GLS-based unit root tests with multiple structural breaks both under the null and the alternative hypotheses*

Josep Lluís Carrion-i-Silvestre[†] Dukpa Kim[‡] Pierre Perron[§]

September 16, 2007

Abstract

Perron (1989) introduced unit root tests that are valid when a break at a known date in the trend function of a time series is present. The motivation was to devise tests invariant to the magnitude of the shift in level and/or slope and, in particular, to allow them to occur under both the null and alternative hypotheses. The subsequent literature aimed to devise procedures valid in the case of an unknown break date. However, in doing so most, in particular the commonly used test of Zivot and Andrews (1992), assumed that if a break occurs, it does so only under the alternative hypothesis of stationarity. This is undesirable for several reasons. Kim and Perron (2006) developed a methodology that allows a break at an unknown time under both the null and alternative hypotheses. Also, when a break is present, the limit distribution of the test is the same as in the case of a known break date, thereby allowing increased power while maintaining the correct size. We extend their work in several directions: 1) we allow for an arbitrary number of changes in both the level and slope of the trend function; 2) we adopt the quasi-GLS detrending method advocated by Elliott et al. (1996) which permits tests that have local asymptotic power functions close to the local asymptotic Gaussian power envelope; 3) we consider a variety of tests, in particular the class of M-tests introduced in Stock (1999) and analyzed in Ng and Perron (2001).

Keywords: multiple structural breaks, unit root, GLS detrending
JEL Codes: C12, C22

---

*This paper was prepared for the conference in honour of Professor Paul Newbold at the University of Nottingham, September 21-22. Perron acknowledges financial support for this work from the National Science Foundation under Grant SES-0649350. Carrion-i-Silvestre acknowledges financial support from the Ministerio de Ciencia y Tecnología under grant SEJ2005-08646/ECON.

[†]AQR Research Group, Departament of Econometrics, Statistics and Spanish Economy, University of Barcelona, Av. Diagonal, 690, Barcelona, 08034 (carrion@ub.edu)

[‡]Department of Economics, University of Virginia, 2015 Ivy Rd, Charlottesville, VA 22903 (dk4p@virginia.edu)

[§]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (perron@bu.edu)

1 Introduction

Professor Paul Newbold has made a number of important contributions in econometrics, in particular in the area of the analysis of non-stationary time series. I (Perron) have been honoured and privileged to have him participate in a research agenda that I put forward in Perron (1989). This paper is, in part, a product motivated and made possible by some of his work.

It is well known that a break in the deterministic trend affects the outcome of unit root tests. Perron (1989) showed that a standard Dickey-Fuller (1979) (DF) type unit root test is not consistent if the alternative is that of a stationary noise component with a break in the slope of the deterministic trend. His main point is that the existence of an exogenous shock which has a permanent effect will lead to a non rejection of the unit root hypothesis even though it is not true. Of interest also is the fact that Leybourne, Mills and Newbold (1998) and Leybourne and Newbold (2000) analyzed the effect of a break on a standard DF test under the unit root null hypothesis and showed that size distortions can occur, especially when the break is early in the sample [1]. Perron (1989, 1990) proposed alternative unit root tests which allow the possibility of a break under both the null and alternative hypotheses. These tests have less power than a standard DF type test when there is no break. Nonetheless, they have a correct size asymptotically and are consistent whether there is a break or not. Moreover, they are invariant to the break parameters and thus their performance does not depend on the magnitude of the break. The most controversial assumption, however, is that its timing is known a priori (see Christiano, 1992).

In order for Perron's (1989, 1990) test procedures to be valid, the break date should be chosen independently of the given data. Whenever a systematic search for a break is done, the limiting distributions in Perron (1989, 1990) are no longer appropriate. Historical facts can often be a good guidance in choosing a break date independently of the given data. Even in that case, it is very likely that an imprecise break date is used. Hecq and Urbain (1993) showed, by simulations, that the use of an incorrect break date in Perron's (1990) test, applicable with non-trending data, causes size distortions and power loss, though as shown by Montañés (1997), this effect disappears asymptotically. Montañés and Olloqui (1999) extended the analysis to Perron's (1989) tests, applicable with trending data, and showed that a loss of power occurs even in large samples. Kim, Leybourne and Newbold

[1] See also Montañés and Reyes (1998, 1999, 2000) who examined the asymptotic behavior of the Augmented Dickey-Fuller test (Dickey and Fuller, 1979, Said and Dickey, 1984) and the Phillips-Perron (1988) test under the crash alternative hypothesis.

(2000) examined the effect of using a wrong break date under the null hypothesis.

The work by Zivot and Andrews (1992) provide methods that treat the occurrence of the break date as unknown, and has become quite popular. However, Professor Newbold and his co-authors have clearly recognized the potential pitfalls of their approach, and in all fairness to that of Perron (1997) which follows a similar path. In this line of work, a break is not allowed under the null hypothesis, only under the alternative, mostly because of the theoretical apparatus adopted. This means, for example, that under the null hypothesis a level shift must be viewed as coming from the tail of the distribution of the data generating process, and a slope change involves errors with a different mean in some sub-samples. This framework is convenient since it allows them to establish various unit root testing procedures; for example by minimizing the t-statistic related to the sum of the autoregressive coefficients over each possible break date. This approach is, however, contrary to Perron (1989)'s original motivation, which was to devise testing procedures that where invariant to the magnitude of the shift in level and/or slope. In particular, if a change is present it is allowed under both the null and alternative hypotheses.

The existence of a structural break in the trend function is a problem of long horizon data; it can happen whether the noise component is stationary or has a unit root. As argued in Nunes, Newbold and Kuan (1997) and Harvey, Leybourne and Newbold (2001) (see also Vogelsang and Perron, 1998, and Lee and Strazicich, 2001), if the noise component has a unit root and a break occurs in the trend function, the Zivot and Andrews' (1992) type test statistics often diverge or are not invariant to break parameters. This is a natural consequence of not permitting a break under the null hypothesis. An added consequence is that this type of tests have substantially less power than Perron's (1989) tests, because they do not fully utilize the information about the break, when one is present.

Despite these shorthcomings, the method of Zivot and Andrews (1992) has remained popular in empirical work, probably in part because of the lack of sound statistical methods that could tackle the problem of allowing for changes in the trend function at unknown times under both the null and alternative hypotheses. Using recent developments on structural change problems related to non-stationary data by Perron and Zhu (2005) and Perron and Yabu (2006), Kim and Perron (2006) developed a methodology to devise new test procedures which allow a break in the trend function at an unknown time under both the null and alternative hypotheses. Also, when a break is present, the limit distribution of the test is the same as in the case of a known break date, thereby allowing increased power while maintaining the correct size. Simulation experiments confirm that it offers an improvement over commonly used methods in small samples.

In this paper, we extend the work of Kim and Perron (2006) in several directions: 1) we allow for an arbitrary number of changes in both the level and slope of the trend function [2]; 2) we adopt the so-called quasi-GLS detrending method advocated by Elliott et al. (1996) which permits tests that have local asymptotic power functions close to the local asymptotic Gaussian power envelope [3]; 3) we consider a variety of tests, in particular the class of M-tests introduced in Stock (1999) and analyzed in Ng and Perron (2001). On the other hand, we restrict our analysis to the case of the so-called AO (additive outlier) models as defined in Perron (1989).

The paper is organized as follows. In Section 2, we present the model that allows for multiple structural breaks in the deterministic trend function. Section 3 discusses the feasible point optimal test with multiple structural breaks assuming the break dates to be known. Section 3.1 provides details on the methods to construct the relevant non-centrality parameter arising in the quasi-differencing procedures used. Section 4 analyzes the M-class and related unit root tests allowing for multiple breaks. Section 5 considers the case with unknown break dates and show that if these are estimated by minimizing the sum of squared residuals from the appropriate GLS regression, the limit distributions of the tests are the same as in the known break date case, provided breaks are present. Section 5.1 provides details on how to compute the estimate of the break dates. Section 6 provides preliminary simulations showing that the tests perform well but that they exhibit important size distortions when no break occurs. A solution to this problem is offered in Section 7 based on a pre-test for changes in the slope of a trend function allowing the noise component to be stationary or integrated based on the work of Perron and Yabu (2006). We show that

[2] Related papers allowing for multiple changes include the following. Lee (1996) and Lumsdaine and Papell (1997), for trending variables, and Carrion, Sanso and Artis (2004), for non-trending variables, generalized the approach in Zivot and Andrews (1992), while Clementes, Montañés and Reyes (1998) extends the work of Perron and Vogelsang (1992a,b) for non—trending variables. Lee and Strazicich (2003) extend Schmidt and Phillips's (1992) LM test to allow for two structural breaks both under the null and the alternative hypotheses. For multiple structural breaks, Ohara (1999) and Kapetanios (2005) generalized the approach in Zivot and Andrews (1992). Gadea, Montañés and Reyes (2004) designed a pseudo F statistic to account for multiple level shifts for non-trending variables. Finally, Bai and Carrion-i-Silvestre (2004) consider the square of the modified Sargan-Bhargava statistic to the presence of multiple structural breaks that might affect either the level and/or the slope of the time trend. Hatanaka and Yamada (1999) is most closely related to the approach of Kim and Perron (2006) and deals with two breaks for the so-called IO (innovational outlier) type model.

[3] Perron and Rodriguez (2003) also consider GLS-detrended type procedures allowing a single break. The approach is, however, quite different. Also, while their treatment yields "optimal" tests when a break is present, this is not the case if a break is absent.

the resulting procedure involving the pre-test has good size and power superior to that of alternative methods. Section 8 offers brief concluding remarks and an appendix the proofs of various theoretical results.

2 The model

Let y_t be a stochastic process generated according to

$$\begin{aligned} y_t &= d_t + u_t && (1) \\ u_t &= \alpha u_{t-1} + v_t,\ t = 0, \ldots, T, && (2) \end{aligned}$$

where $\{u_t\}$ is an unobserved mean-zero process. We assume that $u_0 = 0$, although the results generally hold for the weaker requirement that $E\left(u_0^2\right) < \infty$. The disturbance term v_t is defined by $v_t = \sum_{i=0}^{\infty} \gamma_i \eta_{t-i}$ with $\sum_{i=0}^{\infty} i \left|\gamma_i\right| < \infty$ and $\{\eta_t\}$ a martingale difference sequence adapted to the filtration $F_t = \sigma - field\{\eta_{t-i}; i \geq 0\}$. We define the long-run and short-run variance as $\sigma^2 = \sigma_\eta^2 \gamma\left(1\right)^2$ and $\sigma_\eta^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E\left(\eta_t^2\right)$, respectively.

We consider three models: Model 0 ("level shift" or "crash"'), Model I ("slope change" or "changing growth"), and Model II ("mixed change") [4]. Let $DU_t(T_j^0) = 1$ and $DT_t^*(T_j^0) = (t - T_j^0)$ for $t > T_j^0$ and 0 elsewhere, with $T_j^0 = [T\lambda_j^0]$ denoting the j-th break date, with $[\cdot]$ the integer part, and $\lambda_j^0 \equiv T_j^0/T \in (0, 1)$ the break fraction parameter. As a matter of notation, all true break fractions and break dates are denoted with a superscript 0. Estimates of the break fractions and break dates are denoted with a hat. We also use the convention that $T_0^0 = 0$ and $T_{m+1}^0 = T$. We collect the m break fraction parameters in the vector $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)'$. For now it is assumed that the break dates are known; this will be relaxed later.

The deterministic component in (1) is given by

$$d_t = z_t'(T_0^0)\psi_0 + z_t'(T_1^0)\psi_1 + \cdots + z_t'(T_m^0)\psi_m \equiv z_t'(\lambda^0)\psi \tag{3}$$

where

$$\begin{aligned} z_t(\lambda^0) &= [z_t'(T_0^0), \ldots, z_t'(T_m^0)]', \\ \psi &= (\psi_0', \ldots, \psi_m')'. \end{aligned}$$

The various deterministic components and associated coefficients are defined by

$$z_t(T_0^0) \equiv z_t(0) = (1, t)',$$

[4] Perron and Rodriguez (2003) also considered Models I and II but only with one break.

with $\psi_0 = (\mu_0, \beta_0)'$ and, for $1 \le j \le m$,

$$z_t(T_j^0) = \begin{cases} DU_t(T_j^0), & \text{in Model 0,} \\ DT_t^*(T_j^0), & \text{in Model I,} \\ (DU_t(T_j^0), DT_t^*(T_j^0))', & \text{in Model II,} \end{cases}$$

with $\psi_j = \mu_j$ in Model 0, $\psi_j = \beta_j$ in Model I, and $\psi_j = (\mu_j, \beta_j)'$ in Model II.

For Models 0 and II, we also consider the case where the magnitude of level shifts get large as the sample size grows, i.e., $(\mu_1, \ldots, \mu_m) = T^{1/2+\eta}(\kappa_1, \ldots, \kappa_m)$, with $\eta > 0$. The models are then labelled as Models 0b and IIb, respectively. These models are useful to obtain better approximations of the properties of the tests in finite samples. In Models 0 and II, the level shifts belong to the class of "slowly evolving trend" defined by Elliott et al. (1996) and have no effect on the asymptotic size and power of the tests. When the magnitude of the shifts are non-negligible, this typically implies that the derived asymptotic distribution is a bad approximation to the finite sample distribution. In Models 0b and IIb, the level shifts do not belong to the class of "slowly evolving trend". As shown in Harvey, Leybourne and Newbold (2001) this framework provides better approximations. A second feature of importance is related to the estimation of the unknown break fractions. As shown by Perron and Zhu (2005), the rate of convergence increases when the level shifts are modeled to increase as the sample size grows. This phenomenon has important implications for the properties of the unit root tests, as will be shown in the following sections.

Remark 1 *It is possible to assume that there is no trending deterministic component in Model 0. We do not explicitly consider this case in the subsequent analysis but it should be understood that most of our results pertaining to Models 0 and 0b can readily be applied to the level shift model with no time trend.*

The so-called GLS detrended unit root test statistics are based on the use of the quasi-differenced variables $y_t^{\bar{\alpha}}$ and $z_t^{\bar{\alpha}}(\lambda^0)$ defined by

$$y_t^{\bar{\alpha}} = (y_1, (1 - \bar{\alpha}L)\, y_t)\,,\ z_t^{\bar{\alpha}}(\lambda^0) = \left(z_1(\lambda^0), (1 - \bar{\alpha}L)\, z_t(\lambda^0)\right),\ t = 2, \ldots, T,$$

with $\bar{\alpha} = 1 + \bar{c}/T$ where $\bar{c}$ is a non-centrality parameter to be defined below. Once the data has been transformed, the parameters ψ, associated with the deterministic components, can be estimated by minimizing the following objective function

$$S^*\left(\psi, \bar{\alpha}, \lambda^0\right) = \sum_{t=1}^{T} \left(y_t^{\bar{\alpha}} - \psi' z_t^{\bar{\alpha}}(\lambda^0)\right)^2. \tag{4}$$

We denote the minimum of this function by $S\left(\bar{\alpha}, \lambda^0\right)$.

3 Feasible point optimal test with multiple structural breaks

The choice of the non-centrality parameter $\bar{c}$ is related to the Gaussian point optimal statistic to test the null hypothesis of $\alpha = 1$ in (2) against the alternative hypothesis that $\alpha = \bar{\alpha}$, as suggested by Elliott et al. (1996). Following their analysis and Perron and Rodriguez (2003), the feasible point optimal statistic is given by

$$P_T^{GLS}\left(\lambda^0\right) = \left\{S\left(\bar{\alpha}, \lambda^0\right) - \bar{\alpha} S\left(1, \lambda^0\right)\right\} / s^2(\lambda^0), \tag{5}$$

where $s^2(\lambda^0)$ is an estimate of the spectral density at frequency zero of v_t. Following Ng and Perron (2001) and Perron and Ng (1998), we use an autoregressive estimate defined by

$$s(\lambda^0)^2 = s_{ek}^2/(1 - \hat{b}\,(1))^2, \tag{6}$$

where $s_{ek}^2 = (T-k)^{-1} \sum_{t=k+1}^{T} \hat{e}_{t,k}^2$, $\hat{b}\,(1) = \sum_{j=1}^{k} \hat{b}_j$, with $\hat{b}_j$ and $\hat{e}_{t,k}$ obtained from the OLS estimation of

$$\Delta \widetilde{y}_t = b_0 \widetilde{y}_{t-1} + \sum_{j=1}^{k} b_j \Delta \widetilde{y}_{t-j} + e_{t,k}, \tag{7}$$

with $\widetilde{y}_t = y_t - \hat{\psi}' z_t(\lambda^0)$, where $\hat{\psi}$ minimizes (4). The order of the autoregression k is selected using the modified information criteria suggested by Ng and Perron (2001) with the modification proposed by Perron and Qu (2007).

Let $W_c\,(r)$ be an Ornstein-Uhlenbeck process, i.e., the solution to the stochastic differential equation $dW_c\,(r) = cW_c\,(r)\,dr + dW\,(r)$, with $W_c\,(0) = 0$ where $W\,(r)$ is the standard Brownian motion. Denoting by "$\Rightarrow$" weak convergence of the associated measure of probability, the limiting distribution of the test $P_T^{GLS}(\lambda^0)$ is given in the following Theorem.

Theorem 1 *Let $\{y_t\}_{t=1}^{T}$ be the stochastic process generated according to (1) and (2) with $\alpha = 1 + c/T$. Let $P_T^{GLS}(\lambda^0)$ be the statistic defined by (5) and $s^2(\lambda^0)$ be a consistent estimate of σ^2.*

(i) For Models 0 and 0b:

$$P_T^{GLS}\left(\lambda^0\right) \Rightarrow \bar{c}^2 \int_0^1 V_{c,\bar{c}}^2(r)dr + (1-\bar{c})V_{c,\bar{c}}^2(1) \equiv K^{P_T^{GLS}}\left(c, \bar{c}\right).$$

where $V_{c,\bar{c}}^2(r) = W_c(r) - r[bW_c(1) + 3(1-b)\int_0^1 sW_c(s)ds]$ and $b = (1-\bar{c})/(1-\bar{c}+\bar{c}^2/3)$.

(ii) For Models I, II, and IIb:

$$P_T^{GLS}\left(\lambda^0\right) \Rightarrow M\left(c,0,\lambda^0\right) - M\left(c,\bar{c},\lambda^0\right) - 2\bar{c}\int_0^1 W_c(r)\,dW(r)$$
$$+\left(\bar{c}^2 - 2\bar{c}c\right)\int_0^1 W_c(r)^2\,dr - \bar{c} \equiv H^{P_T^{GLS}}\left(c,\bar{c},\lambda^0\right),$$

where $M\left(c,\bar{c},\lambda^0\right) = \bar{V}\left(\lambda^0\right)' A(\lambda^0)^{-1}\bar{V}\left(\lambda^0\right)$, $\bar{V}\left(\lambda^0\right) = \left(V\left(\lambda_0^0\right),\ldots,V\left(\lambda_m^0\right)\right)'$ *with*

$$V(\lambda_i^0) = (1+\bar{c}\lambda_i^0)[W(1) - W(\lambda_i^0) + (c-\bar{c})\int_{\lambda_i^0}^1 W_c(r)dr]$$
$$-\bar{c}\int_{\lambda_i^0}^1 r\,dW(r) - (c-\bar{c})\bar{c}\int_{\lambda_i^0}^1 rW_c(r)dr,$$

and $A(\lambda^0)$ *a symmetric matrix defined by*

$$A(\lambda^0) = \begin{bmatrix} a\left(\lambda_0^0,\lambda_0^0\right) & a\left(\lambda_0^0,\lambda_1^0\right) & \cdots & a\left(\lambda_0^0,\lambda_m^0\right) \\ & a\left(\lambda_1^0,\lambda_1^0\right) & \cdots & a\left(\lambda_1^0,\lambda_m^0\right) \\ & & \ddots & \vdots \\ & & & a\left(\lambda_m^0,\lambda_m^0\right) \end{bmatrix} \quad (8)$$

with

$$a(\lambda_i^0,\lambda_j^0) = (1/6)(1-\lambda_j)\left[6\bar{c}(\lambda_i - 1) + \bar{c}^2\{\lambda_j(3\lambda_i - 1) - 3\lambda_i - \lambda_j^2 + 2\} + 6\right]$$

The proof for Model II is given in the Appendix. Since Models 0 and I can be viewed as a special case of Model II, no separate proof is provided. The limiting distribution in *(i)* is the same as that of the linear time trend model with no break, which can be found in Elliott et al. (1996). Because the break dates are assumed known here, the test statistic $P_T^{GLS}(\lambda^0)$ is exactly invariant to the value of the coefficients associated with all regressors including those pertaining to the change in the trend. Hence, there is no distinction between Models 0 and 0b, and between Models II and IIb. The limiting distribution of the test statistic for Models I, II, and IIb depends both on the number of structural breaks and on the vector of break fractions.

3.1 The choice of the non-centrality parameter $\bar{c}$

From the limiting distributions in Theorem 1, we can obtain the local Gaussian power envelope for the various cases. For Models I, II and IIb, it is defined by $\pi^*(c) = \Pr[H^{P_T^{GLS}}(c,c,\lambda^0) <$

$b^{P_T^{GLS}}(c,\lambda^0)]$, where, with υ the size of the test, $b^{P_T^{GLS}}(c,\lambda^0)$ is such that $\Pr[H^{P_T^{GLS}}(0,c,\lambda^0) < b^{P_T^{GLS}}(c,\lambda^0)] = \upsilon$. Furthermore, the power envelope allows us to find the "optimal" non-centrality parameter $\bar{c}$ for our models. Elliott et al. (1996) recommended to choose the value $\bar{c}$ such that the asymptotic power of the test is 50%, i.e., $\bar{c}$ is such that $\Pr[H^{P_T^{GLS}}(\bar{c},\bar{c}) < b^{P_T^{GLS}}(\bar{c})] = 0.5$.

For Models I, II, and IIb the parameter $\bar{c}$ depends on the number of structural breaks and their positions. Instead of reporting extensive tables of values, we opted to summarize the relevant information via a response surface analysis. In order to do so, we obtained by simulations the parameter $\bar{c}$ for up to $m = 5$ structural break points for all possible combinations of break fraction vectors $\lambda^0 = (\lambda_1^0, \cdots, \lambda_m^0)'$, with $\lambda_i^0 = \{0.1, 0.2, \cdots, 0.9\}$ (we used 1,000 steps to approximate the Wiener process and 10,000 replications). This gives us 382 cases and the response surface presented will allow an accurate approximation for these and other cases. Visual inspection of the results obtained revealed U-shaped pattern between the estimated parameters $\bar{c}$ and the different vectors λ^0. Furthermore, the estimated values of $\bar{c}$ showed symmetry around $\lambda_i^0 = 0.5$. Therefore, we adopted a functional form that accounts for these two features through the introduction of powers of λ^0 and $|\lambda_i^0 - \lambda_j^0|$, $(i,j = 1,\cdots,m)$ as regressors. The functional form we settled upon is given by

$$\bar{c}\left(\lambda_k^0\right) = \beta_{0,0} + \sum_{l=1}^{4}\sum_{i=1}^{m}\beta_{i,l}(\lambda_{i,k}^0)^l + \sum_{l=1}^{4}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\gamma_{i,j,l}\left|\lambda_{i,k}^0 - \lambda_{j,k}^0\right|^l + \varepsilon_k. \tag{9}$$

The estimates of the coefficients of (9) obtained from the 382 cases simulated are reported in Table 1. A simple Gauss program to compute the values of $\bar{c}(\lambda_k^0)$ for a given vector $\lambda^0 = (\lambda_1^0, \cdots, \lambda_m^0)'$ is available on the authors' web pages. Note that for Models 0 and 0b, a similar analysis holds, though much simpler since the limit distribution is invariant to any break parameters. The results are described in Elliott et al. (1996) and $\bar{c} = -13.5$ (if no trend is present $\bar{c} = -7$).

4 The M-class and related unit root tests

Following Perron and Rodriguez (2003), we suggest the use of the so-called M-class of tests analyzed in Ng and Perron (2001) allowing for multiple structural breaks. The tests are defined by

$$MZ_{\alpha}^{GLS}\left(\lambda^{0}\right) = (T^{-1}\widetilde{y}_{T}^{2}-s\left(\lambda^{0}\right)^{2})(2T^{-2}\sum_{t=1}^{T}\widetilde{y}_{t-1}^{2})^{-1}, \quad (10)$$

$$MSB^{GLS}\left(\lambda^{0}\right) = (s\left(\lambda^{0}\right)^{-2}T^{-2}\sum_{t=1}^{T}\widetilde{y}_{t-1}^{2})^{1/2}, \quad (11)$$

$$MZ_{t}^{GLS}\left(\lambda^{0}\right) = (T^{-1}\widetilde{y}_{T}^{2}-s\left(\lambda^{0}\right)^{2})(4s\left(\lambda^{0}\right)^{2}T^{-2}\sum_{t=1}^{T}\widetilde{y}_{t-1}^{2})^{-1/2}, \quad (12)$$

with $\widetilde{y}_t = y_t - \hat{\psi}' z_t(\lambda^0)$, where $\hat{\psi}$ minimizes (4) and $s(\lambda^0)^2$ is defined in (6). Note that we can also test the unit root hypothesis using the t-ratio statistic for $b_0 = 0$ in (7). This is akin to an extension of the Dickey and Fuller (1979) test and is denoted $ADF^{GLS}(\lambda^0)$. Another statistic considered in Ng and Perron (2001) is a modification of the feasible point optimal test. In our context with breaks, it is defined by

$$MP_{T}^{GLS}\left(\lambda^{0}\right)=[\bar{c}^{2}T^{-2}\sum_{t=1}^{T}\widetilde{y}_{t-1}^{2}+(1-\bar{c})T^{-1}\widetilde{y}_{T}^{2}]/s\left(\lambda^{0}\right)^{2}.$$

This test is based on the same motivation that leads to the definition of the M tests in Stock (1999), namely, to provide functionals of sample moments that have the same asymptotic distributions as well known unit root tests. The $MP_T^{GLS}(\lambda^0)$ is important because its limiting distribution coincides with that of the feasible point optimal test. The following Theorem provides the limit null distribution of the various tests considered.

Theorem 2 *Let $\{y_t\}_{t=1}^{T}$ be the stochastic process generated according to (1) and (2) with $\alpha = 1 + c/T$. Then, provided $s^2(\lambda^0)$ be a consistent estimate of σ^2:*

(i) For Models 0 and 0b:

$$MZ_{\alpha}^{GLS}(\lambda^{0}) \Rightarrow 0.5\left(V_{c,\bar{c}}(1)^{2}-1\right)(\int_{0}^{1}V_{c,\bar{c}}(r)^{2}dr)^{-1}$$

$$MSB^{GLS}(\lambda^{0}) \Rightarrow (\int_{0}^{1}V_{c,\bar{c}}(r)^{2}dr)^{1/2}$$

where $V_{c,\bar{c}}^{2}(r) = W_c(r) - r[bW_c(1) + 3(1-b)\int_0^1 sW_c(s)ds]$ and $b = (1-\bar{c})/(1-\bar{c}+\bar{c}^2/3)$.

(ii) For Models I, II and IIb:

$$MZ_{\alpha}^{GLS}(\lambda^{0}) \Rightarrow 0.5\left(V_{c,\bar{c}}(1,\lambda^{0})^{2}-1\right)(\int_{0}^{1}V_{c,\bar{c}}(r,\lambda^{0})^{2}dr)^{-1} \equiv H^{MZ^{GLS}}(c,\bar{c},\lambda^{0})$$

$$MSB^{GLS}(\lambda^{0}) \Rightarrow (\int_{0}^{1}V_{c,\bar{c}}(r,\lambda^{0})^{2}dr)^{1/2} \equiv H^{MSB^{GLS}}(c,\bar{c},\lambda^{0})$$

where $V_{c,\bar{c}}(r,\lambda^0) = W_c(r) - z_2(r)\, A(\lambda^0)^{-1}\bar{V}(\lambda^0)$ *with* $A(\lambda^0)$ *and* $\bar{V}(\lambda^0)$ *as defined in Theorem 1, and* $z_2(r) = (r,\ (r-\lambda_1^0)1(r > \lambda_1^0),\ \ldots,\ (r-\lambda_m^0)1(r > \lambda_m^0))$*, where* $1(\cdot)$ *is the indicator function.*

(iii) The limiting distribution of $MZ_t^{GLS}(\lambda^0)$ *in all models can be obtained in view of the fact that* $MZ_t^{GLS}(\lambda^0) = MZ_\alpha^{GLS}(\lambda^0) \cdot MSB^{GLS}(\lambda^0)$*, which in turn is also the limiting distribution of the test* $ADF^{GLS}(\lambda^0)$*, denoted by* $H^{ADF^{GLS}}(c,\bar{c},\lambda^0)$*.*

Again, the limiting distribution in *(i)* is the same as that of the linear time trend model with no break given in Ng and Perron (2001). Thus, note that the invariance to the break parameters holds for all test statistics for Models 0 and 0b. This is not the case for Models I, II, and IIb, where their limiting distribution depends on the number and location of the break points. The proof of Model II is given in the Appendix, while the proof for the other models easily follows (the proof for the test $ADF^{GLS}(\lambda^0)$ is more tedious but follows the same steps as in Ng and Perron, 2001). This generalizes the results of Perron and Rodriguez (2003) who showed, for the case of a single break, that each of the M tests and the ADF statistic has the same limiting distribution across Models I and II.

For the case of Models I, II and IIb, the limit distributions depend on the number of breaks and their positions. Instead of reporting extensive tables of values, we again opted to summarize the relevant information via a response surface analysis. As above for the non-centrality parameter $\bar{c}$, we obtained by simulations the 1, 2.5, 5 and 10% percentiles of the limit distributions of the various tests for up to $m = 5$ structural break points for all possible combinations of break fraction vectors $\lambda^0 = (\lambda_1^0, \cdots, \lambda_m^0)'$ with $\lambda_i^0 = \{0.1, 0.2, \cdots, 0.9\}$ (again, we used 1,000 steps to approximate the Wiener process and 10,000 replications). This gives us 382 cases and the response surface presented will allow an accurate approximation for these and other cases. The functional form adopted is given by

$$\begin{aligned} cv\left(\lambda_k^0\right) &= \beta_{0,0} + \sum_{l=1}^{2}\sum_{i=1}^{m}\beta_{l,i}(\lambda_{i,k}^0)^l + \sum_{l=1}^{2}(\gamma_{l,0} + \sum_{i=1}^{m}\gamma_{l,i}\lambda_{i,k}^0)\bar{c}\left(\lambda_k^0\right)^l \\ &\quad + \sum_{l=1}^{4}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\delta_{i,j,l}\left|\lambda_{i,k}^0 - \lambda_{j,k}^0\right|^l \bar{c}\left(\lambda_k^0\right) + \varepsilon_k. \end{aligned}$$

The estimates of the coefficients of the response surfaces are reported in Tables 2 to 5. A simple Gauss program to compute the values of $cv\left(\lambda_k^0\right)$ for a given vector $\lambda^0 = (\lambda_1^0, \cdots, \lambda_m^0)'$ is available on the authors' web pages.

The asymptotic power functions of the tests are defined by $\pi^*_{J^{GLS}}(c,\bar{c},\lambda^0) = \Pr[\, H^{J^{GLS}}(c,\bar{c},\lambda^0) < b^{J^{GLS}}(\bar{c},\lambda^0)]$ for $J = MZ_\alpha,\ MSB,\ ADF$ with $H^{J^{GLS}}(c,\bar{c},\lambda^0)$ defined in Theorem 2. The constants $b^{J^{GLS}}(\bar{c},\lambda^0)$ are such that $\Pr[H^{J^{GLS}}(0,\bar{c},\lambda^0) < b^{J^{GLS}}(\bar{c},\lambda^0)] = \upsilon$, the size of the tests. To assess the efficiency in terms of local asymptotic power of the various tests, we consider the case with a single break occurring at $\lambda^0 = 0.3$, 0.5 and 0.7. The asymptotic power functions are shown in Figure 1 where the solid line is the power envelope. As can be seen, the local power functions are nearly identical and indeed very close to the Gaussian power envelope. So from this local asymptotic power perspective, all tests are nearly efficient. For the case of Models 0 and 0b, we already know from Ng and Perron (2001) that the local power functions of the various tests considered are very close to the power envelope.

5 The case with unknown break dates

The analysis so far assumed that the timing of the structural breaks is known [5]. We need to establish a procedure to estimate them and deduce what is the effect on the limit distribution of the various unit root tests. We propose to estimate the break dates via a global minimization of the sum of squared residuals (SSR) of the GLS-detrended model [6], i.e., $\hat{\lambda} = \arg\min_{\lambda\in\Lambda(\varepsilon)} S(\bar{\alpha},\lambda)$, so that

$$S(\bar{\alpha},\hat{\lambda}) = \min_{\lambda\in\Lambda(\varepsilon)} S(\bar{\alpha},\lambda), \tag{13}$$

where the infimum is taken over all possible break fraction vectors defined on the set

$$\Lambda(\varepsilon)=\{(\lambda_1,...,\lambda_m);\ |\lambda_{i+1}-\lambda_i| \geq \varepsilon\ (i=1,...,m-1), \lambda_1 \geq \varepsilon, \lambda_m \leq 1-\varepsilon\},$$

with ε some trimming parameter that dictates the minimal length of a segment. A common value in the related literature is $\varepsilon = 0.15$. The following Proposition establishes the consistency and rate of convergence of the estimate of the vector of break fractions λ^0.

Proposition 1 *Let $\{y_t\}_{t=1}^T$ be the stochastic process generated according to (1) and (2) with $\alpha = 1$. Assume that $m > 0$ and $\psi_j \neq 0$ ($j = 1,\ldots,m$), so that there are structural breaks affecting y_t under the null hypothesis. Let $\hat{\lambda} = \arg\min_{\lambda\in\Lambda(\varepsilon)} S(\bar{\alpha},\lambda)$, then, as $T \to \infty$:*

(i) in Models I and II:

$$||\hat{\lambda}-\lambda^0|| = O_p\left(T^{-1}\right),$$

(ii) in Models 0b and IIb:

$$||\hat{\lambda}-\lambda^0|| = o_p\left(T^{-1}\right).$$

[5] In this paper, we shall not address the issue of estimating the number of breaks. One natural possibility is to use an information criterion such as the BIC as suggested by Yao (1988).

[6] Note that this approach is different from the one adopted in Perron and Rodriguez (2003), who estimated the location of the break point through the minimization of the SSR under the null and alternative hypotheses.

The proof is given in the Appendix. Following Kim and Perron (2006), the next step is to derive the limit distribution of the unit root tests when this estimate of the break fractions is used instead of the true values. The next Proposition, proved in the Appendix, demonstrates that the rate of convergence is fast enough to guarantee that we recover the same limit distribution as in the known break date case.

Proposition 2 *Let $\{y_t\}_{t=1}^T$ be the stochastic process generated according to (1) and (2) with $\alpha = 1$. Assume that $m > 0$ and $\psi_j \neq 0$ ($j = 1, ..., m$) and that $s(\hat{\lambda})^2$ is a consistent estimate of σ^2. Let $\hat{\lambda} = \arg\min_{\lambda \in \Lambda(\varepsilon)} S(\bar{\alpha}, \lambda)$, then the limit distributions of $P_T^{GLS}(\hat{\lambda})$, $MP_T^{GLS}(\hat{\lambda})$, $MZ_\alpha^{GLS}(\hat{\lambda})$, $MSB^{GLS}(\hat{\lambda})$, $MZ_t^{GLS}(\hat{\lambda})$ and $ADF^{GLS}(\hat{\lambda})$ are the same as those of $P_T^{GLS}(\lambda^0)$, $MP_T^{GLS}(\lambda^0)$, $MZ_\alpha^{GLS}(\lambda^0)$, $MSB^{GLS}(\lambda^0)$, $MZ_t^{GLS}(\lambda^0)$ and $ADF^{GLS}(\lambda^0)$, respectively, for all Models.*

This result is important since even if the break dates are unknown, the use of the particular estimate $\hat{\lambda}$ considered allows us to obtain unit root tests with the same limit distribution as in the known break date case. Since the latter have a local asymptotic power function close to the power envelop, this implies that whether the break dates are known or not the same optimality properties hold, and that in the case of Gaussian errors we cannot do better in terms of local asymptotic power. Note that the result is different from that of Kim and Perron (2006) who considered a framework using OLS based on a regression involving the raw variables (not quasi-detrended). For Models I and II, the rate of convergence was not fast enough to obtain such an equivalence and to solve the problem they proposed a procedure involving data trimmed around the estimate of the break date. No such device is needed here with the GLS-type procedure.

5.1 Computation of the estimates of the break dates

In practice, the computation of the estimates of the break dates defined in (13) is computationaly prohibitive using a regular grid search when $m > 2$. Following Bai and Perron (2003), we propose to use a dynamic programming approach. The procedure involves, however, an additional layer of difficulty since the quasi-differencing used in (13) destroys the block-diagonality of the matrix of regressors. We can, nevertheless, recover a block diagonal matrix provided appropriate restrictions are imposed on the coefficients. We start by noting that the matrix of regressors is given by (assuming only two breaks):

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,2 & 0 & 0 & 0 & 0 \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,3 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,T_1^0 & 0 & 0 & 0 & 0 \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_1^0+1) & 1 & 1 & 0 & 0 \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_1^0+2) & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,T_2^0 & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_2^0-T_1^0) & 0 & 0 \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_2^0+1) & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_2^0-T_1^0+1) & 1 & 1 \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_2^0+2) & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_2^0-T_1^0+2) & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,T & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T-T_1^0) & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T-T_2^0) \end{bmatrix}.$$

For $t = 1$, we have

$$y_1^{\bar{\alpha}} = \mu_0 + \beta_0 + u_1,$$

which can be expressed as

$$\begin{aligned} y_1^{\bar{\alpha}} &= \mu_0 + \beta_0 \pm \mu_0\,(1-\bar{\alpha}) \pm \beta_0\,(1-\bar{\alpha}L)\,1 + u_1 \\ &= \mu_0\,(1-\bar{\alpha}) + \beta_0\,(1-\bar{\alpha}L)\,1 + [\mu_0 + \beta_0 - \mu_0\,(1-\bar{\alpha}) - \beta_0\,(1-\bar{\alpha}L)\,1] + u_1 \\ &= \mu_0\,(1-\bar{\alpha}) + \beta_0\,(1-\bar{\alpha}L)\,1 + [\mu_0 + \beta_0 - \mu_0\,(1-\bar{\alpha}) - \beta_0\,(1-\bar{\alpha}L)\,1]\,D_1\,(T_0) + u_1, \end{aligned}$$

where $D_1\,(T_j)$ denotes an impulse dummy variable defined as $D_t\,(T_j) = 1$ for $t = T_j + 1$ and zero otherwise, with the convention that $T_0 = 0$. Therefore, the model for $1 \leq t \leq T_1^0$ can be written as

$$y_t^{\bar{\alpha}} = \mu_0 1^{\bar{\alpha}} + \beta_0 t^{\bar{\alpha}} + \gamma_0 D_t\left(T_0^0\right) + u_t^{\bar{\alpha}},$$

with $\gamma_0 = \mu_0\bar{\alpha} + \beta_0 - \beta_0(1-\bar{\alpha}L)\cdot 1 = \mu_0\bar{\alpha}$, using the fact that $(1-\bar{\alpha}L)t = t-(1+\bar{c}/T)(t-1) = 1-(\bar{c}/T)(t-1)$. When $t = T_1^0 + 1$ the quasi-differenced variable is given by

$$\begin{aligned}
y_{T_1^0+1}^{\bar{\alpha}} &= \mu_0(1-\bar{\alpha}) + \beta_0(1-\bar{\alpha}L)\left(T_1^0+1\right) + \mu_1 + \beta_1 + u_{T_1^0+1}^{\bar{\alpha}} \\
&= \mu_0(1-\bar{\alpha}) + \beta_0(1-\bar{\alpha}L)\left(T_1^0+1\right) \\
&\quad \pm\mu_1(1-\bar{\alpha}) \pm \beta_1(1-\bar{\alpha}L)\left(T_1^0+1\right) + \mu_1 + \beta_1 + u_{T_1^0+1}^{\bar{\alpha}} \\
&= \mu_0(1-\bar{\alpha}) + \beta_0(1-\bar{\alpha}L)\left(T_1^0+1\right) \\
&\quad +\mu_1(1-\bar{\alpha}) + \beta_1(1-\bar{\alpha}L)\left(T_1^0+1\right) \\
&\quad +\left[\mu_1 + \beta_1 - \mu_1(1-\bar{\alpha}) - \beta_1(1-\bar{\alpha}L)\left(T_1^0+1\right)\right] D_{T_1^0+1}\left(T_1^0\right) + u_{T_1^0+1}^{\bar{\alpha}} \\
&= \left(\mu_0 + \mu_1 - \beta_1 T_1^0\right)(1-\bar{\alpha}) + (\beta_0+\beta_1)(1-\bar{\alpha}L)\left(T_1^0+1\right) \\
&\quad +\bar{\alpha}\mu_1 D_{T_1^0+1}\left(T_1^0\right) + u_{T_1^0+1}^{\bar{\alpha}},
\end{aligned}$$

and for $T_1^0 < t \leq T_2^0$ it is given by

$$y_t^{\bar{\alpha}} = \left(\mu_0 + \mu_1 - \beta_1 T_1^0\right)(1-\bar{\alpha}) + (\beta_0+\beta_1)(1-\bar{\alpha}L)t + \gamma_1 D_t\left(T_1^0\right) + u_t^{\bar{\alpha}},$$

where $\gamma_1 = \bar{\alpha}\mu_1$. A similar expression can be obtained for $T_2^0 < t \leq T$, viz.,

$$y_t^{\bar{\alpha}} = \left(\mu_0 + \mu_1 + \mu_2 - \beta_1 T_1^0 - \beta_2 T_2^0\right)(1-\bar{\alpha}) + (\beta_0+\beta_1+\beta_2)(1-\bar{\alpha}L)t + \gamma_2 D_t\left(T_2^0\right) + u_t^{\bar{\alpha}},$$

with $\gamma_2 = \bar{\alpha}\mu_2$. In general, $y_t^{\bar{\alpha}}$ in the j^{th} regime ($T_{j-1}^0 < t \leq T_j^0$), $j = 1,\ldots,m+1$, can therefore be written as

$$y_t^{\bar{\alpha}} = \mu^*(1-\bar{\alpha}) + \beta^*(1-\bar{\alpha}L)t + \gamma_{j-1}D_t(T_{j-1}^0) + u_t^{\bar{\alpha}}, \tag{14}$$

where the coefficients satisfy the following restrictions

$$\begin{aligned}
\mu^* &= \left(\sum_{i=0}^{j-1}\mu_i - \sum_{i=1}^{j-1}\beta_i T_i^0\right), \\
\beta^* &= \left(\sum_{i=0}^{j-1}\beta_i\right), \\
\gamma_{j-1} &= \bar{\alpha}\mu_{j-1}.
\end{aligned}$$

Therefore, the moment matrix of the regressors can be expressed via the following block diagonal matrix,

$$\begin{bmatrix} (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,1 & 1 & & & & & & & 0 \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,2 & 0 & & & & & & & \\ \vdots & \vdots & \vdots & & & & & & & \\ (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,T_1^0 & 0 & & & & & & & \\ & & & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_1^0+1) & 1 & & & & \\ & & & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_1^0+2) & 0 & & & & \\ & & & \vdots & \vdots & \vdots & & & & \\ & & & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,T_2^0 & 0 & & & & \\ & & & & & & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_2^0+1) & 1 & \\ & & & & & & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,(T_2^0+2) & 0 & \\ & & & & & & \vdots & \vdots & \vdots & \\ 0 & & & & & & (1-\bar{\alpha}) & (1-\bar{\alpha}L)\,T & 0 & \end{bmatrix},$$

and the dynamic programming algorithm can be used provided the appropriate restrictions on the coefficients of the system (14) are imposed. With this specification, the estimation of the break dates can be done using an iterative procedure similar to that of Perron and Qu (2006). The exact steps are as follows.

1. Compute initial estimates of the break dates and associated break fractions $\hat{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_m)$ and coefficients, $\hat{\psi} = (\hat{\psi}_0', \hat{\psi}_1', \ldots, \hat{\psi}_m')'$ using an OLS method applied to (1). This involves a standard application of the algorithm described in Bai and Perron (2003).

2. For a given set of preliminary estimates of the break dates, obtain an initial value of $\bar{c}(\hat{\lambda})$ using (9).

3. Let $T^*(\psi, r, n) = (T_1^*(\psi, r, n), \ldots, T_r^*(\psi, r, n))$ be the vector of the optimal r break dates using the first n observations for a given vector of coefficients ψ, and $RSSR(T^*(\psi, r, n))$ be the associated restricted sum of squared residuals. Then, compute the restricted sum of squared residuals $RSSR(T^*(\psi, 1, n))$ for $2h \leq n \leq T - (m-1)h$ by

$$RSSR(T^*(\hat{\psi}, 1, n)) = \min_{h \leq j \leq n-h} [RSSR^1(1, j) + RSSR^2(j+1, n)]$$

and

$$T^*(\hat{\psi},1,n) = \arg\min_{h\leq j\leq n-h}[RSSR^1(1,j) + RSSR^2(j+1,n)],$$

where

$$RSSR^1(1,j) = \sum_{t=1}^{j}(y_t^{\bar{\alpha}} - z_t^{\bar{\alpha}}(T_0)'\hat{\phi}_0 - \hat{\gamma}_0 D_t(T_0))^2,$$

with $\hat{\phi}_0 = \hat{\psi}_0$ and $\hat{\gamma}_0 = \bar{\alpha}\hat{\mu}_0$, and

$$RSSR^2(j+1,n) = \sum_{t=j+1}^{n}(y_t^{\bar{\alpha}} - z_t^{\bar{\alpha}}(T_0)'\hat{\phi}_1(j) - \hat{\gamma}_1 D_t(j))^2,$$

with $\hat{\phi}_1(j) = (\hat{\mu}_0 + \hat{\mu}_1 - \hat{\beta}_1 j, \hat{\beta}_0 + \hat{\beta}_1)'$ and $\hat{\gamma}_1 = \bar{\alpha}\hat{\mu}_1$. Then, sequentially compute and store $RSSR(T^*(\hat{\psi},r,n))$ for $r = 2,\ldots,m-1$, with n ranging from $(r+1)h$ to $T-(m-r)h$. This is done solving

$$RSSR(T^*(\hat{\psi},r,n)) = \min_{rh\leq j\leq n-h}[RSSR(T^*(\hat{\psi},r-1,n)) + RSSR^{r+1}(j+1,n)].$$

The last r^{th} element is

$$T^*(\hat{\psi},r,n) = \arg\min_{rh\leq j\leq n-h}[RSSR(T^*(\hat{\psi},r-1,n)) + RSSR^{r+1}(j+1,n)],$$

where

$$RSSR^{r+1}(j+1,n) = \sum_{t=j+1}^{n}(y_t^{\bar{\alpha}} - z_t^{\bar{\alpha}}(T_0)'\hat{\phi}_r(j) - \hat{\gamma}_r D_t(j))^2, \tag{15}$$

with $\hat{\phi}_r(j) = (\sum_{i=0}^{r}\hat{\mu}_i - \sum_{i=1}^{r-1}\hat{\beta}_i T_i^*(\hat{\psi},r-1,j) - \beta_r j, \sum_{i=0}^{r}\hat{\beta}_i)'$ and $\hat{\gamma}_r = \bar{\alpha}\mu_r$. Finally compute

$$RSSR(T^*(\hat{\psi},m-1,n)) = \min_{mh\leq j\leq T-h}[RSSR(T^*(\hat{\psi},m-1,n)) + RSSR^{m+1}(j+1,T)],$$

where $RSSR^{m+1}(j+1,T)$ is computed as in (15). Then store the estimated break dates and update $\bar{c}(\hat{\lambda})$ accordingly.

4. Repeat steps 2 and 3 until convergence.

6 Preliminary simulations

Even though we have assumed the existence of a break in the trend function so far, it is instructive to analyze the properties of the tests proposed. We consider cases with a single break without loss of generality. The data generating process used in the simulations is given by

$$y_t = d_t + u_t \quad (16)$$
$$d_t = \mu_b DU_t(T_1^0) + \beta_b DT_t^*(T_1^0) \quad (17)$$
$$u_t = \alpha u_{t-1} + v_t, \quad (18)$$

$v_t \sim i.i.d.\ N(0,1)$, $u_0 = 0$. We specified four values for the magnitude of the level shift $\mu_b = \{0, 0.5, 1, 5\}$. For each value of μ_b, we considered values of β_b ranging from -4 to 4 in increments of 0.2. We investigated the sensitivity of the results using three different values of the fraction $\lambda^0 = 0.3$, 0.5, and 0.7. The sample size is set at $T = \{100, 200, 300\}$ and the results are based on 1,000 replications. The empirical size is analyzed setting $\alpha = 1$, while the power is evaluated with $\alpha = \bar{\alpha} = 1 + \bar{c}/T$, where the parameter $\bar{c}$ depends on the break fractions and is obtained from (9). Hence, in large samples, the power is 50% in all cases. In this paper, we only report results for $\lambda^0 = 0.5$ and for the tests $P_T^{GLS}(\hat{\lambda})$, $MZ_\alpha^{GLS}(\hat{\lambda})$ and $ADF^{GLS}(\hat{\lambda})$ (the properties of $MP_T^{GLS}(\hat{\lambda})$ are similar to those of $P_T^{GLS}(\hat{\lambda})$, and the properties of $MSB^{GLS}(\hat{\lambda})$ and $MZ_t^{GLS}(\hat{\lambda})$ similar to those of $MZ_\alpha^{GLS}(\hat{\lambda})$). The full set of results are available on the authors' web sites.

Figures 2 to 4 present the results of the empirical size using the asymptotic critical values at the 5% level of significance drawn from the estimated response surfaces, while Figures 5 to 7 present results for power. For size, the following features are worthy of note. First, for all tests, the exact size is close to the nominal size as $|\beta_b|$ increases and more so the larger the sample. This is in accordance with our theoretical results. When $|\beta_b|$ and $|\mu_b|$ are small, the tests show liberal size distortions and more so as the sample size increases. This is due to the fact that our theoretical results so far assumed $\beta_b \neq 0$ so that a break is present. The limit null distributions when no break occurs are indeed different since the estimate of the break fraction has a non-degenerate limit distribution on the interval $[0, 1]$ instead of converging to either 0 or 1. When comparing the tests, the $P_T^{GLS}(\hat{\lambda})$ and $MZ_\alpha^{GLS}(\hat{\lambda})$ tests perform similarly while the $ADF^{GLS}(\hat{\lambda})$ test clearly exhibits more size distortions even when $|\beta_b|$ is very large. What is of special interest is the fact that when $|\mu_b|$ is large, the exact size of the $P_T^{GLS}(\hat{\lambda})$ and $MZ_\alpha^{GLS}(\hat{\lambda})$ tests are close to the nominal 5% level irrespective of the value of β_b. This suggests that a correction will be needed to account for cases with both $|\beta_b|$ and $|\mu_b|$ small.

With respect to power, the following features emerge. First, the power of the tests $P_T^{GLS}(\hat{\lambda})$ and $MZ_\alpha^{GLS}(\hat{\lambda})$ quickly approaches the 50% limit value suggested by the local asymptotic power analysis unless T is small. This is so even if $|\beta_b|$ is not very large. For very small value of $|\beta_b|$, the power is above 50% but only because of the presence of important size distortions. The power of the $ADF^{GLS}(\hat{\lambda})$ is also above the target 50% level suggesting a superior performance but again this is simply the consequence of important size distortions.

7 Extension to the case where a break need not occur

Up to this point, we assumed that a break occurs under both the null and alternative hypotheses. When no break occurs, the asymptotic results described in the previous sections do not hold because, under the null hypothesis of a unit root, the estimates of the break fractions have a non-degenerate limit distribution on the interval $[0, 1]$ instead of converging to either 0 or 1 (see, Nunes, Kuan and Newbold, 1995, Bai, 1998). If there is no break in the trend, then the proper unit root test procedure is to simply apply a standard Dickey-Fuller (1979) type test with no break dummies. Hence, what is needed is a pre-test to assess whether a break is present or not. This pre-test for a break should have the correct size if the noise is integrated but should also be powerful whether the noise is stationary or has a unit root in order to ensure a specification that allows a unit root test procedure with good power. This testing problem has recently been addressed by Vogeslang (2001) and Perron and Yabu (2006). Since the procedure of Perron and Yabu (2006) has better size and power, we shall use it as the pre-test. It is based on a quasi-GLS approach using an autoregression for the noise component, with a truncation to 1 when the sum of the autoregressive coefficients is in some neighborhood of 1, along with a bias correction. For given break dates, one constructs the F-test for the null hypothesis of no structural change in the deterministic components. The final statistic uses the *Exp* functional of Andrews and Ploberger (1994). The test has virtually the same asymptotic size whether the noise component is stationary or integrated. We label this test as $Exp\text{-}W_{FS}$ and define the alternative estimate of the break fraction, $\widetilde{\lambda}$, as

$$\widetilde{\lambda} = \hat{\lambda} \cdot 1(Exp\text{-}W_{FS} > cv)$$

where cv is the critical value for a test with, say, nominal size p%. When there is a break, $\widetilde{\lambda}$ consistently estimates λ^0, as long as $\hat{\lambda}$ consistently estimates λ^0, given that $Exp\text{-}W_{FS}$ is a consistent test. When there is no break, $\widetilde{\lambda}$ will yield a non zero estimate p% of times even with infinitely many observations, if cv is fixed. Now, suppose that $Exp\text{-}W_{FS} = O_p(T^{\varpi})$, $\varpi > 0$ under the alternative of a break. Let $cv = cT^{\varpi-\varepsilon}$, $0 < \varepsilon < \varpi$. While such an increasing sequence of critical values does not harm the consistency of the test, the size converges to zero as T increases, which ensures the consistency of $\widetilde{\lambda}$ for $\lambda^0 \in [0,1]$. Given the consistency of $\widetilde{\lambda}$, the modified test procedure is to use a DF type test when $\widetilde{\lambda} = 0$ and the procedure described above when $\widetilde{\lambda} \neq 0$.

7.1 Simulation evidence

We performed simulations with the same data generating process and design as used in Section 6 but this time for the version of our statistics involving the pre-test. The results are presented in Figures 8 to 10 (for size) and 11 to 13 (for power). With respect to the size of the test, one can see that the size distortions for values of $|\beta_b|$ near zero disappear to a large extent. The tests $P_T^{GLS}(\hat{\lambda})$ and $MZ_\alpha^{GLS}(\hat{\lambda})$ with the pre-test for break in trend have an exact size close to the nominal level in all cases. The $ADF^{GLS}(\hat{\lambda})$ test does, however, still show important liberal size distortions and is therefore not recommended for practical implementations.

The power function of $P_T^{GLS}(\hat{\lambda})$ and $MZ_\alpha^{GLS}(\hat{\lambda})$ with the pre-test for break approaches the target 50% level quickly as T increases when $|\beta_b|$ is moderate to large. When $|\beta_b|$ is near zero, the power is actually higher than 50%. This is due to the fact that with $\beta_b = 0$ the test performed without allowing for a change in trend has a local power function that is actually higher than the test that does allow for breaks. So there is improvements in terms of power as well. The only drawback is for values of $|\beta_b|$ neither moderate nor small. Here, the power dips down. This is due to the fact that the pre-test for break is not powerful enough for this range of values for $|\beta_b|$ so that a standard test without breaks is applied. Yet, the break is large enough to affect the power of the test which decreases for the reasons explained in Perron (1989).

Overall, the performance of the $P_T^{GLS}(\hat{\lambda})$ and $MZ_\alpha^{GLS}(\hat{\lambda})$ tests is quite satisfactory and offers an improvement in terms of both size and power over existing procedures. It avoids the large size distortions of the Zivot and Andrews (1992) procedure while still allowing for much improved power by taking advantage of the information about the presence or absence of a break whether it be under the null or the alternative hypotheses. The performance of the tests $MP_T^{GLS}(\hat{\lambda})$, $MSB^{GLS}(\hat{\lambda})$ and $MZ_t^{GLS}(\hat{\lambda})$ is similar as evidenced by results unreported.

8 Conclusion

Following the work of Kim and Perron (2006), the procedures suggested in this paper solve many of the problems raised in the introduction that plague most existing unit root tests designed for series with a breaking trend function. First, the tests allow for a break under both the null and the alternative hypotheses. This is desirable for several reasons. It imposes a symmetric treatment when allowing for a break, so that the tests do not reject when the noise is integrated but the trend is changing. Also, if a break is present, this information is exploited to improve power. Second, when a break is present, the limit distributions of the tests are the same as in the case of a known break date, thereby allowing increased power while maintaining the correct size. Our paper used the quasi-GLS procedure suggested by Elliott et al. (1996) to obtain tests that have a local asymptotic power close to the power envelop, except perhaps is a small neighborhood where the change in slope is small but non-zero. We have also extended the analysis to the multiple break case made possible by a modification of the dynamic programming algorithm as described used in Bai and Perron (2003) and Perron and Zhu (2006). Simulation experiments confirm that our procedures offer an improvement over commonly used methods in small samples. Our tests should therefore be useful in empirical applications.

Appendix

We start by presenting some results that will be used throughout. We have, by definition

$$DT^{*\bar{\alpha}}(T_i) = \begin{cases} 0, & t \leq T_i \\ 1 - \bar{c}(t - T_i - 1)/T, & t \geq T_i + 1 \end{cases} \quad \text{and} \quad DU^{\bar{\alpha}}(T_i) = \begin{cases} 0, & t \leq T_i \\ 1, & t = T_i + 1 \\ -\bar{c}/T, & t \geq T_i + 2 \end{cases}$$

Also, for $T_j > T_i$,

$$DT^{*\bar{\alpha}}(T_j) - DT^{*\bar{\alpha}}(T_i) = \begin{cases} 0, & t \leq T_i \\ -1 + \bar{c}(t - T_i - 1)/T, & T_i + 1 \leq t \leq T_j \\ \bar{c}(\lambda_j - \lambda_i), & T_j + 1 \leq t \end{cases}$$

and

$$DU^{\bar{\alpha}}(T_j) - DU^{\bar{\alpha}}(T_i) = \begin{cases} -1, & t = T_i + 1 \\ \bar{c}/T, & T_i + 2 \leq t \leq T_j \\ 1 + \bar{c}/T, & t = T_j + 1 \\ 0, & \text{otherwise} \end{cases}$$

We can then show that, with $T_i \leq T_j$,

$$\begin{aligned} & T^{-1} DT^{*\bar{\alpha}}(T_i)' DT^{*\bar{\alpha}}(T_j) \\ = \; & (1/6)(1 - \lambda_j)\left[6\bar{c}(\lambda_i - 1) + \bar{c}^2\{\lambda_j(3\lambda_i - 1) - 3\lambda_i - \lambda_j^2 + 2\} + 6\right] + o(1) \\ \equiv \; & a(\lambda_i, \lambda_j) + o(1) \end{aligned} \tag{A.1}$$

and

$$DU^{\bar{\alpha}}(T_i)' DU^{\bar{\alpha}}(T_j) = 1(|T_i - T_j| = 0) + o(1). \tag{A.2}$$

Also,

$$\begin{cases} DU^{\bar{\alpha}}(T_i)' DT^{*\bar{\alpha}}(T_j) = 1(|T_i - T_j| = 0) - \bar{c}(1 - \lambda_j) + \bar{c}^2(1 - \lambda_j)^2/2 + o(1), \\ DU^{\bar{\alpha}}(T_j)' DT^{*\bar{\alpha}}(T_i) = 1 - \bar{c}(\lambda_j - \lambda_i) - \bar{c}(1 - \lambda_j) + \bar{c}^2(1 + \lambda_j - 2\lambda_i)(1 - \lambda_j)/2 + o(1), \end{cases} \tag{A.3}$$

where $1(\cdot)$ denotes the indicator function. From (A.1)-(A.3), we have, with $T_i \leq T_j$,

$$\begin{cases} DU^{\bar{\alpha}}(T_j)'[DU^{\bar{\alpha}}(T_j) - DU^{\bar{\alpha}}(T_i)] = 1(|T_i - T_j| \neq 0) + o(1), \\ DU^{\bar{\alpha}}(T_i)'[DU^{\bar{\alpha}}(T_j) - DU^{\bar{\alpha}}(T_i)] = -1(|T_i - T_j| \neq 0) + o(1), \end{cases} \tag{A.4}$$

$$\begin{cases} DU^{\bar{\alpha}}(T_j)'[DT^{*\bar{\alpha}}(T_j) - DT^{*\bar{\alpha}}(T_i)] = \bar{c}(\lambda_j - \lambda_i) - \bar{c}^2(\lambda_j - \lambda_i)(1-\lambda_j) + o(1) \\ DU^{\bar{\alpha}}(T_i)'[DT^{*\bar{\alpha}}(T_j) - DT^{*\bar{\alpha}}(T_i)] \\ = -1(|T_i - T_j| \neq 0) + \bar{c}(\lambda_j - \lambda_i) - \bar{c}^2(\lambda_j - \lambda_i)^2/2 - \bar{c}^2(\lambda_j - \lambda_i)(1-\lambda_j) + o(1), \end{cases} \quad \text{(A.5)}$$

$$\begin{cases} DT^{*\bar{\alpha}}(T_j)'[DU^{\bar{\alpha}}(T_j) - DU^{\bar{\alpha}}(T_i)] = 1(|T_i - T_j| \neq 0) + o(1) \\ DT^{*\bar{\alpha}}(T_i)'[DU^{\bar{\alpha}}(T_j) - DU^{\bar{\alpha}}(T_i)] = -\bar{c}^2(\lambda_j - \lambda_i)^2/2 + o(1), \end{cases} \quad \text{(A.6)}$$

and

$$\begin{cases} T^{-1}DT^{*\bar{\alpha}}(T_j)'[DT^{*\bar{\alpha}}(T_j) - DT^{*\bar{\alpha}}(T_i)] = \bar{c}(\lambda_j - \lambda_i)[(1-\lambda_j) - \bar{c}(1-\lambda_j)^2/2] + o(1) \\ T^{-1}DT^{*\bar{\alpha}}(T_i)'[DT^{*\bar{\alpha}}(T_j) - DT^{*\bar{\alpha}}(T_i)] \\ = -(\lambda_j - \lambda_i) + \bar{c}(\lambda_j - \lambda_i)(1-\lambda_i) - \bar{c}^2(\lambda_j - \lambda_i)\{3(1-\lambda_i)^2 - (\lambda_j - \lambda_i)^2\}/6 + o(1). \end{cases} \quad \text{(A.7)}$$

From (A.4)-(A.7), we have (with $T_j \geq T_i$)

$$||DU^{\bar{\alpha}}(T_j) - DU^{\bar{\alpha}}(T_i)||^2 = 2 \cdot 1(|T_j - T_i| \neq 0) + o(1), \quad \text{(A.8)}$$

$$\begin{aligned} & T^{-1}||DT^{*\bar{\alpha}}(T_j) - DT^{*\bar{\alpha}}(T_i)||^2 \quad \text{(A.9)} \\ = \; & (\lambda_j - \lambda_i) - \bar{c}(\lambda_j - \lambda_i)^2 + \bar{c}^2(\lambda_j - \lambda_i)^3/3 + \bar{c}^2(\lambda_j - \lambda_i)^2(1-\lambda_j) + o(1), \end{aligned}$$

and

$$\begin{aligned} & (DU^{\bar{\alpha}}(T_j) - DU^{\bar{\alpha}}(T_i))'(DT^{*\bar{\alpha}}(T_j) - DT^{*\bar{\alpha}}(T_i)) \quad \text{(A.10)} \\ = \; & 1(|T_i - T_j| \neq 0) + \bar{c}(\lambda_j - \lambda_i)^2/2 + o(1). \end{aligned}$$

Noting that $u_1^{\bar{\alpha}} = v_1$ and $u_t^{\bar{\alpha}} = v_t + T^{-1}(c - \bar{c})u_{t-1}$ $(t = 2, \ldots, T)$, we deduce that

$$\begin{aligned} DU^{\bar{\alpha}}(T_i)'u^{\bar{\alpha}} &= v_{T_i+1} + T^{-1}(c-\bar{c})u_{T_i} + \bar{c}T^{-1}\sum_{t=T_i+2}^{T}(v_t + T^{-1}(c-\bar{c})u_{t-1}) \quad \text{(A.11)} \\ &= v_{T_i+1} + o_p(1) \end{aligned}$$

and

$$T^{-1/2}DT^{*\bar{\alpha}}(T_i)'u^{\bar{\alpha}} = T^{-1/2}\sum_{T_i+1}^{T}(v_t + T^{-1}(c-\bar{c})u_{t-1})(1 - \bar{c}(t - T_i - 1)/T) \Rightarrow \sigma V(\lambda_i), \quad \text{(A.12)}$$

where

$$\begin{aligned} V(\lambda_i) &= (1 + \bar{c}\lambda_i)[W(1) - W(\lambda_i) + (c-\bar{c})\int_{\lambda_i}^1 W_c(r)dr] \quad \text{(A.13)} \\ & \quad -\bar{c}\int_{\lambda_i}^1 r dW(r) - (c-\bar{c})\bar{c}\int_{\lambda_i}^1 rW_c(r)dr. \end{aligned}$$

Proof of Theorem 1: Define the quadratic form $M_T\left(c,\bar{c},\lambda^0\right)=\left(u^{\bar{\alpha}\prime}z^{\bar{\alpha}}\right)\left(z^{\bar{\alpha}\prime}z^{\bar{\alpha}}\right)^{-1}\left(z^{\bar{\alpha}\prime}u^{\bar{\alpha}}\right)$, where $z^{\bar{\alpha}}$ is the vector of the quasi-differences of $z_t(\lambda^0)$ defined in (3). From Elliott et al. (1996) and Perron and Rodriguez (2003), we have

$$s^2(\lambda^0)P_T^{GLS}\left(c,\bar{c},\lambda^0\right) = M_T\left(c,0,\lambda^0\right)-M_T\left(c,\bar{c},\lambda^0\right)-2\bar{c}T^{-1}\sum_{t=2}^{T}u_{t-1}v_t + \left(\bar{c}^2-2\bar{c}c\right)T^{-2}\sum_{t=2}^{T}u_{t-1}^2-\bar{c}T^{-1}u^{1\prime}u^1+o_p\left(1\right). \quad (A.14)$$

From the invariance principle and the Continuous Mapping Theorem (CMT), $T^{-1}\sum_{t=2}^{T}u_{t-1}v_t \Rightarrow \sigma^2[\int_0^1 W_c\left(r\right)dW\left(r\right)+\gamma]$ and $T^{-2}\sum_{t=2}^{T}u_{t-1}^2 \Rightarrow \sigma^2\int_0^1 W_c\left(r\right)^2 dr$ where $\gamma=(\sigma^2-\sigma_v^2)/(2\sigma^2)$ and $\sigma_v^2=Ev_t^2$. By the law of large numbers, $p\lim T^{-1}u^{1\prime}u^1=\sigma_v^2$. Note that expression given in (A.14) is similar to that in Elliott et al. (1996) and Perron and Rodriguez (2003), where the only difference comes from the definition of the quadratic forms $M_T\left(c,0,\lambda^0\right)$ and $M_T\left(c,\bar{c},\lambda^0\right)$. Let $z_t'=\left(z_{t,1}',z_{t,2}'\right)$ be a rearranged version of $z_t(\lambda^0)$ so that $z_{t,1}$ collects the $m+1$ regressors that correspond to the constant and the impulse dummy variables, and $z_{t,2}$ the $m+1$ trending regressors. Also, let $z_t^{\bar{\alpha}\prime}=(z_{t,1}^{\bar{\alpha}\prime},z_{t,2}^{\bar{\alpha}\prime})$ be the quasi-differenced z_t. Then, the scaled matrix $M_T\left(c,\bar{c},\lambda^0\right)$ can be expressed as

$$M_T\left(c,\bar{c},\lambda^0\right)=\left(u^{\bar{\alpha}\prime}z^{\bar{\alpha}}D_T\right)\left(D_Tz^{\bar{\alpha}\prime}z^{\bar{\alpha}}D_T\right)^{-1}\left(D_Tz^{\bar{\alpha}\prime}u^{\bar{\alpha}}\right),$$

where $D_T=diag\{D_{1,T},D_{2,T}\}=diag\left(1,\cdots,1,T^{-1/2},\cdots,T^{-1/2}\right)$. From (A.11) and (A.12), we have

$$D_Tz^{\bar{\alpha}\prime}u^{\bar{\alpha}}=\sigma(v_1/\sigma,...,v_{T_j+1}/\sigma,...,v_{T_m+1}/\sigma,V(\lambda_0^0),...,V(\lambda_j^0),...,V(\lambda_m^0))'+o_p(1). \quad (A.15)$$

Using (A.1)-(A.3), the limit of $D_Tz^{\bar{\alpha}\prime}z^{\bar{\alpha}}D_T$ is given by the following block diagonal matrix:

$$D_Tz^{\bar{\alpha}\prime}z^{\bar{\alpha}}D_T \rightarrow \begin{bmatrix} I_{m+1} & 0 \\ 0 & A(\lambda^0)\end{bmatrix}, \quad (A.16)$$

where I_{m+1} is the identity matrix of order $m+1$ and $A(\lambda^0)$ is a symmetric matrix defined by

$$A(\lambda^0)=\begin{bmatrix} a\left(\lambda_0^0,\lambda_0^0\right) & a\left(\lambda_0^0,\lambda_1^0\right) & \cdots & a\left(\lambda_0^0,\lambda_m^0\right) \\ & a\left(\lambda_1^0,\lambda_1^0\right) & \cdots & a\left(\lambda_1^0,\lambda_m^0\right) \\ & & \ddots & \vdots \\ & & & a\left(\lambda_m^0,\lambda_m^0\right)\end{bmatrix}. \quad (A.17)$$

with $a\left(\lambda_i^0,\lambda_j^0\right)$ as given in (A.1). Therefore, we have

$$M_T\left(c,\bar{c},\lambda^0\right)=v_1^2+\sum_{j=1}^{m}v_{T_j+1}^2+\sigma^2M\left(c,\bar{c},\lambda^0\right)+o_p(1),$$

where $M\left(c,\bar{c},\lambda^0\right)=\bar{V}\left(\lambda^0\right)' A(\lambda^0)^{-1}\bar{V}\left(\lambda^0\right)$, with $\bar{V}\left(\lambda^0\right)=\left(V\left(\lambda_0^0\right),\ldots,V\left(\lambda_m^0\right)\right)'$ and $V(\lambda_j^0)$ as defined in (A.13). This weak convergence result also holds when $\bar{c}=0$ and, thus,

$$M_T\left(c,0,\lambda^0\right)=v_1^2+\sum_{j=1}^{m}v_{T_j+1}^2+\sigma^2 M\left(c,0,\lambda^0\right)+o_p(1).$$

Finally, it follows that the limiting distribution of the test statistic is

$$\begin{aligned} P_T^{GLS}\left(c,\bar{c},\lambda^0\right) &\Rightarrow M\left(c,0,\lambda^0\right)-M\left(c,\bar{c},\lambda^0\right)-2\bar{c}\int_0^1 W_c\left(r\right)dW\left(r\right) \\ &\quad+\left(\bar{c}^2-2\bar{c}c\right)\int_0^1 W_c\left(r\right)^2 dr-\bar{c}\equiv H^{P_T^{GLS}}\left(c,\bar{c},\lambda^0\right). \end{aligned}$$

Proof of Theorem 2: Note that the scaled detrended variable $\widetilde{y}_t$ is given by

$$T^{-1/2}\widetilde{y}_t=T^{-1/2}u_t-T^{-1/2}z_t' D_T\left(D_T z^{\bar{\alpha}\prime}z^{\bar{\alpha}}D_T\right)^{-1}D_T z^{\bar{\alpha}\prime}u^{\bar{\alpha}}.$$

Using (A.15), (A.16) and the asymptotic block diagonality of $\left(D_T z^{\bar{\alpha}\prime}z^{\bar{\alpha}}D_T\right)^{-1}$, we obtain

$$\begin{aligned} T^{-1/2}\widetilde{y}_t &= T^{-1/2}u_t-T^{-1/2}z_{t,1}'\left(z_1^{\bar{\alpha}\prime}z_1^{\bar{\alpha}}\right)^{-1}z_1^{\bar{\alpha}\prime}u^{\bar{\alpha}} \\ &\quad -T^{-1/2}z_{t,2}'D_{2,T}\left(D_{2,T}z_2^{\bar{\alpha}\prime}z_2^{\bar{\alpha}}D_{2,T}\right)^{-1}D_{2,T}z_2^{\bar{\alpha}\prime}u^{\bar{\alpha}}+o_p(1). \end{aligned}$$

Note that $T^{-1/2}z_{[Tr],1}'\left(z_1^{\bar{\alpha}\prime}z_1^{\bar{\alpha}}\right)^{-1}z_1^{\bar{\alpha}\prime}u^{\bar{\alpha}}\xrightarrow{p}0$, provided that $T^{-1/2}v_{T_{b,j}+1}/\sigma\xrightarrow{p}0$, $\forall j=0,\ldots,m$. Also, $T^{-1/2}D_{2,T}z_{[Tr],2}\rightarrow z_2\left(r\right)$ uniformly on $r\in[0,1]$, where $z_2\left(r\right)=(r,\,(r-\lambda_1^0)1(r>\lambda_1^0),$ $\ldots,\,(r-\lambda_m^0)1(r>\lambda_m^0))$. Therefore,

$$T^{-1/2}\widetilde{y}_{[Tr]}\Rightarrow\sigma\left[W_c\left(r\right)-z_2\left(r\right)A(\lambda^0)^{-1}\bar{V}(\lambda^0)\right]=\sigma V_{c,\bar{c}}(r,\lambda^0),$$

with $A(\lambda^0)$ and $\bar{V}\left(\lambda^0\right)$ as defined above. Using this weak convergence result and the CMT, we have

$$MSB^{GLS}\left(\lambda^0\right)=(s\left(\lambda^0\right)^{-2}T^{-2}\sum_{t=1}^{T}\widetilde{y}_{t-1}^2)^{1/2}\Rightarrow(\int_0^1 V_{c,\bar{c}}(r,\lambda^0)^2 dr)^{1/2},$$

provided $s\left(\lambda^0\right)^2\xrightarrow{p}\sigma^2$. For $MZ_\alpha^{GLS}\left(\lambda^0\right)$, we have

$$\begin{aligned} MZ_\alpha^{GLS}\left(\lambda^0\right) &= (T^{-1}\widetilde{y}_T^2-s\left(\lambda^0\right)^2)(2T^{-2}\sum_{t=1}^{T}\widetilde{y}_{t-1}^2)^{-1} \\ &\Rightarrow 0.5\left(V_{c,\bar{c}}(1,\lambda^0)^2-1\right)(\int_0^1 V_{c,\bar{c}}(r,\lambda^0)^2 dr)^{-1}, \end{aligned}$$

and for $MZ_t^{GLS}\left(\lambda^0\right)$,

$$\begin{aligned} MZ_t^{GLS}\left(\lambda^0\right) &= (T^{-1}\widetilde{y}_T^2 - s\left(\lambda^0\right)^2)(4s\left(\lambda^0\right)^2 T^{-2}\sum_{t=1}^{T}\widetilde{y}_{t-1}^2)^{-1/2} \\ &\Rightarrow 0.5\left(V_{c,\bar{c}}(1,\lambda^0)^2 - 1\right)(\int_0^1 V_{c,\bar{c}}(r,\lambda^0)^2 dr)^{-1/2}. \end{aligned}$$

Finally, it can be shown that the $ADF^{GLS}\left(\lambda^0\right)$ test has the same limiting distribution as the $MZ_t^{GLS}\left(\lambda^0\right)$ test.

Proof of Proposition 1: Let λ^0 denotes the true break fraction and λ a generic one. Except when indicated, $\bar{\alpha}$ is computed using λ. Note that

$$\begin{aligned} S(\bar{\alpha},\lambda) &= (z^{\bar{\alpha}}(\lambda^0)\psi + u^{\bar{\alpha}})' M_\lambda^{\bar{\alpha}}(z^{\bar{\alpha}}(\lambda^0)\psi + u^{\bar{\alpha}}) \\ &= (d^{\bar{\alpha}}(\lambda)\psi + u^{\bar{\alpha}})' M_\lambda^{\bar{\alpha}}(d^{\bar{\alpha}}(\lambda)\psi + u^{\bar{\alpha}}) \\ &= \psi' d^{\bar{\alpha}\prime}(\lambda) M_\lambda^{\bar{\alpha}} d^{\bar{\alpha}}(\lambda)\psi - 2u^{\bar{\alpha}\prime} M_\lambda^{\bar{\alpha}} d^{\bar{\alpha}}(\lambda)\psi + u^{\bar{\alpha}\prime} M_\lambda^{\bar{\alpha}} u^{\bar{\alpha}} \\ &\equiv Q^{\bar{\alpha}}(\lambda) - 2G^{\bar{\alpha}}(\lambda) + u^{\bar{\alpha}\prime} M_\lambda^{\bar{\alpha}} u^{\bar{\alpha}} \end{aligned} \tag{A.18}$$

where $d^{\bar{\alpha}}(\lambda) = z^{\bar{\alpha}}(\lambda) - z^{\bar{\alpha}}(\lambda^0)$, $M_\lambda^{\bar{\alpha}} = I - P_\lambda^{\bar{\alpha}}$, and

$$P_\lambda^{\bar{\alpha}} = z^{\bar{\alpha}}(\lambda) D_T \left(D_T z^{\bar{\alpha}\prime}(\lambda) z^{\bar{\alpha}}(\lambda) D_T\right)^{-1} D_T z^{\bar{\alpha}\prime}(\lambda).$$

When the break fraction is correctly specified, $\lambda = \lambda^0$ and $S(\bar{\alpha}_0,\lambda^0) = u^{\bar{\alpha}_0\prime} M_{\lambda^0}^{\bar{\alpha}_0} u^{\bar{\alpha}_0}$. Note that the above expression consists of $u^{\bar{\alpha}_0}$ and $z^{\bar{\alpha}_0}(\lambda^0)$, which are quasi-differenced with $\bar{\alpha}_0$, i.e. the parameter $\bar{\alpha}$ computed using the true break fraction λ^0. The difference between these two sum of squared residuals is given by

$$S(\bar{\alpha},\lambda) - S(\bar{\alpha}_0,\lambda^0) = Q^{\bar{\alpha}}(\lambda) - 2G^{\bar{\alpha}}(\lambda) + u^{\bar{\alpha}\prime} M_\lambda^{\bar{\alpha}} u^{\bar{\alpha}} - u^{\bar{\alpha}_0\prime} M_{\lambda^0}^{\bar{\alpha}_0} u^{\bar{\alpha}_0}.$$

Consider first

$$Q^{\bar{\alpha}}(\lambda) = \psi' d^{\bar{\alpha}\prime}(\lambda) d^{\bar{\alpha}}(\lambda)\psi - \psi' d^{\bar{\alpha}\prime}(\lambda) P_\lambda^{\bar{\alpha}} d^{\bar{\alpha}}(\lambda)\psi.$$

For simplicity, assume that there is only one break. In Models II and IIb,

$$\begin{aligned} \psi' d^{\bar{\alpha}\prime}(\lambda) d^{\bar{\alpha}}(\lambda)\psi &= \mu_1^2 ||DU^{\bar{\alpha}}(T_1) - DU^{\bar{\alpha}}(T_1^0)||^2 + \beta_1^2 ||DT^{*\bar{\alpha}}(T_1) - DT^{*\bar{\alpha}}(T_1^0)||^2 \\ &\quad + 2\mu_1\beta_1 [DU^{\bar{\alpha}}(T_1) - DU^{\bar{\alpha}}(T_1^0)]'[DT^{*\bar{\alpha}}(T_1) - DT^{*\bar{\alpha}}(T_1^0)]. \end{aligned}$$

From (A.8)-(A.10), we have, in Model II, (if $\lambda > \lambda^0$)

$$\begin{aligned} & T^{-1}\psi' d^{\bar{\alpha}\prime}(\lambda) d^{\bar{\alpha}}(\lambda)\psi \\ = {} & \beta_1^2\left[(\lambda-\lambda^0) - \bar{c}(\lambda-\lambda^0)^2 + \bar{c}^2(\lambda-\lambda^0)^3/3 + \bar{c}^2(\lambda-\lambda^0)^2(1-\lambda)\right] + o(1) \end{aligned}$$

and, in Model IIb,

$$T^{-1-2\eta}\psi' d^{\bar{\alpha}\prime}(\lambda) d^{\bar{\alpha}}(\lambda)\psi = 2\kappa_1^2 \cdot 1(|T_1 - T_1^0| \neq 0) + o(1).$$

Also,

$$d^{\bar{\alpha}\prime}(\lambda) P_{\lambda}^{\bar{\alpha}} d^{\bar{\alpha}}(\lambda) = d^{\bar{\alpha}\prime}(\lambda) z^{\bar{\alpha}}(\lambda) D_T \left(D_T z^{\bar{\alpha}\prime}(\lambda) z^{\bar{\alpha}}(\lambda) D_T\right)^{-1} D_T z^{\bar{\alpha}\prime}(\lambda) d^{\bar{\alpha}}(\lambda).$$

We have shown that $\left(D_T z^{\bar{\alpha}\prime}(\lambda) z^{\bar{\alpha}}(\lambda) D_T\right)^{-1} = O(1)$ with the limit a block-diagonal matrix. It also follows from (A.2)-(A.7) that

$$\begin{aligned} & D_T z^{\bar{\alpha}\prime}(\lambda) d^{\bar{\alpha}}(\lambda) \psi \\ = & [\iota^{\bar{\alpha}}, DU^{\bar{\alpha}}(T_1), T^{-1/2}\tau^{\bar{\alpha}}, T^{-1/2} DT^{*\bar{\alpha}}(T_1)]' \\ & \times [DU^{\bar{\alpha}}(T_1) - DU^{\bar{\alpha}}\left(T_1^0\right), DT^{*\bar{\alpha}}(T_1) - DT^{*\bar{\alpha}}\left(T_1^0\right)](\mu_1, \beta_1)' \\ = & \begin{pmatrix} o(1) & O(T^{-1}|T_1 - T_1^0|) \\ 1 + o(1) & 1(T_1 < T_1^0) + O(T^{-1}|T_1 - T_1^0|) \\ o(1) & O(T^{-1/2}|T_1 - T_1^0|) \\ o(1) & O(T^{-1/2}|T_1 - T_1^0|) \end{pmatrix} \begin{pmatrix} \mu_1 \\ \beta_1 \end{pmatrix}. \end{aligned} \tag{A.19}$$

Hence, we have in Model II

$$\begin{aligned} & T^{-1}\psi' d^{\bar{\alpha}\prime}(\lambda) P_{\lambda}^{\bar{\alpha}} d^{\bar{\alpha}}(\lambda) \psi \\ = & \beta_1^2 \begin{pmatrix} a(0,\lambda) - a(0,\lambda^0) \\ a(\lambda,\lambda) - a(\lambda,\lambda^0) \end{pmatrix}' A(\lambda)^{-1} \begin{pmatrix} a(0,\lambda) - a(0,\lambda^0) \\ a(\lambda,\lambda) - a(\lambda,\lambda^0) \end{pmatrix} + o(1) \\ = & \left|T_1 - T_1^0\right|^2 O(T^{-2}). \end{aligned}$$

with $a(\lambda_i, \lambda_j)$ as defined in (A.1) and, in Model IIb, $T^{-1-2\eta}\psi' d^{\bar{\alpha}\prime}(\lambda) P_{\lambda}^{\bar{\alpha}} d^{\bar{\alpha}}(\lambda) \psi = o(1)$. An entirely analogous argument applies to Models 0b and I. Therefore, collecting terms,

$$Q^{\bar{\alpha}}(\lambda) = \begin{cases} O(T^{1+2\eta}) \cdot \widetilde{I}, & \text{in Models 0b and IIb,} \\ |T_1 - T_1^0| \, O(1), & \text{in Models I and II,} \end{cases} \tag{A.20}$$

where $\widetilde{I} = 1(|T_1 - T_1^0| \neq 0)$. Consider now

$$G^{\bar{\alpha}}(\lambda) = u^{\bar{\alpha}\prime} d^{\bar{\alpha}}(\lambda) \psi - u^{\bar{\alpha}\prime} P_{\lambda}^{\bar{\alpha}} d^{\bar{\alpha}}(\lambda) \psi.$$

As before, we first analyze the order of magnitude of $u^{\bar{\alpha}\prime} d^{\bar{\alpha}}(\lambda) \psi$. Note first that

$$\begin{aligned} & u^{\bar{\alpha}\prime}\left(DU^{\bar{\alpha}}(T_1) - DU^{\bar{\alpha}}\left(T_1^0\right)\right) \\ = & -u_{\lambda^0 T+1}^{\bar{\alpha}} + \bar{c}T^{-1} \sum_{t=\lambda^0 T+2}^{\lambda T} u_t^{\bar{\alpha}} + \left(1 + \bar{c}T^{-1}\right) u_{\lambda T+1}^{\bar{\alpha}} \\ = & -\left(v_{\lambda^0 T+1} + T^{-1}(c - \bar{c}) u_{\lambda^0 T}\right) + \bar{c}T^{-1} \sum_{t=\lambda^0 T+2}^{\lambda T} \left(v_t + T^{-1}(c - \bar{c}) u_{t-1}\right) \\ & + \left(1 + \bar{c}T^{-1}\right)\left(v_{\lambda T+1} + T^{-1}(c - \bar{c}) u_{\lambda T}\right) \\ = & -v_{\lambda^0 T+1} + v_{\lambda T+1} + o_p(1) = O_p(1) \cdot \widetilde{I}, \end{aligned}$$

and

$$
\begin{aligned}
& u^{\bar{\alpha}\prime}\left(DT^{*\bar{\alpha}}\left(T_1\right)-DT^{*\bar{\alpha}}\left(T_1^0\right)\right) \\
= & -u^{\bar{\alpha}}_{\lambda^0 T+1}-\sum_{t=\lambda^0 T+2}^{\lambda T+1}\left(-\bar{c}T^{-1}t+1+\lambda^0\bar{c}+\bar{c}T^{-1}\right)u_t^{\bar{\alpha}}+u^{\bar{\alpha}}_{\lambda T+1} \\
& +\sum_{t=\lambda T+2}^{T}\left(\left(-\bar{c}T^{-1}t+1+\lambda\bar{c}+\bar{c}T^{-1}\right)-\left(-\bar{c}T^{-1}t+1+\lambda^0\bar{c}+\bar{c}T^{-1}\right)\right)u_t^{\bar{\alpha}} \\
= & u^{\bar{\alpha}}_{\lambda T+1}-u^{\bar{\alpha}}_{\lambda^0 T+1}-\sum_{t=\lambda^0 T+2}^{\lambda T+1}\left(-\bar{c}T^{-1}t+1+\lambda^0\bar{c}\right)u_t^{\bar{\alpha}}+\left(\lambda-\lambda^0\right)\bar{c}\sum_{t=\lambda T+2}^{T}(v_t+\frac{c-\bar{c}}{T}u_{t-1})+o_p(1) \\
= & |T_1-T_1^0|O_p(T^{-1/2}).
\end{aligned}
$$

Then,

$$
u^{\bar{\alpha}\prime}d^{\bar{\alpha}}\left(\lambda\right)\psi=\begin{cases} O_p(T^{1/2+\eta})\cdot\widetilde{I}, & \text{in Models 0b and IIb,} \\ |T_1-T_1^0|\,O_p(T^{-1/2}), & \text{in Models I and II.}\end{cases}
$$

Now,

$$
\begin{aligned}
u^{\bar{\alpha}\prime}P_\lambda^{\bar{\alpha}}d^{\bar{\alpha}}\left(\lambda\right)\psi &= u^{\bar{\alpha}\prime}z^{\bar{\alpha}}\left(\lambda\right)D_T\left(D_T z^{\bar{\alpha}\prime}\left(\lambda\right)z^{\bar{\alpha}}\left(\lambda\right)D_T\right)^{-1}D_T z^{\bar{\alpha}\prime}\left(\lambda\right)d^{\bar{\alpha}}\left(\lambda\right)\psi \\
&= O_p(1)D_T z^{\bar{\alpha}\prime}\left(\lambda\right)d^{\bar{\alpha}}\left(\lambda\right)\psi \\
&= \begin{cases} o_p(T^{1/2+\eta})\cdot\widetilde{I}, & \text{in Models 0b and IIb,} \\ |T_1-T_1^0|\,O_p(T^{-1/2}), & \text{in Models I and II.}\end{cases}
\end{aligned}
$$

Therefore, we have

$$
G^{\bar{\alpha}}\left(\lambda\right)=\begin{cases} O_p(T^{1/2+\eta})\cdot\widetilde{I}, & \text{in Models 0b and IIb,} \\ |T_1-T_1^0|\,O_p(T^{-1/2}), & \text{in Models I and II.}\end{cases} \tag{A.21}
$$

For the last term is follows from previous results that

$$
u^{\bar{\alpha}\prime}M_\lambda^{\bar{\alpha}}u^{\bar{\alpha}}-u^{\bar{\alpha}_0\prime}M_{\lambda^0}^{\bar{\alpha}_0}u^{\bar{\alpha}_0}\leq O_p(1)\cdot\widetilde{I}.
$$

Therefore, collecting terms,

$$
T^{-1}\left(S(\bar{\alpha},\lambda)-S(\bar{\alpha}_0,\lambda^0)\right)=T^{-1}Q^{\bar{\alpha}}\left(\lambda\right)+o_p\left(1\right).
$$

Furthermore, note that for a given estimate of the break fraction vector, $\hat{\lambda}$, the inequality $T^{-1}S(\bar{\alpha},\hat{\lambda})\leq T^{-1}S(\bar{\alpha}_0,\lambda^0)$ is always satisfied. Now suppose that $\hat{\lambda}\nrightarrow_p\lambda^0$. Then, according to this inequality, we will need that, for large T, $T^{-1}Q^{\bar{\alpha}}(\hat{\lambda})\leq 0$, but we have shown that

$Q^{\bar{\alpha}}(\lambda) > 0$ when $\lambda \neq \lambda^0$. Hence, a contradiction and the only way that the inequality is satisfied is when $\hat{\lambda} \to_p \lambda^0$. Thus, the minimization of $S(\bar{\alpha}, \lambda)$ over all possible values of the break fraction vector λ results in a consistent estimate of the break fractions. Although our derivations have considered only one structural break, the arguments are also valid for the multiple break case.

To establish the convergence rate for Models I and II, we first define the sets $V_\epsilon = \{T_k : |T_k - T_1^0| < \epsilon T\}$ for $\epsilon \in (0, 1)$ and $V_\epsilon(C) = \{T_k : |T_k - T_1^0| < \epsilon T, |T_k - T_1^0| > C\}$ for $C > 0$, so that $V_\epsilon(C) \subset V_\epsilon$. Note that $S(\bar{\alpha}, \hat{\lambda}) \leq S(\bar{\alpha}_0, \lambda^0)$ with probability 1 and $\Pr(\hat{T}_1 \in V_\epsilon) \to 1$, as $T \to \infty$. Given the previous results, there is a constant $C > 0$ such that

$$\Pr\left(\min_{\lambda T \in V_\epsilon(C)} \frac{S(\bar{\alpha}, \lambda) - S(\bar{\alpha}_0, \lambda^0)}{|\lambda - \lambda^0| T} \leq 0\right) < \zeta$$

for some small $\zeta > 0$, because when C is properly chosen,

$$\begin{aligned} & \Pr\left(\min_{\lambda T \in V_\epsilon(C)} \frac{S(\bar{\alpha}, \lambda) - S(\bar{\alpha}_0, \lambda^0)}{|\lambda - \lambda^0| T} \leq 0\right) \\ = \; & \Pr\left(\min_{\lambda T \in V_\epsilon(C)} \frac{Q^{\bar{\alpha}}(\lambda) - 2G^{\bar{\alpha}}(\lambda) + u^{\bar{\alpha}\prime} M_\lambda^{\bar{\alpha}} u^{\bar{\alpha}} - u^{\bar{\alpha}_0\prime} M_{\lambda^0}^{\bar{\alpha}_0} u^{\bar{\alpha}_0}}{|\lambda - \lambda^0| T} \leq 0\right) < \zeta \end{aligned}$$

since

$$\frac{Q^{\bar{\alpha}}(\lambda) - 2G^{\bar{\alpha}}(\lambda) + u^{\bar{\alpha}\prime} M_\lambda^{\bar{\alpha}} u^{\bar{\alpha}} - u^{\bar{\alpha}_0\prime} M_{\lambda^0}^{\bar{\alpha}_0} u^{\bar{\alpha}_0}}{|\lambda - \lambda^0| T} = O(1) - 2O_p(T^{-1/2}) + \frac{O_p(1) \cdot \widetilde{I}}{C}$$

with the $O(1)$ and $O_p(1)$ terms positive. Hence,

$$\frac{S(\bar{\alpha}, \lambda) - S(\bar{\alpha}_0, \lambda^0)}{|\lambda - \lambda^0| T} > 0$$

on $V_\epsilon(C)$ with large probability. This implies that the minimum cannot be achieved on $V_\epsilon(C)$ and, thus, $\Pr(T|\hat{\lambda} - \lambda^0| \geq C) \leq \zeta$, so that $(\hat{\lambda} - \lambda^0) = O_p(T^{-1})$. For Models 0b and IIb, a similar argument can be applied to show that $(\hat{\lambda} - \lambda^0) = o_p(T^{-1})$.

Proof of Proposition 2: Consider first the statistic $P_T^{GLS}(c, \bar{c}, \hat{\lambda})$ in Models I and II.

$$\begin{aligned} & s(\hat{\lambda})^2 P_T^{GLS}(c, \bar{c}, \hat{\lambda}) \\ = \; & S(\bar{\alpha}, \lambda) - \bar{\alpha} S(1, \hat{\lambda}) \\ = \; & Q^{\bar{\alpha}}(\hat{\lambda}) - 2G^{\bar{\alpha}}(\hat{\lambda}) + u^{\bar{\alpha}\prime} M_{\hat{\lambda}}^{\bar{\alpha}} u^{\bar{\alpha}} - (1 + \bar{c}/T)[Q^1(\hat{\lambda}) - 2G^1(\hat{\lambda}) + u^{1\prime} M_{\hat{\lambda}}^1 u^1] \\ = \; & [Q^{\bar{\alpha}}(\hat{\lambda}) - Q^1(\hat{\lambda})] - \bar{c} T^{-1} Q^1(\hat{\lambda}) - 2[G^{\bar{\alpha}}(\hat{\lambda}) - (1 + \bar{c}/T) G^1(\hat{\lambda})] \\ & + [u^{\bar{\alpha}\prime} M_{\hat{\lambda}}^{\bar{\alpha}} u^{\bar{\alpha}} - (1 + \bar{c}/T) u^{1\prime} M_{\hat{\lambda}}^1 u^1] \\ = \; & [Q^{\bar{\alpha}}(\hat{\lambda}) - Q^1(\hat{\lambda})] + [u^{\bar{\alpha}\prime} M_{\hat{\lambda}}^{\bar{\alpha}} u^{\bar{\alpha}} - (1 + \bar{c}/T) u^{1\prime} M_{\hat{\lambda}}^1 u^1] + o_p(1). \end{aligned}$$

The last equality follows from (A.20), (A.21) and Proposition 1. From this expression, we can see that if $[Q^{\bar{\alpha}}(\hat{\lambda}) - Q^1(\hat{\lambda})]$ is asymptotically negligible, the limit distribution of $s(\hat{\lambda})^2 P_T^{GLS}(c, \bar{c}, \hat{\lambda})$ will be the same as that of $s(\lambda^0)^2 P_T^{GLS}(c, \bar{c}, \lambda^0)$. Now,

$$\begin{aligned} Q^{\bar{\alpha}}(\hat{\lambda}) - Q^1(\hat{\lambda}) &= \psi' d^{\bar{\alpha}\prime}(\hat{\lambda}) d^{\bar{\alpha}}(\hat{\lambda})\psi - \psi' d^{\bar{\alpha}\prime}(\hat{\lambda}) P^{\bar{\alpha}}_{\hat{\lambda}} d^{\bar{\alpha}}(\hat{\lambda})\psi \\ &\quad - \psi' d^{1\prime}(\hat{\lambda}) d^1(\hat{\lambda})\psi + \psi' d^{1\prime}(\hat{\lambda}) P^1_{\hat{\lambda}} d^1(\hat{\lambda})\psi \end{aligned}$$

From (A.8)-(A.10),

$$\psi' d^{\bar{\alpha}\prime}(\hat{\lambda}) d^{\bar{\alpha}}(\hat{\lambda})\psi - \psi' d^{1\prime}(\hat{\lambda}) d^1(\hat{\lambda})\psi = o_p(1).$$

and we recall that

$$d^{\bar{\alpha}\prime}(\hat{\lambda}) P^{\bar{\alpha}}_{\hat{\lambda}} d^{\bar{\alpha}}(\hat{\lambda}) = d^{\bar{\alpha}\prime}(\hat{\lambda}) z^{\bar{\alpha}}(\hat{\lambda}) D_T \left(D_T z^{\bar{\alpha}\prime}(\hat{\lambda}) z^{\bar{\alpha}}(\hat{\lambda}) D_T \right)^{-1} D_T z^{\bar{\alpha}\prime}(\hat{\lambda}) d^{\bar{\alpha}}(\hat{\lambda}).$$

Then, from (A.19) and Proposition 1, we have (assuming for simplicity a single break)

$$\begin{aligned} & D_T z^{\bar{\alpha}\prime}(\hat{\lambda}) d^{\bar{\alpha}}(\hat{\lambda})\psi \\ =\ & [\iota^{\bar{\alpha}}, DU^{\bar{\alpha}}(\hat{T}_1), T^{-1/2}\tau^{\bar{\alpha}}, T^{-1/2} DT^{*\bar{\alpha}}(\hat{T}_1)]' \\ & \times [DU^{\bar{\alpha}}(\hat{T}_1) - DU^{\bar{\alpha}}\left(T_1^0\right), DT^{*\bar{\alpha}}(\hat{T}_1) - DT^{*\bar{\alpha}}\left(T_1^0\right)](\mu_1, \beta_1)' \\ =\ & \begin{pmatrix} o_p(1) & o_p(1) \\ 1 + o_p(1) & 1(T_1 < T_1^0) + o_p(1) \\ o_p(1) & o_p(1) \\ o_p(1) & o_p(1) \end{pmatrix} \begin{pmatrix} \mu_1 \\ \beta_1 \end{pmatrix}, \end{aligned}$$

and, similarly,

$$D_T z^{1\prime}(\hat{\lambda}) d^1(\hat{\lambda})\psi = \begin{pmatrix} o_p(1) & o_p(1) \\ 1 + o_p(1) & 1(T_1 < T_1^0) + o_p(1) \\ o_p(1) & o_p(1) \\ o_p(1) & o_p(1) \end{pmatrix} \begin{pmatrix} \mu_1 \\ \beta_1 \end{pmatrix}.$$

Given the asymptotic block diagonality of $D_T z^{\bar{\alpha}\prime}(\hat{\lambda}) z^{\bar{\alpha}}(\hat{\lambda}) D_T$ and $D_T z^{1\prime}(\hat{\lambda}) z^1(\hat{\lambda}) D_T$, we have

$$\psi' d^{\bar{\alpha}\prime}(\hat{\lambda}) P^{\bar{\alpha}}_{\hat{\lambda}} d^{\bar{\alpha}}(\hat{\lambda})\psi - \psi' d^{1\prime}(\hat{\lambda}) P^1_{\hat{\lambda}} d^1(\hat{\lambda})\psi = o_p(1)$$

We focus on $MSB^{GLS}(\lambda)$ defined by (11). Note that $\widetilde{y}_{-1}'\widetilde{y}_{-1}$ is of the same order of magnitude as $\widetilde{y}'\widetilde{y}$, so we concentrate on $\widetilde{y}'\widetilde{y}$. First, note that, in Models I and II,

$$\begin{aligned} \hat{\psi} &= (z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' y^{\bar{\alpha}} \\ &= (z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\lambda^0)\psi + (z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' u^{\bar{\alpha}} \\ &= \psi + (z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' u^{\bar{\alpha}} \\ &\quad + (z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' [z^{\bar{\alpha}}(\lambda^0) - z^{\bar{\alpha}}(\hat{\lambda})]\psi, \end{aligned}$$

and

$$
\begin{aligned}
\widetilde{y} &= y - z(\hat{\lambda})\hat{\psi} \\
&= z(\lambda^0)(\psi - \hat{\psi}) + u + [z(\lambda^0) - z(\hat{\lambda})]\hat{\psi} \\
&= u - z(\lambda^0)(z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' u^{\bar{\alpha}} \\
&\quad - z(\lambda^0)(z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' [z^{\bar{a}}(\lambda^0) - z^{\bar{a}}(\hat{\lambda})]\psi + [z(\lambda^0) - z(\hat{\lambda})]\hat{\psi}.
\end{aligned}
$$

We need to show that the last two terms are asymptotically negligible compared to the first two. Note that the first two terms are such that

$$
||u - z(\lambda^0)(z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' u^{\bar{\alpha}}|| = O_p(T).
$$

For the third term,

$$
\begin{aligned}
& ||z(\lambda^0)(z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}))^{-1} z^{\bar{\alpha}}(\hat{\lambda})' [z^{\bar{a}}(\lambda^0) - z^{\bar{a}}(\hat{\lambda})]\psi|| \\
=\ & (\psi [z^{\bar{a}}(\lambda^0) - z^{\bar{a}}(\hat{\lambda})] z^{\bar{\alpha}}(\hat{\lambda}) D_T S_T D_T z(\lambda^0)' z(\lambda^0) D_T S_T D_T z^{\bar{\alpha}}(\hat{\lambda})' [z^{\bar{a}}(\lambda^0) - z^{\bar{a}}(\hat{\lambda})]\psi)^{1/2} \\
=\ & (\psi [z^{\bar{a}}(\lambda^0) - z^{\bar{a}}(\hat{\lambda})] z^{\bar{\alpha}}(\hat{\lambda}) D_T S_T D_T z(\lambda^0)' z(\lambda^0) D_T S_T D_T z^{\bar{\alpha}}(\hat{\lambda})' [z^{\bar{a}}(\lambda^0) - z^{\bar{a}}(\hat{\lambda})]\psi)^{1/2} \\
=\ & O_p(T^{1/2})
\end{aligned}
$$

where $S_T = (D_T z^{\bar{\alpha}}(\hat{\lambda})' z^{\bar{\alpha}}(\hat{\lambda}) D_T)^{-1}$. The last equality follows from the fact that $S_T = O_p(1)$, $D_T z(\lambda^0)' z(\lambda^0) D_T = O_p(T^2)$ and, from the previous proof, $D_T z^{\bar{\alpha}}(\lambda)' [z^{\bar{a}}(\lambda^0) - z^{\bar{a}}(\hat{\lambda})]\psi = O_p(T^{-1/2})$. For the last term,

$$
\begin{aligned}
||[z(\lambda^0) - z(\hat{\lambda})]\hat{\psi}|| &= (\hat{\psi}' [z(\lambda^0) - z(\hat{\lambda})]' [z(\lambda^0) - z(\hat{\lambda})]\hat{\psi})^{1/2} \\
&= |\hat{\lambda} - \lambda^0| O_p(T^{3/2}) = O_p(T^{1/2})
\end{aligned}
$$

The CMT with the consistency of $s^2(\lambda)$ completes the proof for $MSB^{GLS}(\lambda)$. An entirely analogous argument applies to $\widetilde{y}_T$ and the proofs for $MZ_{\alpha}^{GLS}(\lambda)$ and $MZ_t^{GLS}(\lambda)$ directly follow.

References

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

Bai, J., 1998. A note on spurious break. Econometric Theory 14, 663-669.

Bai, J., Carrion-i-Silvestre, J.L., 2004. Structural changes, common stochastic trends, and unit roots in panel data. DOCT. AQR. Departament d'Econometria, Estadística i Economia Espanyola. Universitat de Barcelona.

Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. Journal of Applied Econometrics 18, 1-22.

Carion-i-Silvestre, J.L., Sansó-i-Rosselló, A.S., Artis, M., 2004. Raíces unitarias y cambios estructurals en las macromagnitudes Españolas. Revista de Economía Aplicada 35, 5-27.

Christiano, L.J., 1992. Searching for a break in GNP. Journal of Business and Economic Statistics 10, 237-250.

Clemente, J., Montañés, A., Reyes, M., 1998. Testing for a unit root in variables with a double change in the mean. Economics Letters 59, 175-182.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427-431.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64, 813-836.

Gadea, M.D., Montañés, A., Reyes, M., 2004. The European union currencies and the US dollar: from post-Bretton-Woods to the Euro. Journal of International Money and Finance 23, 1109-1136.

Hatanaka, M., Yamada, K., 1999. A unit root test in the presence of structural changes in I(1) and I(0) models. In R.F. Engle and H. White (eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger, Oxford University Press.

Harvey, D.I., Leybourne, S.J., Newbold, P., 2001. Innovational outlier unit root tests with an endogeneously determined break in level. Oxford Bulletin of Economics and Statistics 63, 559-575.

Hecq, A., Urbain, J.P., 1993. Misspecification tests, unit roots and level shifts. Economics Letters 43, 129-135.

Kapetanios, G., 2005. Unit-root testing against the alternative hypothesis of up to m structural breaks. Journal of Time Series Analysis 26, 123-133.

Kim, D., Perron P., 2006. Unit root tests allowing for a break in the trend function under both the null and alternative hypotheses. Unpublished Manuscript, Department of Economics, Boston University.

Kim, T., Leybourne, S.J., Newbold, P., 2000. Spurious rejections by Perron tests in the presence of a break. Oxford Bulletin of Economics and Statistics 62, 433-444.

Lee, J., 1996. Testing for a unit root in the time series with trend breaks. Journal of Macroeconomics 18, 503-519.

Lee, J., Strazicich, M.C., 2001. Break point estimation and spurious rejections with endogenous unit root tests. Oxford Bulletin of Economics and Statistics 63, 535-558.

Lee, J., Strazicich, M.C., 2003. Minimum Lagrange multiplier unit root test with two structural breaks. The Review of Economics and Statistics 85, 1082-1089.

Leybourne, S.J., Mills, T.C., Newbold, P., 1998. Spurious rejections by Dickey-Fuller tests in the presence of a break under the null. Journal of Econometrics 87, 191-203.

Leybourne, S.J., Newbold, P., 2000. Behavior of the standard and symmetric Dickey-Fuller-type tests when there is a break under the null hypothesis. Econometrics Journal 3, 1-15.

Lumsdaine, R.L., Papell, D.H., 1997. Multiple trend breaks and the unit root hypothesis. Review of Economics and Statistics 79, 212-218.

Montañés, A., 1997. Level shifts, unit roots and misspecification of the breaking date. Economics Letters 54, 7-13.

Montañés, A., Olloqui, I., 1999. Misspecification of the breaking date in segmented trend variables: effect on the unit root tests. Economics Letters 65, 301-307.

Montañés, A., Reyes, M., 1998. Effect of a shift in the trend function on Dickey-Fuller unit root tests. Econometric Theory 14, 355-363.

Montañés, A., Reyes, M., 1999. The asymptotic behaviour of the Dickey-Fuller tests under the crash hypothesis. Statistics & Probability Letters 42, 81-89.

Montañés, A., Reyes, M., 2000. Structural breaks, unit roots and methods for removing the autocorrelation pattern. Statistics & Probability Letters 48, 401-409.

Ng, S., Perron, P., 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, 1519-1554.

Nunes, L. C., Kuan, C.-M., Newbold, P., 1995. Spurious Break, Econometric Theory 11, 736-749.

Nunes, L.C., Newbold, P., Kuan, C., 1997. Testing for unit root with breaks: evidence on the great crash and the unit root hypothesis reconsidered. Oxford Bulletin of Economics and Statistics 59, 435-448.

Ohara, H.I., 1999. A unit root test with multiple trend breaks: a theory and application to US and Japanese macroeconomic time-series. The Japanese Economic Review 50, 266-290.

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis. Econometrica 57, 1361-1401.

Perron, P., 1990. Testing for a unit root in a time series with a changing mean. Journal of Business and Economic Statistics 8, 153-162.

Perron, P., 1997. Further evidence on breaking trend functions in macroeconomic variables. Journal of Econometrics 80, 355-385.

Perron, P., Ng, S., 1998. An autoregressive spectral density estimator at frequency zero for nonstationarity tests," Econometric Theory 14, 560-603.

Perron, P., Qu, Z., 2006. Estimating restricted structural change models. Journal of Econometrics 134, 373-399.

Perron, P., Qu, Z., 2007. A simple modification to improve the finite sample properties of Ng and Perron's unit root tests. Economics Letters 94, 12-19.

Perron, P., Rodríguez, G.H., 2003. GLS detrending, efficient unit root tests and structural change. Journal of Econometrics 115, 1-27.

Perron, P., Vogelsang, T.J., 1992a. Nonstationarity and level shifts with an application to purchasing power parity. Journal of Business and Economic Statistics 10, 301-320.

Perron, P., Vogelsang, T.J., 1992b, Testing for a unit root in a time series with a changing mean: correction and extensions. Journal of Business and Economic Statistics 10, 467-470.

Perron, P., Yabu, T., 2006. Testing for shifts in trend with an integrated or stationary noise component. Unpublished Manuscript, Department of Economics, Boston University.

Perron, P., Zhu, X., 2005. Structural breaks with deterministic and stochastic trends. Journal of Econometrics 129, 65-119.

Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335-346.

Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive moving average models with unknown order. Biometrika 71, 599-607.

Schmidt, P., Phillips, P.C.B., 1992. LM tests for a unit root in the presence of deterministic trends. Oxford Bulletin of Economics and Statistics 54, 257-287.

Stock, J.H., 1999. A class of tests for integration and cointegration. In Cointegration, Causality, and Forecasting: A Festschrift for Clive W.J. Granger. Oxford: Oxford University Press, 135-167.

Vogelsang, T.J., 2001. Testing for a shift in trend when serial correlation is of unknown form. Unpublished Manuscript, Department of Economics, Cornell University.

Vogelsang, T.J., Perron, P., 1998. Additional tests for a unit root allowing for a break in the trend function at an unknown time. International Economic Review 39, 1073-1100.

Yao, Y-C., 1988. Estimating the number of change-points via Schwarz' criterion. Statistics and Probability Letters 6, 181-189.

Zivot, E., 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 10, 251-270.

Table 1: Response surface for the non-centrality parameter $\bar{c}$.
(calibrated using up to five breaks)

$$\bar{c}\left(\lambda_k^0\right) = \beta_{0,0} + \sum_{l=1}^{4}\sum_{i=1}^{m}\beta_{i,l}(\lambda_{i,k}^0)^l + \sum_{l=1}^{4}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}\gamma_{i,j,l}\left|\lambda_{i,k}^0 - \lambda_{j,k}^0\right|^l + \varepsilon_k$$

$\hat{\beta}_{0,0}$	-13.18286	$\hat{\gamma}_{1,2,2}$	97.82579
$\hat{\beta}_{1,1}$	0	$\hat{\gamma}_{1,3,2}$	-80.1268
$\hat{\beta}_{2,1}$	-15.15442	$\hat{\gamma}_{1,4,2}$	0
$\hat{\beta}_{3,1}$	13.31425	$\hat{\gamma}_{1,5,2}$	102.8916
$\hat{\beta}_{4,1}$	0	$\hat{\gamma}_{2,3,2}$	105.6941
$\hat{\beta}_{5,1}$	-16.22433	$\hat{\gamma}_{2,4,2}$	-4.307576
$\hat{\beta}_{1,2}$	-33.78199	$\hat{\gamma}_{2,5,2}$	-132.0563
$\hat{\beta}_{2,2}$	22.58485	$\hat{\gamma}_{3,4,2}$	97.34824
$\hat{\beta}_{3,2}$	-36.2509	$\hat{\gamma}_{3,5,2}$	0
$\hat{\beta}_{4,2}$	22.46891	$\hat{\gamma}_{4,5,2}$	41.693
$\hat{\beta}_{5,2}$	0	$\hat{\gamma}_{1,2,3}$	-135.8274
$\hat{\beta}_{1,3}$	91.03556	$\hat{\gamma}_{1,3,3}$	124.6953
$\hat{\beta}_{2,3}$	0	$\hat{\gamma}_{1,4,3}$	0
$\hat{\beta}_{3,3}$	47.48599	$\hat{\gamma}_{1,5,3}$	-164.1035
$\hat{\beta}_{4,3}$	-64.27211	$\hat{\gamma}_{2,3,3}$	-149.0709
$\hat{\beta}_{5,3}$	16.75663	$\hat{\gamma}_{2,4,3}$	0
$\hat{\beta}_{1,4}$	-59.56566	$\hat{\gamma}_{2,5,3}$	176.9245
$\hat{\beta}_{2,4}$	-12.5647	$\hat{\gamma}_{3,4,3}$	-129.1821
$\hat{\beta}_{3,4}$	-26.72178	$\hat{\gamma}_{3,5,3}$	7.431114
$\hat{\beta}_{4,4}$	45.11246	$\hat{\gamma}_{4,5,3}$	-19.47382
$\hat{\beta}_{5,4}$	0	$\hat{\gamma}_{1,2,4}$	70.8171
$\hat{\gamma}_{1,2,1}$	-30.94839	$\hat{\gamma}_{1,3,4}$	-64.20535
$\hat{\gamma}_{1,3,1}$	18.72925	$\hat{\gamma}_{1,4,4}$	0
$\hat{\gamma}_{1,4,1}$	0	$\hat{\gamma}_{1,5,4}$	84.80697
$\hat{\gamma}_{1,5,1}$	-22.3687	$\hat{\gamma}_{2,3,4}$	76.87244
$\hat{\gamma}_{2,3,1}$	-32.9363	$\hat{\gamma}_{2,4,4}$	8.074889
$\hat{\gamma}_{2,4,1}$	0	$\hat{\gamma}_{2,5,4}$	-84.9944
$\hat{\gamma}_{2,5,1}$	40.8027	$\hat{\gamma}_{3,4,4}$	64.40906
$\hat{\gamma}_{3,4,1}$	-32.43146	$\hat{\gamma}_{3,5,4}$	0
$\hat{\gamma}_{3,5,1}$	-5.192179	$\hat{\gamma}_{4,5,4}$	0
$\hat{\gamma}_{4,5,1}$	-24.78274		

All parameters are statistically significant at the 10% level of significance, based on the Newey-West robust estimates of the standard errors. The fit of the regression is $\bar{R}^2 = 0.99$.

Table 2: Response surfaces for the asymptotic critical values of the $P_T^{GLS}(\lambda^0)$ and $MP_T^{GLS}(\lambda^0)$ tests.

	1%	2.5%	5%	10%		1%	2.5%	5%	10%
$\hat{\beta}_{0,0}$	4.055894	4.859365	5.161226	6.150693	$\hat{\delta}_{1,2,2}$	1.23375	1.100153	0.940944	0.892073
$\hat{\beta}_{1,1}$	11.51475	12.63712	8.518757	0	$\hat{\delta}_{1,3,2}$	-1.68662	-1.734786	-1.679684	-2.458478
$\hat{\beta}_{1,2}$	3.115981	3.588129	6.946073	7.278113	$\hat{\delta}_{1,4,2}$	0	0	0	0.110596
$\hat{\beta}_{1,3}$	-1.306252	-5.281776	-3.707751	-7.923907	$\hat{\delta}_{1,5,2}$	0.921445	1.001851	0.471427	1.251522
$\hat{\beta}_{1,4}$	3.407912	2.324983	0	0	$\hat{\delta}_{2,3,2}$	0.451639	0.743685	0.663688	0.879386
$\hat{\beta}_{1,5}$	3.620994	18.03549	11.12886	3.137351	$\hat{\delta}_{2,4,2}$	-0.311286	-0.295055	-0.095371	-0.106705
$\hat{\beta}_{2,1}$	-2.655818	-3.078526	0	0	$\hat{\delta}_{2,5,2}$	-1.53942	-1.975341	-2.421515	-3.001195
$\hat{\beta}_{2,2}$	-3.250386	-3.860014	-5.835792	-6.181602	$\hat{\delta}_{3,4,2}$	0.388264	0.621287	0.19538	0.606314
$\hat{\beta}_{2,3}$	2.614622	2.497937	1.813979	2.516522	$\hat{\delta}_{3,5,2}$	0	-0.255309	-0.123388	-0.265931
$\hat{\beta}_{2,4}$	-2.274934	-2.057081	0	-0.330465	$\hat{\delta}_{4,5,2}$	0.727344	0.739337	0.479379	0.939676
$\hat{\beta}_{2,5}$	-4.417777	-5.213466	-3.474126	-3.990784	$\hat{\delta}_{1,2,3}$	-1.728258	-1.406565	-1.224558	-1.142063
$\hat{\gamma}_{1,0}$	0	0	0	0	$\hat{\delta}_{1,3,3}$	2.586695	2.611613	2.457355	3.701562
$\hat{\gamma}_{1,1}$	0.955323	1.085752	0.983646	0.154939	$\hat{\delta}_{1,4,3}$	-0.048782	0	0	-0.064941
$\hat{\gamma}_{1,2}$	0	0	0	0	$\hat{\delta}_{1,5,3}$	-1.183585	-1.443684	-0.244345	-1.67659
$\hat{\gamma}_{1,3}$	0	-0.381929	-0.29598	-0.642286	$\hat{\delta}_{2,3,3}$	-0.257283	-0.791042	-0.744158	-1.14751
$\hat{\gamma}_{1,4}$	0.077336	0.03169	0	0	$\hat{\delta}_{2,4,3}$	0.278711	0.28051	0.11773	0.132598
$\hat{\gamma}_{1,5}$	0	1.055363	0.628211	0	$\hat{\delta}_{2,5,3}$	2.155698	2.829167	3.131708	3.963276
$\hat{\gamma}_{2,0}$	0	0	0.003024	0.004403	$\hat{\delta}_{3,4,3}$	-0.155567	-0.637582	0	-0.742643
$\hat{\gamma}_{2,1}$	0.022188	0.026568	0.025276	0.007184	$\hat{\delta}_{3,5,3}$	0	0.476813	0.166595	0.509547
$\hat{\gamma}_{2,2}$	0	0	-0.004154	-0.00437	$\hat{\delta}_{4,5,3}$	-0.88688	-0.904828	-0.408659	-1.251418
$\hat{\gamma}_{2,3}$	-0.001744	-0.010641	-0.0086	-0.016969	$\hat{\delta}_{1,2,4}$	0.917618	0.692854	0.648943	0.595435
$\hat{\gamma}_{2,4}$	0	0	0	0	$\hat{\delta}_{1,3,4}$	-1.343122	-1.316134	-1.201767	-1.890047
$\hat{\gamma}_{2,5}$	0.001693	0.021468	0.012495	0.001007	$\hat{\delta}_{1,4,4}$	0.051256	0	0	0
$\hat{\delta}_{1,2,1}$	-0.396601	-0.385792	-0.349046	-0.330487	$\hat{\delta}_{1,5,4}$	0.545985	0.717176	0	0.767657
$\hat{\delta}_{1,3,1}$	0.414869	0.427932	0.42681	0.616974	$\hat{\delta}_{2,3,4}$	0	0.324915	0.296849	0.519356
$\hat{\delta}_{1,4,1}$	0	0	0	-0.054503	$\hat{\delta}_{2,4,4}$	0	0	0	0
$\hat{\delta}_{1,5,1}$	-0.266495	-0.26978	-0.253912	-0.351478	$\hat{\delta}_{2,5,4}$	-1.080734	-1.469398	-1.483281	-1.913977
$\hat{\delta}_{2,3,1}$	-0.247819	-0.310463	-0.254685	-0.283661	$\hat{\delta}_{3,4,4}$	0	0.306692	-0.039153	0.366979
$\hat{\delta}_{2,4,1}$	0.084386	0.066374	0	0	$\hat{\delta}_{3,5,4}$	0.104506	-0.186638	0	-0.203043
$\hat{\delta}_{2,5,1}$	0.44952	0.578628	0.792345	0.961128	$\hat{\delta}_{4,5,4}$	0.42572	0.445571	0.188279	0.618035
$\hat{\delta}_{3,4,1}$	-0.245902	-0.28617	-0.166305	-0.226597					
$\hat{\delta}_{3,5,1}$	-0.053076	0	0	0					
$\hat{\delta}_{4,5,1}$	-0.274101	-0.282583	-0.248772	-0.299152					

All parameters are statistically significant at the 10% level of significance, based on the Newey-West robust estimates of the standard errors. The fit of the regression is $\bar{R}^2 = 0.99$.

Table 3: Response surfaces for the asymptotic critical values of the $ADF(\lambda^0)$ and $MZ_t^{GLS}(\lambda^0)$ tests.

	1%	2.5%	5%	10%
$\hat{\beta}_{0,0}$	-1.51581	-1.126706	-1.123596	-1.038084
$\hat{\beta}_{1,1}$	-2.178373	-4.857866	-5.798566	-5.669399
$\hat{\beta}_{1,2}$	0	0	0	0.300595
$\hat{\beta}_{1,3}$	0	0	1.653186	0
$\hat{\beta}_{1,4}$	0.124969	0	-2.196662	-0.535774
$\hat{\beta}_{1,5}$	-1.311585	-4.891795	-4.511033	-5.597538
$\hat{\beta}_{2,1}$	0.591577	1.241862	1.657028	2.062049
$\hat{\beta}_{2,2}$	0.818683	0.645065	0.588629	0
$\hat{\beta}_{2,3}$	0.58507	0	0	0
$\hat{\beta}_{2,4}$	0	0.170083	0	0.526901
$\hat{\beta}_{2,5}$	1.234374	2.171167	2.529067	2.485652
$\hat{\gamma}_{1,0}$	0.18779	0.216186	0.196932	0.183799
$\hat{\gamma}_{1,1}$	-0.190572	-0.401023	-0.435843	-0.393639
$\hat{\gamma}_{1,2}$	0.136419	0.099371	0.062566	0.022889
$\hat{\gamma}_{1,3}$	0.050768	0.002468	0.163426	0
$\hat{\gamma}_{1,4}$	0	0	-0.200122	0
$\hat{\gamma}_{1,5}$	0	-0.215186	-0.158317	-0.252778
$\hat{\gamma}_{2,0}$	0.003162	0.004896	0.004811	0.004938
$\hat{\gamma}_{2,1}$	-0.004964	-0.00985	-0.010007	-0.009021
$\hat{\gamma}_{2,2}$	0.004586	0.003079	0.001324	0
$\hat{\gamma}_{2,3}$	0.001166	0	0.003972	0
$\hat{\gamma}_{2,4}$	-0.000475	-0.000469	-0.004671	0
$\hat{\gamma}_{2,5}$	0	-0.004392	-0.003355	-0.005326
$\hat{\delta}_{1,2,1}$	0.022779	0.051586	0.061712	0.078724
$\hat{\delta}_{1,3,1}$	0	0	0	0
$\hat{\delta}_{1,4,1}$	0	0.062864	0.020838	0
$\hat{\delta}_{1,5,1}$	0.055794	0	0.0227	0.029488
$\hat{\delta}_{2,3,1}$	0.089138	0.093459	0.106417	0.095085
$\hat{\delta}_{2,4,1}$	0.037545	0.020494	0.013095	0.016575
$\hat{\delta}_{2,5,1}$	-0.097963	-0.06841	-0.066707	0
$\hat{\delta}_{3,4,1}$	0.060277	0.069592	0.071404	0.092617
$\hat{\delta}_{3,5,1}$	0	0.033671	0.066019	0.029572
$\hat{\delta}_{4,5,1}$	0.065901	0.065087	0.081222	0.078205

	1%	2.5%	5%	10%
$\hat{\delta}_{1,2,2}$	0	-0.104406	-0.137608	-0.211755
$\hat{\delta}_{1,3,2}$	0.020127	0.015155	0.016302	0.0395
$\hat{\delta}_{1,4,2}$	0.02695	-0.189128	-0.037523	0
$\hat{\delta}_{1,5,2}$	-0.278515	-0.12443	-0.236102	-0.251287
$\hat{\delta}_{2,3,2}$	-0.372539	-0.317224	-0.329327	-0.251578
$\hat{\delta}_{2,4,2}$	-0.096979	-0.045713	-0.014796	-0.018481
$\hat{\delta}_{2,5,2}$	0.266535	0.194159	0.221897	0.03972
$\hat{\delta}_{3,4,2}$	-0.220092	-0.20317	-0.19438	-0.245119
$\hat{\delta}_{3,5,2}$	0	-0.054685	-0.176222	-0.042657
$\hat{\delta}_{4,5,2}$	-0.243996	-0.200769	-0.214864	-0.211097
$\hat{\delta}_{1,2,3}$	-0.104316	0.061932	0.089751	0.225271
$\hat{\delta}_{1,3,3}$	-0.024748	-0.018097	-0.020127	-0.090121
$\hat{\delta}_{1,4,3}$	-0.081937	0.228386	0.016277	0
$\hat{\delta}_{1,5,3}$	0.504402	0.307872	0.488632	0.511341
$\hat{\delta}_{2,3,3}$	0.61749	0.472924	0.47128	0.326607
$\hat{\delta}_{2,4,3}$	0.070219	0.028355	0	0
$\hat{\delta}_{2,5,3}$	-0.323152	-0.238925	-0.324188	-0.102894
$\hat{\delta}_{3,4,3}$	0.343959	0.27736	0.246337	0.308088
$\hat{\delta}_{3,5,3}$	-0.019057	0.024773	0.20944	0.016733
$\hat{\delta}_{4,5,3}$	0.375573	0.293829	0.29008	0.292033
$\hat{\delta}_{1,2,4}$	0.098532	0	0	-0.078764
$\hat{\delta}_{1,3,4}$	0	0	0	0.053363
$\hat{\delta}_{1,4,4}$	0.057367	-0.103705	0	0
$\hat{\delta}_{1,5,4}$	-0.285799	-0.189041	-0.28817	-0.301665
$\hat{\delta}_{2,3,4}$	-0.345561	-0.253067	-0.252921	-0.17318
$\hat{\delta}_{2,4,4}$	0	0	0	0
$\hat{\delta}_{2,5,4}$	0.145831	0.109136	0.177514	0.070759
$\hat{\delta}_{3,4,4}$	-0.184917	-0.143545	-0.121028	-0.154975
$\hat{\delta}_{3,5,4}$	0.027204	0	-0.098903	0
$\hat{\delta}_{4,5,4}$	-0.199539	-0.157302	-0.156001	-0.157956

All parameters are statistically significant at the 10% level of significance, based on the Newey-West robust estimates of the standard errors. The fit of the regression is $\bar{R}^2 = 0.99$.

Table 4: Response surfaces for the asymptotic critical values of the $MSB(\lambda^0)$ test.

	1%	2.5%	5%	10%		1%	2.5%	5%	10%
$\hat{\beta}_{0,0}$	0.235156	0.27254	0.298518	0.339452	$\hat{\delta}_{1,2,2}$	0	-0.002252	-0.002201	-0.004365
$\hat{\beta}_{1,1}$	-0.053846	-0.100709	-0.132435	-0.206157	$\hat{\delta}_{1,3,2}$	0.000523	0.000337	0.000468	0.000435
$\hat{\beta}_{1,2}$	-0.034397	-0.045047	-0.028955	-0.023594	$\hat{\delta}_{1,4,2}$	-0.00158	-0.004256	-1.05E-03	0
$\hat{\beta}_{1,3}$	-0.00149	-0.002293	0	0	$\hat{\delta}_{1,5,2}$	-0.004808	-0.002652	-0.005415	-0.003924
$\hat{\beta}_{1,4}$	0	-0.030204	-0.082574	-0.060301	$\hat{\delta}_{2,3,2}$	-0.008026	-0.007591	-0.008165	-0.008996
$\hat{\beta}_{1,5}$	-0.122035	-0.13815	-0.150719	-0.22106	$\hat{\delta}_{2,4,2}$	-0.001759	-0.001075	-0.005693	-2.57E-04
$\hat{\beta}_{2,1}$	0.016143	0.02972	0.042159	0.070304	$\hat{\delta}_{2,5,2}$	0.002139	0.001692	0.006675	0
$\hat{\beta}_{2,2}$	0.01831	0.013661	0.015549	0.011052	$\hat{\delta}_{3,4,2}$	-0.00482	-0.005143	-0.006215	-0.007964
$\hat{\beta}_{2,3}$	0	0	0	0	$\hat{\delta}_{3,5,2}$	0	-0.00043	-0.003768	-0.000345
$\hat{\beta}_{2,4}$	0	0	0	0	$\hat{\delta}_{4,5,2}$	-0.005092	-0.005456	-0.00635	-0.006938
$\hat{\beta}_{2,5}$	0.032573	0.042876	0.054106	0.068993	$\hat{\delta}_{1,2,3}$	-0.002506	0.001396	0	3.14E-03
$\hat{\gamma}_{1,0}$	0.009329	0.011935	0.013541	0.016024	$\hat{\delta}_{1,3,3}$	-0.000615	-0.000395	-0.000543	-0.000569
$\hat{\gamma}_{1,1}$	-0.004717	-0.007873	-0.00957	-0.014127	$\hat{\delta}_{1,4,3}$	0.000869	0.005294	0.000459	-9.61E-05
$\hat{\gamma}_{1,2}$	0	-0.001497	0	0	$\hat{\delta}_{1,5,3}$	0.009846	0.006583	0.011188	9.86E-03
$\hat{\gamma}_{1,3}$	0	0	0	8.96E-05	$\hat{\delta}_{2,3,3}$	0.013073	0.011614	0.011948	0.012671
$\hat{\gamma}_{1,4}$	-0.000477	-0.00303	-0.007561	-0.005529	$\hat{\delta}_{2,4,3}$	0.001308	0.000675	0.008277	0.00E+00
$\hat{\gamma}_{1,5}$	-0.007053	-0.007677	-0.007735	-0.01235	$\hat{\delta}_{2,5,3}$	-0.001289	-0.001064	-0.009523	-4.66E-05
$\hat{\gamma}_{2,0}$	0.000188	0.000251	0.000296	0.000348	$\hat{\delta}_{3,4,3}$	0.007706	0.007412	0.008731	0.011019
$\hat{\gamma}_{2,1}$	-0.000127	-0.000187	-0.000214	-0.000302	$\hat{\delta}_{3,5,3}$	-0.001114	0	0.004648	0.00E+00
$\hat{\gamma}_{2,2}$	4.10E-05	0	2.66E-05	2.07E-05	$\hat{\delta}_{4,5,3}$	0.008022	0.008192	0.009249	0.009561
$\hat{\gamma}_{2,3}$	0	0	0	0	$\hat{\delta}_{1,2,4}$	0.00237	0	0.0014	0.00E+00
$\hat{\gamma}_{2,4}$	-2.60E-05	-7.79E-05	-0.000176	-0.000131	$\hat{\delta}_{1,3,4}$	0	0	0	0
$\hat{\gamma}_{2,5}$	-0.000138	-0.000156	-0.000158	-0.000251	$\hat{\delta}_{1,4,4}$	0	-0.002474	0	0.00E+00
$\hat{\delta}_{1,2,1}$	0.000549	0.001053	0.001247	0.001755	$\hat{\delta}_{1,5,4}$	-0.005788	-0.004077	-0.006638	-0.006235
$\hat{\delta}_{1,3,1}$	0	0	0	0	$\hat{\delta}_{2,3,4}$	-0.007103	-0.006111	-0.006246	-0.006539
$\hat{\delta}_{1,4,1}$	0.00072	1.38E-03	0.000566	0	$\hat{\delta}_{2,4,4}$	0	0	-0.004467	0
$\hat{\delta}_{1,5,1}$	0.000666	0	0.000548	0	$\hat{\delta}_{2,5,4}$	0	0	0.005018	0
$\hat{\delta}_{2,3,1}$	0.001845	0.001951	0.002329	2.70E-03	$\hat{\delta}_{3,4,4}$	-0.004255	-0.00377	-0.004446	-0.005689
$\hat{\delta}_{2,4,1}$	0.000654	4.61E-04	1.67E-03	0.000201	$\hat{\delta}_{3,5,4}$	0.000914	0	-0.002247	0
$\hat{\delta}_{2,5,1}$	-0.001146	-0.00086	-0.002054	0	$\hat{\delta}_{4,5,4}$	-0.004267	-0.004338	-0.004949	-0.00518
$\hat{\delta}_{3,4,1}$	0.001265	0.001488	0.001908	0.00255					
$\hat{\delta}_{3,5,1}$	0.000309	0.00041	0.001396	0.000357					
$\hat{\delta}_{4,5,1}$	0.001273	0.001564	2.01E-03	0.002465					

All parameters are statistically significant at the 10% level of significance, based on the Newey-West robust estimates of the standard errors. The fit of the regression is $\bar{R}^2 = 0.99$.

Table 5: Response surfaces for the asymptotic critical values of the $MZ_{\alpha}^{GLS}(\lambda^0)$ test.

	1%	2.5%	5%	10%		1%	2.5%	5%	10%
$\hat{\beta}_{0,0}$	0	0	0	0	$\hat{\delta}_{1,2,2}$	0	-1.68057	-2.034615	-3.0747
$\hat{\beta}_{1,1}$	-49.1033	-80.67152	-94.27681	-98.84623	$\hat{\delta}_{1,3,2}$	4.03E-01	0.276249	0.510663	0.579879
$\hat{\beta}_{1,2}$	0	0	12.49052	20.9028	$\hat{\delta}_{1,4,2}$	-1.215439	-3.085302	-6.20E-01	0
$\hat{\beta}_{1,3}$	0	0	23.0744	0	$\hat{\delta}_{1,5,2}$	-3.836227	-2.058806	-4.221065	-4.01E+00
$\hat{\beta}_{1,4}$	2.918982	0	-31.38412	-3.7837	$\hat{\delta}_{2,3,2}$	-6.37027	-4.771726	-5.320661	-4.366065
$\hat{\beta}_{1,5}$	-63.22623	-77.26193	-71.25859	-82.4666	$\hat{\delta}_{2,4,2}$	-1.436885	-0.506785	-0.228244	-0.241011
$\hat{\beta}_{2,1}$	8.784032	20.5504	26.39792	28.95034	$\hat{\delta}_{2,5,2}$	4.865784	3.309329	4.088078	3.160059
$\hat{\beta}_{2,2}$	13.59995	12.24087	9.484542	5.060179	$\hat{\delta}_{3,4,2}$	-3.466125	-3.609057	-3.347742	-3.774296
$\hat{\beta}_{2,3}$	10.66119	0	0	3.414974	$\hat{\delta}_{3,5,2}$	0	-1.22858	-2.505554	-4.84E-01
$\hat{\beta}_{2,4}$	0	2.326462	0	4.519052	$\hat{\delta}_{4,5,2}$	-3.883586	-3.34232	-3.544715	-3.137332
$\hat{\beta}_{2,5}$	27.43303	39.56724	44.62572	40.67509	$\hat{\delta}_{1,2,3}$	-1.870192	9.48E-01	1.249212	3.239319
$\hat{\gamma}_{1,0}$	2.375973	2.195192	1.999918	1.703449	$\hat{\delta}_{1,3,3}$	-0.478586	-0.319396	-1.033132	-1.277379
$\hat{\gamma}_{1,1}$	-4.867309	-6.865216	-7.250681	-7.417863	$\hat{\delta}_{1,4,3}$	0.689558	3.66918	0.270516	0
$\hat{\gamma}_{1,2}$	2.118534	1.659905	2.182285	2.477604	$\hat{\delta}_{1,5,3}$	7.50E+00	5.136119	8.72E+00	8.090656
$\hat{\gamma}_{1,3}$	0.90565	0	2.109357	0.135494	$\hat{\delta}_{2,3,3}$	10.35363	6.897478	7.766915	6.068159
$\hat{\gamma}_{1,4}$	0	0	-2.871627	0	$\hat{\delta}_{2,4,3}$	1.022642	0	0	0
$\hat{\gamma}_{1,5}$	-2.634858	-2.940626	-2.176966	-3.451044	$\hat{\delta}_{2,5,3}$	-6.274483	-3.896799	-6.192235	-5.130324
$\hat{\gamma}_{2,0}$	0.044459	0.053587	0.055731	0.050724	$\hat{\delta}_{3,4,3}$	5.14785	5.012848	4.208394	4.743969
$\hat{\gamma}_{2,1}$	-1.30E-01	-0.173241	-0.169697	-0.169337	$\hat{\delta}_{3,5,3}$	-3.49E-01	0.579267	2.928145	0.25584
$\hat{\gamma}_{2,2}$	0.068723	0.047938	0.048509	0.051411	$\hat{\delta}_{4,5,3}$	5.740102	4.77502	4.656132	4.089741
$\hat{\gamma}_{2,3}$	0.020546	0	0.04765	0	$\hat{\delta}_{1,2,4}$	1.69E+00	0	0	-1.159203
$\hat{\gamma}_{2,4}$	-0.010131	-0.007096	-0.067457	-0.002892	$\hat{\delta}_{1,3,4}$	0	0	0.509457	7.05E-01
$\hat{\gamma}_{2,5}$	-0.048732	-0.058498	-0.046657	-0.073174	$\hat{\delta}_{1,4,4}$	0.00E+00	-1.645225	0	-0.023476
$\hat{\delta}_{1,2,1}$	4.35E-01	0.866429	0.9477	1.120894	$\hat{\delta}_{1,5,4}$	-4.41E+00	-3.160568	-5.134323	-4.801537
$\hat{\delta}_{1,3,1}$	0	0	0	0	$\hat{\delta}_{2,3,4}$	-5.830718	-3.713127	-4.264494	-3.337409
$\hat{\delta}_{1,4,1}$	0.551608	1.045441	0.348485	0	$\hat{\delta}_{2,4,4}$	0.00E+00	2.72E-01	0	0
$\hat{\delta}_{1,5,1}$	0.69457	0	0.413073	0.513954	$\hat{\delta}_{2,5,4}$	3.016158	1.752061	3.435033	2.920029
$\hat{\delta}_{2,3,1}$	1.62511	1.51E+00	1.700349	1.526902	$\hat{\delta}_{3,4,4}$	-2.81E+00	-2.632964	-2.103161	-2.40406
$\hat{\delta}_{2,4,1}$	0.567561	0.298429	0.201404	0.214627	$\hat{\delta}_{3,5,4}$	0.504965	0	-1.352122	0
$\hat{\delta}_{2,5,1}$	-1.685065	-1.231651	-1.165914	-0.806751	$\hat{\delta}_{4,5,4}$	-3.071827	-2.561458	-2.508772	-2.244213
$\hat{\delta}_{3,4,1}$	1.089693	1.209827	1.25E+00	1.428157					
$\hat{\delta}_{3,5,1}$	0	0.742667	0.957106	0.30942					
$\hat{\delta}_{4,5,1}$	1.166946	1.12E+00	1.380954	1.273161					

All parameters are statistically significant at the 10% level of significance, based on the Newey-West robust estimates of the standard errors. The fit of the regression is $\bar{R}^2 = 0.99$.

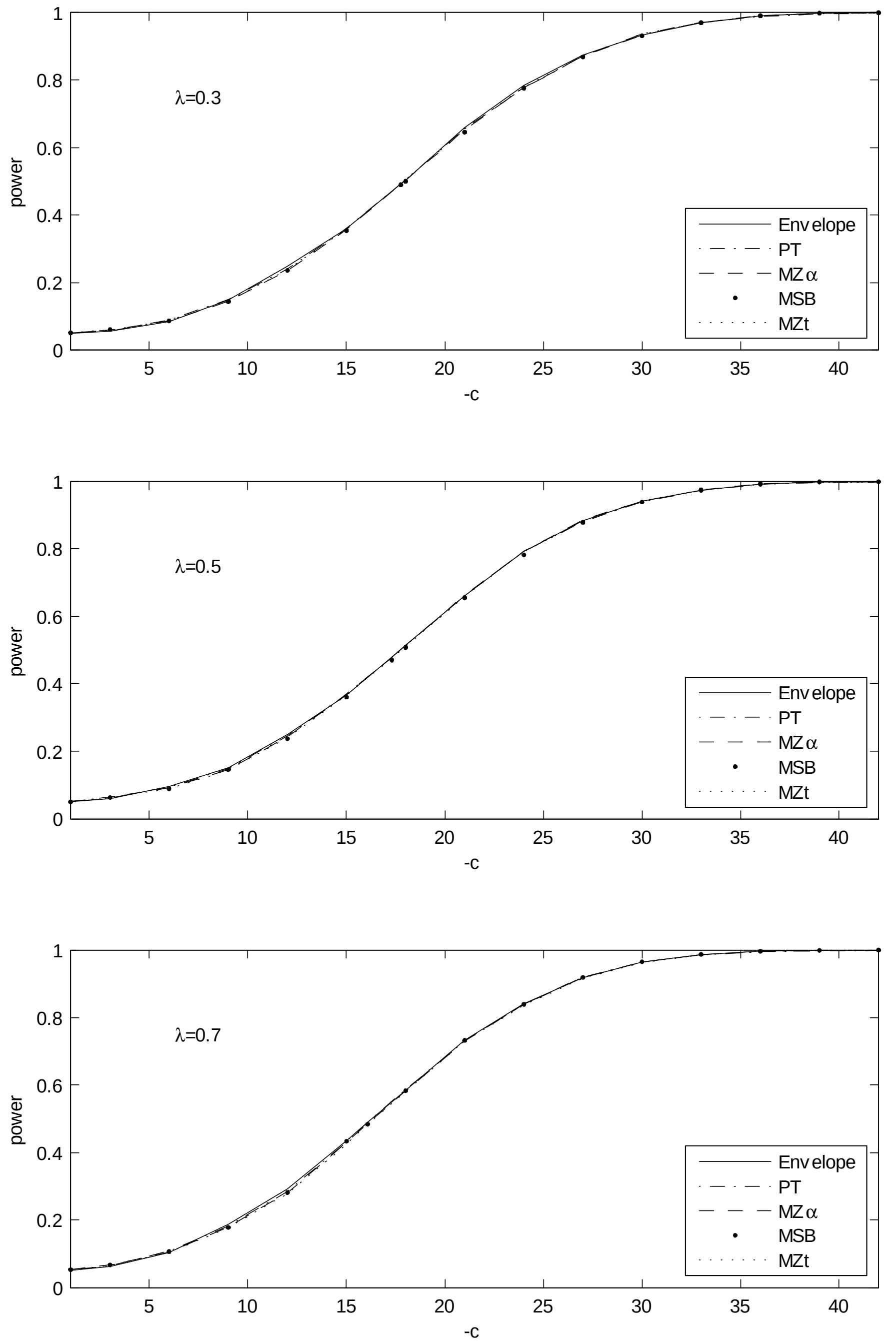

Figure 1: Gaussian Local Power envelope and the Asymptotic Local Power Functions

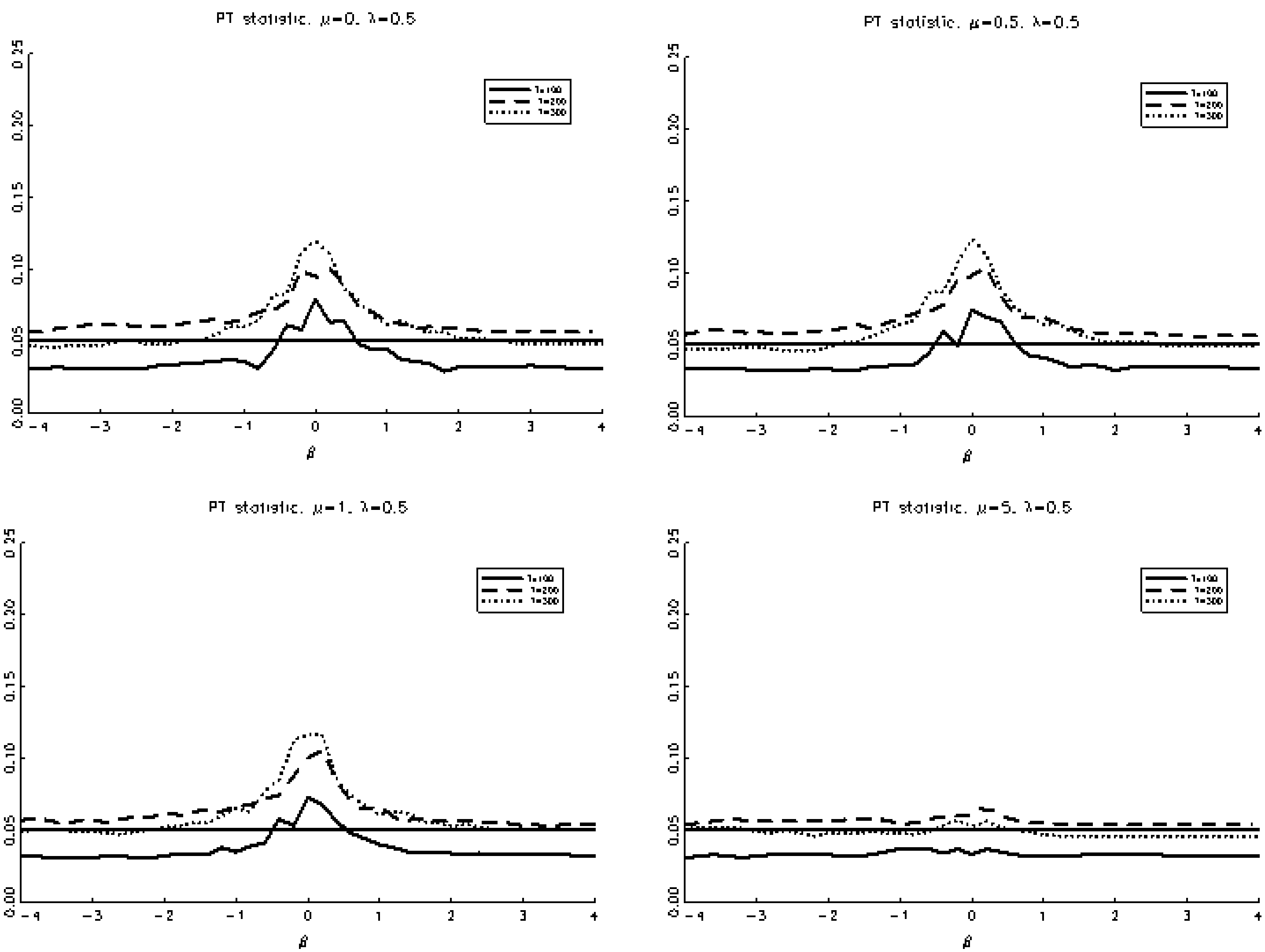

Figure 2: Empirical size of the $P_T^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$

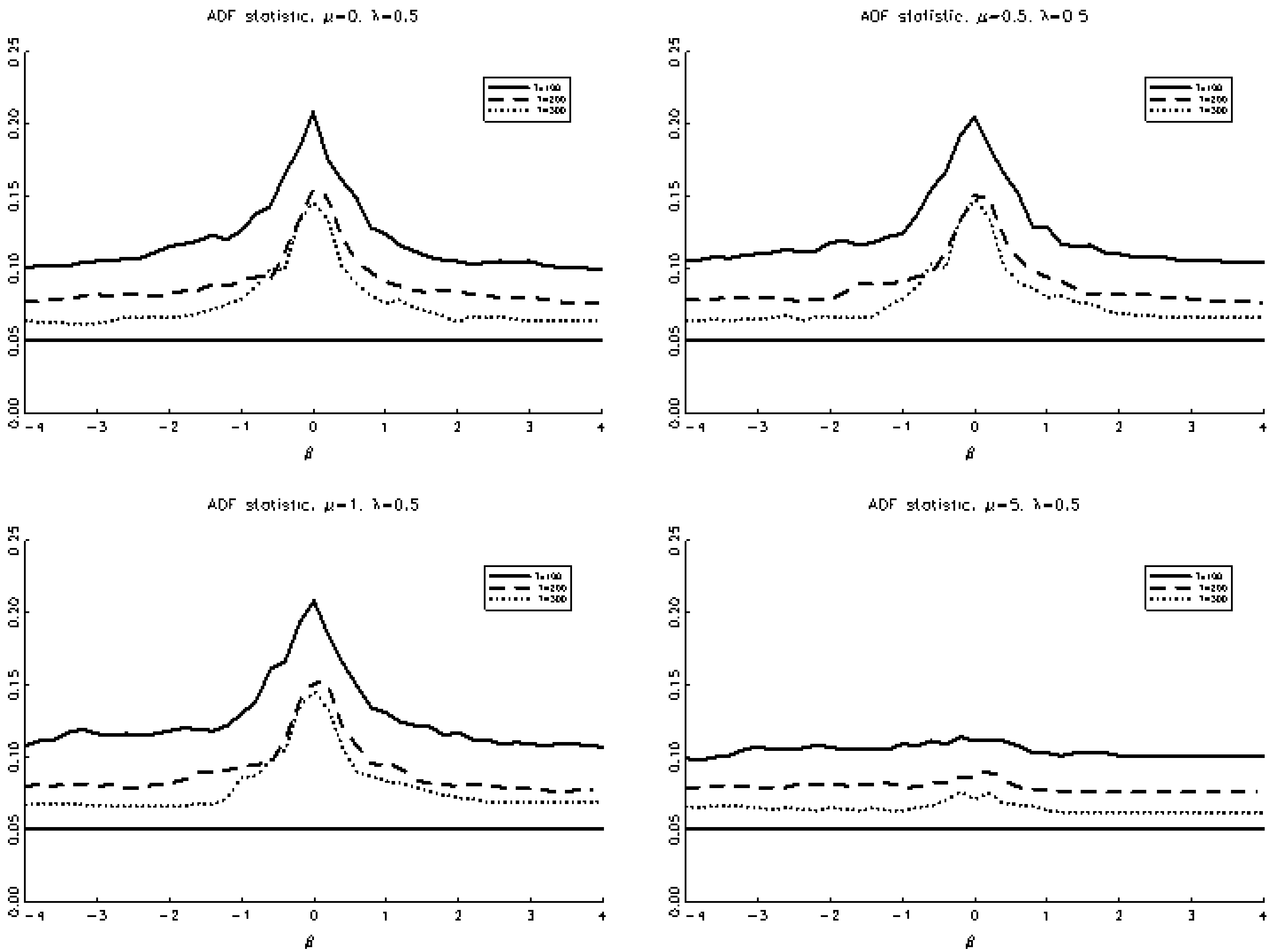

Figure 3: Empirical size of the $ADF^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$

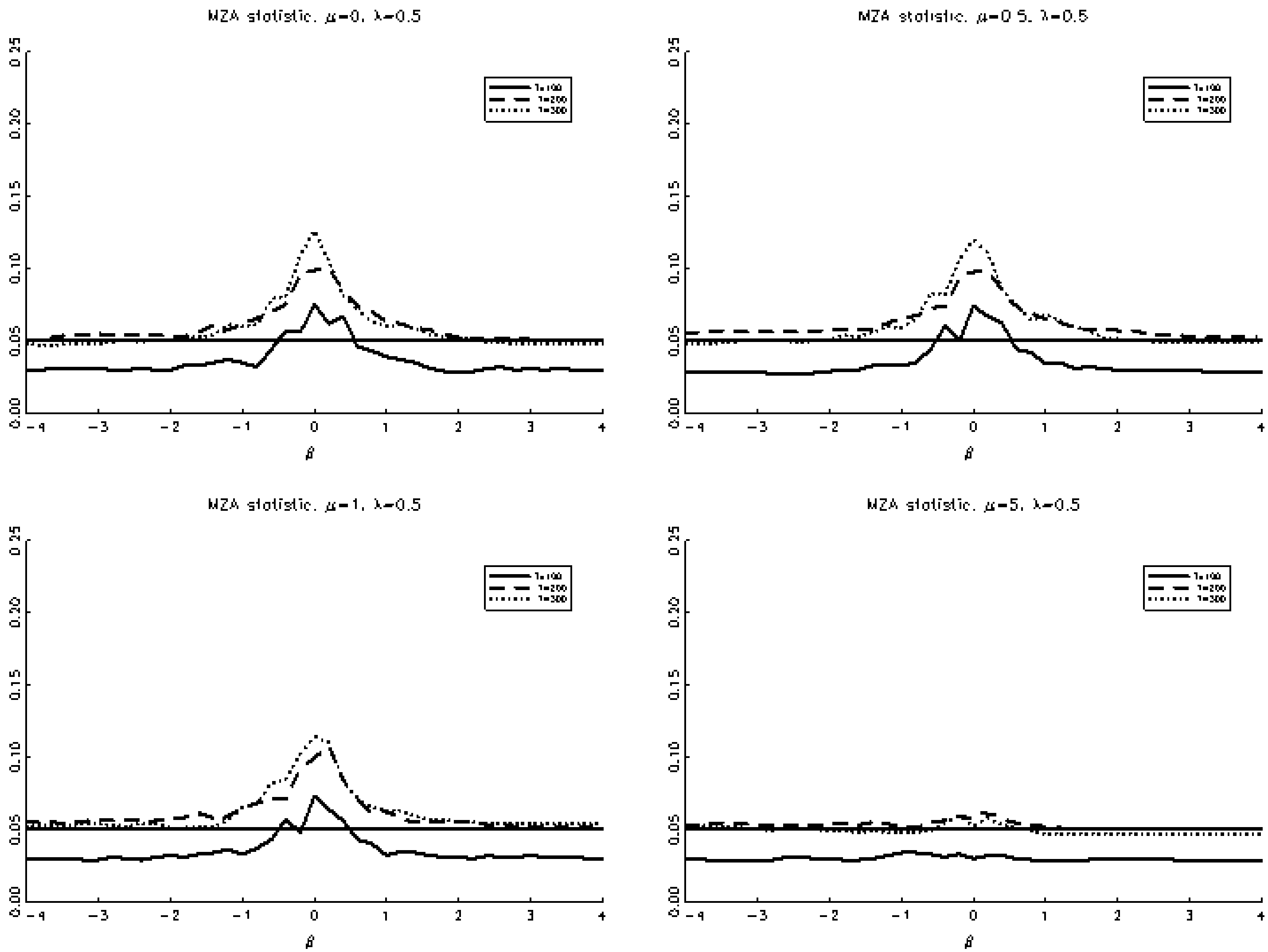

Figure 4: Empirical size of the $MZ_{\alpha}^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$

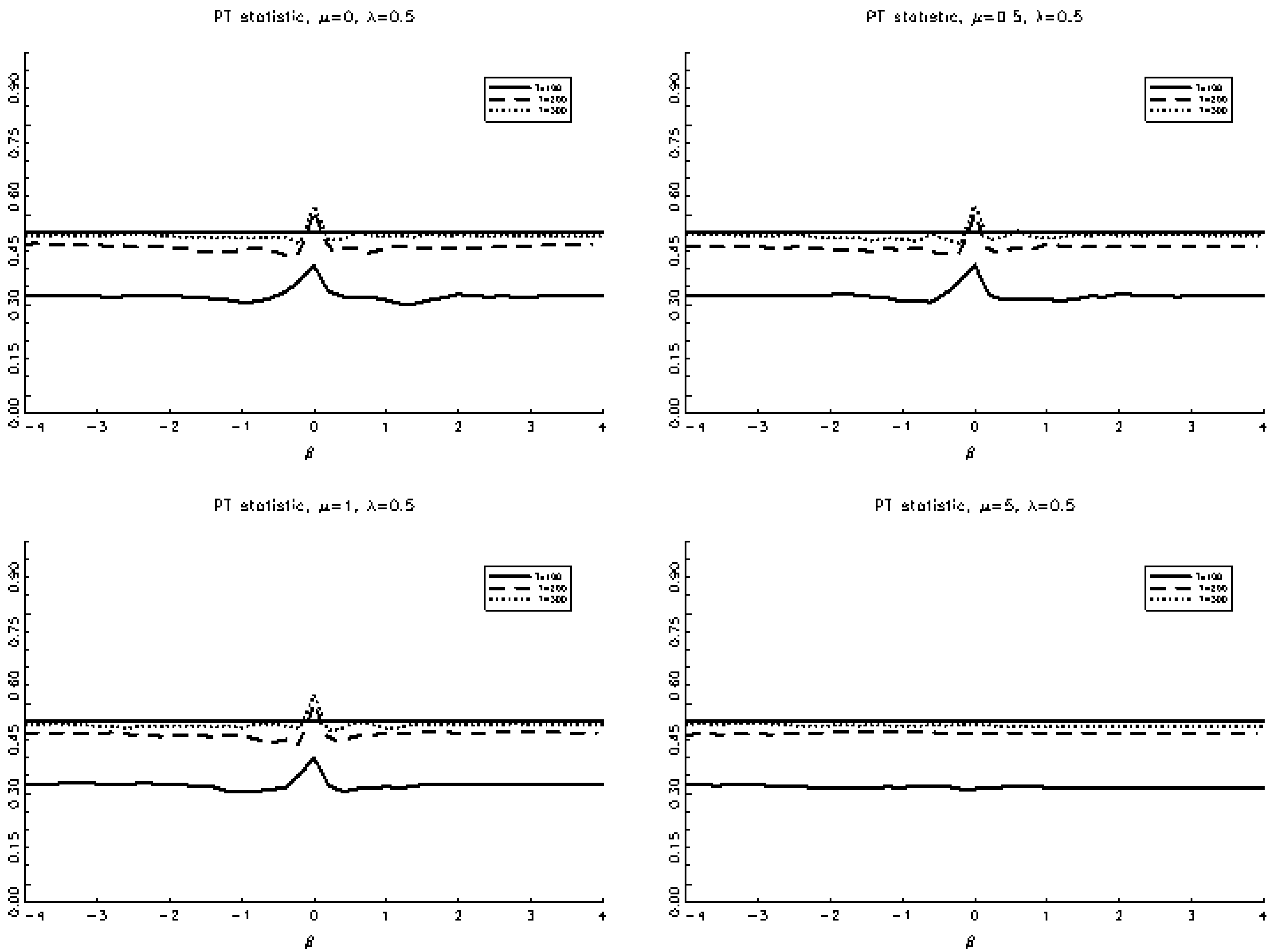

Figure 5: Empirical power of the $P_T^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$

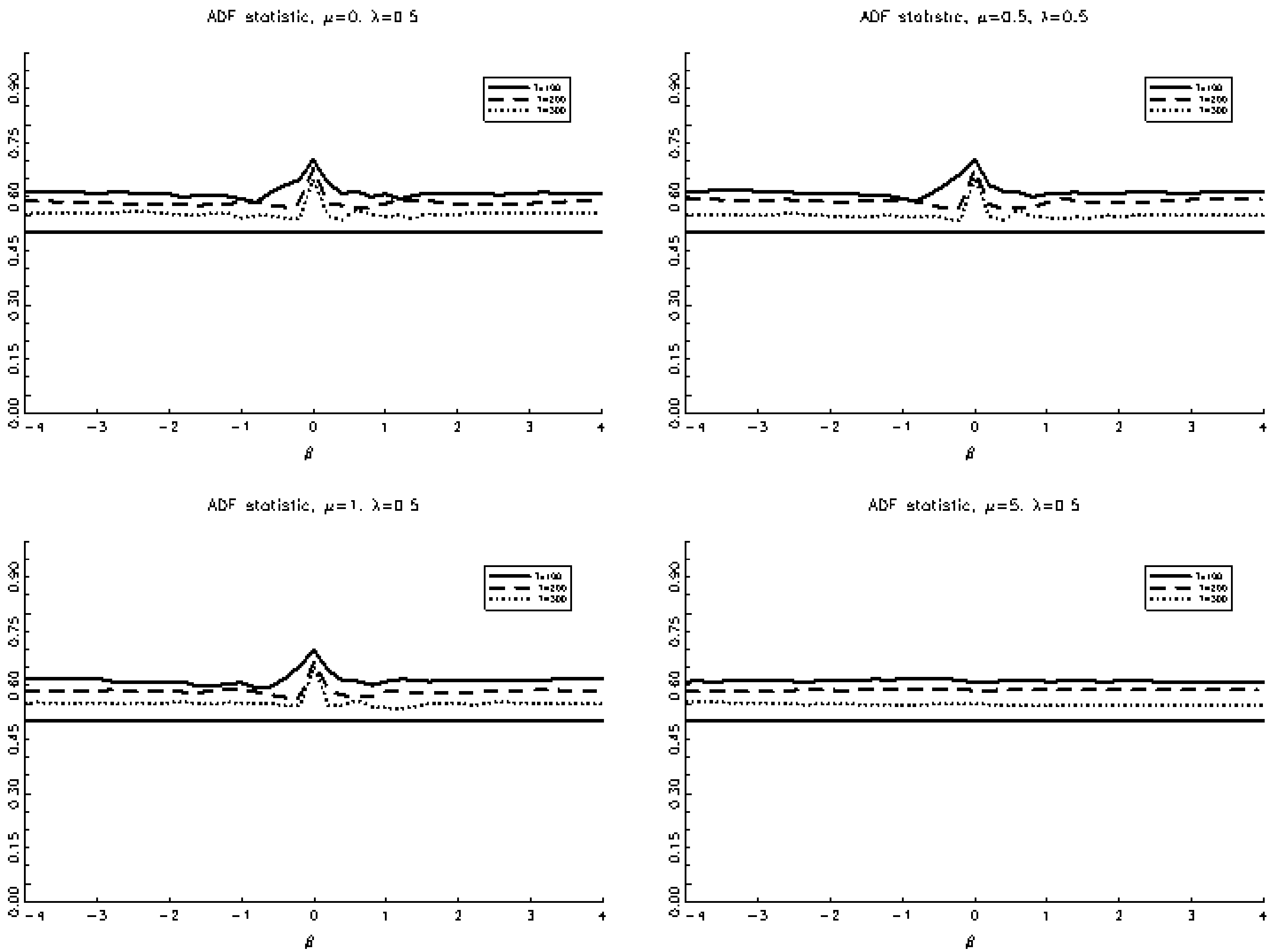

Figure 6: Empirical power of the $ADF^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$

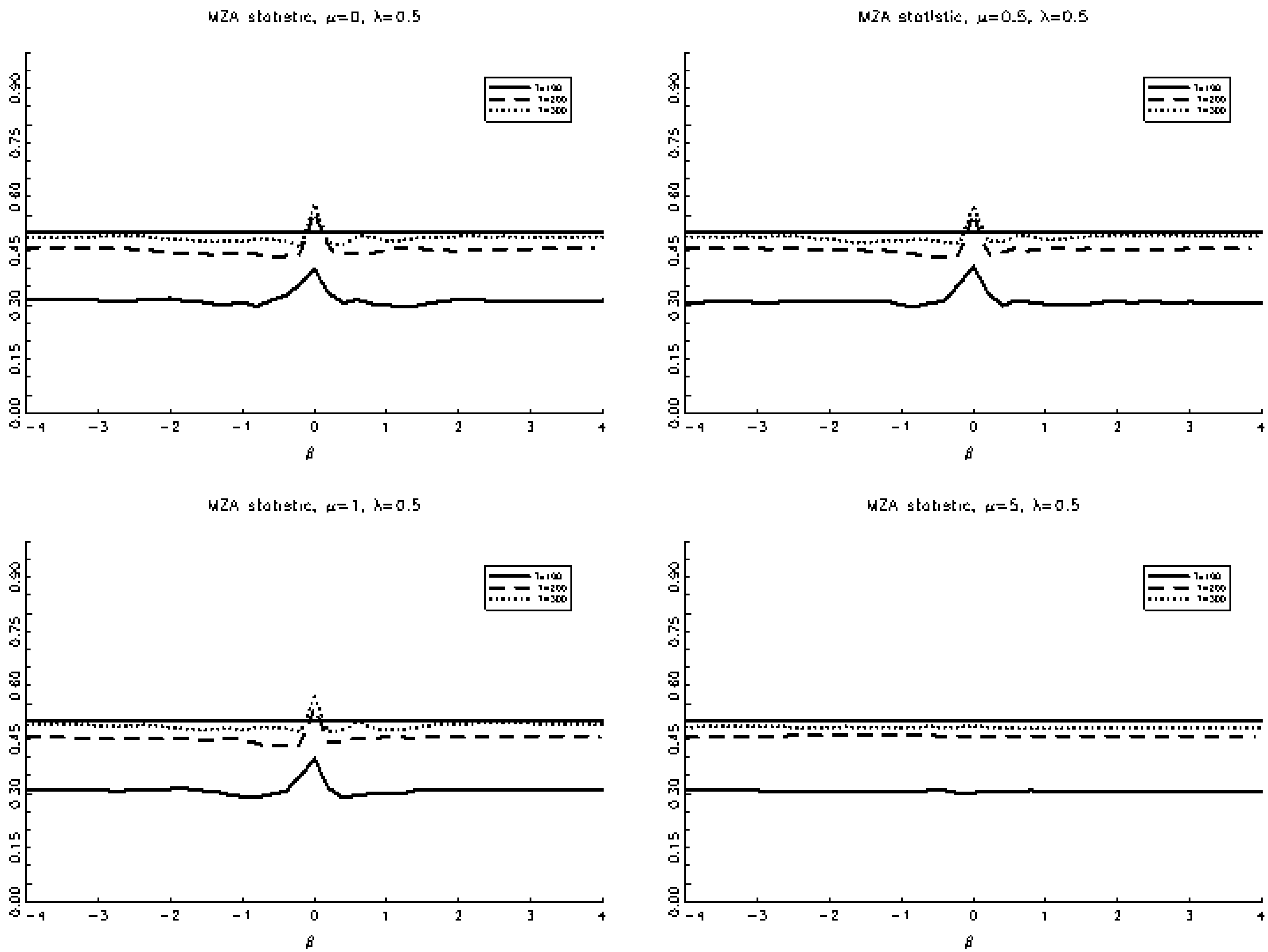

Figure 7: Empirical power of the $MZ_\alpha^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$

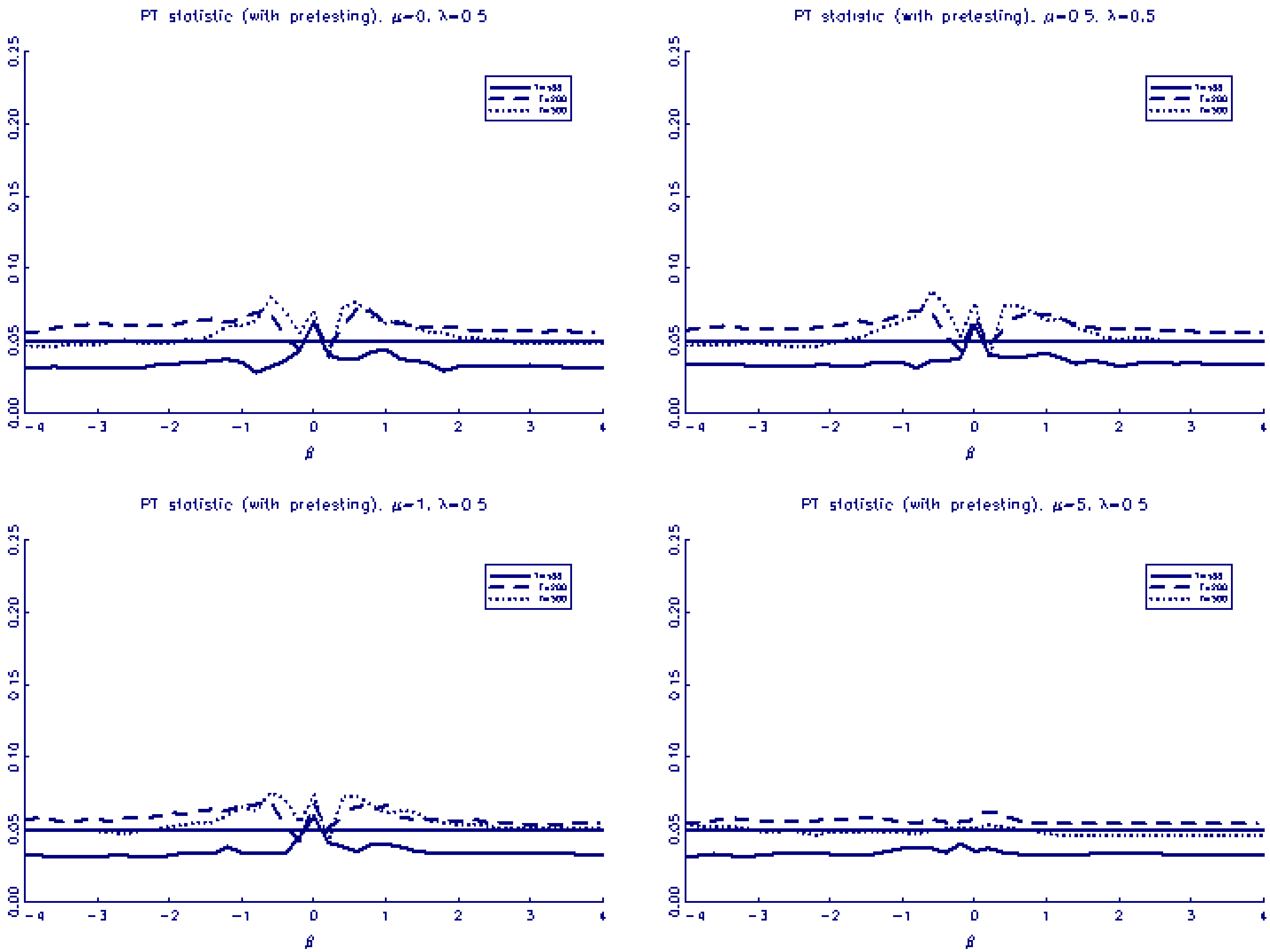

Figure 8: Empirical size of the $P_T^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$, with pretesting

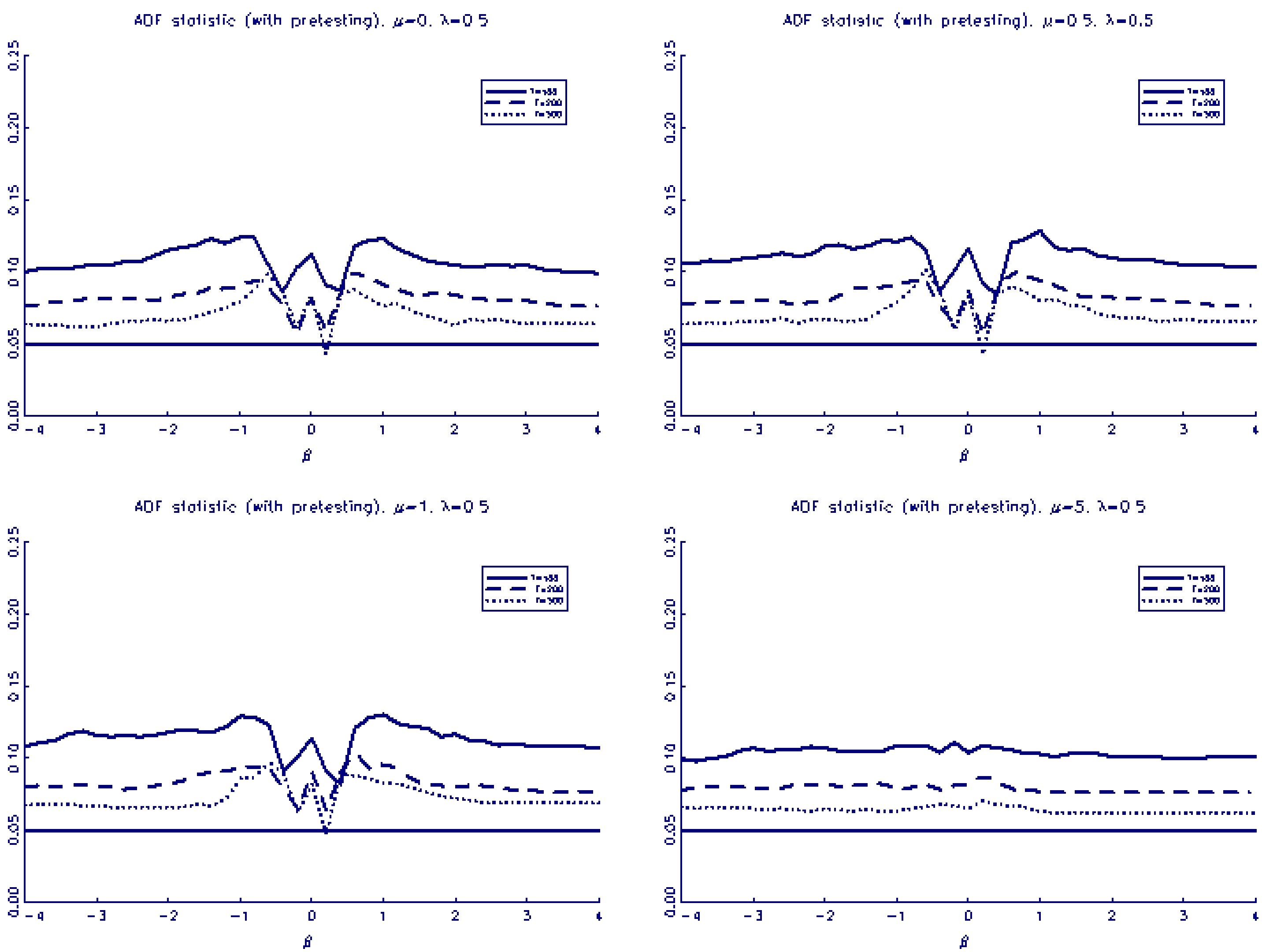

Figure 9: Empirical size of the $ADF^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$, with pretesting

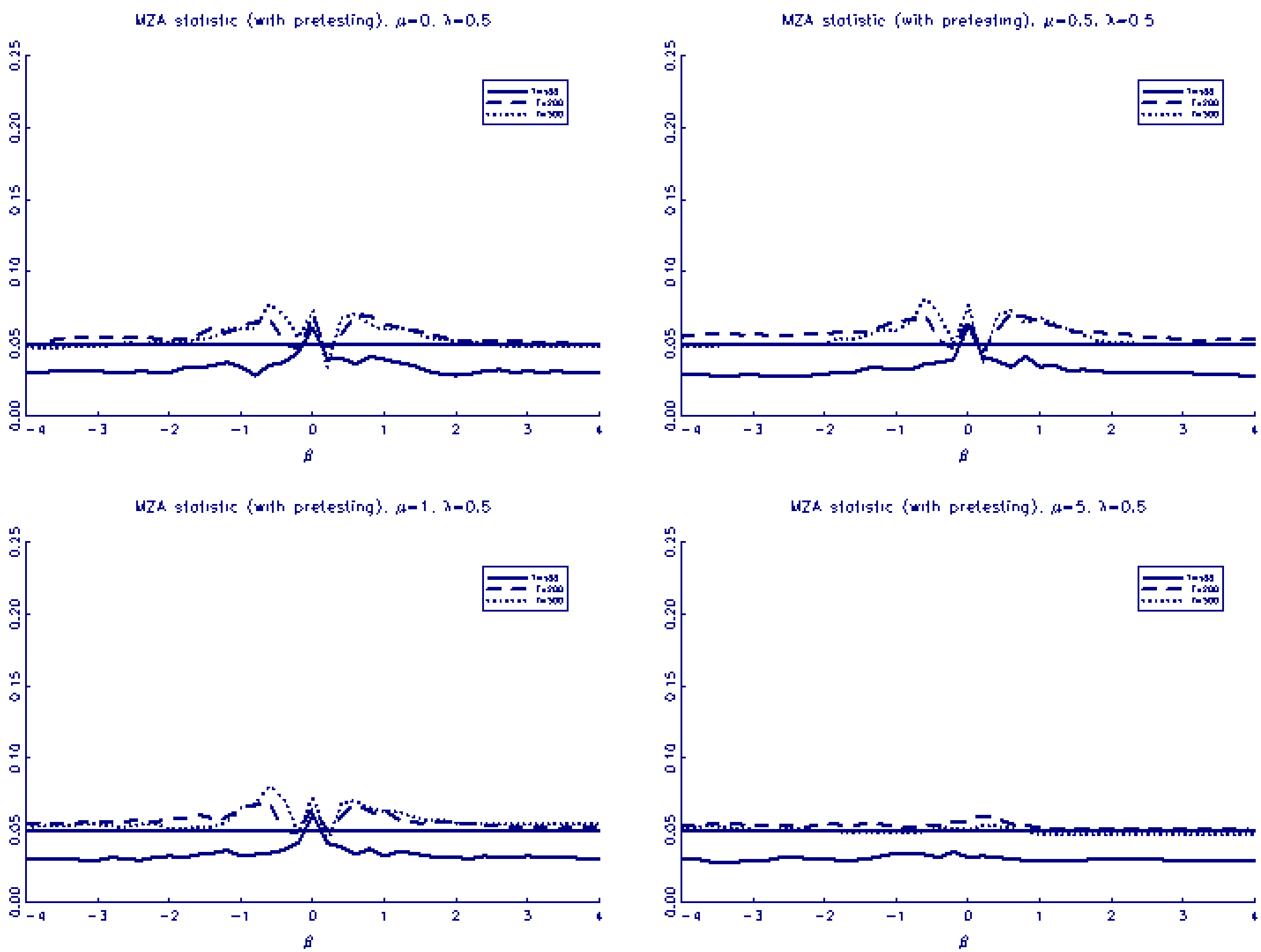

Figure 10: Empirical size of the $MZ_\alpha^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$, with pretesting

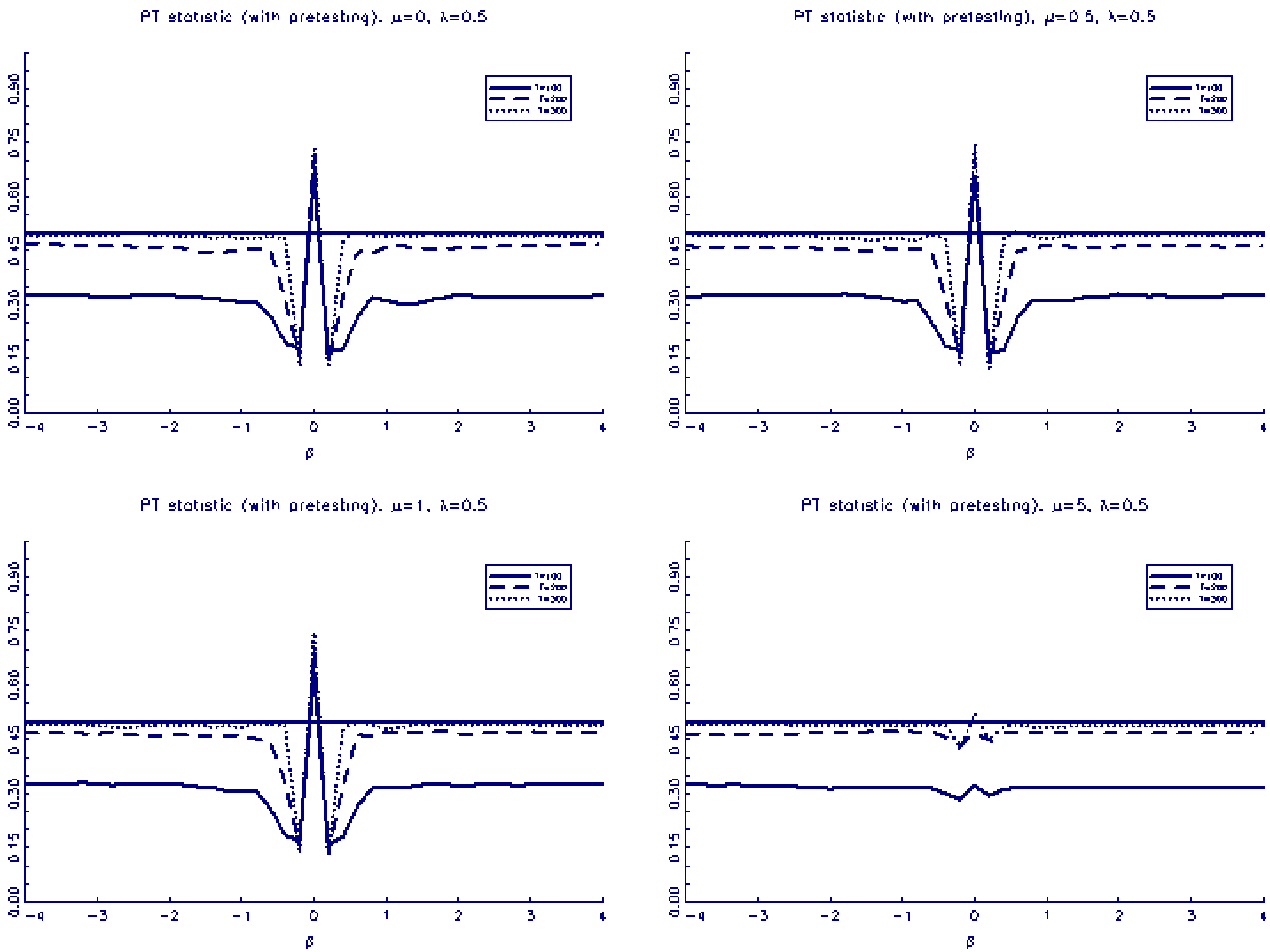

Figure 11: Empirical power of the $P_T^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$, with pretesting

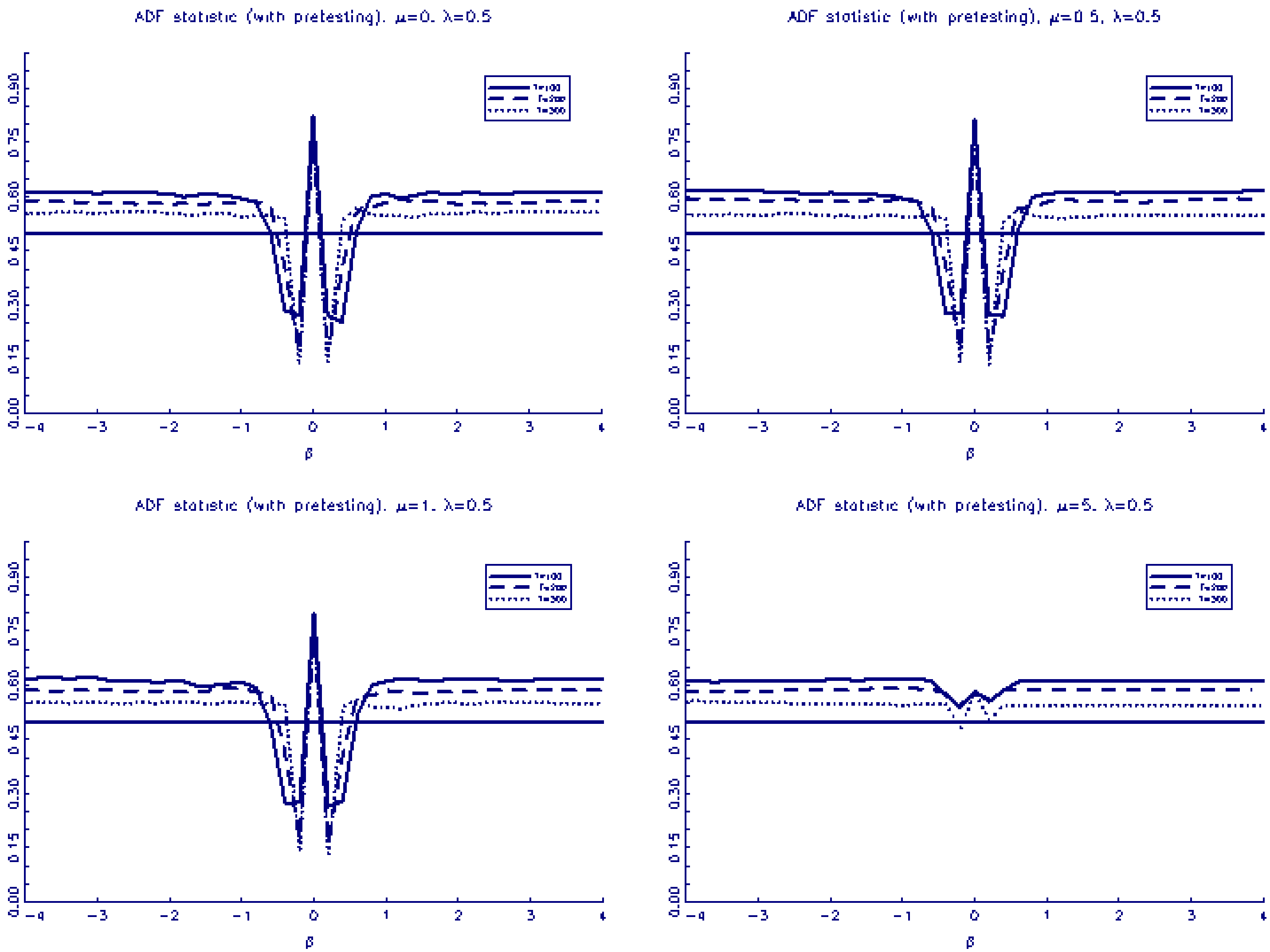

Figure 12: Empirical power of the $ADF^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$, with pretesting

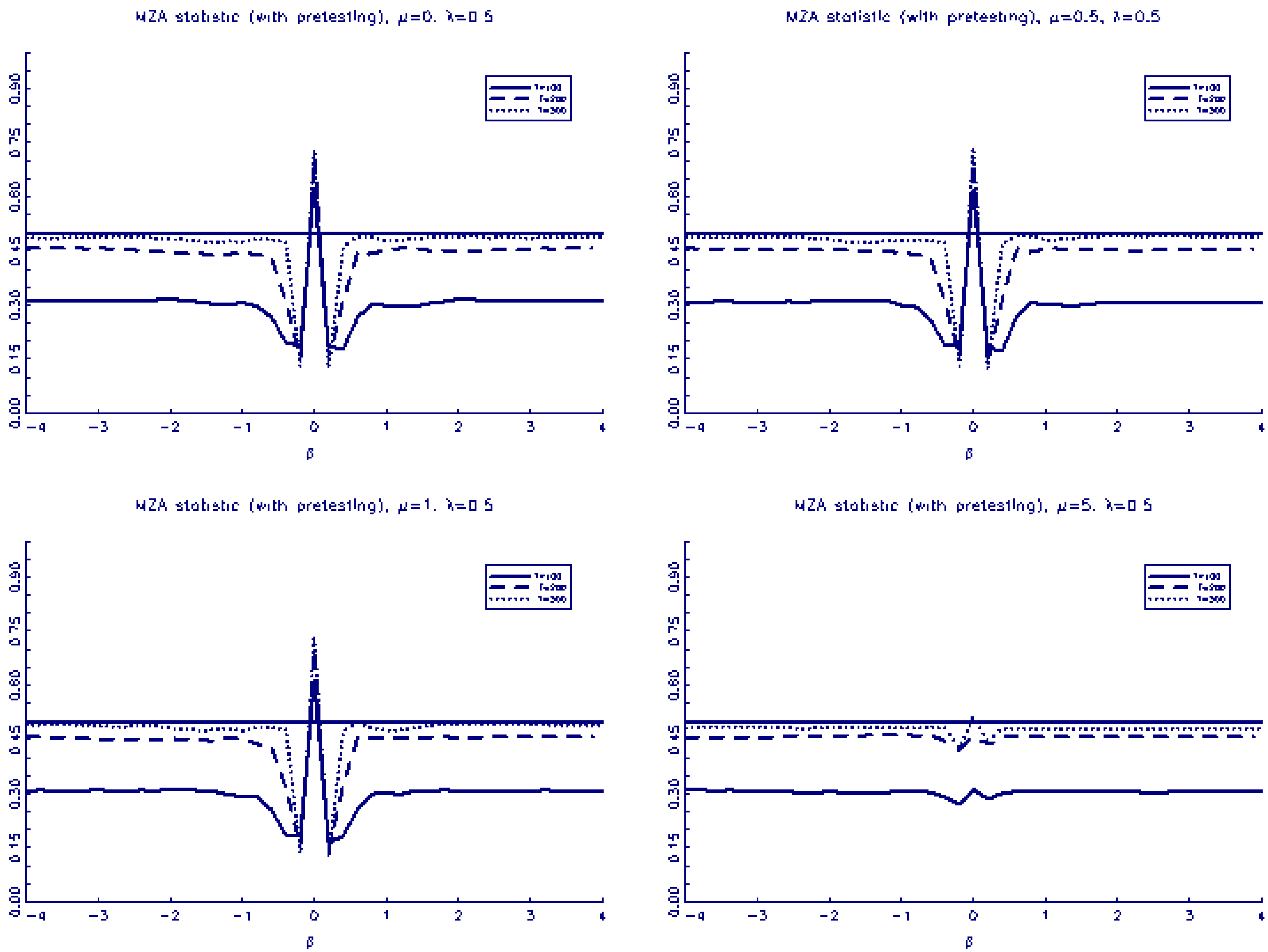

Figure 13: Empirical power of the $MZ_{\alpha}^{GLS}(\hat{\lambda})$ test, $\lambda^0 = 0.5$, with pretesting