	4.3549** [1st week, 12/2007]						
$10y$	2.1967	4.4386** [last week, 08/2008]						

Table 7. 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.

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

Table 8. Proportion of the total variance explained by principal components (λ_1, λ_2 and λ_3 refer to the level, slope and curvature respectively) for each subsample. The samples are split based on the results in Table 7. When considering monthly data, the sample was split at January 2008; when using weekly data, at the first week of December 2007.
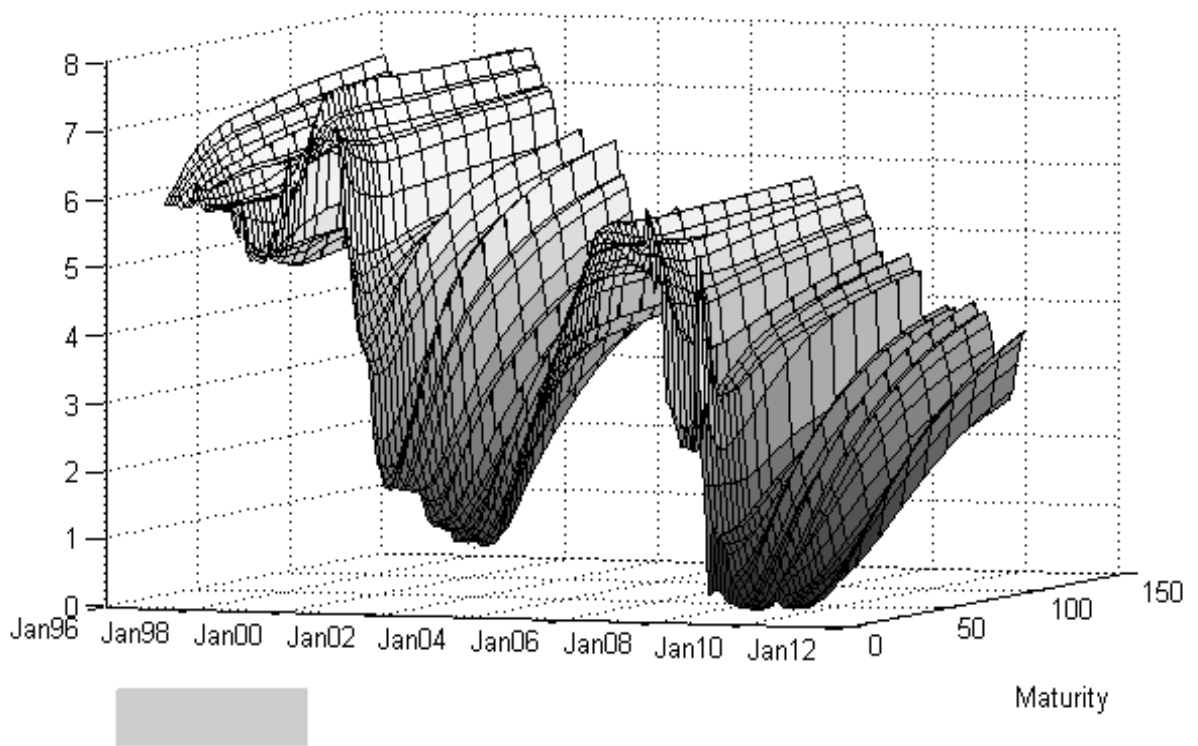
Figure 1. Term structure of the US interest rates. Maturities correspond to 1m, 3m, 6m, 9m, 12m, 15m, 18m, 21m, 24m, 30m, 3y, 4y, 5y, 6y, 7y, 8y, 9y, 10y over the period April 1997-November 2010.
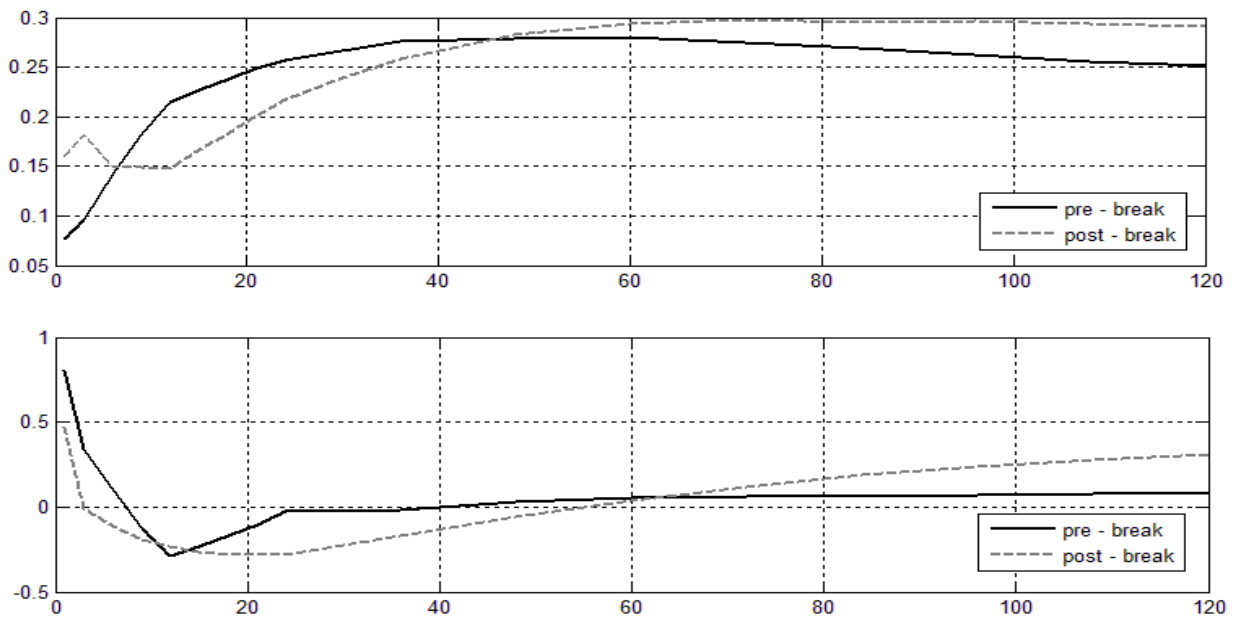

Figure 2: Loadings (eigenvectors) of the first and third principal components, before and after mid-April 2008.