



**HAL**  
open science

# An Omnibus Test to Detect Time-Heterogeneity in Time Series

Dominique Guegan, Philippe de Peretti

► **To cite this version:**

Dominique Guegan, Philippe de Peretti. An Omnibus Test to Detect Time-Heterogeneity in Time Series. 2010. halshs-00560221v1

**HAL Id: halshs-00560221**

**<https://shs.hal.science/halshs-00560221v1>**

Submitted on 27 Jan 2011 (v1), last revised 22 Dec 2011 (v2)

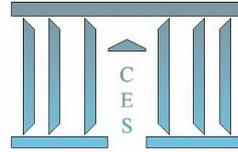
**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Documents de Travail du Centre d'Économie de la Sorbonne

C  
E  
S  
W  
o  
r  
k  
i  
n  
g  
P  
a  
p  
e  
r  
s



## **An Omnibus Test to Detect Time-Heterogeneity in Time Series**

Dominique GUEGAN, Philippe de PERETTI

**2010.98**



# An Omnibus Test to Detect Time-Heterogeneity in Time Series

Dominique Guégan\*      Philippe de Peretti†

December 10, 2010

## Abstract

In this paper, we present a procedure that tests for the null of time-homogeneity of the first two moments of a time series. The test introduced here is a priori to modelling, and is an omnibus test in the sense that, compared to the existing literature, it has power against a broad alternative including for instance *i*) Discrete shifts, *ii*) Smooth transition, *iii*) Time-varying moments, *iv*) probability-driven breaks, *v*) GARCH or Stochastic Volatility Models for the variance. Our test uses the recently introduced maximum entropy bootstrap, designed to capture both time-dependency and time-heterogeneity. Running simulations, our procedure appears to be quite powerful. To some extent, our paper can be seen as a development of Heracleous, Koutris and Spanos (2008).

**Keywords:** Test ; Time-homogeneity ; Maximum Entropy Bootstrap.

**JEL Classification:** C01, C12, C15.

## 1 Introduction

Over the past two decades, unit roots have become the corner stone of modern time-series econometrics. Indeed, prior to estimation, testing for unit roots is

---

\*Centre d'Economie de la Sorbonne (CES), Université Paris Panthéon-Sorbonne, Maison des Sciences Economiques, 106-112 Boulevard de l'Hôpital, 75013 Paris. Email: dominique.guegan@univ-paris1.fr

†Centre d'Economie de la Sorbonne (CES), Université Paris Panthéon-Sorbonne, Maison des Sciences Economiques, 106-112 Boulevard de l'Hôpital, 75013 Paris. Email: philippe.de-peretti@univ-paris1.fr

required to avoid spurious results. In such an approach, unit roots are often seen as the main source of the heterogeneity of the first two moments of a time series. Nevertheless, unit roots are only one kind of time-heterogeneity and then of non-stationarity. Following Spanos (1999), time-heterogeneity is a broader concept also including for instance discrete breaks, smooth transition, and other structural changes. Time-heterogeneity is therefore a more realistic concept, especially in macro-econometrics and in finance.

Perron (2005) reviews the huge literature dedicated to structural changes testing procedures. At least two lessons are to be understood: *i*) Testing for structural changes is a prior to modelling and testing, or alternatively one can both model and take into account structural changes (see e.g. Andrews and Ploberger [1994], Bai [1999], Bai and Perron [1998, 2003], or Altissimo and Corradi [2003]). On the one hand, structural changes are a source of global non-stationarity (Granger and Starica [2005] and Guégan [2010]), and on the other hand, they are likely to bias tests for stationarity (Perron [1989]), and for long memory (Diebold and Inoue [2001], Charffedine and Guégan [2011]). Hence, in addition to causing parameter instability and spurious results, time-heterogeneity, or structural changes, may lead to erroneous statistical inference and thus to incorrect modelling, *ii*) Often, the alternative considered is very narrow, focusing either on discrete shifts, or smooth transition, thus forcing the econometrician to implement different kinds of tests for different assumptions. Hence, the need for an omnibus test to detect time-heterogeneity arising from various sources.

In a recent contribution, Heracleous, Koutris and Spanos (2008) (hereafter HKS), have proposed such an omnibus test, considering discrete shifts, smooth transition or trended moments under the alternative. On a “de-memorized” series, i.e. residuals of an ARMA model, they compute accurate rolling windows estimators of the first two moments using the Maximum Entropy BOOTstrap (MEBOOT) method (Vinod [2004, 2006]). They then model the rolling windows estimators as a first-order auto-regressive process, plus an orthogonal Bernstein polynomial to capture a possible time-heterogeneity. With  $k$ , a sufficiently large polynomial degree, testing for the null of time-homogeneity amounts to testing  $H_0 : k = 0$  against  $H_A : k \geq 1$ , using a standard F-test. Implementing Monte-

Carlo simulations, they show that their test behaves well.

In this paper, we extend the HKS approach, and propose an alternative procedure to test for the null of time-homogeneity having power against a broad alternative, including for instance *i*) Discrete structural breaks, *ii*) Smooth transition, *iii*) time-varying moments, *iv*) probability-driven breaks, i.e. Markov switching models and *v*) GARCH and Stochastic Volatility Models for the second moment. To test for the null, we use a signal plus noise model, and check that the signal, defined as an orthogonal Bernstein polynomial, is constant over time. Our test differs from HKS in the following points: *i*) All tests are conducted on the overall series, and not locally, thus avoiding the difficult choice of choosing a particular window, which may dramatically alter the results, *ii*) The test procedure does not use standard F-tests, but is rather based on bootstrap tailor-made statistics using the Minkowski distance between the forecasts of two auxiliary regressions (see e.g. Davison and Hinkley [1998]). The tailor-made statistics are computed using the MEBOOT procedure, *iii*) In one of the two auxiliary regressions, the degree of the Bernstein polynomial is chosen using the *AICu* criterion, ensuring an optimal trade-off between fitting and smoothing (McQuarrie and Tsai [1998]).

The paper is organized as follows. Section 2 discusses time-heterogeneity and introduces the test. Section 3 implements Monte-Carlo simulations to estimate the empirical size and power under various alternatives. Section 4 presents an empirical application, and Section 5 concludes.

## 2 Testing for the null of time-homogeneity

Let  $\{y_t\}_{t=1}^T$  be an observed time series, and let  $\{\varepsilon_t\}_{t=1}^T$  be the associated “de-memorized” series, as in Andrews and Ploberger<sup>1</sup> (1994) or HKS. For instance, assume that  $\{\varepsilon_t\}_{t=1}^T$  are the residuals of an ARMA model for  $\{y_t\}_{t=1}^T$ .

**Definition 1.** *The process  $\{y_t\}_{t=1}^T$  is said to be weakly stationary if the three conditions are fulfilled: *i*)  $E(y_t) = \mu_t = \mu \forall t \in [1, T]$ , *ii*)  $V(y_t) = \sigma_t^2 = \sigma^2 \forall t \in [1, T]$ , *iii*)  $cov(y_t, y_{t+h}) = g(h)$ .*

---

<sup>1</sup>See also Andrews (1993) and Andrews, Lee and Ploberger (1996).

For instance in the simple Data Generating Process (DGP) defined by (1), Definition 1 implies that to achieve stationarity, we must have: *i*)  $\mu_t = \mu$ , *ii*)  $\sigma_t^2 = \sigma^2$  and *iii*)  $-1 < \phi_t = \phi < 1 \forall t \in [1, T]$ . If additionally  $\phi = 0$ , then data are said to be identically and independently distributed (iid).

$$y_t = \mu_t + \phi_t y_{t-1} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma_t^2) \quad (1)$$

In what follows, we focus only on the residuals  $\{\varepsilon_t\}_{t=1}^T$  i.e. on the filtered series. Our general DGP is given by (2):

$$\varepsilon_t = \mu_t + \sigma_t \eta_t \quad (2)$$

where:  $\eta_t$  is a zero-mean iid term with unit variance,  $\mu_t$  and  $\sigma_t$  are possibly time varying parameters.

**Definition 2.** *There is time-homogeneity if  $\mu_t$  and  $\sigma_t$  are constant over time.*

What we want, is to track the process driving  $\mu_t$  and  $\sigma_t$  over time, and analyze if such a process returns constant values, i.e.  $\mu_t = \mu$  and  $\sigma_t = \sigma \forall t \in [1, T]$ . Therefore, the assumption we want to test is, for a given moment, for instance for the mean:  $H_0 : \mu_t = \mu, \forall t \in [1, T]$  against the broad alternative:  $H_A : \exists \lambda_i \in [1, T-1] : \mu_{\lambda_i} = \mu_{\lambda_{i+1}}$ . Clearly, the alternative spans several processes, as *i*) Discrete shifts, *ii*) Smooth transition, *iii*) time-varying moments, *iv*) probability-driven breaks, i.e. Markov switching as well as complex and functions of time as *v*) GARCH and Stochastic Volatility Models.

To track the underlying process, the idea is to represent it as a signal plus noise model, and then test for the time-invariance of the signal.

In what follows, we chose extracting the signal from the “de-memorized” series, by estimating the following model:

$$\varepsilon_t = B_k(t) + v_t \quad (3)$$

Where:  $B_k(t)$  is a Bernstein polynomial<sup>2</sup> of degree  $k$  defined as:

$$B_k(t) = \sum_{j=0}^k \alpha_j \binom{k}{j} t^j (1-t)^{k-j}, \quad 0 \leq t \leq 1$$

---

<sup>2</sup>Using polynomials to detect time-heterogeneity has been first suggested by MacNeill (1978) and extended by Perron (1991) and Tang and MacNeill (1993).

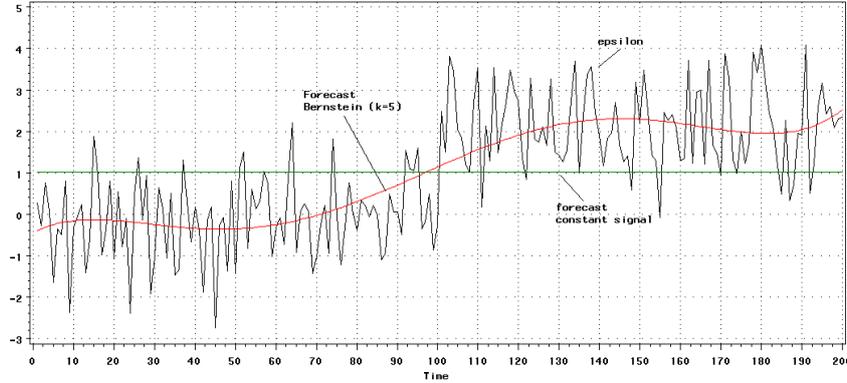


Figure 1: Simulation of an iid sequence ( $T = 200$ ), with a single rupture in mean:  $\mu_t = 0$  for  $t \leq 100$  and  $\mu_t = 2$  otherwise, together with a constant signal and a Bernstein polynomial ( $k = 5$ ).

Where:  $v_t$  is an iid noise, and the  $\alpha'_j$ 's are parameters to be estimated.

For instance, Figure 1 plots a series with a rupture in mean in  $t = 100$ , together with a Bernstein polynomial of order  $k = 5$  (*AICu* selected degree, see below). Note that the polynomial perfectly captures the heterogeneity in mean, given the information set  $IS_T = \{\varepsilon_t\}_{t=1}^T$ .

To test for the null of time-homogeneity, focusing on the first moment, we proceed as follows:

1. Estimate (3) forcing  $k = 0$ , i.e. estimate:

$$\varepsilon_t = \mu + v_t \quad (4)$$

2. Estimate (3) with  $k$  defined according to the *AICu* criterion (5), McQuarrie and Tsai (1998):

$$AICu = \log \left( \frac{\mathbf{v}'\mathbf{v}}{T - k_1} \right) + \frac{T + k_1}{T - k_1 - 2}, \quad (5)$$

where:  $\mathbf{v} = (v_1, v_2, \dots, v_T)$  and  $k_1 = k + 1$ ;

3. Compute the Minkowski distance on the forecasts:

$$D_p = \left( \sum_{i=1}^T \left| \hat{\varepsilon}_i^1 - \hat{\varepsilon}_i^2 \right|^p \right)^{1/p} \quad (6)$$

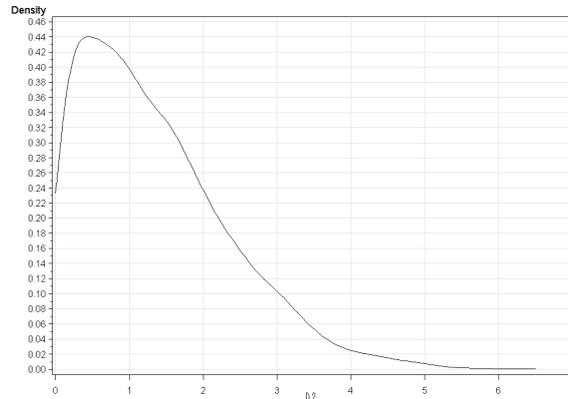


Figure 2: Estimation of the Kernel density of (6) with  $p = 2$ , for an iid sequence ( $T = 100$ ) with a single rupture in mean ( $\mu_t = 0$  for  $t \leq 50$  and  $\mu_t = 2$  otherwise). Estimation is computed by bootstrapping under the null using MEBOOT, 10.000 replications.

where:

$\hat{\varepsilon}_1^1 = E[\varepsilon_1|IS_T], \hat{\varepsilon}_2^1 = E[\varepsilon_2|IS_T], \dots, \hat{\varepsilon}_T^1 = E[\varepsilon_T|IS_T]$ : are the forecasts of model (4),

$\hat{\varepsilon}_1^2 = E[\varepsilon_1|IS_T], \hat{\varepsilon}_2^2 = E[\varepsilon_2|IS_T], \dots, \hat{\varepsilon}_T^2 = E[\varepsilon_T|IS_T]$ : are the forecasts of model (3) when the degree of the polynomial is chosen according to the  $AICu^3$  criterion.

4. Test the significance of the Minkowski distance, i.e.  $H_0 : D_p = 0$  against  $H_A : D_p > 0$  for a given value of  $p$  ( $p = 1, 2, \infty$ ), a non-significance of the distance supporting the null of time-homogeneity.

Testing the significance of the Minkowski distance is achieved by computing bootstrap (tailor-made) p-values using the MEBOOT procedure<sup>4</sup>. Interestingly,

<sup>3</sup>See Appendix A for an empirical comparison of three different information criteria ( $AIC$ ,  $AICc$  and  $AICu$ ) in the Bernstein polynomial framework.

<sup>4</sup>The MEBOOT procedure for a series  $\{\varepsilon_t\}_{t=1}^T$  is as follows:

- a) Sort  $\{\varepsilon_t\}_{t=1}^T$  by ascending order. Let  $\{\varepsilon_{[t]}\}_{t=1}^T$  be the sorted series. Define  $\{ind_t\}_{t=1}^T$  as an index containing the ordering of the original series.
- b) Compute intermediate points  $z_t = \frac{1}{2}(\varepsilon_{[t]} + \varepsilon_{[t+1]})$ ,  $t = 1, \dots, T - 1$ .
- c) Define intervals  $I_1, \dots, I_T$  with equiprobability as well as bounds for the first and last intervals. For this compute the trimmed mean of  $|\Delta\varepsilon_t|$  and compute the lower bound

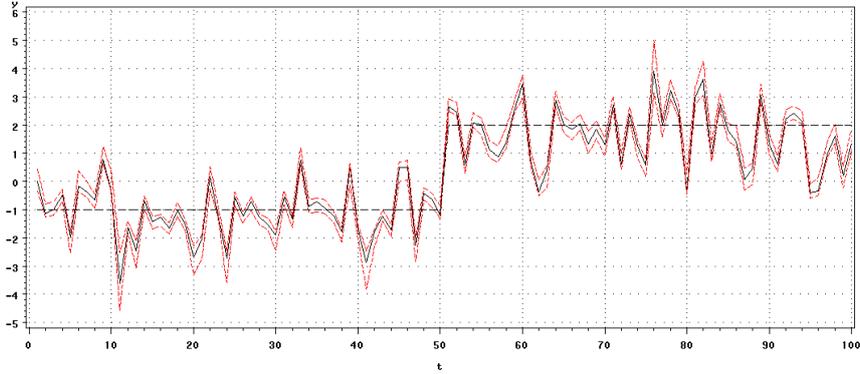


Figure 3: Sequence ( $T = 100$ ) of iid terms, with a single rupture in mean ( $\mu_t = -1$  for  $t \leq 50$  and  $\mu_t = 2$  otherwise), together with a 95% confidence interval built by bootstrapping the series.

this bootstrap procedure is designed to capture both time-dependency and time-heterogeneity. For instance, Figure 3 plots a series with a discrete break in mean (in  $t = 50$ ), as well as a 95 % confidence interval using the bootstrap procedure. The procedure clearly captures the break. A similar pattern appears in Figure 4, showing a discrete break in the variance (in  $t = 50$ ).

To build bootstrap p-values we sample, independently  $\{\varepsilon\}_{t=1}^T$  two times, and compute  $D_p$  under the null as follows:

1. Sample  $\{\varepsilon\}_{t=1}^T$  using the maximum entropy bootstrap and estimate (4).  
Let  $\{\widehat{\varepsilon}^{b1}\}_{t=1}^T$  be the forecast,
2. Sample  $\{\varepsilon\}_{t=1}^T$  using the maximum entropy bootstrap and estimate (4).  
by removing this mean to  $z_1$  and the upper bound by adding this mean to  $z_{T-1}$ .
- d) On each interval compute the desired means defined as  $m_1 = 0.75\varepsilon_{[1]}(1) + 0.25\varepsilon_{[2]}$  for  $I_1$ ,  $m_T = 0.25\varepsilon_{[T-1]} + 0.75\varepsilon_{[T]}$  for  $I_T$  and  $m_k = 0.25\varepsilon_{[k-1]} + 0.5\varepsilon_{[k]} + 0.25\varepsilon_{[k+1]}$  for intermediate values.
- e) Draw uniform realizations on  $[0,1]$  and compute the associated quantiles for  $\{\varepsilon_{[t]}\}_{t=1}^T$  by linear interpolation.
- g) Adjust the quantiles using  $z_t$  to preserve the means, and reorder the adjusted quantiles according to  $ind_t$ . This returns a bootstrap realization for  $\{\varepsilon_t\}_{t=1}^T$ .
- h) Repeat steps a) to g) a large number of times.

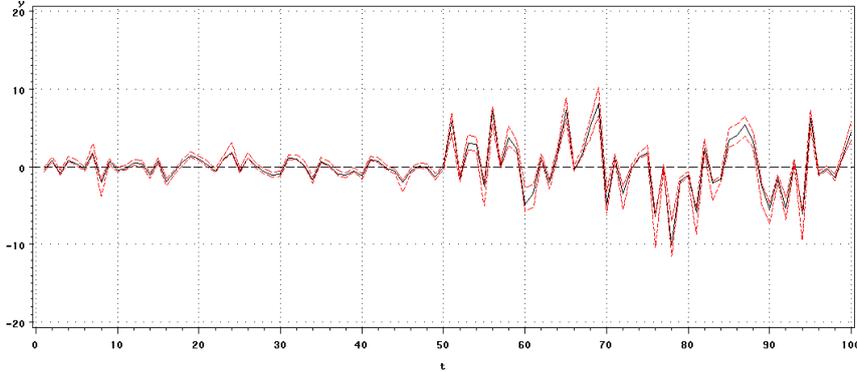


Figure 4: Sequence ( $T = 100$ ) of iid terms, with a single rupture in variance ( $\sigma_t = 1$  for  $t \leq 50$  and  $\sigma_t = 4$  otherwise), together with a 95% confidence interval built by bootstrapping the series.

Let  $\{\hat{\varepsilon}^{b2}\}_{t=1}^T$  be the forecast,

3. Compute a bootstrap<sup>5</sup> value for  $D_p$ , say  $D_p^b = \left( \sum_{i=1}^T \left| \hat{\varepsilon}_i^{b1} - \hat{\varepsilon}_i^{b2} \right|^p \right)^{1/p}$ ,
4. Repeat steps 1 to 3 a large number of times, and compute the empirical distribution of  $D_p^b$ , and then the tailor-made p-value. Reject the null if the p-value is less than a given threshold.

Figure 2 illustrates the above four-step procedure<sup>6</sup>. It returns the estimated density of  $D_2^b$  by Kernel method (10.000 replications) for a series ( $T = 100$ ) presenting a single shift in mean (from 0 to 2) at the middle of the sample size. In our simulation, the value of  $D_2$  equals 6.97, with a p-value of 0, which is deeply coherent with our DGP.

We next turn to intensive Monte-Carlo simulations.

<sup>5</sup>Let  $\{\varepsilon^{jb}\}_{t=1}^T$  be a bootstrapped sample of  $\{\varepsilon\}_{t=1}^T$ . Then, under the null, the Minkowski distance becomes:  $D_1 = T |\bar{\varepsilon}^{1b} - \bar{\varepsilon}^{2b}|$  and  $D_2 = T^{1/2} |\bar{\varepsilon}^{1b} - \bar{\varepsilon}^{2b}|$  where  $\bar{\varepsilon}^{jb} = \frac{1}{T} \sum_{i=1}^T \varepsilon^{jb}$ . For  $D_\infty$ , it is just, in absolute value the difference between the two means.

<sup>6</sup>Alternatively, one can bootstrap  $D_p$  for the estimated  $k$ . Empirically, this returns very similar results.

### 3 Monte-Carlo simulations

In this section, we implement intensive Monte-Carlo simulations to estimate the size and power of the procedure. We first introduce our different DGPs. The first one (case 0) is used to study the empirical size of the tests. The DGP is then altered to estimate the power under various alternatives. The basic DGP is as follows:

#### Case 0 [mean and variance]: iid process

1. Generate  $T$  zero-mean normally<sup>7</sup> distributed terms with unit variance  $\{\varepsilon_t\}_{t=1}^T$ ,
2. Estimate (3) for various values of  $k \in [0, 15]$ ,
3. Choose the model minimizing the *AICu* criterion,
4. If  $k = 0$  then report a success, if  $k \neq 0$  then test for the significance of  $D_p$  using bootstrap p-values.
5. Repeat steps 1 to 4, 1000 times.

We next alter the above DGP to study the power of the procedure.

#### Case 1 [mean and variance]: Single discrete shift in mean or in variance

Our DGP is the same except for steps 1 and 4 replaced by:

1. For the mean, define  $\{\varepsilon_t\}_{t=1}^T$  as  $\varepsilon_t = \mu + \mu_1 \cdot 1_{\{t > \lambda\}} + v_t$ ,  $v_t \sim N(0, 1)$ , where  $1(\cdot)$  is the standard indicator function,  $\lambda \in [T/3, 2T/3]$  is the breaking date, and is randomly drawn with equiprobability within the interval at each replication,  $\mu = 0$  and  $\mu_1 = 2$ .
1. For the variance, define  $\{\varepsilon_t\}_{t=1}^T$  as  $\varepsilon_t = (\sigma + \sigma_1 \cdot 1_{\{t > \lambda\}})v_t$ ,  $v_t \sim N(0, 1)$ , where  $1(\cdot)$  is the standard indicator function,  $\lambda \in [T/3, 2T/3]$  is the breaking date, and is randomly drawn with equiprobability within the interval at each replication,  $\sigma = 1$  and  $\sigma_1 = 2$ .
4. For both scenarii, report if it is a success, that is if  $k > 0$  and  $D_p$  is significantly different from 0.

---

<sup>7</sup>In this section, we generate  $\{\varepsilon_t\}_{t=1}^T$  as normally distributed terms. Nevertheless, since we don't use a residual-based bootstrap, any law with stable first two moments can be used.

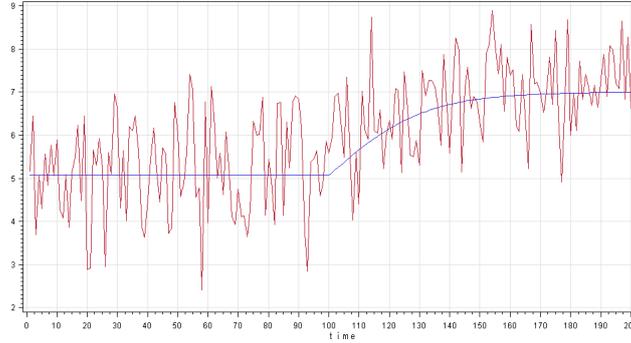


Figure 5: Sequence ( $T = 200$ ) of iid terms, with a constant mean for  $t \leq 100$ , and logistic driven otherwise.

**Case 2 [mean]. Smooth rupture in mean** We now allow the rupture in mean to be driven by a smooth process. Here a standard logistic function is used. Our DGP is the same as above, except that the Step 1 for the mean is replaced by:

1. Define  $\{\varepsilon_t\}_{t=1}^T$  as  $\varepsilon_t = \mu + \frac{\alpha}{1+\exp(-\delta/(T/10-\theta))} + \mu_1 + v_t, v_t \sim N(0, 1)$ ,  $\mu = 0$ ,  $\mu_1 = 3$ ,  $\alpha = 4$ ,  $\delta = 1$  if  $t < \lambda$  and  $\delta = (t - \lambda)$  otherwise,  $\theta = 4$ ,  $\lambda \in [T/3, 2T/3]$  is the breaking date, and is randomly drawn with equiprobability within the interval at each replication.

Thus, for  $t \leq \lambda$  the process returns a constant mean, and for  $t > \lambda$  a time-varying mean reaching a bound.

A typical realization of such a process is given by Figure 5, with  $\lambda = 0.5$  and  $T = 200$ .

**Case 3 [mean]. Stop-Break model** The stop-break model can be seen as a variant of the local trend model, and allows one to have a continuous time-varying mean according to the state-space representation:

$$\varepsilon_t = \mu_t + v_t \quad (7)$$

$$\mu_t = \mu_{t-1} + \frac{v_{t-1}^2}{\gamma + v_{t-1}^2} v_{t-1} \quad (8)$$

Such a model has been introduced by Engle and Smith (1999) and studied by Diebold and Inoue (2001) within the long-memory process framework. Our

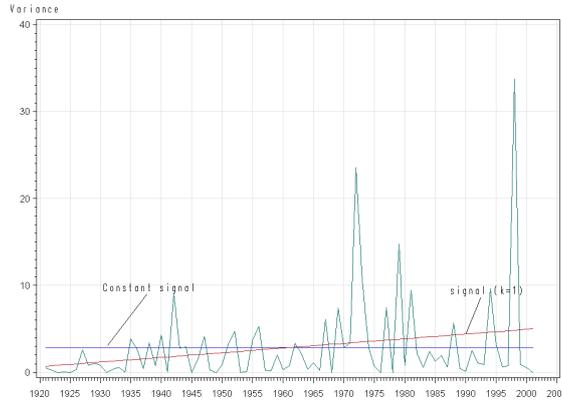


Figure 6: Variance of the yearly US industrial production (USINDPRO) series, together with two Bernstein polynomials with degree  $k = 0$  and  $k = 1$ , the latter being selected using the  $AICu$  criterion.

DGP is still the same, except that the above step 1 is replaced by the following:

1. Define  $\{\varepsilon_t\}_{t=1}^T$  as  $\varepsilon_t = \mu_t + v_t$ , where the dynamics for  $\mu_t$  is given by (8), with  $\gamma = 0.05$ ,  $\eta_t \sim N(0, 1)$ .

**Case 4 [mean]: Markov-switching models** Up to now, under the alternative, we have considered discrete, smooth and time-varying processes for the mean. Here, we consider a two-regime probability-driven switching model (Hamilton [1990]). The modified DGP for  $\{\varepsilon_t\}_{t=1}^T$  is as follows:

$$\varepsilon_t = \mu_{S_t} + \eta_t \quad (9)$$

where:  $S_t$  is a Markov chain, and the stochastic transition matrix is given by:  $T = \begin{bmatrix} p_{ii} & (1 - p_{ii}) \\ (1 - p_{jj}) & p_{jj} \end{bmatrix}$ , with  $p_{ij} = P(S_t = j | S_{t-1} = i)$ . The step 1 is as follows,

1. Define  $\{\varepsilon_t\}_{t=1}^T$  as  $\varepsilon_t = \mu_{S_t} + \eta_t$ ,  $\eta_t \sim N(0, 1)$ ,  $\mu = (0, 2)$ ,  $p_{ii} = 0.8$  and  $p_{jj} = 0.7$ .

**Case 5 [variance]: GARCH errors** Focusing now on the variance, We allow our process to have Generalized Autoregressive Conditional Heteroskedastic (GARCH) errors . Compared to case 3, the step 1 is as follows:

Table 1: Empirical size of tailor-made statistics at the 5% nominal size for the DGP introduced in case 0

$T$	$H_0 : \mu \text{ constant}$			$H_0 : \sigma^2 \text{ constant}$		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	0.016	0.026	0.093	0.025	0.049	0.098
100	0.017	0.030	0.092	0.027	0.051	0.101
150	0.022	0.045	0.134	0.039	0.066	0.122
200	0.011	0.032	0.095	0.025	0.043	0.095

Note: Entries represent the rejection frequencies when the null is true, over 1000 replications.

1. Define  $\{\varepsilon_t\}_{t=1}^T$  as  $\varepsilon_t = \sigma_t \eta_t$ , where  $\eta_t \sim N(0, 1)$  and the dynamics for  $\sigma_t$  is given by (10)

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \quad (10)$$

where:  $q = p = 1$ ,  $\alpha_0 = 0.01$ ,  $\alpha_1 = 0.1$  and  $\beta_1 = 0.8$ .

**Case 6 [variance]: Stochastic Volatility Models** The above GARCH(1,1) returns a stationary conditional variance. We focus on a competing model, which is the Stochastic Volatility Model. Among others, Stochastic Volatility Models have been considered by Jacquier, Polson and Rossi (1994) and Ghysels, Harvey and Renault (1996) within the state-space framework. To generate a variance following such a process, we alter our step 1 as:

1. Define  $\{\varepsilon_t\}_{t=1}^T$  as  $\varepsilon_t = \exp(\sigma_t/2)\eta_t$ , where  $\eta_t \sim N(0, 1)$  and the dynamics for  $\sigma_t$  is given by (11)

$$\sigma_t = \alpha + \beta \sigma_{t-1} + \nu_t \quad (11)$$

where:  $\alpha = 0.05$  and  $\beta = 1$ .

We now discuss the results of the various simulations. The empirical size, at the 5% nominal size<sup>8</sup> of the procedure is reported in Table 1. For  $p = 1$  ( $L1$  norm) and  $p = 2$  ( $L2$  norm), the empirical sizes are all less than 5% or slightly above (for instance for  $T = 150$  and for the variance). For  $p = \infty$ ,

<sup>8</sup>For clarity of the exposé, we only report the empirical size and power at the 5% nominal size. Size and power at the 1%, 10% and 15% are available under request.

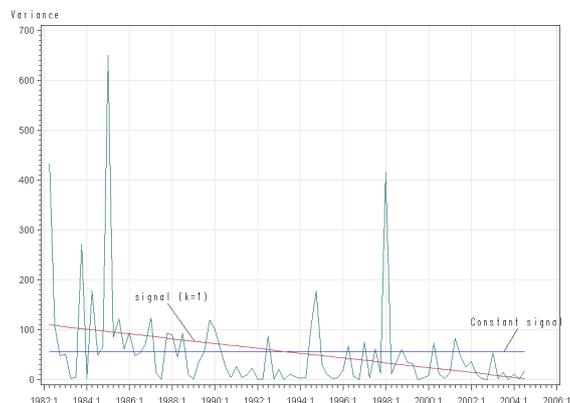


Figure 7: Variance of the quarterly US/Japan Exchange rate (USJEXR), together with two Bernstein polynomials with degree  $k = 0$  and  $k = 1$ , the latter being selected using the  $AIC_u$  criterion.

the empirical size is slightly higher but remains within an acceptable range. Thus, the procedure perfectly recognizes time-homogeneity when there is time-homogeneity. Note that the size does not decrease with the sample size.

Tables 2 and 3 present the power (and size) of the procedure. Table 2 is relative to models with structural changes in mean. Two results are to be emphasized: *i*) For a single discrete shift, a smooth transition, or for the Stop-break model, the power of the procedure is unity. For a Markov-switching model, the power lies between 0.710 and 0.788, which is still acceptable. Hence, the model accurately captures the modifications in mean. *ii*) Analyzing the results of the variance (size), since the second moment is empirically built using the first one, modifications of the mean deeply alter the size of the test, which is fact quite trivial. The size is maximal for the Stop-Break model. This imposes running the test sequentially, first test for the stationarity of the mean, and next, if not rejected, for the stationarity of the variance. Concerning Table 3, i.e. constant mean but structural changes in variance, the test appears to be also quite powerful for the three considered models: discrete shift, GARCH and Stochastic Volatility Model. For a discrete shift in variance, the size of the test for the mean is not altered, as it is for the GARCH model and the Stochastic Volatility Model. Nevertheless, it remains within an acceptable range.

Table 2: Empirical size and power at the 5% nominal size for models exhibiting structural changes in mean

Single discrete shift in mean, $\lambda \in [T/3, 2T/3]$						
$T$	$H_0 : \mu$ constant			$H_0 : \sigma^2$ constant		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	1.000	1.000	1.000	0.026	0.108	0.300
100	1.000	1.000	1.000	0.330	0.330	0.330
150	1.000	1.000	1.000	0.442	0.442	0.442
200	1.000	1.000	1.000	0.460	0.043	0.095

Smooth transition in mean, $\lambda \in [T/3, 2T/3]$						
$T$	$H_0 : \mu$ constant			$H_0 : \sigma^2$ constant		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	1.000	1.000	1.000	0.118	0.118	0.118
100	1.000	1.000	1.000	0.156	0.156	0.156
150	1.000	1.000	1.000	0.166	0.184	0.184
200	1.000	1.000	1.000	0.220	0.244	0.272

Stop-Break model						
$T$	$H_0 : \mu$ constant			$H_0 : \sigma^2$ constant		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	0.998	1.000	1.000	0.952	0.962	0.972
100	1.000	1.000	1.000	0.984	0.984	0.988
150	1.000	1.000	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	0.980	0.980	1.000

Two-regime Markov switching model						
$T$	$H_0 : \mu$ constant			$H_0 : \sigma^2$ constant		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	0.724	0.724	0.724	0.856	0.856	0.856
100	0.710	0.710	0.710	0.842	0.842	0.842
150	0.774	0.782	0.782	0.890	0.864	0.862
200	0.758	0.780	0.788	0.910	0.884	0.854

Note 1: Entries for ' $H_0 : \mu$  constant', represent the rejection frequencies when the null is not true, over 1000 replications.

Note 2: Entries for ' $H_0 : \sigma^2$  constant', represent the rejection frequencies when the null is true, over 1000 replications.

## 4 A real data example

We now implement the test on real data. We implement the procedure on three of the various series used in HKS, namely, the yearly US industrial production (USINDPRO), from 1921 to 2004, the quarterly US/Japan Exchange rate (US-JEXR), from 1982Q2 to 2005Q2 and the quarterly US investment (USINVEST),

Table 3: Empirical size and power at the 5% nominal size for models exhibiting structural changes in variance

Single discrete shift in variance, $\lambda \in [T/3, 2T/3]$						
$T$	$H_0 : \mu$ constant			$H_0 : \sigma^2$ constant		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	0.122	0.122	0.122	0.908	0.908	0.908
100	0.128	0.128	0.128	0.962	0.962	0.962
150	0.114	0.114	0.114	1.000	1.000	1.000
200	0.102	0.110	0.110	1.000	1.000	1.000

GARCH model						
$T$	$H_0 : \mu$ constant			$H_0 : \sigma^2$ constant		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	0.286	0.286	0.286	0.962	0.962	0.962
100	0.276	0.276	0.276	0.962	0.962	0.962
150	0.254	0.254	0.254	0.950	0.950	0.950
200	0.290	0.290	0.290	0.924	0.924	0.924

Stochastic Volatility Model						
$T$	$H_0 : \mu$ constant			$H_0 : \sigma^2$ constant		
	$p = 1$	$p = 2$	$p = \infty$	$p = 1$	$p = 2$	$p = \infty$
80	0.238	0.238	0.238	0.860	0.860	0.860
100	0.170	0.170	0.170	0.882	0.882	0.882
150	0.286	0.286	0.286	0.938	0.938	0.938
200	0.152	0.262	0.318	0.958	0.981	0.986

Note 1: Entries for ' $H_0 : \sigma^2$  constant', represent the rejection frequencies when the null is not true, over 1000 replications.

Note 2: Entries for ' $H_0 : \mu$  constant', represent the rejection frequencies when the null is true, over 1000 replications.

Table 4: Values of (6) ( $p = 2$ ) for the three (de-memorized) series, USINDPRO, USJEXR and USINVEST

Series	$H_0 : \mu$ constant		$H_0 : \sigma$ constant	
	$k$	$D_2$	$k$	$D_2$
USINDPRO	0	0 (1)	1	11.38 (0.072)
USJEXR	0	0 (1)	1	297.49 (0.015)
USINVEST	0	0 (1)	2	0.41 (0.001)

Note: P-values between parenthesis

from 1963Q2 to 1982Q4. HKS found out strong evidence of a structural change in the variance for USINDPRO and USJEXR, and a structural change in mean for USINVEST.

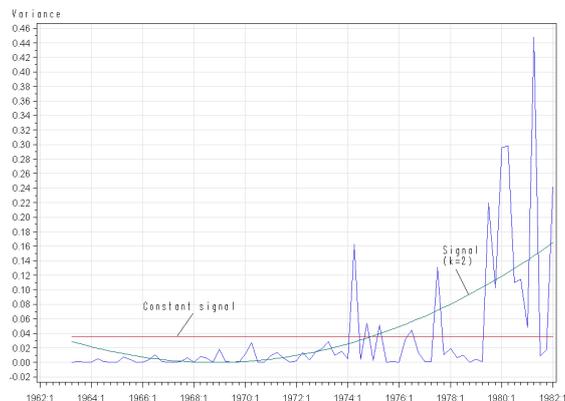


Figure 8: Variance of the quarterly US investment (USINVEST) series, together with two Bernstein polynomials with degree  $k = 0$  and  $k = 2$ , the latter being selected using the  $AICu$  criterion.

Following Koutris (2005), we first filter the series using an auto-regressive process (using the same order), and then run the test on the filtered series. Our results are presented in Table 4. The analysis leads to accept the null of mean-homogeneity for the three series, while rejecting variance-homogeneity (for USINDPRO at 10%). Our results therefore slightly differ from HKS.

## 5 Conclusion

In this paper we have proposed an omnibus test to detect global non-stationarity in time series, when the non-stationarity is caused by structural changes. The test proposed here is to be used prior to modelling. Note that the procedure is used to test if the moments are constant, but if the null is rejected, it returns no information about the process driving the changes in the moments, and can not be used to draw inference about possible breaking dates in case of discrete shifts. Using Monte-Carlo simulations, we have showed that the test was powerful against a broad alternative including for instance *i*) Discrete structural breaks, *ii*) Smooth transition, *iii*) time-varying moments, *iv*) probability-driven breaks, i.e. Markov switching models and *v*) GARCH and Stochastic Volatility Models. Nevertheless, it needs to be implemented sequentially: First test for the time-

homogeneity of the first moment, and if not rejected, for the time-homogeneity of the second moment. The procedure can also be used as a specification test in a linear or non-linear model to check the iid assumption for the residuals, and has its place in the econometrician toolbox.

There is an avenue for further researches using orthogonal polynomials in this field. One possible extension would be altering the procedure to detect the breaking dates and/or the different regimes.

## References

- [1] Altissimo, F. and V. Corradi (2003), Strong Rules for Detecting the Number of Breaks in a Time Series, *Journal of Econometrics* 117, p. 207-244.
- [2] Akaike, H. (1974), Stochastic Theory of Minimal Realization, I.E.E.E. Trans on Auto. Control, p. 667-674.
- [3] Andrews, W.K. (1993), Tests for Parameter Instability and Structural Change With Unknown Change Point, *Econometrica* 61, p. 821-856.
- [4] Andrews, W.K. and W. Ploberger, (1994), Optimal Tests when a Nuisance Parameter is Present only Under the Alternative, *Econometrica* 62, p. 1383-1414.
- [5] Andrews, W.K., Lee, I. and W. Ploberger (1996), Optimal Change Point Tests for Normal Linear Regression, *Journal of Econometrics* 70, p. 9-38.
- [6] Bai, J. (1999), Likelihood Ratio Tests for Multiple Structural Changes, *Journal of Econometrics* 91, p. 299-323.
- [7] Bai, J. and P. Perron (1998), Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica* 66, p. 47-78.
- [8] Bai, J. and P. Perron (2003), Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics* 18, p. 1-22.
- [9] Bai, J. and P. Perron (2005), Multiple Structural Change Model: A Simulation Analysis. In *Econometric Essays*, D. Corbea, S. Durlauf and B.E. Hansen (eds.), Cambridge University Press.
- [10] Burnham, K. P., and D. R. Anderson (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. Springer-Verlag.
- [11] Charfeddine, L. and D. Guegan (2011), Which is the Best Model for the US Inflation Rate : A Structural Changes Model or a Long Memory Process ? to appear in *Journal of Applied Econometrics*.
- [12] Davison, A.C. and D.V. Hinkley (1998), *Bootstrap Methods and Their Application*, Cambridge University Press, Cambridge.

- [13] Diebold, F.X. and A. Inoue (2001), Long Memory and Regime Switching, *Journal of Econometrics* 105, p. 131-159
- [14] Engle, R.F. and A.D. Smith (1999), Stochastic Permanent Breaks, *The Review of Economics and Statistics* 81, p. 553-574,
- [15] Ghysels, E., Harvey, A.C. and E. Renault (1996), Stochastic Volatility. *In* C.R. Rao and G.S. Maddala (Eds.), *Statistical Method in Finance*, p. 119-191, Amesterdam: North-Holland.
- [16] Guégan, D. (2010), Non-Stationary Samples and Meta-Distribution, *in* ISI Platinum Jubilee Volume: Statistical Sciences and Interdisciplinary research, Basu, A., Samanta, T. and A SenGupta (eds.) ICSPRAR World Scientific Review, in press.
- [17] Hamilton, J. (1990), Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics* 45.
- [18] Heracleous, M.S., Koutris, A and A. Spanos (2008), Testing for Nonstationarity Using Maximum Entropy Resampling: A Misspecification Testing Perspective, *Econometric Reviews* 27, p. 363-384.
- [19] Granger, C. and C. Starica (2005), Nonstationarities in Stock Returns, *The Review of Economics and Statistics* 87, p. 503-522.
- [20] Jacquier, E., Polson, N.G. and P.E. Rossi (1994), Bayesian Analysis of Stochastic Volatility Models (with discussion), *Journal of Business and Economic Statistics* 12, p. 371-417.
- [21] Koutris, A., (2005), Testing for Structural Change: Evaluation of the Current Methodologies, a Misspecification Testing Perspective and Applications, Ph.D. Dissertation, Virginia Polytechnic Institute and State University.
- [22] MacNeill, I.B. (1978), Properties of Sequences of Partial Sums of Polynomial Regression Residuals With Applications to Tests for Change of Regression at Unknown Times, *The Annals of Statistics* 6, p. 422-433.
- [23] McQuarrie, A.D.R and C.-L. Tsai (1998), *Regression and Time Series Model Selection*, World Scientific.

- [24] Perron, P. (1989), The Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Econometrica* 57, p. 1361-1401.
- [25] Perron, P. (1991), A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series, Research Memorandum No. 363, Econometric Research Program, Princeton University.
- [26] Perron, P. (2005), Dealing with Structural Breaks, *in* Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory.
- [27] Spanos, A. (1999), Probability Theory and Statistical Inference, Cambridge University Press, Cambridge.
- [28] Tang, S. M. and I.B. MacNeill (1993), The effect of Serial Correlation on Tests for Parameter Change at Unknown Time, *Annals of Statistics* 21, p. 552-575.
- [29] Vinod, H.D. (2004), Ranking Mutual Funds Using Unconventional Utility Theory and Stochastic Dominance, *Journal of Empirical Finance* 11, p. 353-377.
- [30] Vinod, H.D. (2006), Maximum Entropy Ensembles for Time Series Inference in Economics, *Journal of Asian Economics* 17, p. 955-978.

## A Appendix

Table 5: Number of times the correct polynomial degree is detected using the corresponding information criterion

$T$	$P(k = 0)$	$P(k \leq 1)$	$P(k > 1)$
<i>AIC</i>			
50	0.658	0.752	0.248
100	0.691	0.802	0.198
150	0.707	0.811	0.189
200	0.695	0.815	0.185
<i>AICc</i>			
50	0.751	0.852	0.148
100	0.735	0.842	0.158
150	0.738	0.841	0.159
200	0.714	0.832	0.168
<i>AICu</i>			
50	0.894	0.962	0.038
100	0.888	0.962	0.038
150	0.886	0.961	0.039
200	0.872	0.952	0.048

In this appendix, we analyze the relative performance of three different information criteria to select the correct degree  $k$  of the orthogonal Bernstein polynomial under the null. The three indicators are the *AIC* (Akaike [1974]), the *AICc* and the *AICu* (McQuarrie and Tsai [1998], Burnham and Anderson [2002]), the *AIC* and *AICc* being defined as:

$$AIC = 2k_1 + T \log(\mathbf{v}'\mathbf{v}) \quad (12)$$

$$AICc = \log\left(\frac{\mathbf{v}'\mathbf{v}}{T}\right) + \frac{T + k_1}{T - k_1 - 2} \quad (13)$$

and *AICu* by (5)

To perform such a task, we run Monte-Carlo simulations. Our DGP is as follows:

1. Generate a series  $\{\varepsilon_t\}_{t=1}^T$ , where  $\varepsilon_t \sim N(0, 1)$ ,
2. Estimate (3) for various values of  $k \in [0, 15]$ ,
3. For each model, compute *AIC*, *AICc* and *AICu*, and find the  $k$  minimizing the corresponding criterion,

4. Repeat steps 1 to 3, 1000 times.

Table 5 reports for different sample sizes, the probabilities for each criterion to return  $k = 0$ ,  $k \leq 1$  and  $k > 1$ , when the correct value is  $k = 0$ , corresponding to a constant signal. Note that  $k = 1$  will return a linear trend. Clearly, using *AIC* and to a less extent *AICc* results in overestimating the polynomial degree, whatever the sample size is. For instance, basing the decision rule on *AIC* will return the correct information in only 65.8 % of the cases for  $T = 50$ , whereas for *AICu* the correct information is returned in 89.4 % of the cases. Also, the number of times the *AIC* returns  $k > 1$  seems quite large, particularly for small samples.

Since our DGP is based on  $H_0$ , basing a stationarity test on the  $k$  minimizing the *AICu*, i.e.  $H_0 : k = 0$  against  $H_A : k \geq 1$  will return a very simple and powerful test, even if not based on statistical inference.