Rochester Center for Economic Research

Chi-square Tests for Parameter Stability

Marine Carrasco

Working Paper No. 508
September 2004

UNIVERSITY OF
ROCHESTER

Chi-square Tests For Parameter Stability*

Marine Carrasco †
University of Rochester

May 2004

Abstract

Testing when a nuisance parameter is identified only under the alternative is problematic because the Likelihood Ratio test converges to a nonstandard distribution that may depend on unknown parameters. Examples include testing parameter stability in Structural Change and Threshold models.

Our article proposes a class of tests that have the advantage of having a standard distribution, namely a chi-square. In this class, we focus mostly on a Lagrange Multiplier test in an auxiliary regression. We derive the asymptotic power of this test against alternatives which differ from the implicit alternative of the test. We show that this test can be used as a diagnostic test for parameter stabilility.

A Monte Carlo study compares the performance of our tests with other frequently used tests and shows that they have a similar power.

JEL classification: C12, C22.

Keywords: Admissibility, smooth transition, structural change test, Threshold autoregressive models.

---

*A former version of this paper has circulated under the name "Chi-square Tests when a Nuisance Parameter is Present only under the Alternative".

†I am particularly grateful to Werner Ploberger for his precious help and suggestions. I wish to thank, for helpful comments, Stephen Cosslett, Stéphane Grégoir and Jean-Pierre Florens, and also the participants of seminars at Crest, Institute for Advanced Studies (Vienna), Lille, LSE, Montreal, Penn State, Pittsburgh, Rochester, UCL, and USC universities and the participants of the Econometric Society Winter Meeting 1998, and ESEM99, and among them especially Donald Andrews. I thank Lucia Fedina and Liang Hu for excellent computing assistance. Of course, all errors are mine. Part of this research was supported by an Ohio State University Seed grant and a National Science Foundation grant. Address: University of Rochester, Department of Economics, Harkness Hall 226, Rochester, NY 14627. Email: csco@troi.cc.rochester.edu. Phone: (585) 275-7307. Fax: (585) 256-2309.

1. Introduction

The problem of testing when a nuisance parameter is present only under the alternative arises in many economic and financial models. Examples include testing for parameter stability in Structural Change, Threshold, and Markov-switching models. An extensive list of examples can be found in Hansen (1996). In this paper, the model of reference is

$$y_t = z_t'\theta + x_t'cg_t(\pi) + \varepsilon_t.$$

We want to test $H_0 : c = 0$. Under H_0, the distribution of the observations does not depend on some nuisance parameter π, whereas under the alternative, it depends on π. Since π is not identified under the null, the traditional Wald, Lagrange Multiplier (LM) and Likelihood Ratio (LR) tests have a nonstandard distribution.

The purpose of this paper is to construct new tests that have the advantage of having chi-square distributions. The idea of these tests is the following. Under the null hypothesis, the expectation of the score function is equal to zero for any value of the nuisance parameter. It results in a continuum of moment conditions. To handle this infinity of moment conditions, we apply the extension of the Generalized Method of Moments to a continuum of moment conditions developed by Carrasco and Florens (2000). The obtained estimator is asymptotically normal and therefore can be used to construct a Wald-type test. This test has a standard chi-square distribution. We show that a member of this class of tests is asymptotically equivalent to the LM test of $H_0 : c = 0$ in the regression

$$y_t = z_t'\theta + x_t'c \int g_t(\pi)\,\mu(d\pi) + \varepsilon_t$$

where μ is some finite measure. In some sense, we deal with the problem of an unidentified nuisance parameter by replacing the unknown term $g_t(\pi)$ with $\int g_t(\pi)\,\mu(d\pi)$, which is known. This LM test has power against a wide range of alternatives and therefore can be used as a diagnostic test for misspecification (Pagan and Hall, 1983).

In particular, we study the power properties of two tests. The first test is obtained from a structural change model, $g_t(\pi) = I\{t > T\pi\}$. In this case, $\int g_t(\pi)\,\mu(d\pi) = t/T$. This test has optimal power against a trending coefficient model and, in addition, it has power against a permanent structural change model. The second test originates from a Threshold model, $g_t(\pi) = I\{u_t > \pi\}$, for some observed random variable u_t. In this case, $\int g_t(\pi)\,F(d\pi) = F(u_t)$. This test has optimal power against a smooth transition alternative and is shown to have power against Threshold and other kinds of smooth transition models, as well as the Markov-switching model. These two tests are easy to implement and are robust to the heteroskedasticity and the autocorrelation of the error term.

The literature on testing in the presence of unidentified nuisance parameters can be split into two categories. The first category corresponds to tests designed to test a specific alternative. They are based on the likelihood and therefore have some optimality properties. These tests are the SupLM and ExpLM tests proposed by Andrews (1993) and Andrews and Ploberger (1994). One difficulty with these tests is that their asymptotic distributions are not standard and often depend on unknown parameters. In the latter case, critical values can not be tabulated. Hansen (1996) gives a method to compute the p-values via simulations in such cases. The second category includes tests that use an auxiliary model, which is supposed to approximate the true model. The auxiliary model is chosen because it is easy to estimate and the resulting test has a standard distribution. Tests of this type are the lack-of-fit test of Gallant (1977), tests based on expansions (Granger and Teräsvirta, 1993), and the RESET test of Ramsey (1969). Our tests fit in the second category.

The paper is organized as follows: Section 2 presents a general class of tests. Section 3 proposes a test that has power against permanent structural changes. Section 3 introduces a test that has power against Threshold and smooth transition alternatives. Section 5 discusses the results of a limited Monte Carlo experiment. The proofs are in the Appendix.

2. A class of tests when a nuisance parameter is not identified under the null

2.1. Model and null hypothesis

Consider the following regression

$$y_t = z_t'\theta + x_t'cg_t(\pi) + \varepsilon_t \tag{2.1}$$

where $g_t(\pi)$ is some scalar function of π and may be random. The null hypothesis of interest is $H_0 : c = 0$. Note that the nuisance parameter π is not identified under H_0, therefore it is impossible to estimate π consistently under H_0 and the usual Lagrange multiplier (LM) test fails to be asymptotically chi-square. Model (2.1) includes the following specifications:

Example 1. (Structural change model)
$g_t(\pi) = I\{t > T\pi\}$
Example 2. (Threshold model)
$g_t(\pi) = I\{u_t > \pi\}$.
Example 3. (Exponential smooth transition model)
$g_t(\pi) = \left(1 - \exp\left(-\gamma (u_t - d)^2\right)\right)$ where $\pi = (\gamma, d)'$.

See Granger and Teräsvirta (1993), for a review of inference and testing with these models. The observations are given by $\{y_t, z_t, t = 1, ...T\}$. The vector z_t can be decomposed as $z_t = (x_t', w_t')'$, where x_t and w_t may contain both lags of y_t and exogenous variables, for instance we may have $x_t = (y_{t-1}, y_{t-2}, ..., y_{t-m}, w_{1t}')'$ and $w_t = (y_{t-m-1}, y_{t-m-2}, ..., y_{t-l}, w_{2t}')'$ where w_{1t} and w_{2t} are exogenous variables. The number of lags of y_t entering in the regression, $l \geq 0$, is assumed to be known. For identification purposes, we assume that the regressors x_t and w_t do not have elements that are perfectly correlated with each other. $\pi \in \Pi \subset \mathbf{R}^l$, $c \in R^p$ and $\theta \in R^{p+q}$.

2.2. Test statistic

In this section, we propose a test statistic that asymptotically follows a chi-square distribution. Let $\gamma = (\theta', c')'$ and define $h_t(\pi, \gamma) = \left(h_{1t}(\pi, \gamma)', h_{2t}(\pi, \gamma)'\right)'$ by

$$\begin{aligned} h_{1t}(\pi, \gamma) &= z_t\left(y_t - z_t'\theta - x_t'cg_t(\pi)\right), \\ h_{2t}(\pi, \gamma) &= g_t(\pi)\, x_t\left(y_t - z_t'\theta - x_t'cg_t(\pi)\right). \end{aligned} \tag{2.2}$$

h_t is proportional to the score functions with respect to γ. For $\gamma_0 = (\theta_0', 0')$, we have a continuum of moment conditions indexed by π :

$$E^{\gamma_0} h_t(\pi, \gamma) = 0 \text{ for all } \pi \in \Pi \Rightarrow \gamma = \gamma_0, \tag{2.3}$$

where E^{γ_0} is the expectation with respect to the distribution indexed by γ_0. These moment conditions are satisfied for all π because π is not identified under H_0. The idea is to use the generalized method of moment estimator $\hat{\gamma} = \left(\hat{\theta}, \hat{c}\right)$ to construct a Wald-type test that will turn out to have a standard, nuisance-parameter-free distribution. Before presenting the test, we need to define a space of reference. Let μ be a finite measure on Π chosen a priori and $L^2(\Pi, \mu)$ be the Hilbert space of $(p+q)-$vectors of functions $f = \left(f_1, ..., f_{(p+q)}\right)'$ such that

$$\|f\|_{L^2}^2 = \sum_{j=1}^{(p+q)} \int_\Pi f_j(\pi)^2 \mu(d\pi) < \infty.$$

We assume that $\{h_t(\pi, \gamma)\}$ belongs to $L^2(\Pi, \mu)$ for all γ. Let B be a bounded operator from $L^2(\Pi, \mu)$ to $L^2(\Pi, \mu)$. A generalized method of moments (GMM) estimator of γ is

$$\hat{\gamma} = \arg\min_\gamma \|Bh_T(., \gamma)\|_{L^2}^2,$$

where $h_T(\pi, \gamma) = \frac{1}{T}\sum_{t=1}^T h_t(\pi, \gamma)$. The properties of GMM estimators based on a continuum of moment conditions are worked out in Carrasco and Florens (2000) for $\Pi = [0, 1]$ and iid data, in Carrasco and Florens (2002) for $\Pi = \mathbf{R}^l$ and iid data, and in Carrasco, Chernov, Florens, and Ghysels (2004) for $\Pi = \mathbf{R}^l$ and weakly dependent data. Under some regularity conditions the GMM estimator of c, $\hat{c}$, is consistent (to 0) and asymptotically normal under H_0. As a result, a Wald-type test,

$$W_T = T\hat{c}'\hat{V}_T^{-1}\hat{c},$$

where $\hat{V}_T$ is an estimator of the covariance matrix of $\hat{c}$, converges asymptotically to a chi-square distribution with p degrees of freedom. Note that W_T is not really a Wald test and should be called a Hausman test because $\hat{c}$ is not a consistent estimator of c under the alternative. Indeed the moment conditions (2.3) are not satisfied for all π under H_1 but only for $\pi = \pi_0$, the true value of the nuisance parameter. W_T constitutes a class of tests that have a standard distribution under H_0. However, one difficulty with W_T is that its power will depend on the choice of B and need to be verified on a case by case basis. It may happen that, for some B, W_T does not have power.

From now on, we assume that B is such that

$$Bh_t(.,\gamma) = \int_\Pi h_t(\pi,\gamma)\,\mu(d\pi).$$

For the moment conditions given by (2.2), we obtain

$$\begin{aligned}
&\|Bh_T(.,\gamma)\|^2_{L^2} \\
&= \left[\int_\Pi h_{1T}(\pi,\gamma)\,\mu(d\pi)\right]'\left[\int_\Pi h_{1T}(\pi,\gamma)\,\mu(d\pi)\right] + \left[\int_\Pi h_{2T}(\pi,\gamma)\,\mu(d\pi)\right]'\left[\int_\Pi h_{2T}(\pi,\gamma)\,\mu(d\pi)\right].
\end{aligned}$$

Minimizing $\|Bh_T(.,\gamma)\|^2_{L^2}$ with respect to γ is equivalent to finding the solution, $\hat\gamma$, of

$$\begin{aligned}
\int_\Pi h_{1T}(\pi,\hat\gamma)\,\mu(d\pi) &= 0, \\
\int_\Pi h_{2T}(\pi,\hat\gamma)\,\mu(d\pi) &= 0.
\end{aligned} \tag{2.4}$$

If moreover, $g_t(\pi)^2 = g_t(\pi)$ as in Examples 1 and 2, $\hat\gamma$ satisfies

$$\begin{aligned}
\sum_{t=1}^T z_t\left(y_t - z_t'\hat\theta - x_t'\hat c\int_\Pi g_t(\pi)\,\mu(d\pi)\right) &= 0, \\
\sum_{t=1}^T \left(\int_\Pi g_t(\pi)\,\mu(d\pi)\right) x_t\left(y_t - z_t'\hat\theta - x_t'\hat c\right) &= 0.
\end{aligned}$$

Hence $\hat\gamma$ is the OLS estimator of γ in the regression

$$y_t = z_t'\theta + x_t'c\int_\Pi g_t(\pi)\,\mu(d\pi) + \varepsilon_t \tag{2.5}$$

and W_T is the Wald test for testing $H_0 : c = 0$ in (2.5).

2.3. Power

In this subsection, we provide sufficient conditions for the test W_T to have power. Here we consider the general case where $g_t(\pi)^2$ may be different from $g_t(\pi)$. Define

$$\begin{aligned}
x_{1t} &= x_t\int_\Pi g_t(\pi)\,\mu(d\pi), \\
x_{2t} &= x_t\frac{\int_\Pi g_t^2(\pi)\,\mu(d\pi)}{\int_\Pi g_t(\pi)\,\mu(d\pi)},
\end{aligned}$$

From the system of equations (2.4), $\hat\theta$ and $\hat c$ are solutions of

$$\begin{aligned}
\sum_t z_t\left(y_t - z_t'\hat\theta - x_{1t}'\hat c\right) &= 0, \\
\sum_t x_{1t}\left(y_t - z_t'\hat\theta - x_{2t}'\hat c\right) &= 0.
\end{aligned} \tag{2.6}$$

We investigate the power of W_T under the local alternative:

$$y_t = z_t'\theta + x_{3t}'\frac{c}{\sqrt{T}} + \varepsilon_t, \tag{2.7}$$

where x_{3t} may be $x_t g_t(\pi)$ but not necessarily. Let $\Sigma_{xz} = p\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} x_t z_t$ and similarly define $\Sigma_{x_1 z}, \Sigma_{x_1 x_2}$ etc. and

$$M_1 = \begin{bmatrix} \Sigma_{zz} & \Sigma_{zx_1} \\ \Sigma_{x_1 z} & \Sigma_{x_1 x_2} \end{bmatrix} \text{ and } M_2 = \begin{bmatrix} \Sigma_{zz} & \Sigma_{zx_3} \\ \Sigma_{x_1 z} & \Sigma_{x_1 x_3} \end{bmatrix}.$$

Assuming that (i) $\{\varepsilon_t\}$ satisfies $E[\varepsilon_t|\mathcal{Z}_t] = 0$ and $E[\varepsilon_t^2|\mathcal{Z}_t] = \sigma^2$ where $\mathcal{Z}_t$ is the $\sigma-$field generated by $\{z_t, g_t(\pi), y_{t-s}, g_{t-s}(\pi), \varepsilon_{t-s}, z_{t-s}, s \geq 1\}$, (ii) $p\lim \frac{1}{T}\sum_{t=1}^{T} z_t z_t' \equiv \Sigma_{zz}$ is invertible, (iii) M_1 is invertible, and (iv) the law of large numbers and central limit theorem hold (a rigorous treatment will be given in Sections 3 and 4), we have

$$\sqrt{T}\hat{c} - \left[\Sigma_{x_1x_2} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_1}\right]^{-1}\left[\Sigma_{x_1x_3} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_3}\right]c \xrightarrow{L}$$
$$\mathcal{N}\left(0, \sigma^2\left[\Sigma_{x_1x_2} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_1}\right]^{-1}\left[\Sigma_{x_1x_1} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_1}\right]\left[\Sigma_{x_1x_2} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_1}\right]^{-1}\right).$$

Hence, W_T asymptotically follows a noncentral chi-square distribution with p degrees of freedom and noncentrality parameter:

$$\sigma^2 c'\left[\Sigma_{x_1x_3} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_3}\right]'\left[\Sigma_{x_1x_1} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_1}\right]^{-1}\left[\Sigma_{x_1x_3} - \Sigma_{x_1z}\Sigma_{zz}^{-1}\Sigma_{zx_3}\right]c.$$

Remarks:

- The test W_T has power against alternatives of type (2.7) provided $\det(M_1) \neq 0$ and $\det(M_2) \neq 0$. These assumptions seem reasonable and are satisfied in the examples 1 and 2 when $x_{3t} = x_{1t}$ holds for all t (see Propositions 3.1 and 4.1 below).

- The noncentrality parameter is maximized when $x_{3t} = x_{1t}$ holds for all t. This is the implicit alternative of the test W_T in the sense of Davidson and McKinnon (1987). For this alternative the test will have maximum power.

- The noncentrality parameter of W_T does not depend on x_{2t}, hence it is the same as that of a LM test of $H_0 : c = 0$ in Model (2.5). Both tests have the same asymptotic efficiency. As the LM test is easier to implement and has better finite sample properties than the Wald test (Dufour, 1997), it is preferable to use the LM test rather than W_T. This test will be denoted LM_T.

- LM_T can be used as a diagnostic test against a wide range of alternatives, in the tradition of the specification tests studied by Pagan and Hall (1983). Gallant (1977) tackles the same issue, namely testing $c = 0$ in Model (2.1). He suggests testing $c = 0$ in the auxiliary model

$$y_t = z_t'\theta + b_t'c + \varepsilon_t,$$

where

$$b_t = [x_t g_t(\pi_1), ..., x_t g_t(\pi_K)]$$

and $\pi_1, .., \pi_K$ are plausible values of π. He also suggests using the first principal components of the vector b_t as regressors (instead of b_t itself). In the case where $K = 1$, his testing strategy coincides with ours for a specific choice of $B : Bh(.) = h(\pi_1)$.

2.4. Admissibility

Here, and in the rest of the paper, we will focus on the properties of LM_T, the LM test of $H_0 : c = 0$ in Model (2.5). First recall some definitions (see Lehmann, 1959). Let ϕ be a test for testing $H_0 : \theta \in \Omega_0$ against $H_1 : \theta \in \Omega_1$. ϕ is said to be of level α if

$$\alpha \leq P[\phi \text{ rejects } H_0|\theta], \, \theta \in \Omega_0.$$

The power function of ϕ is denoted

$$\beta_\phi(\theta) \equiv P[\phi \text{ rejects } H_0|\theta].$$

A level-α test ϕ is unbiased if

$$\begin{aligned} \beta_\phi(\theta) &\leq \alpha, \, \theta \in \Omega_0, \\ \beta_\phi(\theta) &\geq \alpha, \, \theta \in \Omega_1. \end{aligned}$$

The level-α unbiased test that is uniformly most powerful, ψ is such that for any other level-α unbiased test ϕ,

$$\beta_\phi(\theta) \leq \beta_\psi(\theta), \, \theta \in \Omega_1.$$

The following proposition establishes the finite sample (T fixed) optimality of LM_T against the specific alternative (2.5). The optimality of LM_T follows from the Neyman-Pearson Lemma and the equivalence between LM and LR tests (Leymann, 1959).

Proposition 2.1. *Suppose $\varepsilon_t|\mathcal{Z}_t \sim iid\mathcal{N}(0, \sigma^2)$. The test LM_T is the uniformly most powerful unbiased test for testing $H_0 : c = 0$ against $H_1 : c \neq 0$ in (2.5).*

Proposition 2.1 implies that LM_T is admissible for testing $H_0 : c = 0$ against an unspecified alternative as it is the best test against a specific alternative (2.5). In the sequel, we will concentrate on Examples 1 and 2.

Example 1 (continued). In Section 3 below, we will study the LM test for structural change. In this case $\Pi = [0, 1]$. For μ uniform on Π, we have

$$\int_\Pi g_t(\pi) \mu(d\pi) = \int_0^1 I\{t > T\pi\} d\pi = \frac{t}{T}.$$

This LM test will be denoted SC_T.

Example 2 (continued). In Section 4 below, we study the LM test for threshold. Here $\Pi = \mathbf{R}$. For $\mu = F$, an arbitrary finite function, we have

$$\int_{\Pi} g_t(\pi)\,\mu(d\pi) = \int I\{u_t > \pi\}\,F(d\pi) = F(u_t)\,.$$

This test will be denoted T_T.

We will derive the properties of SC_T and T_T under alternatives, which differ from their implicit alternatives.

3. A diagnostic test for parameter stability

3.1. Asymptotic properties of the test

In this section we study the properties of the LM test of $H_0 : c = 0$ in

$$y_t = z_t'\theta + x_t' c \frac{t}{T} + \varepsilon_t. \tag{3.1}$$

Let $\tilde{\theta}$ be the OLS estimator of θ in the regression restricted under H_0 :

$$y_t = z_t'\theta + \varepsilon_t.$$

Assumption 3.1. *$\{\varepsilon_t\}$ satisfies $E[\varepsilon_t|\mathcal{Z}_t] = 0$ and $E[\varepsilon_t^2|\mathcal{Z}_t] = \sigma^2$ where $\mathcal{Z}_t$ is the $\sigma-$field generated by $\{z_t, y_{t-s}, \varepsilon_{t-s}, z_{t-s}, s \geq 1\}$.*

Assumption 3.2. *$\{z_t\}$ satisfies:*

(i) $p\lim \frac{1}{T}\sum_{t=1}^T z_t z_t' = \lim_{T\to\infty} E\left(\frac{1}{T}\sum_{t=1}^T z_t z_t'\right) \equiv \Sigma_{zz}$ where Σ_{zz} is some positive-definite, $(p+q)\times(p+q)$ matrix,

(ii) the matrix $\Sigma_{xx} \equiv p\lim \frac{1}{T}\sum_{t=1}^T x_t x_t'$ is a positive-definite $p \times p$ matrix,

(iii) $limsup_{T\to\infty}\frac{1}{T}\sum_{t=1}^T E\|z_t\|^{2+\nu} < \infty$ for some $\nu > 0$.

We make some remarks pertaining to the assumptions. Assumption 3.1 imposes that $\{\varepsilon_t\}$ is a homoskedastic martingale difference sequence. It will be relaxed in Assumption 3.4 below. Assumption 3.2 allows for random explanatory variables but rules out trending regressors. In the sequel, we use the notation

$$\begin{aligned}
\Sigma_{zz} &= \begin{bmatrix} \Sigma_{zx} & \Sigma_{zw} \end{bmatrix} = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xw} \\ \Sigma_{wx} & \Sigma_{ww} \end{bmatrix},\ \Sigma_{xz} = \Sigma_{zx}', \\
S_T &= \sum_{t=1}^T \frac{t}{T} x_t \left(y_t - z_t'\tilde{\theta}\right), \\
\Sigma_T &= \frac{1}{T}\begin{bmatrix} \sum z_t z_t' & \sum \frac{t}{T} z_t x_t' \\ \sum \frac{t}{T} x_t z_t' & \sum \left(\frac{t}{T}\right)^2 x_t x_t' \end{bmatrix}, \\
\Lambda &= (\mathbf{0}_{p\times q}, \mathbf{I}_p)'.
\end{aligned}$$

The test we will use in the homoskedastic case is

$$SC_T = \frac{1}{T\hat{\sigma}^2} S_T' \left(\Lambda' \Sigma_T^{-1} \Lambda\right) S_T.$$

where $\hat{\sigma}^2 = \sum \left(y_t - z_t'\tilde{\theta}\right)^2 / T$. Remark that S_T can be rewritten as

$$S_T = \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{t} x_s \left(y_s - z_s'\tilde{\theta}\right)$$

To see this, denote $A_t = (t+1)/T$ and $B_t - B_{t-1} = x_t \left(y_t - z_t'\tilde{\theta}\right)$. Using the relationship $\sum A_{t-1} (B_t - B_{t-1}) + \sum B_t (A_t - A_{t-1}) = B_T A_T - A_0 B_0$, we have

$$\begin{aligned} \sum_{t=1}^{T} \frac{t}{T} x_t \left(y_t - z_t'\tilde{\theta}\right) &= -\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{t} x_s \left(y_s - z_s'\tilde{\theta}\right) + \frac{(T+1)}{T} \sum_{t=1}^{T} x_t \left(y_t - z_t'\tilde{\theta}\right) \\ &= -\frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{t} x_s \left(y_s - z_s'\tilde{\theta}\right), \end{aligned}$$

because $\{x_t\}$ belongs to the regressors $\{z_t\}$ and $\tilde{\theta}$ is the constrained estimator obtained from $\sum_{t=1}^{T} z_t \left(y_t - z_t'\tilde{\theta}\right) = 0$.

Failure to account for serial correlation or conditional heteroskedasticity may result in wrong conclusions about the parameter stability. This problem is well documented in Tang and MacNeill (1993). Therefore, we also derive SCh_T a misspecification-robust version of the LM test proposed by White (1980) in the iid case and Newey and West (1987a) in the dynamic case. First define Ω_T, the kernel estimator of a long-run covariance matrix:

$$\Omega_T = \sum_{j=-T}^{T} \omega\left(\frac{j}{L_T}\right) \hat{\Gamma}_j, \tag{3.2}$$

where L_T is a bandwidth, $\hat{\Gamma}_j = \sum_{t=j+1}^{T} h_t\left(\tilde{\theta}\right) h_{t-j}'\left(\tilde{\theta}\right)/T$, for $j \geq 0$, $\hat{\Gamma}_j = \hat{\Gamma}_{-j}'$ for $j < 0$ and

$$h_t(\theta) = \begin{bmatrix} z_t (y_t - z_t'\theta) \\ \frac{t}{T} x_t (y_t - z_t'\theta) \end{bmatrix}.$$

The term $\omega(x)$ is a kernel, it may be equal to $(1 + |x|) I \{|x| \leq 1\}$ if one adopts the Newey and West estimator (1987b). Other kernels are studied by Andrews (1991). A heteroskedasticity-robust version of the test is

$$SCh_T = \frac{1}{T} S_T' \left(\Lambda' \Omega_T^{-1} \Lambda\right) S_T.$$

Assumption 3.3. *The kernel weight, ω, is either the Bartlett, Parzen, Tukey-Hanning or Quadratic spectral kernel studied by Andrews (1991). Let ν be the parameter that characterizes the smoothness of $\omega : \nu = 1$ for the Bartlett kernel and $\nu = 2$ for the three other kernels. The bandwidth L_T satisfies $L_T \to \infty$ and $L_T^{1+2\nu}/T = O(1)$.*

Assumption 3.4. *(a)* $\{z_t\varepsilon_t\}$ *is strict stationary,* $\alpha-$ *mixing with mixing coefficient* α_m *and satisfies* $E(z_t\varepsilon_t)=0$*. For some* $r\in(2,4]$, $r>2+1/\nu$, *and some* $s>r$,

$$\begin{aligned}\sum_{m=1}^{\infty}\alpha_m^{2(1/r-1/s)} &< \infty, \\ \|z_t\varepsilon_t\|_s &< \infty,\end{aligned} \tag{3.3}$$

where $\|z_t\varepsilon_t\|_s=\left(\sum_j E|z_{tj}\varepsilon_t|^s\right)^{1/s}$ *and* z_{tj} *denotes the* j*th element of* $z_t=\left(z_{t1},...,z_{t(p+q)}\right)'$.

(b) $E\left[\|z_tz_t'\|^2\right]<\infty$ *and* $E\left[\|z_ty_t\|^2\right]<\infty$.

Assumption 3.4 allows for correlation and conditional heteroskedasticity of $\{\varepsilon_t\}$. Assumption 3.4(a) is condition (V1) of Hansen (1992) plus strict stationarity. It guarantees that the covariance matrix estimator, Ω_T, is consistent. It also implies that $\{z_t\varepsilon_t\}$ satisfies a functional central limit theorem. By Doukhan (1994, page 47), we have $\frac{1}{\sqrt{T}}\sum_{t=1}^{[T\pi]}z_t\varepsilon_t$ converges to a Gaussian process with covariance

$$k(\pi_1,\pi_2)=(\pi_1\wedge\pi_2)\sum_{j=-\infty}^{\infty}E\left[z_tz_{t-j}'\varepsilon_t\varepsilon_{t-j}\right]\equiv(\pi_1\wedge\pi_2)\,C_{zz},$$

where $\pi_1\wedge\pi_2$ denotes the minimum of π_1 and π_2. We will use the notation

$$C_{zz}=\begin{bmatrix}C_{zx} & C_{zw}\end{bmatrix}=\begin{bmatrix}C_{xx} & C_{xw}\\ C_{wx} & C_{ww}\end{bmatrix}\text{ and } C_{xz}=C_{zx}'.$$

We now investigate the power of the tests SC_T and SCh_T against a local alternative of the form

$$H_{1T}:y_t=z_t'\theta+x_t'\frac{1}{\sqrt{T}}g\left(\frac{t}{T}\right)+\varepsilon_t,$$

where g is an arbitrary, $p\times1-$dimensional function defined on the $(0,1)-$ interval. Such a specification includes, as a special case, the single structural change when $g(v)=cI(v\geq\pi_0)$ and the multiple structural change when $g(v)=\sum_{i=1}^{m}c_iI(\pi_{i+1}>v\geq\pi_i)$ and $\pi_{m+1}=\infty$.

Proposition 3.1. *(a) If Assumptions 3.1 and 3.2 hold, we have under* H_{1T} :

$$SC_T\xrightarrow{d}\chi^2\left(p,\varphi'V^{-1}\varphi\right),$$

where

$$\begin{aligned}V &= \frac{\sigma^2\Sigma_{xx}}{12},\\ \varphi &= \int_0^1\left(\int_0^{\pi}\Sigma_{xx}g(v)\,dv\right)d\pi-\frac{1}{2}\int_0^1\Sigma_{xx}g(v)\,dv.\end{aligned}$$

(b) If Assumptions 3.2, 3.3, and 3.4 hold, we have under H_{1T} :

$$SCh_T\xrightarrow{d}\chi^2\left(p,\varphi'V_h^{-1}\varphi\right),$$

where

$$V_h = \frac{1}{12} C_{xx}.$$

Under H_0, both tests asymptotically have a standard distribution, which is a chi-square with p degrees of freedom. Remark that the power of SC_T and SCh_T does not depend on the regressors, w_t. Consider the case where there is a common, one-time break point so that $g(v) = cI(v > \pi_0)$. π_0 is the location of the break, while c is its amplitude. In this case, we have

$$\varphi = \frac{\pi_0 (1 - \pi_0)}{2} \Sigma_{xx} c.$$

It means that the test SC_T always has power against a one-time structural change and achieves its maximal power when the break occurs in the middle of the sample, $\pi_0 = 1/2$. However, our SC_T test will only have trivial power against alternatives with $\varphi = 0$. We construct an example where $\varphi = 0$. Let $g(v) = cg_1(v)$ where c is a $p-$dimensional vector and $g_1(v)$ is a scalar function,

$$g_1(v) = \begin{cases} -1, \ 0 < v \leq \pi_1, \\ 1, \ \pi_1 < v \leq \pi_2, \\ -1, \ \pi_2 < v \leq 1. \end{cases}$$

If $\pi_2 = 2\pi_1 = 2/\sqrt{3}$, $\varphi = 0$ and the limiting rejection probability of SC_T is the same as under H_0. However, the SC_T test will, in general, have power against multiple-break alternatives unless these breaks compensate each other, rendering $\varphi = 0$. This lack of power against some multiple-change alternatives is not specific to our test. The CUSUM test (Krämer et al. 1988), supLR (Andrews, 1993) and ExpLR tests (Andrews and Ploberger, 1994) also lack power against some alternatives. All these tests, including SC_T, have no power against alternatives where the process is stationary, this includes the case where y_t follows a Threshold autoregressive model (Tong, 1990), see e.g. Carrasco (2002).

So far, we chose to leave the specification of the alternative hypothesis in a rather vague form. The reason is that, in practice, we rarely know whether g_t experiences only one change, multiple changes, or whether g_t is a random coefficient. Of course, if the alternative was completely specified, there might be tests that are more powerful than ours.

3.2. Related Literature

In this subsection, we review the most popular tests for structural change. Brown, Durbin, and Evans (1975) propose several tests of parameter stability, one is based on the cusum of recursive residuals (the now famous CUSUM test) and another is based on the squares of these residuals. The power properties of these tests have been investigated by Krämer, Ploberger, and Alt (1988) and Ploberger and Krämer (1992). The CUSUM tests are not robust to the autocorrelation of the error in the regression model. As pointed out by Pagan and Hall (1983), a way to detect parameter inconstancy is to test for heteroskedasticity. Szroeter (1978) proposes a test for heteroscedasticity that is similar in spirit to our test, indeed, he devises a test for an alternative of the type $\sigma_t^2 = t$. His test can not be used for testing parameter stability if there is heteroscedasticity in the model that is not due to the parameter inconstancy. Our test is more closely related to the AvgLM test of Andrews and Ploberger (1994). Consider Model (2.1) in the case of a one-time structural change, that is $g_t(\pi) = I\{t > T\pi\}$. For π given, the Lagrange Multiplier test for testing $H_0 : y_t = z_t'\theta + \varepsilon_t$ against $H_1 : y_t = z_t'\theta + x_t'cI\{t > T\pi\} + \varepsilon_t$ is

$$\widetilde{LM}_T(\pi) = \left[\sum_{s=[T\pi]+1}^{T} x_s\left(y_s - z_s'\tilde{\theta}\right)\right]'\left(\Lambda'\tilde{I}_T(\pi)^{-1}\Lambda\right)\left[\sum_{s=[T\pi]+1}^{T} x_s\left(y_s - z_s'\tilde{\theta}\right)\right],$$

where $\tilde{I}_T(\pi)$ is the estimator of the covariance matrix of the moment conditions,

$$h_T(\theta_0) = \begin{bmatrix} \sum_{t=1}^{T} z_t(y_t - z_t'\theta_0) \\ \sum_{t=[T\pi]+1}^{T} x_t(y_t - z_t'\theta_0) \end{bmatrix},$$

evaluated at $\theta = \tilde{\theta}$. To handle the case where π is unknown, Davies (1977) proposes to use

$$\text{SupLM} = \sup_{\pi\in\Pi} \widetilde{LM}_T(\pi),$$

where Π is a set with closure in (0,1). The properties of the supLM are studied in Andrews (1993). Andrews and Ploberger (1994) propose a class of admissible tests that take the form

$$\text{ExpLM} = (1+a)^{-p/2}\int_{\Pi}\exp\left(\frac{1}{2}\frac{a}{1+a}\widetilde{LM}_T(\pi)\right)dJ(\pi),$$

for some constant $a > 0$ and some distribution J on Π. These tests are shown to be optimal for testing $H_0 : c = 0$ against a Bayesian alternative, in which the location of the break, π, is distributed according to J and its amplitude, c, follows a normal distribution. When $a \to 0$, the test becomes

$$\text{AvgLM} = \int_{\Pi}\widetilde{LM}_T(\pi)\,dJ(\pi).$$

Nyblom (1989)'s test is closely related to AvgLM. The main difference between our test and AvgLM lies in the treatment of π. In AvgLM, π is integrated out at the very end, whereas, in SC_T, π is integrated out when calculating S_T. Indeed S_T can be rewritten as

$$S_T = \int_0^1 \sum_{s=[T\pi]+1}^{T} x_s\left(y_s - z_s'\tilde{\theta}\right)d\pi.$$

As a result SC_T has a standard distribution and is nuisance parameter free. All the tests discussed so far have non-standard distributions. The critical values of SupLM and ExpLM have been tabulated by Andrews (1993) and Andrews and Ploberger (1994) for the case where there are no exogenous variables in the model. In the presence of exogenous variables, the distributions of the tests will depend on them and have to be tabulated on a case by case basis. However, the distributions do not depend on nuisance parameters, see Andrews, Lee, and Ploberger (1996) and Forchini (2002). As the number of admissible tests available in the literature is large, one criterion for choosing among admissible tests should be the ease of application. Our test is as easy to implement as the ExpLM test and has the extra advantage of having a standard distribution. Remark that the Sup and Exp tests require selecting the interval Π where π is supposed to lie. Our test does not have this requirement, no trimming of the interval (0,1) is needed. With the same goal to provide an easy to apply test, Altissimo and Corradi (2002) propose a LIL test that has the advantage of having a known critical value, moreover, it can handle misspecification and heterogeneity. Their approach differs from ours because their test is completely consistent, that is, its level goes to zero and its power goes to one.

4. A diagnostic test for threshold alternatives

4.1. Properties of the test

In Section 3, we proposed a test that is specifically designed for alternatives, in which changes in g_t are a deterministic function of the time, t. However, changes in g_t may be triggered by the values taken by some observable variable u_t. This is the case in the threshold model (Tong, 1990) and the smooth transition regression (Teräsvirta, 1994). In this section, we study the LM test of $H_0 : c = 0$ in the following model

$$y_t = z_t'\theta + x_t'cF(u_t) + \varepsilon_t, \tag{4.1}$$

where F is a function chosen a priori, for instance, it may be the c.d.f. of the standard normal distribution. The switching variable, u_t, may be a lagged value of y_t or an exogenous variable, including a subset of the regressors z_t. The observations are given by $\{y_t, z_t, u_t, v_t, t = 1, ...T\}$. The variable v_t and the function g appear only under the local alternative (4.3).

Assumption 4.1. *$\{\varepsilon_t\}$ is strictly stationary, ergodic and satisfies $E[\varepsilon_t|\mathcal{F}_t] = 0$ and $E[\varepsilon_t^2|\mathcal{F}_t] = \sigma^2$ where $\mathcal{F}_t$ is the $\sigma-$field generated by $\{z_t, u_t, \varepsilon_{t-s}, z_{t-s}, u_{t-s}, s \geq 1\}$.*

Assumption 4.2. *$\{z_t, u_t, v_t\}$ is strictly stationary, ergodic and (i) there is a constant $C < \infty$ so that $|F(u_t)| < C$ and $\|g(v_t)\| < \infty$, (ii) $Ez_tz_t' < \infty$ is positive definite.*

Let

$$S_T = \sum_{t=1}^{T} F(u_t)\, x_t \left(y_t - z_t'\tilde{\theta}\right)$$

where $\tilde{\theta} = \left(\sum z_t z_t'\right)^{-1} \sum z_t y_t$.

Our test (denoted as T_T for "Threshold") is

$$T_T = \frac{1}{T} S_T' V_T^{-1} S_T,$$

with

$$V_T = \hat{\sigma}^2 \left\{ \hat{\Sigma}_{xxF2} - \hat{\Sigma}_{xzF} \hat{\Sigma}_{zz}^{-1} \hat{\Sigma}_{zxF} \right\},$$

where $\hat{\Sigma}_{zz}$ denotes the sample estimates of Σ_{zz} and $\hat{\Sigma}_{xzF}$ and $\hat{\Sigma}_{xxF2}$ are the sample estimates of $\Sigma_{xzF} = E\left[x_t z_t' F(u_t)\right]$ and $\Sigma_{xxF2} = E\left[x_t x_t' F^2(u_t)\right]$, respectively. Moreover $\hat{\Sigma}_{zxF} = \hat{\Sigma}_{xzF}'$. To construct a test robust to the heteroskedasticity and the serial correlation of the errors, we use a kernel estimator, Ω_T, as defined in (3.2) with

$$h_t(\theta) = \begin{bmatrix} z_t \left(y_t - z_t'\theta\right) \\ F(u_t)\, x_t \left(y_t - z_t'\theta\right) \end{bmatrix}.$$

A heteroskedasticity-robust version of the test is

$$Th_T = \frac{1}{T} S_T' \left(\Lambda' \Omega_T^{-1} \Lambda\right) S_T.$$

Assumption 4.3. *(a)* $\{z_t, u_t, \varepsilon_t\}$ *is strict stationary,* $\alpha - mixing$ *with mixing coefficient* α_m *and* $E(z_t\varepsilon_t) = 0, E(F(u_t)\, x_t\varepsilon_t) = 0$. *For some* $r \in (2, 4]$, $r > 2 + 1/\nu$, *and some* $s > r$,

$$\begin{aligned} \sum_{m=1}^{\infty} \alpha_m^{2(1/r-1/s)} &< \infty, \\ \left\| z_t\varepsilon_t \right\|_s &< \infty, \end{aligned} \tag{4.2}$$

where $\left\| z_t\varepsilon_t \right\|_s = \left(\sum_j E\left| z_{tj}\varepsilon_t \right|^s \right)^{1/s}$ *and* z_{tj} *denotes the* j*th element of* $z_t = \left(z_{t1}, ..., z_{t(p+q)}\right)'$.

(b) $E\left[\left\| z_t z_t' \right\|^2\right] < \infty$ *and* $E\left[\left\| z_t y_t \right\|^2\right] < \infty$.

Under Assumption 4.3, we have

$$\begin{pmatrix} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} F(u_t)\, x_t\varepsilon_t \\ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_t\varepsilon_t \end{pmatrix} \xrightarrow{L} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_{xxF2} & C_{Fxz} \\ C_{zxF} & C_{zz} \end{pmatrix} \right),$$

where C_{xxF2}, and C_{zz} are the long-run covariances of $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} F(u_t)\, x_t\varepsilon_t$ and $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_t\varepsilon_t$, $C_{Fxz} = \lim_{T\to\infty} \sum_{j=-T}^{T} E\left(F(u_t)\,\varepsilon_t\varepsilon_{t-j} x_t z_{t-j}'\right)$, and $C_{zxF} = C_{Fxz}'$. We examine the power of T_T and Th_T against local alternatives of the form

$$H_{1T} : y_t = z_t'\theta + x_t' \frac{g(v_t)}{\sqrt{T}} + \varepsilon_t, \tag{4.3}$$

where g is a $p \times 1$ vector function of v_t, and v_t is a stationary random variable that may or may not coincide with u_t. This alternative includes (i) the threshold regression, $g(v_t) = cI\{v_t \geq r_0\}$, for some unknown threshold r_0 (ii) the smooth transition regression (STR), $g(v_t) = cG(v_t)$, where G is some cdf (iii) the exponential STR, $g(v_t) = c\left(1 - \exp\left(-\gamma (v_t - d)^2\right)\right)$, (iv) the logistic STR, $g(v_t) = c\left(1 + \exp\left(-\gamma (v_t - d)\right)\right)^{-1}$, (v) the Markov switching model (Hamilton, 1989) where v_t is an exogenous two-state Markov chain.

Proposition 4.1. *Under Assumptions 4.1, 4.2, and H_{1T},*

$$T_T \xrightarrow{L} \chi^2\left(p, \varphi' V^{-1} \varphi\right)$$

with $V = \sigma^2\left(\Sigma_{xxF2} - \Sigma_{xzF}\Sigma_{zz}^{-1}\Sigma_{zxF}\right)$ *and*

$$\varphi = E\left[F\left(u_t\right) x_t x_t' g\left(v_t\right)\right] - E\left[F\left(u_t\right) x_t z_t'\right] \Sigma_{zz}^{-1} E\left[z_t x_t' g\left(v_t\right)\right].$$

Under Assumptions 4.2, 4.3, 3.3, and H_{1T},

$$Th_T \xrightarrow{L} \chi^2\left(p, \varphi' V_h^{-1} \varphi\right)$$

with $V_h = \sigma^2\left(C_{xxF2} - \Sigma_{xzF}\Sigma_{zz}^{-1}C_{zz}\Sigma_{zz}^{-1}\Sigma_{zxF} - C_{Fxz}\Sigma_{zz}^{-1}\Sigma_{zxF} - \Sigma_{Fxz}\Sigma_{zz}^{-1}C_{zxF}\right).$

Note that the power of T_T depends on the variables ω_t, if u_t is correlated with x_t, and w_t. We can establish the following results:

- If F and g are independent from each other, then $\varphi = 0$ and the test does not have power.
- If F is independent of z, then $\varphi = cov\left(F, xx'g\right)$.
- If g is independent of z, then $\varphi = cov\left(Fxx', g\right)$.
- In particular, if F and g are independent of z and $g\left(v_t\right) = cG\left(v_t\right)$, where G is a scalar function, then $\varphi = cov\left(F\left(u_t\right), G\left(v_t\right)\right)\Sigma_{xx}c$. Hence, if F and G are correlated, as in examples (i) to (iv) (with $v_t = u_t$), T_T has always power. Maximal power is achieved when $F = G$.
- Consider the case where v_t is an exogenous Markov chain. Let $g\left(v_t\right) = v_t$, $u_t = y_{t-1}$ and $x_t = y_{t-1}$. In this case, $\varphi \neq 0$ because v_t is correlated with v_{t-1} and hence y_{t-1}. Therefore, T_T will have power against a Markov-switching alternative. Note that if v_t were an independent sequence, as it is in the independent mixture model, T_T would have no power.

T_T coincides with the Lagrange Multiplier test for testing an alternative STR of type (ii) with $g = cF$ and, therefore, it is optimal for this alternative. According to Godfrey (1988, Chapter 3), T_T is also optimal against a class of alternatives that are asymptotically equivalent to $H_1 : c \neq 0$ in (4.1). Members of this class of alternatives are $H_2 : y_t = z_t'\theta + x_t'g\left(u_t, c\right) + \varepsilon_t$ that satisfy

$$\begin{cases} F(u_t) = \left.\frac{\partial g(u_t,c)}{\partial c}\right|_{c=0}, \\ g\left(u_t, 0\right) = 0. \end{cases}$$

For instance, $g\left(u_t, c\right) = \left(1 - \exp\left(-cF\left(u_t\right)\right)\right)$ is part of this class. The sample value of the LM statistic will be the same whether the alternative is H_1 or H_2, even though, they appear dissimilar. The LR and Wald statistics for locally equivalent alternatives are only asymptotically equivalent.

4.2. Advantages over other Threshold tests

Tsay (1989) proposes a test for threshold nonlinearity that asymptotically follows a chi-square distribution. Tsay's test is based on arranged autoregression and predictive residuals. It does not apply to the case of detecting a shift in the intercept in the absence of regressors. Moreover, it is not robust to the serial correlation of the errors. Luukkonen, Saikkonen, and Teräsvirta (1988) propose to use a LM test based on an augmented regression, see Granger and Teräsvirta (1993). One way to compare the performances of these various tests against a specific alternative would be to compute the asymptotic relative efficiency (see Davidson and McKinnon, 1987).

The test that is the closest to ours is the AvgLM test. Let $\widetilde{LM}_T(r)$ be the Lagrange Multiplier test to test $H_0 : c = 0$ against $H_1 : y_t = z_t'\theta + x_t'cI\{u_t > r\} + \varepsilon_t$ for a given value of r. When r_0 is unknown, Chan (1990) proposes to use, as a test, $\sup_{r\in R}\widetilde{LM}_T(r)$ where R is a bounded interval within which r_0 is supposed to lie. Admissible versions of this test are proposed by Andrews and Ploberger (1994). As for the structural change model, one can define ExpLM and AvgLM tests by integrating over the nuisance parameter r. The main problem with these tests is that the critical values depend on unknown parameters and can not be tabulated. Hansen (1996) gives a method to compute the p-values in such cases. On the contrary, our test has a standard distribution.

5. Monte Carlo experiment

5.1. Power of SC_T and SCh_T tests

We investigate the power performance of the tests SC_T and SCh_T in two cases. First, in the case where the data generating process is a one-time structural change model:

$$y_t = 0.1y_{t-1} + cy_{t-1}I\{t > T\pi\} + \varepsilon_t, \tag{5.1}$$

where ε_t is generated from a GARCH$(1,1)$ model with

$$\begin{aligned} \varepsilon_t &= \sqrt{h_t}\eta_t, \\ h_t &= \frac{1}{2} + \frac{1}{2}\varepsilon_{t-1}^2 + \frac{1}{4}h_{t-1}, \end{aligned} \tag{5.2}$$

and $\{\eta_t\}$ are i.i.d. standard normal. We set the true change point $\pi_0 = 0.5$ and $c = 0.5$, 0.8 under the alternative.

In the second case, the data generating process is the explicit alternative of the tests SC_T and SCh_T, namely,

$$y_t = 0.1y_{t-1} + cy_{t-1}\frac{t}{T} + \varepsilon_t, \tag{5.3}$$

where ε_t is as above and c is chosen to be 0.5 and 0.8 under the alternative.

Simulations are programmed in GAUSS 3.2. First, data are generated according to the Data Generating Process above. Since we want to start from a stationary process, we generate 200 extra data and then discard the first 200 data. For comparison, we also perform the heteroskedasticity-robust versions of Andrews and Ploberger's SupLM, AvgLM and ExpLM tests. The interval $[\underline{\pi}, \overline{\pi}]$ for computing these three tests has been selected to be $[0.15, 0.85]$ as suggested by Andrews (1993). The data are generated using the same seed for all tests. The seed used to compute the empirical critical values corresponds to rndseed 39700802 and the seed used to compute the power corresponds to rndseed 39700803. We compute both the empirical power and the size-corrected power. The critical values used for the empirical power are the values tabulated by Andrews (1993) for the SupLM test and Andrews and Ploberger(1994) for the AvgLM and ExpLM tests. We use the cut-off point of the chi-square with 1 degree of freedom for SC_T and SCh_T. Note that SC_T does not converge to a $\chi^2(1)$ due to the presence of heteroskedasticity. We still report the empirical power of SC_T as an indication of what one would get if the heteroskedasticity is ignored.

To calculate the long-run covariance matrix,

$$\Omega_T = \sum_{j=-T}^{T} \omega(\frac{j}{L_T})\hat{\Gamma}_j,$$

for the SCh_T test statistic, we followed the methods suggested in Andrews (1991) and Andrews and Monahan (1992) for Heteroskedasticity and Autocorrelation Consistent (HAC) covariance matrix estimators. We used the Quadratic Spectral Kernel, a first-order VAR prewhitening procedure, and an automatic bandwidth selection. To adjust for singularity, we picked the cutoff point to be 0.95 (see the footnote on page 957 of Andrews and Monahan (1992) for an analysis of it). In fact, we tried different cutoff points near 0.90 and it turned out that our results were insensitive to it. For the automatic selection of the bandwidth parameter, we use a first order autoregressive model to approximate the parametric models here. The specified weights are all chosen to be 1's.

For both the empirical and size-corrected values tabulated, we adopted the same number of iterations and the same sample size. For the samples of size $T = 60$ and $T = 100$, we used 2000 iterations. For the samples of size $T = 200$, 500, and 1000, we used 1000 iterations.

Discussion of the results:

Table 1 displays the empirical and size-corrected powers for the five tests when the DGP is (5.1). We can see that:

i) The size distortion is smaller for SC_T and SCh_T than for Sup, Avg and Exp LM tests. Diebold and Chen (1996) previously documented that the homoskedastic versions of the Sup tests exhibit important size distortions in small samples when the persistence in the data is high.

ii) SC_T has the highest empirical power for all sample sizes. Its size-corrected power is the highest in small samples but it is dominated by that of AvgLM and ExpLM for $T = 200$.

iii) The empirical power of SCh_T is bigger than that of Sup, Avg and Exp LM tests for small samples, the reverse is true for large samples. The size-corrected power of SCh_T is comparable to that of other tests when $T = 60$, but it is dominated by that of AvgLM and ExpLM for larger sample sizes. This makes sense since the AvgLM and ExpLM are the optimal tests for structural change alternatives.

iv) Since our data is generated by a GARCH model, it is surprising to find that the SC_T test performs better than the robust SCh_T test in some cases. The reason may be that the correlation and the heteroskedasticity are not big enough (we choose the coefficients to be $\frac{1}{2}$ and $\frac{1}{4}$). But we can see that as the sample size goes to 500, or even larger, 1000, the SCh_T test catches up in terms of power.

Table 2 includes the empirical and size-corrected powers for the five tests when the DGP is (5.3). We make the following observations:

i) SCh_T performs better than Sup, Avg and Exp LM tests in terms of both the empirical power and the size-corrected power. This is consistent with our theory, as the test SCh_T is optimal for the alternative (5.3).

ii) SC_T has the highest empirical power for all T. Its size-corrected power is the highest in small samples, but then is dominated by SCh_T and $AvgLM$.

iii) The size-corrected power of SCh_T is lower than that of SC_T in small samples but the reverse is true for samples of size $T = 200$ and larger.

5.2. Power of T_T test

Assume that the DGP is a TAR model

$$y_t = 0.1y_{t-1} + cy_{t-1}I\left(y_{t-1} > r\right) + \varepsilon_t, \tag{5.4}$$

where ε_t follows a GARCH(1,1) with parameters described in (5.2) and the true value of r is set to $r = 0.05452$ (the median of y_t under H_0). Consider the tests T_T and Th_T, which are the LM test of $H_0 : c = 0$ against the alternative

$$y_t = \alpha y_{t-1} + cy_{t-1}\Phi\left(\frac{y_{t-1}}{\hat{\sigma}}\right) + \varepsilon_t,$$

where $\hat{\sigma}$ is the sample standard error of y_t. Note that the presence of $\hat{\sigma}$ does not alter the asymptotic distribution of the test. We compare the power of T_T and Th_T with that of the heteroskedasticity-robust SupLM, ExpLM, and AvgLM of $H_0 : c = 0$ against $H_1 : c \neq 0$ in (5.4). For r, we use the interval $\left[y_{(0.15*T)}, y_{(0.85*T)}\right]$ where $y_{(0.15*T)}$ and $y_{(0.85*T)}$ are the 15th and 85th percentiles of the empirical distribution of y_t respectively. The tests SupLM, ExpLM, and AvgLM do not have a pivotal distribution (see Tong, 1990). Hence we compute their asymptotic p-values using the method described in Hansen (1996) based on $J = 300$ artificial observations. From these p-values, we compute the empirical power of the tests. As the distributions of SupLM, ExpLM, and AvgLM depends on unknown parameters, there is no simple way to compute the size-corrected powers of these tests, therefore we will not report them here. We obtain the power of T_T and Th_T using the chi-square distribution with one degree of freedom. It is worth noting that while $\chi^2(1)$ is the asymptotic distribution of Th_T, it is not that of T_T because of the heteroskedasticity. However, the results of T_T are reported as a benchmark for what one would get if one omits to take into account the heteroskedasticity.

As before, we generate 200 extra data and discard them to ensure stationarity. Simulations are programmed in Gauss 5.0. The sample sizes under consideration are $T = 50, 100, 200, 500, 1000$. For each sample size, we iterate 1000 times. The results are reported in Table 3 for two values of c, 0.4 and 0.8.

From Table 3, we see that the test Th_T has greater power than SupLM, AvgLM, and ExpLM for all values of c and sample sizes, except for $T = 1000$ and $c = 0.4$ where ExpLM dominates. This shows that Th_T is a good test to detect threshold nonlinearity.

6. Conclusion

This paper describes a class of tests for testing problems in which a nuisance parameter exists under the alternative hypothesis but not under the null. We chose to focus on a particular member of this class, which is simply a Lagrange Multiplier test. However, the method proposed in Section 2 permits to construct other tests by choosing other formulations of the operator B. The LM test we investigate in detail has the advantages of (i) being simple to implement, (ii) having a standard chi-square distribution, and (iii) having power against a wide range of alternatives. Moreover, a Monte Carlo experiment shows that the tests SC_T and SCh_T have little size distortion and have good (empirical and size-corrected) power in small samples compared to competing tests. This makes these tests particularly attractive when working in small samples. Another Monte Carlo experiment shows that the test Th_T has a high power against a Threshold autoregressive alternative. As the p-values of Th_T are much faster to compute than those of the competing tests like the SupLM test, Th_T can be used as a quick way to detect threshold nonlinearities.

A. Proofs

Proof of Proposition 3.1. Let $D[0,1]$ be the set of all real valued functions on the $[0,1]$ −interval that are right continuous and have left limits. Denote $D[0,1]^p = D[0,1] \times ... \times D[0,1]$ the product metric space. We will establish the weak convergence in $D[0,1]$ endowed with the Skorohod metric (see Billingsley, 1968).

(a) Note that

$$\begin{aligned} \frac{1}{\sqrt{T}} S_T &= \frac{1}{T} \sum_{t=1}^{T} \frac{1}{\sqrt{T}} \sum_{s=1}^{t} x_s \left(y_s - z_s' \tilde{\theta} \right) \\ &= \int_0^1 \frac{1}{\sqrt{T}} \sum_{s=1}^{[T\pi]} x_s \left(y_s - z_s' \tilde{\theta} \right) d\pi, \end{aligned}$$

where $[T\pi]$ denotes the integer part of $T\pi$. Replacing y_s by its expression, we obtain

$$\begin{aligned} \frac{1}{\sqrt{T}} S_T &= \int_0^1 \frac{1}{T} \sum_{s=1}^{[T\pi]} x_s x_s' g\left(\frac{s}{T} \right) d\pi \\ &\quad - \left(\int_0^1 \frac{1}{T} \sum_{s=1}^{[T\pi]} x_s z_s' d\pi \right) \sqrt{T} \left(\tilde{\theta} - \theta_0 \right) \\ &\quad + \int_0^1 \frac{1}{\sqrt{T}} \sum_{s=1}^{[T\pi]} x_s \varepsilon_s d\pi. \end{aligned}$$

Replacing $\sqrt{T}\left(\tilde{\theta} - \theta_0\right)$ by its expression, we have

$$\frac{1}{\sqrt{T}} S_T$$

$$= \int_0^1 \frac{1}{T} \sum_{s=1}^{[T\pi]} x_s x_s' g\left(\frac{s}{T} \right) d\pi \tag{A.1}$$

$$- \left(\int_0^1 \frac{1}{T} \sum_{s=1}^{[T\pi]} x_s z_s' d\pi \right) \left(\frac{1}{T} \sum z_t z_t' \right)^{-1} \frac{1}{T} \sum z_t x_t' g\left(\frac{t}{T} \right) \tag{A.2}$$

$$+ \int_0^1 Z_T(\pi) \, d\pi$$

with

$$Z_T(\pi) = \frac{1}{\sqrt{T}} \sum_{s=1}^{[T\pi]} x_s \varepsilon_s - \left(\int_0^1 \frac{1}{T} \sum_{s=1}^{[T\pi]} x_s z_s' d\pi \right) \left(\frac{1}{T} \sum_{s=1}^{T} z_t z_t' \right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_t \varepsilon_t.$$

Moreover by Lemma 4 of Krämer, Ploberger, and Alt (1988) and by Assumption 3.2(i), the following relationships hold uniformly in π :

$$\frac{1}{T} \sum_{s=1}^{[T\pi]} x_s z_s' \xrightarrow{P} \pi \Sigma_{xz},$$

$$\frac{1}{T}\sum_{s=1}^{[T\pi]} x_s x_s' g\left(\frac{s}{T}\right) \xrightarrow{P} \int_0^{\pi} \Sigma_{xx} g(v)\, dv,$$
$$\frac{1}{T}\sum z_t x_t' g\left(\frac{t}{T}\right) \xrightarrow{P} \int_0^{1} \Sigma_{zx} g(v)\, dv.$$

It follows that

$$(A.1) + (A.2) \xrightarrow{P} \int_0^1 \left(\int_0^{\pi} \Sigma_{xx} g(v)\, dv\right) d\pi - \frac{1}{2}\Sigma_{xz}\Sigma_{zz}^{-1}\int_0^1 \Sigma_{zx} g(v)\, dv.$$

Moreover, using the matrix inversion formula, we have the simplification $\Sigma_{xx} = \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx}$.

Now, we turn our attention to the term $Z_T(\pi)$. By Lemma 3 of Krämer, Ploberger, and Alt (1988) and under Assumptions 3.1 and 3.2, weak convergence holds in $D[0,1]$. The process $\frac{1}{\sqrt{T}}\sum_{s=1}^{[T\pi]} x_s\varepsilon_s$ converges in distribution to a $p-$dimensional Gaussian process with covariance $\sigma^2\pi\Sigma_{xx}$. $\frac{1}{\sqrt{T}}\sum_{s=1}^{T} z_t\varepsilon_t$ converges in distribution to a centered normal with covariance $\sigma^2\Sigma_{zz}$. By the continuous mapping theorem (CMT), $Z_T(.)$ converges in $D[0,1]^p$ to a Gaussian process, $Z(.) \sim \mathcal{N}(0,K)$ where K is the covariance operator with kernel:

$$\begin{aligned} & E\left(Z_T(\pi_1) Z_T(\pi_2)'\right) \\ = \; & \sigma^2\left\{\pi_1 \wedge \pi_2 \Sigma_{xx} - \frac{1}{2}(\pi_1+\pi_2)\Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx} + \frac{1}{4}\Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx}\right\} \\ \equiv \; & k(\pi_1,\pi_2) \end{aligned}$$

where $\pi_1 \wedge \pi_2$ is the minimum of π_1 and π_2. Again by the CMT, $\int_0^1 Z_T(\pi)\, d\pi$ converges in distribution to $\int_0^1 Z(\pi)\, d\pi$. As $Z(.)$ is a function of a Wiener process, it belongs a.s. to $C^p[0,1]$ the space of continuous functions defined on $[0,1]$. By Shorack and Wellner (1986, page 42), $\int_0^1 Z(\pi)\, d\pi = (1, Z(.))$ is normally distributed with mean zero and covariance:

$$\begin{aligned} (1, K1) &= \int_0^1\int_0^1 k(\pi_1,\pi_2)\, d\pi_1 d\pi_2 \\ &= \sigma^2\left\{\frac{1}{3}\Sigma_{xx} - \frac{1}{4}\Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx}\right\} \\ &= \frac{\sigma^2}{12}\Sigma_{xx}. \end{aligned}$$

To complete the proof, it remains to show that

$$\frac{1}{\hat{\sigma}^2}\Lambda'\Sigma_T^{-1}\Lambda \xrightarrow{P} \left\{\frac{\sigma^2}{12}\Sigma_{xx}\right\}^{-1}.$$

This follows again from Lemma 4 of Krämer et al. (1988).

(b) We first need to establish that Ω_T converges in probability to

$$\Omega = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{T} E\left(h_t(\theta_0)\, h_{t-j}'(\theta_0)\right)$$

with

$$h_t(\theta_0) = \left[\begin{array}{c} z_t\varepsilon_t \\ \left(\frac{t}{T}\right) x_t\varepsilon_t \end{array}\right].$$

Note that

$$\Omega = \left[\begin{array}{cc} C_{zz} & \frac{1}{2}C_{zx} \\ \frac{1}{2}C_{xz} & \frac{1}{3}C_{xx} \end{array}\right].$$

Note that $\{h_t(\theta_0)\}$ is not covariance stationary because of the term t/T, hence the conditions of Andrews (1991) are not satisfied. However, the conditions of Hansen (1992) do not require covariance stationarity, instead $\{h_t(\theta_0)\}$ needs to be $\alpha-$mixing with α_m satisfying (3.3) in Assumption 3.4. As t/T is bounded, $\{h_t(\theta_0)\}$ is $\alpha-$mixing with the same coefficient as that of $\{z_t\varepsilon_t\}$. Moreover, we have $\sqrt{T}\left(\tilde{\theta}-\theta_0\right) = O_p(1)$ and $E\sup_{\theta\in\mathcal{N}}\|h_t(\theta)\|^2 < \infty$ where $\mathcal{N}$ is a neighborhood of θ_0 by Assumption 3.4(b). The conditions of Theorem 2 of Hansen (1992) are fulfilled and it follows that $\Omega_T \xrightarrow{P} \Omega$.

We start the proof as in (a). Under Assumption 3.4 and by Doukhan (1994), $Z_T(\pi)$ converges in distribution in $D^p[0,1]$ to a Gaussian process $\tilde{Z}(\pi)$ with covariance kernel

$$\begin{aligned}\tilde{k}(\pi_1,\pi_2) &= \pi_1\wedge\pi_2 C_{xx} \\ &\quad -\frac{1}{2}\pi_1\Sigma_{xz}\Sigma_{zz}^{-1}C_{zx} - \frac{1}{2}\pi_2 C_{xz}\Sigma_{zz}^{-1}\Sigma_{zx} \\ &\quad +\frac{1}{4}\Sigma_{xz}\Sigma_{zz}^{-1}C_{zz}\Sigma_{zz}^{-1}\Sigma_{zx}.\end{aligned}$$

Moreover $\int_0^1 \tilde{Z}(\pi)\,d\pi$ is normally distributed with mean 0 and covariance

$$\begin{aligned}&\frac{1}{3}C_{xx} - \frac{1}{4}\left\{\left(\Sigma_{xz}\Sigma_{zz}^{-1}C_{zx}\right) + \left(\Sigma_{xz}\Sigma_{zz}^{-1}C_{zx}\right)'\right\} + \frac{1}{4}\Sigma_{xz}\Sigma_{zz}^{-1}C_{zz}\Sigma_{zz}^{-1}\Sigma_{zx} \\ =\ &\frac{1}{12}C_{xx}. \qquad \text{(A.3)}\end{aligned}$$

The simplification in (A.3) follows from the fact that $\Sigma_{xz}\Sigma_{zz}^{-1} = (I_p\ \mathbf{O}_{p\times q})$.

Proof of Proposition 4.1 Under H_{1T}, we have

$$\begin{aligned}&\frac{1}{\sqrt{T}}S_T \\ =\ &\frac{1}{T}\sum_{t=1}^{T}F(u_t)\,x_t\left[x_t'g(v_t) - z_t'\left(\sum_{t=1}^{T}z_tz_t'\right)^{-1}\sum_{s=1}^{T}z_sx_s'g(v_s)\right] \qquad \text{(A.4)} \\ &+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F(u_t)\,x_t\left[\varepsilon_t - z_t'\left(\sum_{t=1}^{T}z_tz_t'\right)^{-1}\sum_{s=1}^{T}z_s\varepsilon_s\right]. \qquad \text{(A.5)}\end{aligned}$$

Because $\{z_t, u_t\}$ is stationary ergodic, the law of large numbers applies and (A.4) converges in probability to φ. By Assumptions 4.1, 4.2, we have

$$\left(\begin{array}{c} \frac{1}{\sqrt{T}}\sum_{t=1}^{T}F(u_t)\,x_t\varepsilon_t \\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T}z_t\varepsilon_t \end{array}\right) \xrightarrow{L} \mathcal{N}\left(\left(\begin{array}{c}0\\0\end{array}\right), \sigma^2\left(\begin{array}{cc}\Sigma_{xxF2} & \Sigma_{xzF} \\ \Sigma_{zxF} & \Sigma_{zz}\end{array}\right)\right).$$

Hence (A.5) converges in distribution to a normal with mean 0 and variance

$$\sigma^2 \left(\Sigma_{xxF2} - \Sigma_{xzF} \Sigma_{zz}^{-1} \Sigma_{zxF} \right).$$

The limiting distribution of Th_T is obtained as in the proof of Proposition 3.1.

B. Tables

Table 1: Empirical and size-corrected power of SupLM, AvgLM, ExpLM, $\mathbf{SC}_T$ and $\mathbf{SCh}_T$ tests for the One-Time Structural Change Model

The DGP is

$$\begin{cases} y_t = 0.1y_{t-1} + cy_{t-1}I\{t > 0.5T\} + \varepsilon_t \\ \varepsilon_t \sim Garch(1,1) \end{cases}$$

T	Test	$c = 0.5$			$c = 0.8$		
	Test size	1%	5%	10%	1%	5%	10%
60	SupLM power	1.6	3.4	6.85	2.65	7.35	19.15
	size correct.	1.05	8.85	20.65	1.6	25.25	49.15
	AvgLM power	1.25	4.85	12.5	2.6	16.35	36.6
	size correct.	2.95	16.65	28.9	9	46.8	65.8
	ExpLM power	2.05	6	13.05	4	18.6	38.75
	size correct.	1.15	11.05	27.35	1.8	35.1	61.75
	SC_T power	16.5	33.7	45.85	48.15	69.65	77.95
	size correct.	8.7	24.65	35.7	33.4	60.15	71.7
	SCh_T power	3.1	17.25	32.15	9.7	42.85	63.55
	size correct.	3.4	18.55	29.75	12.1	44.5	59.95
100	SupLM power	2.05	8.15	18.15	8	35.5	60.7
	size correct.	6.75	25.5	41.7	29.4	69.75	82.7
	AvgLM power	1.6	11.85	27.2	9.35	48.75	73.7
	size correct.	9.15	30.9	48.2	41.45	77.7	90
	ExpLM power	3.3	15	30.85	16.9	57.8	77.45
	size correct.	6	28.85	46.1	30.95	75.7	89
	SC_T power	31.55	51.2	62.25	75.3	88.85	92.65
	size correct.	13.15	33.7	47.8	52.6	78.2	86.8
	SCh_T power	7.05	34.15	51.55	34.45	73.25	87.6
	size correct.	7	28.65	44.65	34.05	68.3	83.65

T	Test	$c = 0.5$			$c = 0.8$		
	Test size	1%	5%	10%	1%	5%	10%
200	SupLM power	11	36.8	54	62.6	88.6	95.2
	size correct.	21.8	48.9	62.7	78.4	94.5	97.6
	AvgLM power	8.8	41.6	59.8	55.8	91.4	98.2
	size correct.	21.6	51.2	67.8	78.7	96.2	99.2
	ExpLM power	16.5	48.2	64.7	74.3	94.4	98.7
	size correct.	22	53.5	68.4	79.7	96.1	99.3
	SC_T power	64.3	79.1	84.4	96.5	98.5	99.2
	size correct.	24.3	55.6	67.9	81.5	95.1	96.6
	SCh_T power	30.7	63	75.8	81.1	96.1	98.2
	size correct.	23.4	51.6	66.9	74.2	92.8	96.9
500	SupLM power	60.3	83.9	90.8	97.3	99.5	100
	size correct.	52.7	86.8	92.3	96.2	99.8	100
	AvgLM power	50.8	81.9	91.2	96.2	99.3	99.8
	size correct.	48.7	88.2	94.3	95.7	99.7	100
	ExpLM power	66.6	87	93.5	98	99.6	99.9
	size correct.	48.4	89.5	94.1	95.5	99.7	100
	SC_T power	92	95.4	96.8	99.6	99.8	99.9
	size correct.	44.5	83.8	92	95.9	99.4	99.6
	SCh_T power	72.5	88.9	93.3	97	99.5	99.7
	size correct.	68.1	86	92.2	96.3	99.1	99.7
1000	SupLM power	91	97.3	99.1	99.7	99.9	100
	size correct.	83.9	98.1	99.1	99.6	99.9	100
	AvgLM power	87.6	96.2	98.5	99.8	99.8	99.9
	size correct.	87.7	97.7	99.1	99.8	99.8	100
	ExpLM power	92.6	97.9	99.2	99.8	99.9	99.9
	size correct.	88.2	98.5	99.2	99.7	99.9	99.9
	SC_T power	99.4	99.7	99.8	100	100	100
	size correct.	86.5	97.2	98.8	99.7	100	100
	SCh_T power	92.8	97.8	98.8	99.5	100	100
	size correct.	87.5	96.8	98.5	99.2	100	100

Table 2: Empirical and size-corrected power of SupLM, AvgLM, ExpLM, $\mathbf{SC}_T$ and $\mathbf{SCh}_T$ tests for the trending coefficient model

The DGP is

$$\begin{cases} y_t = 0.1y_{t-1} + c\frac{t}{T}y_{t-1} + \varepsilon_t \\ \varepsilon_t \sim Garch(1,1) \end{cases}$$

T	Test	$c = 0.5$			$c = 0.8$		
	Test size	1%	5%	10%	1%	5%	10%
60	SupLM power	1.6	2.8	5.05	2.3	5.55	10.35
	size correct.	1	6.2	13.75	1.5	13.5	26.15
	AvgLM power	0.9	3.2	8.1	2.05	10.25	21.1
	size correct.	2	10.6	19	6.05	26.25	39.2
	ExpLM power	1.75	4.35	8.8	2.9	10.9	20.75
	size correct.	1.1	7.8	17.1	1.75	18.2	35.2
	SC_T power	8.25	20.45	29.1	24.85	45.25	55.7
	size correct.	4	13	21.95	15.6	33.95	47.15
	SCh_T power	2.3	10	21.3	5.65	24.9	42.15
	size correct.	2.6	10.9	19.15	7	26.4	39.05
100	SupLM power	1.2	4.15	9.85	3.05	13.85	28.4
	size correct.	3.35	13.65	23	11.5	36.85	50.45
	AvgLM power	0.95	6.3	14.5	5.05	24.55	42.4
	size correct.	4.55	16.2	27	19.6	45.65	60.35
	ExpLM power	1.65	7.6	15.8	6.55	26.05	43.1
	size correct.	2.85	14.7	25.7	12.15	41.05	57.55
	SC_T power	15.3	30.65	40	44.55	64.6	74.65
	size correct.	6.25	16.85	27.8	23.5	47.25	61.4
	SCh_T power	3.3	17	30.95	15.6	46.35	63.8
	size correct.	3.2	14.45	25.6	15.55	41.7	57.3

T	Test	$c = 0.5$			$c = 0.8$		
	Test size	1%	5%	10%	1%	5%	10%
200	SupLM power	2	12.2	24.3	19.4	48.7	64.9
	size correct.	6	20.7	31.3	33.2	61.3	73.2
	AvgLM power	2.9	15.1	29.9	26.4	60.3	74.8
	size correct.	7.9	22.7	38.2	43.2	67.3	82.1
	ExpLM power	4.3	17.8	33.2	31.1	62.2	77.6
	size correct.	6.6	22	37.1	37.7	67.5	80.2
	SC_T power	34.3	53.1	61.7	78.6	89.6	93
	size correct.	7.3	26.5	38.5	44.7	72.8	79.7
	SCh_T power	11.3	34.4	49.2	49.8	77.5	87.6
	size correct.	7.6	24.9	39.3	41	69.1	80.7
500	SupLM power	16.9	41.8	57.9	73.6	90.1	96
	size correct.	13.4	48.9	61.2	66.6	92.9	96.7
	AvgLM power	14.7	41.3	58.5	76.2	93	96.7
	size correct.	13.8	52.2	68.6	75.5	95.6	97.9
	ExpLM power	21.1	47.9	64	81.5	94.2	97.5
	size correct.	11.4	53.3	66.9	66.3	95.5	97.9
	SC_T power	68.4	79.7	85.4	97.8	98.7	99.1
	size correct.	12.3	46.9	66.6	76.1	93.8	97.6
	SCh_T power	34.3	62	73	88.8	97.1	98.4
	size correct.	28.6	56.3	69.4	85.7	95.7	98.2
1000	SupLM power	45.9	71.5	83.3	96.1	99.2	99.7
	size correct.	29.9	74.1	83.9	91.7	99.5	99.7
	AvgLM power	42.1	70.6	83.1	96.1	99.2	99.5
	size correct.	43.5	78.4	86.6	96.2	99.4	99.6
	ExpLM power	52	76.8	86.1	97.2	99.4	99.7
	size correct.	38.9	79.1	86.3	95.4	99.6	99.7
	SC_T power	88.9	93.7	95.7	99.9	99.9	99.9
	size correct.	40.3	70.3	84	97.9	99.8	99.9
	SCh_T power	64.1	83.5	89.6	98.1	99.6	99.6
	size correct.	49.3	80	88.4	96.5	99.3	99.6

Table 3: Empirical power of SupLM, AvgLM, ExpLM, T_T and Th_T for a TAR model

The DGP is

$$\begin{cases} y_t = 0.1 y_{t-1} + c y_{t-1} I\left(y_{t-1} > .05452\right) + \varepsilon_t \\ \varepsilon_t \sim GARCH(1,1) \end{cases}$$

T	Test	c = 0.4			c = 0.8		
Test Size		1%	5%	10%	1%	5%	10%
50	SupLM	0.3	6.4	15.9	0.6	7.8	17.1
	AvgLM	0.3	8.1	19.7	0.6	11	22.7
	ExpLM	0.2	7.8	19.2	0.3	9.7	20.4
	T_T	0.1	2.8	6.2	0.6	3.6	10.7
	Th_T	1.8	10.4	24.6	1.1	11.2	26
100	SupLM	3.7	20.4	30.7	4.8	23.3	40
	AvgLM	6.1	22.3	35.9	6.6	30.7	51.1
	ExpLM	5.7	22.3	35.4	6	29.4	50.5
	T_T	3.2	13.3	24	5.5	22.1	39
	Th_T	7.1	26.7	41.8	7.3	35.8	57.1
200	SupLM	12.9	38.3	52.9	22.2	55.9	73.6
	AvgLM	16	42.5	57.2	30.2	63.6	79.4
	ExpLM	16	42.8	57.4	30.2	64.7	81.1
	T_T	11.8	34	46.1	29.7	61.9	75.3
	Th_T	22.4	49.2	61.6	36.3	74	86.3
500	SupLM	41.6	66.7	76.4	70.9	89.3	95.2
	AvgLM	46.2	68.9	78.7	74.4	91.6	97
	ExpLM	46.6	69.2	79.3	77	92	96.9
	T_T	45.3	66.7	77.9	89.1	97.4	99.2
	Th_T	52.8	71.6	80.4	83.7	96.1	98.4
1000	SupLM	70.9	87.8	92.7	90.3	96.5	98.3
	AvgLM	73.7	89.3	93.1	91.3	96.9	98.6
	ExpLM	73.9	89.5	93.3	91.4	97	98.6
	T_T	80.4	92.1	95.1	99.7	99.9	99.9
	Th_T	77.2	88.8	92.9	95.1	99.1	99.8

References

Altissimo, F. and V. Corradi (2002) "Bounds for inference with nuisance parameters present only under the alternative", *Econometrics Journal*, 5, 494-519.

Andrews, D. (1991) "Asymptotic normality of series estimators for for nonparametric and semiparametric regression models", *Econometrica*, **59**, No. 2, 307-345.

Andrews, D.W.K (1993) "Tests for Parameter Instability and Structural Change Point",

Econometrica, 61, 821-856.
Andrews, D.W.K, I. Lee, and W. Ploberger (1996) "Optimal changepoint tests for normal linear regression", *Journal of Econometrics*, **70**, 9-38.
Andrews, D.W.K. and Monahan, J. C. (1992) "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator" *Econometrica,* **60**, 953-66.
Andrews, D.W.K. and W. Ploberger (1994) "Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative", *Econometrica*, **62**, No 6, 1383-1414.
Billingsley, P. (1968) *Convergence of Probability Measures*, Wiley, New York.
Brown, R. L., J. Durbin, and J. M. Evans (1975) "Techniques for Testing the Constancy of Regression Relationships over Time", *Journal or the Royal Statistical Society*, B, 37, 149-163.
Carrasco, M. (2002) "Misspecified Structural Change, Threshold, and Markov Switching Models", *Journal of Econometrics,* **109**, 239-273.
Carrasco, M. and J. P. Florens (2000) "Generalization of GMM to a continuum of moment conditions", *Econometric Theory*, **16**, 797-834.
Carrasco, M. and J. P. Florens (2002) "Efficient GMM Estimation Using the Empirical Characteristic Function", mimeo, University of Rochester.
Carrasco, M., M. Chernov, J. P. Florens and E. Ghysels (2004) "Efficient Estimation of Jump Diffusions and General Dynamic Models with a Continuum of Moment Conditions", mimeo, University of Rochester.
Davidson, R. and J. MacKinnon (1987) "Implicit Alternatives and the Local Power of Test Statistics", *Econometrica*, **55**, 1305-1329.
Davies, R.B. (1977) "Hypothesis testing when a nuisance parameter is present only under the alternative" *Biometrika*, **64**, 2, 247-254.
Diebold, F.X. and C. Chen (1996) "Testing structural stability with endogenous breakpoint: A size comparison of analytic and bootstrap procedures", *Journal of Econometrics*, **70**, 221-241.
Doukhan, P. (1994) *Mixing: Properties and Examples*, Springer Verlag, New-York.
Dufour, J. M. (1997), "Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models," *Econometrica*, 65, 1365–1387.
Forchini, G. (2002) "Optimal Similar Tests for Structural Change for the Linear Regression Model", *Econometric Theory*, **18**, 853-867.
Gallant, A.R. (1977) "Testing a Nonlinear Regression Specification: A Nonregular Case", *Journal of the American Statistical Association*, **72**, 523-530.
Godfrey, L. G. (1988) *Misspecification Tests in Econometrics*, Cambridge University Press, Cambridge.
Granger, C.W.J. and T. Teräsvirta (1993) *Modelling Nonlinear Economic Relationships*, Oxford University Press, New York.
Hamilton, J.D. (1989) "A new Approach to the Economic Analysis of Nonstationary Time

Series and the Business Cycle", *Econometrica*, **57**, 357-384.
Hansen, B.E. (1992) "Consistent Covariance Matrix estimation for Dependent Heterogeneous Processes", *Econometrica*, **60**, 967-972.
Hansen, B.E. (1996) "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis", *Econometrica*, **64**, No. 2, 413-430.
Krämer, W., W. Ploberger, and R. Alt (1988) "Testing for Structural Change in Dynamic Models", *Econometrica*, **56**, 1355-1369.
Lehmann, E. (1959) *Testing statistical hypotheses*, Wiley, New York.
Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988) "Testing linearity against smooth transition autoregressive models", *Biometrika*, **75**, 491-499.
Newey, W. and K. West (1987a) "Hypothesis Testing with Efficient Method of Moments Estimation", *International Economic Review*, **28**, 777-787.
Newey, W. and K. West (1987b) "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica*, **55**, No. 3, 703-708.
Nyblom, J. (1989)"Testing for the Constancy of Parameters Over Time", *Journal of the American Statistical Association*, **84**, 223-230.
Pagan, A.R. and A.D. Hall (1983) "Diagnostic Tests as Residual Analysis", *Econometric Reviews*, 2 (2), 159-218.
Ploberger, W. and W. Krämer (1992) "A tend-resistant test for strutural change based on OLS residuals", *Journal of Econometrics*, 70, 175-185.
Ramsey, J.B. (1969) "Tests for specification errors in classical linear least-squares regression analysis", *Journal of the Royal Statistical Society*, B, 31, 350-371.
Shorak, G. and J. Wellner (1986) *Empirical Processes with Applications to Statistics*, John Wiley & Sons, New York.
Szroeter, J. (1978) "A Class of Parametric Tests for Heteroscedasticity in Linear Econometric Models", *Econometrica*, **46**, 1311-1327.
Tang, S. M. and I. B. MacNeill (1993) "The effect if serial correlation on tests for parameter change at unknown time", *The Annals of Statistics*, **21**, 552-575.
Teräsvirta, T. (1994) "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models", *Journal of the American Statistical Association*, 89, 208-217.
Tong, H. (1990) *Non-linear Time Series*. Oxford University Press.
Tsay, R. (1989) "Testing and Modeling Threshold Autoregressive Processes", *Journal of the American Statistical Association*, **84**, No. 405, 231-240.
White, H. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, **48**, 817-838.