AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

Testing Constancy of the Error Covariance Matrix in Vector Models against Parametric Alternatives using a Spectral Decomposition

Yukai Yang

CREATES Research Paper 2014-11

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Testing Constancy of the Error Covariance Matrix in Vector Models against Parametric Alternatives using a Spectral Decomposition

Yukai Yang
CREATES, Aarhus University
CORE, Université catholique de Louvain

4 April 2014

Abstract

I consider multivariate (vector) time series models in which the error covariance matrix may be time-varying. I derive a test of constancy of the error covariance matrix against the alternative that the covariance matrix changes over time. I design a new family of Lagrange-multiplier tests against the alternative hypothesis that the innovations are time-varying according to several parametric specifications. I investigate the size and power properties of these tests and find that the test with smooth transition specification has satisfactory size properties. The tests are informative and may suggest to consider multivariate volatility modelling.

JEL Classification: C32, C52.

Keywords: Covariance constancy; Error covariance structure; Lagrange multiplier test; Spectral decomposition; Auxiliary regression; Model misspecification; Monte Carlo simulation.

Acknowledgements: This research has been supported by CREATES, funded by the Danish National Research Foundation, Grant Number DNFR78. Material from this paper has been presented at EEA-ESEM Conference, Malaga, August 2012; 23rd (EC)2 Conference, Maastricht, December 2012. Comments from participants of these occasions are gratefully acknowledged. I also wish to thank Timo Teräsvirta, Rickard Sandberg, Dick van Dijk and Christian Hafner for many constructive remarks. Responsibility for any errors or shortcomings in this work remains mine.

1 Introduction

In univariate time series analysis, testing the adequacy of the estimated model has long been standard practice, see Box and Jenkins (1970). In vector models, most of the tests available in the literature, for checking the specification, have been designed to test the conditional mean. However, the error covariance matrix can also be nonconstant over time, even when the conditional mean is correctly specified, see Lütkepohl (2004) for a detailed discussion. Tests exist for testing the constancy of the error variance in univariate models, whereas less has been done in multivariate models.

The most recent work for testing the constancy of the error covariance matrix can be found in Eklund and Teräsvirta (2007). They derived a family of test statistics against various types of misspecification making the use of the constant conditional correlation (CCC) framework of Bollerslev (1990). Typically, the tests derived from specific parametric models may have the highest power against the assumed alternative, and hence, offer guidance to the model builders. It is, however, desirable to have more tests based on different assumptions about alternatives. The purpose of this paper is to develop a new multivariate heteroskedasticity test as an alternative to the one proposed in Eklund and Teräsvirta (2007).

The basic idea is to derive tests for the null hypothesis of constant covariance against parsimoniously parameterised alternatives, such that the tests would still be powerful against many kinds of departure from parameter constancy. The constancy tests in this paper are of the Lagrange-multiplier type. Under the null hypothesis the covariance matrix is constant over time, whereas under the alternative, the evolution of the covariance matrix through time is fully specified.

The constancy tests in this work are based on the spectral decomposition of the error covariance matrix. I develop several tests which allow for various possible specifications under the spectral decomposition assumption. This considerably reduces the dimension of the null hypothesis compared to the case where all the elements in the half-vectorization of the covariance matrix can vary freely under the alternative hypothesis.

The constancy tests can be extremely useful, for instance, in the structural vector models where the constant conditional correlation (CCC) assumption of Bollerslev (1990) is not plausible, or the multivariate volatility models where the matrix exponential assumption of Kawakatsu (2006) has been made. They are informative in the sense that they suggest specifications for modelling the multivariate time-varying covariance matrix, especially when some test results are significant while the others are not.

Nonlinearity and misspecification tests in multivariate models, such as the ones developed against the smooth transition alternative in Teräsvirta and Yang (2014a), may suffer from the possible heteroskedasticity, because the time-varying covariance results in a strong size-distortion. A solution to that problem is to apply the wild bootstrap version of these tests, see the applications in Teräsvirta and Yang (2014b). However, the bootstrap can be time-consuming especially in nonlinear models. Thus, it is essential to have a joint test against heteroskedasticity for multivariate models before using the bootstrap.

Following Eklund and Teräsvirta (2007), three types of alternatives to constancy are considered. The first one may be viewed as a multivariate generalization of the heteroskedasticity test of White (1980), and the second one generalizes the test against autoregressive conditional heteroskedasticity of Engle (1982). The third variant of the test generalizes the univariate constant variance test of Medeiros and Veiga (2003), in which it is assumed that under the alternative hypothesis the variance changes smoothly over time.

In addition to the spectral decomposition assumption, I assume that the time-varying eigenvalues in the covariance matrix are functions of linear combinations of possible exogenous variables. The simulation-based experiments show that the test based on the smooth transition specification has satisfactory size among the others. All tests have good power properties even in high-dimensional vector models when the alternative hypothesis is true.

The plan of the paper is as follows. The statistical model is introduced in Section 2. The Lagrange-multiplier type test statistic is derived in Section 3. In Section 4, I discuss different kinds of specification. The finite sample properties of the tests are investigated in Section 5. Section 6 concludes.

2 The model

Consider the following multivariate (vector) model:

$$\boldsymbol{y}_t = \boldsymbol{f}(\boldsymbol{x}_t) + \boldsymbol{u}_t, \tag{2.1}$$

where $\boldsymbol{y}_t = (y_{1t}, ..., y_{pt})'$ is a $p \times 1$ vector of observable variables, $\boldsymbol{u}_t$ are serially uncorrelated errors with mean zero and time-varying covariance matrix $\boldsymbol{\Sigma}_t$, $\boldsymbol{f}$ is a vector of functions, and $\boldsymbol{x}_t$ is a vector of variables which may contain lags of the dependent variable $\boldsymbol{y}_t$, the intercept, deterministic dummy variables and exogenous variables. The model (2.1) may be nonlinear.

The covariance $\boldsymbol{\Sigma}_t$ is a symmetric positive definite matrix, conditional on all the information available at time t. I make the following assumption:

Assumption 2.1. *(Spectral decomposition) The time-varying conditional covariance matrix $\boldsymbol{\Sigma}_t$ can be decomposed as follows:*

$$\boldsymbol{\Sigma}_t = \boldsymbol{P}\boldsymbol{\Lambda}_t\boldsymbol{P}', \tag{2.2}$$

where the time-invariant matrix $\boldsymbol{P}$ *satisfies* $\boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I}_p$, $\boldsymbol{I}_p$ *being an identity matrix, and* $\boldsymbol{\Lambda}_t = \mathrm{diag}(\lambda_{1t}, ..., \lambda_{pt})$ *whose elements are all positive.*

Assumption 2.1 implies that the conditional error covariance matrix is time-varying in the way that the eigenvectors remain constant while the corresponding eigenvalues can vary over time.

For better understanding of this assumption, consider the random variable $\boldsymbol{\eta} \sim N(\mathbf{0}, \boldsymbol{I}_p)$ in $\mathbb{R}^p$, where the covariance $\boldsymbol{I}_p$ is an identity matrix. Any vector Gaussian distribution $N(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ has a spectral decomposition $\boldsymbol{\Sigma} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'$, can be represented through left-multiplying $\boldsymbol{\eta}$ by $\boldsymbol{P}\boldsymbol{\Lambda}^{1/2}$, where $\boldsymbol{P}$ is a rotation matrix in $\mathbb{R}^p$ and $\boldsymbol{\Lambda}$ is a scaling matrix. Note that the ordering of the column vectors in $\boldsymbol{P}$ is not unique, but each element in the diagonal of $\boldsymbol{\Lambda}$ corresponds to a certain column vector in $\boldsymbol{P}$. Assumption 2.1 implies that, for a certain ordering of the column vectors in $\boldsymbol{P}$, the rotation matrix $\boldsymbol{P}$ is constant over time, whereas the scaling matrix $\boldsymbol{\Lambda}$ can be time-varying.

Assumption 2.1 is a sufficient condition for the matrix exponential modelling of the multivariate volatilities, see for example the matrix exponential GARCH model in Kawakatsu (2006). The matrix exponential transformation making the eigenvalues to be exponential functions has the advantage that it ensures positive definiteness of the covariance matrix. It is not only useful for the multivariate GARCH models, but allows for many other possible forms as well.

Assumption 2.1 is also a sufficient condition for the existence of a matrix (several ordered linear combinations) such that left-multiplying both sides of (2.1) by this matrix (the ordered linear combinations) removes the contemporaneous correlation. Thus, the assumption is applicable in the structural vector models, which identify the shocks and allow for heteroskedasticity. See, for example, Lanne and Lütkepohl (2008) and Lanne et al. (2009). In this case, the vector of eigenvalues is simply the vector of variances of the structural model with identifed shocks, and hence may be heteroskedastic. Compared to Assumption 2.1, although the CCC assumption implies a constant correlation structure, the correlation between errors cannot be removed if the variances are time-varying.

Note that Assumption 2.1 is different from the constant conditional correlation (CCC) decomposition in Bollerslev (1990). Under the CCC assumption, the contemporaneous correlation structure of the errors is assumed time-invariant, while under Assumption 2.1, both the correlations and the variances of the error vector are time-varying.

Under Assumption 2.1, the log-likelihood function for observation $t = 1, ..., T$ based on vector Gaussian distributed errors is:

$$\begin{aligned}\log L_t &= c - \frac{1}{2}\log|\boldsymbol{\Sigma}_t| - \frac{1}{2}\boldsymbol{u}_t'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{u}_t \\ &= c - \frac{1}{2}\log|\boldsymbol{\Lambda}_t| - \frac{1}{2}\boldsymbol{w}_t'\boldsymbol{\Lambda}_t^{-1}\boldsymbol{w}_t \\ &= c - \frac{1}{2}\sum_{i=1}^{p}\left(\log\lambda_{it} + w_{it}^2\lambda_{it}^{-1}\right), \end{aligned} \tag{2.3}$$

where $\boldsymbol{w}_t = \boldsymbol{P}'\boldsymbol{u}_t = (w_{1t}, ..., w_{pt})'$ contains the errors after a certain rotation $\boldsymbol{P}$. When the error vector is not Gaussian, (2.3) is the *quasi Gaussian log-likelihood function* for observation t. Let $\boldsymbol{\varphi}_i$ be the vector of parameters in λ_{it} for $i = 1, ..., p$, and define $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1', ..., \boldsymbol{\varphi}_p')'$. Let $\boldsymbol{\phi}$ be the vector of the parameters in the conditional mean. Consequently, $\boldsymbol{\theta} = (\boldsymbol{\varphi}', \boldsymbol{\phi}')'$ is the vector of all parameters except the ones in $\boldsymbol{P}$. Under Assumption 2.1, the matrix $\boldsymbol{P}$ does not contribute to maximizing the log-likelihood function, but serves to identify $\boldsymbol{\Lambda}_t$ in $\boldsymbol{\Sigma}_t = \boldsymbol{P}\boldsymbol{\Lambda}_t\boldsymbol{P}'$. Therefore $\boldsymbol{\theta}$ excludes $\boldsymbol{P}$.

Based on Assumption 2.1, I make the following assumption:

Assumption 2.2. *The time-varying components are functions of x_t, $\lambda_{it} = h_i(x_t)$, $i = 1, ..., p$, where $h_i(x_t)$ is a positive function. The function $h_i(x_t)$ is at least second-order differentiable almost everywhere. Furthermore, the argument $x_t = \boldsymbol{\varphi}_i'\boldsymbol{z}_{it}$, where $\boldsymbol{\varphi}_i$ is a vector of parameters and $\boldsymbol{z}_{it}$ is a vector of predetermined variables with respect to the information available at time t.*

Assuming h_i to be at least second-order differentiable ensures the existence of the corresponding information matrix. Assumption 2.2 allows for a wide variety of error covariance structures. The exponential function $h_i(x_t) = \exp(x_t)$ is one possibility, which ensures the function is strictly positive-valued. Although the functional form of h_i is quite flexible, it does not play a role in deriving the test statistic. In the following, h_i' is the first-order and h_i'' the second-order derivative of the the function h_i with respect to x_t. It can be see later in the following sections, the elements of $\boldsymbol{z}_{it}$ are determined by the specification of heteroskedasticity.

3 LM test statistic

Our focus is on testing the constancy of the whole covariance matrix when the alternative is characterized by Assumptions 2.1 and 2.2. The null hypothesis to be tested is thus:

$$H_0: \quad \lambda_{it} = \lambda_i, \quad i = 1, ..., p. \tag{3.1}$$

or, put differently,

$$H_0: \quad \boldsymbol{\varphi} = (\varphi_{i0}, 0, ..., 0)', \tag{3.2}$$

where φ_{i0} is the coefficient of the intercept $z_{i0} = 1$. That is, the vector $\boldsymbol{\varphi}$ has only one non-zero element under the null hypothesis.

The tests to be considered here are Lagrange-multiplier tests. This family of tests has the advantage that there is no need to estimate the model under the alternative hypothesis. Consequently, I only have to estimate the model under the null hypothesis (3.1). The log-likelihood function for observation t appears in (2.3). I define the average score vector and the average information matrix of the quasi log-likelihood function as follows:

$$\boldsymbol{s}(\boldsymbol{\theta}) = T^{-1}\sum_{t=1}^{T}\frac{\partial \log L_t}{\partial \boldsymbol{\theta}} \tag{3.3}$$

$$\boldsymbol{I}(\boldsymbol{\theta}) = -T^{-1}\sum_{t=1}^{T}\mathrm{E}\left[\frac{\partial^2 \log L_t}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right], \tag{3.4}$$

where $\log L_t$ has been defined in (2.3). Let $\tilde{\boldsymbol{\theta}}$, $\tilde{\boldsymbol{\phi}}$, $\tilde{\boldsymbol{\varphi}}$ and $\tilde{\boldsymbol{\varphi}}_i$, $i = 1, ..., p$, be the estimates of the parameters under the null hypothesis. Thus, I have the average score vector $\mathbf{s}(\tilde{\boldsymbol{\theta}})$ and the average information matrix $\mathbf{I}(\tilde{\boldsymbol{\theta}})$ evaluated under the null hypothesis. The LM test statistic takes the form

$$LM = T\mathbf{s}(\tilde{\boldsymbol{\theta}})'\boldsymbol{I}^{-1}(\tilde{\boldsymbol{\theta}})\mathbf{s}(\tilde{\boldsymbol{\theta}}) \tag{3.5}$$

and is asymptotically χ^2 distributed, with the the degrees of freedom equal to the number of restrictions, when the null hypothesis is valid.

The information matrix $\boldsymbol{I}(\tilde{\boldsymbol{\theta}})$ evaluated under the null hypothesis is a block diagonal matrix. Thus, I define the corresponding blocks of the average score vector and of the average information matrix of the quasi log-likelihood function as follows:

$$\boldsymbol{s}_{\varphi}(\boldsymbol{\theta}) = T^{-1}\sum_{t=1}^{T}\frac{\partial \log L_t}{\partial \boldsymbol{\varphi}} \tag{3.6}$$

$$\boldsymbol{I}_{\varphi}(\boldsymbol{\theta}) = -T^{-1}\sum_{t=1}^{T}\mathrm{E}\left[\frac{\partial^2 \log L_t}{\partial \boldsymbol{\varphi}\partial \boldsymbol{\varphi}'}\right]. \tag{3.7}$$

Under Assumption 2.2, the Lagrange-multiplier test (3.5) can be equivalently applied as follows:

$$LM = T\mathbf{s}_{\varphi}(\tilde{\boldsymbol{\theta}})'\boldsymbol{I}_{\varphi}^{-1}(\tilde{\boldsymbol{\theta}})\mathbf{s}_{\varphi}(\tilde{\boldsymbol{\theta}}), \tag{3.8}$$

see Godfrey (1978), Breusch and Pagan (1978) and Breusch and Pagan (1980) for details. I have the following theorem:

Theorem 3.1. *Under Assumption 2.1 and 2.2, the corresponding blocks of the average score vector and of the average information matrix of the quasi Gaussian log-likelihood* *based in* (2.3) *are*

$$\boldsymbol{s}_{\varphi}(\tilde{\boldsymbol{\theta}}) = (2T)^{-1}\sum_{t=1}^{T}\left[\tilde{\zeta}_1\tilde{g}_{1t}\tilde{\boldsymbol{z}}'_{1t}, \ \dots \ , \ \tilde{\zeta}_p\tilde{g}_{pt}\tilde{\boldsymbol{z}}'_{pt}\right]' \tag{3.9}$$

$$\boldsymbol{I}_{\varphi,i}(\tilde{\boldsymbol{\theta}}) = (2T)^{-1}\sum_{t=1}^{T}\tilde{\zeta}_i^2\mathrm{E}\left[\tilde{\boldsymbol{z}}_{it}\tilde{\boldsymbol{z}}'_{it}\right], \tag{3.10}$$

where $\tilde{\zeta}_i = \tilde{h}'_i\tilde{\lambda}_i^{-1}$, $\tilde{g}_{it} = \tilde{w}_{it}^2/\tilde{\lambda}_i - 1$, h'_i *is the scalar first-order derivative of the positive function* h_i, *and they are evaluated under the null hypothesis of constancy. The LM test statistic* (3.5) *has the following form:*

$$LM = \frac{1}{2}\sum_{i=1}^{p}\left[\left(\sum_{t=1}^{T}\tilde{g}_{it}\tilde{\boldsymbol{z}}'_{it}\right)\left(\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{it}\tilde{\boldsymbol{z}}'_{it}\right)^{-1}\left(\sum_{t=1}^{T}\tilde{g}_{it}\tilde{\boldsymbol{z}}_{it}\right)\right]. \tag{3.11}$$

Under regularity conditions, the LM statistic in (3.11) *is asymptotically* χ^2 *distributed with degrees of freedom equal to the number of restrictions.*

Proof. See Appendix A. □

$\tilde{\boldsymbol{z}}_{it} = \boldsymbol{z}_{it}$ only if $\boldsymbol{z}_{it}$ is observable, or $\tilde{\boldsymbol{z}}_{it}$ is the ML estimate of $\boldsymbol{z}_{it}$ under the null hypothesis. In the following section, it is shown that $\tilde{\boldsymbol{z}}_{it}$ may contain the transformed error term $\tilde{\boldsymbol{w}}_t$ estimated from the restricted model. The number of restrictions is the number of zeros in (3.2). Moreover, it is seen from (3.11) that the general positive function h_i and its derivative have been cancelled out as the argument of h_i is a constant under H_0. There is thus no need to uniquely define the functional form of h_i when setting up the test.

Consider the fact that $T^{-1}\sum_{t=1}^{T}\tilde{g}_{it}^2$ converges to 2 in probability under the null hypothesis and that the errors are Gaussian. Denote

$$R_i^2 = \left(\sum_{t=1}^{T}\tilde{g}_{it}^2\right)^{-1}\left(\sum_{t=1}^{T}\tilde{g}_{it}\tilde{\boldsymbol{z}}'_{it}\right)\left(\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{it}\tilde{\boldsymbol{z}}'_{it}\right)^{-1}\left(\sum_{t=1}^{T}\tilde{g}_{it}\tilde{\boldsymbol{z}}_{it}\right), \tag{3.12}$$

for $i = 1, ..., p$. Computing R_i^2 is quite easy. After obtaining the sequence $\{\tilde{g}_{it}\}_{t=1}^{T}$ for $i = 1, ..., p$, run a simple auxiliary regression of $\tilde{g}_{it}$ on $\tilde{\boldsymbol{z}}_{it}$ and collect the residuals. Denote the SSG_i the sum of squared $\tilde{g}_{it}$, and the RSS_i the corresponding residual sum of squares in the auxiliary regression. It follows that

$$R_i^2 = \frac{SSG_i - RSS_i}{SSG_i}. \tag{3.13}$$

Thus,

$$\sum_{i=1}^{p}T\frac{SSG_i - RSS_i}{SSG_i} = \sum_{i=1}^{p}TR_i^2 \tag{3.14}$$

is asymptotically equivalent to the LM statistic (3.5). The test can be carried out as follows:

- Estimate the vector model (2.1) under the null hypothesis of constant covariances. Collect the residuals $\tilde{\boldsymbol{u}}_t$, $t = 1, ..., T$. Compute the empirical covariance matrix $\tilde{\boldsymbol{\Sigma}}$, and the eigenvalue decomposition $\tilde{\boldsymbol{\Sigma}} = \tilde{\boldsymbol{P}}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{P}}'$, where $\tilde{\boldsymbol{\Lambda}} = \text{diag}(\tilde{\lambda}_1, ..., \tilde{\lambda}_p)$.

- Compute the transformed residuals $\tilde{\boldsymbol{w}}_t = \tilde{\boldsymbol{P}}'\tilde{\boldsymbol{u}}_t$, and $\tilde{g}_{it} = \tilde{w}_{it}^2/\tilde{\lambda}_i - 1$, for $t = 1, ..., T$, $i = 1, ..., p$.

- For each equation, regress $\tilde{g}_{it}$ on $\tilde{\boldsymbol{z}}_{it}$ and compute the corresponding TR_i^2. Compute the LM test $\sum_{i=1}^{p} TR_i^2$.

In the next section, I will discuss different specifications of $\tilde{\boldsymbol{z}}_{it}$.

4 Specifications for heteroskedastic residuals

There are a number of possible specifications for heteroskedasticity in the errors. I will consider three useful covariance specifications against the homoskedasticity in the following. They have already been considered in Eklund and Teräsvirta (2007), but as already mentioned, the decomposition of $\boldsymbol{\Sigma}_t$ is different from theirs.

The first time-varying variance specification (White specification), proposed in a single-equation case by White (1980) as an alternative to homoskedasticity, is obtained by defining:

$$\lambda_{it} = h_i(\sigma_i^2 + \boldsymbol{\delta}_i'\text{vech}(\boldsymbol{x}_t\boldsymbol{x}_t')) \tag{4.1}$$

where vech() represents the half-vectorization which collects the lower triangular elements of a symmetric matrix; $\boldsymbol{\delta}_i$, $i = 1, ..., p$, are $q(q+1)/2$-dimension column vectors of parameters; and the column vector $\boldsymbol{x}_t$ defined in (2.1), has dimension q. The null hypothesis of a constant covariance matrix in (3.1) is

$$H_0: \quad \boldsymbol{\delta}_i = \mathbf{0}, \quad i = 1, ..., p. \tag{4.2}$$

The corresponding number of degrees of freedom of the LM test is $q^2(q+1)/2$.

The second variance specification (ARCH specification) is obtained by defining

$$\lambda_{it} = h_i(\sigma_i^2 + \sum_{j=1}^{q} \alpha_{ij} w_{i,t-j}^2) \tag{4.3}$$

Note that we use the transformed error $w_{i,t-j}$ instead of $u_{i,t-j}$, because $\boldsymbol{\Lambda}_t$ is the covariance matrix of $\boldsymbol{w}_t$. The null hypothesis corresponding to (3.1) is

$$H_0: \quad \alpha_{ij} = 0, \quad i = 1, ..., p, \quad j = 1, ..., q. \tag{4.4}$$

The corresponding number of degrees of freedom of the LM test is pq.

The third (smooth transition) specification is obtained by assuming $\boldsymbol{u}_t$ to be a heteroskedastic error term with a smoothly changing covariance matrix, that is,

$$\boldsymbol{\Sigma}_t = \mathrm{E}_t(\boldsymbol{u}_t\boldsymbol{u}_t') = \boldsymbol{\Sigma}_1 + G(s_t)\boldsymbol{\Sigma}_2 \tag{4.5}$$

where $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are symmetric matrices, and $G(s_t)$ is called a transition function whose value is controlled by an observable transition variable s_t. We assume that the transition variable s_t is a weakly stationary random variable, but it can also be a time trend.

Assume that the transition function is a real-valued, bounded, monotonically increasing and at least second-order differentiable function, e.g. a logistic function:

$$G(s_t) = G(s_t; \gamma, c) = (1 + \exp(-\gamma(s_t - c)))^{-1} \tag{4.6}$$

where the parameter $\gamma > 0$ determines the smoothness of the transition, and c is the location parameter. It is seen from (4.5) and (4.6) that the covariance matrix changes smoothly from $\boldsymbol{\Sigma}_1$ to $\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$ as a function of s_t. Both $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2$ must be positive definite matrices.

Following Assumption 2.1 and Equation (4.5), write $\boldsymbol{\Sigma}_1 = \boldsymbol{P}\boldsymbol{\Lambda}_1\boldsymbol{P}'$ and $\boldsymbol{\Sigma}_2 = \boldsymbol{P}\boldsymbol{\Lambda}_2\boldsymbol{P}'$. It is obvious that

$$\boldsymbol{\Sigma}_t = \boldsymbol{P}\left(\boldsymbol{\Lambda}_1 + G(s_t)\boldsymbol{\Lambda}_2\right)\boldsymbol{P}' = \boldsymbol{P}\boldsymbol{\Lambda}_t\boldsymbol{P}' \tag{4.7}$$

where

$$\begin{aligned}
\boldsymbol{\Lambda}_1 &= \mathrm{diag}\left(\lambda_{11}, ..., \lambda_{1p}\right) \\
\boldsymbol{\Lambda}_2 &= \mathrm{diag}\left(\lambda_{21}, ..., \lambda_{2p}\right) \\
\boldsymbol{\Lambda}_t &= \mathrm{diag}\left(\lambda_{1t}, ..., \lambda_{pt}\right) \\
\lambda_{it} &= \lambda_{1i} + G(s_t)\lambda_{2i}, \\
\text{s.t.} \quad & \lambda_{1i} > 0, \quad \lambda_{1i} + \lambda_{2i} > 0,
\end{aligned} \tag{4.8}$$

for $i = 1, ..., p$.

The null hypothesis under the specification (4.6), (4.7) and (4.8) is: $H_0 : \gamma = 0$. It is seen that under the null hypothesis $G(s_t) = 1/2$ and hence the parameters in $\boldsymbol{\Sigma}_2$ are not identified. In order to solve the identification problem, Luukkonen et al. (1988) suggested to approximate the transition function (4.6) by its first-order Taylor expansion around $\gamma = 0$. This means writing

$$\lambda_{it} = \lambda_{1i} + (as_t + b + r_t)\lambda_{2i} \approx \lambda_{0i}^* + \lambda_{1i}^* s_t, \tag{4.9}$$

where a and b are constants, and r_t is the remainder. In this case, the null hypothesis is:

$$H_0 : \quad \lambda_{1i}^* = 0, \quad i = 1, ..., p. \tag{4.10}$$

The corresponding number of degrees of freedom of the LM test is p.

The transition function can also be approximated by a higher-order Taylor expansion. This may often increase the power of the test. For a Taylor expansion of order $N > 1$, (4.9) can be extended to:

$$\lambda_{it} \approx \sum_{n=0}^{N} \lambda_{ni}^{*} s_t^n. \tag{4.11}$$

The null hypothesis is

$$H_0: \quad \lambda_{ni}^{*} = 0, \quad i = 1, ..., p, \quad n = 1, ..., N. \tag{4.12}$$

In this case, the number of degrees of freedom of the LM test is pN. However, in the following the focus will be on the first-order approximation to the logistic function (4.6).

5 Finite sample properties of the test

When investigating the properties of a classical test statistic, two aspects are of prime importance. First, I have to check whether the empirical size of the test (the probability of rejecting the null when it is true) is close to the nominal size (used to calculate the asymptotic critical values) at typical sample sizes. Given that empirical size is a reasonable approximation to the nominal size, I then have to investigate the empirical power of the test (the probability of rejecting the null when it is false) for a number of different alternative hypotheses.

In order to investigate the size and power properties of the test in finite samples, I conduct a series of Monte Carlo simulations. I consider the bivariate case $p = 2$, the trivariate case $p = 3$ and a high-dimensional case $p = 5$. The data generating process is a special case of (2.1):

$$y_{i,t} = 0.8 y_{i,t-1} + u_{i,t}, \quad i = 1, ..., p, \tag{5.1}$$

where the error term $u_{i,t}$ is independent and identically distributed. This is a simple design in the sense that the variables in the VAR model are only linked through the covariance matrix. The finite sample sizes I investigate in the size experiments are $T = 100$ and $T = 500$. The autoregressive model in (5.1) is exactly the same as the one in Eklund and Teräsvirta (2007), which makes it easy to compare the size properties of the two tests under the null hypothesis of constant covariance matrix. Thus, I will not repeat their size experiments here.

I report my results using the size discrepancy and power plots recommended by Davidson and MacKinnon (1998). The number of replications of the Monte Carlo simulations is $N = 10000$.

For space reasons, only a fraction of the results are shown. The remaining ones are available at:
`http://creates.au.dk/research/research-papers/supplementary-downloads/`.

5.1 Size experiments

In investigating the finite sample size behaviour of the test statistics, I assume $u_{i,t}$ to be either i.i.d. Gaussian or $t(5)$ distribution in the basic data generating process (5.1). The LM type tests are derived assuming that the errors are Gaussian. The t(5)-distribution contradicts this assumption, but I consider it to see what kind of an effect a fat-tailed error distribution may have on the empirical size of the test.

In the bivariate case, the covariance matrix has the structure:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{5.2}$$

where $\rho = 0.9$, 0 and -0.9. Eklund and Teräsvirta (2007) used the same design for the bivariate case. In the trivariate case, The covariance matrix has the structure:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}, \tag{5.3}$$

where $\rho = 0.9$ and 0. For the high-dimensional case $p = 5$, I only report the results from $\rho = 0$ for space reasons. It will be seen that for $p \geq 3$ the correlation may affect the size properties. In the following, I conduct the three LM tests by setting

White specification: $\boldsymbol{z}_{it} = (1, \mathrm{vech}(\boldsymbol{x}_t\boldsymbol{x}_t')')'$, where $\boldsymbol{x}_t = \boldsymbol{y}_{t-1}$;

ARCH(5) specification: $\tilde{\boldsymbol{z}}_{it} = (1, \tilde{w}_{i,t-1}^2, ..., \tilde{w}_{i,t-5}^2)'$, where $\tilde{w}_{i,t-q}$, $q = 1, ..., 5$ are estimates of the transformed errors under the null hypothesis of constancy;

Smooth transition specification: $\boldsymbol{z}_{it} = (1, \tau)'$, where $\tau = t/T$.

In the smooth transition specification I choose $s_t = \tau$ to be the transition variable. This is because I just focus on testing whether the covariance matrix changes over time.

I compare the size properties of the three LM tests when $\rho = 0$. I report the results from the case $p = 5$ in Figures 1 to 4. My finding is that the empirical distributions of all three tests converge towards their limiting distribution when T increases. The test against the White specification over-rejects in all the cases, especially when the errors are $t(5)$, whereas the test against the ARCH(5) specification seems to under-reject. It is seen from Figures 3 to 4 that when the errors are $t(5)$, the tests against the White specification and the ARCH(5) specification have greater size distortion than test against the smooth transition specification. The latter test performs well in almost all cases.

To see whether the correlation ρ plays a role, I conducted experiments to investigate the actually sizes of the tests with respect to different sample sizes. The sample sizes I used for these experiments are $T = 25$, 50, 100, 250, 500 and 1000, respectively. Not all the plots are given here in the paper, but the reader can reach them by visiting the link given above.

The finding is that the correlation hardly affects the size properties in the bivariate case, but it may do so in trivariate case. The results from the White specification show slightly different convergence rate of the tests. Figure 5 shows that the test against the White specification converges towards its limiting distribution a bit faster when there is no correlation in errors in finite sample case from $T = 25$ to 50. It is the same for the case when the errors are $t(5)$, see Figure 6. There is no clear sign that the correlation affects the size of the test against the ARCH(5) specification. This may be due to the fact that the ARCH specification (4.3) does not allow for any cross-effects between different equations. Moreover, the correlation does not affect the size of the test against the smooth transition specification. It is seen that this test is free from size distortion in almost all cases considered.

5.2 Power experiments

In power simulations I assume that the data generating process $h_i(\boldsymbol{\varphi}_i' \boldsymbol{z}_{it}) = \boldsymbol{\varphi}_i' \boldsymbol{z}_{it}$ for simplicity. I only consider the bivariate case $p = 2$ and the high-dimensional case $p = 5$. The data generating process is still (5.1), but now the covariance matrix will change over time. The errors are drawn from the conditional vector Gaussian distribution. I will conduct the following three kinds of power simulations.

5.2.1 Power simulations under Assumption 2.1

In this case, the covariance matrix of the errors will change under Assumption 2.1. That is,

$$\boldsymbol{\Sigma}_t = \boldsymbol{P} \boldsymbol{\Lambda}_t \boldsymbol{P}'. \tag{5.4}$$

First, I consider the case that the covariance matrix changes once through time and the transition is threshold-like. The threshold point is at $T/2$, that is, $c = 0.5$. Then I consider the special case that the covariance matrix changes once through time but the transition is rather smooth. The smooth function takes the form (4.6) where $\gamma = 12$ and $c = 0.5$. The transition variable is the normalized time $s_t = t/T$. Figures 7 and 8 show the results from the experiments where $p = 5$, $T = 500$. The former one has abrupt change, while the latter one has smooth change. The smooth change implies that Λ_t is actually changing everywhere over the time horizon, not only once. It is seen that the test against the smooth transition specification performs always the best in both the case in which the transition is abrupt and the case in which the transition is rather smooth. This is not surprising because the DGP is just the smooth transition specification. Moreover, the sample size $T = 500$ is big enough to see that the power of the other tests do not increase fast, and this implies that the p-values from these tests may suggest the specification of the covariance change.

Next I investigate the power of the tests when λ_{it} in the covariance matrix evolves through time using the ARCH specification (4.3). More specifically, I assume an ARCH(2) process for all eigenvalues with parameters $\sigma_i^2 = 1$, $\alpha_{i1} = 0.25$ and $\alpha_{i2} = 0.2$. Figure 9 shows the results from the typical case $p = 5$, $T = 500$. It is not surprising that the test against the ARCH specification outperforms the others. However, it is seen that the other tests have just a little power even when the sample size is 500.

Finally, I consider the case when λ_{it} in the covariance matrix evolves through time using the White specification (4.1). The parameters are $\sigma_i^2 = 1$ and $\boldsymbol{\delta}_i = (1, ..., 1)'$ is a $p(p+1)/2 \times 1$ vector. I still report the results from the typical case $p = 5$, $T = 500$, see Figure 10. This time the test against the White specification is the best performer. It is seen that the other tests have very little power in this case even in large samples.The tests can be very informative and suggest the specification of the covariance change, especially when the true DGP follows ARCH or White specifications.

5.2.2 Power simulations when Assumption 2.1 is violated

It is important to investigate the consequences of violating Assumption 2.1. Since Assumption 2.1 is very restrictive, one may argue that if the null hypothesis of constant covariance matrix is rejected, it would be difficult without any further investigation to distinguish between a rejection due to time-varying $\boldsymbol{\Lambda}$, time-varying $\boldsymbol{P}$ or a combination of the two. Here I investigate the case that

$$\boldsymbol{\Sigma}_t = \boldsymbol{P}_t \boldsymbol{\Lambda} \boldsymbol{P}_t'. \tag{5.5}$$

The data generating process takes the form (5.1). $\boldsymbol{\Lambda} = \text{diag}(0.50, 0.40)$ for $p = 2$, and $\boldsymbol{\Lambda} = \text{diag}(0.50, 0.40, 0.30, 0.25, 0.20)$ for $p = 5$.

I consider the case that $\boldsymbol{P}$ changes once through time. The transition is threshold-like, and the threshold point is at $T/2$. The way to determine the values in the matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ is tricky. Let $\boldsymbol{U}_i$, $i = 1, 2$, be a $p \times p$ matrix whose elements are a sample of independent draws from a standard Gaussian distribution. Let $\boldsymbol{P}_i$ be the eigenvectors of

$\boldsymbol{U}_i\boldsymbol{U}_i'$. Then, the rotation (orthonormal basis) $\boldsymbol{P}_i$ are uniformly distributed over the set of all rotation matrices (orthonormal bases).

The case that $\boldsymbol{P}$ changes smoothly through time will also be considered. I use the way mentioned in the previous paragraph to pick $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$. However, the conditional covariance matrix at time t should be computed as follows:

$$\boldsymbol{\Sigma}_t = \boldsymbol{P}_1\boldsymbol{\Lambda}\boldsymbol{P}_1'(1 - G(\tau;\gamma,c)) + \boldsymbol{P}_2\boldsymbol{\Lambda}\boldsymbol{P}_2'G(\tau;\gamma,c) \tag{5.6}$$

where G has been defined in (4.6), $\tau = t/T$, $\gamma = 12$ and $c = 0.5$.

Figures 11 to 13 depict the rejection frequencies for several typical cases. I also compare these tests with the test in Eklund and Teräsvirta (2007), since Assumption 2.1 is violated. The results in Figure 11 show that in low-dimensional case ($p = 2$) all the tests from Assumption 2.1 have little power except the test in Eklund and Teräsvirta (2007). It is interesting to see that, when the dimension increases, the power of my test against smooth transition specification approaches that of the test in Eklund and Teräsvirta (2007), see Figure 12. And Figure 13 gives a ranking of the power performance for $p = 5$ when the sample size increases. The conclusion is that my tests can detect the change in $\boldsymbol{\Lambda}$, but it is not very sensitive to the change in the rotation matrix $\boldsymbol{P}$.

5.2.3 Power simulations under the constant conditional correlation assumption

I also simulated the situation when the constant conditional correlation (CCC) assumption is satisfied. Under the CCC assumption, the time-varying covariance matrix can be decomposed as follows:

$$\boldsymbol{\Sigma}_t = \boldsymbol{D}_t\boldsymbol{Q}\boldsymbol{D}_t', \tag{5.7}$$

where

$$\boldsymbol{D}_t = \mathrm{diag}(\omega_{1t}^{1/2}, ..., \omega_{pt}^{1/2}) \tag{5.8}$$

$$\tag{5.9}$$

is a diagonal matrix of error standard deviations, and $\boldsymbol{Q} = [\rho_{ij}]$ is the corresponding correlation matrix. The value of ρ_{ij} is chosen in following way. Let $\boldsymbol{U}$ be a $p \times p$ matrix whose elements are a sample of independent draws from a standard Gaussian distribution, and denote u_{ij} the element of $\boldsymbol{U}_i\boldsymbol{U}_i'$ where i is the row number and j is the column number. $\rho_{ij} = u_{ij}/\sqrt{u_{ii}u_{jj}}$.

I consider first that the error variances ω_{it} change once through time and the transition is threshold-like. The threshold point is as $T/2$, that is, $c = 0.5$. Second, I consider the case in which the error variances ω_{it} change smoothly through time from ω_{i1} to ω_{i2}. The transition function takes the form (4.6). Let $s_t = t/T$, $\gamma = 12$ and $c = 0.5$.

It is seen from the results that the tests derived from Assumption 2.1 have very satisfactory power, though the model is misspecified. It is surprising to see that the test against the smooth transition specification has the best performance not only in the threshold case but in the smooth transition case as well. It performs even better than the test in Eklund and Teräsvirta (2007) especially in the high-dimensional finite sample case, see Figures 14.

6 Concluding remarks

In this work, I develop a test of constancy of the error covariance matrix against the alternative that the covariance matrix changes over time. The test is based on the spectral decomposition of the covariance matrix. This implies that the conditional error covariance is time-varying in the way that the eigenvectors remain constant through time and only the corresponding eigenvalues are time-varying. There exist linear combinations which make error vectors in the corresponding structural vector model contemporaneously uncorrelated but still heteroskedastic. I design a family of LM tests against the alternative hypothesis that the errors are time-varying and follow parametric specifications.

Three specifications are considered. They are: the White specification which generalizes the heteroskedasticity test of White (1980), the ARCH specification which generalizes the test against autoregressive conditional heteroskedasticity of Engle (1982) and the smooth transition specification which generalizes the test against smoothly changing variance of Medeiros and Veiga (2003). The test of constancy of the error covariance matrix is very easy to implement and use. From the simulation experiments it is seen that the test has satisfactory size and power properties even in vector models.

References

Bollerslev, T.: 1990, Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model, *Review of Economics and Statistics* **72**, 498–505.

Box, G. and Jenkins, G.: 1970, Time series analysis: Forecasting and control, *San Francisco: Holden-Day* .

Breusch, T. S. and Pagan, A. R.: 1978, A simple test for heteroscedasticity and random coefficient variation, *Econometrica* **46**, 1287–1294.

Breusch, T. S. and Pagan, A. R.: 1980, The Lagrange multiplier test and its applications to model specification in econometrics, *Review of Economic Studies* **47**, 239–253.

Davidson, R. and MacKinnon, J. G.: 1998, Graphical methods for investigating the size and power of hypothesis tests, *The Manchester School of Economic & Social Studies* **66**, 1–26.

Eklund, B. and Teräsvirta, T.: 2007, Testing constancy of the error covariance matrix in vector models, *Journal of Econometrics* **140**, 753–780.

Engle, R. F.: 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kindom inflation, *Econometrica* **50**, 987–1007.

Godfrey, L. G.: 1978, Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables, *Econometrica* **46**, 1293–1302.

Kawakatsu, H.: 2006, Matrix exponential GARCH, *Journal of Econometrics* **134**, 95–128.

Lanne, M. and Lütkepohl, H.: 2008, Stock Prices and Economic Fluctuations: A Markov Switching Structural Vector Autoregressive Analysis, *European University Institute, Economics Working Papers* (ECO2008/29).

Lanne, M., Lütkepohl, H. and Maciejowska, K.: 2009, Structural Vector Autoregressions with Markov Switching, *European University Institute, Economics Working Papers* (ECO2009/06).

Lütkepohl, H.: 2004, Vector autoregressive and vector error correction models, *in Applied Time Series Econometrics, H. Lütkepohl and M. Krätzig, eds.* pp. 86–158. Cambridge, Cambridge University Press.

Luukkonen, R., Saikkonen, P. and Teräsvirta, T.: 1988, Testing linearity against smooth transition autoregressive models, *Biometrika* **75**, 491–499.

Medeiros, M. C. and Veiga, A.: 2003, Diagnostic checking in a flexible nonlinear time series model, *Journal of Time Series Analysis* **24**, 461–482.

Teräsvirta, T. and Yang, Y.: 2014a, Linearity and misspecification tests for vector smooth transition regression models, *Research Paper 2014-04*, CREATES, Aarhus University.

Teräsvirta, T. and Yang, Y.: 2014b, Specification, estimation and evaluation of vector smooth transition autoregressive models with applications, *Research Paper 2014-08*, CREATES, Aarhus University.

White, H.: 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* **48**, 817–838.

A Proof for Theorem 3.1

Proof. Based in (2.3), assume that $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1', ..., \boldsymbol{\varphi}_p')'$ and $\lambda_{it} = h_i(\boldsymbol{\varphi}_i'\boldsymbol{z}_{it})$, $i = 1, ..., p$. I have

$$\frac{\partial \log L_t}{\partial \boldsymbol{\varphi}_i} = \frac{\partial \log L_t}{\partial \lambda_{it}} \frac{\partial \lambda_{it}}{\partial \boldsymbol{\varphi}_i} \tag{A.1}$$

$$\frac{\partial^2 \log L_t}{\partial \boldsymbol{\varphi}_i \partial \boldsymbol{\varphi}_i'} = \frac{\partial^2 \log L_t}{\partial \lambda_{it}^2} \frac{\partial \lambda_{it}}{\partial \boldsymbol{\varphi}_i} \frac{\partial \lambda_{it}}{\partial \boldsymbol{\varphi}_i'} + \frac{\partial \log L_t}{\partial \lambda_{it}} \frac{\partial^2 \lambda_{it}}{\partial \boldsymbol{\varphi}_i \partial \boldsymbol{\varphi}_i'} \tag{A.2}$$

$$\frac{\partial^2 \log L_t}{\partial \boldsymbol{\varphi}_i \partial \boldsymbol{\varphi}_j'} = \mathbf{0} \quad \text{for} \quad i \neq j. \tag{A.3}$$

Furthermore, in (A.1), (A.2) and (A.3), I have

$$\frac{\partial \log L_t}{\partial \lambda_{it}} = \frac{1}{2\lambda_{it}} g_{it} \tag{A.4}$$

$$\frac{\partial^2 \log L_t}{\partial \lambda_{it}^2} = \frac{1}{2\lambda_{it}^2} \left(1 - 2w_{it}^2 \lambda_{it}^{-1}\right) \tag{A.5}$$

$$\frac{\partial \lambda_{it}}{\partial \boldsymbol{\varphi}_i} = h_i' \boldsymbol{z}_{it} \tag{A.6}$$

$$\frac{\partial^2 \lambda_{it}}{\partial \boldsymbol{\varphi}_i \partial \boldsymbol{\varphi}_i'} = h_i'' \boldsymbol{z}_{it} \boldsymbol{z}_{it}', \tag{A.7}$$

where $g_{it} = w_{it}^2/\lambda_{it} - 1$, h_i' and h_i'' are the scalor first-order and second-order derivatives of the positive function h_i, respectively.

The corresponding blocks of the average score vector and of the average information matrix of the quasi log-likelihood are defined to be:

$$\boldsymbol{s}_{\varphi}(\boldsymbol{\theta}) = T^{-1} \sum_{t=1}^{T} \frac{\partial \log L_t}{\partial \boldsymbol{\varphi}} \tag{A.8}$$

$$\boldsymbol{I}_{\varphi}(\boldsymbol{\theta}) = -T^{-1} \sum_{t=1}^{T} \mathrm{E} \left[\frac{\partial^2 \log L_t}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} \right]. \tag{A.9}$$

It can be seen from (A.3) that the corresponding hession matrix $(\partial^2 \log L_t / \partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}')$ is block diagonal, and so is $\boldsymbol{I}_{\varphi}(\boldsymbol{\theta})$ in (A.9).

From (A.1), (A.4) and (A.6), it is seen that:

$$\frac{\partial \log L_t}{\partial \boldsymbol{\varphi}_i} = \left(\frac{h_i'}{2\lambda_{it}} \right) g_{it} \boldsymbol{z}_{it} \tag{A.10}$$

Thus, under the null hypothesis of constant covariance over time, I have the average score vector:

$$\boldsymbol{s}_{\varphi}(\tilde{\boldsymbol{\theta}}) = T^{-1} \sum_{t=1}^{T} \left[\left(\frac{\tilde{h}_1'}{2\tilde{\lambda}_1} \right) \tilde{g}_{1t} \tilde{\boldsymbol{z}}_{1t}', \ \ldots \ , \ \left(\frac{\tilde{h}_p'}{2\tilde{\lambda}_p} \right) \tilde{g}_{pt} \tilde{\boldsymbol{z}}_{pt}' \right]'. \tag{A.11}$$

where $\tilde{g}_{it} = \tilde{w}_{it}^2 / \tilde{\lambda}_i - 1$.

From (A.2), and (A.4) to (A.7), I have that

$$\begin{aligned}
\mathrm{E}\left[\frac{\partial^2 \log L_t}{\partial \boldsymbol{\varphi}_i \partial \boldsymbol{\varphi}_i'}\right] &= \mathrm{E}\left[\frac{\partial^2 \log L_t}{\partial \lambda_{it}^2}\frac{\partial \lambda_{it}}{\partial \boldsymbol{\varphi}_i}\frac{\partial \lambda_{it}}{\partial \boldsymbol{\varphi}_i'}\right] + \mathrm{E}\left[\frac{\partial \log L_t}{\partial \lambda_{it}}\frac{\partial^2 \lambda_{it}}{\partial \boldsymbol{\varphi}_i \partial \boldsymbol{\varphi}_i'}\right] \\
&= \mathrm{E}\left[\frac{1}{2}\left(\frac{h_i'}{\lambda_{it}}\right)^2 \left(1 - 2w_{it}^2 \lambda_{it}^{-1}\right) \boldsymbol{z}_{it}\boldsymbol{z}_{it}'\right] + \mathrm{E}\left[\frac{1}{2\lambda_{it}} g_{it} h_i'' \boldsymbol{z}_{it}\boldsymbol{z}_{it}'\right] \\
&= -\frac{1}{2}\left(\frac{h_i'}{\lambda_{it}}\right)^2 \mathrm{E}\left[\boldsymbol{z}_{it}\boldsymbol{z}_{it}'\right], \qquad \text{(A.12)}
\end{aligned}$$

due to the fact that $\mathrm{E}\left[1 - 2w_{it}^2\lambda_{it}^{-1}\right] = -1$ and $\mathrm{E}\left[g_{it}\right] = 0$.

Thus, under the null hypothesis of constant covariance over time, the diagonal block i of the average information matrix takes the form:

$$\boldsymbol{I}_{\varphi,i}(\tilde{\boldsymbol{\theta}}) = \frac{1}{2T}\sum_{t=1}^{T}\left(\frac{\tilde{h}_i'}{\tilde{\lambda}_i}\right)^2 \mathrm{E}\left[\tilde{\boldsymbol{z}}_{it}\tilde{\boldsymbol{z}}_{it}'\right]. \qquad \text{(A.13)}$$

The LM test can be consistently estimated as follows:

$$LM = \frac{1}{2}\sum_{i=1}^{p}\left[\left(\sum_{t=1}^{T}\tilde{g}_{it}\tilde{\boldsymbol{z}}_{it}'\right)\left(\sum_{t=1}^{T}\tilde{\boldsymbol{z}}_{it}\tilde{\boldsymbol{z}}_{it}'\right)^{-1}\left(\sum_{t=1}^{T}\tilde{g}_{it}\tilde{\boldsymbol{z}}_{it}\right)\right], \qquad \text{(A.14)}$$

where (h_i'/λ_{it}) has been cancelled out. $\square$

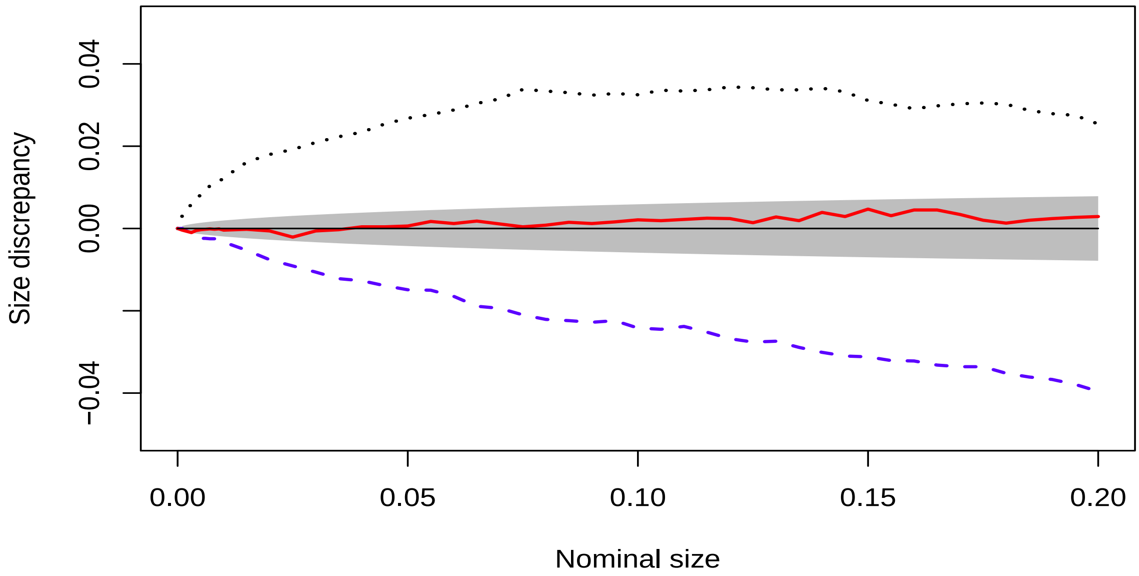


Figure 1: The size discrepancy plot: Gaussian errors $p = 5$, $T = 100$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed) and LM test against White specification (dotted). The grey area represents the 95% confidence region.

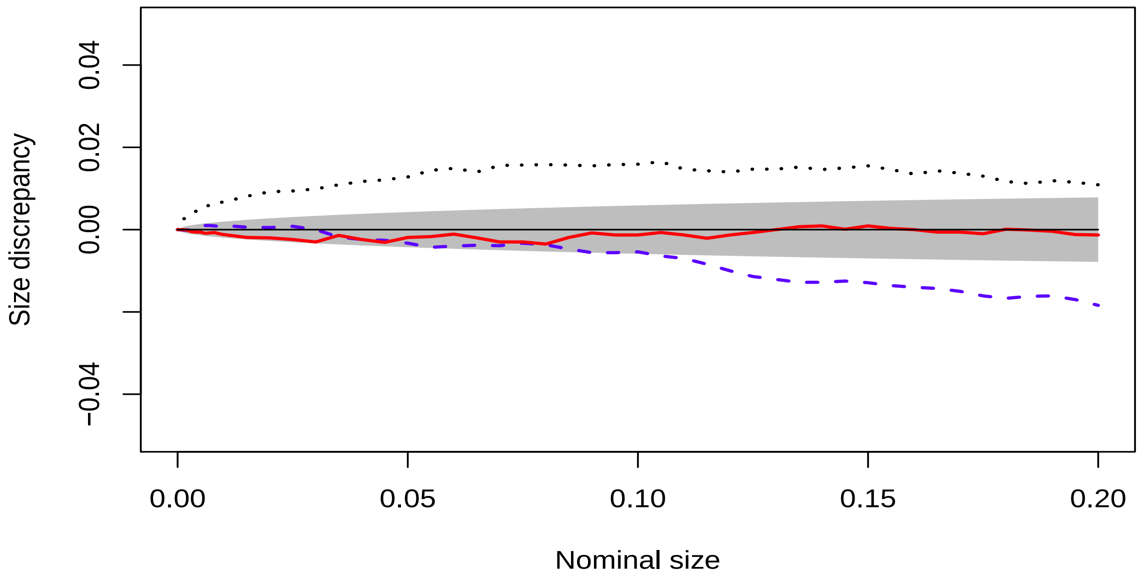


Figure 2: The size discrepancy plot: Gaussian errors $p = 5$, $T = 500$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed) and LM test against White specification (dotted). The grey area represents the 95% confidence region.

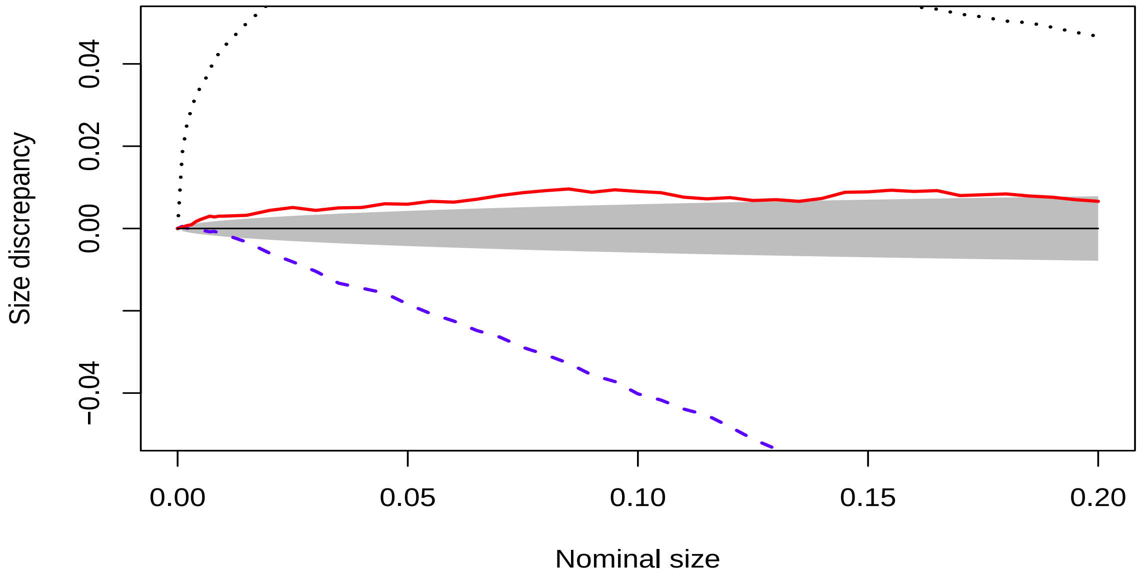


Figure 3: The size discrepancy plot: t(5) errors $p = 5$, $T = 100$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed) and LM test against White specification (dotted). The grey area represents the 95% confidence region.

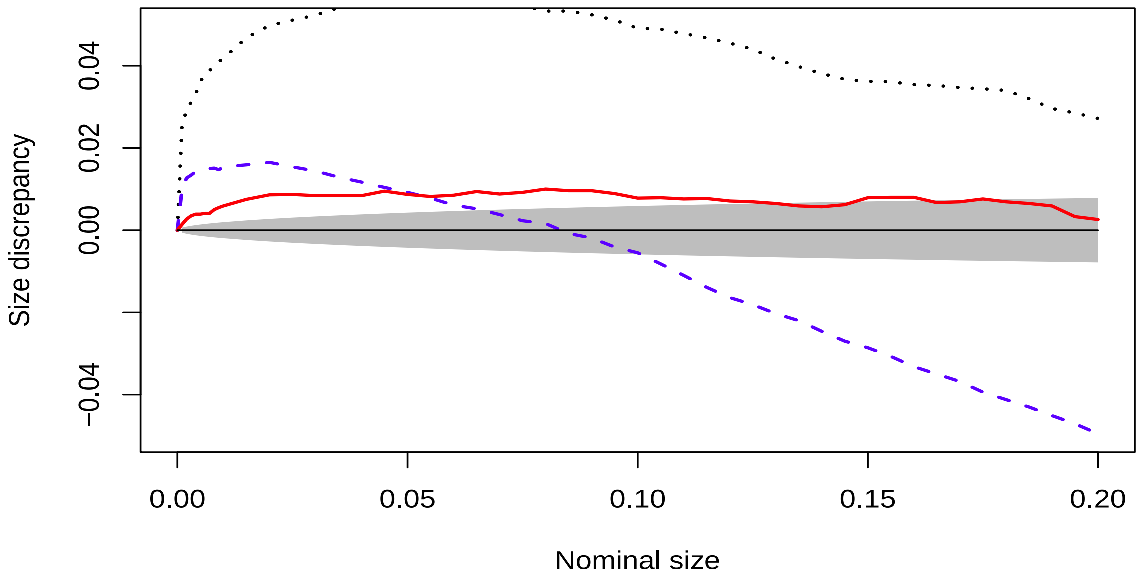


Figure 4: The size discrepancy plot: t(5) errors $p = 5$, $T = 500$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed) and LM test against White specification (dotted). The grey area represents the 95% confidence region.

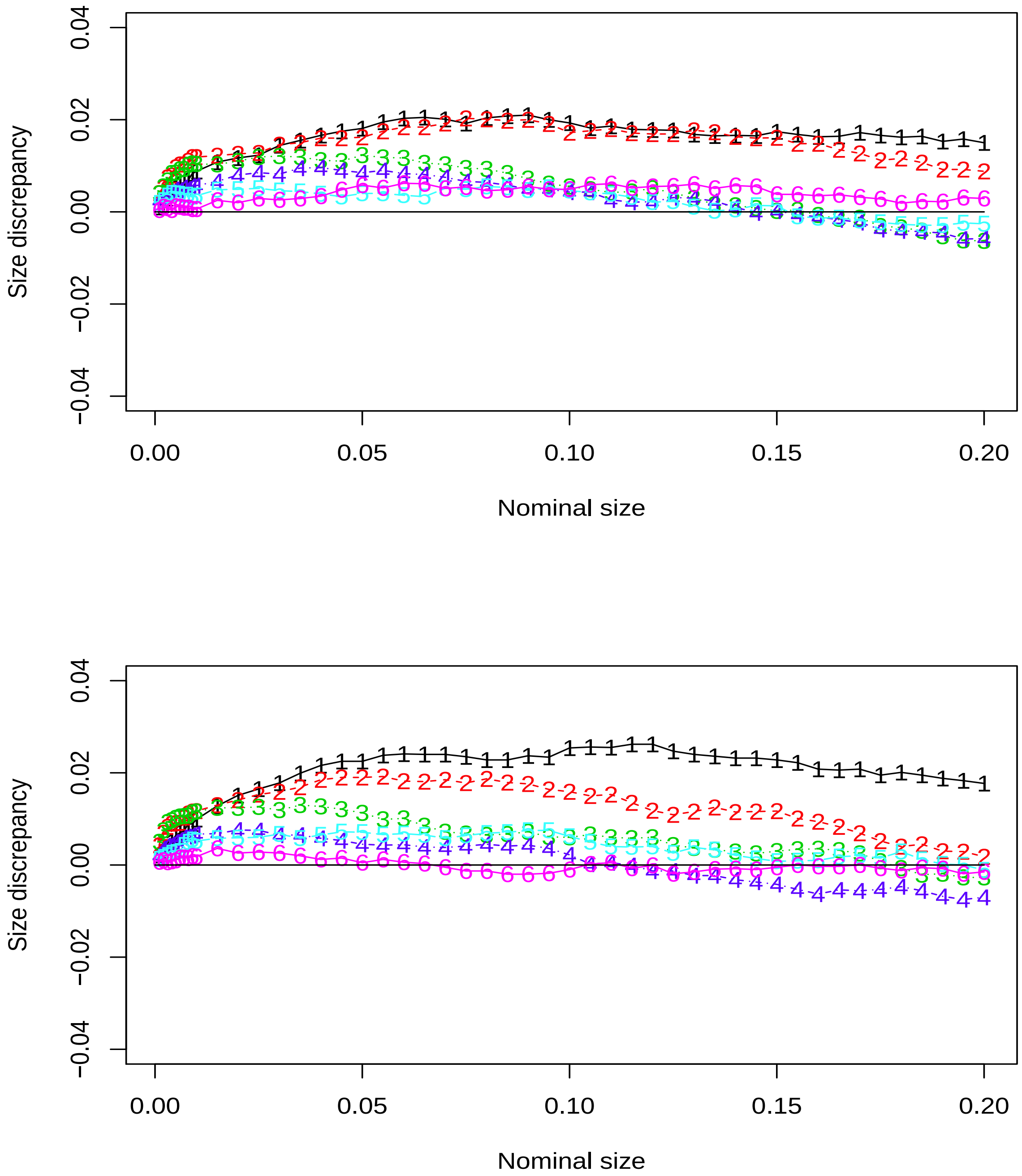
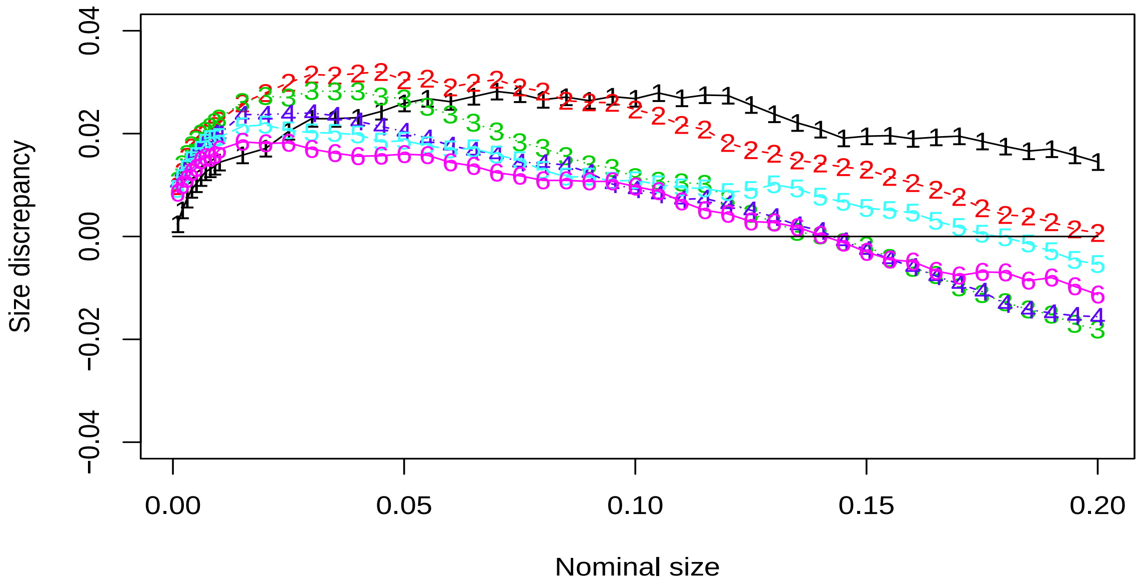


Figure 5: The size discrepancy plot of LM test against White specification: Gaussian errors $p = 3$, $T = 25, 50, 100, 250, 500, 1000$ from 1 to 6 and $N = 10000$. From top to bottom: $\rho = 0.9, 0$.



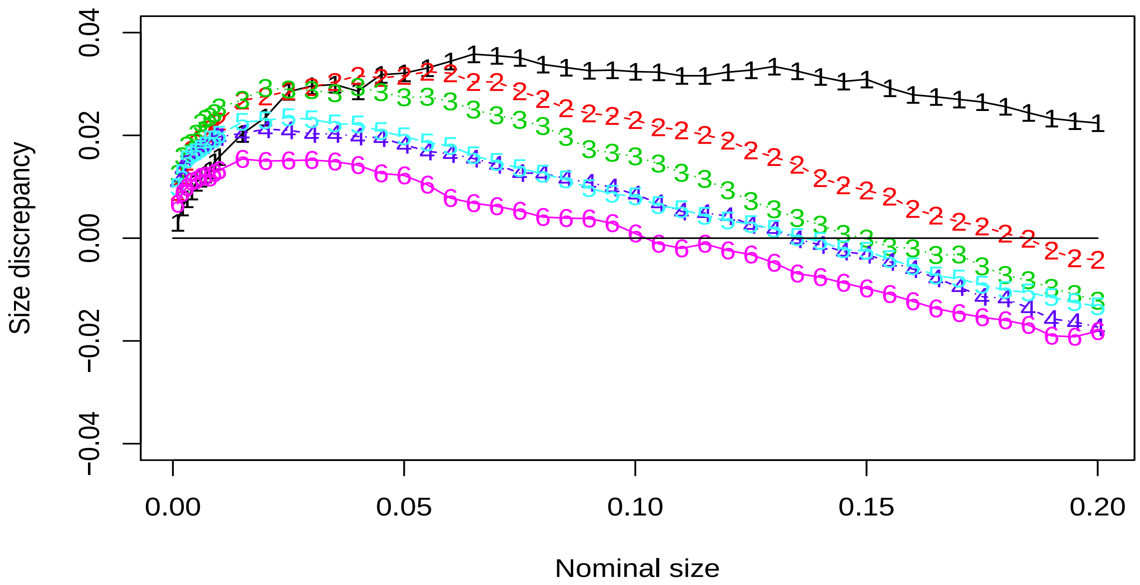


Figure 6: The size discrepancy plot of LM test against White specification: t(5) errors $p = 3$, $T = 25, 50, 100, 250, 500, 1000$ from 1 to 6 and $N = 10000$. From top to bottom: $\rho = 0.9, 0$.

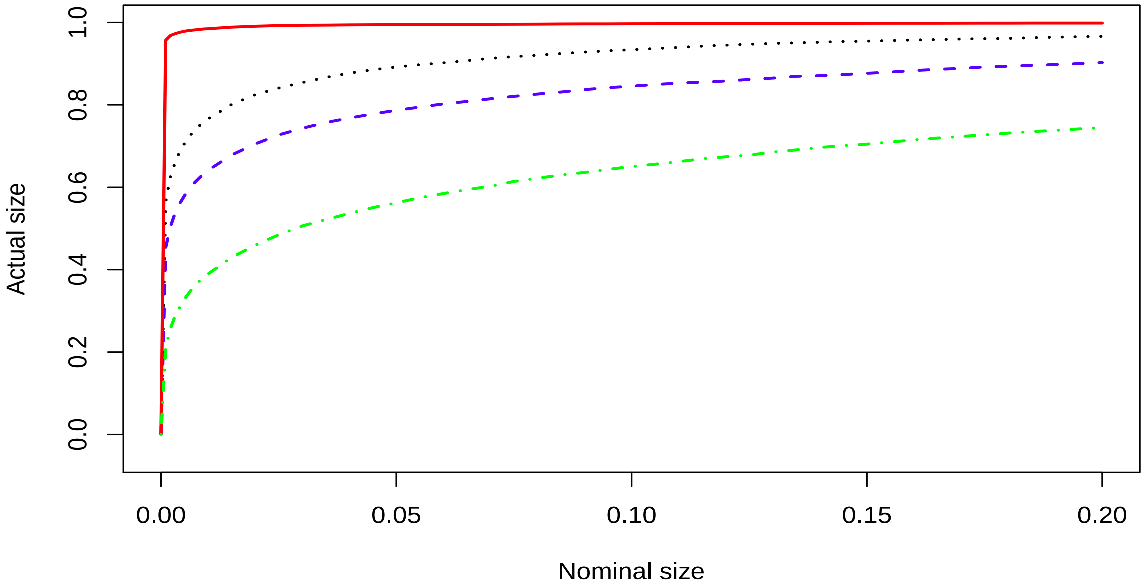


Figure 7: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{P}\boldsymbol{\Lambda}_t\boldsymbol{P}'$ with threshold change at $T/2$, $p = 5$, $T = 100$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed), LM test against White specification (dotted) and test in Eklund and Teräsvirta (2007) (dot-dashed).

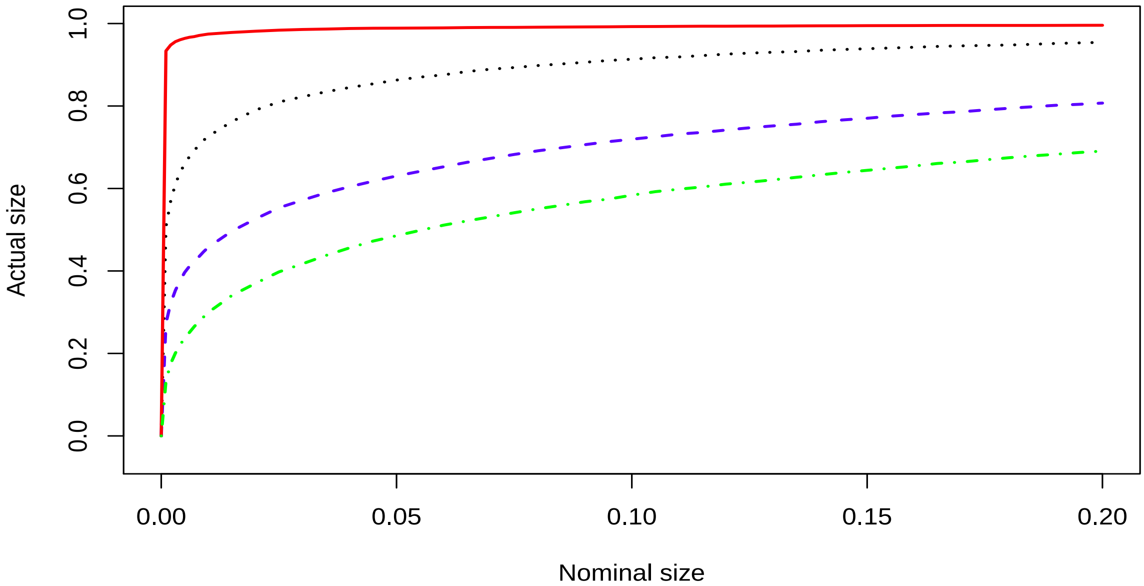


Figure 8: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{P}\boldsymbol{\Lambda}_t\boldsymbol{P}'$ with smooth change at $T/2$, $\gamma = 12$, $p = 5$, $T = 100$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed), LM test against White specification (dotted) and test in Eklund and Teräsvirta (2007) (dot-dashed).

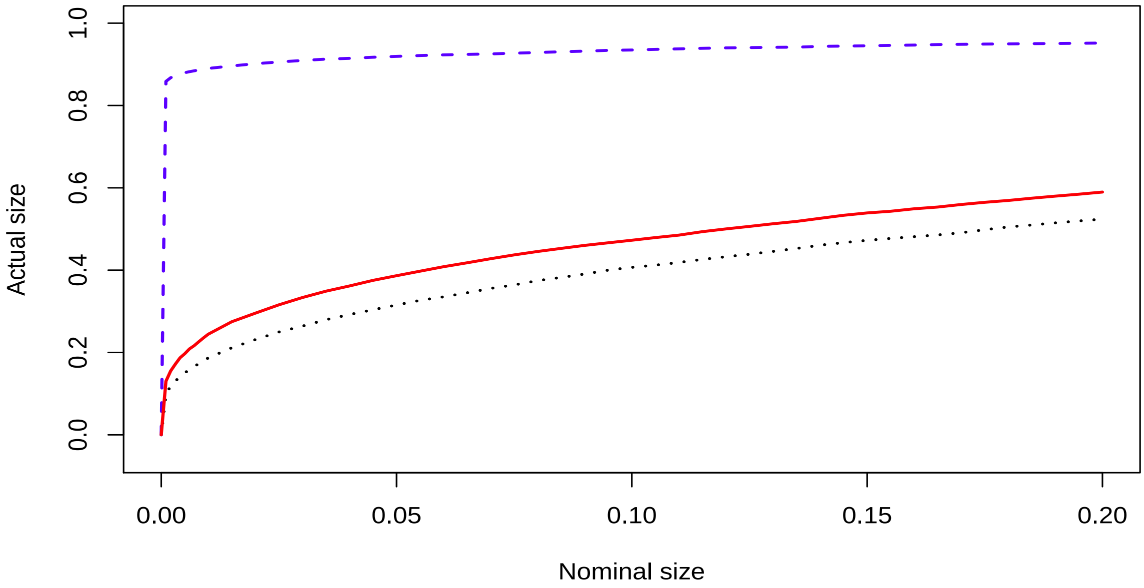


Figure 9: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{P}\boldsymbol{\Lambda}_t\boldsymbol{P}'$ with ARCH specification, $p = 5$, $T = 500$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed) and LM test against White specification (dotted).

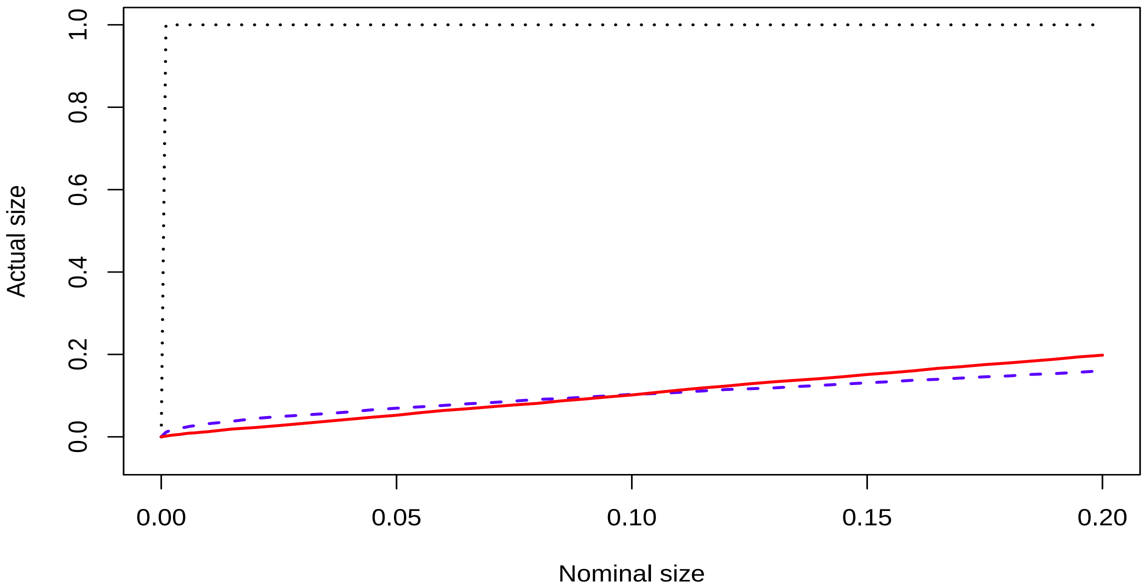


Figure 10: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{P}\boldsymbol{\Lambda}_t\boldsymbol{P}'$ with White specification, $p = 5$, $T = 500$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed) and LM test against White specification (dotted).

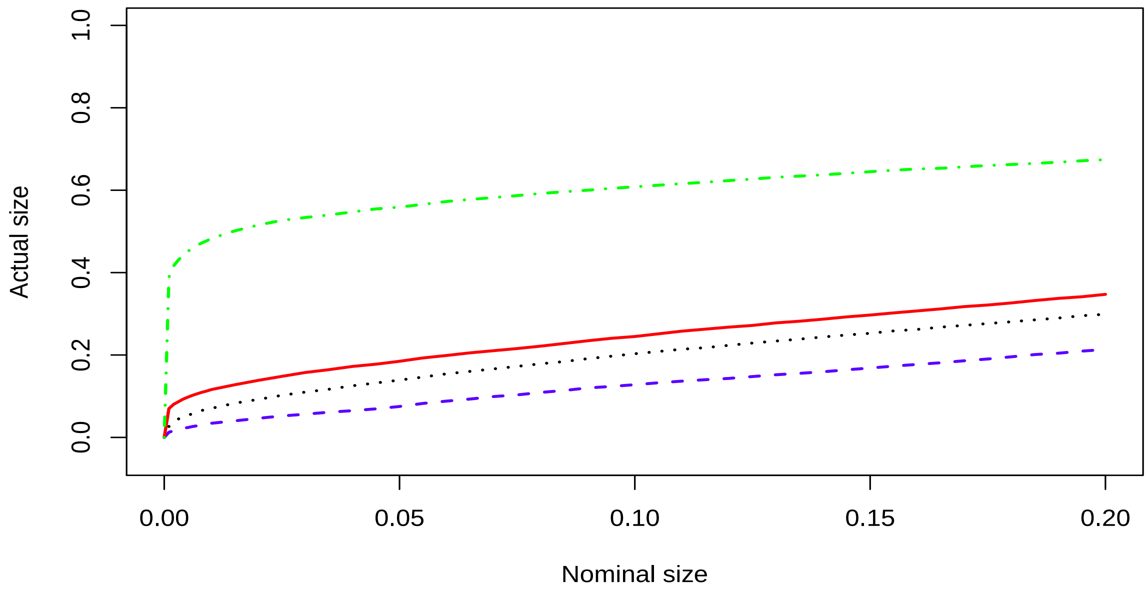


Figure 11: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{P}_t \boldsymbol{\Lambda} \boldsymbol{P}_t'$ with smooth change at $T/2$, $\gamma = 12$, $p = 2$, $T = 100$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed), LM test against White specification (dotted) and test in Eklund and Teräsvirta (2007) (dot-dashed).

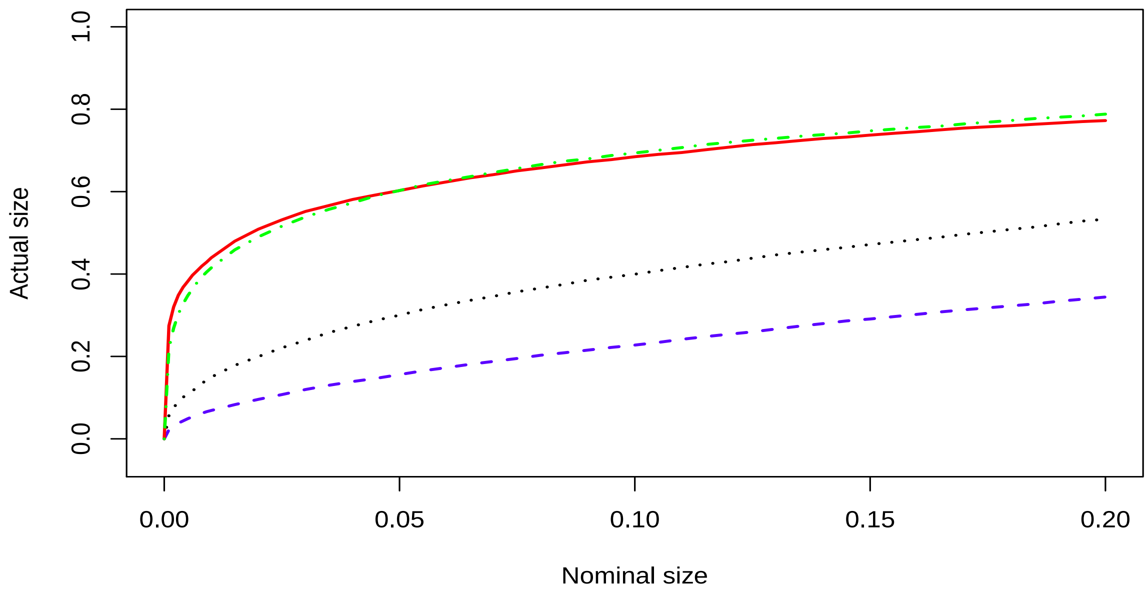


Figure 12: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{P}_t \boldsymbol{\Lambda} \boldsymbol{P}_t'$ with smooth change at $T/2$, $\gamma = 12$, $p = 5$, $T = 100$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed), LM test against White specification (dotted) and test in Eklund and Teräsvirta (2007) (dot-dashed).

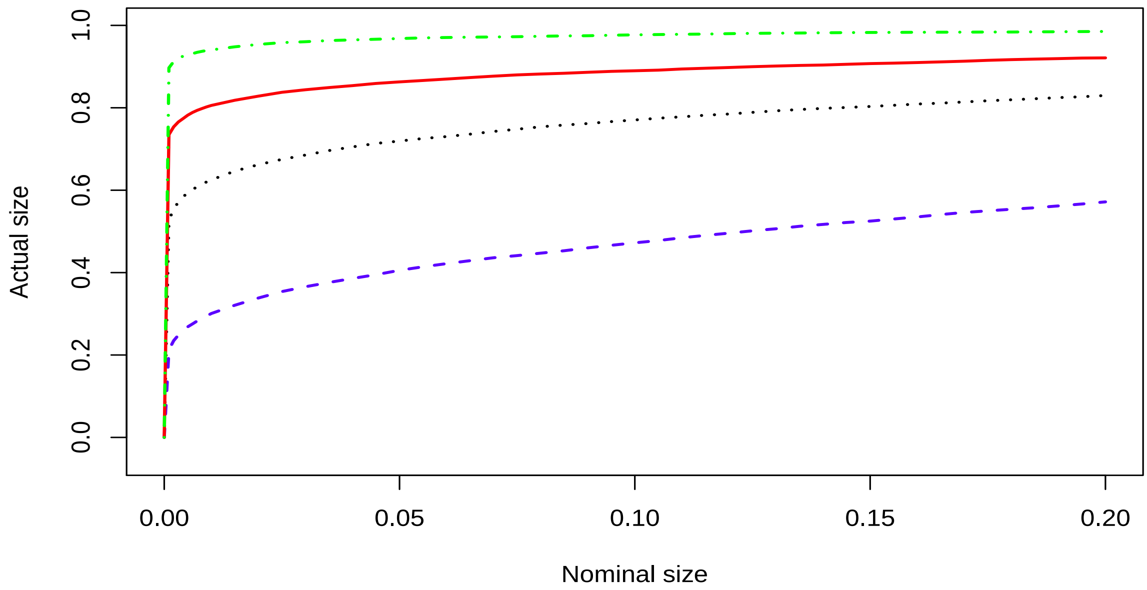


Figure 13: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{P}_t \boldsymbol{\Lambda} \boldsymbol{P}_t'$ with smooth change at $T/2$, $\gamma = 12$, $p = 5$, $T = 500$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed), LM test against White specification (dotted) and test in Eklund and Teräsvirta (2007) (dot-dashed).

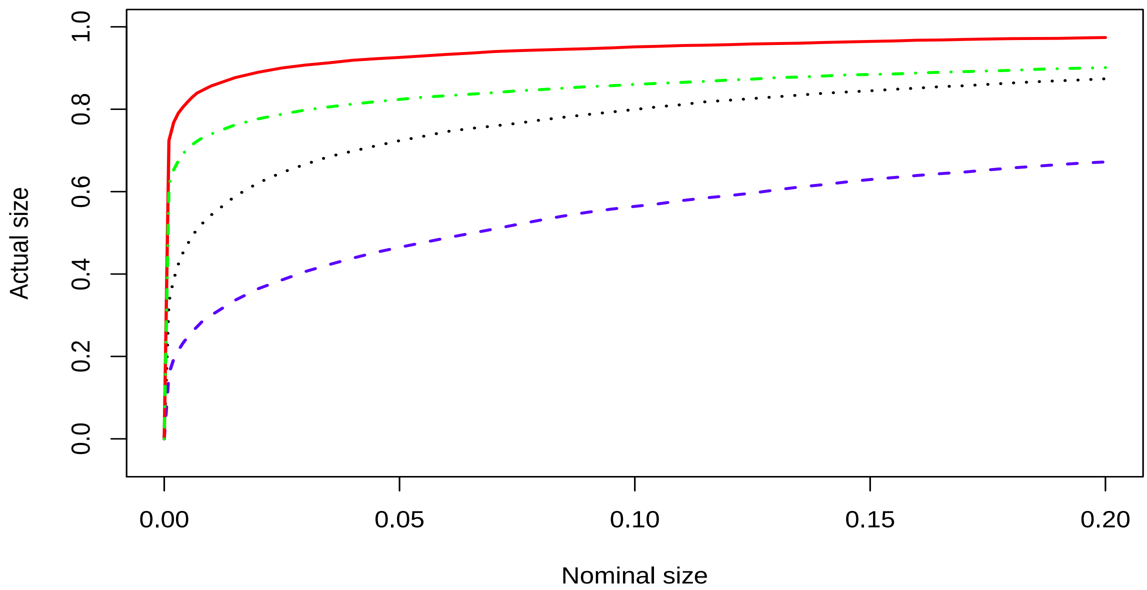


Figure 14: The power plot: $\boldsymbol{\Sigma}_t = \boldsymbol{D}_t \boldsymbol{Q} \boldsymbol{D}_t'$ with smooth change at $T/2$, $\gamma = 12$, $p = 5$, $T = 100$ and $N = 10000$. LM test against smooth transition specification (solid), LM test against ARCH specification (dashed), LM test against White specification (dotted) and test in Eklund and Teräsvirta (2007) (dot-dashed).

Research Papers 2013

Center for Research in Econometric Analysis of Time Series

2013-46:	Peter Christoffersen, Kris Jacobs, Xisong Jin and Hugues Langlois: Dynamic Diversification in Corporate Credit
2013-47:	Peter Christoffersen, Mathieu Fournier and Kris Jacobs: The Factor Structure in Equity Options
2013-48:	Peter Christoffersen, Ruslan Goyenko, Kris Jacobs and Mehdi Karoui: Illiquidity Premia in the Equity Options Market
2013-49:	Peter Christoffersen, Vihang R. Errunza, Kris Jacobs and Xisong Jin: Correlation Dynamics and International Diversification Benefits
2013-50:	Georgios Effraimidis and Christian M. Dahl: Nonparametric Estimation of Cumulative Incidence Functions for Competing Risks Data with Missing Cause of Failure
2013-51:	Mehmet Caner and Anders Bredahl Kock: Oracle Inequalities for Convex Loss Functions with Non-Linear Targets
2013-52:	Torben G. Andersen, Oleg Bondarenko, Viktor Todorov and George Tauchen: The Fine Structure of Equity-Index Option Dynamics
2014-01	Manuel Lukas and Eric Hillebrand: Bagging Weak Predictors
2014-02:	Barbara Annicchiarico, Anna Rita Bennato and Emilio Zanetti Chini: 150 Years of Italian CO2 Emissions and Economic Growth
2014-03:	Paul Catani, Timo Teräsvirta and Meiqun Yin: A Lagrange Multiplier Test for Testing the Adequacy of the Constant Conditional Correlation GARCH Model
2014-04:	Timo Teräsvirta and Yukai Yang: Linearity and Misspecification Tests for Vector Smooth Transition Regression Models
2014-05:	Kris Boudt, Sébastien Laurent, Asger Lunde and Rogier Quaedvlieg: Positive Semidefinite Integrated Covariance Estimation, Factorizations and Asynchronicity
2014-06:	Debopam Bhattacharya, Shin Kanaya and Margaret Stevens: Are University Admissions Academically Fair?
2014-07:	Markku Lanne and Jani Luoto: Noncausal Bayesian Vector Autoregression
2014-08:	Timo Teräsvirta and Yukai Yang: Specification, Estimation and Evaluation of Vector Smooth Transition Autoregressive Models with Applications
2014-09:	A.S. Hurn, Annastiina Silvennoinen and Timo Teräsvirta: A Smooth Transition Logit Model of the Effects of Deregulation in the Electricity Market
2014-10:	Marcelo Fernandes and Cristina M. Scherrer: Price discovery in dual-class shares across multiple markets
2014-11:	Barbara Annicchiarico and Yukai Yang: Testing Constancy of the Error Covariance Matrix in Vector Models against Parametric Alternatives using a Spectral Decomposition