Linearity and Misspecification Tests for Vector Smooth Transition Regression Models

Timo Teräsvirta and Yukai Yang

CREATES Research Paper 2014-4

(revised July 2015)

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

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

Linearity and Misspecification Tests for Vector Smooth Transition Regression Models

Timo Teräsvirta[†]   Yukai Yang[†‡]

[†]CREATES, Aarhus University
[‡]FINS, Vrije Universiteit Brussel

Abstract

The purpose of the paper is to derive Lagrange multiplier and Lagrange multiplier type specification and misspecification tests for vector smooth transition regression models. We report results from simulation studies in which the size and power properties of the proposed asymptotic tests in small samples are considered. The results on simulating the size show that these tests generally suffer from positive size distortion. We find that both Wilks's Λ and Rao's F statistic, the latter in particular, have satisfactory size properties and can generally be recommended for empirical use. The local asymptotic power and finite sample power properties of these tests are studied as well.

[†] Timo Teräsvirta, email: tterasvirta@creates.au.dk
[†‡] Yukai Yang, email: yukai.yang@gmail.com

1   Introduction

Hypothesis testing is an essential part of building and evaluating nonlinear time series models. Many nonlinear models such as the smooth transition regression or switching regression model or their univariate counterparts nest a linear model and are not identified if this linear model has generated the observations. This is why testing linearity is essential before fitting a nonlinear model. Evaluating any time series model before using it, typically for forecasting, is important to ensure the relevance of the empirical results. In this paper we consider testing a linear vector autoregressive (VAR) model against a nonlinear logistic vector smooth transition autoregressive (LVSTAR) or regression (LVSTR) models. Furthermore, we derive various misspecification tests for estimated LVSTAR or LVSTR models. They are needed for the strategy for building LVSTAR models developed and applied in Teräsvirta and Yang (2014). For a recent survey of Vector Threshold Autoregressive (VTAR) and LVSTAR models, see Hubrich and Teräsvirta (2013).

Eitrheim and Teräsvirta (1996) constructed misspecification tests for univariate STAR models. They included the test of no error autocorrelation, based on considerations in Godfrey (1988, Section 4.4), a test of the hypothesis of no additional nonlinearity, and a test against parameter nonconstancy. The last two tests contain the linearity test and the parameter constancy test in the linear VAR model as special cases. They build on the idea of circumventing the identification problem present in testing as in Saikkonen and Luukkonen (1988) and Luukkonen et al. (1988). Camacho (2004) generalised the test of no error autocorrelation to a bivariate STAR model.

In this paper we work further on linearity and misspecification tests in the LVSTR framework. We allow the dimension of the model exceed two and, furthermore, do not restrict the number of transitions to one. We focus on two cases. In the first one, the LVSTR model only has one transition variable, that is, the same transition variable is controlling nonlinearity in all equations. In the second case, the transition variable need not be the same for all equations. In some applications, there may exist underlying theory propositions determining these transition variables and thus justifying this type of test.

The asymptotic theory of our tests requires that the log-likelihood function is at least twice continuously differentiable in a neighbourhood of the null hypothesis. Theoretically this means that the tests are not valid for vector threshold autoregressive models such as the model by Tsay (1998). In practice, our tests do have power even against threshold-

type alternatives, see Strikholm and Teräsvirta (2006) and Teräsvirta et al. (2010, Section 16.4) for a discussion of this in the univariate threshold autoregressive case.

The plan of the paper is as follows. The LVSTAR model introduced in Teräsvirta and Yang (2014) is presented in Section 2. In Section 3 we consider linearity tests for two separate cases. In the first one, a single transition variable controls regime changes in all equations of the system. The second, somewhat more complicated, case is the one in which each equation can have its own transition function and transition variable. The misspecification tests for model evaluation are derived in Section 4. Small sample properties of the LM and other tests are investigated by simulation and results reported in Sections 5 and 6. Section 7 concludes. Some derivations can be found in the Appendix.

2   The logistic vector smooth transition model

The Logistic Vector Smooth Transition Regression (LVSTR) model with k lags is defined as follows:

$$\mathbf{y}_t = (\mathbf{B}_1' + \mathbf{G}_t^1\mathbf{B}_2' + ... + \mathbf{G}_t^{m-1}\mathbf{B}_m')\mathbf{x}_t + \boldsymbol{\varepsilon}_t = \boldsymbol{\Psi}_t'\mathbf{B}'\mathbf{x}_t + \boldsymbol{\varepsilon}_t \qquad (2.1)$$

where $\mathbf{y}_t = (y_{1t}, ..., y_{pt})'$ is a $p \times 1$ column vector, $\mathbf{x}_t = (\mathbf{y}_{t-1}', ..., \mathbf{y}_{t-k}', \mathbf{d}_t')'$ is a $(kp+q) \times 1$ vector in which $\mathbf{d}_t$ is a $q \times 1$ vector of a constant and $q-1$ stationary strongly exogenous variables. Furthermore, $\mathbf{B}_1, ..., \mathbf{B}_m$ are $(kp+q) \times p$ parameter matrices, $\mathbf{G}_t^i$ is a $p \times p$ diagonal matrix of transition functions,

$$\mathbf{G}_t^i = \text{diag}\left\{ \ g(s_{1t}|\gamma_{i1}, c_{i1}), ..., \ g(s_{pt}|\gamma_{ip}, c_{ip}) \ \right\} \qquad (2.2)$$

for $i = 1, ..., m-1$, where $s_{1t}, ..., s_{pt}$ are stationary observable random variables, and

$$g(s_{jt}|\gamma_{ij}, c_{ij}) = (1 + \exp\{-\gamma_{ij}(s_{jt} - c_{ij})\})^{-1}, \ \gamma_{ij} > 0 \qquad (2.3)$$

for $i = 1, ..., m-1$ and $j = 1, ..., p$ are logistic functions. In addition, $\boldsymbol{\Psi}_t = (\mathbf{I}_p, \mathbf{G}_t^1, ..., \mathbf{G}_t^{m-1})'$ is an $mp \times p$ full rank matrix and $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_m)$ is a $(kp+q) \times mp$ matrix of parameters. The $p \times 1$ error vectors $\boldsymbol{\varepsilon}_t$ are assumed independent and identically distributed (iid) with mean zero and a positive definite covariance matrix Ω. When $\mathbf{d}_t$ only contains the intercept, the model (2.1) is the Logistic Vector Smooth Transition Autoregressive (LVSTAR) model.

When $s_{1t} = s_{2t} = ... = s_{pt} = s_t$, $\gamma_{ij} \to \infty$, and $\mathbf{d}_t$ is the intercept vector, one obtains the vector threshold autoregressive (VTAR) model. In this model it is typically assumed that $c_{ij} = c_i$, $j = 1, ..., p$, for $i = 1, ..., m-1$, so $\mathbf{G}_t^i = \ g(s_t|\gamma_i, c_i)\mathbf{I}_p$. This gives the VTAR model with m regimes in Tsay (1998). It is also possible to simplify the LVSTAR model in the similar fashion, should the phenomenon to be modelled allow it.

3 Testing linearity against LVSTAR models

3.1 Testing when the equations have the same transition variable

Testing linearity is a central part of building not only vector STAR models but many other nonlinear models that nest a linear model as well. If the null hypothesis is not rejected, one can proceed with a linear vector autoregressive model with straightforward asymptotic theory and inference; see Lütkepohl (2006). Models such as the LVSTAR model or VTAR model are not identified under linearity, which makes it necessary to test linearity before fitting either one of these nonlinear models to data.

We begin by considering the case where $m = 2$, so the LVSTAR model (2.1) becomes

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{B}_1'\mathbf{x}_t + \mathbf{G}_t\mathbf{B}_2'\mathbf{x}_t + \boldsymbol{\varepsilon}_t \\
&= \sum_{j=1}^{k}\mathbf{B}_{1j}'\mathbf{y}_{t-j} + \boldsymbol{\mu}_1 + \mathbf{G}_t(\sum_{j=1}^{k}\mathbf{B}_{2j}'\mathbf{y}_{t-j} + \boldsymbol{\mu}_2) + \boldsymbol{\varepsilon}_t.
\end{aligned}
\tag{3.1}
$$

where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the intercept vectors. If each equation of (3.1) has a different transition variable, testing is often carried out equation by equation as in Camacho (2004), but we shall consider this situation in detail in Section 3.2. We now focus on the case in which the model has only a single transition variable, that is, $s_{1t} = s_{2t} = ... = s_{pt} = s_t$. This contains the special case $\mathbf{G}_t = g(s_t|\gamma, c)\mathbf{I}_p$.

Due to lack of identification under linearity, the null hypothesis may be stated in two ways. Either H$_0$: $\mathbf{B}_2 = \mathbf{0}$, in which case γ_i and c_i, $i = 1,...,p$, in (3.1) are the unidentified parameters, or H$_0$: $\gamma_i = 0$, $i = 1,...,p$. The latter case implies $\mathbf{G}_t \equiv (1/2)\mathbf{I}_p$, so that $\mathbf{B}_1$ and $\mathbf{B}_2$ in $\mathbf{B}_0 = \mathbf{B}_1 + (1/2)\mathbf{B}_2$ and each c_i are unidentified. In the former case one can follow Hansen (1996), construct a statistic independent of the $2p$ nuisance parameters and obtain its empirical null distribution by simulation or bootstrap. The last step is necessary because the null distribution of this test statistic is generally unknown even asymptotically. This method is computationally demanding, in particular when $p > 1$ and when each equation has its own transition variable. It is therefore expected to be feasible only if it is assumed that $\mathbf{G}_t = g(s_t|\gamma, c)\mathbf{I}$ so that the number of nuisance parameter is restricted to two.

We adopt a different method, suggested in the univariate case by Saikkonen and Luukkonen (1988) and Luukkonen et al. (1988). It leads to a computationally simple test and standard asymptotic inference applies. Following these authors, we solve the identification problem by approximating the logistic function $g(s_t|\gamma_j, c_j)$ with an nth

4

order Taylor approximation around $\gamma_j = 0$. We have

$$g(s_t|\gamma_j, c_j) = \sum_{i=0}^{n} d_{ji} s_t^i + r_{jt}$$

where $d_{j0} = 1/2$, $d_{ji} = \gamma^i \widetilde{d}_{ji}$, $\widetilde{d}_{ji} \neq 0$, $i = 1, ..., n$, and r_{jt} is the remainder. Rewrite $\mathbf{G}_t$ in (3.1) as follows:

$$\begin{aligned}
\mathbf{G}_t &= \text{diag}\{\sum_{i=0}^{n} d_{1i} s_t^i + r_{1t}, \ ..., \ \sum_{i=0}^{n} d_{pi} s_t^i + r_{pt}\} \\
&= \sum_{i=0}^{n} \mathbf{D}_i s_t^i + \mathbf{R}_t \tag{3.2}
\end{aligned}$$

where $\mathbf{D}_i = \text{diag}(d_{1i}, ..., d_{pi})$ and $\mathbf{R}_t = \text{diag}(r_{1t}, ..., r_{pt})$. Inserting (3.2) into (3.1) yields the following auxiliary model:

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{B}_1' \mathbf{x}_t + (\sum_{i=1}^{n} \mathbf{D}_i s_t^i + \mathbf{R}_t) \mathbf{B}_2' \mathbf{x}_t + \boldsymbol{\varepsilon}_t \\
&= (\mathbf{B}_1' + (1/2)\mathbf{B}_2') \mathbf{x}_t + \sum_{i=1}^{n} \mathbf{D}_i \mathbf{B}_2' \mathbf{x}_t s_t^i + \mathbf{R}_t \mathbf{B}_2' \mathbf{x}_t + \boldsymbol{\varepsilon}_t \\
&= \mathbf{B}_0' \mathbf{x}_t + \sum_{i=1}^{n} \boldsymbol{\Theta}_i' \mathbf{x}_t s_t^i + \boldsymbol{\varepsilon}_t^* \tag{3.3}
\end{aligned}$$

where $\boldsymbol{\Theta}_i = \mathbf{B}_2 \mathbf{D}_i$, $i = 1, ..., n$, and $\boldsymbol{\varepsilon}_t^* = \mathbf{R}_t \mathbf{B}_2' \mathbf{x}_t + \boldsymbol{\varepsilon}_t$. The null hypothesis implies $\mathbf{D}_i = \mathbf{0}$, $i = 1, ..., n$, and $\mathbf{R}_t = \mathbf{0}$ in (3.2). Thus, the auxiliary VAR model (3.3) is linear and H$_0$: $\boldsymbol{\Theta}_1 = ... = \boldsymbol{\Theta}_n = \mathbf{0}$. Moreover, when the null hypothesis holds, $\boldsymbol{\varepsilon}_t^* = \boldsymbol{\varepsilon}_t$ because then $\mathbf{R}_t = \mathbf{0}$. Since the Lagrange multiplier test only requires estimating the model under the null hypothesis, the remainder that is present under the alternative does not affect the asymptotic inference.

Denoting $\mathbf{Y} = (\mathbf{y}_1, ..., \mathbf{y}_T)'$, $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_T)'$, $\mathbf{E}^* = (\boldsymbol{\varepsilon}_1^*, ..., \boldsymbol{\varepsilon}_T^*)'$, $\boldsymbol{\Theta}^{(n)} = (\boldsymbol{\Theta}_1', ..., \boldsymbol{\Theta}_n')'$, and the $T \times np(kp + q)$ matrix $\mathbf{Z}_n = [\mathbf{x}_t' s_t^i]$, (3.3) can be rewritten as

$$\mathbf{Y} = \mathbf{X} \mathbf{B}_0 + \mathbf{Z}_n \boldsymbol{\Theta}^{(n)} + \mathbf{E}^*. \tag{3.4}$$

The approximate log-likelihood function of the auxiliary model (3.4) with T observations is, ignoring $\mathbf{R}_t$,

$$L_T(\theta) = k - (T/2) \log |\boldsymbol{\Omega}| - (1/2) \sum_{t=1}^{T} \boldsymbol{\varepsilon}_t' \boldsymbol{\Omega}^{-1} \boldsymbol{\varepsilon}_t.$$

Setting $\mathbf{s}_t^{(n)} = (s_t, s_t^2, ..., s_t^n)'$, the error term $\boldsymbol{\varepsilon}_t$ equals

$$\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \mathbf{B}_0' \mathbf{x}_t - \sum_{i=1}^{n} \boldsymbol{\Theta}_i' \mathbf{x}_t s_t^i = \mathbf{y}_t - \mathbf{B}_0' \mathbf{x}_t - \boldsymbol{\Theta}^{(n)'} \mathbf{z}_t$$

where $\mathbf{z}_t = (\mathbf{x}_t' s_t, \mathbf{x}_t' s_t^2, ..., \mathbf{x}_t' s_t^n)' = (\mathbf{x}_t \otimes \mathbf{s}_t^{(n)})$. Let $\mathbf{Z}_n = (\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_T)'$ and $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_T)'$. In order to derive an asymptotically valid test statistic we make the following assumptions:

Assumption 1. *The linear null model* $\mathbf{y}_t = \mathbf{B}_0'\mathbf{x}_t + \boldsymbol{\varepsilon}_t$ *is stationary.*

Assumption 2. *The elements of the error sequence* $\{\boldsymbol{\varepsilon}_t\}$ *are independent normal with mean zero and a positive definite covariance matrix* $\boldsymbol{\Omega}$.

Assumption 3. $\mathbf{X}'\mathbf{X}$ *and* $\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n$ *are positive definite matrices.*

Assumption 2 strengthens the previous assumption that the errors are iid by adding normality. Assumption 3 is a moment condition. For example, if the model is an LVSTAR model and $s_t = y_{i,t-d}$, $d > 0$, this implies that $\mathbf{y}_t$ has a finite $2(n+1)$th moment.

The block of the score vector involving the parameters under test can be written as follows:

$$
\begin{aligned}
\frac{\partial \log L(\tilde{\theta})}{\partial \boldsymbol{\Theta}^{(n)}} &= -\frac{\partial}{\partial \boldsymbol{\Theta}^{(n)}}(1/2)\sum_{t=1}^T \boldsymbol{\varepsilon}_t'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon}_t = \frac{\partial}{\partial \boldsymbol{\Theta}^{(n)}}\sum_{t=1}^T \mathbf{z}_t'\boldsymbol{\Theta}^{(n)}\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon}_t \\
&= \sum_{t=1}^T \mathbf{z}_t\boldsymbol{\varepsilon}_t'\boldsymbol{\Omega}^{-1} = \mathbf{Z}_n'\mathbf{E}\boldsymbol{\Omega}^{-1}
\end{aligned}
\tag{3.5}
$$

see for example Lütkepohl (1996, p. 171). Evaluated under H_0, (3.5) becomes

$$
\frac{\partial \log L(\tilde{\theta})}{\partial \boldsymbol{\Theta}^{(n)}}\Big|_{\mathrm{H}_0} = \sum_{t=1}^T \mathbf{z}_t\tilde{\boldsymbol{\varepsilon}}_t'\tilde{\boldsymbol{\Omega}}^{-1} = \mathbf{Z}_n'\tilde{\mathbf{E}}\tilde{\boldsymbol{\Omega}}^{-1}
\tag{3.6}
$$

where $\tilde{\mathbf{E}} = (\tilde{\boldsymbol{\varepsilon}}_1, \tilde{\boldsymbol{\varepsilon}}_2, ..., \tilde{\boldsymbol{\varepsilon}}_T)'$, $\tilde{\boldsymbol{\varepsilon}}_t = \mathbf{y}_t - \widetilde{\mathbf{B}}_0'\mathbf{x}_t$, and $\tilde{\boldsymbol{\Omega}}^{-1} = \{(1/T)\sum_{t=1}^T \tilde{\boldsymbol{\varepsilon}}_t\tilde{\boldsymbol{\varepsilon}}_t'\}^{-1}$. The matrix $\widetilde{\mathbf{B}}_0$ is the maximum likelihood estimator of $\mathbf{B}_0$ under the null hypothesis. From Assumptions 1 and 2 it follows that the vectorised score matrix is asymptotically normally distributed with $p \times \mathrm{cd}(\mathbf{Z}_n)$ degrees of freedom ($\mathrm{cd}(\mathbf{Z}_n)$ is the column dimension of $\mathbf{Z}_n$). See for example Breusch and Pagan (1980). As the score is normal and $\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n$ is positive definite, the vectorised LM test statistic

$$
LM_n = \mathrm{vec}(\tilde{\mathbf{E}}'\mathbf{Z}_n)'\{(\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n) \otimes \tilde{\boldsymbol{\Omega}}\}^{-1}\mathrm{vec}(\tilde{\mathbf{E}}'\mathbf{Z}_n)
\tag{3.7}
$$

where $\mathbf{P}_\mathbf{X} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, has an asymptotic χ^2-distribution with $p \times \mathrm{cd}(\mathbf{Z}_n)$ degrees of freedom when the null hypothesis holds.

The statistic (3.7) can also be written as follows:

$$
\begin{aligned}
LM_n &= \mathrm{vec}(\tilde{\mathbf{E}}'\mathbf{Z}_n)'\{(\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n) \otimes \tilde{\boldsymbol{\Omega}}\}^{-1}\mathrm{vec}(\tilde{\mathbf{E}}'\mathbf{Z}_n) \\
&= \mathrm{vec}(\tilde{\mathbf{E}}'\mathbf{Z}_n)'\{(\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n)^{-1} \otimes \tilde{\boldsymbol{\Omega}}^{-1}\}\mathrm{vec}(\tilde{\mathbf{E}}'\mathbf{Z}_n) \\
&= \mathrm{vec}(\tilde{\mathbf{E}}'\mathbf{Z}_n)'\mathrm{vec}\{\tilde{\boldsymbol{\Omega}}^{-1}\tilde{\mathbf{E}}'\mathbf{Z}_n(\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n)^{-1}\} \\
&= \mathrm{tr}\{\mathbf{Z}_n'\tilde{\mathbf{E}}\tilde{\boldsymbol{\Omega}}^{-1}\tilde{\mathbf{E}}'\mathbf{Z}_n(\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n)^{-1}\} \\
&= \mathrm{tr}\{\tilde{\boldsymbol{\Omega}}^{-1}\tilde{\mathbf{E}}'\mathbf{Z}_n[\mathbf{Z}_n'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_n]^{-1}\mathbf{Z}_n'\tilde{\mathbf{E}}\}.
\end{aligned}
\tag{3.8}
$$

Vectorisation and Kronecker products in (3.7) are avoided in (3.8). We have the following result:

Theorem 1. *Consider the model* (3.1) *and its approximation* (3.4). *Suppose that As-sumptions 1–3 hold. The LM test statistic for testing the null hypothesis* $H_0 : \gamma_j = 0$, $j = 1, ..., p$, *in* (3.1), *or equivalently,* $H_0 : \mathbf{\Theta}^{(n)} = \mathbf{0}$ *in* (3.4), *then equals*

$$LM_n = \text{tr}\{\tilde{\mathbf{\Omega}}^{-1}\tilde{\mathbf{E}}'\mathbf{Z}_n\,[\mathbf{Z}_n'(\mathbf{I}_T - \mathbf{P_X})\mathbf{Z}_n]^{-1}\,\mathbf{Z}_n'\tilde{\mathbf{E}}\} \tag{3.9}$$

where $\mathbf{P_X} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. *When the null hypothesis is valid, the LM statistic* (3.9) *has an asymptotic* χ^2-*distribution with* $np(kp+1)$ *degrees of freedom.*

The LM type statistic (3.9) can also be computed in three steps using the 'multivariate TR^2 form' as follows:

1. Estimate the restricted model. Collect the residuals $\tilde{\varepsilon}_t$ from this model and form the residual matrix $\tilde{\mathbf{E}}$. Compute the matrix residual sum of squares $\tilde{\mathbf{E}}'\tilde{\mathbf{E}}$.

2. Regress $\tilde{\mathbf{E}}$ on $\mathbf{X}$ and $\mathbf{Z}$. Form the residual matrix $\tilde{\mathbf{\Xi}}$ and the matrix residual sum of squares $\tilde{\mathbf{\Xi}}'\tilde{\mathbf{\Xi}}$.

3. Compute the test statistic

$$\begin{aligned} LM_{TR2} &= T\,\text{tr}\{(\tilde{\mathbf{E}}'\tilde{\mathbf{E}})^{-1}(\tilde{\mathbf{E}}'\tilde{\mathbf{E}} - \tilde{\mathbf{\Xi}}'\tilde{\mathbf{\Xi}})\} \\ &= T(p - \text{tr}\{(\tilde{\mathbf{E}}'\tilde{\mathbf{E}})^{-1}\tilde{\mathbf{\Xi}}'\tilde{\mathbf{\Xi}}\}). \end{aligned} \tag{3.10}$$

The statistics (3.9) and (3.10) are asymptotically equivalent. The former one collapses into the univariate LM-type linearity test statistic when $p = 1$ while the latter one is its TR^2-version. The test can also be applied to any subset of equations in the system, for instance, to check whether some equations in the system are nonlinear with a common transition variable while the rest are assumed linear.

The choice of the order of the Taylor expansion n may depend on the length of the time series and the dimension of the model. A higher order will increase the column dimension of $\mathbf{Z}_n$. But then, rejecting the null hypothesis may become easier, since a higher order often increases the power of the test. On the other hand, a lower order, for example $n = 1$, may in small samples lead to a test with better size properties, because the null hypothesis consists of fewer parameters than when $n > 1$.

However, choosing $n = 1$ is not a good choice when $s_t = y_{t-d,j}$ for some $1 \le d \le k$, $1 \le j \le p$ and only the intercept vector is varying nonlinearly with this transition variable. Then, as Luukkonen et al. (1988) pointed out in the univariate case, the LM_1 statistic has only trivial power against this alternative. The problem is typically solved by choosing $n = 3$ instead of $n = 1$, because in some situations $n = 2$ may still lead to a test that lacks power.

The statistic (3.9) does have power against the VTAR model, although the asymptotic theory ceases to apply when $\gamma_j \to \infty$, $j = 1, ..., p$. In this case, the Hansen approach is expected to be superior to the Taylor series approximation -based test because there is just one nuisance parameter present under the null hypothesis. Simulations in Hansen (1996) showed that the power loss when switching from the former test to the latter was distinct but not overly large. Those simulations were conducted with univariate models.

3.2 Testing when the equations have different transition variables

We turn to the LVSTAR model in which every equation has its own stationary transition variable s_{jt}, $j = 1, ..., p$. As already mentioned, linearity testing is in that situation often carried out equation by equation. However, an investigator may want to use a joint test in order to control the overall significance level of the test(s). Besides, he may want to make use of error correlations across the equations. It is therefore reasonable to consider a joint linearity test of the model (3.1). Following Luukkonen et al. (1988) and approximating each transition function in $\mathbf{G}_t$ by its first-order (for notational simplicity) Taylor approximation yields

$$\mathbf{y}_t = \mathbf{B}_0' \mathbf{x}_t + \mathbf{S}_t \mathbf{\Theta}_1' \mathbf{x}_t + \boldsymbol{\varepsilon}_t^* \tag{3.11}$$

where $\mathbf{S}_t = \text{diag}(s_{1t},\ s_{2t},\ ...,\ s_{pt})$ and $\boldsymbol{\varepsilon}_t^*$ is defined as in (3.3). The null hypothesis to be tested is $\mathbf{\Theta}_1 = \mathbf{0}$ in (3.11). The block of the score vector containing the partial derivatives of $\mathbf{\Theta}_1$ and evaluated under H_0 now has the form

$$\frac{\partial \log L(\tilde{\theta})}{\partial \mathbf{\Theta}_1}|_{H_0} = \sum_{t=1}^{T} \mathbf{z}_t \tilde{\boldsymbol{\varepsilon}}_t' \tilde{\mathbf{\Omega}}^{-1} \mathbf{S}_t. \tag{3.12}$$

where $\tilde{\boldsymbol{\varepsilon}}_t$ has been defined in (3.5). The LM statistic (3.7) can still be used for testing $\mathbf{\Theta}_1 = \mathbf{0}$. To see this, write

$$\begin{aligned}
\mathbf{S}_t \mathbf{\Theta}_1' \mathbf{x}_t &= \text{vec}(\mathbf{x}_t' \mathbf{\Theta}_1 \mathbf{S}_t) = (\mathbf{S}_t \otimes \mathbf{x}_t)' \text{vec}(\mathbf{\Theta}_1) = \text{vec}\left(\text{vec}(\mathbf{\Theta}_1)'(\mathbf{S}_t \otimes \mathbf{x}_t)\right) \\
&= (\mathbf{I}_p \otimes \text{vec}(\mathbf{\Theta}_1))' \, \text{vec}\,(\mathbf{S}_t \otimes \mathbf{x}_t) = \breve{\mathbf{\Theta}}_1' \text{vec}\,(\mathbf{s}_t \otimes \mathbf{x}_t) \tag{3.13}
\end{aligned}$$

where $\mathbf{s}_t = (s_{1t},\ s_{2t},\ ...,\ s_{pt})'$ and $\breve{\mathbf{\Theta}}_1$ corresponds to the nonzero elements of $\text{vec}\,(\mathbf{S}_t \otimes \mathbf{x}_t)$ after eliminating the zeroes. The number of degrees of freedom must thus be equal to the number of parameters in $\breve{\mathbf{\Theta}}_1$. The null hypothesis can be rewritten as H_0: $\breve{\mathbf{\Theta}}_1 = \mathbf{0}$. Using this reparameterisation, the corresponding block of the score matrix is (3.5) with $\mathbf{z}_t = \text{vec}\,(\mathbf{s}_t \otimes \mathbf{x}_t)$.

An LM-type statistic based on (3.11) is obtained by replacing $\mathbf{Z}_n$ in (3.8) by

$$\mathbf{Z} = (\text{vec}^*(\mathbf{s}_1 \otimes \mathbf{x}_1), \text{vec}^*(\mathbf{s}_2 \otimes \mathbf{x}_2), ..., \text{vec}^*(\mathbf{s}_T \otimes \mathbf{x}_T))'. \tag{3.14}$$

The notation vec* implies that the operator removes duplicate variables from $\mathbf{z}_t$, so $\mathbf{Z}$ has a full column rank. For example, if $\mathbf{x}_t = \mathbf{y}_{t-1} = (y_{1,t-1}, ..., y_{p,t-1})'$ in (3.14), and the ith equation has $y_{i,t-1}$ as its transition variable, there exist $p(p-1)/2$ pairs of variables of type $y_{i,t-1}y_{j,t-1}$, $i \neq j$ in $\mathbf{z}_t = \text{vec}(\mathbf{s}_t \otimes \mathbf{x}_t)$. One has to be removed from each pair. The degrees of freedom of the test statistic have to be adjusted accordingly, so in this example there are $p^2(p+1)/2$ degrees of freedom. The test can also be carried out in stages as discussed in the previous section.

Another way of performing a joint test is to sum up the individual asymptotically χ^2-distributed statistics and use the sum assuming that it is χ^2-distributed. This assumption only holds when the errors are not correlated. When they are correlated, the size of the test may be adjusted by bootstrap but the problem is that the statistic is not pivotal. Small sample properties of these two test statistics are studied by simulation in Sections 5 and 6.

4   Evaluation tests

4.1   Serial correlation in the error process

The evaluation stage of the LVSTAR modelling strategy designed and applied in Teräsvirta and Yang (2014) makes use of misspecification tests of the estimated LVSTAR model. We shall consider three such tests that are either LM or LM type tests. First, we extend the LM test of no serial correlation of Eitrheim and Teräsvirta (1996) to the multivariate case. Camacho (2004) considered this extension in the bivariate STR model. Assume the p-dimensional m-regime nonlinear LVSTAR model (2.1) with stationary autocorrelated errors:

$$\mathbf{y}_t = (\mathbf{B}_1' + \mathbf{G}_t^1\mathbf{B}_2' + ... + \mathbf{G}_t^{m-1}\mathbf{B}_m')\mathbf{x}_t + \mathbf{u}_t = \mathbf{\Psi}_t'\mathbf{B}'\mathbf{x}_t + \mathbf{u}_t \tag{4.1}$$

where

$$\mathbf{u}_t = \sum_{j=1}^{J} \mathbf{P}_j'\mathbf{u}_{t-j} + \boldsymbol{\varepsilon}_t = \mathbf{P}(\mathsf{L})\mathbf{u}_t + \boldsymbol{\varepsilon}_t. \tag{4.2}$$

In (4.2), each $\mathbf{P}_j$ is a $p \times p$ matrix, $\mathbf{P}(\mathsf{L}) = \sum_{j=1}^{J} \mathbf{P}_j'\mathsf{L}^j$, L is the lag operator, J is the lag length, and $\boldsymbol{\varepsilon}_t \sim \text{iid}\mathcal{N}(0, \boldsymbol{\Omega})$ is a $p \times 1$ vector. Stationarity implies that the roots of the polynomial $|\mathbf{I}_p - \sum_{j=1}^{J} \mathbf{P}_j'z^j|$ lie outside the unit circle. The null hypothesis equals H_0: $\mathbf{P}_1 = \mathbf{P}_2 = ... = \mathbf{P}_J = \mathbf{0}$ in (4.2).

Left-multiplying (4.1) by $\mathbf{I}_p - \mathbf{P}(\mathsf{L})$ yields

$$\begin{aligned}\mathbf{y}_t &= \mathbf{P}(\mathsf{L})\mathbf{u}_t + \mathbf{\Psi}_t'\mathbf{B}'\mathbf{x}_t + \boldsymbol{\varepsilon}_t \\ &= \mathbf{P}'\mathbf{z}_t + \mathbf{\Psi}_t'\mathbf{B}'\mathbf{x}_t + \boldsymbol{\varepsilon}_t,\end{aligned} \tag{4.3}$$

see Godfrey (1988, Section 4.4), where $\mathbf{P} = \begin{bmatrix} \mathbf{P}'_1 & \mathbf{P}'_2 & ... & \mathbf{P}'_J \end{bmatrix}'$ and $\mathbf{z}_t = \begin{bmatrix} \mathbf{u}'_{t-1} & \mathbf{u}'_{t-2} & ... & \mathbf{u}'_{t-J} \end{bmatrix}'$. The log-likelihood of the model (4.1) is

$$
\begin{aligned}
\log L &= -((T-J)p/2)\log 2\pi - ((T-J)/2)\log|\mathbf{\Omega}| \\
&\quad -(1/2)\sum_{t=J+1}^{T}(\mathbf{y}_t - \mathbf{P}'\mathbf{z}_t - \mathbf{\Psi}'_t\mathbf{B}'\mathbf{x}_t)'\,\mathbf{\Omega}^{-1}\,(\mathbf{y}_t - \mathbf{P}'\mathbf{z}_t - \mathbf{\Psi}'_t\mathbf{B}'\mathbf{x}_t) \\
&= -((T-J)p/2)\log 2\pi - ((T-J)/2)\log|\mathbf{\Omega}| \\
&\quad -(1/2)\sum_{t=J+1}^{T}(\mathbf{u}_t - \mathbf{P}'\mathbf{z}_t)'\,\mathbf{\Omega}^{-1}\,(\mathbf{u}_t - \mathbf{P}'\mathbf{z}_t)\,.
\end{aligned}
$$

The LM test is based on the score evaluated under the null hypothesis:

$$
\frac{\partial \log L(\tilde{\theta})}{\partial \mathbf{P}} = \sum_{t=J+1}^{T}\left[\tilde{\mathbf{z}}_t\tilde{\mathbf{u}}'_t\tilde{\mathbf{\Omega}}^{-1}\right] = \tilde{\mathbf{Z}}'\tilde{\mathbf{U}}\tilde{\mathbf{\Omega}}^{-1} \tag{4.4}
$$

where $\tilde{\mathbf{Z}}' = \begin{bmatrix} \tilde{\mathbf{z}}_{J+1} & \tilde{\mathbf{z}}_{J+2} & ... & \tilde{\mathbf{z}}_T \end{bmatrix}$ and $\tilde{\mathbf{U}}' = \begin{bmatrix} \tilde{\mathbf{u}}_{J+1} & \tilde{\mathbf{u}}_{J+2} & ... & \tilde{\mathbf{u}}_T \end{bmatrix}$, and $\tilde{\mathbf{z}}_t$, $\tilde{\mathbf{u}}_t$ and $\tilde{\mathbf{\Omega}}$ are maximum likelihood estimators under the null hypothesis. Setting $\tilde{\mathbf{E}} = \tilde{\mathbf{U}}$, $\tilde{\mathbf{Z}} = \begin{bmatrix} \tilde{\mathbf{z}}_{J+1} & \tilde{\mathbf{z}}_{J+2} & ... & \tilde{\mathbf{z}}_T \end{bmatrix}$, where $\mathbf{z}_t = \begin{bmatrix} \mathbf{u}'_{t-1} & \mathbf{u}'_{t-2} & ... & \mathbf{u}'_{t-J} \end{bmatrix}'$ and $\mathbf{P}_{\mathbf{K}} = \tilde{\mathbf{K}}(\tilde{\mathbf{K}}'\tilde{\mathbf{K}})^{-1}\tilde{\mathbf{K}}'$ with

$$
\tilde{\mathbf{K}} = \begin{bmatrix} \text{vec}[\partial(\tilde{\mathbf{\Psi}}'_{J+1}\tilde{\mathbf{B}}'\mathbf{x}_{J+1})/\partial\theta] & \text{vec}[\partial(\tilde{\mathbf{\Psi}}'_{J+2}\tilde{\mathbf{B}}'\mathbf{x}_{J+2})/\partial\theta] & ... & \text{vec}[\partial(\tilde{\mathbf{\Psi}}'_T\tilde{\mathbf{B}}'\mathbf{x}_T)/\partial\theta] \end{bmatrix} \tag{4.5}
$$

We make the following assumptions:

Assumption 4. *The LVSTAR model (4.1) is stationary and ergodic under H_0: $\mathbf{P}_1 = \mathbf{P}_2 = ... = \mathbf{P}_J = \mathbf{0}$ such that its parameters can be estimated consistently by maximum likelihood when the null hypothesis holds.*

Assumption 5. $\tilde{\mathbf{K}}'\tilde{\mathbf{K}}$ *and* $\tilde{\mathbf{Z}}'(\mathbf{I}_{T-J} - \mathbf{P}_{\mathbf{K}})\tilde{\mathbf{Z}}$ *are positive definite matrices.*

Assumption 4 is a high-level assumption, as general asymptotic theory for maximum likelihood estimators of stationary LVSTAR models does not yet exist. For a discussion of stability (but not stationarity) results for a class of vector nonlinear models including the LVSTAR model with $\mathbf{G}_t \equiv g(s_t|\gamma, c)\mathbf{I}_p$, see Saikkonen (2008).

With these assumptions and noting that Assumption 2 secures asymptotic normality of the score (4.4), we obtain a vectorised LM test statistic corresponding to (3.7). Analogously to the derivations leading to (3.8) we can state the following result:

Theorem 2. *Consider the model (4.1). Suppose that Assumptions 2, 4 and 5 hold. The LM test statistic for testing the null hypothesis H_0: $\mathbf{P}_1 = \mathbf{P}_2 = ... = \mathbf{P}_J = \mathbf{0}$ equals*

$$
LM = \text{tr}\{\tilde{\mathbf{\Omega}}^{-1}\tilde{\mathbf{U}}'\tilde{\mathbf{Z}}[\tilde{\mathbf{Z}}'(\mathbf{I}_{T-J} - \mathbf{P}_{\mathbf{K}})\tilde{\mathbf{Z}}]^{-1}\tilde{\mathbf{Z}}'\tilde{\mathbf{U}}\}. \tag{4.6}
$$

When the null hypothesis is valid, the LM statistic (4.6) has an asymptotic χ^2 distribution with Jp^2 degrees of freedom.

The vectorised first-order partial derivatives of $\boldsymbol{\Psi}_t'\mathbf{B}'\mathbf{x}_t$ with respect to parameters in θ can be found in Appendix A. The column dimension of $\tilde{\mathbf{K}}$ is $[(pk+q)m+2(m-1)]p^2$ which grows very rapidly as a function of m and p. If either p or m or both are large compared to T, inverting $\tilde{\mathbf{K}}'\tilde{\mathbf{K}}$ requires care. The matrix may be near-singular and the inversion slow. A necessary condition for the test statistic to exist is $T-J > (pk+q)m+2(m-1)]p^2$, which in short and moderate samples is bound to restrict the dimension of the null hypothesis.

The test can also be performed using the TR^2-form with a slight modification. Omission of the first J residuals causes positive size distortion because $\tilde{\mathbf{U}}'\tilde{\mathbf{K}} = \mathbf{0}$ only holds when all T residuals are included in $\tilde{\mathbf{U}}$. Furthermore, sometimes a nonlinear estimation algorithm may be terminated slightly too early such that $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{K}}$ are not exactly orthogonal. To remedy the situation, Eitrheim and Teräsvirta (1996) suggested replacing $\tilde{\mathbf{U}}$ without the first J residuals by its orthogonal part to the space spanned by $\tilde{\mathbf{K}}$, i.e., by $\tilde{\mathbf{V}} = (\mathbf{I} - \tilde{\mathbf{K}}(\tilde{\mathbf{K}}'\tilde{\mathbf{K}})^{-1}\tilde{\mathbf{K}}')\tilde{\mathbf{U}}$. In the multivariate case, doing so leads to the following procedure:

1. Estimate the LVSTAR model under the null hypothesis of no serial correlation. Choose a lag length J of the serial correlation in residuals. Regress the residuals $\tilde{\mathbf{U}}$ on $\tilde{\mathbf{K}}$ from $t = J+1$ to T. Collect the residuals $\tilde{\mathbf{V}}$ and compute the matrix residual sum of squares $\tilde{\mathbf{V}}'\tilde{\mathbf{V}}$.

2. Regress $\tilde{\mathbf{V}}$ on (shortened; the first J observations are removed) $\tilde{\mathbf{K}}$ and $\tilde{\mathbf{Z}}$. Collect the residuals $\tilde{\tilde{\boldsymbol{\Xi}}}$, and form the matrix residual sum of squares $\tilde{\tilde{\boldsymbol{\Xi}}}'\tilde{\tilde{\boldsymbol{\Xi}}}$.

3. Compute the value of the test statistic (3.10).

It should be noted that in small and moderate samples the statistic (4.6) can be seriously size distorted despite this adjustment. How to deal with this problem will be discussed in Section 5.

This test is also valid when the estimated model is a VTAR model. The reason is that the threshold parameter if $m = 1$, or parameters if $m > 1$, in these models are estimated super consistently and can thus be assumed known in the tests. The test of no error autocorrelation thus becomes analogous to the corresponding test in the linear VAR model, see Strikholm and Teräsvirta (2006) for discussion.

4.2   Additive nonlinearity

We shall now consider the alternative hypothesis that after fitting a LVSTAR model, there is still nonlinearity left unmodelled. Following Eitrheim and Teräsvirta (1996), it is specified as another logistic smooth transition component that enters the model additively. When the null hypothesis of no additive nonlinearity is rejected, there are at least two

alternatives exist. First, one may accept the alternative and estimate a LVSTAR model with two transitions. Since the reasons for a rejection usually remain unknown, it is also possible to conclude that the model does not fit the data sufficiently well and either respecify the whole model or switch to another family of models.

The model under the alternative is a p-dimensional additive LVSTAR model

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{B}_1'\mathbf{x}_t + \mathbf{G}_t^1\mathbf{B}_2'\mathbf{x}_t + ... + \mathbf{G}_t^{m-1}\mathbf{B}_m'\mathbf{x}_t + \boldsymbol{\varepsilon}_t + \mathbf{G}_t^m\mathbf{B}_{m+1}'\mathbf{x}_t + \boldsymbol{\varepsilon}_t \\
&= \boldsymbol{\Psi}_t'\mathbf{B}'\mathbf{x}_t + \boldsymbol{\varepsilon}_t + \mathbf{G}_t^m\mathbf{B}_{m+1}'\mathbf{x}_t + \boldsymbol{\varepsilon}_t
\end{aligned}
\tag{4.7}
$$

where $\boldsymbol{\varepsilon}_t \sim \text{iid}\mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$. The null hypothesis $\gamma_{mj} = 0$, $j = 1, ..., p$, is tested against the alternative that at least one $\gamma_{mj} > 0$. The null hypothesis implies that $\mathbf{G}_t^m = (1/2)\mathbf{I}_p$. If $s_{1t} = ... = s_{pt} = s_t$ in $\mathbf{G}_t^m$ under the alternative, the identification problem is solved by approximating $\mathbf{G}_t^m$ as in (3.2) so that it appears as in (3.3). The derivation of the LM test statistic is the same as that in Section 3.1, except that

$$
\tilde{\mathbf{E}} = \mathbf{Y} - \mathbf{X}\tilde{\mathbf{B}}_0 - \tilde{\mathbf{G}}_t^1\tilde{\mathbf{B}}_2'\mathbf{x}_t - ... - \tilde{\mathbf{G}}_t^{m-1}\tilde{\mathbf{B}}_m'\mathbf{x}_t
\tag{4.8}
$$

where each of $\tilde{\mathbf{G}}_t^1, ..., \tilde{\mathbf{G}}_t^{m-1}$ is a function of the slope and location parameters estimated under H_0 and that $\mathbf{P_K} = \tilde{\mathbf{K}}(\tilde{\mathbf{K}}'\tilde{\mathbf{K}})^{-1}\tilde{\mathbf{K}}'$ where $\tilde{\mathbf{K}}$ is defined as in (4.5) with $J = 0$. We make the following assumption:

Assumption 6. $\tilde{\mathbf{K}}'\tilde{\mathbf{K}}$ *and* $\mathbf{Z}_n'(\mathbf{I}_T - \mathbf{P_K})\mathbf{Z}_n$ *are positive definite matrices.*

Following the arguments used in deriving the test statistic (4.6), we can state the following result:

Theorem 3. *Consider the LVSTAR model (4.7). Suppose that Assumptions 2, 4 and 6 hold. The LM test statistic for testing the null hypothesis $H_0 : \gamma_{mj} = 0$, $j = 1, ..., p$, or $\boldsymbol{\Theta}^{(n)} = 0$ as in (3.4) equals*

$$
LM = \text{tr}\{\tilde{\boldsymbol{\Omega}}^{-1}\tilde{\mathbf{E}}'\mathbf{Z}_n[\mathbf{Z}_n'(\mathbf{I}_T - \mathbf{P_K})\mathbf{Z}_n]^{-1}\mathbf{Z}_n'\tilde{\mathbf{E}}\}
\tag{4.9}
$$

with $\tilde{\mathbf{E}}$ defined as in (4.8) and $\tilde{\mathbf{K}}$ as in (4.5) with $J = 0$. When the null hypothesis is valid, the LM statistic (4.9) has an asymptotic χ^2 distribution with $pn(kp + 1)$ degrees of freedom.

The test statistic is similar to the one in Section 3 except for the fact that $\mathbf{X}$ in the definition of $\mathbf{P_X}$ is replaced by $\tilde{\mathbf{K}}$. If $m = 1$ in (4.7) ($\mathbf{G}_t^0 = \mathbf{I}_p$), the test becomes the linearity test discussed in that section. The test can also be performed in the TR^2-form with the modification discussed in Section 4.1. When $p = 1$, it collapses into the corresponding test in Eitrheim and Teräsvirta (1996).

If the transition variables are not the same across equations, the test statistic (4.9) is still valid with $\mathbf{Z}_1$ defined in (3.14). If the model is a VTAR model, the test collapses into a linearity test. This is because, as noted in the previous section, the least squares estimators of the threshold parameters are super consistent and can be assumed fixed in performing the test. It can be regarded as a generalization of the corresponding test for self-exciting TAR models (Tong, 1990) by Strikholm and Teräsvirta (2006).

The statistic (4.9) may also be viewed as a specification test. This is due to the fact that in building LVSTAR models one must proceed 'bottom up', the identification problem being the reason for this. If a 'top down' strategy is applied, specification may well begin with a situation in which both the null model and the alternative are unidentified so that the LM type test cannot be applied.

4.3   Parameter constancy

The LVSTAR model (as any standard VAR model) is estimated assuming that its parameters are constant over time. This assumption is testable. As in the univariate case, considered in Eitrheim and Teräsvirta (1996), a useful alternative is that the parameters change smoothly over time. In the alternative model the parameter change is characterised using a logistic function. In this case, rescaled time $\tau = t/T$ is the transition variable for all equations. Consider the following LVSTAR model:

$$\mathbf{y}_t = \mathbf{\Psi}_t' \mathbf{B}(\tau)' \mathbf{x}_t + \boldsymbol{\varepsilon}_t, \tag{4.10}$$

where $\mathbf{B}(\tau) = \mathbf{B}_a + \mathbf{B}_b \lambda(\tau|\gamma, c)$. It is assumed that $\lambda(\tau|\gamma, c)$ can take one of the three following forms:

$$\lambda_1(\tau|\gamma, c) = (1 + \exp(-\gamma(\tau - c)))^{-1} - 1/2 \tag{4.11}$$

$$\lambda_2(\tau|\gamma, c) = (1 + \exp(-\gamma(\tau - c_1)(\tau - c_2)))^{-1} - 1/2 \tag{4.12}$$

$$\lambda_3(\tau|\gamma, c) = \left(1 + \exp\left(-\gamma \prod_{j=1}^{3}(\tau - c_j)\right)\right)^{-1} - 1/2. \tag{4.13}$$

The null hypothesis of parameter constancy is $\lambda_j(\tau|\gamma, c) = 0$, $j = 1$, 2, or 3. Function (4.11) postulates a smooth monotonic parameter change and function (4.12) a nonmonotonic change that is symmetric around $(c_1 + c_2)/2$. Function (4.13) describes an even more flexible, generally nonmonotonic, parameter change, but monotonic change appears as a special case when $c_1 = c_2 = c_3$. When $\gamma \to \infty$ in (4.11), the alternative becomes a single structural break. When the same occurs in (4.12), one obtains a special case of a double break if $c_1 \neq c_2$, whereas (4.13) implies a triple break but only two extreme regimes if $c_1 \neq c_2 \neq c_3$.

For notational simplicity, we only consider the function (4.11) here. The first order Taylor expansion of (4.11) becomes

$$\lambda_1(\tau|\gamma, c) = (\gamma/4)(\tau - c) + r_t \tag{4.14}$$

where r_t is the remainder. The model (4.10) is approximated and reparameterised as follows:

$$\mathbf{y}_t = \mathbf{\Psi}_t' \mathbf{B}_a' \mathbf{x}_t + \mathbf{\Psi}_t' \mathbf{B}_b' \mathbf{x}_t \tau + \boldsymbol{\varepsilon}_t^* \tag{4.15}$$

where $\boldsymbol{\varepsilon}_t^*$ also contains the remainder r_t from the Taylor expansion (4.14). Note, however, that under H$_0$, $\boldsymbol{\varepsilon}_t^* = \boldsymbol{\varepsilon}_t$. The new null hypothesis is

$$\text{H}_0\text{: } \mathbf{B}_b = \mathbf{0}. \tag{4.16}$$

The Lagrange multiplier test is derived from the score evaluated under the null hypothesis

$$\frac{\partial \log L(\tilde{\boldsymbol{\theta}})}{\partial \mathbf{B}_b} = \sum_{t=1}^T \left[\mathbf{x}_t \tau \left(\mathbf{y}_t - \tilde{\mathbf{\Psi}}_t' \tilde{\mathbf{B}}_a' \mathbf{x}_t \right)' \tilde{\mathbf{\Omega}}^{-1} \tilde{\mathbf{\Psi}}_t' \right],$$

where $\tilde{\mathbf{\Psi}}$, $\tilde{\mathbf{B}}_a$, and $\tilde{\mathbf{\Omega}}$ are estimates under the null. Denote $\tilde{\boldsymbol{\varepsilon}}_t = \mathbf{y}_t - \tilde{\mathbf{\Psi}}_t' \tilde{\mathbf{B}}_a' \mathbf{x}_t$, and $\tilde{\mathbf{E}} = (\tilde{\boldsymbol{\varepsilon}}_1, \tilde{\boldsymbol{\varepsilon}}_2, ..., \tilde{\boldsymbol{\varepsilon}}_T)'$. Let

$$\tilde{\mathbf{Z}} = \left[\begin{array}{cccc} \text{vec}(\tilde{\mathbf{\Psi}}_1 \otimes \mathbf{x}_1(1/T))' & \text{vec}(\tilde{\mathbf{\Psi}}_2 \otimes \mathbf{x}_2(2/T))' & ... & \text{vec}(\tilde{\mathbf{\Psi}}_T \otimes \mathbf{x}_T)' \end{array} \right]' \tag{4.17}$$

and

$$\tilde{\mathbf{K}} = \left[\begin{array}{cccc} \text{vec}[(\partial \tilde{\mathbf{\Psi}}_1' \tilde{\mathbf{B}}_a' \mathbf{x}_1 / \partial \theta)]' & \text{vec}[(\partial \tilde{\mathbf{\Psi}}_2' \tilde{\mathbf{B}}_a' \mathbf{x}_2 / \partial \theta)]' & ... & \text{vec}[(\partial \tilde{\mathbf{\Psi}}_T' \tilde{\mathbf{B}}_a' \mathbf{x}_T / \partial \theta)]' \end{array} \right]. \tag{4.18}$$

Again, we denote $\mathbf{P_K} = \tilde{\mathbf{K}}(\tilde{\mathbf{K}}'\tilde{\mathbf{K}})^{-1}\tilde{\mathbf{K}}'$.

We make the following assumption:

Assumption 7. $\tilde{\mathbf{K}}'\tilde{\mathbf{K}}$ *and* $\tilde{\mathbf{Z}}'(\mathbf{I}_T - \mathbf{P_K})\tilde{\mathbf{Z}}$ *are positive definite matrices.*

Arguments analogous to ones in Sections 4.1 and 4.2 lead to the following result:

Theorem 4. *Consider the model in (4.10) with* $\lambda_1(\tau|\gamma, c)$. *Suppose that Assumptions 2, 4 and 7 hold. The LM test statistic for testing the null hypothesis* $H_0 : \lambda_1(\tau) = 0$ *in (4.16) equals*

$$LM = \text{tr}\{\tilde{\mathbf{\Omega}}^{-1} \tilde{\mathbf{E}}' \tilde{\mathbf{Z}} [\tilde{\mathbf{Z}}'(\mathbf{I}_T - \mathbf{P_K})\tilde{\mathbf{Z}}]^{-1} \tilde{\mathbf{Z}}' \tilde{\mathbf{E}}\}. \tag{4.19}$$

When the null hypothesis is valid, the LM statistic (4.19) has an asymptotic χ^2 *distribution with* $mp^2(kp + 1)$ *degrees of freedom.*

The test can also be performed using the multiplicative TR^2-form in Section 3 by replacing $\mathbf{X}$ by $\tilde{\mathbf{K}}$ and $\mathbf{Z}_n$ by $\tilde{\mathbf{Z}}$. The empirical size distortion problem due to the non-orthogonality between $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{K}}$ can be alleviated as discussed in Section 4.1.

1. Estimate the LVSTAR model under the null hypothesis of constant parameters over time. Regress the residuals $\tilde{\mathbf{E}}$ on $\tilde{\mathbf{K}}$. Collect the residuals $\tilde{\mathbf{V}}$, and compute the matrix residual sum of squares $\tilde{\mathbf{V}}'\tilde{\mathbf{V}}$.

2. Run the auxiliary regression of $\tilde{\mathbf{V}}$ on $(\tilde{\mathbf{K}}, \tilde{\mathbf{Z}})$. Collect the residuals $\tilde{\tilde{\boldsymbol{\Xi}}}$, and the matrix residual sum of squares $\tilde{\tilde{\boldsymbol{\Xi}}}'\tilde{\tilde{\boldsymbol{\Xi}}}$.

3. Compute the test statistic (3.10).

This test can also be applied to subsets, which does not only mean subsets of equations or even single equations. It is often useful to focus on certain types of coefficients. For example, in a single equation it may be useful to test the constancy of the intercepts or other linear or nonlinear parameters separately, see Teräsvirta (1998) for discussion. This helps the modeller to locate possible weaknesses in the specification of the estimated model. This is particularly useful when the joint test rejects parameter constancy. Moreover, when the conditional mean (4.10) is linear, the test collapses into the corresponding parameter constancy test in a linear VAR model, see He et al. (2009).

5   Size simulations

5.1   The problem

A recurring problem in testing vector models is that the standard LM or LM type tests can be strongly oversized when the null hypothesis contains a large number of parameters and when at the same time the time series are not very long. Bartlett and Bartlett-type corrections have been used as a remedy to this problem. The idea is to rescale the degrees of freedom of the test and apply an F-statistic, see Laitinen (1978) and Meisner (1979). The Monte Carlo results of Bera et al. (1981) showed that the Laitinen-Meisner correction is likely to overcorrect the size.

We simulate both the LM test and its rescaled F-variant. The other two tests considered are Wilks's lambda and Rao's F-test, Rao (1951; 1965, Section 8c.5). The former is based on Wilks's Λ-distribution, see for example Anderson (2003, Section 8.3), while the latter is a function of Wilks's lambda. In Appendix B it is shown that Wilks's lambda is applicable in our testing situation and how the test is carried out in our framework.

5.2   *P*-value plot and *p*-value discrepancy plot

As already noted, the LM or LM-type test statistic has the advantage that estimation of the alternative model is avoided. In reporting results, we make use of the graphical

methods by Davidson and MacKinnon (1998). These authors suggested p-value and p-value discrepancy plots for the purpose. Consider a Monte Carlo experiment in which N realizations of some test statistic S are generated using a data-generating process (DGP) that is a special case of the null hypothesis (size experiments), or of the alternative (power experiments). Let $p_j = p(S_j)$ denote the p-value of the jth realization S_j, $j = 1, ..., N$. Both of the p-value and the p-value discrepancy plot are based on the empirical distribution function (EDF) of the p-values of the test statistic

$$\hat{F}(x_i) = \frac{1}{N} \sum_{j=1}^{N} I(p_j \leq x_i) \qquad (5.1)$$

where $I(p_j \leq x_i)$ is an indicator function, $x_i \in (0, x_{\max})$. In this work $x_{\max} = 0.2$ and the set of M values of x_i is

$$\{x_i\}_{i=1}^{M} = \{0.001, 0.002, ..., 0.010, 0.015, ..., 0.195, 0.200\}$$

where $M = 68$.

The p-value discrepancy plot is a scatterplot of $\hat{F}(x_i) - x_i$ against x_i, whereas the p-value plot is a scatterplot of $\hat{F}(x_i)$ against x_i. The former one is used in size experiments to describe the deviation of the empirical distribution from the nominal one, and the latter one in power simulations. For cases in which N, the number of replications, is not large, Davidson and MacKinnon (1998) suggested to smoothen the graphs. This will not be necessary here, as $N = 5000$ in our experiments.

We have to consider the randomness in the p-value discrepancy plot caused by a finite N. Davidson and MacKinnon (1998) employed the Kolmogorov-Smirnov (KS) statistic for the purpose. This statistic tends to be rather conservative, however, and we use a 95% two-sided confidence band assuming that the nominal distribution of the test statistic is normal. Under this assumption, for each $j = 1, ..., N$ the value of the indicator function $I(p_j \leq x_i)$ is a realization of a Bernoulli distributed random variable with parameter x_i. The distribution of the p-value discrepancy $\hat{F}(x_i) - x_i$ is thus approximated by a normal distribution with zero mean and variance $N^{-1} x_i (1 - x_i)$.

5.3 Linearity tests I: same transition variable for all equations

In order to study the finite sample size behaviour of our linearity tests we conduct a set of Monte Carlo experiments in which 5000 realizations of the four aforementioned test statistics are generated using DGPs that are special cases of the general null hypothesis $\Theta^{(n)} = \mathbf{0}$ in (3.4). More specifically, the DGP consists of p equations

$$y_{i,t} = \sum_{j=1}^{k} \rho^i y_{i,t-j} + \varepsilon_{i,t}, \ i = 1, ..., p \qquad (5.2)$$

where $\rho = 0.4$ and $\mathbf{\Omega} = \mathbf{I}$. The six designs to be considered are as follows:

$$
\begin{aligned}
&\text{Design 1:} \quad p = 2, \quad\ \ k = 1 \quad \text{and} \quad T = 30; \\
&\text{Design 2:} \quad p = 2, \quad\ \ k = 1 \quad \text{and} \quad T = 100; \\
&\text{Design 3:} \quad p = 2, \quad\ \ k = 2 \quad \text{and} \quad T = 30; \\
&\text{Design 4:} \quad p = 2, \quad\ \ k = 5 \quad \text{and} \quad T = 30; \\
&\text{Design 5:} \quad p = 5, \quad\ \ k = 1 \quad \text{and} \quad T = 50; \\
&\text{Design 6:} \quad p = 10, \quad k = 1 \quad \text{and} \quad T = 50.
\end{aligned}
$$

The p-value discrepancy plots for Designs 1 and 2 appear in Figure 1. Both designs are bivariate, and $k = 1$. The only difference between them is the sample size T. As for Design 1, both the LM test and the rescaled LM test, denoted by F in all graphs, are size-distorted. The former one over-rejects, whereas the latter under-rejects. This agrees with previous results in the literature. The empirical size of Wilks's Λ is very close to that of Rao's F-test, and neither is size-distorted. When $T = 100$, the LM test has an acceptable empirical size, whereas the rescaled LM test still under-rejects. This finding agrees with the Monte Carlo results of Bera et al. (1981) on the behaviour of the latter test in univariate models.

The p-value discrepancy plots for Designs 3 and 4 can be found in Figure 2. Both designs are bivariate and $T = 30$. The lag lengths equal two and five. The lag length has a strong impact on results. While the rescaled test and the LM test behave badly in both cases, Rao's F-test is the only one that still has no size distortion when $k = 5$. Wilks's Λ, a good performer in the first three designs, is now rather strongly oversized.

In Designs 5 and 6 the main object of interest is p, the dimension of the VAR model. Again, while the behaviour of Wilks's Λ is acceptable for $p = 5$, the test over-rejects when $p = 10$. The empirical size of Rao's F-test is practically unaffected by the change in the lag length from five to ten. Our conclusion is that of the tests inspected one should prefer Rao's F-test to the alternatives. This accords with the results in Edgerton and Shukur (1999) who considered testing for error autocorrelation in linear VAR models. In Shukur and Edgerton (2002) the test under scrutiny was the functional form specification test, RESET, by Ramsey (1969), and the conclusion was similar.

It is interesting to see whether the order of the Taylor expansion affects the empirical size of the best performing Rao's F-test. We report the p-value discrepancy plots of this test for Design 5 in Figure 4 with orders from one to four. The empirical size of the test remains rather stable when the order is increased. Further increases would of course ultimately lead to the lack of degrees of freedom and make the test unavailable.

Yet another way of correcting the size of the LM type test would be to obtain the critical values of the statistic by simulation. Since this automatically corrects the size, we

do not report any results of this procedure. In Teräsvirta and Yang (2014), the authors make the correction using the wild bootstrap, proposed originally by Wu (1986), because the errors in their applications are likely to contain conditional heteroskedasticity. This is a valid procedure if the appropriate conditions, see Gonçalves and Kilian (2004), are satisfied when the null hypothesis holds.

5.4 Linearity tests II: different transition variable for each equation

As already discussed, it is customary to test linearity equation by equation when each equation of the alternative LVSTAR model has a different transition variable. In Section 3.2 we have derived a joint test for the purpose. Yet another alternative, assuming that the errors of the null model are multivariate normal, would be to use the sum of the individual χ^2-statistics as the test statistic. This test, called the 'sum test' is, however, valid only if the errors are not correlated. In this section we simulate both the 'sum test' and the joint test without correcting the size.

The null model is the following five-dimensional stationary first-order VAR model:

$$\mathbf{y}_t = \mathbf{B}_1' \mathbf{y}_{t-1} + \varepsilon_t$$

where

$$\mathbf{B}_1' = \begin{bmatrix} 0.7 & -0.2 & 0.2 & -0.2 & 0.2 \\ 0.2 & 0.7 & -0.2 & 0.2 & -0.2 \\ -0.2 & 0.2 & 0.7 & -0.2 & 0.2 \\ -0.2 & 0.2 & -0.2 & 0.7 & 0.2 \\ 0.2 & -0.2 & 0.2 & -0.2 & 0.7 \end{bmatrix} \tag{5.3}$$

and the error correlation matrix is either an identity matrix, or

$$\mathrm{cov}(\varepsilon_t) = \begin{bmatrix} 1 & 2/\sqrt{5} & \sqrt{3/5} & \sqrt{2/5} & 1/\sqrt{5} \\ & 1 & \sqrt{3}/2 & 1/\sqrt{2} & 1/2 \\ & & 1 & \sqrt{2/3} & 1/\sqrt{3} \\ & & & 1 & 1/\sqrt{2} \\ & & & & 1 \end{bmatrix}. \tag{5.4}$$

The transition variable of the ith equation is assumed $y_{i,t-1}$, $i = 1, ..., 5$, so $\mathbf{S}_t = \mathrm{diag}(y_{1,t-1}, ..., y_{p,t-1})$ in (3.11). We generate 10000 replications from the null model at different sample sizes. We include the individual tests for each equation and their sum. The joint test is computed both from (3.10) using (3.14) and in stages. The results for the case where the covariance matrix is an identity matrix can be found in Table 1. The sum of the five asymptotically χ^2-distributed statistics with five degrees of freedom each

has an asymptotic χ^2-distribution with 25 degrees of freedom. Due to the large number of degrees of freedom the 'sum test' is oversized at smallest sample sizes but has a correct size for $T \geq 250$. The joint test is more oversized as its asymptotic null distribution has 75 degrees of freedom, so it takes longer for the empirical size to get close to the nominal one. Unlike the 'sum test' it does not make use of the fact that the errors are independent. We also carried out the joint test in stages, and the results were identical to the ones in the 'Joint' column of Table 1.

- insert Table 1 about here -

The results for the case where the error correlation matrix is (5.4) appear in Table 2. The sum statistic and the joint test are equally oversized for $T = 100$, but increasing the sample size reveals, as expected, that the former will over-reject even asymptotically due to ignoring the error correlations. Comparing the results in this table to the ones in Table 1 shows that the empirical size of the joint test is not affected by cross-correlated errors.

- insert Table 2 about here -

5.5   Misspecification tests

The basic DGP is a stable p-dimensional single transition LVSTAR model (3.1) with k lags such that $\mathbf{B}_{1j} = 0.4^j \mathbf{I}_p$, $\mathbf{B}_{2j} = -(0.2^j)\mathbf{I}_p$, and $\mathbf{G}_t = \text{diag}(g(y_{1,t-1}|\gamma_1, c_1), ..., g(y_{p,t-1}|\gamma_p, c_p))$, so each equation has its own transition variable. The slope and the location parameters are $\gamma_j = c_j = 1$, $j = 1, ..., p$, and $\mathbf{\Omega} = \mathbf{I}_p$. For each of the 5000 realizations, an LVSTAR model is estimated and a residual matrix computed. The values of the test statistic and the corresponding p-values are computed using the TR^2 form as in Section 4.

We consider the size distortion of the four tests. The performance of the misspecification tests depends on how accurate the parameter estimates are. Since LVSTAR models are estimated numerically, relatively large time series are required for reasonable estimation accuracy, and we choose $T = 200$ and $T = 500$.

P-value discrepancy plots of the test of no serial correlation appear in Figure 5. As can be expected, the LM test over-rejects, whereas the rescaled test under-rejects. The empirical size of Wilks's Λ test is very close to that of Rao's F-test, and neither test is size-distorted. With the increase of sample size from $T = 200$ to $T = 500$, the performance of both the LM test and the rescaled test improves but the improvement is not very large.

Figure 6 shows the p-value discrepancy plots for test of no additive nonlinearity. Both Wilks's Λ and Rao's F-test have a satisfactory empirical size. The rescaled test performs better than the LM test and shows no size distortion for $T = 500$. The corresponding plots for the test of parameter constancy can be found in Figure 7. For $T = 200$, Wilks's

Λ is slightly oversized, while Rao's F-test shows hardly any size distortion. The rescaled test and the LM test perform even worse.

Computing the three misspecification tests requires constructing the matrices $\mathbf{Z}$ and $\mathbf{K}$. In small samples their column dimensions affect the empirical size of the corresponding test statistic. In the three tests, $\mathbf{K}$ has the same column dimension, whereas those of $\mathbf{Z}$ are different. Choosing a large lag length J for testing serial correlation slows down the convergence of the standard LM test statistic to its limiting distribution. The column dimension of the matrix $\mathbf{Z}$ of the parameter constancy test is the squared size of a Kronecker product, and the test thus has the least favourable small sample performance. Differences in size properties between the tests can be explained by different column dimensions of $\mathbf{Z}$. However, Rao's F-test and the Wilks's Λ test are performing well throughout.

6   Power simulations

6.1   Each equation has the same transition variable

In order to study power properties of our linearity tests, we consider the two main testing situations. First, each equation has the same transition variable and second, each equation has a different transition variable. We compare the power performance of the joint test with its univariate counterpart and the 'sum test' with and without error correlations. For comparison we also include results on testing each equation individually.

In the first situation, our model is a five-dimensional first-order LVSTAR model (3.1) in which the parameter matrix $\mathbf{B}'_1$ is defined by (5.3), $\mathbf{B}'_1 + \mathbf{B}'_2 = 0.8\mathbf{I}_5$, $\mathbf{G}_t = g(s_t|\gamma, c)\mathbf{I}_5$, and the single transition function is

$$g(s_t|\gamma, c) = (1 + \exp\{-\gamma(s_t - 0.5)\})^{-1}.$$

The transition variable is generated by an exogenous first-order AR process $s_t = 0.95 s_{t-1} + \varepsilon_t^{(s)}$, where $\varepsilon_t^{(s)}$ is standard normal. The slope parameter γ of the transition function equals 0.5 when the errors are uncorrelated and 0.25 when the correlation matrix is (5.4). The location parameter $c = 0.5$ in both cases and $N = 10000$. Since we are interested in power we size-correct the joint test by using Rao's F statistic.

Table 3 contains the results for uncorrelated errors. The 'sum test' has higher power than Rao's F-test when $T \leq 100$. The individual tests have weaker power than the joint test, the test for the fifth equation being an exception. The joint test, as may be expected, is preferable to the individual ones.

- insert Table 3 about here -

Table 4 contains the results for correlated errors. In this case, the 'sum test' is size distorted. Nevertheless, its power is not comparable to that of Rao's F test. This test now has a clear edge over the 'sum test' even if the latter is not size corrected. The decrease in power of individual tests when moving from $T = 50$ to $T = 100$ is due to the fact that the the asymptotic χ^2-statistics are oversized for $T = 50$. In small samples, a standard F test could be preferred to the asymptotic one.

- insert Table 4 about here -

6.2   Local asymptotic power of the test in a special case

In order to consider the local asymptotic power of our linearity test (B) we choose a special case with $n = 1$, so H_0: $\boldsymbol{\Theta}_1 = \mathbf{B}_2\mathbf{D}_1 = \mathbf{0}$. For simplicity, assume that under H_1 the parameters of the transition function are $\gamma_i = \widetilde{\gamma}/\sqrt{T}$, and $c_i = c$, $i = 1, .., p$, where $\widetilde{\gamma} > 0$. Since

$$g(s_t|\widetilde{\gamma}/\sqrt{T}, c) = 1/2 + (4\sqrt{T})^{-1}(s_t - c)\widetilde{\gamma} + r_{jt} \tag{6.1}$$

this implies, ignoring $r_{jt} = O(T^{-1})$, that H_1: $\boldsymbol{\Theta}_1 = \mathbf{B}_2\mathbf{D}_1 = \widetilde{\gamma}(4\sqrt{T})^{-1}\mathbf{B}_2\widetilde{\mathbf{D}}_1$, where $\widetilde{\mathbf{D}}_1 = \mathrm{diag}(\widetilde{d}_{11}, ..., \widetilde{d}_{p1})$ with $\widetilde{d}_{i1} \neq 0$, $i = 1, .., p$. Assuming as before that the errors are normal, we have

$$\mathrm{vec}(\widetilde{\mathbf{E}}') \sim \mathcal{N}(\widetilde{\gamma}(4\sqrt{T})^{-1}\mathrm{vec}(\widetilde{\mathbf{D}}_1\mathbf{B}_2'), \mathbf{I}_T \otimes \boldsymbol{\Omega}).$$

Under H_1, $\mathrm{vec}(\widetilde{\mathbf{E}}')$ deviates locally from the null vector, and the test statistic has the form

$$
\begin{aligned}
LM &= \mathrm{vec}(\widetilde{\mathbf{E}}'\mathbf{Z}_T)' \left((\mathbf{Z}_T'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}) \otimes \widetilde{\boldsymbol{\Omega}} \right)^{-1} \mathrm{vec}(\widetilde{\mathbf{E}}'\mathbf{Z}_T) \\
&= \widetilde{\gamma}^2(16T)^{-1}\mathrm{vec}(\widetilde{\mathbf{D}}_1\mathbf{B}_2'\mathbf{Z}_T)' \left((\mathbf{Z}_T'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_T) \otimes \widetilde{\boldsymbol{\Omega}} \right)^{-1} \mathrm{vec}(\widetilde{\mathbf{D}}_1\mathbf{B}_2'\mathbf{Z}_T) \\
&= (\widetilde{\gamma}^2/16)\mathrm{vec}(T^{-1}\widetilde{\mathbf{D}}_1\mathbf{B}_2'\mathbf{Z}_T)' \left(T^{-1}(\mathbf{Z}_T'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_T) \otimes \widetilde{\boldsymbol{\Omega}} \right)^{-1} \mathrm{vec}(T^{-1}\widetilde{\mathbf{D}}_1\mathbf{B}_2'\mathbf{Z}_T).
\end{aligned}
\tag{6.2}
$$

The notation $\mathbf{Z}_T$ implies that the intercept in $\mathbf{Z}_T$ is a function of T as (6.1) contains the asymptotically vanishing term $-(4\sqrt{T})^{-1}c\widetilde{\gamma}$ of order zero. The asymptotic distribution of (6.2) is a noncentral χ^2-distribution with $np(kp + 1)$ degrees of freedom and the noncentrality parameter $(\widetilde{\gamma}^2/16)\mathbf{m}'(\mathbf{M} \otimes \boldsymbol{\Omega})^{-1}\mathbf{m}$, where $\mathbf{m} = \mathrm{vec}(\mathrm{plim}_{T\to\infty} T^{-1}\widetilde{\mathbf{D}}_1\mathbf{B}_2'\mathbf{Z}_T)$ and $\mathbf{M} = \mathrm{plim}_{T\to\infty} T^{-1}(\mathbf{Z}_T'(\mathbf{I} - \mathbf{P}_\mathbf{X})\mathbf{Z}_T)$. Locally, other things equal, the asymptotic power of the test is an increasing function of the slope parameter $\widetilde{\gamma}$ but also depends on the parameter matrix $\mathbf{B}_2$.

6.3   Each equation has a different transition variable

In these simulations the model is still a five-dimensional first-order LVSTAR model (3.1). We employ the same setup as in Section 6.1 except that

$$\mathbf{G}_t = \text{diag}\{g(y_{1,t-1}|\gamma,c), ..., g(y_{5,t-1}|\gamma,c)\}$$

and the transition functions are

$$g(y_{i,t-1}|\gamma,c) = (1 + \exp\{-\gamma(y_{i,t-1} - 0.5)\})^{-1}$$

for $i = 1, ..., 5$. This implies $\mathbf{S}_t = \text{diag}(y_{1,t-1}, ..., y_{5,t-1})$ in the auxiliary VAR equation (3.11). The errors are either independent standard normal or correlated with the covariance matrix (5.4). In the former case, $\gamma = 0.5$, in the latter, $\gamma = 0.25$. The results, based on $N = 10000$, appear in Table 5 for the uncorrelated and in Table 6 for the correlated errors. As in Section 5.4, the power of the individual tests, that of their sum, and the power of the joint test are reported. The 'sum test' is not size-corrected. The estimated bias of the test in this simulation can be deduced from Table 2.

<div align="center">

- insert Table 5 about here -
- insert Table 6 about here -

</div>

The results for the model with uncorrelated errors can be found in Table 5. The decrease in the power from $T = 50$ to $T = 100$ is again explained by the fact that the tests are somewhat size distorted at the smallest sample size. The same seed is used in both size and power simulations, so Tables 1 and 5 are fully comparable. It appears that even Rao's F-test is slightly size distorted for $T = 50$ but less than its χ^2-counterpart (results not reported). The empirical powers of these two tests are identical when $T \geq 250$. Table 5 shows that the gain in power from knowing that the errors are uncorrelated only becomes evident for $T \geq 250$.

Table 6 contains the results for the case when the error correlation matrix is (5.4). They indicate that the joint test is vastly superior to the 'sum test' despite the fact that the latter is not size corrected. The conclusion is that it is an advantage to test all equations jointly, especially if it cannot assumed that the errors are uncorrelated. The only problem is that in some applications the number of degrees of freedom may not be sufficient for carrying out the joint test. In that case, testing linearity of each equation individually is the remaining possibility. The 'sum test' is inconsistent whenever the errors are correlated.

7   Concluding remarks

In this paper, we propose LM and LM type linearity and misspecification tests for the LVSTR framework. We allow the dimension of the model to exceed two and, furthermore, do not restrict the number of transitions to one. We consider both the case in which the LVSTAR model only has a single transition variable, and the case in which every equation has its own (known) transition variable.

We extend three misspecification tests for univariate STAR models to the vector case. These are the test of no serial error correlation, the test of no additive nonlinearity and the parameter constancy test. They are either LM or LM type tests. The first one of these has already been considered in the bivariate case by Camacho (2004). We generalize the univariate misspecification tests in Eitrheim and Teräsvirta (1996) to multivariate joint tests. These tests should form an important part of the toolbox of practitioners building LVSTAR models. Many of them may also be used in building VTAR models.

Small-sample properties of the tests are of interest because they are affected by the dimension of the model. We find that the standard LM tests are severely size-distorted when the dimension of the system increases. Wilks's Λ statistic and Rao's F statistic that have satisfying size properties, the latter in particular, are recommended for empirical use. Nevertheless, the size of the LM test can be corrected by an appropriate bootstrapped version, although this has not been explicitly considered in this work. The power experiments demonstrate that the joint test is more powerful in finite sample if the errors are correlated than any other alternative.

References

Anderson, T. W.: 2003, *An Introduction to Multivariate Statistical Analysis*, 3rd edn, Wiley, New York.

Bartlett, M. S.: 1954, A note on multiplying factors for various χ^2 approximations, *Journal of the Royal Statistical Society. Series B* **16**, 296–298.

Bera, A. K., Byron, R. P. and Jarque, C. M.: 1981, Further evidence on asymptotic tests for homogeneity and symmetry in large demand systems, *Economics Letters* **8**, 101–105.

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.

Camacho, M.: 2004, Vector smooth transition regression models for US GDP and the composite index of leading indicators, *Journal of Forecasting* **23**, 173–196.

Davidson, R. and MacKinnon, J. G.: 1998, Graphical methods for investigating the size

References

and power of hypothesis tests, *The Manchester School of Economic & Social Studies* **66**, 1–26.

Edgerton, D. and Shukur, G.: 1999, Testing autocorrelation in a system perspective, *Econometric Reviews* **18**, 343–386.

Eitrheim, Ø. and Teräsvirta, T.: 1996, Testing the adequacy of smooth transition autoregressive models, *Journal of Econometrics* **74**, 59–75.

Godfrey, L. G.: 1988, *Misspecification tests in econometrics*, Cambridge University Press, Cambridge.

Gonçalves, S. and Kilian, L.: 2004, Bootstrapping autoregressions with conditional heteroskedasticity of unknown form, *Journal of Econometrics* **123**, 89–120.

Hansen, B. E.: 1996, Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* **64**, 413–430.

He, C., Teräsvirta, T. and González, A.: 2009, Testing parameter constancy in vector autoregressive models against continuous change, *Econometric Reviews* **28**, 225–245.

Hubrich, K. and Teräsvirta, T.: 2013, Thresholds and smooth transitions in vector autoregressive models, *Advances in Econometrics* **32**, 273–326.

Laitinen, K.: 1978, Why is demand homogeneity so often rejected?, *Economics Letters* **1**, 187–191.

Lütkepohl, H.: 1996, *Handbook of Matrices*, Wiley, New York.

Lütkepohl, H.: 2006, *New Introduction to Multiple Time Series Analysis*, Springer, New York.

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

Mardia, K. V., Kent, J. T. and Bibby, J. M.: 1979, *Multivariate Analysis*, Academic Press, San Diego.

Meisner, J. F.: 1979, The sad fate of the asymptotic Slutsky symmetry test for large systems, *Economics Letters* **2**, 231–233.

Ramsey, J. B.: 1969, Test for specification error in classical linear least squares regression analysis, *Journal of the Royal Statistical Society, Series B* **31**, 350–371.

Rao, C. R.: 1951, An asymptotic expansion of the distribution of Wilks' criterion, *Bulletin of the International Statistical Institute* **33**, 177–180.

Rao, C. R.: 1965, *Linear Statistical Inference and Its Applications*, 1st edn, New York: Wiley.

Saikkonen, P.: 2008, Stability of regime switching error correction models under linear cointegration, *Econometric Theory* **24**, 294–318.

Saikkonen, P. and Luukkonen, R.: 1988, Lagrange multiplier tests for testing non-linearities in time series models, *Scandinavian Journal of Statistics* **15**, 55–68.

Shukur, G. and Edgerton, D.: 2002, The small sample properties of the reset test as applied to systems of equations, *Journal of Statistical Computation and Simulation* **72**, 909–924.

Strikholm, B. and Teräsvirta, T.: 2006, A sequential procedure for determining the number of regimes in a threshold autoregressive model, *Econometrics Journal* **9**, 472–491.

Teräsvirta, T.: 1998, Modeling economic relationships with smooth transition regressions, *in* A. Ullah and D. E. Giles (eds), *Handbook of Applied Economic Statistics*, Dekker, New York, pp. 507–552.

Teräsvirta, T., Tjøstheim, D. and Granger, C. W. J.: 2010, *Modelling Nonlinear Economic Time Series*, Oxford University Press, Oxford.

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

Tong, H.: 1990, *Non-linear time series: a dynamical system approach*, Oxford University Press, Oxford.

Tsay, R. S.: 1998, Testing and modeling multivariate threshold models, *Journal of the American Statistical Association* **93**, 1188–1202.

Wu, C. F. J.: 1986, Jackknife, bootstrap and other resampling methods in regression analysis, *The Annals of Statistics* **14**, 1261–1295.

A   The first-order partial derivatives of $\mathbf{\Psi}'_t\mathbf{B}'\mathbf{x}_t$

The vectorised first order derivative of $\mathbf{\Psi}'_t\mathbf{B}'\mathbf{x}_t$ with respect to parameters θ can be easily found in both univariate and multivariate cases, see Eitrheim and Teräsvirta (1996). The set of parameters θ consists of $\mathbf{B}$, $\mathbf{\Omega}$, $\mathbf{\Gamma}$ and $\mathbf{C}$, where $\mathbf{B} = [b_{ij}]$, $\mathbf{\Gamma} = [\gamma_{ij}]$ and $\mathbf{C} = [c_{ij}]$.

For parameter $\mathbf{B} = [b_{ij}]$, we have

$$\frac{\partial \mathbf{\Psi}'_t\mathbf{B}'\mathbf{x}_t}{\partial b_{ij}} \;=\; \mathbf{\Psi}'_t\mathbf{H}'_{ij}\mathbf{x}_t \tag{A.1}$$

where $\mathbf{H}_{ij} = [h_{kl}]$ is a matrix in which $h_{ij} = 1$ and $h_{kl} = 0$ for $k \neq i$ and $l \neq j$. Vector (A.1) is the directional derivative of the vector $\mathbf{\Psi}'_t\mathbf{B}'\mathbf{x}_t$ with respect to the unit length matrix $\mathbf{H}_{ij}$.

For the parameter matrices $\mathbf{\Gamma} = [\gamma_{ij}]$ and $\mathbf{C} = [c_{ij}]$, letting $\delta_{ij} = \gamma_{ij}, c_{ij}$, we have

$$\frac{\partial \mathbf{\Psi}'_t\mathbf{B}'\mathbf{x}_t}{\partial \delta_{ij}} \;=\; (\; \mathbf{0}_p \;,\; ... \;,\; \frac{\partial \mathbf{G}^i_t}{\partial \delta_{ij}} \;,\; ...\mathbf{0}_p \;)\mathbf{B}'\mathbf{x}_t = \frac{\partial \mathbf{G}^i_t}{\partial \delta_{ij}}\mathbf{B}'_{i+1}\mathbf{x}_t \tag{A.2}$$

for $i = 1, ..., m-1$, where

$$\frac{\partial \mathbf{G}^i_t}{\partial \delta_{ij}} \;=\; \mathrm{diag}\{\; 0 \;,\; ... \;,\; \frac{\partial g^{ij}_t}{\partial \delta_{ij}} \;,... \;,\; 0 \;\} \tag{A.3}$$

for $j = 1, ..., p$. When $\delta_{ij} = \gamma_{ij}$,

$$\frac{\partial g_t^{ij}}{\partial \gamma_{ij}} = (g_t^{ij})^2 \exp\{-\gamma_{ij}(s_t - c_{ij})\}(s_t - c_{ij}) = (s_t - c_{ij})\, g_t^{ij}(1 - g_t^{ij}) \tag{A.4}$$

and when $\delta_{ij} = c_{ij}$,

$$\frac{\partial g_t^{ij}}{\partial c_{ij}} = -(g_t^{ij})^2 \exp\{-\gamma_{ij}(s_t - c_{ij})\}\gamma_{ij} = -\gamma_{ij}\, g_t^{ij}(1 - g_t^{ij}). \tag{A.5}$$

Finally,

$$\frac{\partial \boldsymbol{\Psi}_t' \mathbf{B}' \mathbf{x}_t}{\partial \boldsymbol{\Omega}} = \mathbf{0}. \tag{A.6}$$

The dimension of the first-order derivative of $\boldsymbol{\Psi}_t' \mathbf{B}' \mathbf{x}_t$ with respect to θ is $p \times [(kp + q)mp + 2(m-1)p]$.

B   Linearity test statistic with improved size

Bartlett and Bartlett-type corrections have been widely used as a possible remedy to size problems in LM-type tests. The Laitinen-Meisner correction consists of a degrees of freedom rescaling of the form $(pT - K)/(G \times pT)$, where p and T are as before, K is the number of parameters, and G the number of restrictions, see Laitinen (1978) and Meisner (1979). The F-type LM test statistic, or rescaled LM test statistic, can be computed as follows

$$F = \frac{(pT - K)}{G \times pT} LM. \tag{B.1}$$

The rescaled test statistic is assumed to follow an $F(G, pT - K)$ distribution. In the following, it will be called the rescaled LM test. The Monte Carlo results of Bera et al. (1981) show that the Laitinen-Meisner correction is likely to overcorrect the size.

We consider an improvement based on the so-called Wilks's Λ-distribution, and we call it Wilks's statistic or Wilks's Λ. Before introducing it we state the following result:

Theorem 5. *Let* $\mathbf{RSS}_j$, $j = 0, 1$, *be the* $p \times p$ *residual sum of squares matrix from the restricted regression ($j = 0$) and the auxiliary regression ($j = 1$). Furthermore, let* $\mathbf{W}_1 = \mathbf{RSS}_0 - \mathbf{RSS}_1$, *and* $\mathbf{W}_2 = \mathbf{RSS}_1$. *Under the null hypothesis of linearity,* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *are two independent Wishart distributed random matrices:*

$$\mathbf{W}_1 \sim \mathcal{W}_p\Big(\boldsymbol{\Omega}, \mathrm{cd}(\mathbf{Z})\Big) \quad \mathbf{W}_2 \sim \mathcal{W}_p\Big(\boldsymbol{\Omega}, T - \mathrm{cd}(\mathbf{X}) - \mathrm{cd}(\mathbf{Z})\Big) \tag{B.2}$$

where $\mathrm{cd}(\mathbf{A})$ *is the column dimension of the matrix* $\mathbf{A}$.

The proof is omitted.

Matrix $\mathbf{Z}$ in (B.2) corresponds to $\mathbf{Z}_n$ in (3.9) and $\boldsymbol{\Omega}$ is the covariance matrix of errors under H_0. It is worth stressing that, in the special case $p = 1$, the two independent Wishart variables $\mathbf{W}_1$ and $\mathbf{W}_2$ become scalars and χ^2-distributed, which implies an F-test.

Wilks's Λ-distribution is defined as follows:

Definition 1. *When* $\mathbf{A} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, m)$ *and* $\mathbf{B} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, n)$ *are independent,* $\boldsymbol{\Sigma}$ *is a* $p \times p$ *positive definite matrix,* $m \geq p$,

$$\Lambda = |\mathbf{A}|/|\mathbf{A} + \mathbf{B}| = |\mathbf{I}_p + \mathbf{A}^{-1}\mathbf{B}|^{-1} \sim \mathcal{L}(p, m, n) \tag{B.3}$$

has a Wilks's Λ*-distribution with parameters* p, m, *and* n.

The above definition is a variant of Definition 3.7.1 in Mardia et al. (1979). Anderson (2003, Section 8.3) and Mardia et al. (1979) contain a detailed discussion of the Wilks's Λ distribution. The distribution is invariant under changes in the covariance matrix $\boldsymbol{\Sigma}$. Wilks's Λ statistic has the following form:

$$\Lambda = |\mathbf{W}_2|/|\mathbf{W}_2 + \mathbf{W}_1| = |\mathbf{RSS}_1|/|\mathbf{RSS}_0|. \tag{B.4}$$

It follows Wilks's Λ-distribution $\mathcal{L}(\, p, T - \mathrm{cd}(\mathbf{X}) - \mathrm{cd}(\mathbf{Z}), \mathrm{cd}(\mathbf{Z}) \,)$ under linearity. If T is large, we may use Bartlett's approximation

$$\lambda = \left(\frac{1}{2}(p + \mathrm{cd}(\mathbf{Z}) + 1) + \mathrm{cd}(\mathbf{X}) - T \right) \log \Lambda \sim \chi^2_{\mathrm{cd}(\mathbf{Z})p} \tag{B.5}$$

see Bartlett (1954) and Anderson (2003, Section 8.3). The value of the test statistic can be computed by performing steps 1 and 2 outlined in the algorithm in Section 3 but computing the value of the test statistic defined in (B.4) and (B.5) instead of step 3.

C   Tables and figures

C   Tables and figures

T	Test						
	1	2	3	4	5	Sum	Joint
50	0.058	0.061	0.070	0.077	0.069	0.096	0.186
100	0.050	0.049	0.055	0.063	0.062	0.067	0.116
250	0.050	0.048	0.049	0.058	0.055	0.053	0.074
500	0.048	0.049	0.052	0.054	0.053	0.053	0.059
1000	0.051	0.053	0.054	0.051	0.050	0.052	0.053

Table 1: p-values of the LM-type linearity test at various sample sizes for individual equations 1-5, the sum of the χ^2-statistics ('Sum') and the joint statistic (3.10) ('Joint'). Same transition variable for each equation. Uncorrelated errors. The nominal significance level equals 0.05.

T	Test						
	1	2	3	4	5	Sum	Joint
50	0.068	0.055	0.058	0.069	0.070	0.149	0.183
100	0.047	0.045	0.050	0.059	0.053	0.104	0.110
250	0.042	0.046	0.042	0.052	0.051	0.095	0.073
500	0.049	0.046	0.046	0.050	0.047	0.093	0.058
1000	0.045	0.051	0.045	0.053	0.049	0.095	0.056

Table 2: p-values of the LM-type linearity test at various sample sizes for individual equations 1-5, the sum of the χ^2-statistics ('Sum'), and the joint statistic (3.10) ('Joint'). Same transition variable for each equation. Correlated errors. The nominal significance level equals 0.05.

T	Test						
	1	2	3	4	5	Sum	Joint
50	0.107	0.109	0.118	0.121	0.363	0.356	0.140
100	0.084	0.099	0.128	0.153	0.771	0.646	0.551
250	0.080	0.130	0.218	0.314	0.998	0.990	0.989
500	0.091	0.239	0.442	0.643	1.000	1.000	1.000
1000	0.120	0.470	0.792	0.941	1.000	1.000	1.000

Table 3: Power of the LM-type linearity test at various sample sizes for individual equations 1-5, the sum of the statistics ('Sum'), and the joint Rao's F statistic ('Joint'). AR(1) exogenous transition variable. Uncorrelated errors. The nominal significance level equals 0.05.

T	Test						
	1	2	3	4	5	Sum	Joint
50	0.107	0.108	0.112	0.110	0.134	0.271	0.160
100	0.080	0.090	0.102	0.096	0.218	0.255	0.631
250	0.068	0.104	0.107	0.182	0.587	0.461	0.995
500	0.074	0.157	0.160	0.179	0.922	0.845	1.000
1000	0.084	0.278	0.314	0.341	0.999	0.999	1.000

Table 4: Power of the LM-type linearity test at various sample sizes for individual equations 1-5, the sum of the statistics ('Sum'), and the joint Rao's F statistic ('Joint'). AR(1) exogenous transition variable. Correlated errors. The nominal significance level equals 0.05.

T	Test						
	1	2	3	4	5	Sum	Joint
50	0.063	0.066	0.076	0.088	0.083	0.122	0.142
100	0.054	0.058	0.073	0.111	0.099	0.128	0.149
250	0.060	0.078	0.119	0.227	0.188	0.280	0.206
500	0.080	0.132	0.233	0.463	0.388	0.635	0.403
1000	0.123	0.245	0.470	0.815	0.730	0.966	0.801

Table 5: Power of the LM-type linearity test at various sample sizes for individual equations 1-5 , the sum of the statistics ('Sum') and the joint Rao's F statistic ('Joint'). The data-generating process is a first-order LVSTAR model with uncorrelated errors. Each equation has its own transition variable. The nominal significance level equals 0.05.

T	Test						
	1	2	3	4	5	Sum	Joint
50	0.071	0.067	0.065	0.072	0.103	0.174	0.198
100	0.052	0.071	0.055	0.078	0.129	0.156	0.282
250	0.054	0.149	0.056	0.128	0.299	0.290	0.684
500	0.084	0.336	0.076	0.237	0.603	0.613	0.979
1000	0.137	0.669	0.115	0.496	0.924	0.970	1.000

Table 6: Power of the LM-type linearity test at various sample sizes for individual equations 1- 5, the sum of the statistics ('Sum'), and the joint Rao's F statistic ('Joint'). The data-generating process is a first-order LVSTAR model with correlated errors. Each equation has its own transition variable. The nominal significance level equals 0.05.

Figure 1: Size discrepancy plots for Design 1(top, $p = 2$, $k = 1$ and $T = 30$) and Design 2 (bottom, $p = 2$, $k = 1$ and $T = 100$). The dotted lines represent the upper 95% confidence bound (top), zero line (middle) and the lower 95% confidence bound (bottom).
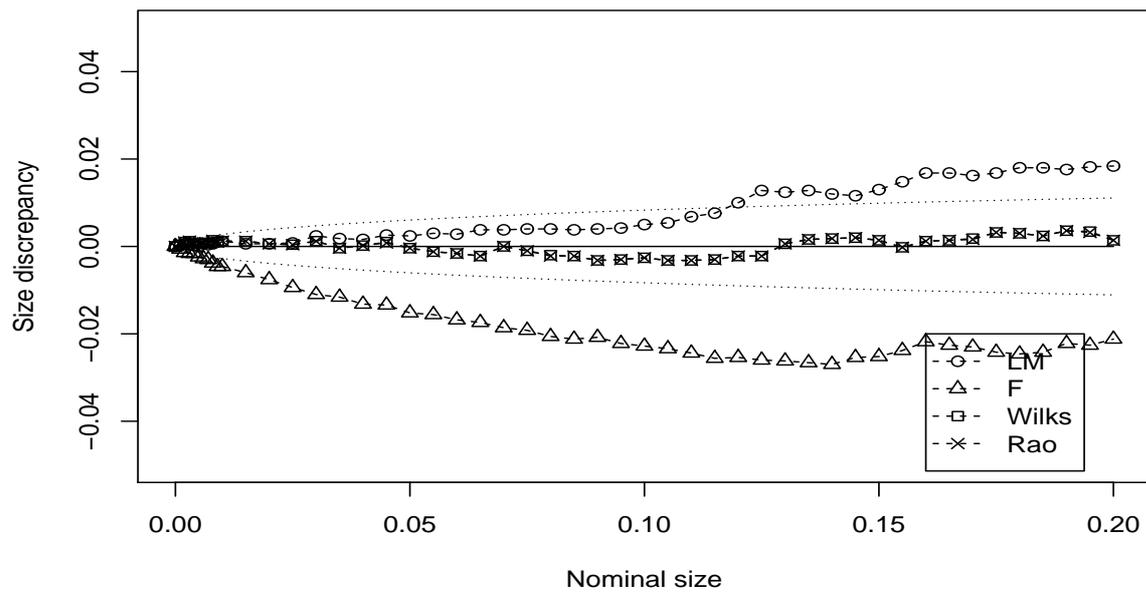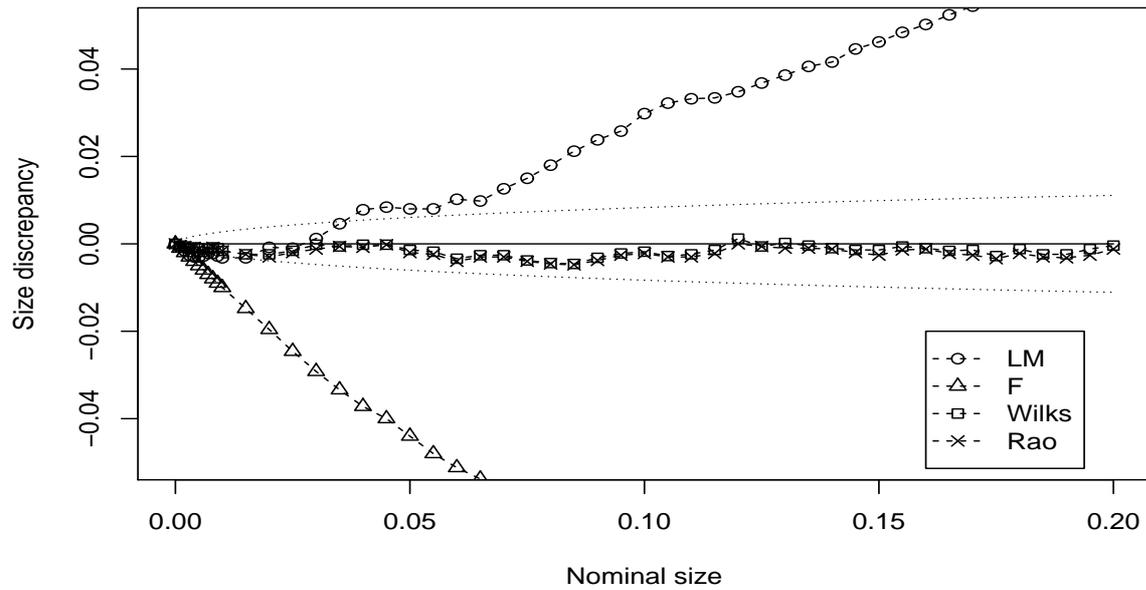
Figure 2: Size discrepancy plots for Design 3 (top, $p = 2$, $k = 2$ and $T = 30$) and Design 4 (bottom, $p = 2$, $k = 5$ and $T = 30$). The dotted lines represent the upper 95% confidence bound (top), zero line (middle) and the lower 95% confidence bound (bottom).
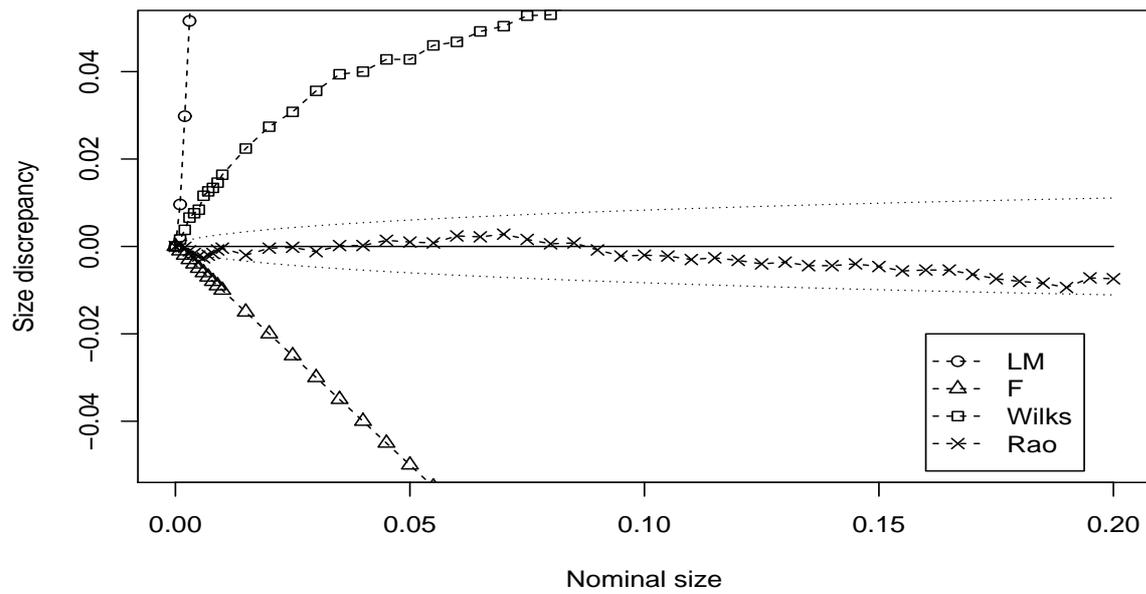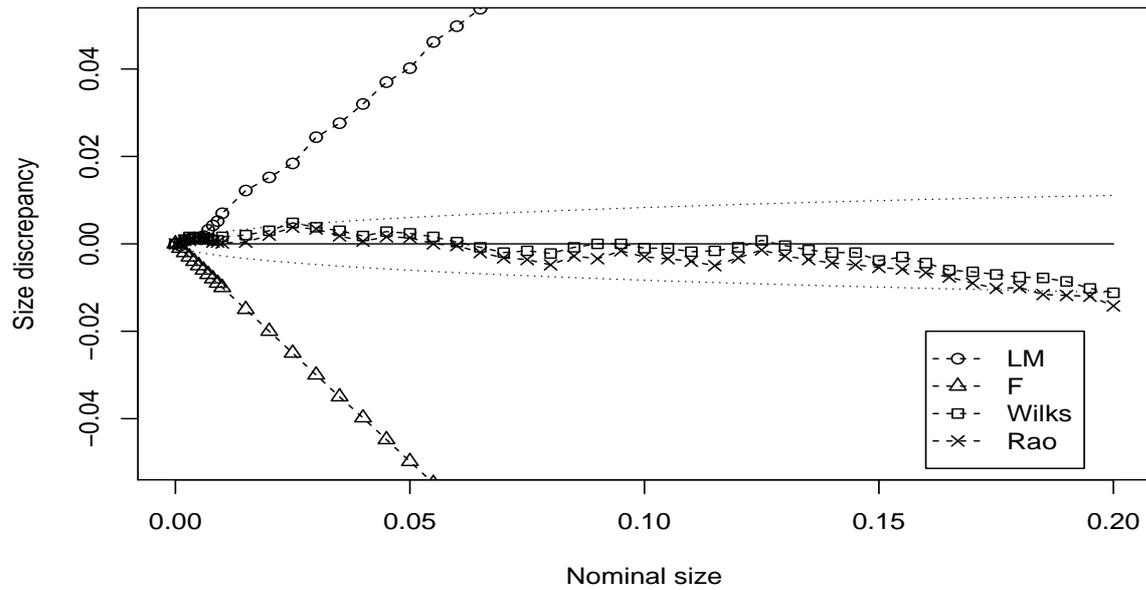
Figure 3: Size discrepancy plots for Design 5 (top, $p = 5$, $k = 1$ and $T = 50$) and Design 6 (bottom, $p = 10$, $k = 1$ and $T = 50$). The dotted lines represent the upper 95% confidence bound (top), zero line (middle) and the lower 95% confidence bound (bottom).
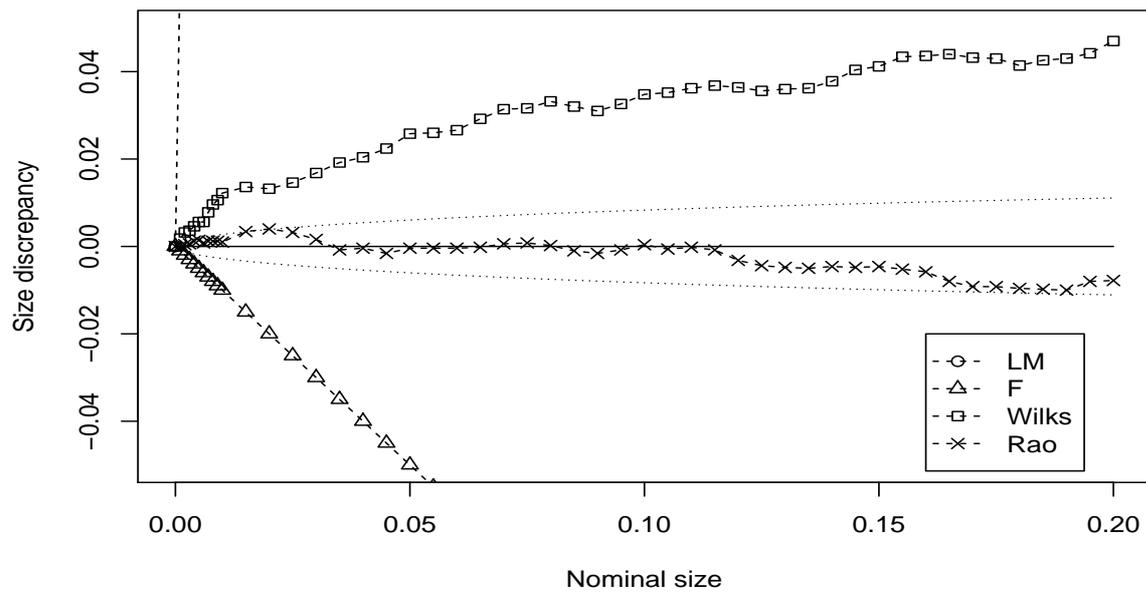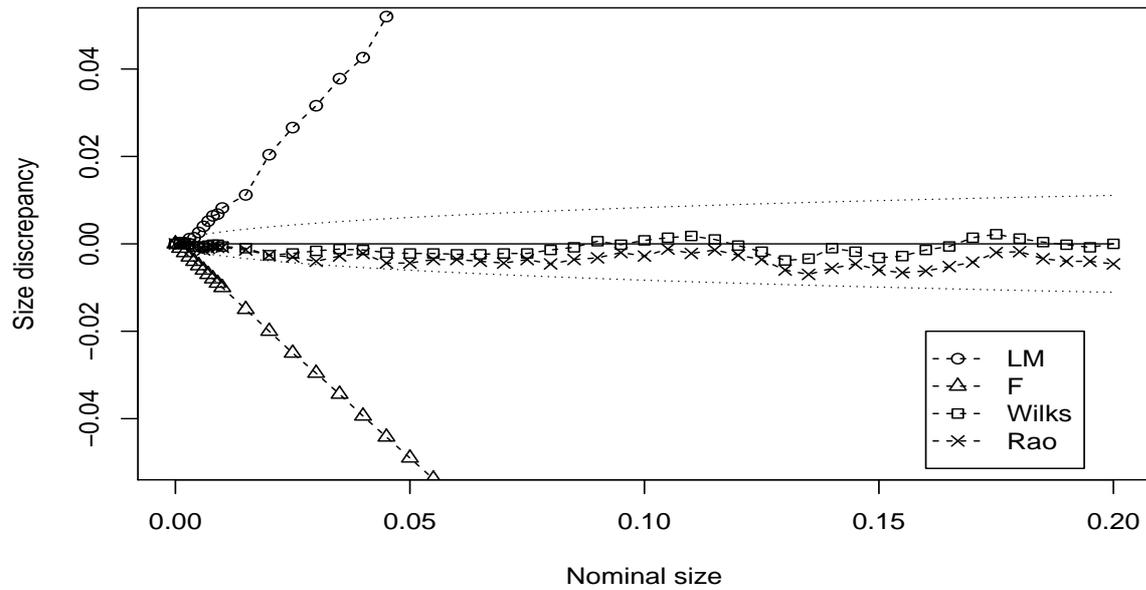
Figure 4: Size discrepancy plot for Design 5 ($p = 5$, $k = 1$ and $T = 50$) with Taylor expansion orders from one to four. Rao's F test. The dotted lines represent the upper 95% confidence bound (top), zero line (middle) and the lower 95% confidence bound (bottom).
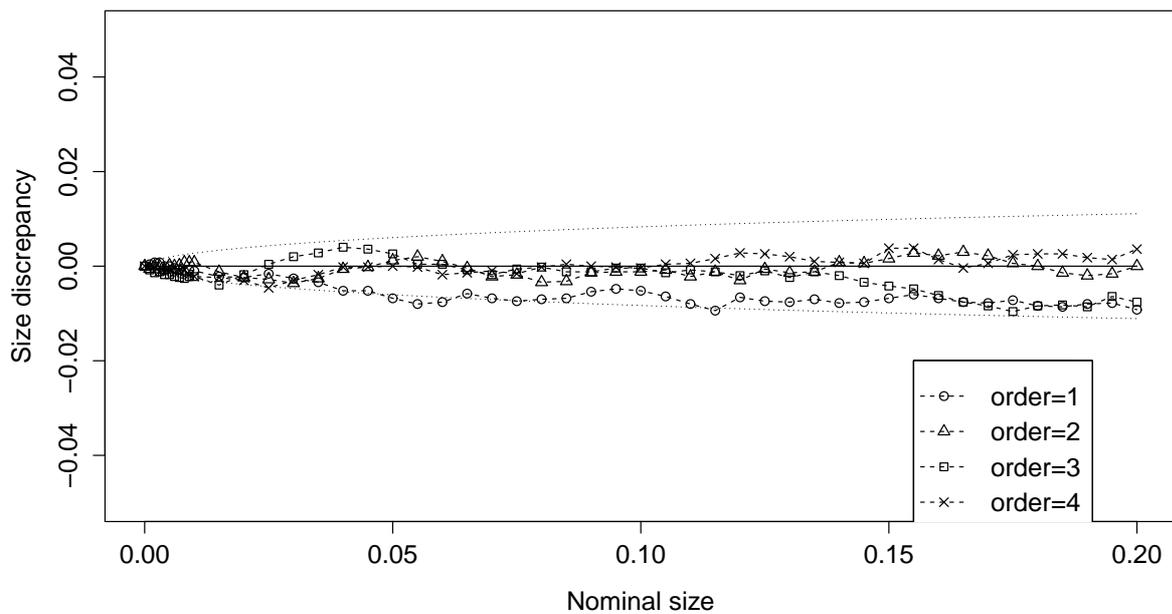
Figure 5: Size discrepancy plots for tests of no serial correlation: $p = 2$, $k = 2$ and $T = 200$ (Top); $T = 500$ (Bottom). The dotted lines represent the upper 95% confidence bound (top), zero line (middle) and the lower 95% confidence bound (bottom).
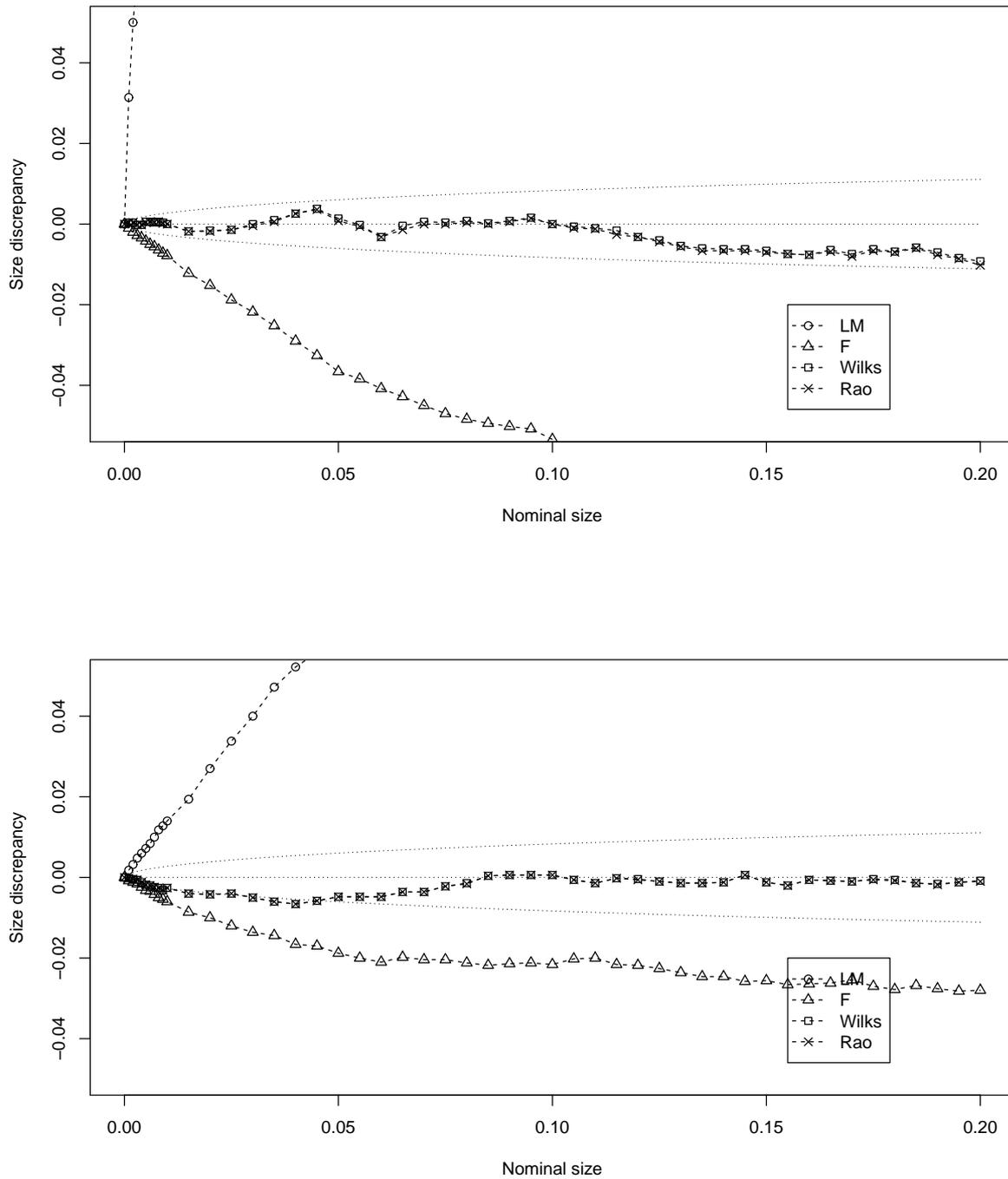
Figure 6: Size discrepancy plots for tests of no additive nonlinearity: $p = 2$, $k = 2$ and $T = 200$ (Top); $T = 500$ (Bottom). The dotted lines represent the upper 95% confidence bound (top), zero line (middle) and the lower 95% confidence bound (bottom).
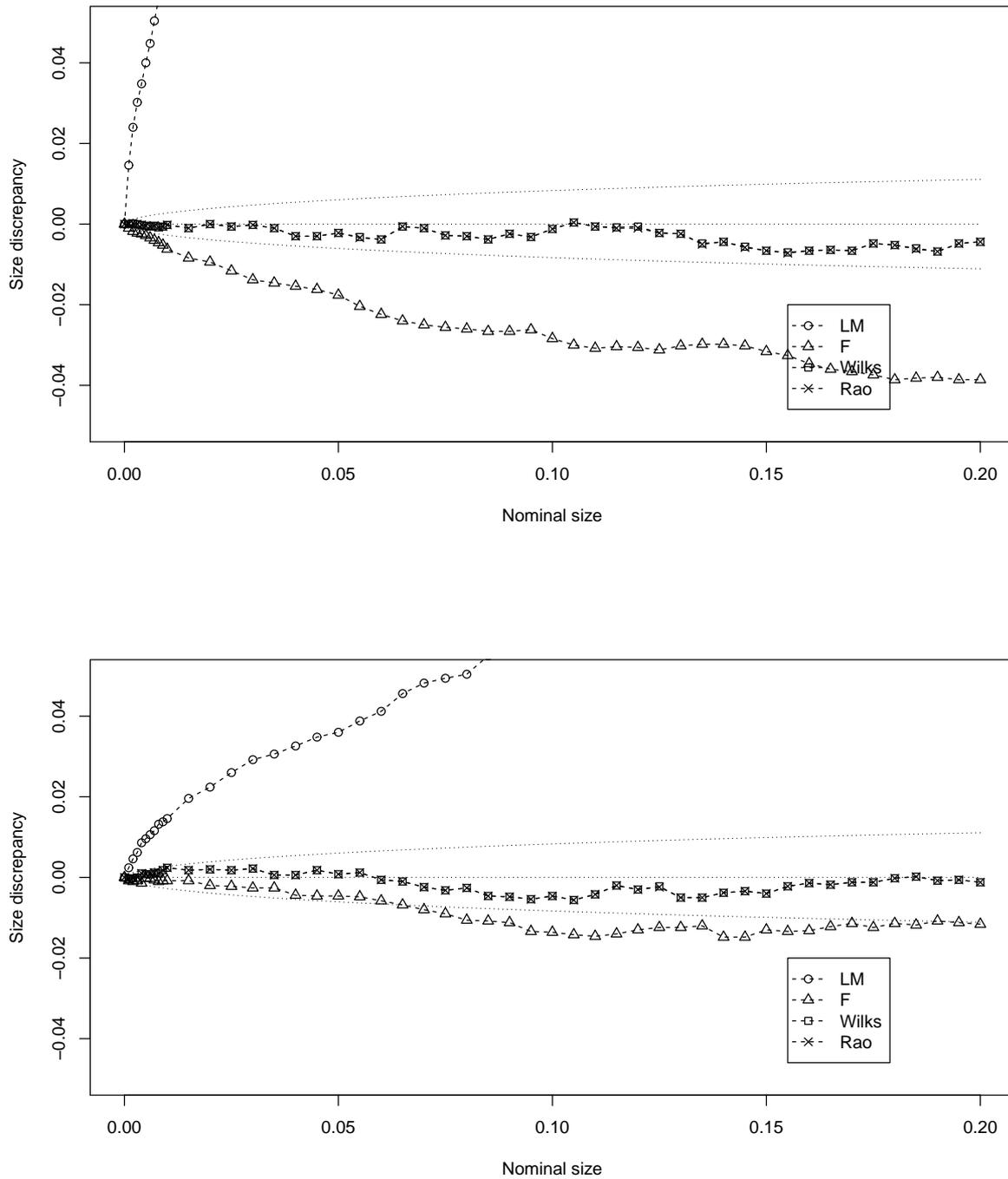
Figure 7: Size discrepancy plots for tests of parameter constancy: $p = 2$, $k = 2$ and $T = 200$ (Top); $T = 500$ (Bottom). The dotted lines represent the upper 95% confidence bound (top), zero line (middle) and the lower 95% confidence bound (bottom).
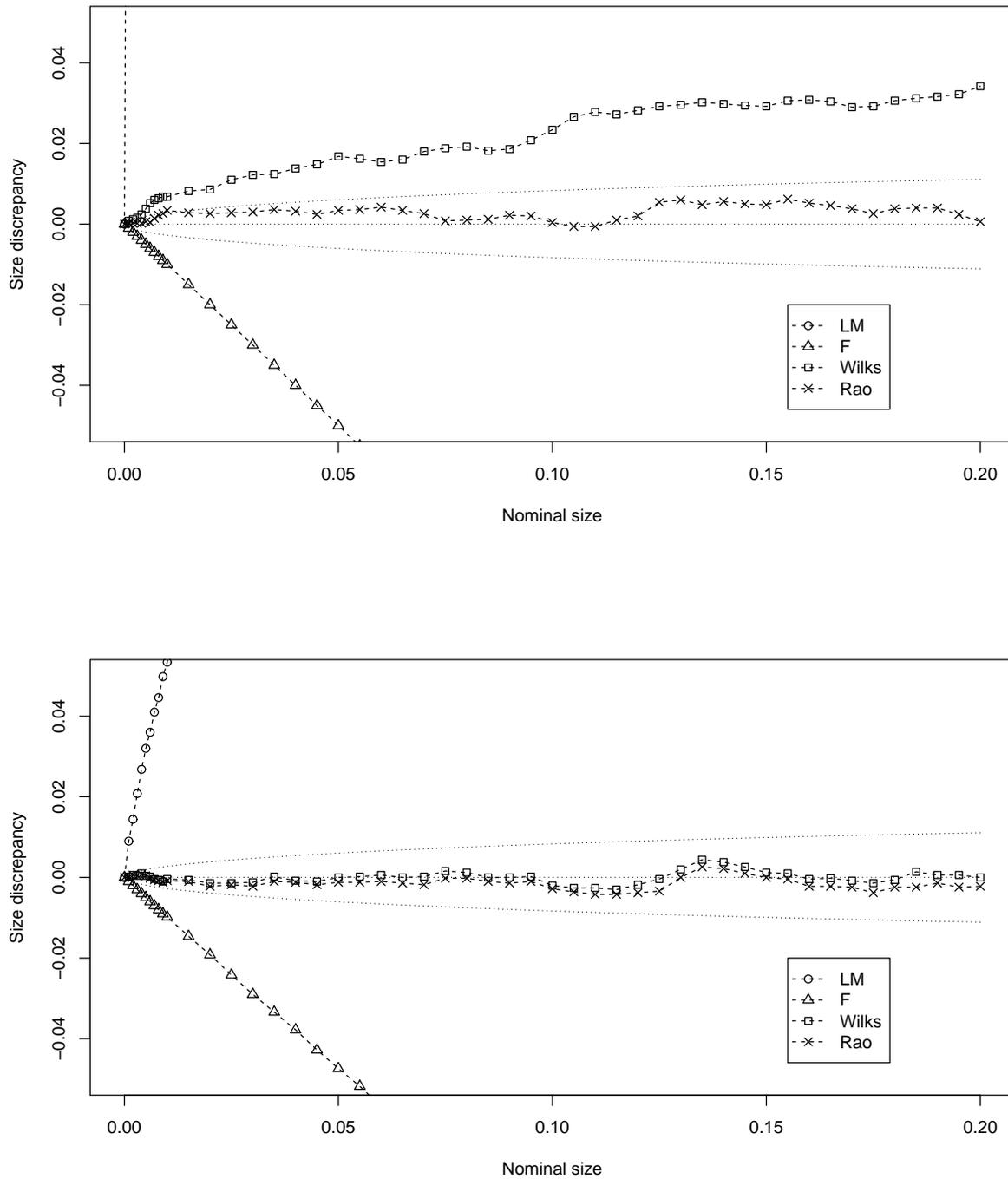
Research Papers
2013

2013-39:   Juan Carlos Parra-Alvarez: A comparison of numerical methods for the solution of continuous-time DSGE models

2013-40:   Daniel Ventosa-Santaulària and Carlos Vladimir Rodríguez-Caballero: Polynomial Regressions and Nonsense Inference

2013-41:   Diego Amaya, Peter Christoffersen, Kris Jacobs and Aurelio Vasquez: Does Realized Skewness Predict the Cross-Section of Equity Returns?

2013-42:   Torben G. Andersen and Oleg Bondarenko: Reflecting on the VPN Dispute

2013-43:   Torben G. Andersen and Oleg Bondarenko: Assessing Measures of Order Flow Toxicity via Perfect Trade Classification

2013-44:   Federico Carlini and Paolo Santucci de Magistris: On the identification of fractionally cointegrated VAR models with the F(d) condition

2013-45:   Peter Christoffersen, Du Du and Redouane Elkamhi: Rare Disasters and Credit Market Puzzles

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 CO_2 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