External bootstrap tests for parameter stability

Miguel A. Delgado*, Inmaculada Fiteni

Universidad Carlos III de Madrid, Madrid 126-128, 28903 GETAFE Madrid, Spain

Abstract

This article considers tests for parameter stability over time in general econometric models, possibly nonlinear-in-variables. Existing test statistics are commonly not asymptotically pivotal under nonstandard conditions. In such cases, the external bootstrap tests proposed in this paper are appealing from a practical viewpoint. We propose to use bootstrap versions of the asymptotic critical values based on a first-order asymptotic expansion of the test statistics under the null hypothesis, which consists of a linear transformation of the unobserved "innovations" partial sum process. The nature of these transformations under nonstandard conditions is discussed for the main testing principles. Also, we investigate the small sample performance of the proposed bootstrap tests by means of a small Monte Carlo experiment. © 2002 Elsevier Science B.V. All rights reserved.

JEL classification: C22

Keywords: Structural stability tests; Nonlinear-in-variables models; Empirical processes; GMM; External bootstrap

1. Introduction and statement of the problem

Asymptotically admissible hypothesis tests on parameter stability can be easily implemented in practice, even when test statistics are not asymptotically pivotal, by using bootstrap estimates of asymptotic critical regions. This paper suggests to apply this approach in the context of general econometric models, possibly nonlinear-in-variables.

Bootstrap tests allow to implement hypothesis tests under fairly general conditions, in situations where test statistics limiting distributions are case dependent and, therefore, tables are not available.

Bootstrap versions of test statistics are usually obtained from a resample of the original data set, after imposing the restrictions to be tested on the null hypothesis,

* Corresponding author. Tel.: +34-91-6249804; fax: +34-91-6249849.
E-mail address: delgado@est-econ.uc3m.es (M.A. Delgado).

but respecting unknown features of the underlying data generating process affecting the limiting distribution under the null. Under suitable resampling schemes, conditional quantiles, given the sample, of these bootstrap versions are consistent estimators of the corresponding asymptotic critical values, and, therefore, form a basis for asymptotically admissible tests. Bootstrap approximations also permit to use simpler test statistics in practice, avoiding necessary scaling for obtaining asymptotically pivotal tests.

Hansen (2000) has suggested a bootstrap approach in the context of Quandt (1960) and Nyblom (1989) parameter stability tests for the linear regression model. He proposes to apply Wu (1990) external bootstrap residual resample, which consists in fixing the exogenous variables and then, resampling the endogenous variables by adding to the OLS fits the corresponding residuals multiplied by an independent standard normal random variable. This procedure is difficult to justify in nonlinear-in-variables models, where innovations enter nonlinearly in "reduced form" equations. Instead of using bootstrap analogs of the original test statistics, we propose to use external bootstrap versions of a first-order expansion, as it has been suggested by Su and Wei (1991) and Hansen (1996), among others, in other contexts.

Let $\mathscr{Z}_n = \{Z_{ni}, i = 1, 2, \ldots n, n = 1, 2, \ldots\}$ be a $\mathbb{R}^q$-*valued* double array of random variables. Consider the nonlinear-in-variables simultaneous equations system,

$$U(Z_{ni}, \theta_{ni}) = U_{ni}, \quad i = 1, \ldots, n; \;\; n \geqslant 1,$$

where θ_{ni} is a *p-valued* parameter vector, $U : \mathbb{R}^q \times \mathbb{R}^p \to \mathbb{R}^k$ are the known functions, and U_{ni} are the innovations satisfying that

$$\mathrm{E}[U_{ni}] = 0 \quad \text{for all } i \geqslant 1.$$

The parameter stability hypothesis is stated in terms of a parameter θ_0, defined as follows:

A1. There exists a parameter vector θ_0, which is an interior point of $\Theta \subset \mathbb{R}^p$, and the only solution to the set of equations

$$\lim_{n\to\infty} \mathrm{E}[T_n^0(1)] = 0,$$

where $T_n^0(\cdot) := T_n(\cdot, \theta_0)$ and $T_n(\cdot, \theta) := n^{-1} \sum_{i=1}^{[n\cdot]} U(Z_{ni}, \theta)$.

These conditions identify the parameter θ_0, and the null hypothesis can be equivalently written as

$$\mathrm{H}_0 : \theta_{ni} = \theta_0 \quad \text{for all } i \geqslant 1 \quad \text{and some } \theta_0 \in \Theta \subset \mathbb{R}^p.$$

Tests are designed to be consistent in the direction of fixed alternatives,

$$\mathrm{H}_{1n}(\tau_0) : \theta_{ni} = \begin{cases} \theta_1(\tau_0), & i = 1, \ldots, [n\tau_0] \\ \theta_2(\tau_0), & i = [n\tau_0] + 1, \ldots, n \end{cases}$$

$$\text{for some } \tau_0 \in [0, 1], \quad \text{and} \quad \theta_1(\tau_0), \theta_2(\tau_0) \in \Theta \subset \mathbb{R}^p,$$

or the more general formulation $\bigcup_{\tau\in\Pi} H_{1n}(\tau)$ for some $\Pi \subset [0, 1]$ with nonzero Lebesgue measure. These tests are also able to detect contiguous alternatives of

the form

$$\mathrm{H}_{Cn}: \theta_{ni} = \theta_0 + \frac{g(i/n)}{\sqrt{n}} \quad \text{for some bounded function } g:[0,1] \to \mathbb{R}^P.$$

As it will be discussed in detail in Section 3, test statistics based on the main testing principles, such as *CUSUM*, likelihood ratio (*LR*), Wald, lagrange multiplier (*LM*) or Predictive tests can be written as

$$\hat{\eta}_n = \varphi(n^{1/2} R_n),$$

where $\{R_n\}$ is an $m \times 1$ vector ($m \geqslant k$) of empirical processes taking their values in the space of right-continuous functions on $[0,1]$ with left-hand limits, the space $D[0,1]$. We will say that $\{R_n\}$ is a random element of $(D[0,1])^m$. The form of $\{R_n\}$ depends on the specific testing principle applied (see Section 3) and $\varphi:(D[0,1])^m \mapsto \mathbb{R}^+$ is some continuous functional. For *CUSUM* tests, popular for structural break testing in linear models, the functionals usually chosen are

$$\varphi(f) = \sup_{\tau \in [0,1]} |f(\tau)| \text{ (Kolmogorov–Smirnov's type tests)}, \tag{1}$$

$$\varphi(f) = \int_0^1 f(\tau)^2 \, \mathrm{d}\tau \text{ (Cramér–von Misses' type tests)}.$$

Some other examples of φ maps, used in tests statistics based on Wald, *LM*, *LR* or Predictive tests principles are

$$\varphi(f) = \sup_{\tau \in [\pi, 1-\pi]} \|f(\tau)\|^2 \text{ (e.g. Andrews, 1993)},$$

$$\varphi(f) = \int_0^1 \|f(\tau)\|^2 \, \mathrm{d}J(\tau) \text{ (e.g. Andrews and Ploberger, 1994)}, \tag{2}$$

$$\varphi(f) = \log \int_0^1 \exp\left(\frac{1}{2} \|f(\tau)\|^2\right) \mathrm{d}J(\tau) \text{ (e.g. Andrews and Ploberger, 1994)},$$

where J is a suitable weight function and, henceforth, for any matrix A, $\|A\| = \bar{\lambda}(A'A)^{1/2}$, where $\bar{\lambda}(\cdot)$ means the greatest eigenvalue.

Under certain regularity conditions, which will be discussed in Section 2, Wald, *LM*, *LR* or Predictive tests belong to the following class:

A2. The class of tests we are interested in is based on test statistics of the form $\hat{\eta}_n = \varphi(n^{1/2} R_n)$ for a given continuous functional $\varphi:(D[0,1])^m \mapsto \mathbb{R}^+$ and a suitable empirical process $\{R_n\}$ taking values in $(D[0,1])^m$, such that, under H_{Cn},

$$\sup_{\tau \in [0,1]} |R_n(\tau) - \mathscr{L} T_n^0(\tau)| = \mathrm{o_p}(n^{-1/2}),$$

where $\mathscr{L}:(D[0,1])^k \mapsto (D[0,1])^m$ is a given linear bounded transformation.

The specific form of the transformation $\mathscr{L}$ for the main testing principles: *CUSUM*, *LM*, *LR*, *Wald* and Predictive tests will be discussed in Section 3.

The empirical process $\{n^{1/2}T_n^0\}$ is a random element of $(D[0,1])^k$. Under fairly general regularity conditions, which will be discussed in Section 2, $\{n^{1/2}T_n^0\}$ converges in distribution to a Gaussian element with continuous sample paths in the space $(D[0,1])^k$ with the Skorohod's topology (see Billingsley, 1968, Chapter 3). Tests statistics are designed in such a way that under more or less restrictive regularity conditions the linear transformation $\mathscr{L}$ satisfies that $\{n^{1/2}\mathscr{L}T_n^0\}$ converges weakly to a vector of independent Brownian motions, and asymptotic critical values are usually approximated by Monte Carlo. However, under departures from the assumed regularity conditions, $\{n^{1/2}\mathscr{L}T_n^0\}$ has covariance function depending on unknown features of the data generating process, and the asymptotic distribution of resulting test statistics cannot be tabulated under general conditions. Some tests are not asymptotically pivotal when innovations are heteroskedastic or autocorrelated, e.g. Quandt (1960) and Nyblom (1989) tests for the linear regression model, and many others, which are robust under heteroskedasticity and autocorrelation (e.g. Andrews, 1993; Sowell, 1983), may have a case-dependent limiting distribution when exogenous variables or innovations are not covariance-stationary. Thus, it seems reasonable to use bootstrap critical values and p-values for preventing situations where test statistics are not asymptotically pivotal under the null hypothesis.

The rest of the paper is organized as follows. The next section discusses the asymptotic behavior of test statistics under departures from standard conditions. Section 3 shows that A2 is satisfied for leading classes of tests. Section 4 proposes bootstrap versions of $\tilde{\eta}_n$, based on an external bootstrap version of $\{T_n^0\}$, and justifies the validity of resulting bootstrap tests, which estimate asymptotic critical values and p-values from conditional quantiles, given the sample $\mathscr{Z}_n$, of these bootstrap statistics. Section 5 presents some remarks on the implementation of the bootstrap testing procedures in linear models. Section 6 illustrates the finite sample performance of proposed bootstrap tests by means of a small Monte Carlo experiment. Proofs of stated results are placed in a final section at the end of the paper.

2. Regularity conditions and asymptotic tests

Henceforth, "$\Rightarrow$" means convergence in distribution in the Skorohod metric space $(D[0,1])^k$, "$\xrightarrow{\mathrm{d}}$" denotes convergence in distribution on Euclidean sets, and "$\overset{\mathrm{d}}{=}$" means equally distributed. The next set of high level regularity conditions will be helpful for discussing the asymptotic behavior of test statistics.

A3. $U_{ni}(\cdot):=U(Z_{ni},\cdot)$ is differentiable in a neighborhood of θ_0, with derivatives $\dot{U}_{ni}(\theta):=\partial U_{ni}(\theta)/\partial\theta'$ such that, for any $\bar{\theta}_n=\theta_0+\mathrm{o_p}(1)$,

$$\sup_{\tau\in[0,1]}\|M_n(\tau,\bar{\theta}_n)-M_n(\tau,\theta_0)\|=\mathrm{o_p}(1) \tag{3}$$

and

$$\sup_{\tau\in[0,1]}\|M_n(\tau,\theta_0)-M(\tau)\|=\mathrm{o_p}(1), \tag{4}$$

where $M_n(\cdot,\theta)=n^{-1}\sum_{i=1}^{[n\cdot]}\dot{U}_{nt}(\theta)$ and $M(\tau)$ is a full rank matrix of continuous functions uniformly in τ.

Condition (3) is satisfied when $\dot{U}_{ni}(\cdot)$ is smooth enough. For instance, a sufficient condition is that, for some $\varepsilon > 0$,

$$\sup_{\{\theta:\ \|\theta-\theta_0\| \leqslant \varepsilon\}} \|\dot{U}_{ni}(\theta) - \dot{U}_{ni}(\theta_0)\| \leqslant \|\theta - \theta_0\| \xi_{ni} \quad \text{for each } i \geqslant 1,$$

such that $n^{-1} \sum_{i=1}^{n} \xi_{ni} = \mathrm{O_p}(1)$. The matrix $M(\cdot)$ of continuous functions in (4) is, typically, such that

$$\sup_{\tau \in [0,1]} \|\mathrm{E}(M_n(\tau, \theta_0)) - M(\tau)\| = \mathrm{o}(1). \tag{5}$$

This condition holds when, for instance, $\mathrm{E}[\dot{U}_{ni}(\theta_0)] = M(i/n)$, where M is a matrix of bounded and continuous functions, which are uniformly Riemman's integrable on intervals $[0, \tau]$ for all $\tau \in [0, 1]$. Thus, (4) follows from (5) and

$$\sup_{\tau \in [0,1]} \|M_n(\tau, \theta_0) - \mathrm{E}[M_n(\tau, \theta_0)]\| = \mathrm{o_p}(1). \tag{6}$$

This last condition is satisfied when $U_{ni}(\theta_0) - \mathrm{E}[U_{ni}(\theta_0)]$ is a martingale difference, and $M_n(1, \theta_0)$ converges in mean square. Note that under these circumstances, by Doob's inequality (see Doob, 1953, p. 317),

$$\mathrm{E}\left[\sup_{\tau \in [0,1]} \|M_n(\tau, \theta_0) - \mathrm{E}[M_n(\tau, \theta_0)]\|^2\right] \leqslant 2\mathrm{E}[\|M_n(1, \theta_0) - \mathrm{E}[M_n(1, \theta_0)]\|^2]. \tag{7}$$

Thus, (6) follows, applying Markov's inequality, if the right-hand side of (7) converges to zero, which is the case when, for instance, $\{\|\dot{U}_{ni}(\theta_0)\|^2, i \geqslant 1\}$ is uniformly integrable, applying Chow's (1971) result (see also Hall and Heyde, 1980, Section 2.7, for discussion on L_p convergence of martingales). The martingale assumption can be relaxed by applying the extension of Doob's inequality in MacLeish (1975a) to the case where $\{\dot{U}_{ni}(\theta_0), i \geqslant 1\}$ is a mixingale.

In order to implement a test, we also need to know the weak limit behavior of $\{n^{1/2} T_n^0\}$. Assume

A4.

$$\sup_{\tau \in [0,1]} \|T_n^0(\tau) - T^0(\tau)\| = \mathrm{o_p}(1) \quad \text{and} \quad \sup_{\tau \in [0,1]} \|n\mathrm{E}[T_n^0(\tau) T_n^0(\tau)'] - S(\tau)\| = \mathrm{o_p}(1),$$

where S is a matrix of continuous functions. Also,

$$n^{1/2}(T_n^0 - T^0) \Rightarrow N_k,$$

where $\{N_k\}$ is a Gaussian process with continuous sample paths, centered at zero and with covariance function $\mathrm{E}(N_k(\tau_1) N_k(\tau_2)') = S(\min(\tau_1, \tau_2)) - T^0(\tau_1) T^0(\tau_2)'$ for any $\tau_1, \tau_2 \in [0, 1]$.

This assumption is satisfied under a variety of circumstances. Billingsley (1968) proves a number of invariance principles for φ—*mixing* and functions of φ—*mixing* random variables, which were extended by Davidov (1968, 1970) to stationary sequences with more than two moments, allowing for α-*mixing* processes as well. Brown (1971) provides an invariance principle for martingales, and MacLeish (1975b, 1977) for nonstationary mixingales.

In the parameter stability literature, there are usually assumed conditions (see e.g. Andrews, 1993; Sowell, 1996) such that:

A3′. A3 holds with $M(\tau) = \tau M(1)$.

A4′. A4 holds with $S(\tau) = \tau S(1)$.

Therefore, under H_0, if A4′ holds, $N_k(\tau) \stackrel{\mathrm{d}}{=} S(1)^{1/2} B_k(\tau)$, where $\{B_k\}$ is a vector $k \times 1$ of independent standard Brownian motions. Distribution free tests are available when A3′ and A4′ are satisfied. However, A3′ may not hold when $\{\dot{U}_{ni}(\theta_0), i \geqslant 1\}$ is not mean stationary, and A4′ may not hold when $\{U_{ni}(\theta_0), i \geqslant 1\}$ is covariance non stationary. This is illustrated in the following examples.

Case I: Change in the distribution of Z_{ni} at an unknown moment of time. For instance,

$$Z_{ni} = Z_i^{(1)} 1(i \leqslant [n\tau_0]) + Z_i^{(2)} 1(i > [n\tau_0]) \quad \text{some } \tau_0 \in [0,1],$$

where $(Z_i^{(1)}, Z_i^{(2)})$ are i.i.d. Thus

$$U_{ni} = U(Z_i^{(1)}, \theta_{ni}) 1(i \leqslant [n\tau_0]) + U(Z_i^{(2)}, \theta_{ni}) 1(i > [n\tau_0])$$

and

$$M(\tau) = \tau M^{(1)} + (\tau - \tau_0)(M^{(2)} - M^{(1)}) 1(\tau > \tau_0),$$

$$S(\tau) = \tau S^{(1)} + (\tau - \tau_0)(S^{(2)} - S^{(1)}) 1(\tau > \tau_0),$$

where $n^{-1} \sum_{i=1}^{n} \dot{U}(Z_i^{(j)}, \theta_0) = M^{(j)} + \mathrm{o_p}(1)$, and $n^{-1} \sum_{i=1}^{n} U(Z_i^{(j)}, \theta_0) U(Z_i^{(j)}, \theta_0)' = S^{(j)} + \mathrm{o_p}(1)$, $j = 1, 2$. Therefore, A3′ and A4′ do not hold when $M^{(1)} \neq M^{(2)}$ and $S^{(1)} \neq S^{(2)}$, respectively.

Case II: Z_{ni} with trending components. For instance, $Z_{ni} = (Z_i, i/n)$, with Z_i i.i.d., and under regularity conditions,

$$M(\tau) = \int_0^\tau \mathrm{E}[\dot{U}(Z_1, u, \theta_0)]\, \mathrm{d}u \quad \text{and} \quad S(\tau) = \int_0^\tau \mathrm{E}[U(Z_1, u, \theta_0) U(Z_1, u, \theta_0)']\, \mathrm{d}u.$$

Another important case discussed by Hansen (2000) in the context of a linear regression model, consists of Z_{ni} having stochastic trending components. Then, $M(\tau)$ may be a stochastic matrix, and the weak behavior of $\{n^{1/2}(T_n^0 - T^0)\}$ will depend on the nonlinear structure of U. An excellent monograph on the asymptotic behavior of nonlinear functionals of random walks can be found in Borodin and Ibragimov (1995). Recently, Park and Phillips (2001) have provided asymptotic theory for nonlinear regression with integrated processes. The study of asymptotic inference in nonlinear-in-variables econometric models with stochastic trending variables deserves more attention in future research.

In order to establish a CLT for the test statistics under contiguous alternatives, H_{Cn}, we need a further assumption.

A5. For any $\bar{\theta}_{ni}$, such that $\|\bar{\theta}_{ni} - \theta_0\| \leqslant n^{-1/2} \|g(i/n)\|$,

$$\sup_{\tau \in [0,1]} \left\| \frac{1}{n} \sum_{i=1}^{[n\tau]} \dot{U}_{ni}(\bar{\theta}_{ni}) g\left(\frac{i}{n}\right) - L(\tau) \right\| = \mathrm{o_p}(1),$$

where $L(\tau) = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{[n\tau]} \mathrm{E}[\dot{U}_{ni}(\theta_0) g(i/n)]$ exists uniformly in $\tau \in [0,1]$.

When A3′ holds, it is expected also that the following condition is satisfied:

A5′. A5 holds and $L(\tau) = M(1)G(\tau)$ with $G(\tau) := \int_0^\tau g(u)\,\mathrm{d}u$.

Next theorem is immediate.

Theorem 1. *Assuming* A1–A5, *under* H_{Cn},

$$\hat{\eta}_n \xrightarrow{\mathrm{d}} \varphi(\mathscr{L}(N_k^0 + L)),$$

where $\{N_k^0\}$ *is a Gaussian process with continuous sample paths, centered at zero, and with covariance function* $\mathrm{E}(N_k^0(\tau_1)N_k^0(\tau_2)') = S(\min(\tau_1, \tau_2))$, *for any* $\tau_1, \tau_2 \in [0, 1]$ (*i.e.* $N_k^0 \overset{\mathrm{d}}{=} N_k$ *when* $T^0(\tau) = 0$ *uniformly in* τ).

Thus, if A1–A3 hold, under H_0, $\hat{\eta}_n \xrightarrow{\mathrm{d}} \eta_\infty \overset{\mathrm{d}}{=} \varphi(\mathscr{L}N_k^0)$, using the fact that $\sup_{\tau\in[0,1]} \|g(\tau)\| = 0$ under H_0, and, thus, $\sup_{\tau\in[0,1]} \|L(\tau)\| = 0$. The distribution of η_∞ is continuous with respect to the uniform metric. Under A3′ and A4′, the linear transformation $\mathscr{L}$ is typically such that, $\mathscr{L}N_k^0(\tau) \overset{\mathrm{d}}{=} s(\tau)[B_k(\tau) - \tau B_k(1)]$, where s is a known scalar function, usually $s(\tau) = 1$ or $s(\tau) = 1/\tau^{1/2}(1-\tau)^{1/2}$, and the distribution of η_∞ can be derived by Monte Carlo.

Next section discusses the nature of such transformations $\mathscr{L}$ in the context of some leading testing principles.

3. Asymptotic representation of tests statistics

Tests statistics can be constructed based on *Wald*, *LM* and *LR* principles (see Andrews, 1993). *CUSUM* tests, designed originally for linear regression models are also very popular and easy to compute. These last tests only require to estimate θ_0 under the restrictions of H_0 and, hence, they are related to *LM* tests. We also discuss predictive (P) tests, which are not related to the above principles.

In order to implement any of these tests we need some estimate of θ_0. The partial sum *GMM* (*PS-GMM*) estimators of $\theta_1(\tau)$ and $\theta_2(\tau)$ are defined as

$$\hat{\theta}_n(\tau) = (\hat{\theta}_{1n}'(\tau), \hat{\theta}_{2n}'(\tau))' = \underset{t_1, t_2 \in \Theta}{\operatorname{argmin}} \{Q_{1n}(\tau, t_1) + Q_{2n}(\tau, t_2)\},$$

where

$$Q_{1n}(\tau, \theta) = T_n(\tau, \theta)'\hat{S}_n(\tau)^{-1}T_n(\tau, \theta),$$

$$Q_{2n}(\tau, \theta) = [T_n(1, \theta) - T_n(\tau, \theta)]'[\hat{S}_n(1) - \hat{S}_n(\tau)]^{-1}[T_n(1, \theta) - T_n(\tau, \theta)],$$

and $\hat{S}_n(\cdot)$ is a consistent estimate of $S(\cdot)$, based on some preliminary estimate of θ_0 (see Velasco and Robinson (1997) for a review of covariance matrices estimates under autocorrelation of unknown form). That is,

A6. $\sup_{\tau\in[0,1]} \|\hat{S}_n(\tau) - S(\tau)\| = \mathrm{o_p}(1)$.

The discussion below is perfectly valid substituting $\hat{S}_n$ by any other bounded-in-probability *p.d* matrix of functions, which may not converge to S, but to any other *p.d*. matrix. The *GMM* estimator restricted to H_0 is $\tilde{\theta}_n =: \hat{\theta}_{1n}(1)$. For notational convenience, define $\bar{M}(\tau) := [M(1) - M(\tau)]$, $\bar{S}(\tau) := S(1) - S(\tau)$, $V(\tau) := M(\tau)'S(\tau)^{-1}M(\tau)$,

and $\bar{V}(\tau):=\bar{M}(\tau)'\bar{S}(\tau)^{-1}\bar{M}(\tau)$. It is reasonable to expect that the following first-order asymptotic approximation hold (see Andrews, 1993):

A7. $V(\cdot)$ and $\bar{V}(\cdot)$ are *p.d.* uniformly in $\tau \in [0,1]$. Then,

$$\|\tilde{\theta}_n - \theta_0 + V(1)^{-1}M(1)S(1)'^{-1}T_n^0(1)\| = \mathrm{o_p}(n^{-1/2})$$

and under H_{Cn},

$$\sup_{\tau\in[0,1]} \|\hat{\theta}_{1n}(\tau) - \theta_0 + V(\tau)^{-1}M(\tau)'S(\tau)^{-1}T_n^0(\tau)\| = \mathrm{o_p}(n^{-1/2}),$$

$$\sup_{\tau\in[0,1]} \|\hat{\theta}_{2n}(\tau) - \theta_0 + \bar{V}(\tau)^{-1}\bar{S}(\tau)^{-1}\bar{M}(\tau)'(T_n^0(1) - T_n^0(\tau))\| = \mathrm{o_p}(n^{-1/2}).$$

The asymptotic expansion of $\tilde{\theta}_n$ can be justified straightforwardly if A1 and A3 are satisfied, as well as some other regularity conditions guaranteeing that $\tilde{\theta}_n$ is a consistent estimator of θ_0 (see Hansen, 1982). The asymptotic expansions of $\hat{\theta}_{1n}(\cdot)$ and $\hat{\theta}_{2n}(\cdot)$ can also be easily justified under A1 and A3, providing that the PS-GMM are consistent estimators of θ_0 under H_{Cn} (see Andrews, 1993, Theorem 1 and A1). The asymptotic expansions of both $\hat{\theta}_{1n}(\cdot)$ and $\hat{\theta}_{2n}(\cdot)$ under $\mathrm{H}_{1n}(\tau_0)$ will depend on the partial sum process evaluated at some parameter values, which does not converges to zero.

W-type statistics are functionals of $(\hat{\theta}_{1n}(\cdot) - \hat{\theta}_{2n}(\cdot))$, suitably standardized. *LM*, *LR*, *CUSUM* and *P* statistics are based on certain functionals of standardized residual partial sums. Next proposition, which is an immediate consequence of the stated assumptions, will be helpful in the discussion of asymptotic representations.

Proposition 1. *Assume* A1, A3–A7. *Then*,

$$\sup_{\tau\in[0,1]} \|T_n(\tau,\tilde{\theta}_n) - T_n^0(\tau) + M(\tau)(\tilde{\theta}_n - \theta_0)\| = \mathrm{o_p}(n^{-1/2}),$$

and under H_{Cn},

$$\sup_{\tau_1,\tau_2\in[0,1]} \|T_n(\tau_1,\hat{\theta}_{1n}(\tau_2)) - T_n^0(\tau_1) + M(\tau_1)(\hat{\theta}_{1n}(\tau_2) - \theta_0)\| = \mathrm{o_p}(n^{-1/2}),$$

$$\sup_{\tau_1,\tau_2\in[0,1]} \|T_n(\tau_1,\hat{\theta}_{2n}(\tau_2)) - T_n^0(\tau_1) + M(\tau_1)(\hat{\theta}_{2n}(\tau_2) - \theta_0)\| = \mathrm{o_p}(n^{-1/2}).$$

3.1. CUSUM tests

These tests are particularly easy to compute and are designed for the linear regression model (see Brown et al. (1975) or, more recently, Ploberger and Krämer, 1992) though they can be employed in general situations. A *CUSUM-type* test has the form

$$\hat{\eta}_n^{\mathrm{CUSUM}} = \varphi\left(n^{1/2}\frac{\mathbf{i}'\tilde{T}_n}{(\mathbf{i}'\hat{S}_n(1)\mathbf{i})^{1/2}}\right),$$

where $\mathbf{i} = (1,0,\ldots,0)'$ and $\tilde{T}_n(\cdot) = T_n(\cdot,\tilde{\theta}_n)$. Next proposition provides the asymptotic representation of $\hat{\eta}_n^{CUSUM}$.

Proposition 2. *Assume* A1, A3–A7. *Then* $\hat{\eta}_n^{CUSUM}$ *satisfies* A2 (*also under* H_{1n}) *with* $\mathscr{L}$ *such that*

$$\mathscr{L} f(\tau) = \frac{\mathbf{i}' f(\tau) - \mathbf{i}' M(\tau) V(1)^{-1} M(1)' S(1)^{-1} f(1)}{(\mathbf{i}' S(1) \mathbf{i})^{1/2}}, \tag{8}$$

and if A3′ *also holds*,

$$\mathscr{L} f(\tau) = \frac{\mathbf{i}' f(\tau) - \mathbf{i}' \tau M(1) V(1)^{-1} M(1)' S(1)^{-1} f(1)}{(\mathbf{i}' S(1) \mathbf{i})^{1/2}}.$$

When the test does not use overidentified restrictions, i.e. $k = p$,

$$M(1) V(1)^{-1} M(1)' S(1)^{-1} = I_p,$$

and when A3′ and A4′ hold,

$$\mathscr{L} f(\tau) = \frac{\mathbf{i}' [f(\tau) - \tau f(1)]}{(\mathbf{i}' S(1) \mathbf{i})^{1/2}}.$$

Therefore, under H_0, A3′ and A4′,

$$\hat{\eta}_n^{CUSUM} \xrightarrow{\mathrm{d}} \varphi(B_1^0),$$

where $B_k^0(\tau) = B_k(\tau) - \tau B_k(1)$ is a $k \times 1$ vector of independent standard Brownian Bridges with time parameter $\tau \in [0, 1]$. Then, if $\varphi(f) = \sup_{\tau \in [0,1]} |f(\tau)|$, $\varphi(B_1^0)$ follows the Kolmogorov–Smirnov distribution, and if $\varphi(f) = \int_0^1 f(\tau)^2 \, \mathrm{d}\tau$, $\varphi(B_1^0)$ follows the Cramér–von Misses distribution. For these choices of φ functionals, the asymptotic expansion in Proposition 2 and Theorem 1 justifies that *CUSUM* tests have nontrivial power in the direction of H_{Cn}. Consistency in the direction of $\mathrm{H}_{1n}(\tau_0)$, for some unknown τ_0, is a trivial consequence of Proposition 3 and A3, since, in this case, T^0 is bounded away from zero on some $[0, 1]$ subset with nonzero Lebesgue measure.

Other tests related to *CUSUM*, also designed for the linear regression model and based on moving OLS estimates can be represented asymptotically, or exactly, as functionals of some linear transformation of $\{n^{1/2} T_n^0\}$ (see Kuan and Hornik (1995) for discussion).

3.2. Lagrange multiplier tests

These tests statistics, in a form related to that proposed by Andrews (1993), have the expression

$$\hat{\eta}_n^{LM} = \varphi(LM_n),$$

where $\varphi : (D[0, 1])^k \mapsto \mathbb{R}^+$ has some of the forms in (2), and

$$LM_n(\tau) = [I_k \vdots - I_k] \begin{bmatrix} n^{1/2} [\tilde{V}_{1n}(\tau)^{-1} + \tilde{V}_{2n}(\tau)^{-1}]^{-1/2} \tilde{V}_{1n}(\tau)^{-1} \partial Q_{1n}(\tau, \tilde{\theta}_n) / \partial \theta' \\ n^{1/2} [\tilde{V}_{1n}(\tau)^{-1} + \tilde{V}_{2n}(\tau)^{-1}]^{-1/2} \tilde{V}_{2n}(\tau)^{-1} \partial Q_{2n}(\tau, \tilde{\theta}_n) / \partial \theta' \end{bmatrix},$$

where

$$\tilde{M}_n(\cdot) =: \tilde{M}_n(\cdot, \tilde{\theta}_n), \quad \tilde{V}_{1n}(\cdot) =: \tilde{M}_n(\cdot)' \hat{S}_n(\cdot)^{-1} \tilde{M}_n(\cdot),$$

$$\tilde{V}_{2n}(\cdot) =: [\tilde{M}_n(1) - \tilde{M}_n(\cdot)]' [\hat{S}_n(1) - \hat{S}_n(\cdot)]^{-1} [\tilde{M}_n(1) - \tilde{M}_n(\cdot)].$$

This LM formulation, in contrast to Andrews (1993, Eq. (4.3)), only uses the restricted estimate of θ_0 everywhere, $\tilde{\theta}_n$, rather than the PS-GMM estimates. Finally, statistics are constructed after imposing A3′ and A4′ for estimating $M(\tau)$ and $S(\tau)$, and instead of $LM_n(\tau)$, the computationally more attractive

$$LM_n^0(\tau) = n^{1/2} \frac{1}{\tau^{1/2}(1-\tau)^{1/2}} \tilde{V}_{1n}(1)^{-1/2} \tilde{M}_n(1)' \hat{S}_n(1)^{-1} \tilde{T}_n(\tau)$$

is used. Note that $CUSUM$ and LM_n^0 tests are very similar. Next proposition provides an asymptotic representation of $\hat{\eta}_n^{LM}$.

Proposition 3. *Assume* A1, A3–A7. *Then,* $\hat{\eta}_n^{LM}$ *satisfies* A2 (*also under* H_{1n}) *with* $\mathscr{L}$ *such that*

$$\mathscr{L} f(\tau) = [V(\tau)^{-1} + \bar{V}(\tau)^{-1}]^{-1/2} [V(\tau)^{-1} M(\tau)' S(\tau)^{-1} + \bar{V}(\tau)^{-1} \bar{M}(\tau)' \bar{S}(\tau)^{-1}]$$
$$\times [f(\tau) - M(\tau) V(1)^{-1} M(1)' S(1)^{-1} f(1)].$$

If A3′ *and* A4′ *also hold,*

$$\mathscr{L} f(\tau) = \frac{1}{\tau^{1/2}(1-\tau)^{1/2}} V(1)^{-1/2} M(1)' S(1)^{-1} [f(\tau) - \tau f(1)]. \tag{9}$$

Thus, if A3′ and A4′ hold, under H_0,

$$n^{1/2} \mathscr{L} T_n^0 \Rightarrow B_k^*, \tag{10}$$

where $B_k^*(\tau) = [B_k(\tau) - \tau B_k(1)]/\tau^{1/2}(1-\tau)^{1/2}$, and the asymptotic distribution of $\hat{\eta}_n^{LM}$, based on some of the φ maps in (2) can be tabulated by Monte Carlo (see tables in Andrews 1993; Andrews and Ploberger, 1994). From this proposition, it is straightforward to show that, under the stated assumptions, $\hat{\eta}_n^{LM^0} := \varphi(LM_n^0)$ also satisfies A2 with

$$\mathscr{L} f(\tau) = \frac{1}{\tau^{1/2}(1-\tau)^{1/2}} V(1)^{-1/2} M(1)' S(1)^{-1}$$
$$\times [f(\tau) - M(\tau) V(1)^{-1} M(1)' S(1)^{-1} f(1)],$$

and if A3′ and A4′ also hold, the transformation is the same as in (9). Therefore, when A3′ and A4′ hold, $\hat{\eta}_n^{LM}$ and $\hat{\eta}_n^{LM^0}$, based on the same φ in (2) are asymptotically weakly equivalent.

Like the $CUSUM$ tests, LM tests based on some of the φ functionals in (2) can detect contiguous alternatives H_{Cn}, and they are consistent in the direction of $H_{1n}(\tau_0)$.

3.3. Wald tests

These tests, as formulated by Andrews (1993), have the expression

$$\hat{\eta}_n^W = \varphi(W_n),$$

where $\varphi : (D[0,1])^k \mapsto \mathbb{R}^+$ has some of the forms in (2) and

$$W_n(\tau) = n^{1/2}[\tilde{V}_{1n}(\tau)^{-1} + \tilde{V}_{2n}(\tau)^{-1}]^{-1/2}[\hat{\theta}_{1n}(\tau) - \hat{\theta}_{2n}(\tau)].$$

Next proposition states the asymptotic representation of these test statistics.

Proposition 4. *Assume* A1, A3–A7. *Then,* $\hat{\eta}_n^W$ *satisfies* A2 *with* $\mathscr{L}$ *such that*

$$\mathscr{L} f(\tau) = [V(\tau)^{-1} + \bar{V}(\tau)^{-1}]^{-1/2}\{\bar{V}(\tau)^{-1}\bar{M}(\tau)\bar{S}(\tau)^{-1} f(1) - [V(\tau)^{-1}M(\tau)'S(\tau)^{-1} + \bar{V}(\tau)^{-1}\bar{M}(\tau)'\bar{S}(\tau)^{-1}] f(\tau)\},$$

if A3′ *also holds,*

$$\mathscr{L} f(\tau) = -\frac{1}{\tau^{1/2}(1-\tau)^{1/2}} V(1)^{-1/2} M(1)' S(1)^{-1}[f(\tau) - \tau f(1)].$$

Therefore, this proposition and Theorem 1 justifies that W tests based on some of the φ functionals in (2) are able to detect contiguous alternatives H_{Cn}. Also, in view of (10), when A3′ and A4′ hold, LM and W tests computed with the same φ in (2) are asymptotically weakly equivalent.

3.4. Likelihood ratio tests

These tests, in Andrews (1993) formulation, are functionals of

$$LR_n(\tau) = n[Q_{1n}(\tau, \tilde{\theta}_n) - Q_{1n}(\tau, \hat{\theta}_{1n}(\tau))] + n[Q_{2n}(\tau, \tilde{\theta}_n) - Q_{2n}(\tau, \hat{\theta}_{2n}(\tau))].$$

Next proposition provides an useful asymptotic representation for LR_n.

Proposition 5. *Assume* A1, A3–A7. *Then,*

$$\sup_{\tau \in [0,1]} |LR_n(\tau) - \|LR_n^0(\tau)\|^2| = \mathrm{o_p}(1),$$

where

$$LR_n^0(\tau) = \begin{bmatrix} n^{1/2}\tilde{V}_{1n}(\tau)^{1/2}(\tilde{\theta}_n - \hat{\theta}_{1n}(\tau)) \\ n^{1/2}\tilde{V}_{2n}(\tau)^{1/2}(\tilde{\theta}_n - \hat{\theta}_{2n}(\tau)) \end{bmatrix}.$$

Therefore, likelihood ratio test statistics are asymptotically weak equivalent to test statistics of the form

$$\hat{\eta}_n^{LR^0} = \varphi(LR_n^0),$$

for some functional $\varphi : (D[0,1])^{2k} \mapsto \mathbb{R}^+$ in (2). Next proposition provides an asymptotic representation of $\hat{\eta}_n^{LR^0}$.

Proposition 6. *Assume* A1, A3–A7. *Then* $\hat{\eta}_n^{LR}$ *satisfies* A2, *with* $\mathscr{L}$ *such that*

$$\mathscr{L} f(\tau) = \begin{pmatrix} V(\tau)^{1/2}[V(\tau)^{-1}M(\tau)'S(\tau)^{-1}f(\tau) \\ -V(1)^{-1}M(1)'S(1)^{-1}f(1)] \\ \bar{V}(\tau)^{1/2}[[\bar{V}(\tau)^{-1}\bar{M}(\tau)'\bar{S}(\tau)^{-1} - V(1)^{-1}M(1)'S(1)^{-1}]f(1) \\ -\bar{V}(\tau)^{-1}\bar{M}(\tau)'\bar{S}(\tau)^{-1}f(\tau)] \end{pmatrix},$$

if A3′ *also holds*,

$$\mathscr{L} f(\tau) = \begin{pmatrix} \frac{1}{\tau^{1/2}} I_k \\ -\frac{1}{(1-\tau)^{1/2}} I_k \end{pmatrix} V(1)^{-1/2}M(1)'S(1)^{-1}[f(\tau) - \tau f(1)].$$

Therefore, this proposition and Theorem 1 justify that *LR* tests based on some of the φ functionals in (2) are able to detect contiguous alternatives H_{Cn}. Also, in view of (10), if A3′ holds, under H_0, *LR*, *W* and *LM* tests, based on the same φ functional in (2), are asymptotically weakly equivalent.

3.5. Predictive tests

These test statistics are designed to test

$$\bar{\mathrm{H}}_0 : \mathrm{E}[U(X_{ni}, \theta_0)] = 0 \quad \text{all } i \geqslant 1.$$

Then, the null hypothesis is formulated in terms of changes in the moment conditions, without attributing necessarily such changes to the parameter vector. These tests have nontrivial power in the direction of contiguous alternatives of the form

$$\bar{\mathrm{H}}_{Cn} : \mathrm{E}[U(X_{ni}, \theta_0)] = \frac{1}{\sqrt{n}} g\left(\frac{i}{n}\right) \quad \text{for some bounded function}$$

$$g : [0,1] \to \mathbb{R}^P \quad \text{for all } i \geqslant 1.$$

Predictive tests, proposed by Ghysels et al. (1997) are designed for detecting these contiguous alternatives. These tests have the form

$$\hat{\eta}_n^P = \varphi(P_n),$$

where $\varphi : (D[0,1])^k \mapsto \mathbb{R}^+$ has some of the forms in (2) and

$$P_n(\tau) = \frac{1}{1-\tau} n^{1/2} \hat{W}_n(\tau)^{-1/2}[T_n(1, \hat{\theta}_{1n}(\tau)) - T_n(\tau, \hat{\theta}_{1n}(\tau))],$$

where, $\sup_{\tau \in [0,1]} \|\hat{W}_n(\tau) - W(\tau)\| = \mathrm{o_p}(1)$, and $W(\tau) = [S(1) + M(1)V(1)^{-1}M(1)'(1-\tau)/\tau]$. The next proposition establishes the asymptotic representation of $\hat{\eta}_n^P$.

Proposition 7. *Assume* A1, A3–A7. *Then* $\hat{\eta}_n^P$ *satisfies* A2, *substituting* H_{Cn} *by* $\bar{\mathrm{H}}_{Cn}$, *with* $\mathscr{L}$ *such that*

$$\mathscr{L} f(\tau) = \frac{1}{1-\tau} W(\tau)^{-1/2}[f(1) - [I_k + \bar{M}(\tau)V(\tau)^{-1}M(\tau)'S(\tau)^{-1}]f(\tau)],$$

and if A3′ *also holds*,

$$\mathscr{L} f(\tau) = \frac{1}{1-\tau} W(1)^{-1/2}\left[f(1) - \left[I_k + \frac{1-\tau}{\tau} M(1)V(1)^{-1}M(1)'S(1)^{-1}\right] f(\tau)\right].$$

Therefore, this proposition and Theorem 1 justify that P tests based on some of the φ functionals in (2) are able to detect contiguous alternatives H_{Cn}. Consider the spectral decomposition,

$$S(1)^{-1/2}M(1)V(1)^{-1}M(1)'S(1)^{-1/2} =: C'\Lambda C$$

where $C'C = CC' = I_k$, and $\Lambda = diag(I_p|0_{k-p})$ with 0_m being an $m \times m$ squared matrix of zeroes. Hence, under A3′,

$$\mathscr{L} f(\tau) = \frac{1}{1-\tau} W(1)^{-1/2}S(1)^{1/2}C$$
$$\times \left[C'S(1)^{-1/2}f(1) - \left[C' + \frac{1-\tau}{\tau}\Lambda C\right] S(1)^{-1/2}f(\tau)\right],$$

and under H_0,

$$n^{1/2}\mathscr{L}T_n^0 \Rightarrow \frac{1}{1-\tau} W(1)^{-1/2}S(1)^{1/2}C\left[B_k(1) - B_k(\tau) - \frac{1-\tau}{\tau}\Lambda B_k(\tau)\right].$$

Noticing that

$$W(1) = \left[S(1) - \frac{(1-\tau)}{\tau} M(1)V(1)^{-1}M(1)'\right]$$
$$= S(1)^{1/2}C'\left[I_k + \frac{(1-\tau)}{\tau}\Lambda\right] CS(1)^{1/2},$$

and that

$$C'S(1)^{1/2}W(1)^{-1}S(1)^{1/2}C = \begin{bmatrix} \tau I_p & 0_{k-p} \\ 0_p & I_{k-p} \end{bmatrix};$$

if A3′ holds, under H_0,

$$\|n^{1/2}\mathscr{L}T_n^0\|^2 \Rightarrow \left\| \begin{matrix} [\tau B_p(1) - B_p(\tau)]'[\tau B_p(1) - B_p(\tau)]/\tau(1-\tau) \\ [B_{k-p}(1) - B_{k-p}(\tau)]'[B_{k-p}(1) - B_{k-p}(\tau)]/(1-\tau) \end{matrix} \right\|^2.$$

Thus, *P-type* tests based on some of the φ maps in (2) are asymptotically pivotal under H_0, when A3′ is satisfied (tables are provided by Ghysels et al., 1997).

3.6. Simpler tests

When critical values are estimated by bootstrap, it is not necessary to scale the tests statistics, since their limiting distribution under the null is always consistently estimated, whatever it could be. Then, it makes sense to use test statistics easier to compute than those presented in the previous subsections. For instance, a test statistic related to *CUSUM* and *LM* tests is

$$\hat{\eta}_n = \varphi(n^{1/2}\tilde{T}_n)$$

for which A2 holds, as a trivial consequence of Proposition 1, with

$$\mathscr{L} f(\tau) = f(\tau) - M(\tau)V(1)^{-1}M(1)'S(1)^{-1}f(1).$$

A test statistic, related to *Wald* tests is

$$\hat{\eta}_n = \varphi(n^{1/2}(\hat{\theta}_{1n} - \hat{\theta}_{2n})),$$

for which A2 holds also as a trivial consequence of Proposition 1, with,

$$\begin{aligned}\mathscr{L} f(\tau) = {} & [V(\tau)^{-1}M(\tau)'S(\tau)^{-1} + \bar{V}(\tau)^{-1}\bar{M}(\tau)'\bar{S}(\tau)^{-1}]f(\tau) \\ & - \bar{V}(\tau)^{-1}\bar{M}(\tau)'\bar{S}(\tau)^{-1}f(1).\end{aligned}$$

And a test statistics, related to *P-test* statistics is

$$\hat{\eta}_n = \varphi(n^{1/2}\dot{T}_n),$$

where $\dot{T}_n(\tau) = T_n(1, \hat{\theta}_{1n}(\tau)) - T_n(\tau, \hat{\theta}_{1n}(\tau))$, for which A2 holds, with H_{Cn} replaced by $\bar{\mathrm{H}}_{Cn}$, as a trivial consequence of Proposition 1, with

$$\mathscr{L} f(\tau) = f(1) - f(\tau) - \bar{M}(\tau)V(\tau)^{-1}M(\tau)'S(\tau)^{-1}f(\tau).$$

None of these simple tests are asymptotically pivotal, even when A3′ and A4′ hold, but they can be easily implemented applying the results in Section 4.

3.7. Some remarks on consistency

As it has been discussed in Subsections 3.1 and 3.2, consistency of *CUSUM* and *LM* tests is trivially justified, since asymptotic expansions of $\tilde{\theta}_n$, and, therefore of $\tilde{T}_n(\tau)$, are available both under H_{Cn} and $\mathrm{H}_{1n}(\tau_0)$. However, further assumptions are required for showing that *Wald* and *LR* tests are consistent in the direction of $\mathrm{H}_{1n}(\tau_0)$, since some knowledge on the limit behavior of $\hat{\theta}_{1n}(\cdot)$ and $\hat{\theta}_{2n}(\cdot)$ is needed. In particular, providing regularity conditions such that, under $\mathrm{H}_{1n}(\tau_0)$,

$$\sup_{\tau\in[0,1]} \|\hat{\theta}_{kn}(\tau) - \bar{\theta}_k(\tau)\| = \mathrm{o_p}(1), \quad k = 1,2 \quad \text{with } \bar{\theta}_1(\tau) \neq \bar{\theta}_2(\tau),$$

it is fairly straightforward to show that Wald and *LR* statistics detect alternatives $\mathrm{H}_{1n}(\tau_0)$ for φ functionals in (2).

4. Bootstrap tests

As we have discussed in the last section, $\mathscr{L}$ depends on $M(\cdot)$ and $S(\cdot)$, which are unknown, but they can be consistently estimated, uniformly in $\tau \in [0,1]$, by $\tilde{M}_n(\cdot)$ and $\hat{S}_n(\cdot)$. Suppose that there exists a feasible transformation $\mathscr{L}_n$, which depends on observed data $\mathscr{Z}_n$, such that,

A8. $\sup_{\tau\in[0,1]} \|(\mathscr{L}_n - \mathscr{L})f(\tau)\| \leqslant \mathrm{o_p}(1) \cdot \sup_{\tau\in[0,1]} \|f(\tau)\|$.

This assumption is very reasonable for all transformations discussed in the last section. For instance, for *CUSUM* tests, a feasible version of $\mathscr{L}$ in (8) is $\mathscr{L}_n$, where

$$\mathscr{L}_n f(\tau) = \frac{\mathbf{i}' f(\tau) - \mathbf{i}'\tilde{M}_n(\tau)\tilde{V}_{1n}(1)^{-1}\tilde{M}_n(1)'\hat{S}_n(1)f(1)}{(\mathbf{i}'\hat{S}_n(1)\mathbf{i}^{1/2})},$$

and $\sup_{\tau\in[0,1]} |(\mathscr{L}_n - \mathscr{L})f(\tau)|$ is equal to

$$\begin{aligned}
&\sup_{\tau\in[0,1]} \left| \mathbf{i}' \left[\frac{\tilde{M}_n(\tau)\tilde{V}_{1n}(1)^{-1}\tilde{M}_n(1)'\hat{S}_n(1)}{(\mathbf{i}'\hat{S}_n(1)\mathbf{i})^{1/2}} - \frac{M(\tau)V(1)^{-1}M(1)'S(1)}{(\mathbf{i}'S(1)\mathbf{i})^{1/2}} \right] f(1) \right| \\
&\quad \leqslant \sup_{\tau\in[0,1]} \left\| \frac{\tilde{M}_n(\tau)'\tilde{V}_{1n}(1)^{-1}\tilde{M}_n(1)'\hat{S}_n(1)}{(\mathbf{i}'\hat{S}_n(1)\mathbf{i})^{1/2}} - \frac{M(\tau)'V(1)^{-1}M(1)'S(1)}{(\mathbf{i}'S(1)\mathbf{i})^{1/2}} \right\| \\
&\qquad \times \sup_{\tau\in[0,1]} \|f(\tau)\| \\
&\quad = \mathrm{o_p}(1) \cdot \sup_{\tau\in[0,1]} \|f(\tau)\|
\end{aligned}$$

by A3 and A6. Similar feasible transformations, holding A8, can be obtained for all test statistics discussed in the last section by just substituting $M(\cdot)$ by $\tilde{M}_n(\cdot)$ and $S(\cdot)$ by $\hat{S}_n(\cdot)$ in their asymptotic expansions (as indicated in Propositions 3–7).

The bootstrap version of $\tilde{\eta}_n = \varphi(n^{1/2}\mathscr{L}T_n^0)$ has the form

$$\tilde{\eta}_n^* := \varphi(n^{1/2}\mathscr{L}_n\tilde{T}_n^*),$$

where $\tilde{T}_n^*(\cdot) := n^{-1}\sum_{i=1}^{[n\cdot]} \tilde{U}_{ni}^*$ and $\{\tilde{U}_{ni}^*, i = 1,\ldots,n\}$ is a resample of $\{\tilde{U}_{ni}, i = 1,\ldots,n\}$, where $\tilde{U}_{ni} := U(Z_{ni}, \tilde{\theta}_n)$. Therefore, the resampling scheme must be such that, for any distance d metricizing weak convergence on the real line to $\eta_\infty := \varphi(\mathscr{L}N_k^0)$, with $\{N_k^0\}$ defined in Theorem 1,

$$d(F_{\tilde{\eta}_n^*}^*, F_{\eta_\infty}) = \mathrm{o_p}(1),$$

where, henceforth, $F_\xi^*(\cdot) = \Pr[\xi \leqslant \cdot\,|\mathscr{Z}_n]$ is the conditional distribution of ξ, given the sample $\mathscr{Z}_n$, and $F_\xi(\cdot) = \Pr(\xi \leqslant \cdot)$ is the marginal distribution of ξ. When F_{η_∞} is continuous, as in our case, a suitable distance is the uniform, by Polya's theorem. Thus, the resampling scheme must be such that

$$\sup_{t\in\mathbb{R}^+} |F_{\tilde{\eta}_n^*}^*(t) - F_{\eta_\infty}(t)| = \mathrm{o_p}(1), \tag{11}$$

which suggests to use as critical values

$$c^*_{n\alpha}:=\inf\{t\colon F^*_{\hat{\eta}^*_n}(t) \geqslant 1-\alpha\}.$$

The null hypothesis is rejected when $\hat{\eta}_n > c^*_{n\alpha}$. Applying Theorem 1 and the Continuous Mapping Theorem, under the null hypothesis,

$$\sup_{t\in\mathbb{R}^+} |F_{\hat{\eta}_n}(t) - F_{\eta_\infty}(t)| = \mathrm{o}(1), \tag{12}$$

and, hence, (11) and (12) imply that under the null,

$$F^*_{\hat{\eta}_n}(c^*_{n\alpha}) = 1-\alpha+\mathrm{o_p}(1) \tag{13}$$

and also,

$$\lim_{n\to\infty} \Pr[\hat{\eta}_n > c^*_{n\alpha}] = \alpha.$$

Under contiguous alternatives,

$$\lim_{n\to\infty} \Pr[\hat{\eta}_n > c^*_{n\alpha}] \geqslant \alpha \tag{14}$$

for suitable φ functionals (e.g. those in (2) for LM, W, LR or P tests, or those provided in (1) for $CUSUM$ tests). And, if the test based on $\hat{\eta}_n$ is consistent, (11) justifies that, under the fixed alternative,

$$\lim_{n\to\infty} \Pr[\hat{\eta}_n > c^*_{n\alpha}] = 1. \tag{15}$$

Therefore, (14) and (15) show that tests based on statistics satisfying A2, with critical values $c^*_{n\alpha}$ are asymptotically admissible, and able to detect contiguous alternatives converging to the null at the parametric rate $n^{-1/2}$.

The resampling scheme must be such that (11) is satisfied, mimicking key features of the underlying data generating process. Not all the resampling procedures are valid. Suppose, for the moment, that

$$\sup_{t\in\mathbb{R}^+} |F^*_{\hat{\eta}^*_n}(t) - F^*_{\tilde{\eta}^{0*}_n}(t)| = \mathrm{o_p}(1), \tag{16}$$

where $\tilde{\eta}^{0*}_n = \varphi(n^{1/2}\mathscr{L}T^{0*}_n)$, $T^{0*}_n(\cdot):=n^{-1}\sum_{i=1}^{[n\cdot]} U^{0*}_{ni}$ and $\{U^{0*}_{ni}, i=1,\ldots,n\}$ is a resample from $\{U^0_{ni}, i=1,\ldots,n\}$, with $U^0_{ni}:=U(Z_{ni},\theta_0)$. Then, (11) is implied by (16) and

$$\sup_{t\in\mathbb{R}^+} |F^*_{\tilde{\eta}^{0*}_n}(t) - F_{\eta_\infty}(t)| = \mathrm{o_p}(1). \tag{17}$$

While (16) is satisfied for most resampling schemes, (17) does not follow under A3 and A4, at least the nonstationarity of $\{U^0_{ni}, i=1,\ldots,n\}$ is taken into account. Two necessary conditions for (17) are

$$\mathrm{E}[T^{0*}_n(\tau) \mid \mathscr{L}_n] = \mathrm{o_p}(1), \tag{18}$$

$$\mathrm{E}[T^{0*}_n(\tau)T^{0*}_n(\tau)' \mid \mathscr{L}_n] = S(\tau) + \mathrm{o_p}(1). \tag{19}$$

For instance, suppose that the U^0_{ni} are as in cases I and II in Section 2. Then, a näive bootstrap consists of a random sample with replacement $\{\hat{U}^*_{ni}, i=1,\ldots,n\}$ from

$\{\hat{U}_{ni}, i=1,\ldots,n\}$, and, hence, $\{U_{ni}^{0*}, i=1,\ldots,n\}$ is a random sample with replacement from $\{U_{ni}^{0}, i=1,\ldots,n\}$, which satisfies (18), but, for each τ,

$$\mathrm{E}[T_n^{0*}(\tau)T_n^{0*}(\tau)' \mid \mathscr{Z}_n] = \tau S(1) + \mathrm{o_p}(1),$$

and (19) does not hold under nonstationarity. The same problem appears by applying the block stationary bootstrap, where the sample is divided into blocks, which are sampled at random with replacement. The blocks, whose size must increase with the sample size, can be nonoverlapping (Carlstein, 1986) or overlapping (Künsch, 1989). These bootstrap procedures are designed to preserve the unknown autocorrelation structure in a stationary context, where asymptotic tests are available. This type of resample can also be applied to the whole data set $\mathscr{Z}_n$, the resulting $\hat{\eta}_n$ bootstrap analog satisfies (11) under A3′ (see Hall and Horowitz, 1996) but it may not be valid under A3 or A4. Subsampling (see Politis et al., 1999, monograph) can be applied in nonstationary situations (e.g. Romano and Wolf, 2001). However, this procedure does not impose the restriction under the null and, as a result, the subsampling test statistics behave like the original one under the null and alternative. Therefore, subsampling critical values estimates are consistent under the null hypothesis, but they diverge to infinity under the alternative, and the resulting tests will be inconsistent.

We propose to use an external bootstrap procedure, where the resample is obtained from an external distribution, rather than the empirical distribution of the data set. That is, $\hat{U}_{ni}^{*} = \hat{U}_{ni}V_i$ and $U_{ni}^{0*} = U_{ni}^{0}V_i$, where

A9. V_i are scalar *i.i.d.* random variables, bounded, independent of $\mathscr{Z}_n$, and such that $\mathrm{E}(V_i) = 0$ and $\mathrm{E}(V_i^2) = 1$.

This resampling method, known as "wild" or "external" bootstrap, was suggested by Wu (1986) in the context of fixed-design regression models with heteroskedastic errors, in order to match, asymptotically, the first and second moments of the statistic distribution under the null hypothesis. Su and Wei (1991) suggested to apply this resampling procedure to asymptotic linearizations of empirical processes, which are a basis for constructing goodness-of-fit test statistics. It has been also applied in other testing problems by Hansen (1996, 2000), Härdle and Mammen (1991), de Jong (1996), Stute et al. (1998), and Delgado and González-Manteiga (2001), amongst others.

Using this resample scheme,

$$\mathrm{E}[T_n^{0*}(\tau)T_n^{0*}(\tau)' \mid \mathscr{Z}_n] = \frac{1}{n}\sum_{i=1}^{[n\tau]} U_{ni}^{0}U_{ni}^{0'} \quad \text{for each } \tau \in [0,1],$$

and (11) is satisfied assuming

A10.

$$\frac{1}{n}\sum_{i=1}^{[n\tau]} U_{ni}^{0}U_{ni}^{0'} - S(\tau) = \mathrm{o}(1) \quad \text{with probability 1 for each } \tau \in [0,1], \tag{20}$$

$$\frac{1}{n}\sum_{i=1}^{n} \|\tilde{U}_{ni} - U_{ni}^{0}\|^2 = \mathrm{o_p}(1). \tag{21}$$

Condition (20) is satisfied even when $\{U_{ni}^0 U_{ni}^{0'}, i \geqslant 1\}$ is a nonstationary mixingale (see MacLeish, 1975a) and (21) is reasonable in view that $\tilde{\theta}_n$ is consistent, (A7) and A3. Assumption A10 suggests to use,

$$\hat{S}_n(\cdot) = \frac{1}{n} \sum_{i=1}^{[n\cdot]} U_{ni}(\tilde{\theta}_n) U_{ni}(\tilde{\theta}_n)'. \tag{22}$$

However, one may wish to use a bootstrap version of $\hat{S}_n$, for instance,

$$\tilde{S}_n^*(\cdot) = \frac{1}{n} \sum_{i=1}^{[n\cdot]} \tilde{U}_{ni}^* \tilde{U}_{ni}^{*'}.$$

The bootstrap test will be based on the bootstrap version of $\mathscr{L}_n$, say $\mathscr{L}_n^*$, where $\hat{S}_n$ in $\mathscr{L}_n$ is substituted by $\tilde{S}_n^*$. Then, the bootstrap statistic is $\tilde{\eta}_n^{**} := \varphi(n^{1/2} \mathscr{L}_n^* \tilde{T}_n^*)$. The consistency of $\tilde{\eta}_n^{**}$ is discussed in Section 5 remark.

The next, condition is required for guaranteeing that a bootstrap Lindeberg's condition is satisfied:

A11. For all $\delta > 0$,

$$\frac{1}{n} \sum_{i=1}^{n} \|U_{ni}^0\|^2 1(\|U_{ni}^0\| > n^{1/2}\delta) \to 0 \quad \text{with probability 1.}$$

A sufficient condition for A11 is that there exists an $\varepsilon > 0$ such that $n^{-1} \sum_{i=1}^{n} \|U_{ni}^0\|^{2+\varepsilon} = \mathrm{O}(1)$ a.s. It does not seem too restrictive in view of A4 and A10. Since V_i are independent, our resampling does not permit autocorrelation of the $U_{ni}^{0'}s$, but it does allow for nonstationarity. Next theorem justifies to use the resulting bootstrap test in practice.

Theorem 2. *Assuming* A1–A11, *the external bootstrap statistic* $\tilde{\eta}_n^*$ *satisfies* (11).

A popular choice of $V_i's$ is

$$V_i = \begin{cases} -(\sqrt{5}-1)/2 & \text{with probability } (\sqrt{5}+1)/2\sqrt{5}, \\ (\sqrt{5}+1)/2 & \text{with probability } (\sqrt{5}-1)/2\sqrt{5}, \end{cases} \tag{23}$$

which has been proposed by Liu (1988) for mimicking the third-order moment structure of regression errors by their bootstrap analogs. Other authors have proposed to take V_i standard normal (e.g. Su and Wei, 1991; Hansen, 2000), which seems very reasonable, since the $\{n^{1/2} T_n^0\}$ limiting distribution is Gaussian. The bootstrap test is also valid under this alternative hypothesis.

A9′. V_i are scalar i.i.d. standard normal random variables, independent of $\mathscr{Z}_n$.

Corollary 1. *Assuming* A1–A8, A9′, *and* A10, *the external bootstrap statistic* $\tilde{\eta}_n^*$ *satisfies* (11).

In practice, $c^*_{n\alpha}$ is difficult to compute exactly, but it can be approximated, as accurately as desired, by Monte Carlo (see Beran et al. (1987) for a formal justification of this statement). That is, we generate independent random numbers $\{V_i^b, i = 1, \ldots, n, b = 1, \ldots, B\}$. Then, instead of $c^*_{n\alpha}$, we use,

$$c^{*B}_{n\alpha} := \inf\left\{ t\colon \frac{1}{B}\sum_{b=1}^{B} \tilde{\eta}^{*b}_n \geqslant 1 - \alpha \right\}, \tag{24}$$

where

$$\tilde{\eta}^{*b}_n := \varphi(n^{1/2}\mathscr{L}_n \tilde{T}^{*b}_n)$$

and

$$\tilde{T}^{*b}_n(\cdot) := \frac{1}{n}\sum_{i=}^{[n\cdot]} \tilde{U}_{ni} V_i^b.$$

The greater the B, the better the Monte Carlo approximation.

Thus, the following steps must be followed in practice for implementing bootstrap tests.

1. Choose a test statistic $\hat{\eta}_n$ from those discussed in Section 3 or any other allowing the representation in A2.
2. Compute the feasible transformation $\mathscr{L}_n$ for the chosen statistic, substituting S by $\hat{S}_n$, and M by $\tilde{M}_n$.
3. Generate, independently, random numbers $\{V_i^b, i = 1, \ldots, n, b = 1, \ldots, B\}$ from a distribution satisfying A9 or A9′, e.g. like (23) or standard normals.
4. Compute $c^{*B}_{n\alpha}$ in (24).
5. Reject the null hypothesis when $\hat{\eta}_n > c^{*B}_{n\alpha}$.

5. Some remarks on linear models

Linear models deserve a detailed discussion. For these models, $U_{ni}(\theta) = \xi_{ni}(\theta) W_{ni}$, where

$$\xi_{ni}(\theta) := Y_{ni} - X'_{ni}\theta$$

with $Z_{ni} = (Y_{ni}, X'_{ni}, W'_{ni})'$ and $Y'_{ni}s$ being scalar. The GMM, or instrumental variables estimator, has the closed form

$$\tilde{\theta}_n = [\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\tilde{M}_n(1)]^{-1}\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\frac{1}{n}\sum_{i=1}^{n} W_{ni}Y_{ni},$$

where now, $\tilde{M}_n(\cdot) := n^{-1}\sum_{i=1}^{[n\cdot]} W_{ni}X'_{ni}$ does not depend on parameters. Then, exactly, for each n,

$$\tilde{T}_n(\tau) = T^0_n(\tau) - [\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\tilde{M}_n(1)]^{-1}\tilde{M}_n(1)'\hat{S}_n(1)^{-1}T^0_n(1),$$

which has an external bootstrap version,

$$\bar{T}_n^*(\tau) = \tilde{T}_n^*(\tau) - [\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\tilde{M}_n(1)]^{-1}\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\bar{T}_n^*(1)$$

$$= \frac{1}{n}\sum_{i=1}^{n} W_{ni}[Y_{ni}^* - X_{ni}'\tilde{\theta}_n^*],$$

where $\{Y_{ni}^* = X_{ni}'\tilde{\theta}_n + \xi_{ni}(\tilde{\theta}_n)V_i, i = 1, \ldots, n\}$ is the external bootstrap resample of $\{Y_{ni}, i = 1, \ldots, n\}$, and

$$\tilde{\theta}_n^* = [\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\tilde{M}_n(1)]^{-1}\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\frac{1}{n}\sum_{i=1}^{n} W_{ni}Y_{ni}^*$$

$$= \tilde{\theta}_n + [\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\tilde{M}_n(1)]^{-1}\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\frac{1}{n}\sum_{i=1}^{n}\tilde{U}_{ni}^*$$

is the external bootstrap analog of $\tilde{\theta}_n$. That is, $\tilde{T}_n^*$ is the bootstrap analog of $\tilde{T}_n$ computed with the resample $\{(Y_{ni}^*, X_{ni}', W_{ni}')', i = 1, \ldots, n\}$. $\hat{S}_n$ can be substituted by $\tilde{S}_n^*$ if desired without affecting the asymptotic properties of the resulting bootstrap tests. Therefore, our bootstrap version of test statistics based on functionals of $\{n^{1/2}\tilde{T}_n\}$, like *CUSUM* and *LM*, can be expressed as the same functional of the external bootstrap analog of $\{n^{1/2}\tilde{T}_n\}$. For the *CUSUM* test

$$\mathscr{L}_n\tilde{T}_n^*(\tau) = \frac{\mathbf{i}'\tilde{T}_n^*(\tau) - \mathbf{i}'[\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\tilde{M}_n(1)]^{-1}\tilde{M}_n(1)'\hat{S}_n(1)^{-1}\tilde{T}_n^*(1)}{(\mathbf{i}'\hat{S}_n(1)\mathbf{i})^{1/2}}$$

$$= \frac{\mathbf{i}'\bar{T}_n^*(\tau)}{(\mathbf{i}'\hat{S}_n(1)\mathbf{i})^{1/2}},$$

and a similar expression can be obtained for the test based on LM_n^0. Our bootstrap statistics are also identical to the corresponding external bootstrap analog for other testing principles. Hansen (2000), in the context of the linear regression model (i.e. $W_{ni} = X_{ni}$), discusses a *Wald* tests statistic, proposing its bootstrap analog for approximating critical values, which produce the same bootstrap statistic than our method, (see Hansen, 2000, proof of Theorem 5, p. 114). That is, for linear models one can compute the bootstrap statistics as the external bootstrap analogs of the original test statistic, which is more attractive from a computational viewpoint.

6. Monte Carlo experiments

Some Monte Carlo experiments have been carried out for studying the level accuracy of bootstrap tests in the context of the Box–Cox model:

$$\xi_{ni}(\theta) = \frac{Y_{ni}^{\lambda} - 1}{\lambda} - \alpha - \beta X_{ni}, \quad i \geqslant 1 \tag{25}$$

with $\theta = (\lambda, \alpha, \beta)'$ and $\lambda \in (0, 1]$. Then,

$$U_{ni}(\theta) = \xi_{ni}(\theta) W_{ni}, \quad i \geqslant 1,$$

where the $W_{ni}'s$ is a vector of instruments.

Model (25) has the reduced form equation

$$Y_{ni} = [1 + \lambda_0(\xi_{ni} + \alpha_0 + \beta_0 X_{ni})]^{1/\lambda_0}$$

with $\xi_{ni} := \xi_{ni}(\theta_0)$. For a given choice of $\theta_0 = (\lambda_0, \alpha_0, \beta_0)'$, we generate $\{Y_{ni}, i = 1, \ldots, n\}$ observations for each of the following ξ_{ni} and X_{ni} designs. Let $\{(\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i}), i = 1, \ldots, n\}$ be vectors of *i.i.d.* variables, where the components of $(\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i})$ are mutually independent. Then, the $\xi_{ni}'s$ are generated according to the following designs:

- Homoskedastic: $\xi_{ni} = \varepsilon_{1i}$.
- Heteroskedastic: $\xi_{ni} = \sigma(X_{ni}) \cdot \varepsilon_{1i}$; $\sigma(X_{ni}) = \sqrt{0.1}|X_{ni}|$.

And the $X_{ni}'s$ have been generated according to the following designs:

- i.i.d.: $X_{ni} = \varepsilon_{2i}$.
- Mean break: $X_{ni} = \varepsilon_{2i} + 5 \cdot 1(t > [n/2])$.
- Variance break: $X_{ni} = \varepsilon_{2i} \cdot 1(t \leqslant [n/2]) + \varepsilon_{3i} \cdot 5 \cdot 1(t > [n/2])$.
- Deterministic trend: $X_{ni} = t/n$.

We run simulations under H_0 with the parameter choice $\theta_0 = (0.5, 2, 2)'$. In all the experiments, we report the proportion of rejections using asymptotic critical values, valid under A3′ and A4′, and their bootstrap counterparts, which are valid under more general conditions.

We consider two cases. First, the nonlinear-in-variables model, where λ_0 is unknown. We choose $W_{ni} = (1,\ X_{ni},\ X_{ni}^2)'$, which has also been chosen by Amemiya and Powell (1981) in simulations of the same model. Also, Robinson (1991) used these instruments in simulations for the *arcsinh* transformation model. Second, we consider a linear model, where λ_0 is known, the model becomes linear and $(\alpha_0, \beta_0)'$ is estimated by OLS, i.e. taking $W_{ni} = (1,\ X_{ni})'$.

Tables report the proportion of rejections in 5000 Monte Carlo experiments. At each experiment, bootstrap critical values are approximated using $B = 1000$ resamples, as explained in Section 4, with $V_i's$ generated as (23), and satisfying condition A9.

Simulations for the nonlinear models are computationally costly. This is why we only report, in this case, results for the simpler version of the LM test (LM^0), which only needs to compute parameter estimates under the null hypothesis. In the linear case, we also report results for the $CUSUM$ and LR.

Table 1 provides the proportion of rejections under H_0 for the nonlinear model using the LM tests. We used the statistic with $\varphi(f) = \sup_{\tau \in [-\pi, 1-\pi]} |f(\tau)' f(\tau)|$ and $\pi = 0.1, 0.3$. The critical values for the asymptotic test, valid under A3′ and A4′, can be found in Andrews (1993). Bootstrap tests exhibit an excellent level accuracy in all cases considered. On the contrary, there are observed quite large size distortions for the "asymptotic" test when $Z_{ni}'s$ are nonstationary with hetoroskedastic $\xi_{ni}'s$, and

Table 1
Lagrange multiplier statistic for the Box–Cox model

	Regressor process							
	i.i.d.		Mean break		Variance break		Trend	
	$n = 100$	$n = 200$	$n = 100$	$n = 200$	$n = 100$	$n = 200$	$n = 100$	$n = 200$
($\Pi = [0.1n; 0.9n]$) Homocedastic error								
Asymptotic								
$\alpha = 0.1$	0.059	0.072	0.077	0.091	0.111	0.133	0.108	0.098
$\alpha = 0.05$	0.028	0.039	0.043	0.052	0.066	0.080	0.072	0.060
Bootstrap								
$\alpha = 0.1$	0.074	0.089	0.095	0.100	0.064	0.083	0.107	0.103
$\alpha = 0.05$	0.028	0.041	0.043	0.047	0.024	0.034	0.046	0.046
Heterocedastic error								
Asymptotic								
$\alpha = 0.1$	0.051	0.071	0.105	0.125	0.374	0.427	0.252	0.278
$\alpha = 0.05$	0.023	0.033	0.058	0.069	0.292	0.349	0.202	0.226
Bootstrap								
$\alpha = 0.1$	0.086	0.092	0.088	0.090	0.092	0.108	0.107	0.103
$\alpha = 0.05$	0.035	0.046	0.041	0.041	0.046	0.054	0.056	0.051
($\Pi = [0.3n; 0.7n]$) Homocedastic error								
Asymptotic								
$\alpha = 0.1$	0.056	0.073	0.023	0.031	0.062	0058	0.013	0.008
$\alpha = 0.05$	0.022	0.033	0.009	0.013	0.036	0.029	0.007	0.004
Bootstrap								
$\alpha = 0.1$	0.107	0.106	0.105	0.108	0.080	0.090	0.102	0.102
$\alpha = 0.05$	0.049	0.056	0.052	0.055	0.036	0.044	0.053	0.052
Heterocedastic error								
Asymptotic								
$\alpha = 0.1$	0.044	0.057	0.069	0.086	0.345	0.397	0.056	0.066
$\alpha = 0.05$	0.017	0.024	0.036	0.059	0.279	0.329	0.031	0.037
Bootstrap								
$\alpha = 0.1$	0.101	0.110	0.107	0.107	0.115	0.119	0.109	0.101
$\alpha = 0.05$	0.048	0.050	0.054	0.054	0.068	0.067	0.056	0.050

taking $\pi = 0.1$. However, sensible empirical sizes are observed in the homoskedastic case, though some bias is appreciated for the variance break model. When we take $\pi = 0.3$, the "asymptotic" tests work very poorly in all cases, but the bootstrap tests are unaffected by the change of π. The size bias of "asymptotic" tests depends on the value of π, as Hansen (2000) in Section 4 discusses for the *Wald* test, in the context of a single linear regression model, under homoskedastic errors and the regressor generated according to the Mean Break model.

Table 2 compares the proportion of rejections for *CUSUM*, *LR* and *LM* statistics for the lineal model (i.e. knowing that $\lambda_0 = 0.5$). The *CUSUM* test in based on the popular choice $\varphi(f) = \sup_{\tau \in [0,1]} |f(\tau)|$, which has the Kolmogorov–Smirnov's limiting distribution when A3′ and A4′ are satisfied. *LM* and *LR* statistics are computed with $\pi = 0.1$ and φ as in Table 1. The bootstrap test exhibit an excellent level accuracy for

Table 2
LM, *CUSUM* and *RV* statistics for the linear regression model

	Regressor process							
	i.i.d.		Mean break		Variance break		Trend	
	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$	$n = 50$	$n = 100$
($\hat{\eta}_n^{LM}$) Homocedastic error								
Asymptotic								
$\alpha = 0.1$	0.052	0.068	0.072	0.089	0.063	0.082	0.132	0.151
$\alpha = 0.05$	0.020	0.028	0.033	0.051	0.028	0.044	0.078	0.097
Bootstrap								
$\alpha = 0.1$	0.083	0.095	0.101	0.109	0.093	0.094	0.111	0.104
$\alpha = 0.05$	0.029	0.040	0.045	0.050	0.035	0.045	0.050	0.052
Heterocedastic error								
Asymptotic								
$\alpha = 0.1$	0.055	0.064	0.105	0.126	0.121	0.159	0.197	0.221
$\alpha = 0.05$	0.025	0.031	0.050	0.070	0.069	0.095	0.133	0.156
Bootstrap								
$\alpha = 0.1$	0.065	0.073	0.084	0.091	0.083	0.099	0.109	0.109
$\alpha = 0.05$	0.020	0.029	0.032	0.040	0.023	0.036	0.050	0.054
($\hat{\eta}_n^{RV}$) Homocedastic error								
Asymptotic								
$\alpha = 0.1$	0.096	0.085	0.115	0.119	0.099	0.094	0.131	0.126
$\alpha = 0.05$	0.054	0.041	0.065	0.059	0.058	0.050	0.075	0.065
Bootstrap								
$\alpha = 0.1$	0.137	0.120	0.118	0.113	0.126	0.115	0.123	0.111
$\alpha = 0.05$	0.072	0.057	0.059	0.057	0.067	0.058	0.059	0.052
Heterocedastic error								
Asymptotic								
$\alpha = 0.1$	0.434	0.481	0.408	0.413	0.687	0.741	0.417	0.421
$\alpha = 0.05$	0.341	0.377	0.321	0.313	0.609	0.656	0.328	0.320
Bootstrap								
$\alpha = 0.1$	0.182	0.139	0.163	0.132	0.259	0.165	0.139	0.119
$\alpha = 0.05$	0.109	0.066	0.091	0.066	0.159	0.088	0.059	0.053
($\hat{\eta}_n^{CUSUM}$) Homocedastic error								
Asymptotic								
$\alpha = 0.1$	0.067	0.074	0.005	0.010	0.069	0.073	0.000	0.000
$\alpha = 0.05$	0.029	0.036	0.001	0.001	0.029	0.034	0.000	0.000
Bootstrap								
$\alpha = 0.1$	0.109	0.111	0.105	0.103	0.112	0.109	0.098	0.107
$\alpha = 0.05$	0.057	0.055	0.046	0.050	0.058	0.053	0.052	0.052
Heterocedastic error								
Asymptotic								
$\alpha = 0.1$	0.109	0.069	0.046	0.067	0.102	0.136	0.002	0.009
$\alpha = 0.05$	0.057	0.032	0.018	0.025	0.050	0.067	0.000	0.002
Bootstrap								
$\alpha = 0.1$	0.052	0.113	0.107	0.110	0.127	0.125	0.117	0.119
$\alpha = 0.05$	0.019	0.058	0.050	0.057	0.065	0.064	0.057	0.058

all test statistics, as expected, but the "asymptotic" test observe large size distortions for all statistics considered under nonstationary designs.

7. Proofs

Proofs of Theorem 1 and Propositions 1–6 are immediate applying the stated conditions and a standard mean value theorem (MVT) argument. We just prove Theorem 2 and Corollary 1.

Let us introduce the following notation. Let $\{\xi_n, n \geqslant 1\}$ be a sequence of random variables. We say that $\xi_n = \mathrm{o}_{\mathrm{p}^*}(1)$ in probability (with probability 1) if for any constant $\delta > 0$,

$$\Pr\{\|\xi_n\| > \delta \mid \mathscr{Z}_n\} \xrightarrow{\mathrm{p}} 0 \ (\to 0 \text{ with probability } 1).$$

Proof of Theorem 1. First, we show that

$$n^{1/2} \sup_{\tau \in [0,1]} \|\tilde{T}_n^*(\tau) - T_n^{0*}(\tau)\| = \mathrm{o}_{\mathrm{p}^*}(1). \tag{26}$$

Note that

$$\tilde{T}_n^*(\tau) - T_n^{0*}(\tau) = \frac{1}{n} \sum_{i=1}^{[n\tau]} [U_{ni}(\tilde{\theta}_n) - U_{ni}(\theta_0)] V_i.$$

Since conditional on the sample, $\mathscr{Z}_n$, $[U_{ni}(\tilde{\theta}_n) - U_{ni}(\theta_0)]V_i$ are independent, an application of Doob's inequality establishes that

$$\begin{aligned} \mathrm{E}\left[\left(n^{1/2} \sup_{\tau \in [0,1]} \|\tilde{T}_n^*(\tau) - T_n^{0*}(\tau)\|^2\right) \middle| \mathscr{Z}_n\right] &\leqslant 2\mathrm{E}[\|n^{1/2}[\tilde{T}_n^*(1) - T_n^{0*}(1)]\|^2 | \mathscr{Z}_n] \\ &= 2\frac{1}{n} \sum_{i=1}^{n} \|U_{ni}(\tilde{\theta}_n) - U_{ni}(\theta_0)\|^2 \\ &= \mathrm{o}_{\mathrm{p}}(1) \end{aligned}$$

by A10, which proves (26), after applying Markov's inequality. Next, Lemmata 1–3 below show that conditional on $\mathscr{Z}_n$,

$$n^{1/2} T_n^{0*} \Rightarrow N_k^0 \quad \text{with probability 1.} \tag{27}$$

We use the strategy of proof in Stute et al. (1998). Lemma 1 shows that the covariance function, given $\mathscr{Z}_n$, of $\{n^{1/2} T_n^{0*}\}$ converges to the covariance function of $\{N_k^0\}$ a.s. Lemma 2 shows that the finite-dimensional distributions (fidis) of $\{n^{1/2} T_n^{0*}\}$, conditional on $\mathscr{Z}_n$, converge to the corresponding distributions of $\{N_k^0\}$ a.s. (i.e. in view of Lemma 1, it suffices to show that the fidis are Gaussian). Finally, Lemma 3 shows that the conditional distribution of $\{n^{1/2} T_n^{0*}\}$ given $\mathscr{Z}_n$ is tight with probability 1 (see

Billingsley (1968, p. 9) for a definition of tightness). Then, the theorem follows from (26) and (27). Note that, applying A2 and A8,

$$\begin{aligned}
&\sup_{\tau\in[0,1]} \|n^{1/2}\mathscr{L}_n\tilde{T}_n^*(\tau) - n^{1/2}\mathscr{L}T_n^{0*}(\tau)\| \\
&\quad\leqslant \sup_{\tau\in[0,1]} \|n^{1/2}\mathscr{L}(\tilde{T}_n^*(\tau) - T_n^{0*}(\tau))\| + \sup_{\tau\in[0,1]} \|n^{1/2}(\mathscr{L}_n - \mathscr{L})\tilde{T}_n^*(\tau)\| \\
&\quad\leqslant Cn^{1/2} \sup_{\tau\in[0,1]} \|\tilde{T}_n^*(\tau) - T_n^{0*}(\tau)\| + \mathrm{o_p}(1)\cdot \sup_{\tau\in[0,1]} \|n^{1/2}\tilde{T}_n^*(\tau)\| \\
&\quad= \mathrm{o_{p^*}}(1) \quad \text{in probability,}
\end{aligned} \tag{28}$$

by (26), and noting that $\sup_{\tau\in[0,1]} \|n^{1/2}\tilde{T}_n^*(\tau)\|$ converges in distribution, conditional on $\mathscr{Z}_n$, to $\sup_{\tau\in[0,1]} \|N_k^0(\tau)\|$ by (27) and (26). Henceforth, C denotes a generic constant. □

Remark. If we use a transformation $\mathscr{L}_n^*$, instead of $\mathscr{L}_n$, such that

$$\sup_{\tau\in[0,1]} \|(\mathscr{L}_n^* - \mathscr{L}_n)f(\tau)\| \leqslant \mathrm{o_{p^*}}(1)\cdot \sup_{\tau\in[0,1]} \|f(\tau)\|,$$

it also follows, under the stated assumptions that

$$n^{1/2} \sup_{\tau\in[0,1]} \|\mathscr{L}_n^*\tilde{T}_n^*(\tau) - \mathscr{L}T_n^{0*}(\tau)\| = \mathrm{o_{p^*}}(1) \quad \text{in probability,}$$

applying the argument in (28). For instance, we may want to use a bootstrap version of $S_n(\cdot)$ in the feasible transform.

Remember that $\tilde{\eta}_n^*:=\varphi(n^{1/2}\mathscr{L}_n\tilde{T}_n^*)$ and $\tilde{\eta}_n^{0*}:=\varphi(n^{1/2}\mathscr{L}T_n^{0*})$. Since φ is continuous, and η_∞ has a continuous distribution, (27) implies, applying the continuous mapping theorem that

$$\sup_{x\in\mathbb{R}} |F^*_{\tilde{\eta}_n^{0*}}(x) - F_{\eta_\infty}(x)| \to 0 \quad \text{with probability 1,} \tag{29}$$

and (28) implies that

$$\tilde{\eta}_n^* = \tilde{\eta}_n^{0*} + \mathrm{o_{p^*}}(1), \tag{30}$$

Therefore, (29) and (30) prove the theorem.

Lemma 1. *Under the assumptions in Theorem* 2, *we have that*

$$Cov(n^{1/2}T_n^{0*}(\tau_1), n^{1/2}T_n^{0*}(\tau_2)|\mathscr{Z}_n) \to S(\min(\tau_1,\tau_2)) \quad \textit{with probability } 1$$

for each $\tau_1, \tau_2 \in [0,1]$.

Proof. Since $V_i's$ have zero mean $\mathrm{E}(T_n^{0*}(\tau), |\mathscr{Z}_n) = 0$ a.s., and since they are i.i.d. with unit variance, for each $\tau_1, \tau_2 \in [0,1]$,

$$\begin{aligned}
Cov(n^{1/2}T_n^{0*}(\tau_1), n^{1/2}T_n^{0*}(\tau_2)'|\mathscr{Z}_n) &= \mathrm{E}(n^{1/2}T_n^{0*}(\tau_1), n^{1/2}T_n^{0*}(\tau_2)'|\mathscr{Z}_n) \\
&= \frac{1}{n}\sum_{i=1}^{[n\tau_1]}\sum_{i=1}^{[n\tau_2]} U_{ni}^0 U_{nj}^{0'} \mathrm{E}(V_iV_j|\mathscr{Z}_n)
\end{aligned}$$

$$= \frac{1}{n} \sum_{i=1}^{[n\min(\tau_1,\tau_2)]} U_{ni}^0 U_{ni}^{0\prime}$$

$$\to S(\min(\tau_1, \tau_2)) \quad \text{with probability 1,}$$

by A10. □

Lemma 2. *Under conditions in Theorem* 2, *the finite-dimensional conditional distributions, given* $\mathscr{Z}_n$, *of* $\{n^{-1/2}T_n^{0*}\}$ *converge to the corresponding finite-dimensional distributions of* $\{N_k^0\}$ *with probability* 1.

Proof. Consider the case $k = 1$. The extension to $k > 1$ is trivial. Fix some $\tau_1, \ldots, \tau_q$ and $a_1, \ldots, a_q$. By the Cramér–Wold device, it suffices to investigate the a.s. limit of the conditional distribution given $\mathscr{Z}_n$ of

$$\frac{1}{n^{1/2}} \sum_{j=1}^{q} a_j T_n^{0*}(\tau_j) = \frac{1}{n^{1/2}} \sum_{i=1}^{n} \sum_{j=1}^{q} a_j 1(i \leqslant [n\tau_j]) U_{ni}^0 V_i$$

$$= \frac{1}{n^{1/2}} \sum_{i=1}^{n} \eta_i U_{ni}^0 V_i,$$

where the $\eta_i' s$ are bounded constants. Because the $V_i' s$ are independent of $\mathscr{Z}_n$, and $\{n^{-1/2}\eta_i U_{ni}^0 V_i, i = 1, \ldots n, n \geqslant 1\}$ is a triangular array of random variables, it remains to verify, in view of Lemma 1, Lindeberg's condition. That is, for each $\delta > 0$,

$$L_n(\delta) := \mathrm{E}\left[\frac{1}{n} \sum_{i=1}^{n} \eta_i^2 U_{ni}^{02} V_i^2 1(|\eta_i U_{ni} V_i| > n^{1/2}\delta) \middle| \mathscr{Z}_n \right] \to 0 \quad \text{with probability 1.}$$

Indeed, since $\sup_{1 \leqslant i \leqslant n} |V_i \eta_i| \leqslant c < \infty$, we must check that

$$L_n(\delta) \leqslant c^2 \frac{1}{n} \sum_{i=1}^{n} U_{ni}^{02} 1\left(|U_{ni}| > \frac{n^{1/2}\delta}{c} \right) \to 0 \quad \text{with probability 1,}$$

which is satisfied by A11. □

Lemma 3. *Under the conditions in Theorem* 2, *the conditional distribution of* $\{n^{1/2}T_n^*, n \geqslant 1\}$ *given* $\mathscr{Z}_n$ *is asymptotically C-tight.*

Proof. We only discuss the case $k = 1$, the extension to $k > 1$ is straightforward. Asymptotic tightness in $C[0, 1]$ would follow from its tightness in $D[0, 1]$ and the fact that a possible limit has continuous sample paths with probability one. For tightness in D, it suffices to check Billingsley (1968, Theorem 15.6). That is we must show that, for $0 \leqslant \tau_0 < \tau_1 < \tau_2 \leqslant 1$,

$$\mathrm{E}[|T_n^{0*}(\tau_1) - T_n^{0*}(\tau_0)|^2 |T_n^{0*}(\tau_2) - T_n^{0*}(\tau_1)|^2 \mid \mathscr{Z}_n] \leqslant |\mathrm{H}_n(\tau_2) - \mathrm{H}_n(\tau_0)|^2, \tag{31}$$

for some monotone function H_n, which may depend on $\mathscr{Z}_n$. The left-hand side of (31) is equal to

$$\frac{1}{n^2}\sum_{i=[n\tau_0]+1}^{[n\tau_1]}\sum_{j=[n\tau_0]+1}^{[n\tau_1]}\sum_{k=[n\tau_1]+1}^{[n\tau_2]}\sum_{l=[n\tau_1]+1}^{[n\tau_2]} U_{ni}^0 U_{nj}^0 U_{nk}^0 U_{nl}^0 \mathrm{E}(V_i V_j V_k V_l | \mathscr{Z}_n)$$

$$= \frac{1}{n^2}\sum_{i=[n\tau_0]+1}^{[n\tau_1]}\sum_{j=[n\tau_1]+1}^{[n\tau_2]} U_{ni}^{02} U_{nj}^{02}$$

$$\leqslant \left[\frac{1}{n}\sum_{i=[n\tau_0]+1}^{[n\tau_2]} U_{ni}^{02}\right]^2$$

$$= |\mathrm{H}_n(\tau_2) - \mathrm{H}_n(\tau_0)|^2,$$

where $\mathrm{H}_n(\cdot) = \frac{1}{n}\sum_{i=1}^{[n\cdot]} U_{ni}^{02}$ is monotonically increasing function. By A10, for each $\tau \in [0,1]$,

$$\mathrm{H}_n(\tau) \to S(\tau) \quad \text{with probability 1,}$$

and S is continuous and monotone, which proves the lemma. □

Proof of Corollary 1. It is an immediate consequence of Theorem 2. Just note that Lemma 2, and then boundedness of V_i and A11, are not required, since $\{n^{1/2}T_n^*\}$ is, conditional on $\mathscr{Z}_n$, Gaussian for each n. □

Acknowledgements

We are very thankful to Ronald Gallant, the Associate Editor, and two anonymous referees for their constructive comments on the first version of this paper, formerly titled "External bootstrap tests for structural stability". Research funded by Spanish "Dirección General de Enseñanza Superior", reference number BEC2001-1270 and a Comunidad de Madrid/European Union postdoctoral fellowship with reference number 4323/2000.

References

Amemiya, T., Powell, J.L., 1981. A comparison of the Box-Cox maximum likelihood estimator and the non-linear two-stage least squares estimator. Journal of Econometrics 17, 351–381.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856.

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.

Beran, R., Le Cam, L., Millar, P.W., 1987. Convergence of stochastic empirical measures. Journal of Multivarite Analysis 23, 159–168.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.
Borodin, A.N., Ibragimov, I.A., 1995. Limit Theorems for Functionals of Random Walks. . Proceedings of the Steklov Institute of Mathematics, Vol. 195. American Mathematical Society, Providence, RI.
Brown, B.M., 1971. Martingale Central Limit Theorems. Annals of Mathematical Statistics 42, 59–66.
Brown, R.L., Durbin, J., Evans, J.M., 1975. Techniques for testing the constancy of regression relationships over time (with discussion). Journal of the Royal Statistical Society, Series B 37, 149–192.
Carlstein, E., 1986. The use of subseries methods for estimating the variance of a general statistic from a stationary time series. Annals of Statistics 14, 1171–1179.
Chow, Y.S., 1971. On the L_p convergence of $n^{-1/p}S_n$, $0 < p < 2$. Annals of Mathematical Statistics 42, 393–394.
Davidov, Y.A., 1968. Convergence of distributions generated by stationary stochastic process. Theory of Probability and its Applications 8, 691–696.
Davidov, Y.A., 1970. The invariance principle for stationary process. Theory of Probability and its Applications 15, 487–498.
de Jong, R.M., 1996. On the Bieren's test under data dependence. Journal of Econometrics 72, 1–32.
Delgado, M.A., González-Manteiga, W., 2001. Significance testing in nonparametric regression based on the bootstrap. Annals of Statistics 29, 1469–1507.
Doob, J.L., 1953. Stochastic Processes. Wiley, New York.
Ghysels, E., Guay, A., Hall, A., 1997. Predictive tests for structural change with unknown breakpoint. Journal of Econometrics 82, 209–233.
Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and its Applications. Academic Press, New York.
Hall, P., Horowitz, J., 1996. Bootstrap critical values for tests based on generalized method-of-moments estimators. Econometrica 64, 891–916.
Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.
Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.
Hansen, B.E., 2000. Testing for structural change in conditional models. Journal of Econometrics 97, 93–115.
Härdle, W., Mammen, E., 1991. Bootstrap methods in nonparametric regression. In: G. Roussas (Ed.) Nonparametric Functional Estimation and Related Topics. Kluwer Academic Publishers, Dordrecht, pp. 111–123.
Kuan, C.-M., Hornik, K., 1995. The generalized fluctuation test: A unifying view Econometric Reviews 14, 135–161.
Künsch, H.R., 1989. The jacknife and the bootstrap for general stationary observations. Annals of Statistics 17, 1217–1241.
Liu, R.Y., 1988. Bootstrap procedures under some non-i.i.d. models. Annals of Statistics 16, 1697–1708.
MacLeish, D.L., 1975a. Invariance principles for dependent variables. Zeitschrift fuer Wahrscheinlichkeitstheorie verw. Gebiete 32, 165–178.
MacLeish, D.L., 1975b. A maximal inequality and dependent strong laws. Annals of Probability 5, 829–839.
MacLeish, D.L., 1977. On the invariance principles for nonstationary mixingales. Annals of Probability 5, 616–621.
Nyblom, J., 1989. Testing for the constancy of parameters over time. Journal of the American Statistical Association 84, 223–230.
Park, J.Y., Phillips, P.B., 2001. Nonlinear regression with integrated time series. Econometrica 69, 117–162.
Ploberger, W., Krämer, W., 1992. The CUSUM test with OLS residuals. Econometrica 60, 271–286.
Politis, N.D., Romano, J.P., Wolf, M., 1999. Subsampling. Springer, Berlin.
Quandt, R.E., 1960. Tests of the hypothesis that a linear regression system obeys two separate regimes. Journal of the American Statistical Association 55, 324–330.
Robinson, P.M., 1991. Best nonlinear three-stage least squares estimation of certain econometric models. Econometrica 59, 755–787.
Romano, J.P., Wolf, M., 2001. Subsampling intervals in autoregressive models with linear time trend. Econometrica 69, 1283–1315.
Sowell, F., 1996. Optimal tests for parameter instability in the generalized method of moments framework. Econometrica 64, 1085–1107.

Stute, W., González-Manteiga, W., Presedo Quindimil, M., 1998. Bootstrap approximations in model checks for regression. Journal of the American Statistical Association 33, 141–149.

Su, J.Q., Wei, L.J., 1991. A lack of fit test for the mean function in a generalized linear model. Journal of the American Statistical Association 86, 420–426.

Velasco, C., Robinson, P.M., 1997. Autocorrelation robust inference. In: Handbook of Statistics 15, 267–298.

Wu, C.F.J., 1986. Jackknife, bootstrap and other resampling methods in regression analysis (with discussion.) Annals of Statistics 14, 1261–1350.