OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 74, 2 (2012) 0305-9049
doi: 10.1111/j.1468-0084.2011.00645.x

Testing for Fractional Integration Versus Short Memory with Structural Breaks*

Laura Mayoral

Institute for Economic Analysis, (CSIC) and Barcelona GSE, Campus UAB-08193 Bellaterra, Barcelona, Spain (e-mail: laura.mayoral@iae.csic.es)

Abstract

Although it is commonly accepted that most macroeconomic variables are non-stationary, it is often difficult to identify the source of the non-stationarity. Integrated processes and short-memory models with trending components, possibly affected by structural breaks, imply similar features in the data and, accordingly, are hard to distinguish. The goal of this article is to extend the classical testing framework of I(1) versus I(0) + trends and/or breaks by considering a more general class of models under the null hypothesis: fractionally integrated (FI) processes. The asymptotic properties of the proposed tests are derived and it is shown that they are very well-behaved in finite samples. An illustration using US inflation data is also provided.

I. Introduction

A standard practice in most macroeconomic applications is to test whether the trend component of a variable is best represented as stochastic or deterministic. Typically, the stochastic trend is characterized as a unit root process with a drift while the deterministic one is represented as the sum of a stochastic short-memory component and some deterministic trends. Perron (1989) contributed to this literature by showing that standard unit root tests could lead to erroneous conclusions if the true data generating process (DGP) was a short-memory –$I(0)$– process containing breaks in the deterministic components. This seminal contribution was the starting point of a myriad of articles on the problem of distinguishing between $I(1)$ and $I(0)+$ breaks processes.

Nevertheless, unit root processes are a very particular class within the group of integrated processes. There is substantial empirical evidence showing that the behaviour of many macroeconomic variables can be better captured by fractional as opposed to integer integration orders, see for instance Haubrich (1993), Michelacci and Zaffaroni (2000) and Mayoral (2006) among others.[1] There are also theoretical underpinnings that justify the

*I would like to thank Anindya Banerjee, two anonymous referees and participants in seminars at LSE, CORE and UB for valuable comments and suggestions. Financial support from the Spanish Government CICYT project n^o SEJ2006-00369, the BGSE Research Network and the Generalitat de Catalunya is gratefully acknowledged. The usual disclaimer applies.
JEL Classification numbers: C22, C12.

[1] See also Henry and Zaffaroni (2002) for a survey of empirical applications in the areas of macroeconomics and finance.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford, 2011. Published by Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

existence of fractional roots in macroeconomic data. They are based on the results of Robinson (1978) and Granger (1980) that show how the aggregation of dynamic heterogeneous processes can give rise to fractional integration. Since the existence of this type of behaviour at the disaggregate level is widely documented, fractionally integrated (FI) processes at the aggregate level are likely to arise in practice. FI models encompass the traditional I(0)-I(1) setup but also offer other interesting possibilities to model the persistence of shocks.

Not surprisingly, it is also difficult to provide an unambiguous answer as to whether a process is best represented as fractionally integrated or as I(0) plus some deterministic components, possibly perturbed by sudden changes, since a similar identification problem as in the I(1) versus I(0) + breaks case is found here. The issue of detecting patterns similar to those of an FI process when the DGP is short-memory containing deterministic terms and breaks has been widely analyzed (see Bhattacharya, Gupta and Waymire, 1983; Teverovsky and Taqqu, 1997; Lobato and Savin, 1998; Diebold and Inoue, 2001; Giraitis, Kokoszka and Lerpus, 2001; Granger and Hyung, 2004; Davidson and Sibbertsen, 2005; Perron and Qu, 2006, etc.).[2] It is generally concluded that the use of standard techniques devised for FI processes could lead to the detection of spurious FI behaviour when applied to short-memory processes containing trends and/or breaks. The opposite effect is also well documented, that is, conventional procedures for detecting and dating structural changes tend to find spurious breaks, usually at the middle of the sample when, in fact there is only fractional integration in the data (see Nunes; Kuan and Newbold, 1995; Hsu, 2001; Krämer and Sibbertsen, 2002).

There is increasing interest in developing techniques to distinguish between fractional integration and I(0) models containing trends and breaks.[3] Most of the papers in this area consider the problem of testing for (stationary) long-memory versus a weakly dependent series with smooth trended components or breaks in the mean (see Künsch, 1986; Heyde and Dai, 1996; Iacone, 2005; Berkes *et al.*, 2006; Giraitis, Leipus and Philippe, 2006; Ohanissian, Russel and Tsay, 2008). Although this problem is of genuine interest, many macroeconomic variables seem to display non-stationary orders of integration, non-smooth trends or breaks in the trending components and, therefore, the above-mentioned techniques would not be useful. Surprisingly, the problem of testing for non-stationary fractional integration versus short-memory + trends (possibly containing breaks) has been less studied. Shimotsu (2006a) presents two techniques to distinguish between FI(d) models, with $d \in (-1/2, 2)$, and other data generating processes (DGP) that can generate spurious persistence. However, these techniques cannot be applied to trending data since the FI models considered in that paper cannot accommodate trends.

The goal of this article is to develop a simple testing device that is able to determine whether the non-stationarity observed in the data is due to strongly persistent shocks, modelled as a non-stationary fractionally integrated variable, or to the existence of deterministic trends, possibly containing breaks, in an otherwise stationary process. Thus, structural breaks will only be allowed under the alternative hypothesis. This approach is similar

[2]Davidson and Sibbertsen also point out that cross-sectional aggregation of a fairly general class of nonlinear processes produces a model that not only has the same correlation patterns as FI processes but is also observationally equivalent to FI, in the sense that the aggregated model is linear and converges to fractional Brownian Motion.

[3]See Banerjee and Urga (2005) and the references therein.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

to that in Zivot and Andrews (1992), Banerjee, Lumsdaine and Stock, (1992) and Perron (1997) in the classical $I(1)$ versus $I(0)+$ structural breaks framework. The test is based on the likelihood ratio principle and the statistic is given by the ratio of two sums of squared residuals, computed under the alternative and the null hypotheses.

The structure of the article is as follows. Section II presents the model and the hypotheses of interest. Section III analyzes the problem of testing for FI vs. $I(0)+$ trended regressors. This framework is extended in section IV by allowing for the presence of breaks (occurring at an unknown time) in the deterministic components. Section V presents the results of some Monte Carlo simulations that evaluate the finite-sample performance of the test introduced in section IV. An application using US inflation data is reported in section VI. Section VII draws some final conclusions. All proofs are gathered in appendix A while critical values for the proposed tests are presented in appendix B.

In what follows, non-stationary FI(d) processes are defined as the cumulation of stationary $FI(d-1)$ variables. The following conventional notation is adopted throughout the article: L is the lag operator, $\Delta=(1-L)$, $\Gamma(.)$ is the gamma function, $B_d(.)$ denotes (standard) fractional Brownian motion (fBM) corresponding to the limit distribution of the standardized partial sums of stationary FI(d) processes,[4] $\xrightarrow{w}$ denotes weak convergence, $\xrightarrow{p}$ means convergence in probability and '$\equiv$' denotes equivalence in distribution. All integrals are taken with respect to the Lebesgue measure.

II. The model and the hypotheses

This section presents the basic framework that will be considered in the article along with the null and the alternative hypotheses of interest.

To capture the features observed in most macroeconomic variables, that is, the existence of trends and/or very persistent innovations, we consider two families of competing models: non-stationary fractionally integrated processes and processes that are the sum of a deterministic component, with parameters that may change over the sample, and a short-memory term. These two classes of models can be nested in the following equation. We assume that the data $y_1,\ldots,y_T$ is generated as

$$y_t=\beta' Z_t+\delta' V_t(\omega)+x_t, \quad t=1,2,\ldots, \tag{1}$$

$$(1-L)^d X_t=u_t, \tag{2}$$

and

$$V_t(\omega)=\begin{cases} Z_{t-T_B} & t>T_B, \\ 0 & \text{otherwise,} \end{cases}$$

where Z_t is a vector of deterministic components given by $Z_t=(1)$ or $Z_t=(1\ t)$, corresponding to the cases where a constant or a constant and a linear trend are included in the model. The parameter T_B represents the time when the break occurs and $\omega=T_B/T\in\Omega=[\omega_L,\omega_H]\subset(0,1)$ denotes the (relative) timing of this change. The processes y_t and Z_t are observable, β and δ are vectors of unknown parameters and u_t is a weakly dependent

[4]According to the notation introduced in Marinucci and Robinson (1999), $B_d(.)$ is a type I fractional Brownian motion.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

process. If δ is different from zero, $V_t(\omega)$ captures the existence of changes in the coefficients associated with the deterministic components Z_t.

Under the null hypothesis, y_t is considered to be a non-stationary FI(d) process with no breaks, so it is assumed that $d = d_0 \in (0.5, 1.5)$ and $\delta = 0$.

The definition of non-stationary FI processes is similar to that employed, for instance, in Velasco and Robinson (2000), that is,

$$(1-L)X_t = a_t, \quad t > 0$$
$$(1-L)^{d_0-1} a_t = u_t, \quad t = 0, \pm 1, \pm 2, \ldots,$$

where a_t is a stationary FI($d_0 - 1$) process and the filter $(1-L)^{\zeta}$ is defined as $(1-L)^{\zeta} = \sum_{i=0}^{\infty} \pi_i(\zeta) L^i$ with

$$\pi_i(\zeta) = \frac{\Gamma(i-\zeta)}{\Gamma(-\zeta)\Gamma(i+1)}. \tag{3}$$

In some economic problems, the order of integration under the null hypothesis, d_0, is known, for instance, in the popular unit root case where d_0 is set equal to 1. Nevertheless, in most cases this value is unknown. Accordingly, the null hypothesis will be simple or composite, that is,

$$H_0 : d = d_0, \quad \delta = 0 \text{ for some } d_0 \in (0.5, 1.5), \tag{4}$$

$$\text{or} \quad H_0' : d \in D_0, \quad \delta = 0, \quad D_0 = [\underline{d}, \bar{d}] \subset (0.5, 1.5), \tag{5}$$

corresponding to the cases where d_0 is known or unknown, respectively.

Under the alternative hypothesis, the process y_t is assumed to be the sum of some deterministic components, whose parameter values may change over the sample, and a short-memory term. Therefore, $d = 0$ is imposed. The case where no breaks in the coefficients of the deterministic components ($\delta = 0$) are allowed is analysed in section III. Section IV deals with the case where δ is (partially or totally) unconstrained, thus allowing for the possibility of a break occurring at an unknown time T_B. More specifically, the alternative hypothesis is formulated as

$$H_1 : d = 0, \quad \delta = 0, \tag{6}$$

$$\text{or} \quad H_1' : d = 0, \quad \delta \text{ unconstrained (totally or partially).} \tag{7}$$

When changes in the parameter values are allowed, attention is exclusively focused on the case where, at most, a single break exists. An extension to a multiple-change environment can be implemented along the lines of Bai (1999) and Bai and Perron (1998).

The following condition will be adopted throughout the article.

Condition 1. The stationary sequence $\{u_t\}$ admits a moving-average representation $u_t = \Psi(L)\varepsilon_t$, where $\Psi(z) = \sum_{j=0}^{\infty} \psi_j z^j$, and the coefficients ψ_j are such that $\sum_{j=0}^{\infty} j|\psi_j| < \infty$. The sequence $\{\varepsilon_t\}$ is an unobserved i.i.d. zero-mean process with unknown variance equal to σ^2. We further assume that $E(uu') = \Sigma$, where $u = (u_1, \ldots, u_T)'$ is a $T \times 1$ vector and Σ is a positive-definite variance-covariance matrix.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

III. Preliminaries: testing fractional integration versus I(0) + trends

This section explores the problem of testing whether the trend component of a process is *stochastic* or *deterministic*. Typically, the former is represented by a unit root process with a drift. This section considers a broader category of models, non-stationary fractionally integrated processes, which nest the unit root class as a particular case. In addition, it establishes the basic testing framework that will be used throughout the article.

Consider the setup introduced in section II and the following set of conditions. These assumptions are only intended to simplify the exposition and motivate the testing strategy and will be weakened soon.

Condition 2. δ is known and equal to zero.
Condition 3. d_0, the degree of integration under H_0, is known.
Condition 4. $u_t = \varepsilon_t$ is an i.i.d. process.
Condition 5. The process $\{\varepsilon_t\}$ is Gaussian.

Condition 2 implies that the coefficients of the deterministic components are stable throughout the sample. Condition 3 indicates that the null hypothesis H_0 is that defined in equation (4). Condition 4 strengthens condition 1 by assuming that u_t is i.i.d. Condition 5 will allow us to give a likelihood-ratio test interpretation to the test proposed in this section. While condition 2 will be maintained throughout this section, conditions 3–5 will be relaxed shortly.

Under conditions 1–5, the problem of testing H_0 against H_1 is straightforward since it is simply a test of a simple hypothesis. Notice that, as δ is assumed to be equal to zero, the null and alternative hypotheses only differ in the value of one parameter: the degree of integration of the stochastic component x_t. Thus, a natural way of testing H_0 against H_1 would be by means of a likelihood ratio test, which, by the Neyman-Pearson lemma, would be the most powerful invariant test. Minus two times the log likelihood is (except for an additive constant) given by

$$L(d_i,\beta)|_{H_i} = \sigma^{-2}(\Delta^{d_i}y - (\Delta^{d_i}Z)\beta)'(\Delta^{d_i}y - (\Delta^{d_i}Z)\beta), \quad i \in \{0,1\},$$

where d_0 and $d_1(=0)$ are the orders of integration under H_0 and H_1, respectively, $\Delta^{d_0}y = (\Delta^{d_0}y_2,\ldots,\Delta^{d_0}y_T)'$, $\Delta^{d_1}y = y = (y_1,\ldots,y_T)'$, $(\Delta^{d_0}Z) = (\Delta^{d_0}Z_2,\ldots,\Delta^{d_0}Z_T)'$ and $\Delta^{d_1}Z = Z = (Z_1,Z_2,\ldots,Z_T)'$.

From the developments in Lehmann (1959), the most powerful invariant test of $d = d_0$ vs. $d = 0$ rejects H_0 for small values of

$$\min_{\beta} L(d,\beta)|_{H_0} - \min_{\beta} L(d,\beta)|_{H_1}.$$

The test statistic is the difference of the sum of squared residuals from two constrained Ordinary least squares (OLS) regressions, one imposing $d = d_0$ and the other $d = 0$. Rearranging terms, it follows that the critical region of the most powerful invariant (MPI) test can be written as

$$R(d_0) = T^{1-2d_0}\frac{(y - Z\tilde{\beta})'(y - Z\tilde{\beta})}{(\Delta^{d_0}y - \Delta^{d_0}Z\hat{\beta})'(\Delta^{d_0}y - \Delta^{d_0}Z\hat{\beta})} < k_T \qquad (8)$$

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

for some k_T, where $\hat{\beta}$ and $\tilde{\beta}$ denote the OLS estimators of β under H_0 and H_1, respectively.[5]

The assumptions imposed above are too unrealistic to be useful in applications. It is not difficult, however, to modify the test in equation (8) so that it can still be employed for the same purposes when conditions 2–5 are relaxed. Although optimality is lost under more general assumptions, section V will show that the test still performs well in finite samples.

If condition 4 is dropped, implying that Σ is unknown and, in general, different from the identity matrix, the ratio $R(d_0)$ will have a limiting distribution depending on the error variances and covariances. However, it is easy to construct a modified statistic that does produce a valid large-sample test.

The simplest way to proceed will be to use a semi-parametric correction that deals with the correlation structure of Σ in such a way that the corrected statistic has the same asymptotic distribution as that described in theorem 1. A feasible statistic can be obtained as

$$R^f(d_0) = T^{1-2d_0}\left(\frac{\hat{\lambda}^2}{\hat{\gamma}_0}\right)^{-1}\frac{(y-Z\tilde{\beta})'(y-Z\tilde{\beta})}{(\Delta^{d_0}y-\Delta^{d_0}Z\hat{\beta})'(\Delta^{d_0}y-\Delta^{d_0}Z\hat{\beta})}, \tag{9}$$

where $\hat{\lambda}$ and $\hat{\gamma}_0$ are consistent estimators of the quantities $\lambda = \sigma\Psi(1)$ and $\gamma_0 = \sigma^2\sum_{i=0}^{\infty}\psi_j^2$, respectively. The factor $(\hat{\lambda}^2/\hat{\gamma}_0)^{-1}$ needed to construct equation (9) can be estimated by non-parametric kernel techniques, analogous to those used in the estimation of the spectral density. Details on the estimation of this quantity will be provided in section V.

In many applications, the order of integration d_0 is unknown and, therefore, the test $R^f(d_0)$ cannot be computed either. In these cases, attention will be focused on the composite null hypothesis H_0', that is, x_t is an FI(d) process with $d \in D_0 = [\underline{d}, \bar{d}] \subset (0.5, 1.5)$ under H_0'. We suggest to use the statistic $R^f(\hat{d}_T)$, which is defined as

$$R^f\left(\hat{d}_T\right) = T^{1-2\hat{d}_T}\left(\frac{\hat{\lambda}^2}{\hat{\gamma}_0}\right)^{-1}\frac{(y-Z\tilde{\beta})'(y-Z\tilde{\beta})}{(\Delta^{\hat{d}_T}y-\Delta^{\hat{d}_T}Z\hat{\beta})'(\Delta^{\hat{d}_T}y-\Delta^{\hat{d}_T}Z\hat{\beta})}, \tag{10}$$

and is obtained by replacing d_0 by $\hat{d}_T$ in equation (9), where $\hat{d}_T$ is a T^{κ}-consistent estimator of d_0, with $\kappa > 0$.

Finally, in applications, the filter Δ^d cannot be directly applied because an infinite number of data points would be needed. Instead, to compute the relevant statistics, we apply a truncated filter, Δ_T^d, that yields

$$\Delta_T^d x_t = \Delta_T^{d-1}\Delta x_t = \sum_{i=0}^{t-1}\pi_i(d-1)a_{t-i}, \quad t = 2, \ldots T$$
$$= u_t - \sum_{i=t}^{\infty}\pi_i(d-1)a_{t-i}.$$

[5]The statistic in equation (8) is similar to the Von-Neumann ratio proposed in the framework of efficient unit root tests (see Sargan and Bhargava, 1983 and Bhargava, 1986). It also bears some similarities with the non-parametric variance ratio unit root test introduced by Breitung (2002).

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

Application of this filter to Z yields

$$\Delta_T^d Z_t = (\Delta_T^d) Z_t = (\tau_t(d) \quad \tau_t(d-1)),$$

where $\tau_t(\zeta) = \sum_{i=0}^{t-1} \pi_i(\zeta)$.[6] To simplify the notation, in the following we will denote the truncated filter by Δ^d.

Theorem 1 presents the asymptotic distributions of $R^f(d_0)$ and $R^f(\hat{d}_T)$ under assumptions 1–3 and 1 and 2, respectively.

Theorem 1. Suppose $\{y_t\}$ is generated by models (1) and (2) with $d = d_0 \in D_0$. Then,

(a) under conditions 1–3 and H_0, it follows that

$$R^f(d_0) \xrightarrow{w} \int_0^1 (B_{d_0}^c(r))^2 \, \mathrm{d}r,$$

where $B_{d_0}^c(r)$, with $c \in \{\mu, \tau\}$, is the L_2 projection residual from the continuous time regression,

$$B_{d_0}(r) = \hat{\beta}' z^c(r) + B_{d_0}^c(r), \; c \in \{\mu, \; \tau\},$$

where $z^\mu(r) = (1)$ and $z^\tau(r) = (1, \; r)$, according to whether Z_t contains a constant or a constant and a linear trend, and $\hat{\beta}$ solve

$$\min_\beta \int_0^1 |B_{d_0(r)} - \beta' z^i(r)|^2 \, \mathrm{d}r.$$

(b) Under conditions 1 and 2 and H_0, if $\hat{d}_T$ is a T^κ-consistent estimator of d_0, with $\kappa > 0$ and $\hat{\mathrm{d}}_T > 0.5$, it follows that

$$R^f(d_0) - R^f(\hat{d}_T) = o_p(1).$$

Several methods for obtaining estimates of d suit the framework considered in this article well. For instance, the semi-parametric exact local Whittle estimator proposed in Shimotsu and Phillips (2005) and Shimotsu (2006b) provides T^κ-consistent estimates of d for non-stationary FI(d) models containing a linear trend.

For large enough T, any consistent estimator of d_0 will always be larger than 0.5 under H_0. However, in finite samples, estimates smaller than this quantity can be obtained. Thus, in applications we propose using the following simple rule for choosing $\hat{d}_T$: define $\hat{d}_T = \max\{d_L, \; \hat{d}\}$, where $d_L = 0.5 + \epsilon$, for some small $\epsilon > 0$ and $\hat{d}$ is the estimated value of d_0 obtained by applying one of the methods available in the literature.

Finally, the following theorem states that the test proposed in this section is consistent.

Theorem 2. Suppose $\{y_t\}$ is generated by models (1) and (2), with $d = \delta = 0$. Then, the test based on the statistic $R^f(d_0)$ ($\mathrm{R}^f(\hat{d}_T)$, for some $\hat{d}_T > 0.5$) rejects the hypothesis of H_0 (H_0') with a probability approaching 1.

It is easy to check that the test is also consistent if the true process is FI(d^*), with $d^* < 0.5$. In this case, both $(y - Z\tilde{\beta})'(y - Z\tilde{\beta})$ and $(\Delta^{\hat{d}_T} y - \Delta^{\hat{d}_T} Z\hat{\beta})'(\Delta^{\hat{d}_T} y - \Delta^{\hat{d}_T} Z\hat{\beta})$ are $O_p(T)$, since they contain the sum of squared residuals from two stationary processes, so the ratio is

[6] Notice that $t = (1-L)^{-1} 1_{(t>0)}$, where $1_{(,)}$ denotes the indicator function.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

$O_p(1)$ (and strictly greater than zero). Then, the statistic is the product of T^{1-2d_0}, where d_0 is the value of d under the null hypothesis, and a term that is $O_p(1)$. It follows that the product tends to zero at a rate T^{1-2d_0}, implying that the probability of rejecting H_0 tends to 1. A similar behaviour has also been observed in other unit root tests that tend to reject their corresponding null hypotheses ($d=1$ or $d=0$) when the true DGP is fractionally integrated of order $d \in (0,1)$ as shown by Diebold and Rudebusch (1991), and Lee and Schmidt (1996), respectively.

Thus, it is convenient to stress that the rejection of the null hypothesis does not imply the acceptance of the alternative since other DGPs than the one postulated under H_1 can also cause rejection of H_0, as described above. If the null hypotheses H_0' is ruled out, one can conclude that the DGP is not a *non-stationary* FI process. Other techniques developed for distinguishing between *stationary* FI models and some types of trends and breaks can be applied at this point to discard the possibility that the data is a stationary FI process, see Berkes *et al.* (2006), Giraitis *et al.* (2006), etc.

IV. Testing fractional integration versus I(0)+ breaking trends

Perron (1989) was the first to point out that tests of stochastic versus deterministic trends tend to favour the former hypothesis in cases where the trend is, in fact, deterministic but contains sudden changes in its parameter values. This section extends the procedure proposed in section III so it can accommodate structural breaks under the alternative hypothesis.

Let y_t be defined as in equations (1) and (2), and assume that condition 1 holds. A likelihood ratio test, in the spirit of Cox (1962), can still be implemented in this case, provided the candidate for break date, T_B, is known. In this case, the vector $V_t(\omega)$ is completely determined since $\omega = T_B/T$ is also known. Thus, following similar steps to those in section III, a feasible statistic for testing for H_0 versus H_1' could be constructed as

$$T^{1-2d_0}\left(\frac{\hat{\lambda}^2}{\hat{\gamma}_0}\right)^{-1}\frac{\min_{\beta,\delta}(y-Z\beta-V(\omega)\delta)'(y-Z\beta-V(\omega))}{\min_{\beta}(\Delta^{d_0}y-\Delta^{d_0}Z\beta)'(\Delta^{d_0}y-\Delta^{d_0}Z\beta)}, \tag{11}$$

where $V(\omega)=(V_1'(\omega),\ V_2'(\omega),\ldots,V(\omega)_T')'$ and $\hat{\lambda}^2$ and $\hat{\gamma}_0$ are consistent estimates under H_0 of $\lambda^2=\sigma^2\Psi(1)^2$ and $\gamma_0=\sigma^2\sum_{i=0}^{\infty}\psi_j^2$, respectively.

However, for the general case where the candidate for break date is unknown, the procedure needs to be modified slightly. Under H_0, ω is a nuisance parameter that is not identified. The usual procedure in these cases consists of, first, computing the feasible test equation (11) for a grid of values of $\omega \in \Omega$ and, then, computing a certain functional of these pointwise statistics (see for instance Andrews and Ploberger, 1994).

We follow this approach here by considering the infimum of a sequence of statistics computed for different values of $\omega \in \Omega \subset (0,1)$. Since considering the whole interval (0,1) would lead to tests with very low power, optimization is carried out in $\omega \in \Omega$, where $\Omega = [\omega_L, \omega_H]$ for some $0<\omega_L<\omega_H<1$. More specifically, we will use the restricted interval $\Omega = [0.15, 0.85]$, as suggested by Andrews (1993). A feasible statistic, $R_b^f(d_0)$, for testing H_0 versus H_1' can be computed as

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

$$R_b^f(d_0) = T^{1-2d_0}\left(\frac{\hat{\lambda}^2}{\hat{\gamma}_0}\right)^{-1} \inf_{\omega\in\Omega} \frac{\min_{\beta,\delta}(y - Z\beta - V(\omega)\delta)'(y - Z\beta - V(\omega)\delta)}{\min_{\beta}(\Delta^{d_0}y - \Delta^{d_0}Z\beta)'(\Delta^{d_0}y - \Delta^{d_0}Z\beta)}, \qquad (12)$$

where $\hat{\beta}$ and $(\tilde{\beta}, \tilde{\delta})$ are the OLS estimates of the parameters β and δ under H_0 and H_1', respectively. The null hypothesis will be rejected for small values of equation (12).

The asymptotic behaviour of $R_b^f(d_0)$ depends on the regressors included in Z_t but, in addition, on the terms that are allowed to break. Four different possibilities, corresponding to the cases considered by Perron (1989) and Zivot and Andrews (1992), have been analysed. Three of these models contain both a constant and a linear trend but differ in the coefficients that are allowed to vary over the sample: Model 1 allows for a break in the level of the series, Model 2 allows for a change in the rate of growth and, finally, Model 3 admits both changes. In addition, we also consider Model 0, where Z_t only contains a constant that is allowed to break once in the sample.[7] More specifically, denoting $\beta = (\beta_0,\ \beta_1)'$ and $\delta = (\delta_0, \delta_1)'$, the four models can be summarized as:

Model 0: β_0 unconstrained, $\beta_1 = 0$; δ_0 unconstrained, $\delta_1 = 0$.
Model 1: $\beta_0,\ \beta_1$ unconstrained; δ_0 unconstrained, $\delta_1 = 0$.
Model 2: $\beta_0,\ \beta_1$ unconstrained; $\delta_0 = 0,\ \delta_1$ unconstrained.
Model 3: $\beta_0,\ \beta_1$ unconstrained; $\delta_0,\ \delta_1$ unconstrained.

In order to compute the statistic $R_{b_i}^f(d_0)$, with $i \in \{0, 1, 2, 3\}$, corresponding to Models 0 to 3, respectively, the values of the constrained parameters are imposed and then, the remaining parameters are estimated by OLS as usual. Theorem 3 describes the large-sample properties of $R_{b_i}^f(d_0)$ for each of the four cases considered.

Theorem 3. Suppose $\{y_t\}$ is generated as in equations (1) and (2) with $d = d_0 > 1/2$. Then, under condition 1 and H_0, the asymptotic distribution of $R_{b_i}^f(d_0)$ for $i \in \{0, 1, 2, 3\}$ is given by

$$R_{b_i}^f(d_0) \xrightarrow{w} \inf_{\omega\in\Omega}\left(\int_0^1 (B_{d_0}^i(r,\omega))^2\,\mathrm{d}r\right),\ i \in \{0, 1, 2, 3\},$$

where $B_{d_0}^i(r,\omega)$ is the L_2 projection residual from the continuous time regressions

$$B_{d_0}(r) = \hat{\beta}' z^i(r) + \hat{\delta}' v^i(r,\omega) + B_{d_0}^i(r,\omega),\ i \in \{0, 1, 2, 3\},$$

where $v^0(r,\omega) = v^1(r,\omega) = (1_{(r>\omega)})$, $v^2(r,\omega) = ((r - \omega)1_{(r>\omega)})$; $v^3(r,\omega) = ((1_{(r>\omega)}) \times (r-\omega)1_{(r>\omega)})$; $z^0(r) = (1)$ and $z^j(r) = (1\ r)$ for $j \in \{1, 2, 3\}$; and $\hat{\beta}$ and $\hat{\delta}$ solve

$$\min_{\beta,\delta} \int_0^1 |B_{d_0(r)} - \beta' z^i(r) - \delta' v^i(r,\omega)|^2\,\mathrm{d}r,\ i \in \{0, 1, 2, 3\},$$

(See appendix A for details).

[7] This case may be of interest when modelling series that do not seem to display a trend, such as inflation or interest rates.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

H_0 will be rejected against H_1' for small values of $R^f_{b_i}(d_0)$. Critical values for each of the four models considered above have been obtained by Monte Carlo simulation and are presented in appendix B, Tables B3–B6.

When composite hypotheses such as H_0' are considered, the test-statistic defined in equation (12) can still be employed if d_0 is replaced by $\hat{d}_T$, a T^{κ}-consistent estimator of d_0 under the null hypothesis, with $\kappa > 0$ and $\hat{d}_T > 0.5$. Theorem 4 states that under these assumptions the statistic $R^f_{b_i}(\hat{d}_T)$ defined as

$$R^f_b(\hat{d}_T) = T^{1-2\hat{d}_T}\left(\frac{\hat{\lambda}^2}{\hat{\gamma}_0}\right)^{-1}\frac{\inf_{\omega\in\Omega}(y - Z\tilde{\beta} - V(\omega)\tilde{\delta})'(y - Z\tilde{\beta} - V(\omega)\tilde{\delta})}{(\Delta^{\hat{d}_T}y - \Delta^{\hat{d}_T}Z\hat{\beta})'(\Delta^{\hat{d}_T}y - \Delta^{\hat{d}_T}Z\hat{\beta})}, \tag{13}$$

has the same distribution as $R^f_{b_i}(d_0)$, for $i = \{0,\ 1,\ 2,\ 3\}$.

Theorem 4. Suppose $\{y_t\}$ is generated by models (1) and (2) and $d = d_0 > 0.5$. If $\hat{d}_T$ is a T^{κ}-consistent estimator of d_0, with $\kappa > 0$ and $\hat{d}_T > 0.5$ and condition 1 and H_0' hold, then

$$R^f_{b_i}(d_0) - R^f_{b_i}(\hat{d}_T) = o_p(1), \quad \text{for} \quad i = \{0,\ 1,\ 2,\ 3\}. \tag{14}$$

It follows that H_0' will be rejected against H_1' when the value of $R^f_{b_i}(\hat{d}_T)$ is smaller than the corresponding critical value.

Finally, the following theorem states the consistency of the proposed test.

Theorem 5. Let y_t be defined as in equations (1) and (2) with $d = 0$ and δ possibly different from zero. Then, the test based on the statistics $R^f_{b_i}(d_0)$ $(R^f_{b_i}(\hat{d}_T))$ rejects the hypothesis of $H_0(H_0')$ with a probability approaching 1.

V. Finite sample results

This section presents the results of some Monte Carlo experiments designed to illustrate the identification problem addressed in this article and to explore the finite-sample performance of the proposed techniques. Two sample sizes ($T = 100$ and $T = 400$) have been considered in all the experiments in this section. The number of replications in each experiment was set at 5,000.

To illustrate that a short-memory process with breaks can be easily confused with an FI(d) process, we have carried out a simple experiment. We have generated variables of the form $y_t = 10 + 0.5t + b_t(\omega) + \varepsilon_t$, where $t = 1, \ldots, T$, $\varepsilon_t \sim$ i.i.d. $N(0, 1)$, and $b_t(\omega) = (\delta_0 1_{(t>T_B)} +$
$\delta_1(t - T_B)1_{(t>T_B)})$ is the term containing the breaks. Different values of δ_0 and δ_1 have been considered, namely, $\delta_0 \in \{0, 1, 2, 4\}$ and $\delta_1 = \{0, 0.1, 0.2, 0.3\}$, and breaks occur at the middle of the sample. The semiparametric exact local Whittle estimator with detrending (Shimotsu, 2006b) has been applied to these processes to obtain estimates of d.[8] Table 1 reports the mean and the standard deviation of the estimates of d over the 5,000 replications, for different combinations of δ_0 and δ_1. When $\delta_0 = \delta_1 = 0$, accurate estimates of $d = 0$ are obtained. However, when breaks in the deterministic components are introduced,

[8]The Whittle estimator is known to have better properties than other techniques such as the R/S or the log-periodogram under some types of trends and breaks (see Iacone, 2005).

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

TABLE 1

*Mean and STD of $\hat{d}_T$ (FELW)**

$DGP\ (H_1): y_t = 10 + 0.5t + \delta_0 1_{(t>\omega T)} + \delta_1 (t - \omega T) 1_{(t>\omega T)} + \varepsilon_t; \varepsilon_t \sim iN(0,1); \omega = 0.5$										
	No break	*Model 1*			*Model 2*			*Model 3*		
δ_0=	0	1	2	4	0	0	0	1	2	4
δ_1=	0	0	0	0	0.1	0.2	0.3	0.1	0.2	0.3
$T = 100$	0.075	0.097	0.339	0.625	0.536	0.816	0.978	0.557	0.823	0.978
	(0.199)	(0.182)	(0.152)	(0.124)	(0.128)	(0.109)	(0.095)	(0.123)	(0.109)	(0.099)
$T = 400$	−0.03	0.206	0.415	0.636	0.923	1.114	1.223	0.923	1.111	1.209
	(0.102)	(0.085)	(0.073)	(0.056)	(0.042)	(0.041)	(0.042)	(0.043)	(0.042)	(0.042)

Notes: *Standard deviations in parentheses.

positive and large values of d are found. The simulations show that estimates are particularly sensitive to breaks in the trend component and that even a small change in the slope brings about a considerable increase in the estimated values of d. Moreover, the size of the break matters a lot: the larger the size, the higher the estimate of d.

Next, we explore the finite-sample properties of the technique proposed in section IV. To study the size of the test, ARFIMA $(1, d, 0)$ processes, for different values of d and ϕ (the autoregressive parameter), have been generated, more specifically, $d \in \{0.7, 0.9, 1.1\}$ and $\phi \in \{0, 0.3, 0.8\}$. Data has been generated as follows. Firstly, the vector of innovations has been obtained as $\varepsilon_t \sim$ i.i.d. $N(0,1)$, for $t = 1, \ldots, T + m$ and $m = 1{,}000$. AR(1) processes have been computed as $u_t = \phi u_{t-1} + \varepsilon_t$, $t = 2, \ldots T + m$, for the values of ϕ specified above. Next, fractionally integrated processes of order $(d - 1)$ have been simulated as

$$x_t^* = \Delta^{-(d-1)} u_t = \sum_{i=0}^{T+m-1} \pi_i (1 - d)\, u_{t-i}, \tag{15}$$

where the coefficients $\pi_i(.)$ are defined in equation (3). Finally, the first $m - 1$ observations are dropped and non-stationary $FI(d)$ processes are generated by cumulating x_t^*, that is

$$x_t = \sum_{j=m}^{m+t-1} x_j^*,\ t = 1, \ldots T.$$

The resulting data has been employed to compute the statistics $R^f_{b_i}(\hat{d}_T),\ i \in \{1, 2, 3\}$, defined in equation (13). Since a value of $d > 0.5$ is needed to perform the test, the value $\hat{d}_T$ has been chosen according to the following rule: $\hat{d}_T = \max\{0.5001, \hat{d}\}$, where $\hat{d}$ is the FELW estimator with detrending. The variance of u_t, γ_0, has been estimated under H_0 by the sample variance, $\hat{\gamma}_0 = T^{-1} \sum (\Delta^{\hat{d}_T} y_t - \Delta^{\hat{d}_T} Z \hat{\beta})^2$.

The parameter $\lambda^2 = \gamma_0 + 2 \sum_{i=1}^{\infty} \gamma_i$, where γ_i denotes the ith-autocovariance of u_t, is usually estimated as $\hat{\lambda}^2 = \hat{\gamma}_0 + 2 \sum_{i=1}^{q} \kappa(j/q) \hat{\gamma}_i$, where $\hat{\gamma}_i = T^{-1} \sum_{t=i+1}^{T} \hat{u}_t \hat{u}_{t-i}$, $\kappa(.)$ is a kernel and q is its bandwidth parameter. Several estimators of this quantity, that differ in the choice of $\kappa(.)$ and q, have been proposed.[9] A common conclusion from many Monte Carlo

[9]Although estimation of λ^2 is routinely carried out in the computation of unit root tests, Pötscher (2002) has shown that it belongs to the so-called ‘ill-posed’ estimation problems. Uniformly consistent estimators of λ^2 can only be achieved if very strong *a priori* assumptions on the set of feasible DGPs are considered. Otherwise, confidence sets may be too large to convey useful information on this quantity.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

experiments is that the choice of kernel is usually not very important for small sample results (Andrews, 1991; Cheung and Lai, 1997). However, correct choice of q has been shown to be critical. In order to investigate the sensitivity of the test to the modelling of short-run dependence, we have employed three different automatic bandwidth selection procedures, more specifically, those introduced in Andrews (1991), q_A, Andrews and Monahan (1992), q_{AM}, and Newey and West (1994), q_{NW}.[10] The Bartlett kernel has been used in all cases so that estimates of $\hat{\lambda}_k^2$ have obtained as

$$\hat{\lambda}_k^2 = \hat{\gamma}_0 + 2\sum_{i=1}^{q_k}(1 - i/(q_k+1))\hat{\gamma}_i, k = \{A,\ AM,\ NW\}. \tag{16}$$

Table 2 reports the corresponding rejection frequencies at the 5% significance level. For moderate sample sizes ($T = 100$), the test is generally slightly oversized when $\hat{\lambda}_{AM}^2$ is used to correct for short-term autocorrelation and undersized when either $\hat{\lambda}_A^2$ or $\hat{\lambda}_{NW}^2$ are employed. When larger samples sizes are considered ($T = 400$), the use of $\hat{\lambda}_{AM}^2$ yields a

TABLE 2

Size of $R^f_{b_i}(\hat{d}_T), i \in \{1,2,3\}$; *S.L.: 5%*

DGP $(H_0): y_t = \Delta^{-d_0}u_t; u_t = \varepsilon_t/(1-\phi L); \varepsilon_t \sim iN(0,1)$										
$T = 100$										
		$R^f_{b_1}(\hat{d}_T)$			$R^f_{b_2}(\hat{d}_T)$			$R^f_{b_3}(\hat{d}_T)$		
$\hat{\lambda}_k^2$		$\phi =$ 0(%)	$\phi =$ 0.3(%)	$\phi =$ 0.8(%)	$\phi =$ 0(%)	$\phi =$ 0.3(%)	$\phi =$ 0.8(%)	$\phi =$ 0(%)	$\phi =$ 0.3(%)	$\phi =$ 0.8(%)
$\hat{\lambda}_{AM}^2$	$d_0 = 0.7$	8.41	8.50	4.30	7.90	8.90	3.10	8.90	11.30	4.20
	$d_0 = 0.9$	6.70	8.40	4.70	6.71	8.70	2.50	7.60	11.20	3.35
	$d_0 = 1.1$	3.64	5.14	3.10	7.23	5.94	2.50	6.70	7.63	2.90
$\hat{\lambda}_A^2$	$d_0 = 0.7$	2.51	3.82	2.13	1.41	2.40	1.93	2.01	5.00	1.97
	$d_0 = 0.9$	2.25	2.76	2.00	1.80	2.50	1.50	2.20	3.10	1.79
	$d_0 = 1.1$	1.90	2.07	1.90	3.30	2.18	1.80	2.60	1.80	1.05
$\hat{\lambda}_{NW}^2$	$d_0 = 0.7$	1.14	1.92	0.81	1.00	1.60	0.70	1.90	2.07	1.21
	$d_0 = 0.9$	1.43	1.26	0.64	1.30	1.50	0.30	1.70	1.44	0.62
	$d_0 = 1.1$	0.52	0.37	0.24	1.90	1.00	0.60	1.90	0.69	0.61
$T = 400$										
$\hat{\lambda}_{AM}^2$	$d_0 = 0.7$	7.21	8.58	6.60	9.80	7.50	6.02	8.20	9.40	6.70
	$d_0 = 0.9$	4.93	4.21	5.84	5.00	4.60	5.23	4.70	5.10	5.60
	$d_0 = 1.1$	5.54	4.34	3.81	7.04	3.10	4.90	7.03	2.40	4.80
$\hat{\lambda}_A^2$	$d_0 = 0.7$	2.71	4.80	4.40	3.21	4.53	4.12	3.43	6.10	6.10
	$d_0 = 0.9$	2.37	1.83	2.30	1.99	2.69	1.30	1.70	1.60	1.41
	$d_0 = 1.1$	1.95	1.70	2.12	4.90	2.01	2.20	3.40	1.90	1.80
$\hat{\lambda}_{NW}^2$	$d_0 = 0.7$	6.60	2.70	7.10	6.80	2.70	5.20	7.30	2.60	6.00
	$d_0 = 0.9$	6.80	2.90	7.60	6.30	2.00	4.40	6.10	2.30	5.60
	$d_0 = 1.1$	4.80	2.30	7.70	6.90	2.10	6.70	6.90	2.30	8.10

[10] See Cheung and Lai (1997) for a description and a Monte Carlo comparison of these techniques in the unit root testing framework.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

TABLE 3

*Power of $R^f_{b_i}(\hat{d}_T)$, $i \in \{1, 2, 3\}$; S.L.: 5%**

$DGP\ (H_1): y_t = 10 + 0.5t + \delta_0 1_{(t>T_B)} + \delta_1 (t - T_B) 1_{(t>T_B)} + u_t;\ u_t = \varepsilon_t/(1-\phi L),\ \varepsilon_t \sim iN(0,1);\ \omega = 0.5$										
		$R^f_{b_1}(\hat{d}_T)$ (%)			$R^f_{b_2}(\hat{d}_T)$ (%)			$R^f_{b_3}(\hat{d}_T)$ (%)		
	$\delta_0 =$	1	2	4	0	0	0	1	2	4
	$\delta_1 =$	0	0	0	0.1	0.2	0.3	0.1	0.2	0.3
$\phi = 0$										
$\hat{\lambda}^2_{AM}$	$T = 100$	96.0	95.4	98.3	93.0	94.3	95.2	93.5	82.5	71.3
	$T = 400$	100	100	100	100	100	100	100	99.2	97.2
$\hat{\lambda}^2_{A}$	$T = 100$	89.1	85.4	84.6	91.1	90.7	91.2	86.3	72.3	69.3
	$T = 400$	100	100	100	100	100	94.4	95.2	89.4	90.1
$\hat{\lambda}^2_{NW}$	$T = 100$	74.1	69.2	71.2	78.1	73.3	70.1	64.2	59.2	45.9
	$T = 400$	93.6	92.3	89.1	96.2	96.4	92.1	84.2	87.6	75.5
$\phi = 0.3$										
$\hat{\lambda}^2_{AM}$	$T = 100$	81.2	83.9	84.3	83.1	86.1	83.1	84.3	71.2	61.0
	$T = 400$	100	100	100	98.9	99.2	92.4	89.3	83.2	76.5
$\hat{\lambda}^2_{A}$	$T = 100$	74.2	71.5	69.2	71.4	70.1	68.9	67.2	66.2	59.2
	$T = 400$	100	100	100	85.2	86.2	74.2	74.2	74.2	71.3
$\hat{\lambda}^2_{NW}$	$T = 100$	63.2	65.4	55.3	61.3	53.4	54.3	45.4	43.6	39.5
	$T = 400$	88.4	90.4	85.3	87.4	83.4	82.3	73.2	63.1	70.2
$\phi = 0.8$										
$\hat{\lambda}^2_{AM}$	$T = 100$	23.2	21.8	26.4	21.4	27.3	31.3	15.4	13.1	21.4
	$T = 400$	59.4	53.6	58.1	59.1	78.9	79.3	55.4	76.1	79.3
$\hat{\lambda}^2_{A}$	$T = 100$	21.3	24.1	19.8	24.2	29.1	29.3	25.1	11.3	11.0
	$T = 400$	61.4	59.8	53.1	51.9	67.4	59.2	51.3	67.1	69.8
$\hat{\lambda}^2_{NW}$	$T = 100$	16.3	18.2	20.2	17.6	19.5	20.0	15.1	11.4	8.8
	$T = 400$	45.3	51.3	49.3	46.2	51.2	45.3	46.1	48.2	45.2

Notes: (*) Size-adjusted power.

size that is close to nominal whereas using $\hat{\lambda}^2_A$ or $\hat{\lambda}^2_{NW}$ still yields tests that are, in general, undersized. Size distortions are generally highest for $\hat{\lambda}^2_{NW}$.

In order to study the power of the test, processes of the form

$$y_t = 10 + 0.5t + b_t(\omega) + u_t, \tag{17}$$

$$u_t = \frac{\varepsilon_t}{1 - \phi L},$$

have been generated, where $b_t(\omega)$ is defined as above, (δ_0, δ_1) take the values $\delta_0 \in \{0, 1, 2, 4\}$ and $\delta_1 = \{0, 0.1, 0.2, 0.3\}$, and $\phi = \{0, 0.3, 0.8\}$. Three different break locations have been considered, namely, $\omega = \{0.30, 0.50, 0.70\}$. For the sake of brevity, since the results were very similar, only figures computed with $\omega = 0.5$ are reported. By combining the values of δ_0 and δ_1, processes under Models 1–3 have been generated. The statistic $R^f_{b_i}(\hat{d}_T)$ has been computed on the data corresponding to Model i, $i \in \{1, 2, 3\}$ for different values of $\hat{\lambda}^2_k$, $k = \{A, AM, NW\}$. Rejection frequencies (computed at the 5% S.L.) are reported in Table 3.

For $T = 100$, power is, in general, high for small values of ϕ ($= 0$ or $= 0.3$) and it deteriorates when higher autocorrelation is introduced to u_t. The highest power is achieved

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

TABLE 4

Power of $R^f_{b_i}(\hat{d}_T)$ $i \in \{1,2,3\}$; *S.L.*: 5% *DGP* $(H_1): \Delta^{d_1} y_t = \varepsilon_t, d_1 < 0.5$

	$R^f_{b_1}(\hat{d}_T)$		$R^f_{b_2}(\hat{d}_T)$		$R^f_{b_3}(\hat{d}_T)$	
	$T=100$ (%)	$T=400$ (%)	$T=100$ (%)	$T=400$ (%)	$T=100$ (%)	$T=400$ (%)
$d_1=0.1$	96.2	100	95.9	100	92.3	100
$d_1=0.2$	79.5	100	78.2	100	74.1	100
$d_1=0.3$	54.5	100	51.1	97.9	47.8	99.3
$d_1=0.4$	31.0	72.5	16.8	69.0	26.0	69.1

when $\hat{\lambda}^2_{AM}$ is used.[11] Power ranges from 70% to 100% and from 61% to 100% for $\phi = 0$ and $\phi = 0.3$ respectively, and it decreases to 13%–31% for $\phi = 0.8$. However, if larger sample sizes are employed ($T = 400$), the power increases considerably, even for high values of ϕ, as shown in Table 3.

As mentioned earlier, the test is consistent when the alternative hypothesis is a stationary FI process. To evaluate the behaviour of the test under this type of alternative, we have generated stationary FI(d) processes as in equation (15), with $d = \{0.1, 0.2, 0.3, 0.4\}$, $\phi = 0$, and $T = \{100, 400\}$. $\hat{d}_T$ has been estimated as described above and we did not perform any short-term correlation correction. The corresponding rejection frequencies are reported in Table 4.

For moderate values of d (0.1, 0.2), power is around 70%–90% for $T = 100$ (around 100% for $T = 400$) but, not surprisingly, it decreases substantially when d approaches 0.5. For $d = 0.4$, power is around 20%–30% for $T = 100$ (70% for $T = 400$).

VI. Empirical illustration

We now apply the techniques introduced in this article to the analysis of inflation data. The study of the statistical properties of this variable has attracted a great deal of attention because it plays a central role in the design of monetary policy and has important implications for the behaviour of private agents. However, in spite of the large number of empirical and theoretical papers on this issue that have appeared recently, there is no consensus in the literature about the most appropriate way to model the inflation rate. On the one hand, there is abundant empirical evidence that post-war inflation in industrial countries exhibits high persistence, close to the unit root behaviour. The papers of Pivetta and Reis (2007) for the USA and O'Reilly and Whelan (2005) for the euro zone are two examples. On the other hand, some authors have argued that the above-mentioned results are very sensitive to the statistical techniques employed. They claim that the observed persistence may be due to the existence of unaccounted breaks, probably stemming from changes in the inflation targets of monetary authorities, different exchange rate regimes or shocks in key prices. For instance, Levin and Piger (2003) have found evidence of a break in the intercept of the inflation equation and, conditional on this break, they argue that inflation shows very low persistence. Finally, Cogley and Sargent (2001, 2005) claim that non-stationary (integrated) representations of inflation

[11] Similar results have been obtained for the Phillips–Perron unit root test (Cheung and Lai, 1997).

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

are implausible from an economic point of view, since they would imply an infinite asymptotic variance, which could never be optimal if the Central Bank's loss function includes the variance of inflation. Thus, they consider inflation to be a short-memory ($I(0)$) process.

The aim of this section is to shed further light on this controversy by applying the techniques developed in this article. There is both economic and statistical support for the hypothesis of FI in inflation. Gadea and Mayoral (2006) provide an economic explanation for the existence of fractional integration in inflation data. They consider a sticky price model as in Rotemberg (1987) and, by considering firms having heterogeneous costs of adjusting their prices, show that inflation can behave as an FI process. From an applied point of view, evidence in favour of FI behaviour in inflation has been reported in several papers (Baillie, Chung and Tieslan, 1996; Doornik and Ooms, 2004; Gadea and Mayoral, 2006, etc.).

The contradicting results described above could be explained if the inflation rate was an FI process. Although unit root tests are, in general, consistent against fractional alternatives, their finite-sample power is known to be small, (see Diebold and Rudebusch, 1991). This might account for the low rejection frequencies of the unit root hypothesis in this type of applications. On the other hand, if inflation is FI and standard techniques for detecting and dating breaks are employed, spurious breaks are likely to be detected. The opposite is also true: if the data contains structural breaks, FI maybe found spuriously too. Thus, we now test whether the high degree of persistence observed in inflation data is true and can be characterized using FI models (that encompass the I(1) case) or is spurious and induced by the existence of structural breaks in the deterministic components in an, otherwise, short-memory process.

To facilitate comparison with previous analysis, the same data set as in Pivetta and Reis (2007) has been employed: the price level, P_t, is measured through the seasonally-adjusted quarterly data on the GDP deflator from the first quarter of 1947 to the last quarter of 2003 (9 observations have been added with respect to their analysis). This data has been obtained from the Bureau of Economic Analysis. Then, inflation is computed as $\pi_t = 400 * \log(P_t/P_{t-1})$, that is, it is the quarterly continuously compounded annualized rate of change of the price level. Figure 1 presents a plot of this data.

To begin the analysis, Table 5 presents the results of some standard tests for unit roots. The first three columns contain the figures obtained by applying three different techniques that take the I(1) model as their null hypothesis: the Dickey–Fuller test with GLS detrending (Elliott, Rothenberg and Stock, 1996, DF-GLS henceforth), the MZ-GLS test (Ng and Perron, 2001) and the P-P test (Phillips and Perron, 1988). To correct for the short-term correlation, different techniques have been employed. The number of lags in the DF-GLS regression was chosen according to the modified AIC (M-AIC) and the modified SIC (M-SIC). Estimates of the spectral density needed for the calculation of both the MZ and the P-P tests were obtained using GLS-detrended autoregressive methods, where the number of lags in the autoregression was selected according to the M-AIC and the M-SIC. Similar techniques were also employed to compute the KPSS. A constant was included as the only deterministic regressor in the regression model.

The results greatly depend on the method employed to perform the short-term correction. When the M-AIC is used, neither the DF-GLS nor the MZ tests can reject the unit root

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

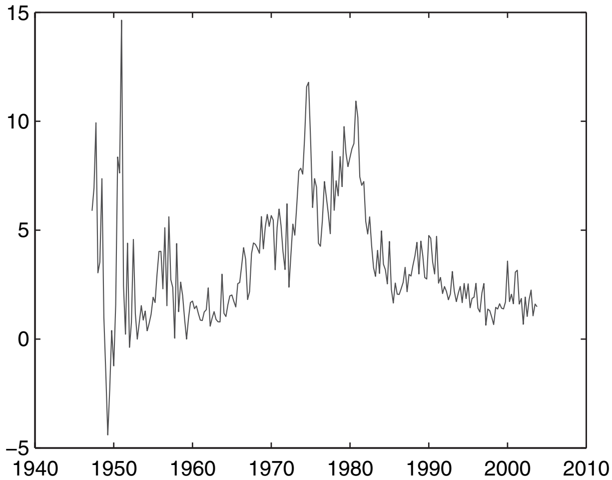

Figure 1. US inflation. 1947.1–2003.4

TABLE 5

Unit root tests on US Inflation

	DF-GLS		*N-P*		*P-P*		*KPSS*	
Inf. criterion	*m-aic*	*m-sic*	*m-aic*	*m-sic*	*m-aic*	*m-sic*	*m-aic*	*m-sic*
Value of the test	−1.31	−2.36*	−1.64	−12.99*	−5.56**	−4.24**	68.2**	11.6**

Notes: (*) (**) denote rejection at the 5% and 1% levels, respectively.

hypothesis. However, the P-P always rejects the I(1) null hypothesis as do the DF-GLS and the MZ tests computed with the M-SIC. On the other hand, the KPSS tests rejects the I(0) hypothesis for both the M-AIC and the M-SIC methods, although opposite results are found when other techniques, such as the Bartlett or the Parzen kernels, are employed to perform the short-term corrections.

These non-conclusive results could be consistent with the existence of both fractional integration and some types of structural breaks. Therefore, we have checked whether there is some evidence of FI in this data set. Table 6 presents the results of estimating d using different techniques: the Feasible Exact local Whittle (FELW, Shimotsu, 2006b), the Exact Maximum likelihood (EML, Sowell, 1992) and the Minimun Distance (MD, Mayoral, 2007) estimators.[12] Fractional values of d greater than 0.5 and relatively 'far' from both the $I(0)$ and the $I(1)$ hypotheses are found: the estimated values of d are around 0.6 for the three techniques employed. Furthermore, if tests of fractional versus integer integration (FI versus $d=0$ or $d=1$) based on confidence intervals around the estimated values of d are considered, the hypothesis of FI versus $d=0$ or $d=1$ cannot be rejected at the 5% significance level.

[12]Models were chosen according to the AIC.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

TABLE 6

*Estimated d**

	FELW	*EML*	*MD*
$\hat{d}_T$	0.62	0.61	0.58
	(0.11)	(0.14)	(0.12)

Note: *Standard errors in parentheses.

TABLE 7

$R^f_{b_i}(\hat{d}_T)$ *tests of* $FI(d)$ *vs* $I(0)$ *+Breaks*; $i=\{0,2,3\}$

		*Crit val.**	$\hat{\lambda}^2_A$	$\hat{\lambda}^2_{AM}$	$\hat{\lambda}^2_{NW}$
$\hat{d}_{T\text{FEWL}}$	Model 0	0.313	0.452	0.426	0.466
	Model 2	0.324	0.479	0.451	0.493
	Model 3	0.288	0.416	0.392	0.428
$\hat{d}_{T\text{EML}}$	Model 0	0.485	0.460	0.497	0.328
	Model 2	0.514	0.487	0.527	0.341
	Model 3	0.446	0.423	0.457	0.306
$\hat{d}_{T\text{MD}}$	Model 0	0.657	0.660	0.645	0.436
	Model 2	0.699	0.699	0.683	0.454
	Model 3	0.699	0.699	0.683	0.409

Note: *Critical values at the 5% S.L.

Finally, we have applied the techniques introduced in this article to check whether the evidence in favour of FI could be due to the existence of unaccounted breaks in the deterministic components. Since the true d_0 is unknown, we take H'_0 as the null hypothesis. As for the alternative hypothesis, Models 0, 2 and 3 have been considered. The former introduces an intercept in the inflation equation that is allowed to break as the only deterministic component, which is the model advocated by Levin and Piger (2003). Another possibility, as Figure 1 suggests, is that inflation could have an upward trend, until about the middle of the sample, followed by a downward trend. Models 2 and 3 can reproduce for this behaviour. Table 7 presents the values of the statistics for testing non-stationary FI against I(0) +breaks. The estimates of d reported in Table 6. have been used to compute $R^f_{b_i}(\hat{d}_T)$, $i\in\{0,2,3\}$, as defined in equation (12). The first column of figures in Table 7 report critical values corresponding to the particular model and value of d employed to run the test. Columns 2 to 4 display the values of the tests computed with several estimates of the long-run variance $\hat{\lambda}^2_k$ as defined in equation (16) for $k=\{AM, A, NW\}$.

Table 7 shows that the finding of FI in inflation data is very robust. According to this table, there is no evidence to reject the null hypothesis of FI versus the alternative of $I(0)+$ breaks for any of the models considered under H'_1. These conclusions are also robust to the use of different values of q_k employed to perform the short-term correlation correction. The economic implication of this finding is clear: no evidence of structural breaks that could induce the observed persistence in inflation data has been found, implying that the inflation rate is a highly persistent variable.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

VII. Conclusions

This article analyses the long-standing issue of determining the source of the non-stationarity observed in economic variables: whether it is the result of innovations that are highly persistent or whether it appears as a consequence of the existence of rare and unexpected events that change the underlying structure of the series (breaks). We depart from the traditional framework that sets the unit root process as the null hypothesis by considering the more general class of non-stationary FI(d) models. The number of interesting testing frameworks that one could consider is, of course, much larger. Another interesting possibility would be to consider models where breaks are allowed under both the null and the alternative hypotheses and one is interested in testing for the degree of integration (see Perron (1989, 2005), for related references for the I(1) versus $I(0)+$ breaks case). In this framework where breaks can occur under both hypotheses, a variance ratio test in the spirit of Breitung (2002) could be implemented. An advantage of using the latter technique is that the problem of estimating the long-run variance could be avoided. More research should be carried out to study this and other possibilities.

Final Manuscript Received: November 2010

References

Andrews, D. W. K. (1991). 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica*, Vol. 59, pp. 817–854.

Andrews, D. W. K. (1993). 'Test for parameter instability and structural change with unknown break point', *Econometrica*, Vol. 61, pp. 821–856.

Andrews, D. W. K. and Monahan, J. C. (1992). 'An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator', *Econometrica*, Vol. 60, pp. 953–966.

Andrews, D.W. K. and Ploberger, W. (1994). 'Optimal tests when a nuisance parameter is present only under the alternative', *Econometrica*, Vol. 62, pp. 1383–1414.

Bai, J. (1999). 'Likelihood ratio tests for multiple structural changes', *Journal of Econometrics*, Vol. 91, pp. 299–323.

Bai, J. and Perron, P. (1998). 'Testing for and estimation of multiple structural changes', *Econometrica*, Vol. 66, pp. 47–79.

Baillie, R., Chung, C.-F. and Tieslau, M. (1996). 'Analyzing inflation by the fractionally integrated ARFIMA-GARCH model', *Journal of Applied Econometrics*, Vol. 11, pp. 23–40.

Banerjee, A. and Urga, G. (2005). 'Modelling structural breaks, long-memory and stock market volatility: an overview', *Journal of Econometrics*, Vol. 129, pp. 1–34.

Banerjee, A., Lumsdaine, R. L. and Stock, J. H. (1992). 'Recursive and sequential tests of the unit root and trend break hypothesis: theory and international evidence', *Journal of Business and Economic Statistics*, Vol. 10, pp. 271–287.

Berkes, I., Horvath, L., Kokoszka, P. S. and Shao, Q.-M. (2006). 'On discriminating between long-range dependence and changes in mean', *Annals of Statistics*, Vol. 34, pp. 1140–1165.

Bhargava, A. (1986). 'On the theory of testing for unit roots in observed time series', *The Review of Economic Studies*, Vol. 53, pp. 369–384.

Bhattacharya, R. N., Gupta, V. K. and Waymire, E. (1983). 'The Hurst efect under trends', *Journal of Applied Probability*, Vol. 20, pp. 649–662.

Breitung, J. (2002). 'Nonparametric tests for unit roots and cointegration', *Journal of Econometrics*, Vol. 108, pp. 343–363.

Cheung, Y. W. and Lai, K. S. (1997). 'Bandwidth selection, prewhitening, and the power of the Phillips–Perron test', *Econometric Theory*, Vol. 13, pp. 679–691.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

Cogley, T. and Sargent, T. J. (2001). 'Evolving post World War II inflation dynamics', *NBER Macroeconomics Annual*, Vol. 16, pp. 331–373.
Cogley, T. and Sargent, T. J. (2005). 'The conquest of U.S. inflation. Learning and robustness to model uncertainty', *Review of Economic Dynamics*, Vol. 8, pp. 528–563.
Cox, D. (1962). 'Further results on tests of separate families of hypotheses', *Journal of the Royal Statistical Society B*, Vol. 24, pp. 406–424.
Davidson, J. (1994). *Stochastic Limit Theory*, Oxford University Press, New York.
Davidson, J. and Sibbertsen, P. (2005). 'Generating schemes for long-memory processes: regimes, aggregation and linearity', *Journal of econometrics*, Vol. 128, pp. 253–282.
Davydov, Y. (1970). 'The invariance principle for stationary processes', *Theory of Probability and its Applications*, Vol. 15, pp. 487–489.
Diebold, F. X. and Inoue, A. (2001). 'Long-memory and regime switching', *Journal of Econometrics*, Vol. 105, pp. 131–159.
Diebold, F. X. and Rudebusch, G. (1991). 'On the power of the Dickey–Fuller tests against fractional alternatives', *Economics Letters*, Vol. 35, pp. 55–160.
Doornik, J. A and Ooms, M. (2004). 'Inference and forecasting for ARFIMA models with an application to U.S. and U.K. Inflation', *Studies in Non linear Dynamics and Econometrics* 8, article 14.
Elliott, G., Rothenberg, E. and Stock, J. (1996). 'Efficient test for an autoregressive unit root', *Econometrica*, Vol. 64, pp. 813–839.
Gadea, L. and Mayoral, L. (2006). 'The persistence of inflation in OECD countries: a fractionally integrated approach', *International Journal of Central Banking*, Vol. 2, pp. 51–104.
Giraitis, L., Kokoszka, P. and Leipus, R. (2001). 'Testing for long-memory in the presence of a general trend', *Journal of Applied Probability*, Vol. 38, pp. 1033–1054.
Giraitis, L., Leipus, R. and Philippe, A. (2006). 'A test for stationarity versus trends and unit roots for a wide class of dependent errors', *Econometric Theory*, Vol. 22, pp. 989–1029.
Granger, C. W. J. (1980). 'Long-memory relationships and the aggregation of dynamic models', *Journal of Econometrics*, Vol. 14, pp. 227–238.
Granger, C. W. J. and Hyung, N. (2004). 'Occasional structural breaks and long-memory with an application to the S&P 500 absolute stock returns', *Journal of Empirical Finance*, Vol. 11, pp. 399–421.
Haubrich, J. G. (1993). 'Consumption and fractional differencing: old and new anomalies', *Review of Economics and Statistics*, Vol. 75, pp. 767–772.
Henry, M. and Zaffaroni, P. (2002). 'The long range dependence paradigm for macroeconomics and finance', in Doukhan P. Oppenheim G. and Taqqu M. (eds), *Long Range Dependence: Theory and Applications*, Birkhäuser, Boston, pp. 417–438.
Heyde, C. C. and Dai, W. (1996). 'On the robustness to small trends of estimation based on the smoothed periodogram', *Journal of Time Series Analysis*, Vol. 17, pp. 141–150.
Hsu, C. C. (2001). 'Change point estimation in regressions with I(d) variables', *Economics Letters*, Vol. 70, pp. 147–155.
Iacone, F. (2005). 'Local Whittle estimation of the memory parameter in presence of deterministic components', Mimeo.
Krämer, W. and Sibbertsen, P. (2002). 'Testing for structural changes in the presence of long-memory', *International Journal of Business and Economics*, Vol. 1, pp. 235–242.
Künsch, H. (1986). 'Discrimination between deterministic trends and long range dependence', *Journal of Applied Probability*, Vol. 23, pp. 1025–1030.
Lee, D. and Schmidt, P. (1996). 'On the power of the KPSS test of stationarity against fractionally-integrated alternatives', *Journal of Econometrics*, Vol. 73, pp. 285–302.
Lehmann, E. L. (1959). *Testing Statistical Hypothesis*, John Wiley and Sons, New York.
Levin, A. and Piger, J. (2003). *Is Inflation Persistence Intrinsic in Industrial Economics*, Federal Reserve Bank of Saint Louis Working Paper 2002-023.
Lobato, I. and Savin, N. E. (1998). 'Real and spurious long-memory properties of the stock market data', *Journal of Business and Economic Statistics*, Vol. 16, pp. 261–283.
Marinucci, D. and Robinson, P. M. (1999). 'Alternative forms of fractional Brownian motion', *Journal of Statistical Planning and Inference*, Vol. 80, pp. 111–122.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

Marinucci, D. and Robinson, P. M. (2000). 'Weak convergence of multivariate fractional processes', *Stochastic Processes and their Applications*, Vol. 86, pp. 103–120.
Mármol, F. and Velasco, C. (2002). 'Trend stationarity versus long-range dependence in time series analysis', *Journal of Econometrics*, Vol. 108, pp. 25–42.
Mayoral, L. (2006). 'Further evidence on the statistical properties of real GNP', *Oxford Bulletin of Economics and Statistics*, Vol. 68, pp. 901–920.
Mayoral, L. (2007). 'Minimum distance estimation of stationary and non-stationary ARFIMA processes', *Econometrics Journal*, Vol. 10, pp. 124–148.
Michelacci, C. and Zaffaroni, P. (2000). 'Fractional beta convergence', *Journal of Monetary Economics*, Vol. 45, pp. 129–153.
Newey, K. N. and West, K. D. (1994). 'Automatic lag selection in covariance matrix estimation', *Review of Economic Studies*, Vol. 61, pp. 631–653.
Ng, S. and Perron, P. (2001). 'Lag length selection and the construction of unit root tests with good size and power', *Econometrica*, Vol. 69, pp. 1519–1554.
Nunes, L. C., Kuan, C. M. and Newbold, P. (1995). 'Spurious breaks', *Econometric Theory*, Vol. 11, pp. 736–749.
Ohanissian, A., Russell, J. R. and Tsay, R. S. (2008). 'True of spurious long-memory? A new test', *Journal of Business & Economic Statistics*, Vol. 26, pp. 161–175.
O'Reilly, G. and Whelan, K. (2005). 'Has Euro-area inflation persistence changed over time?' *Review of economics and statistics*, Vol. 87, pp. 709–720.
Perron, P. (1989). 'The great crash, the oil price shock and the unit root hypothesis', *Econometrica*, Vol. 58, pp. 1361–1401.
Perron, P. (1997). 'Further evidence from breaking trend functions in macroeconomic variables', *Journal of Econometrics*, Vol. 80, pp. 355–385.
Perron, P. (2006). 'Dealing with structural breaks', in Patterson K., and Mills T. C. (eds), *Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory*, Palgrave Macmillan, Chippenham and Eastbourne, pp. 278–352.
Perron, P. and Qu, Z. (2006). *An Analytical Evaluation of the Log-Periodogram Estimate in the Presence of Level Shifts and its Implications for Stock Returns Volatility*, Boston University Working Papers Series WP2006-016.
Phillips, P. C. B. and Perron, P. (1988). 'Testing for a unit root in time series regression', *Biometrika*, Vol. 75, pp. 335–346.
Pivetta, F. and Reis, R. (2007). 'The persistence of inflation in the United States', *Journal of Economic Dynamics and Control*, Vol. 31, pp. 1326–1358.
Pötscher, B. M. (2002). 'Lower risk bounds and properties of confidence sets for ill-posed estimation problems with applications to spectral density and persistence estimation, unit roots, and estimation of long-memory parameters', *Econometrica*, Vol. 70, pp. 1035–1065.
Robinson, P. M. (1978). 'Statistical inference for a random coefficient autoregressive model', *Scandinavian Journal of Statistics*, Vol. 5, pp. 163–168.
Rotemberg, J. (1987). 'The new Keynesian microfoundations', *Macroeconomics Annual*, Vol. 2, pp. 69–104.
Sargan, J. D. and Bhargava, A. (1983). 'Testing residuals from least squares regression for being generated by the Gaussian random walk', *Econometrica*, Vol. 51, pp. 153–174.
Shimotsu, K. (2006a). *Simple (but effective) Tests of Long-Memory Versus Structural Breaks*, Queen's Economics Department Working Paper No. 1101.
Shimotsu, K. (2006b). *Exact Local Whittle Estimation of Fractional Integration with Unknown Mean and Time Trend*, Queens Economics Department Working Paper No. 1061. *Econometric Theory*, Vol. 26, pp. 501–540.
Shimotsu, K. and Phillips, P. C. B. (2005). 'Exact local whittle estimation of fractional integration', *Annals of Statistics*, Vol. 33, pp. 1890–1933.
Sowell, F. B. (1992). 'Maximum likelihood estimation of stationary univariate fractionally-integrated time-series models', *Journal of Econometrics*, Vol. 53, pp. 165–188.
Teverovsky, V. and Taqqu, M. (1997). 'Testing for long-range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator', *Journal of Time Series Analysis*, Vol. 18, pp. 279–304.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

Velasco, C. and Robinson, P. M. (2000). 'Whittle pseudo-maximum likelihood estimates of non-stationary time series', *Journal of the American Statistical Association*, Vol. 95, pp. 1229–1243.
Zivot, E. and Andrews, D. W. K. (1992). 'Further evidence on the Great Crash, the oil-price shock and the unit root hypothesis', *Journal of Business and Economic Statistics*, Vol. 10, pp. 251–270.

Appendix A
Proof of theorem 1

(a) Consider first the case where condition C holds.
If $Z_t=(1)$, then

$$(y-Z\tilde{\beta})'(y-Z\tilde{\beta})=\sum_{t=1}^{T}x_t^2-T^{-1}\left(\sum_{t=1}^{T}x_t\right)^2,$$

where x_t is defined in equation (2). From the functional central limit theorem (see Davydov, 1970) and the continuous mapping theorem, it follows that

$$T^{-2d_0}\sum_{t=1}^{T}x_t^2-\left(T^{-(1/2+d_0)}\sum_{t=1}^{T}x_t\right)^2\xrightarrow{w}\lambda^2\int_0^1(B_{d_0}^{\mu}(r))^2\,\mathrm{d}r$$
$$=\lambda^2\int_0^1\left(B_{d_0}-\left(\int_0^1 B_{d_0}(r)\,\mathrm{d}r\right)\right)^2,$$

where $\lambda^2=\sigma^2\Psi(1)^2$ is the long-run variance.

Convergence of the numerator of equation (8) for the case where $Z_t=(1\ t)$ can be obtained using a similar strategy as Marmol and Velasco (2002, p. 38). In this case

$$T^{-2d_0}(y-Z\tilde{\beta})'(y-Z\tilde{\beta})\xrightarrow{w}\lambda^2\int_0^1(B_{d_0}^{\tau}(r))^2\,\mathrm{d}r$$
$$=\lambda^2\left(\int_0^1 B_{d_0}^2(r)\,\mathrm{d}r-\left(\int_0^1 B_{d_0}(r)\,\mathrm{d}r\right)^2-12\left(\int_0^1(r-1/2)B_{d_0}(r)\,\mathrm{d}r\right)^2\right).$$

With respect to the denominator of equation (8), notice that $\Delta^{d_0}y-\Delta^{d_0}Z\hat{\beta}=\varepsilon^*-\Delta^{d_0}Z(\beta-\hat{\beta})$, where $\varepsilon_t^*=\varepsilon_t-\sum_{i=t-1}^{\infty}\pi_i(d_0-1)\Delta x_t=\varepsilon_t+O_p(t^{d_0/2-3/4})$ (see Marinucci and Robinson, 2000). Also notice that $\Delta^{d_0}t=\tau_t(d_0-1)\approx kt^{1-d_0}$. Consider first the case where $Z_t=t$. Then,

$$T^{-1}(\Delta^{d_0}y-\Delta^{d_0}Z\hat{\beta})'(\Delta^{d_0}y-\Delta^{d_0}Z\hat{\beta})$$
$$=T^{-1}\left(\sum_{t=2}^{T}\varepsilon_t^{*2}+(\beta-\hat{\beta})^2\sum_{t=2}^{T}t^{2(1-d_0)}-2(\beta-\hat{\beta})\sum_{t=2}^{T}\varepsilon_t^{*}t^{1-d_0}\right)+o_p(1). \quad (18)$$

The first term in equation (18) converges in probability to σ^2 given that

$$T^{-1}\sum_{t=2}^{T}\varepsilon_t^{*2}=T^{-1}\sum_{t=2}^{T}\varepsilon_t^2+T^{-1}\sum_{t=2}^{T}O_p(t^{d_0-3/2})+o_p(1)$$
$$=\sigma^2+O_p(T^{d_0-3/2})+o_p(1)=\sigma^2+o_p(1), \quad (19)$$

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

since $d_0 < 3/2$. We now show that the second and third terms in equation (18) converge to zero. Using standard results, it is easy to show that $\sum_{t=2}^{T} \varepsilon_t t^{1-d_0} = O_p(T^{3/2-d_0})$ and that $\hat{\beta}$ is a $T^{3/2-d_0}$-consistent estimator of β (see Hamilton, 1994, pp. 459). Thus, the second term in equation (18) is

$$T^{-1}(\beta-\hat{\beta})^2 \sum_{t=2}^{T} t^{2(1-d_0)} = T^{-1} O_p(T^{-(3-2d_0)}) O(T^{3-2d_0}) + o_p(1) \xrightarrow{p} 0. \tag{20}$$

Finally,

$$\begin{aligned} T^{-1}(\beta-\hat{\beta}) \sum_{t=2}^{T} \varepsilon_t^* t^{1-d_0} &= T^{-1} O_p(T^{-(3-2d_0)}) \left(\sum_{t=2}^{T} \varepsilon_t t^{1-d_0} - \sum_{t=2}^{T} O_p(t^{d_0/2-3/4}) t^{1-d_0} \right) \\ &= T^{-1} O_p(T^{-3/2+d_0}) + T^{-1} O_p(T^{3/4-d_0/2}) \xrightarrow{p} 0. \end{aligned} \tag{21}$$

Expressions (19), (20) and (21) imply that equation (18) converges to σ^2 in probability.

The case where $Z_t = (1\ t)$ can be analyzed in a similar manner. In this case, $(\hat{\beta} - \beta) = O_p(T^{-(3/2-d_0)})$ and $(\hat{\alpha} - \alpha) = O_p(1)$. Thus

$$T^{-1}(\Delta^{d_0} y - \Delta^{d_0} Z\hat{\beta})'(\Delta^{d_0} y - \Delta^{d_0} Z\hat{\beta}) \tag{22}$$

$$= \sigma^2 + T^{-1} \left((\alpha-\hat{\alpha})^2 \sum_{t=2}^{T} \tau_t^2(d_0) + 2(\alpha-\hat{\alpha})(\hat{\beta}-\beta) \sum_{t=2}^{T} \tau_t(d_0)\tau_t(d_0-1) \right) \tag{23}$$

$$+ T^{-1} \left(2(\alpha-\hat{\alpha}) \sum_{t=2}^{T} \varepsilon_t \tau_t(d_0) \right) + o_p(1), \tag{24}$$

where the $o_p(1)$ component above gathers the terms that have been shown above to converge to zero. Using the fact that $\sum_{t=2}^{T} \tau_t(d_0)\tau_t(d_0-1) = O(T^{2-2d_0})$ and that $\sum_{t=2}^{T} \varepsilon_t \tau_t(d_0) = O_p(T^{1/2})$, it is straightforward to show that the second and third terms in the RHS of equation (22) are also $o_p(1)$. If condition 3 is relaxed, standard use of the continuous mapping theorem (CMT) yields

$$T^{-2d_0}(y - Z\tilde{\beta})'(y - Z\tilde{\beta}) \xrightarrow{w} \lambda^2 \left(\int_0^1 (B_{d_0}^c(r)) \right)^2 \mathrm{d}r, c \in \{\mu, \tau\},$$

where $B_{d_0}^{\mu}(r)$ and $B_{d_0}^{\tau}(r)$ correspond to the cases where $Z_t = (1)$ or $Z_t = (1, t)$, respectively, and $\lambda^2 = \sigma^2 \Psi(1)^2$ is the long-run variance. Finally, using similar arguments as above, it is easy to show that

$$T^{-1}(\Delta^{d_0} y - \Delta^{d_0} Z\hat{\beta})'(\Delta^{d_0} y - \Delta^{d_0} Z\hat{\beta}) \xrightarrow{p} \gamma_0,$$

where γ_0 is the variance of u_t.

(b) Consider now the case where condition 2 is dropped. The limit of the numerator of $R^f(\hat{d}_T)$ can be easily obtained since

$$\frac{T^{-2\hat{d}_T}}{T^{-2d_0}} T^{-2d_0}(y - Z\tilde{\beta})'(y - Z\tilde{\beta}) = T^{-2d_0}(y - Z\tilde{\beta})'(y - Z\tilde{\beta}) + o_p(1), \tag{25}$$

as $T^{-2\hat{d}_T}/T^{-2d_0} \to 1$ if $\hat{d}_T$ is a T^{κ}-consistent estimator of d_0.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

Next, denote $G_T(d)=T^{-1}(\Delta^d y-\Delta^d Z\beta)'(\Delta^d y-\Delta^d Z\beta)$. Sufficient conditions that guarantee that $G_T(\hat{d}_T)-G_T(d_0)=o_p(1)$ are the consistency of $\hat{d}_T$ and the fact that $G_T(d)\xrightarrow{p}G(d)$ uniformly on an open set B_0 containing d_0. If $G_T(d)\xrightarrow{p}G(d)$ pointwise and $\{G_T\}$ is stochastically equicontinuous in B_0, then $G_T(d)\xrightarrow{p}G(d)$ uniformly on B_0. From theorem 21.10 in Davidson (1994), it follows that $\{G_T\}$ is stochastically equicontinuous provided

$$G'(d)=\sup_{d^*\in B_0^*}|G'(d)|_{d=d^*}|=O_p(1),$$

where $G'(d)|_{d=d^*}$ is the derivative of $G(d)$ evaluated at $d^*\in B_0^*$, where B_0^* is an open convex set containing B_0. The derivative of $G(d)$ is given by

$$G'(d)|_{d=d^*}=T^{-1}\sum_{t=2}^{T}(\Delta^{d^*}y_t-\beta'\Delta^{d^*}Z_t)(\log(1-L)(\Delta^{d^*}y-\Delta^{d^*}Z\beta')). \tag{26}$$

Noticing that $(\Delta^{d^*}y_t-\beta'\Delta^{d^*}Z)=\Delta^{d^*-d_0}u_t^*$, where $u_t^*=u_t-\sum_{i=t-1}^{\infty}\pi_i(d_0-1)\Delta x_t$, and that $\log(1-L)=-(L+L^2/2+L^3/3+\cdots)$, equation (26) can be written as

$$\begin{aligned}G_T'(d)|_{d=d^*}&=-T^{-1}\sum_{t=2}^{T}(\Delta^{d^*-d_0}u_t^*)\left(\Delta^{d^*-d_0}u_{t-1}^*+\frac{\Delta^{d^*-d_0}u_{t-2}^*}{2}+\frac{\Delta^{d^*-d_0}u_{t-3}^*}{3}+\cdots\right)\\&=-\sum_{j=1}^{T-2}\frac{\hat{\gamma}_{\Delta^{d^*-d_0}u_t^*}(j)}{j},\end{aligned} \tag{27}$$

where $\hat{\gamma}_{\Delta^{d^*-d_0}u_t^*}(j)=T^{-1}\sum_{t=j+1}^{T}(\Delta^{d^*-d_0}u_t^*)(\Delta^{d^*-d_0}u_{t-j}^*)$ is the jth sample autocorrelation of $\Delta^{d^*-d_0}u_t^*$. Let B_0^* be an open ball of radius $\epsilon<1/4$ centred at d_0 . Using a similar strategy as before, it can be shown that $\Delta^{d^*-d_0}u_t^*$ is the sum of an FI(d_0-d^*) process, $\Delta^{d^*-d_0}u_t$, and a process whose sample autocovariance function tends to zero in probability. Also notice that $\Delta^{d^*-d_0}u_t$ is stationary for any $d^*\in B_0^*$, since $|d^*-d_0|<1/4$. Then, the smaller d^*, the higher the order of integration of $\Delta^{d^*-d_0}u_t$ and, for a given u_t, the higher the value of the sum of the sample covariances, $\sum_{j=1}^{T-2}|\hat{\gamma}_{\Delta^{d^*-d_0}u_t}(j)/j|$. Notice that

$$\sup_{d^*\in B_0^*}|G'(d)|_{d=d^*}|\leq\sup_{d^*\in B_0^*}\sum_{j=1}^{T-2}\left|\frac{\hat{\gamma}_{\Delta^{d^*-d_0}u_t}(j)}{j}\right| \tag{28}$$

and that the arg sup of the right-hand side of equation (28) is $d^*=d_0-1/4$ (in which case, $\Delta^{d^*-d_0}u_t$ is an FI(1/4) process). Finally, notice that equation (28) is bounded in probability since $\hat{\gamma}_{\Delta^{-1/4}u_t}(j)/j\xrightarrow{p}\gamma_{\Delta^{-1/4}u_t}(j)/j\approx kj^{-3/2}$, for some constant k and, therefore, the sum of these terms is well-defined. It follows that $T^{-1}(\Delta^{\hat{d}_T}y-\Delta^{\hat{d}_T}Z\beta)'(\Delta^{\hat{d}_T}y-\Delta^{\hat{d}_T}Z\beta)\xrightarrow{p}\gamma_0$. When β is replaced by $\hat{\beta}$, similar arguments as those employed in (a) can be used to show that $T^{-1}(\Delta^{\hat{d}_T}y-\Delta^{\hat{d}_T}Z\hat{\beta})'(\Delta^{\hat{d}_T}y-\Delta^{\hat{d}_T}Z\hat{\beta})\xrightarrow{p}\gamma_0$.

Proof of theorem 2

Under H_1, the terms $T^{-1}(\Delta^{d_0}y-\Delta^{d_0}Z\hat{\beta})'(\Delta^{d_0}y-\Delta^{d_0}Z\hat{\beta})$ and $T^{-1}(y-Z\tilde{\beta})'(y-Z\tilde{\beta})$ are the sample variances of stationary and ergodic processes and they tend in probability to

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

the corresponding population variances (which are strictly greater than zero). Therefore, $(y - Z\tilde{\beta})'(y - Z\tilde{\beta})/(\Delta^{d_0}y - \Delta^{d_0}Z\hat{\beta})'(\Delta^{d_0}y - \Delta^{d_0}Z\hat{\beta})$ is $O_p(1)$. Similarly, $\hat{\lambda}^2$ and $\hat{\gamma}_0$ are also sample moments of a stationary and ergodic process; therefore, that they are both $O_p(1)$.[13] This implies that equation (8) tends to zero at a rate T^{1-2d_0} and, therefore, the probability of rejecting H_1 tends to 1 if $T \to \infty$. The same arguments are valid if $d_0$is replaced by $\hat{d}_T$, since $\hat{d}_T$ satisfies that $\hat{d}_T > 0.5$.

Proof of theorem 3

The proof of this theorem closely follows the proof of theorem 1 in Perron (1997), (PE hereafter) which, in turn, borrows from the proof of theorem 1 in Zivot and Andrews (1992). (Z&A, henceforth). To simplify cross-references, the same notation as in those papers has been adopted. For the sake of brevity, proofs of some intermediate results whose proofs can be found in PE or in Z&A are not reproduced below but precise references are provided.

We consider, first, the case where $u_t = \varepsilon_t$ is an i.i.d. sequence. Let $S_t = \sum_{j=0}^{t-1} \pi_j(-d)\varepsilon_{t-j}$ ($S_0 = 0$), and $X_T(r)$ be the partial sum process defined as,

$$X_T(r) = T^{1/2-d_0}\sigma^{-1}S_{\lfloor Tr \rfloor},\ (j-1)/T < r < (j+1)/T \qquad \text{for} \qquad j = 1, \ldots, T,$$

where $\lfloor . \rfloor$ denotes integer part. Let $z^i_{tT}(\omega)$ for $i = \{0, 1, 2, 3\}$ be a vector containing the deterministic components for model i of the numerator of equation (12). $z^i_{tT}(\omega)$ explicitly depends on the break fraction (ω) and T, the sample size. For instance, if $i = 1$, then $z^1_{tT}(\omega)' = (1\ t\ BC_t(\omega))$, where $BC_t = 1$, if $t > T_B$ and 0 otherwise. The vector $Z^i_T(\omega, r)$ represents a rescaled version of the deterministic regressors, i.e. $Z^i_T(\omega, r) = \theta^i_T z^i_{[Tr]T}(\omega)$, where θ^i_T is a diagonal matrix of weights.[14] We also define the limiting functions $Z^0(\omega, r) = (1\ bc(\omega, r))$ where $bc(\omega, r) = 1_{(r>\omega)}, Z^1(\omega, r) = (1\ r\ bc(\omega, r))$, $Z^2(\omega, r) = (1, r, bt(\omega, r))$, where $bt(\omega, r) = (r - \omega)1_{(r>\omega)}$, and $Z^3(\omega, r) = (1\ r\ bc(\omega, r)\ bt(\omega, r))$.

Let $P_{z_T(\omega)} = [P_{z_{1,T}}(\omega), \ldots, P_{z_{T,T}}(\omega)]$ be the linear map projecting onto the space spanned by the columns of $z_T(\omega)' = (z_{1,T}, \ldots, z_{T,T}(\omega))$, that is,

$$P_{z_T}(\omega) = z_T(\omega)(z_T(\omega)'z_T(\omega)')^{-}z_T(\omega)'$$

where $(.)^{-}$ denotes generalized inverse. Premultiplying by $M_{z_T}(\omega) = (I - P_{z_T}(\omega))$. Equation (1) can be rewritten, in matrix notation, as $M_{z_T}(\omega)y = M_{z_T}x_t$. It follows that $R^f_{b_i}(d_0)$ is given by

$$R^f_{b_i}(d_0) = \inf_{\omega\in\Omega} \frac{T^{-2d_0}y'M_{z_T}(\omega)y}{\hat{\sigma}^2_T}, \qquad \text{for} \qquad i = \{0, 1, 2, 3\}, \tag{29}$$

where $\hat{\sigma}^2_T = T^{-1}(\Delta^{d_0}y - \Delta^{d_0}Z\beta)'(\Delta^{d_0}y - \Delta^{d_0}Z\beta)$.

[13] Notice that, for consistency of the test, it is not necessary that $\hat{\lambda}$ or $\hat{\gamma}_0$ are consistent under H_1. They only need to be bounded.

[14] For instance, in Model 1,

$$\theta^1_T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & T & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

Henceforth, only Model 1 will be considered. Proofs for models $\{0,2,3\}$ are analogous and, therefore, are omitted. For simplicity, the superscript denoting the model is dropped hereafter.

The proof will be completed in three steps that closely follow Perron's approach. The first one shows that the numerator of equation (29) can be written as a functional g that is a composition of functionals depending on $X_T(.)$ and $Z_T(.,.)$. Next, some joint convergence results are presented. Finally, it is shown that g is a composition of continuous functionals and, therefore, it is also continuous. The proof of the theorem is completed by applying the continuous mapping theorem (CMT).

First step. Simple algebra shows (see expression (A.3) in PE) that

$$T^{-2d_0}y'M_{z_T}(\omega)y=\sigma^2\int_0^1\{X_T(r)-P_{z_T}(\omega)X_T(r)\}^2\,\mathrm{d}r+o_{p\omega}(1) \tag{30}$$

where $o_{p\omega}(1)$ denotes a random variable that converges in probability to zero uniformly in ω, and

$$P_{z_T}(\omega)X_T(r)=Z_T(\omega,r)'[Z_T(\omega,r)Z_T(\omega,r)']^{-1}Z_T(\omega,r)\int_0^1 Z_T(\omega,s)X_T(s)\,\mathrm{d}s.$$

Thus, the statistic $R^f_{b_i}(d_0)$ can be expressed as a composite functional

$$\inf_{\omega\in\Omega}\frac{T^{-2d_0}y'M_{Z_T}(\omega)y}{s_T^2}=g(X_T,P_{z_t}(\omega)X_T(r),S_T^2),$$

where

$$g=h^*[h(F_1[X_T,P_{z_t}(\omega)X_T(r)],\hat{\sigma}_T^2)],$$

with $h^*(m)=\inf_{\omega\in\Omega}m(\omega)$ for any real function $m=m(.)$ on Ω and, for any real functions $m_1(.)$ and $m_2(.)$, $h(m_1(\omega),m_2)=m_1(\omega)/m_2$. The functional F_1 is defined by equation (30).

Second step. This step establishes some joint convergence results. The following Lemma is part of Lemma A.1 in PE.

Lemma 1. (Lemma A.1., PE) The following convergence results hold jointly

$$X_T(.)\xrightarrow{w}B_{d_0}(.), \tag{31}$$

$$\begin{aligned}P_{z_T}(\omega)X_T(r)&\xrightarrow{w}P_Z(\omega)B_{d_0}(r)\\&\equiv Z(\omega,r)'\left[\int_0^1 Z(\omega,s)Z(\omega,s)'\,\mathrm{d}s\right]^{-1}\int_0^1 Z(\omega,s)B_{d_0}(s)\,\mathrm{d}s,\end{aligned} \tag{32}$$

and

$$\hat{\sigma}_T^2\xrightarrow{p}\sigma^2.$$

Proof. See the proof of Lemma A.1 in PE, p. 380.

Third step. The final step is to show the continuity of various functionals.

Lemma 2. The functionals h, h^* and F_1 are continuous.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

Proof. Continuity of h and h^* has been shown by Z&A (Lemmas A.3 and A.4) while continuity of F_1 is shown in Perron (Lemma A.2, p. 383).

The continuity of these three functionals implies that g is also continuous. This result, combined with equations (31), (32) and the CMT imply that

$$\frac{T^{-2d_0}y'M_{z_T}(\omega)y}{\hat{\sigma}_T^2} \xrightarrow{w} \frac{\inf_{\omega\in\Omega}\left(\sigma^2\int_0^1\{B_{d_0}(r)-P_Z(\omega)B_{d_0}(r)\}^2\,\mathrm{d}r\right)}{\sigma^2}. \tag{33}$$

For the general case where u_t is allowed to present autocorrelation, using standard results it is possible to obtain that the numerator of equation (29) tends to $\inf_{\omega\in\Omega}(\lambda^2\int_0^1\{B_{d_0}(r)-P_Z(\omega)B_{d_0}(r)\}^2\,\mathrm{d}r)$ whereas the denominator converges to γ_0.

Proof of theorem 4

The numerator of $R^f_{b_i}(\hat{d}_T)$ can be written as

$$\left(\frac{T^{-2\hat{d}_T}}{T^{-2d_0}}\right)T^{-2d_0}\inf_{\omega\in\Omega}y'M_{z_T}(\omega)y. \tag{34}$$

If $\hat{d}_T$ is a T^κ-consistent estimator of d_0, with $\kappa>0$, then $(T^{-2\hat{d}_T}/T^{-2d_0})\xrightarrow{p}1$, implying that the distribution of equation (34) is given in equation (33). The denominator of $R^f_{b_i}(\hat{d}_T)$ is the same as that of $R^f(\hat{d}_T)$ and, therefore, theorem 1 applies. □

Proof of theorem 5

The proof of this theorem can be constructed along similar lines to that of theorem 2 and, therefore, is omitted.

Appendix B

The asymptotic distributions defined in the main text have been simulated for 50 values of d, from $d=0.51$ up to $d=1.49$ with an increment of 0.02 between each consecutive value. The number of replications is 20,000 and the innovations are drawn from independent Gaussian series.

The results are summarized in Tables B1 to B6 by means of the coefficients of polynomial OLS regressions of the 1%, 5%, and 10% sample quantiles of the corresponding statistic on a polynominal of degree 7, namely, $(1\ d\ d^2\ d^3\ d^4\ d^5\ d^6\ d^7)$. This particular transformation was used to obtain approximately homoskedastic errors in the polynomial regression. This table can be used with great precision to obtain critical values for any value of d (or $\hat{d}_T)\in(0.5,1.5)$. Tables B1 and B2 present the critical values of the test presented in section III, where R^μ and R^τ refer to the test statistics corresponding to a model where a constant and a constant and a linear trend, respectively, were included. Tables B3–B6 report the critical values of the tests of structural breaks corresponding to models 0–3.

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

TABLE B1

Critical values $R^{\mu}(d)$ test

	$T=100$			$T=400$		
	1% *S.L.*	5% *S.L.*	10% *S.L.*	1% *S.L.*	5% *S.L.*	10% *S.L.*
c	35.07	36.86	35.56	79.36	8.1555	8.222
d	−207.2	−217.7	−204.8	−505.6	−518.3	−519.2
d^2	534.3	563.5	516.8	1,391.2	1,427.8	1,424.2
d^3	−774.2	−821.1	−733.2	−2,130.4	−2,193.3	−2,181.2
d^4	677.2	723.1	628.2	1,953.9	2,020.2	2,004.5
d^5	−356.3	−383.2	−323.7	−1,070.7	−1,112.3	−1,101.7
d^6	104.1	112.8	92.69	324.0	338.3	334.6
d^7	−13.01	−14.20	−11.32	−41.75	−43.81	−43.28

TABLE B2

Critical values $R^{\tau}(d)$ test

	$T=100$			$T=400$		
	1% *S.L.*	5% *S.L.*	10% *S.L.*	1% *S.L.*	5% *S.L.*	10% *S.L.*
c	32.98	35.87	38.96	74.40	88.90	88.94
d	−192.8	−211.6	−233.6	−471.6	−577.18	−573.5
d^2	490.4	545.0	613.4	1,289.7	1,617.9	1,600.4
d^3	−699.4	−788.8	−906.1	−1,962.1	−2,524.4	−2,489.2
d^4	601.2	689.32	808.7	1,787.4	2,357.7	2,318.8
d^5	−310.6	−362.3	−434.1	−972.8	−1,314.5	−1,290.5
d^6	89.10	105.8	129.4	292.4	404.4	396.3
d^7	−10.92	−13.21	16.48	−37.43	−52.34	51.72

TABLE B3

Critical values $R_{b_0}(d)$ test; (Model 0)

	$T=100$			$T=400$		
	1% *S.L.*	5% *S.L.*	10% *S.L.*	1% *S.L.*	5% *S.L.*	10% *S.L.*
c	34.14	33.15	33.68	76.91	78.71	80.84
d	−206.6	−195.8	−198.2	−499.2	−508.9	−522.3
d^2	544.2	503.4	508.2	1,396.6	1,420.1	1,457.0
d^3	−803.4	−725.1	−730.6	−2,172.8	−2,204.9	−2,262.0
d^4	714.4	629.0	633.2	2,022.9	2,049.5	2,102.7
d^5	−381.4	−327.5	−329.6	−1,124.4	−1,137.7	−1,167.3
d^6	112.9	94.59	95.20	345.0	348.65	357.7
d^7	−14.27	−11.66	−11.75	−45.02	−45.45	−46.64

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011

TABLE B4

Critical values $R_{b_1}(d)$ test; (Model 1)

	$T=100$			$T=400$		
	1% *S.L.*	5% *S.L.*	10% *S.L.*	1% *S.L.*	5% *S.L.*	10% *S.L.*
c	34.25	36.14	36.29	71.67	76.37	78.31
d	−208.7	−220.8	−220.3	−461.3	−493.0	−505.5
d^2	552.9	587.4	582.9	1,279.8	1,372.5	1,408.0
d^3	−819.9	−876.0	−864.6	−1,974.5	2,125.5	−2,182.0
d^4	732.0	786.9	772.7	1,823.6	1,970.5	2,024.6
d^5	−392.0	−424.3	−414.6	−1,006.0	−1,091.1	−1,122.0
d^6	116.3	126.8	123.3	306.5	333.4	592.3
d^7	−14.74	−16.18	−15.66	−39.74	−43.39	−44.69

TABLE B5

Critical values $R_{b_2}(d)$ test; (Model 2)

	$T=100$			$T=400$		
	1% *S.L.*	5% *S.L.*	10% *S.L.*	1% *S.L.*	5% *S.L.*	10% *S.L.*
c	36.37	37.41	37.96	73.19	78.27	78.28
d	−223.2	−228.3	−230.6	−470.4	−505.0	−503.0
d^2	595.8	607.1	610.7	1,303.1	1,405.2	1,395.1
d^3	−890.8	−905.3	−907.4	−2,007.2	−2,175.0	−2,153.1
d^4	801.9	813.5	812.7	1,850.7	2,015.9	1,991.7
d^5	−433.1	−438.9	−437.1	−1,019.3	−1,115.7	−1,100.3
d^6	129.6	131.2	130.4	310.0	340.9	335.7
d^7	−16.56	−16.76	−16.61	−40.13	−44.32	−43.60

TABLE B6

Critical values $R_{b_3}(d)$ test; (Model 3)

	$T=100$			$T=400$		
	1% *S.L.*	5% *S.L.*	10% *S.L.*	1% *S.L.*	5% *S.L.*	10% *S.L.*
c	32.30	35.37	36.17	69.97	73.01	76.40
d	−195.30	−216.9	−221.4	−449.8	−469.4	−492.9
d^2	512.4	578.9	590.2	1,245.8	1,301.4	1,371.5
d^3	−752.2	−865.8	−881.7	−1,918.2	−2,006.5	−2,122.5
d^4	664.3	779.7	793.3	1,767.9	1,852.2	1,966.3
d^5	−352.0	421.4	−428.4	−973.2	−1,021.3	−1,088.0
d^6	103.3	126.2	128.2	295.8	311.0	332.4
d^7	−12.96	−16.14	−16.38	−38.28	−40.31	−43.21

© Blackwell Publishing Ltd and the Department of Economics, University of Oxford 2011