The Econometrics Journal

Econometrics Journal (2021), volume **24**, pp. 1–22.
doi: 10.1093/ectj/utaa008

Testing identification via heteroskedasticity in structural vector autoregressive models

HELMUT LÜTKEPOHL[*], MIKA MEITZ[†], ALEKSEI NETŠUNAJEV[‡] AND PENTTI SAIKKONEN[§]

[*]*DIW Berlin and School of Business and Economics, Freie Universität Berlin, Mohrenstr. 58, 10117 Berlin, Germany.*
Email: hluetkepohl@diw.de

[†]*Discipline of Economics, University of Helsinki, P. O. Box 17, FI-00014 University of Helsinki, Finland.*
Email: mika.meitz@helsinki.fi

[‡]*School of Business and Governance, Department of Economics and Finance, Tallinn University of Technology, Akademia tee 3, 12618 Tallinn, Estonia.*
Email: aleksei.netsunajev@ttu.ee

[§]*Department of Mathematics and Statistics, University of Helsinki, P. O. Box 68, FI-00014 University of Helsinki, Finland.*
Email: pentti.saikkonen@helsinki.fi

First version received: 1 October 2019; final version accepted: 3 March 2020.

Summary: Tests for identification through heteroskedasticity in structural vector autoregressive analysis are developed for models with two volatility states where the time point of volatility change is known. The tests are Wald-type tests for which only the unrestricted model, including the covariance matrices of the two volatility states, has to be estimated. The residuals of the model are assumed to be from the class of elliptical distributions, which includes Gaussian models. The asymptotic null distributions of the test statistics are derived, and simulations are used to explore their small-sample properties. Two empirical examples illustrate the usefulness of the tests in applied work.

Keywords: *Heteroskedasticity*, *structural identification*, *vector autoregressive process*.

JEL codes: *C32*.

1. INTRODUCTION

Identification by heteroskedasticity of structural shocks has become a standard tool in structural vector autoregressive (VAR) analysis (see, e.g., Kilian and Lütkepohl, 2017, Chapter 14). Heteroskedasticity can complement identifying restrictions based on economic theory or subject matter knowledge. The underlying idea is that if the variance of the structural shocks changes during the sample period and there is heterogeneity in the variance changes of different shocks, this feature can be used to distinguish (identify) the shocks. The objective of the present study is to

present a formal statistical test for the required variance heterogeneity and hence the identification of the shocks.

Of course, identifying structural shocks purely through their statistical properties implies that a further step is necessary to associate the identified shocks with economic shocks of interest. Therefore, identification through heteroskedasticity is often complemented by conventional identifying restrictions on, for example, the impact effects of the shocks. In particular, if competing identifying restrictions are available, which are just-identifying in a conventional setting, then identifying information from heteroskedasticity can be used as overidentifying restrictions, which opens up the possibility of formally testing identifying restrictions that are otherwise not testable. Such tests are not considered in the present paper but are discussed in, for example, Lanne and Lütkepohl (2008), Netšunajev (2013), and Lütkepohl and Netšunajev (2017).

A main advantage of identification via heteroskedasticity is that the data are in principle informative on the conditions for identification. Thus, identification can be investigated by statistical tests. The problem in developing such tests is that the model is typically not identified under the null hypothesis of no identification, which complicates the derivation of the asymptotic distributions of standard tests. Some authors still use standard Wald and likelihood ratio (LR) tests for identification through heteroskedasticity and approximate the distribution under the null hypothesis by the usual χ^2 distributions. Examples are Lanne et al. (2010), Herwartz and Lütkepohl (2014), Lütkepohl and Velinov (2016), Velinov and Chen (2015), Netšunajev (2013), and Lütkepohl and Netšunajev (2014). However, so far the asymptotic distributions of these tests have not been derived formally, and it is unlikely that the assumed χ^2 distributions provide precise approximations to the true asymptotic distributions of the test statistics. Alternatively, some authors have proposed Bayesian methods for assessing identification in this context (e.g., Woźniak and Droumaguet, 2015, and Lütkepohl and Woźniak, 2020).

In the following, we will develop formal frequentist tests that can help in assessing identification through heteroskedasticity for the special case of stable VAR models with two volatility regimes of the residuals. Such simple models for the change in volatility have been considered by, for example, Rigobon (2003), Lanne and Lütkepohl (2008, 2014), and Lütkepohl and Schlaak (2018). (The actual models used for empirical analysis are more sophisticated in some of these articles, where more volatility states and alternative volatility models are considered as well. We focus on the simpler case to make the problem tractable and leave more general models for future research.) For developing our tests, we assume that the distribution of the residuals is elliptically symmetric, which covers the case of Gaussian VAR processes but also models where the residuals have t distributions or mixtures of normal distributions. We develop Wald-type tests for which we can derive the asymptotic distribution under the null hypothesis of no identification. Our results shed further doubts on the previously assumed test distributions for related statistics. Of course, if the test indicates identification through heteroskedasticity, the identifying information may still be limited. In other words, there may still be weak identification. This issue in the context of identification through heteroskedasticity has been discussed by Lewis (2018b), and more general results on weak identification have been given by Andrews (2018). Such procedures may be used in addition to or as an alternative to our approach.

Our tests may indicate only that there is some identifying information through heteroskedasticity but may not suggest that the structural model is fully identified. We discuss a sequential testing procedure that can be helpful in this context. Our tests can be used at different stages of such a procedure. We show by simulation that the asymptotic theory is a good guide for small-sample performance of the individual tests, if the sample size is sufficiently large. Finally, we present examples that illustrate the usefulness of our tests for applied work.

The remainder of this paper is structured as follows. The model is set up in the following section. Section 3 presents the tests for identification and their asymptotic properties. Section 4 considers the small-sample properties of the tests. Two empirical examples are discussed in Section 5. The final section concludes. The proofs of the asymptotic results for the test statistics are provided in Appendix A.

2. THE MODEL

Consider a stable K-dimensional reduced-form VAR(p) model,

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \tag{2.1}$$

where ν is an intercept term, and A_j ($j = 1, \ldots, p$) are ($K \times K$) VAR slope coefficient matrices satisfying the usual stability condition,

$$\det\left(I_K - A_1 z - \cdots - A_p z^p\right) \neq 0 \quad \text{for} \quad |z| \leq 1.$$

The error process u_t is independent white noise, with u_t independent of u_s for $s \neq t$, zero mean, $\mathbb{E}(u_t) = 0$, and (positive definite) covariance matrices,

$$\mathbb{E}(u_t u_t') = \begin{cases} \Sigma_1 & \text{for} \quad t \in \mathcal{T}_1 = \{1, \ldots, T_1\}, \\ \Sigma_2 & \text{for} \quad t \in \mathcal{T}_2 = \{T_1 + 1, \ldots, T\}, \end{cases} \tag{2.2}$$

where T signifies the sample size. Thus, the errors of the model are assumed to be heteroskedastic so that the covariance matrix changes from Σ_1 to Σ_2 at time $T_1 + 1$, which we assume to be known. Moreover, we assume that for some fixed fraction $\tau \in (0, 1)$, T_1 is the integer part of τT; that is, $T_1 = [\tau T]$, so that the sample size for both volatility regimes goes to infinity as $T \to \infty$.

The setup in (2.2) is used here for convenience. For our asymptotic analysis, it is in fact sufficient that $\mathcal{T}_1$ contains a fraction τ of sample periods, whereas $\mathcal{T}_2$ contains the remaining fraction of $1 - \tau$ periods in the sample. In other words, $\mathcal{T}_1$ and $\mathcal{T}_2$ do not have to contain consecutive parts of the sample, but the size of both sets has to go to infinity with the sample size in such a way that both Σ_1 and Σ_2 can be estimated consistently.

We consider the case in which the error term u_t has an elliptically symmetric distribution or briefly an elliptical distribution possessing a density $\left(\sqrt{\det \Sigma_m}\right)^{-1} g(u_t' \Sigma_m^{-1} u_t)$, where Σ_m is a symmetric positive definite matrix, $g(\cdot)$ is a positive function such that the density integrates to one, and the fourth moments of the distribution exist (see, e.g., Anderson, 2003, Section 2.7, for further discussion of elliptical distributions). Note that the elliptical distributions are such that all components of u_t have the same kurtosis parameter. More precisely, denoting the k^{th} diagonal element of Σ_m by σ_{km}^2, the kurtosis parameter $[\mathbb{E}(u_{kt}^4)/3\sigma_{km}^4] - 1$ is the same for $k = 1, \ldots, K$ (see also Anderson, 2003, p. 54, Equation 36). We explicitly allow for the possibility that the kurtosis parameter may be different for the different volatility regimes and define

$$[\mathbb{E}(u_{kt}^4)/3\sigma_{km}^4] - 1 = \kappa_m \quad \text{for} \quad t \in \mathcal{T}_m, \quad m = 1, 2.$$

Notice, however, that the case of Gaussian residuals is obtained as a special case by choosing the kurtosis parameter equal to zero. Thus, even if the variance changes across the sample, we may have $\kappa_1 = \kappa_2$ if, for example, the sample is Gaussian. Other distribution families covered by our assumptions are t distributions and mixtures of normal distributions. We need the elliptical distributions to apply limiting results from Anderson (2003) in our derivation of the test for identification in Section 3.

The covariance matrices in (2.2) can be decomposed as follows:

$$\Sigma_1 = BB', \quad \Sigma_2 = B\Lambda B', \tag{2.3}$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_K)$ is a $(K \times K)$ diagonal matrix with positive diagonal elements, and B is a nonsingular $(K \times K)$ matrix (see Lütkepohl, 2013). A standard assumption in the related structural VAR (SVAR) literature is that only the volatility of the shocks changes, whereas the responses of the variables remain time invariant. This is accomplished by obtaining the structural shocks from the reduced-form errors as $\varepsilon_t = B^{-1}u_t$, such that B is the matrix of impact effects of the shocks, and the covariance matrices of the structural errors are given by

$$\mathbb{E}(\varepsilon_t \varepsilon_t') = \begin{cases} I_K & \text{for} \quad t \in \mathcal{T}_1, \\ \Lambda & \text{for} \quad t \in \mathcal{T}_2. \end{cases}$$

Thus, the structural errors are instantaneously uncorrelated in both volatility regimes. Replacing the reduced-form errors u_t in (2.1) by the structural errors $B\varepsilon_t$ yields the SVAR(p) model

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + B\varepsilon_t. \tag{2.4}$$

For the statistical results to be obtained later, we assume that the structural errors ε_t or, equivalently, the reduced-form errors u_t are temporally independent.

It is well known (see, e.g., Theorem A9.9 and its proof in Muirhead, 1982) that the diagonal elements of the matrix Λ in (2.3) are the eigenvalues of the matrix $\Sigma_2\Sigma_1^{-1}$ so that they satisfy the (generalised) eigenvalue equations

$$\det\left(\Sigma_1^{-1} - \lambda_k \Sigma_2^{-1}\right) = 0, \quad k = 1, \ldots, K, \tag{2.5}$$

whereas the columns of the matrix $B = [b_1 \colon \cdots \colon b_K]$ are the corresponding (generalised) eigenvectors that satisfy

$$\left(\Sigma_1^{-1} - \lambda_k \Sigma_2^{-1}\right) b_k = 0, \quad k = 1, \ldots, K. \tag{2.6}$$

Furthermore, if the eigenvalues $\lambda_1, \ldots, \lambda_K$ are distinct, the matrix B is unique apart from permutations and sign reversals of its columns (see the aforementioned theorem of Muirhead, 1982, or Lanne et al., 2010, Proposition 1). In what follows, we assume (without loss of generality) that the eigenvalues $\lambda_1, \ldots, \lambda_K$ are ordered from largest to smallest so that $\lambda_1 \geq \cdots \geq \lambda_K > 0$ holds. If the matrix B is not unique, we have an identification problem in the SVAR(p) model (2.4). Testing for a possible lack of identification is therefore of interest and will be discussed in the next section.

Clearly, our model is special in that it assumes two volatility regimes only. In practice, there may be more than two volatility regimes (see, e.g., Rigobon and Sack, 2003, and Lanne and Lütkepohl, 2008). In that case, the identification conditions become more elaborate and more difficult to test formally (e.g., Kilian and Lütkepohl, 2017, Section 14.3.1). We leave this case for future research. Moreover, more general volatility models have been considered in the literature. For example, Lanne et al. (2010) and Herwartz and Lütkepohl (2014) used a Markov-switching process to model endogenously changing volatility regimes. Unfortunately, we do not know whether the tests developed in the following can be extended to that case even in situations in which very similar identification conditions are of interest. Again, we leave extensions to such models for future research.

3. A TEST PROCEDURE FOR IDENTIFICATION OF B

3.1. The testing problem

Given that the diagonal elements of the matrix Λ are ordered from largest to smallest, uniqueness of the matrix B obtains if $\lambda_1 > \cdots > \lambda_K$ and the possibility of sign reversals in B is eliminated. One possibility to fix the column signs to be used in this study is to require that the first nonzero element of each column of B be positive. In order to test for lack of identification, we consider the pair of hypotheses

$$\mathbb{H}_0 : \lambda_{s+1} = \lambda_{s+2} = \cdots = \lambda_{s+r} \ (= \lambda_0) \quad \text{versus} \quad \mathbb{H}_1 : \neg\mathbb{H}_0$$

for $s \in \{0, \ldots, K-2\}$ and $r \in \{2, \ldots, K-s\}$. Thus, under the null hypothesis, r consecutive eigenvalues of Λ are equal to a value λ_0, implying lack of identification. The remaining eigenvalues, $\lambda_1, \ldots, \lambda_s, \lambda_{s+r+1}, \ldots, \lambda_K$, may have multiplicities larger than one but have to be different from λ_0, the common value under $\mathbb{H}_0$.

Let $y_{-p+1}, \ldots, y_0, y_1, \ldots, y_T$ be the available data. The reduced-form Gaussian log-likelihood function (apart from a constant, and conditioning on the first p observations $y_{-p+1}, \ldots, y_0$) is given by

$$\begin{aligned} l(\boldsymbol{\vartheta}, \boldsymbol{\sigma}) = &-\frac{1}{2}\sum_{t=1}^{T_1} \log\det(\Sigma_1) - \frac{1}{2}\sum_{t=1}^{T_1} u_t(\boldsymbol{\vartheta})'\Sigma_1^{-1}u_t(\boldsymbol{\vartheta}) \\ &-\frac{1}{2}\sum_{t=T_1+1}^{T} \log\det(\Sigma_2) - \frac{1}{2}\sum_{t=T_1+1}^{T} u_t(\boldsymbol{\vartheta})'\Sigma_2^{-1}u_t(\boldsymbol{\vartheta}), \end{aligned}$$

where $\boldsymbol{\vartheta} = \mathrm{vec}(\nu, A_1, \ldots, A_p)$, $u_t(\boldsymbol{\vartheta})$ signifies u_t in expression (2.1) when these quantities are interpreted as functions of the underlying parameters and $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)$ with $\boldsymbol{\sigma}_m = \mathrm{vech}(\Sigma_m)$ $(m = 1, 2)$. Here, 'vec' denotes the usual column-stacking operator, and 'vech' denotes the operator that stacks the columns of a square matrix from the main diagonal downward. If the data generating process (DGP) is Gaussian, maximising $l(\boldsymbol{\vartheta}, \boldsymbol{\sigma})$ with respect to the parameters gives the maximum likelihood (ML) estimators, and if the true distribution is not Gaussian but of a more general elliptical form, the resulting estimators are quasi-ML estimators.

Instead of ML estimation, one may use a feasible generalised least squares (GLS) procedure. In that case, (2.1) is estimated with equationwise ordinary least squares in a first step. The residuals $\hat{u}_t$ obtained in that way are then used for estimating the covariance matrices as $\widehat{\Sigma}_m = T_m^{-1}\sum_{t\in\mathcal{T}_m}\hat{u}_t\hat{u}_t'$, $m = 1, 2$, where $T_2 = T - T_1$. In a further step, the GLS estimator,

$$\tilde{\boldsymbol{\vartheta}} = \left(\sum_{t=1}^{T} Z_{t-1}Z_{t-1}' \otimes \widehat{\Sigma}_t^{-1}\right)^{-1}\left(\sum_{t=1}^{T}(Z_{t-1}\otimes\widehat{\Sigma}_t^{-1})y_t\right),$$

is computed, where $Z_{t-1} = (1, y_{t-1}', \ldots, y_{t-p}')'$ and $\widehat{\Sigma}_t = \widehat{\Sigma}_m$ for $t \in \mathcal{T}_m$ $(m = 1, 2)$. If the VAR process is stable, these estimators have standard asymptotic properties and can be used accordingly (see Lütkepohl, 2005, Chapter 17). Then, the GLS residuals can be used to estimate the covariance matrices Σ_1 and Σ_2. In what follows, $\tilde{\boldsymbol{\vartheta}}$ can be any estimator of $\boldsymbol{\vartheta}$ such that $\tilde{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta} = O_p(T^{-1/2})$.

Then, one readily finds that $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$ are asymptotically equivalent to their (unfeasible) counterparts based on the reduced-form errors or, specifically,

$$\tilde{\Sigma}_1 = \frac{1}{T_1}\sum_{t=1}^{T_1}\tilde{u}_t\tilde{u}_t' = \frac{1}{T_1}\sum_{t=1}^{T_1}u_t u_t' + o_p(T^{-1/2}), \tag{3.1}$$

$$\tilde{\Sigma}_2 = \frac{1}{T-T_1}\sum_{t=T_1+1}^{T}\tilde{u}_t\tilde{u}_t' = \frac{1}{T-T_1}\sum_{t=T_1+1}^{T}u_t u_t' + o_p(T^{-1/2}), \tag{3.2}$$

where $\tilde{u}_t$ signifies the GLS residuals described above; that is, $\tilde{u}_t = y_t - \tilde{\nu} - \tilde{A}_1 y_{t-1} - \cdots - \tilde{A}_p y_{t-p}$ (cf. Proposition 3.2 in Lütkepohl, 2005). Replacing the theoretical covariance matrices Σ_1 and Σ_2 in equations (2.5) and (2.6) with the estimators $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$, we obtain the vector of eigenvalues $\tilde{\boldsymbol{\lambda}} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_K)$ and the matrix of eigenvectors $\tilde{B} = [\tilde{b}_1 : \cdots : \tilde{b}_K]$. Similarly to their theoretical counterparts, the estimated eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_K$ are ordered from largest to smallest, and, because they are distinct with probability one, we have $\tilde{\lambda}_1 > \cdots > \tilde{\lambda}_K > 0$ almost surely. Eliminating the possibility of sign reversals in $\tilde{B}$ in the same way as in B, we therefore have a one-to-one continuous correspondence between the estimators $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$ and the elements of the matrix $\tilde{B}$ and the vector $\tilde{\boldsymbol{\lambda}}$. Thus, $\tilde{B}$ and $\tilde{\boldsymbol{\lambda}}$ can be viewed as unrestricted estimators of B and $\boldsymbol{\lambda}$.

Deriving the asymptotic properties of estimated eigenvalues is known to be a complicated problem when the theoretical eigenvalues are not distinct, which is the case under our null hypothesis. In the context of principal component analysis, where the population eigenvalues satisfy equation (2.5) with $\Sigma_1 = I_K$, and with independent observations, a complete solution to this problem was provided by Anderson (1963) (see also Anderson, 2003, Section 11.7.3, and Muirhead, 1982, Sections 9.5 and 9.6), whereas Anderson (2003, Section 13.6.3) treated the case of a general Σ_1 (again with independent observations). In what follows, we adopt Anderson's approach to our problem.

For setting up our test statistics, we also need consistent estimates of the kurtosis parameters. One possible estimator is discussed in Schott (2001, p. 33),

$$\tilde{\kappa}_m = \frac{1}{3K}\sum_{k=1}^{K}\frac{z_k^m}{w_k^m} - 1, \quad m = 1, 2,$$

where

$$z_k^m = \frac{\sum_{t\in\mathcal{T}_m}(\tilde{u}_{kt} - \bar{u}_k^m)^4 - 6\tilde{\sigma}_k^4}{T_m - 4} \quad \text{and} \quad w_k^m = \frac{T_m}{T_m - 1}\left(\tilde{\sigma}_k^4 - \frac{z_k^m}{T_m}\right).$$

Here, $\bar{u}_k^m = T_m^{-1}\sum_{t\in\mathcal{T}_m}\tilde{u}_{kt}$ is the mean of the residuals associated with the m^{th} volatility regime. Of course, if the u_t are Gaussian and this fact is known to the analyst, the kurtosis parameters can simply be replaced by zero; that is, $\tilde{\kappa}_1 = \tilde{\kappa}_2 = 0$ in the test statistic. Similarly, if the distribution is such that $\kappa_1 = \kappa_2$, the kurtosis parameter can be estimated from the full sample by using the formulas as above based on the full sample.

3.2. The test statistic

We base our test statistic on the eigenvalues $\tilde{\lambda}_{s+1}, \ldots, \tilde{\lambda}_{s+r}$. In principal component analysis with Gaussian independent and identically distributed (i.i.d.) data, the LR test for testing the equality of eigenvalues is based on the ratio of the geometric mean and arithmetic mean of the ML

estimators of the eigenvalues assumed to be identical under the null hypothesis (see Anderson, 1963, or Anderson, 2003, Section 11.7.3). Proceeding according to this pattern, we consider the test statistic

$$Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2) = -c(\tau, \tilde{\kappa}_1, \tilde{\kappa}_2)^2 Tr \log\left(\frac{\prod_{k=s+1}^{s+r} \tilde{\lambda}_k^{1/r}}{\frac{1}{r}\sum_{k=s+1}^{s+r} \tilde{\lambda}_k}\right) \tag{3.3}$$

$$= c(\tau, \tilde{\kappa}_1, \tilde{\kappa}_2)^2 \left[-T \sum_{k=s+1}^{s+r} \log(\tilde{\lambda}_k) + Tr \log\left(\frac{1}{r}\sum_{k=s+1}^{s+r} \tilde{\lambda}_k\right)\right],$$

where $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ are consistent estimators of the kurtosis parameters, and the term

$$c(\tau, \tilde{\kappa}_1, \tilde{\kappa}_2)^2 = \left(\frac{1+\tilde{\kappa}_1}{\tau} + \frac{1+\tilde{\kappa}_2}{1-\tau}\right)^{-1}$$

is included to obtain a convenient limiting distribution. Because the test statistic involves unrestricted estimators only, the test is akin to a Wald test. Of course, other distance measures could also be considered. The following proposition gives the asymptotic distribution of the test statistic under the null hypothesis. As can be seen in the proof in Appendix A, the asymptotic distribution consists of two independent terms having χ^2 distributions. The first one has $r-1$ degrees of freedom and reflects the $r-1$ restrictions on the λ_ks under $\mathbb{H}_0$. The second term has $\frac{1}{2}r(r-1)$ degrees of freedom and reflects the number of restrictions needed to identify the eigenvectors corresponding to the restricted λ_ks. Adding the two independent χ^2 terms results in an overall χ^2 distribution with $\frac{1}{2}(r+2)(r-1)$ degrees of freedom.

PROPOSITION 3.1. *Let u_t be independent white noise with elliptical distribution possessing a density as well as finite fourth moments with kurtosis parameters κ_m for $t \in \mathcal{T}_m$ $(m = 1, 2)$, where $\mathcal{T}_1 = \{1, \ldots, T_1 = [\tau T]\}$, $\mathcal{T}_2 = \{T_1 + 1, \ldots, T\}$, and the fraction $\tau \in (0, 1)$is assumed to be known and fixed. Furthermore, let $\lambda_1 \geq \cdots \geq \lambda_K$ be ordered from largest to smallest, and let $Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2)$ be the test statistic defined in equation (3.3) for testing the pair of hypotheses*

$$\mathbb{H}_0 : \lambda_{s+1} = \lambda_{s+2} = \cdots = \lambda_{s+r}(= \lambda_0) \quad \textit{versus} \quad \mathbb{H}_1 : \neg\mathbb{H}_0$$

for $s \in \{0, \ldots, K-1\}$ and $r \in \{2, \ldots, K-s\}$. Suppose that $\lambda_s \neq \lambda_{s+1}$ and $\lambda_{s+r} \neq \lambda_{s+r+1}$. Furthermore, let $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ be consistent estimators of κ_1 and κ_2, respectively. Then,

$$Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2) \xrightarrow{d} \chi^2\left(\tfrac{1}{2}(r+2)(r-1)\right),$$

where $\xrightarrow{d}$ denotes convergence in distribution.

The χ^2-limiting distribution requires that r eigenvalues are equal to λ_0 and all other eigenvalues are different from λ_0 (i.e., $\lambda_s \neq \lambda_{s+1}$ and $\lambda_{s+r} \neq \lambda_{s+r+1}$). In order to ensure this condition, the following considerations may be helpful. If

$$\mathbb{H}_0 : \lambda_1 = \cdots = \lambda_K$$

does not hold, we know that $\lambda_1 \neq \lambda_K$, and Proposition 3.1 enables us to test

$$\mathbb{H}_0 : \lambda_1 = \cdots = \lambda_{K-1} \quad \text{and} \quad \mathbb{H}_0 : \lambda_2 = \cdots = \lambda_K.$$

If these two null hypotheses are false, we can test all null hypotheses involving $K-2$ consecutive eigenvalues, etc. If all null hypotheses in this sequence are false, we can finally test

$$\mathbb{H}_0 : \lambda_1 = \lambda_2, \ldots, \mathbb{H}_0 : \lambda_{K-1} = \lambda_K.$$

If all the null hypotheses are rejected, the tests support that all the structural parameters are identified via heteroskedasticity.

For example, for $K = 3$, Proposition 3.1 implies that we can test the null hypothesis

$$\mathbb{H}_0 : \lambda_1 = \lambda_2 = \lambda_3 \tag{3.4}$$

by using $Q_3(\tilde{\kappa}_1, \tilde{\kappa}_2)$ with a $\chi^2(5)$ distribution. If the null hypothesis is false, it follows that $\lambda_1 \neq \lambda_3$, so that we can test

$$\mathbb{H}_0 : \lambda_1 = \lambda_2 \quad \text{and} \quad \mathbb{H}_0 : \lambda_2 = \lambda_3 \tag{3.5}$$

by using $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$ statistics with a $\chi^2(2)$ distribution. Rejecting the latter two null hypotheses is evidence of a fully identified structural model via heteroskedasticity.

Note that the two null hypotheses in (3.5) are each tested independently of the outcome of the other test. In other words, $\mathbb{H}_0 : \lambda_2 = \lambda_3$ is not conditioned on the outcome of the test of $\mathbb{H}_0 : \lambda_1 = \lambda_2$. Even if $\mathbb{H}_0 : \lambda_1 = \lambda_2$ is not rejected, this may be a problem of insufficient power against the hypothesis, and hence, testing $\mathbb{H}_0 : \lambda_2 = \lambda_3$ in addition is not inconsistent with at least two of the three λ_ks being distinct.

As this example shows, a sequential testing procedure is needed if the dimension of the VAR process is greater than two (i.e., $K > 2$). As usual, considering the sequential testing procedure results in different Type I and Type II errors than the individual tests. For example, $\mathbb{H}_0$ in (3.5) is considered only if (3.4) is rejected. Thus, the properties of the sequential testing procedure depend on the power of the tests. We will explore the small-sample power of the tests and the properties of sequential tests in Section 4.

In the previous literature, a related Wald test for equality of two eigenvalues of a similar type was sometimes used with a $\chi^2(1)$ distribution (e.g., Lanne et al., 2010, and Velinov and Chen, 2015). Although somewhat different volatility models were used in these publications, Proposition 3.1 suggests that the $\chi^2(1)$ distribution is a poor approximation to the actual asymptotic distributions of the test statistics. An adjustment of the degrees-of-freedom parameter is likely to be useful. Note that increasing the degrees-of-freedom parameter increases the correspondingly assumed p-values and hence may reduce the number of rejections.

One may wonder whether other procedures for assessing identification could be used in the context of our model. Other procedures have indeed been used for other volatility models. For example, Lütkepohl and Milunovich (2016) used a different frequentist testing strategy for VAR models with generalised autoregressive conditional heteroskedasticity (GARCH) errors. Likewise, Lewis (2018a) proposed a rank test for identification through heteroskedasticity. Moreover, Woźniak and Droumaguet (2015) and Lütkepohl and Woźniak (2020) considered Bayesian procedures for assessing identification through heteroskedasticity in related models. Clearly, the objective of the present paper is to develop frequentist tests. The advantage of our tests based on the relative variances in the second volatility state is that they provide detailed information on which shocks are identified. Thus, they are also useful in situations where only partial identification through heteroskedasticity is obtained. Moreover, it is quite common in the related literature to assess identification by considering the relative variances; see, for example, Lanne et al. (2010), Herwartz and Lütkepohl (2014), Lütkepohl and Velinov (2016), Velinov and Chen (2015), and Netšunajev (2013).

4. SMALL-SAMPLE PROPERTIES OF TESTS FOR IDENTIFICATION

4.1. Experimental design

We consider a range of DGPs to investigate the small-sample properties of our tests. All DGPs have zero intercept, $\nu = 0$. We still fit VARs with intercept. The error distributions are either Gaussian, $u_t \sim \mathcal{N}(0, \Sigma_m)$, have t distributions with five degrees of freedom [$t(5)$ distributions], or are χ^2-distributed as specified below. Some of the simulation results are only summarised in the following, and numbers are provided in the Online Appendix to this article, in which further details of some of the DGPs are also presented. We report results for the following specific DGPs.

DGP1: A bivariate ($K = 2$) VAR(0) process, $y_t = u_t$, with volatility change at $T_1 = \tau T$, where $0 < \tau < 1$. The errors u_t are independent Gaussian and have $t(5)$ or χ^2 distributions with covariance matrix $\Sigma_1 = I_2$ for $t = 1, \ldots, T_1 = \tau T$ and $\Sigma_2 = \text{diag}(\lambda_1, \lambda_2)$ for $t > T_1$, where $(\lambda_1, \lambda_2) = (2, 2), (2, 1)$. The $t(5)$-distributed u_t are temporally and contemporaneously independent and are generated as $\sqrt{\lambda_k} \times t(5)$ distributions for $t > T_1$ and $k = 1, 2$. The χ^2-distributed errors are temporally and contemporaneously independently distributed, with

$$u_t \sim \left(\frac{\chi^2(2) - 2}{2}, \frac{\chi^2(5) - 5}{\sqrt{10}} \right)' \quad \text{and} \quad u_t \sim \left(\sqrt{\lambda_1} \frac{\chi^2(2) - 2}{2}, \sqrt{\lambda_2} \frac{\chi^2(5) - 5}{\sqrt{10}} \right)'$$

for $t = 1, \ldots, T_1 = \tau T$ and for $t > T_1$, respectively.

DGP2: A Gaussian bivariate VAR(2) process,

$$y_t = \begin{bmatrix} 0.190 \\ 0.523 \end{bmatrix} + \begin{bmatrix} -0.036 & -0.705 \\ -0.093 & 1.211 \end{bmatrix} y_{t-1} + \begin{bmatrix} 0.090 & 0.796 \\ -0.085 & -0.276 \end{bmatrix} y_{t-2} + u_t,$$

where the slope coefficients are the estimated values for the first example process considered in Section 5.1. The error process consists of independent Gaussian vectors, $u_t \sim \mathcal{N}(0, \Sigma_m)$ for $t \in \mathcal{T}_m$, $m = 1, 2$, and $\tau = 0.3$, which also corresponds to the empirical value for the first example process. The covariance matrices are formed as in equation (2.3), with

$$B = \begin{bmatrix} 0.317 & 1.059 \\ 0.242 & -0.450 \end{bmatrix},$$

and $\Lambda = \text{diag}(\lambda_1, \lambda_2)$, with $(\lambda_1, \lambda_2) = (0.5, 0.5), (0.5, 0.1)$. The latter choice is also inspired by the first example process. Thus, DGP2 has features similar to the first example process.

Estimation of the VAR slope coefficients is done by GLS, and then the λ_k are obtained as generalised eigenvalues by using (2.5) with estimated covariance matrices $\tilde{\Sigma}_1 = T_1^{-1} \sum_{t=1}^{T_1} \tilde{u}_t \tilde{u}_t'$ and $\tilde{\Sigma}_2 = (T - T_1)^{-1} \sum_{t=T_1+1}^{T} \tilde{u}_t \tilde{u}_t'$, where $\tilde{u}_t$ are the GLS residuals. Even for the Gaussian processes, we pretend that we do not know the true distribution and fit models with possibly two distinct kurtosis parameters. We also vary the sample size because it is expected to affect the properties of the tests. Specifically, $T = 100$, $T = 250$, and $T = 500$ are used. The number of replications of all simulation experiments is 1,000.

Table 1. Relative rejection frequencies of tests for Gaussian DGP1 with different assumed or estimated volatility shift points (nominal significance level 5%).

		$\tau = 0.5$ (true value)		$\tau = 0.4$		$\tau = 0.3$		Estimated τ	
(λ_1, λ_2)	T	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$
(2,2)	100	0.054	0.063	0.064	0.072	0.056	0.062	0.166	0.182
(size)	250	0.045	0.046	0.046	0.050	0.057	0.055	0.087	0.087
	500	0.046	0.050	0.057	0.053	0.049	0.050	0.063	0.069
(2,1)	100	0.331	0.342	0.258	0.262	0.174	0.190	0.526	0.569
(power)	250	0.703	0.709	0.549	0.557	0.348	0.348	0.782	0.783
	500	0.949	0.950	0.846	0.844	0.716	0.720	0.960	0.962

Note: DGP1 is a bivariate Gaussian VAR(0) process with $\tau = 0.5$, independent $u_t \sim \mathcal{N}(0, I_2)$ for $t = 1, \ldots, 0.5T$, and independent $u_t \sim \mathcal{N}(0, \text{diag}(\lambda_1, \lambda_2))$ for $t = 0.5T + 1, \ldots, T$.

Table 2. Relative rejection frequencies of tests for DGP1 with different distributions (nominal significance level 5%).

		Gaussian DGP		t-Distributed DGP		χ^2-Distributed DGP	
(λ_1, λ_2)	T	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$
(2,2)	100	0.054	0.063	0.173	0.047	0.224	0.086
(size)	250	0.045	0.046	0.187	0.048	0.232	0.054
	500	0.046	0.050	0.214	0.048	0.239	0.053
(2,1)	100	0.331	0.342	0.400	0.195	0.388	0.181
(power)	250	0.703	0.709	0.618	0.391	0.647	0.358
	500	0.949	0.950	0.802	0.598	0.837	0.588

Note: The underlying DGPs are a VAR(0) process with independent Gaussian, independent t-distributed, and independent χ^2-distributed errors, respectively.

4.2. Simulation results for individual tests

The results for the bivariate DGP1 and DGP2 are discussed first. They are presented in Tables 1–3 and are also complemented by further results in the Online Appendix. Moreover, additional results for other processes are briefly summarised in Section 4.2.2.

4.2.1. Bivariate DGPs. In Table 1, we report results for a Gaussian DGP1, not only for the case where the change point of the volatility is specified correctly ($\tau = 0.5$), but also the situation of a misspecified or estimated volatility change point.[1] In the panel for $(\lambda_1, \lambda_2) = (2, 2)$ in Table 1, it can be seen that the relative rejection frequencies of the tests in finite samples are roughly in line with the nominal size of 5% if the true or a slightly misspecified volatility change point is used ($\tau = 0.5, 0.4, 0.3$). Even for sample size $T = 100$, the relative rejection frequencies are reasonably close to 5%, regardless of the kurtosis parameters used. In other words, $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$ has very similar rejection frequencies to $Q_2(0, 0)$ for a Gaussian process. Furthermore, the relative

[1] Estimation of the change point is done by minimising the objective function $T_1 \log\det \widehat{\Sigma}_1 + (T - T_1) \log\det \widehat{\Sigma}_2$ over $\tau \in [0.15, 0.85]$, where $\widehat{\Sigma}_1 = T_1^{-1} \sum_{t=1}^{T_1} \hat{u}_t \hat{u}_t'$ and $\widehat{\Sigma}_2 = (T - T_1)^{-1} \sum_{t=T_1+1}^{T} \hat{u}_t \hat{u}_t'$ are based on ordinary least squares residuals $\hat{u}_t$ and $T_1 = [\tau T]$.

Table 3. Relative rejection frequencies of tests for DGP2 (nominal significance level 5%).

		VAR(2)		VAR(1)	
(λ_1, λ_2)	T	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$	$Q_2(0, 0)$	$Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$
(0.5,0.5)	100	0.076	0.100	0.059	0.065
(size)	250	0.064	0.062	0.043	0.046
	500	0.065	0.066	0.039	0.041
(0.5,0.1)	100	0.330	0.360	0.255	0.272
(power)	250	0.644	0.650	0.561	0.570
	500	0.910	0.909	0.869	0.866
(0.5,0.1)	100	0.253	0.243	0.234	0.194
(size-adjusted power)	250	0.613	0.622	0.584	0.580
	500	0.904	0.892	0.876	0.879

Note: DGP2 is a bivariate Gaussian VAR(2) process with independent errors and $\tau = 0.3$.

rejection frequencies under the null hypothesis are not much affected by using a misspecified volatility change point (see the columns corresponding to $\tau = 0.4$ and 0.3 in Table 1).

In this respect, the situation is quite different for the power of the tests. In the lower part of Table 1, it is seen that the tests are more powerful if τ is specified correctly ($\tau = 0.5$), and the power declines when the volatility change point is misspecified (see $\tau = 0.4$ or 0.3 in Table 1). In fact, the farther away the assumed change point is from the true change point, the lower the power. It is again worth noting that it does not seem to matter much for the power of the tests whether the true kurtosis parameters are known or estimated (compare the corresponding results for $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$ and $Q_2(0, 0)$ in Table 1).

The picture changes if τ is estimated (see the last two columns in Table 1). Note that estimating the volatility change point is a challenge in our setup because the volatility does not change much. The standard deviation of the errors in the second volatility regime is only about 1.4 times larger than in the first regime for a $\lambda_k = 2$. Therefore, a precise estimation of the volatility change point cannot be expected. This feature translates into larger rejection frequencies in small samples. Even for $T = 500$, the rejection frequencies are a bit larger than 5% under the null hypothesis $(\lambda_1, \lambda_2) = (2, 2)$. Because the power in Table 1 is not size adjusted, it is also not surprising that the rejection frequencies for $(\lambda_1, \lambda_2) = (2, 1)$ are larger when τ is estimated than for fixed τ values.

To explore the impact of misspecifying the kurtosis parameters, we have also simulated DGP1 with t-distributed errors u_t and compare the results in Table 2 to results for Gaussian errors. Clearly, if the errors are t-distributed and the kurtosis parameters are mistakenly set to zero as for Gaussian processes, the corresponding test statistic $Q_2(0, 0)$ is considerably oversized even in large samples. For example, for $T = 500$, the relative rejection frequency is 0.214 for $Q_2(0, 0)$ in Table 2, instead of the desired 5% if the null hypothesis is true $[(\lambda_1, \lambda_2) = (2, 2)]$. In contrast, for Gaussian and t-distributed errors, the relative rejection frequencies of $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$ are very close to 5% if the null hypothesis is true. Note, however, that the power of our test is considerably lower for t-distributed errors than for Gaussian processes if estimated kurtosis parameters are used. The substantial size distortions when incorrect kurtosis parameters are used suggest that it is advisable to use the test statistics with estimated kurtosis parameters if, as is usual in practice, the true distribution is unknown. Nothing much can be gained from using the true kurtosis parameters if the distribution is known to be Gaussian.

We have also performed simulations of DGP1 with a non-elliptical error distribution to investigate the importance of the distributional assumption in Proposition 3.1. The results for a process with χ^2-distributed errors are also shown in Table 2. When the incorrect Gaussian kurtosis parameters are used, the rejection frequencies are again larger than the nominal 5%. However, the situation improves markedly if estimated kurtosis parameters are used. In that case, the rejection frequencies for larger samples ($T = 250, 500$) are close to the nominal 5%. This result is reassuring for using the tests in applied work, although there may, of course, be distributions where the situation is different. As in the case of t-distributed errors, the power is, however, substantially lower than for a Gaussian process.

So far we have considered only the situation in which the true DGP is white noise [VAR(0)]. From a practical point of view, this situation is not of much interest, of course, and in Table 3 we present results for DGP2, which is a Gaussian VAR(2) process based on a real-life dataset. The table presents the relative rejection frequencies for the situation where the VAR order is correctly specified to be $p = 2$ and also for the case where VAR(1) processes are fitted and, hence, the order is underspecified. Obviously, this situation may occur in practice and is therefore of interest here.

For the case of a correctly specified VAR lag order, the tests are slightly oversized in small samples because of the larger dimensional parameter space. For example, for $T = 100$ and $(\lambda_1, \lambda_2) = (0.5, 0.5)$, the test based on $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$ has a relative rejection frequency of 0.100 in Table 3, which is double the nominal significance level of 0.05. In this case, using the true instead of estimated kurtosis parameters [i.e., using $Q_2(0, 0)$ instead of $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$] improves the situation slightly. We have also fitted VAR(4) processes to DGP1 and present the results in Tables S1 and S2 of the Online Appendix, where it can be seen that the tests are even more oversized for larger models with more parameters if the sample size is small ($T = 100$). As in Table 3, the relative rejection frequencies are much closer to 5% for the larger sample sizes, however.

By comparing to the corresponding results in Table 1, it can also be seen in Table 3 that the size-adjusted power declines for larger models. In Table 3, the unadjusted power of the test based on $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$ is 0.360 for the VAR(2) and $T = 100$, which is similar to the power of the test for the corresponding case in Table 1 for a VAR(0). However, if we account for the fact that the test is oversized in Table 3 for the VAR(2) and compute the size-adjusted power, the relative rejection frequency is only 0.243. Thus, the results in Table 3 clearly show that the tests are oversized and the actual power of the tests is reduced for larger models in small samples when $T = 100$. Fortunately, the small-sample distortions in size and power largely disappear for the larger sample sizes ($T = 250$ and 500).

If we now consider the results for the misspecified VAR(1) process in Table 3, it turns out that in this situation the tests can even be somewhat too conservative, with relative rejection frequencies below the nominal 5% for larger samples. Also, the size-adjusted power may be distorted and tends to be lower than for the correctly specified VAR(2) process.

In summary, on the basis of our specific bivariate DGPs, it appears that the number of lags and, hence, the size of the model affects the rejection frequencies. Larger models result in oversized tests in small samples with reduced actual power. Size and power distortions appear in small samples if the VAR order is underspecified. If a slightly misspecified but fixed volatility change point is used, the size of the tests is not much affected, but the power is reduced. Having to estimate the volatility change point may affect the size of the tests, however. If the true distribution of the DGP is not known to be Gaussian, then it makes sense to use the test statistics based on estimated kurtosis parameters because they display similar rejection frequencies in the Gaussian case to the test statistics based on known kurtosis parameters, and their empirical size is much closer to the nominal size if the true distribution is non-Gaussian.

4.2.2. Extensions. We have also considered a number of extensions of the simulations reported so far and present further detailed results in the Online Appendix to this article. In Tables S1 and S2 of the Online Appendix, we explore the impact of the location of the volatility shift within the sample. More precisely, we compare results based on DGP1 with break fractions $\tau = 0.5$ and $\tau = 0.2$. In other words, we investigate the implication of the volatility change happening closer to the beginning of the sample. The results for both Gaussian and t-distributed processes are very similar for $\tau = 0.5$ and $\tau = 0.2$. Thus, even if the break point is not in the middle of the sample, this seems to have no substantial impact on the small-sample properties of our tests. Note that in this case, we have always fitted processes with correctly specified break dates.

In Table S3 of the Online Appendix, the properties for two Gaussian VAR(1) processes with different persistence are compared. One of the processes has an autoregressive root much closer to the unit circle than the other process. Again, this does not affect the small-sample properties of our tests much.

Finally, in Table S4 of the Online Appendix, results for a five-dimensional Gaussian VAR(1) process are presented, which indicate that the tests tend to reject too often under the null hypothesis for larger models. This, of course, is in line with our previous finding in Table 3. Note that individual tests rather than sequential tests are considered in Table S4. The magnitude of the distortions is quite substantial for such a large model, and even for sample size $T = 500$ there is still some bias. We have not found satisfactory small-sample corrections for our tests in large models and leave the issue for future research.

4.3. Sequential testing

So far we have explored the properties of our tests when they are applied to a single null hypothesis. However, in practice it is tempting to use them sequentially for higher-dimensional processes, as discussed in Section 3.2. Although we emphasise that our asymptotic results do not relate to this situation, we have also considered the possibility of applying the tests sequentially in a simulation exercise. To this end, we have generated the following three-dimensional Gaussian VAR(0) process.

> **DGP3**: A three-dimensional ($K = 3$) Gaussian VAR(0) process, $y_t = u_t$, with volatility change at $T_1 = 0.5T$ (i.e., $\tau = 0.5$). The errors are independent over t, with $u_t \sim \mathcal{N}(0, \Sigma_1 = I_3)$ for $t = 1, \ldots, \tau T_1$ and $u_t \sim \mathcal{N}(0, \Sigma_2 = \text{diag}(\lambda_1, \lambda_2, \lambda_3))$ for $t = T_1 + 1, \ldots, T$, where $(\lambda_1, \lambda_2, \lambda_3) = (2, 2, 2)$, $(3, 2, 1)$ or $(3,2,2)$.

Simulation results for this case are provided in Table 4. In that table, the notation $\mathbb{H}_{0i}|\neg\mathbb{H}_{0j}$ means that the null hypothesis $\mathbb{H}_{0i}$ is tested conditionally on $\mathbb{H}_{0j}$ being rejected. Thus, for example, $\mathbb{H}_{02}|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$ means that the null hypothesis $\mathbb{H}_{02} : \lambda_1 = \lambda_2$ is tested only for those simulations in which $\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$ has been rejected.

The results for $T = 100$ and $(\lambda_1, \lambda_2, \lambda_3) = (2, 2, 2)$ confirm that the test is slightly oversized for larger models. It rejects the true hypothesis $\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$ in roughly 8% of the cases if $Q_3(\tilde{\kappa}_1, \tilde{\kappa}_2)$ is used with a nominal significance level of 5%. All other tests in the same panel are conditional on this outcome, and, not surprisingly, false conclusions about the identification of the process are rare. In particular, the false conclusion that the process is fully identified via heteroskedasticity is never reached when $Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2)$ is used. As one would expect, the same is true if the sample size is increased and $(\lambda_1, \lambda_2, \lambda_3) = (2, 2, 2)$.

However, in the panels for $(\lambda_1, \lambda_2, \lambda_3) = (3, 2, 1)$, a full identification is also found rarely if $T = 100$, although the model is clearly identified. More precisely, $Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2)$ finds full identification

Table 4. Relative rejection frequencies of tests for three-dimensional Gaussian DGP3 (nominal individual significance level 5%).

T	$(\lambda_1, \lambda_2, \lambda_3)$	Hypotheses	$Q_r(0, 0)$	$Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2)$
100	(2,2,2)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.073	0.081
	(size)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.020	0.022
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.008	0.012
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0.001	0
	(3,2,1)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.548	0.570
	(power)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.114	0.128
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.226	0.240
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0.008	0.013
	(3,2,2)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.146	0.153
	(power)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.035	0.043
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.026	0.027
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0	0
250	(2,2,2)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.056	0.053
	(size)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.019	0.018
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.010	0.009
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0	0
	(3,2,1)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.934	0.936
	(power)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.270	0.272
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.612	0.616
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0.112	0.110
	(3,2,2)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.252	0.264
	(power)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.103	0.106
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.031	0.031
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0	0
500	(2,2,2)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.055	0.054
	(size)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.018	0.017
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.009	0.010
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0	0
	(3,2,1)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	1	1
	(power)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.509	0.515
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.932	0.929
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0.480	0.451
	(3,2,2)	$\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$	0.530	0.527
	(power)	$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$	0.309	0.305
		$\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01} : \lambda_2 = \lambda_3$	0.036	0.038
		$\mathbb{H}_{01}$ and $\mathbb{H}_{02}\|\neg\mathbb{H}_{01}$ and $\mathbb{H}_{03}\|\neg\mathbb{H}_{01}$	0.004	0.004

Note: The notation $\mathbb{H}_{0i}|\neg\mathbb{H}_{0j}$ means that the null hypothesis $\mathbb{H}_{0i}$ is tested conditionally on $\mathbb{H}_{0j}$ being rejected.

in only 1.3% of the cases for $T = 100$. Fortunately, the relative frequency goes up to 45.1% for $T = 500$. Of course, even that means that full identification is not even found in half of the cases. In other words, the tests do have some power in the sequential procedure, but it is rather small even for larger samples. Note, however, that some identifying information through heteroskedasticity is always found in this case because, for $T = 500$ and $(\lambda_1, \lambda_2, \lambda_3) = (3, 2, 1)$, the false null hypothesis $\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$ is always rejected.

It is perhaps also of interest to consider the case $(\lambda_1, \lambda_2, \lambda_3) = (3, 2, 2)$, which corresponds to a partially identified model. In this case, it is in fact rather unlikely that full identification will be diagnosed, even in small samples. In other words, even for $T = 100$, the three hypotheses $\mathbb{H}_{01}$, $\mathbb{H}_{02}|\neg\mathbb{H}_{01}$, and $\mathbb{H}_{03}|\neg\mathbb{H}_{01}$ are never rejected jointly. Also, the true conditional null hypothesis $\mathbb{H}_{03}|\neg\mathbb{H}_{01}$ alone is not rejected very often. More precisely, the relative rejection frequency in the testing sequence of $Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2)$ for $T = 100$ is 0.027. In fact, this rejection frequency increases when the sample size increases because the false null hypothesis $\mathbb{H}_{01} : \lambda_1 = \lambda_2 = \lambda_3$ is rejected more often when the sample size increases. Thus, the actual rejection frequency for $\mathbb{H}_{03}|\neg\mathbb{H}_{01}$ moves closer to the nominal 5%. On the other hand, the rejection frequency of the false null hypothesis $\mathbb{H}_{02}|\neg\mathbb{H}_{01} : \lambda_1 = \lambda_2$ is also not rejected very often for sample size $T = 250$. For this case, the relative rejection frequency of $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$ is 0.106.

The overall conclusion for the sequential testing procedure is that for our still relatively small three-dimensional example process, the test is valuable, but one needs to be careful in drawing conclusions about full and partial identification of the underlying model because the overall procedure is obviously less powerful than the individual tests, and it is generally more prone to error. Given the results for the individual tests, one may speculate that it becomes even more difficult to reach correct conclusions if the process is larger (higher-dimensional or larger lag order).

5. EMPIRICAL EXAMPLES

We present two empirical examples to illustrate the use of our tests for identification. The first one reconsiders a bivariate model for US data originally proposed by Blanchard and Quah (1989), and the second one looks at a model from Kilian (2009) for the crude oil market.

5.1. Blanchard-Quah model

Blanchard and Quah (1989) identified demand and supply shocks in a bivariate macro model for US economic growth and unemployment by assuming that the demand shocks have no long-run effects on output. Their model has become a textbook example for identification by restrictions on the long-run effects of the structural shocks; see, for example, Breitung et al. (2004), Lütkepohl (2005, Chapter 9), and Kilian and Lütkepohl (2017, Chapter 10). Chen and Netšunajev (2016) used seasonally adjusted quarterly data for the period 1970q1–2007q4 and used identification through heteroskedasticity to investigate the validity of the long-run neutrality of demand shocks in a VAR(2) model for $y_t = (\Delta gnp_t, U_t)$, where gnp_t denotes the log of gross national product (GNP) and U_t is the unemployment rate. They modeled volatility changes by a smooth transition in the reduced-form error covariance matrices. Their estimated change in the variances turned out to be a decline in the error variances around 1983q1, which was roughly the time when the Great Moderation started in the US (see also Figure 1 of Chen and Netšunajev, 2016). Therefore, it is plausible to use the VAR model (2.1) with a change in the residual covariance matrix in period 1983q1.

We have used the data from Chen and Netšunajev (2016) and estimated a VAR(2) model with error covariance change as in expression (2.2), with $T_1 = 1982q4$. Because we have a sample size of $T = 152$, the corresponding sample fraction of the break is $\tau = 0.34$. The estimated relative variances are $\tilde{\lambda}_1 = 0.457$ (0.154) and $\tilde{\lambda}_2 = 0.152$ (0.041), with the estimated standard

Table 5. Estimated relative variances of oil market model.

Relative variance	Estimate	Standard deviation
λ_1	2.174	0.275
λ_2	0.358	0.050
λ_3	0.186	0.025

errors given in parentheses. Both $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ are smaller than one so that the second part of the sample is associated with lower residual volatility.

The estimated λ_ks are clearly distinct, and on the basis of the standard errors, one may expect that they are significantly different. This informal evidence was in fact used by Chen and Netšunajev (2016) to justify the assumption of distinct relative variances. Using our test statistic $Q_2(\tilde{\kappa}_1, \tilde{\kappa}_2)$, we can now formally test the null hypothesis $\mathbb{H}_0 : \lambda_1 = \lambda_2$. This test statistic equals 8.600, follows a $\chi^2(2)$ distribution, and has a p-value of 0.014, so that $\mathbb{H}_0$ is rejected at a common level of significance. Thus, we support the assumption underlying the analysis of Chen and Netšunajev (2016). Note that we use the test statistic with estimated kurtosis parameters to avoid the assumption of a Gaussian error distribution.

5.2. *Oil market model*

Our second example is based on a benchmark study by Kilian (2009), who investigated the interactions on the global market for crude oil by using a VAR model for $y_t = (\Delta prod_t, q_t, p_t)'$, where $\Delta prod_t$ represents the percent change in global crude oil production, q_t is the logarithm of a detrended index of real economic activity, and p_t is the logarithm of the real price of oil. Kilian used a recursive scheme (corresponding to a lower-triangular B matrix) to identify an oil supply shock, an aggregate demand shock, and an oil-market–specific demand shock.

Using monthly data for 1973m2–2006m12 and a VAR(3) model, Lütkepohl and Netšunajev (2014) investigated Kilian's model in the context of identification through heteroskedasticity. Although these authors used different volatility models based on a Markov-switching mechanism, their analysis suggests that a VAR model with a volatility change in the second half of 1987 may also provide a reasonable approximation to the DGP. Therefore, we use their dataset with $T = 407$ and a VAR(3) model and estimate the volatility change point with the estimation procedure used for DGP1 in Table 1 (see Footnote 1). The estimated change point turns out to be in October 1987, which corresponds to $\tau = 0.43$. We have applied our identification tests using this model and volatility change point.

The estimated relative variances are presented in Table 5. Because the estimates $\tilde{\lambda}_1$, $\tilde{\lambda}_2$, and $\tilde{\lambda}_3$ are clearly distinct and their estimated standard errors suggest that the underlying λ_ks are also different, it may be worthwhile to perform formal statistical tests to confirm what might be concluded from an informal inspection. The results of our tests are given in Table 6. The first test of the null hypothesis $\mathbb{H}_0 : \lambda_1 = \lambda_2 = \lambda_3$ presents strong evidence against all three relative variances' being equal. However, the last test in Table 6 shows that there is little evidence against $\lambda_2 = \lambda_3$. Thus, our formal tests show that an informal inspection may lead to misleading conclusions.

In fact, on the basis of statistical procedures that are not backed by formal asymptotic theory, Lütkepohl and Netšunajev (2014) concluded that their SVAR model was identified through

Table 6. Identification tests for oil market model.

$\mathbb{H}_0$	$Q_r(\tilde{\kappa}_1, \tilde{\kappa}_2)$	Degrees of freedom	p-value
$\lambda_1 = \lambda_2 = \lambda_3$	44.367	5	1.950×10^{-8}
$\lambda_1 = \lambda_2$	19.580	2	5.599×10^{-5}
$\lambda_2 = \lambda_3$	2.859	2	0.239

Table 7. Test for lower-triangularity of $B = [b_{ij}]$.

$\mathbb{H}_0$	LR statistic	Degrees of freedom	p-value
$b_{12} = b_{13} = b_{23} = 0$	6.056	3	0.109
		2	0.048

heteroskedasticity, and on the basis of that assumption, they tested triangularity of B. They could not reject the triangularity of B. However, such a result may be due to a lack of identification. In fact, if no identifying information is available from volatility changes, a model with lower-triangular B matrix would be just identified and would have the same likelihood maximum as the reduced form and, hence, an LR test value of zero.

Because the analysis in Lütkepohl and Netšunajev (2014) was based on a different volatility model, we also perform a corresponding LR test of the null hypothesis that B is triangular in the context of our model. The results are reported in Table 7. If full identification of B through heteroskedasticity is assumed, the test statistic should have an asymptotic $\chi^2(3)$ distribution because we are testing three zero restrictions on B, which are all overidentifying. The corresponding p-value is 0.109, which may be taken as support for a recursive model with triangular B matrix. Our test results for the relative variances do not support full identification through heteroskedasticity, however. Hence, the conclusion that B is lower-triangular is based on soft grounds, because not rejecting triangularity of B may just be a reflection of insufficient identifying information in the data. Consequently, the actual number of degrees of freedom of the LR test may be lower than three. Note that even if there is no full identification through heteroskedasticity, we still have a fully identified model under $\mathbb{H}_0$ because the zero restrictions imposed under $\mathbb{H}_0$ identify the structural matrix B. Thus, the standard asymptotic theory that implies an asymptotic χ^2 distribution of the LR statistic is not in question here, but only its degrees of freedom. Assuming that two of the λ_ks are identical such that heteroskedasticity does not fully identify the B matrix and the LR test for lower-triangularity has perhaps only two degrees of freedom would result in a p-value of 0.048. Hence, one would reject a lower-triangular B at a 5% significance level. This outcome shows that whether all or only some of the λ_ks are distinct can make a difference for the conclusions drawn from an empirical analysis.

6. CONCLUSIONS

In the present study, we have developed frequentist tests for identification through heteroskedasticity in SVAR models. We consider VAR models with two volatility states. The change point

of the volatility is assumed to be known. The tests are Wald-type tests, such that only the unrestricted model has to be estimated. The model errors are assumed to be from the class of elliptical distributions. This class of distributions includes the Gaussian distribution, as well as *t* and mixed normal distributions. We propose test versions where the kurtosis of the distribution is assumed to be known and also allow for the possibility that the kurtosis is estimated.

The asymptotic null distributions of the test statistics are derived and are shown to be χ^2 distributions, although the models are not identified under the null hypothesis. We have also explored the small-sample properties of the tests by Monte Carlo simulations, and we have found that the tests are oversized for large models when the sample size is small. However, for larger samples and smaller models, the size and power of the tests are reasonable, and the properties of the tests do not depend on the timing of the volatility break. However, misspecification of the volatility change point is found to reduce power, and underspecification of the lag order may lead to size and power distortions. Moreover, if the change point has to be estimated, this may lead to size distortions. In most situations, the small-sample properties are not much affected by estimating the kurtosis parameters. Thus, in practice we recommend the use of the test versions that are based on estimated kurtosis parameters.

Two empirical examples are considered to illustrate the usefulness of the tests. The first example considers a bivariate model for US data. Our tests support the assumption of earlier studies that the model is identified by heteroskedasticity. The second example is based on a three-dimensional model for the international crude oil market. We find that there is some identifying information from heteroskedasticity, but there is little support for a full identification.

There are a number of desirable extensions of our tests. First, it would be useful if tests for more than two volatility regimes could be developed. Moreover, the volatility model is very special. It assumes that the change in volatility is extraneously generated. Other models have been used in the literature on identification through heteroskedasticity. It would be desirable to have tests for identification for other related models, as well.

ACKNOWLEDGEMENTS

Part of the research for this paper was done while the first author was a Fernand Braudel Fellow in the Economics Department of the European University Institute, Florence, Italy. The second and fourth authors thank the Academy of Finland (grant number 1308628) for financial support. We thank Savi Virolainen, two anonymous referees and the editor, Dennis Kristensen, for helpful comments that led to improvements of the exposition of our paper. The paper was presented at a number of seminars and conferences. We thank the participants for helpful comments.

REFERENCES

Amemiya, Y. (1990). A note of the limiting distribution of certain characteristic roots. *Statistics and Probability Letters 9*, 465–70.

Anderson, T. W. (1963). Asymptotic theory for principal component analysis. *Annals of Mathematical Statistics 34*, 122–48.

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis*. (3rd ed.). New York, NY: John Wiley.

Andrews, I. (2018). Valid two-step identification-robust confidence sets for GMM. *Review of Economics and Statistics 100*, 337–48.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review 79*, 4, 655–73.

Breitung, J., R. Brüggemann and H. Lütkepohl (2004). Structural vector autoregressive modeling and impulse responses. In H. Lütkepohl and M. Krätzig (Eds.), *Applied Time Series Econometrics*, pp. 159–96. Cambridge: Cambridge University Press.

Chen, W. and A. Netšunajev (2016). On the long-run neutrality of demand shocks. *Economics Letters 139*, 57–60.

Herwartz, H. and H. Lütkepohl (2014). Structural vector autoregressions with Markov switching: combining conventional with statistical identification of shocks. *Journal of Econometrics 183*, 104–16.

Kilian, L. (2009). Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. *American Economic Review 99*, 3, 1053–69.

Kilian, L. and H. Lütkepohl (2017). *Structural Vector Autoregressive Analysis*. Cambridge, UK: Cambridge University Press.

Lanne, M. and H. Lütkepohl (2008). Identifying monetary policy shocks via changes in volatility. *Journal of Money, Credit and Banking 40*, 1131–49.

Lanne, M. and H. Lütkepohl (2014). A statistical comparison of alternative identification schemes for monetary policy shocks. In J. Knif and B. Pape (Eds.), *Contributions to Mathematics, Statistics, Econometrics, and Finance - Essays in Honour of Professor Seppo Pynnönen, Acta Wasaensia 296, Statistics 7*, pp. 137–52. Vaasa, Finland: University of Vaasa.

Lanne, M., H. Lütkepohl and K. Maciejowska (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control 34*, 121–31.

Lewis, D. J. (2018a). Identifying shocks via time-varying volatility. Technical Report 871, Federal Reserve Bank of New York, New York, NY, USA.

Lewis, D. J. (2018b). Robust inference in models identified via heteroskedasticity. Technical Report Staff Report No. 876, Federal Reserve Bank of New York, New York, NY, USA.

Lütkepohl, H. (2005). , *New Introduction to Multiple Time Series Analysis*. Berlin, Germany: Springer-Verlag.

Lütkepohl, H. (2013). Identifying structural vector autoregressions via changes in volatility. *Advances in Econometrics 32*, 169–203.

Lütkepohl, H. and G. Milunovich (2016). Testing for identification in SVAR-GARCH models. *Journal of Economic Dynamics and Control 73*, 241–58.

Lütkepohl, H. and A. Netšunajev (2014). Disentangling demand and supply shocks in the crude oil market: how to check sign restrictions in structural VARs. *Journal of Applied Econometrics 29*, 479–96.

Lütkepohl, H. and A. Netšunajev (2017). Structural vector autoregressions with heteroskedasticy: a review of different volatility models. *Econometrics and Statistics 1*, 2–18.

Lütkepohl, H. and T. Schlaak (2018). Choosing between different time-varying volatility models for structural vector autoregressive analysis. *Oxford Bulletin of Economics and Statistics 80*, 715–35.

Lütkepohl, H. and A. Velinov (2016). Structural vector autoregressions: checking identifying long-run restrictions via heteroskedasticity. *Journal of Economic Surveys 30*, 377–92.

Lütkepohl, H. and T. Woźniak (2020). Bayesian inference for structural vector autoregressions identified by Markov-switching heteroskedasticity. *Journal of Economic Dynamics and Control 113*, doi.org/10.1016/j.jedc.2020.103862.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Hoboken, NJ: John Wiley & Sons.

Netšunajev, A. (2013). Reaction to technology shocks in Markov-switching structural VARs: identification via heteroskedasticity. *Journal of Macroeconomics 36*, 51–62.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. 2nd ed., New York, NY: John Wiley.
Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics 85*, 777–92.
Rigobon, R. and B. Sack (2003). Measuring the reaction of monetary policy to the stock market. *Quarterly Journal of Economics 118*, 639–69.
Schott, J. R. (2001). Some tests for the equality of covariance matrices. *Journal of Statistical Planning and Inference 94*, 25–36.
Tyler, D. E. (1983). Robustness and efficiency properties of scatter matrices. *Biometrika 70*, 411–20.
Velinov, A. and W. Chen (2015). Do stock prices reflect their fundamentals? New evidence in the aftermath of the financial crisis. *Journal of Economics and Business 80*, 1–20.
Woźniak, T. and M. Droumaguet (2015). Assessing monetary policy models: Bayesian inference for heteroskedastic structural VARs. Technical report, University of Melbourne, Melbourne, Australia.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Online Appendix
Replication Package

Co-editor Dennis Kristensen handled this manuscript.

APPENDIX A: PROOF OF PROPOSITION 3.1

We assume that the u_t have an elliptical distribution possessing a density as well as finite fourth moments as in Proposition 3.1. Such assumptions are needed because we are using limiting results for elliptical distributions from Anderson (2003).

We study $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_K$, the eigenvalues of $\tilde{\Sigma}_2\tilde{\Sigma}_1^{-1}$, and follow the pattern of proof in Anderson (2003, Sections 13.6.1 and 13.6.2). As in Anderson (2003, Equation 9 on p. 550), for the theoretical developments that follow, it will be convenient to transform the estimators $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$ and consider the matrices

$$\tilde{\Omega}_1 = B^{-1}\tilde{\Sigma}_1 B'^{-1} \text{ and } \tilde{\Omega}_2 = B^{-1}\tilde{\Sigma}_2 B'^{-1}.$$

(As before, we here assume that the first nonzero element on each column of B is positive.) With this transformation, the asymptotic distributions of $\tilde{\Omega}_1$ and $\tilde{\Omega}_2$ below will depend only on Λ and not on B (note also that the theoretical counterparts of $\tilde{\Omega}_1$ and $\tilde{\Omega}_2$ are $B^{-1}\Sigma_1 B'^{-1} = I_K$ and $B^{-1}\Sigma_2 B'^{-1} = \Lambda$). Furthermore, because $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_K$ are the eigenvalues of $\tilde{\Sigma}_2\tilde{\Sigma}_1^{-1}$, they are also the eigenvalues of $\tilde{\Omega}_2\tilde{\Omega}_1^{-1}$ or, equivalently, the eigenvalues of $\tilde{\Omega}_1^{-1/2}\tilde{\Omega}_2\tilde{\Omega}_1^{-1/2}$. Thus, as far as asymptotic properties of the eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_K$ or their functions are concerned, we can use the matrices $\tilde{\Omega}_1$ and $\tilde{\Omega}_2$ instead of $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$.

From (3.1) and (3.2), it follows that the asymptotic distributions of $\tilde{\Omega}_1$ and $\tilde{\Omega}_2$ can be derived by using the (independent) errors u_t in place of the residuals in the definitions of $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$. For simplicity, denote $T_2 = T - T_1$, and note that, because of the assumption $T_1 = [\tau T]$ for some $\tau \in (0, 1)$, both $T_1 \to \infty$ and $T_2 \to \infty$ when $T \to \infty$. From Theorem 3.6.2 in Anderson (2003, p. 102), we can thus conclude that $T_1^{1/2}(\tilde{\Omega}_1 - I_K) = \tilde{Z}_1$ and $T_2^{1/2}(\tilde{\Omega}_2 - \Lambda) = \tilde{Z}_2$, say, converge jointly in distribution, as $T \to \infty$, to the matrices $Z_1 = [z_{1,ij}]$ and $Z_2 = [z_{2,ij}]$ $(i, j = 1, \ldots, K)$. Here, Z_1 and Z_2 are independent, their elements are jointly normally distributed, and their functionally independent elements are statistically independent.

Furthermore, their elements have mean zero and covariance structure given by

$$Cov[\text{vec}(Z_1)] = (1+\kappa_1)(I_{K^2}+\mathbf{K})(I_K \otimes I_K) + \kappa_1 \text{vec}(I_K)\text{vec}(I_K)'$$

and

$$Cov[\text{vec}(Z_2)] = (1+\kappa_2)(I_{K^2}+\mathbf{K})(\Lambda \otimes \Lambda) + \kappa_2 \text{vec}(\Lambda)\text{vec}(\Lambda)',$$

where $\mathbf{K}$ ($K^2 \times K^2$) is a commutation matrix and $\otimes$ denotes the Kronecker product. The Gaussian case is obtained as a special case by choosing $\kappa_1 = \kappa_2 = 0$. In what follows, the null hypothesis is assumed to hold unless otherwise stated.

As in Tyler (1983, p. 413, in the paragraph after Equation 1), we can describe the elements of $Cov[\text{vec}(Z_1)]$ as follows. The distinct off-diagonal elements of $Cov[\text{vec}(Z_1)]$ are uncorrelated with each other and uncorrelated with the diagonal elements, and each of them has variance $1+\kappa_1$. All diagonal elements have variance $2+3\kappa_1$, and the covariance between any two diagonal elements is κ_1. In the special case where $\Lambda = \lambda_0 I_K$, the same description clearly applies to the elements of $Cov[\text{vec}(Z_2)]$, with κ_1 replaced by κ_2, provided the variances and covariances are multiplied by λ_0^2, and, by the definition of the commutation matrix, the same is true when Z_2 is replaced by the matrix $[z_{2,ij}]_{i,j=s+1}^{s+r}$ and Λ is replaced by $\Lambda_2 = \lambda_0 I_r$.

Theorem 1 of Amemiya (1990) implies that $T^{1/2}(\tilde{\lambda}_{s+1} - \lambda_0, \ldots, \tilde{\lambda}_{s+r} - \lambda_0)$ converges in distribution to an $(r \times 1)$ random vector consisting of the eigenvalues of the matrix $U = [u_{ij}]_{i,j=1}^{r} = [(1-\tau)^{-1/2} z_{2,ij} - \lambda_0 \tau^{-1/2} z_{1,ij}]_{i,j=s+1}^{s+r}$. The elements of U are jointly normally distributed with mean zero and covariances given in the following equations, where $c(\tau,\kappa_1,\kappa_2)^2 = \left(\frac{1+\kappa_1}{\tau} + \frac{1+\kappa_2}{1-\tau}\right)^{-1}$ and $i,j = s+1,\ldots,s+r$:

$$\mathbb{E}[u_{ij}^2] = \frac{(1+\kappa_2)\lambda_0^2}{1-\tau} + \frac{(1+\kappa_1)\lambda_0^2}{\tau} = \lambda_0^2 c(\tau,\kappa_1,\kappa_2)^{-2} \text{ for } i \neq j$$

$$\mathbb{E}[u_{ii}^2] = \frac{(2+3\kappa_2)\lambda_0^2}{1-\tau} + \frac{(2+3\kappa_1)\lambda_0^2}{\tau} = 2\lambda_0^2 c(\tau,\kappa_1,\kappa_2)^{-2} + \lambda_0^2\left(\frac{\kappa_2}{1-\tau} + \frac{\kappa_1}{\tau}\right)$$

$$\mathbb{E}[u_{ii}u_{jj}] = \lambda_0^2\left(\frac{\kappa_2}{1-\tau} + \frac{\kappa_1}{\tau}\right) \text{ for } i \neq j.$$

Distinct off-diagonal elements of U are independent of each other, and the off-diagonal and diagonal elements of U are independent.

Now define the (infeasible) test statistic,

$$Q_r(\kappa_1,\kappa_2) = c(\tau,\kappa_1,\kappa_2)^2\left[-T\sum_{k=s+1}^{s+r}\log(\tilde{\lambda}_k) + Tr\log\left(\frac{1}{r}\sum_{k=s+1}^{s+r}\tilde{\lambda}_k\right)\right],$$

for which we have

$$Q_r(\kappa_1,\kappa_2) \xrightarrow{d} \frac{c(\tau,\kappa_1,\kappa_2)^2}{\lambda_0^2}\sum_{i<j}u_{ij}^2 + \frac{c(\tau,\kappa_1,\kappa_2)^2}{2\lambda_0^2}\left[\sum_{i=s+1}^{s+r}u_{ii}^2 - \frac{1}{r}\left(\sum_{i=s+1}^{s+r}u_{ii}\right)^2\right]$$

$$\stackrel{def}{=} Q_{1,r}^*(\kappa_1,\kappa_2) + Q_{2,r}^*(\kappa_1,\kappa_2).$$

Here, $Q_{1,r}^*(\kappa_1,\kappa_2)$ and $Q_{2,r}^*(\kappa_1,\kappa_2)$ are independent, and $Q_{1,r}^*(\kappa_1,\kappa_2)$ has a χ^2 distribution with $\frac{1}{2}r(r-1)$ degrees of freedom. As to $Q_{2,r}^*(\kappa_1,\kappa_2)$, defining $\boldsymbol{w}_s$ as

$$\boldsymbol{w}_s = \frac{c(\tau,\kappa_1,\kappa_2)}{\sqrt{2}\lambda_0}(u_{s+1,s+1},\ldots,u_{s+r,s+r})',$$

and the $(r \times r)$ projection matrix P_r as $P_r = I_r - \frac{1}{r}\mathbf{1}_r\mathbf{1}_r'$, where $\mathbf{1}_r = (1,\ldots,1)'$ is an $(r \times 1)$ vector, we have $Q_{2,r}^*(\kappa_1,\kappa_2) = \boldsymbol{w}_s' P_r \boldsymbol{w}_s$. Hence, it follows that the random vector $\boldsymbol{w}_s$ is normally distributed with zero

mean and covariance matrix (see the above expressions of $\mathbb{E}[u_{ii}^2]$ and $\mathbb{E}[u_{ii}u_{jj}]$ $(i \neq j)$)

$$Cov[\boldsymbol{w}_s] = I_r + \frac{c(\tau,\kappa_1,\kappa_2)^2}{2}\left(\frac{\kappa_2}{1-\tau}+\frac{\kappa_1}{\tau}\right) I_r + \frac{c(\tau,\kappa_1,\kappa_2)^2}{2}\left(\frac{\kappa_2}{1-\tau}+\frac{\kappa_1}{\tau}\right)(\mathbf{1}_r\mathbf{1}_r' - I_r)$$

$$= I_r + \frac{c(\tau,\kappa_1,\kappa_2)^2}{2}\left(\frac{\kappa_2}{1-\tau}+\frac{\kappa_1}{\tau}\right)\mathbf{1}_r\mathbf{1}_r'.$$

Thus, we have $P_r Cov[\boldsymbol{w}_s] = P_r$, and we find that $Q^*_{2,r}(\kappa_1,\kappa_2)$ has a χ^2 distribution with $r-1$ degrees of freedom. This fact can be justified by a well-known result of quadratic forms of normal random vectors; see, for example, result (vii) in Rao (1973, p. 188).

From the preceding discussion, we can now conclude that $Q_r(\kappa_1,\kappa_2) \xrightarrow{d} Q^*_{1,r}(\kappa_1,\kappa_2) + Q^*_{2,r}(\kappa_1,\kappa_2)$, where $Q^*_{1,r}(\kappa_1,\kappa_2)$ and $Q^*_{2,r}(\kappa_1,\kappa_2)$ are independent and have χ^2 distributions with degrees of freedom $\frac{1}{2}r(r-1)$ and $r-1$. Therefore, the infeasible test statistic $Q_r(\kappa_1,\kappa_2)$ has an asymptotic χ^2 distribution with $\frac{1}{2}(r+2)(r-1)$ degrees of freedom, and the same is true for its feasible version $Q_r(\tilde{\kappa}_1,\tilde{\kappa}_2)$, where $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$ are consistent estimators of κ_1 and κ_2, respectively. This proves Proposition 3.1.