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‡

March 16, 2007

Abstract

This paper proposes M-class unit root tests that allow for the presence of multiple structural breaks that might affect the level and/or the slope of the time series both under the null and the alternative hypotheses. We show that the minimization of the sum of the squared residuals of the quasi-differenced model gives consistent estimates of the break fractions with a rate of convergence that ensures that the unit root test statistics have the same limiting distribution as when the break points are assumed to be known. Furthermore, we propose a dynamic algorithm to estimate the multiple break points in an efficient way. The paper provides with response surfaces to approximate the parameter that is used in the quasi-differencing and the percentiles of interest for each test statistic. The performance of the statistics is investigated through Monte Carlo simulations.

Keywords: multiple structural breaks, unit root, GLS detrending

JEL Codes: C12, C22

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

†Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (dpkim@bu.edu)

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

1 Introduction

Non-stationarity in variance analysis that accounts for the presence of structural breaks has devoted a great interest in time series analysis. Now it is well known that misspecification of deterministic function of the auxiliary regression that is used for either testing the null hypothesis of unit root or variance stationarity can lead to conclude in favour of variance non-stationarity – see Perron (1989), Montañés and Reyes (1998), and Lee, Huang, and Shin (1997). This implied the design of test statistics that can accommodate the presence of structural breaks. Earlier proposals were designed to account for one structural break, which could affect either the level and/or the slope of the deterministic time trend.

Since the allowance of one structural break would not be enough in all situations, some proposals in the literature have extended the analysis to the presence of multiple structural breaks. First steps addressed the presence of two structural breaks. Garcia and Perron (1996) propose a pseudo F statistic that accounts for two structural breaks. Lee (1996) and Lumsdaine and Papell (1997) – trending variables – and Carrion-i-Silvestre, Sansó, and Artís (2004) – non-trending variables – generalize the approach in Zivot and Andrews (1992), while Clemente, Montañés, and Reyes (1998) deal with the approach in Perron and Vogelsang (1992) – non-trending variables allowing for structural breaks both under the null and alternative hypotheses. Finally, Lee and Strazicich (2003) extend the statistic in Schmidt and Phillips (1992) to allow for two structural breaks both under the null and the alternative hypotheses. Notwithstanding, recent proposals have generalized the framework to account for multiple structural breaks. Ohara (1999) and Kapetanios (2005) generalize the approach in Zivot and Andrews (1992) through the consideration of multiple structural breaks using the DF statistic to test the unit root hypothesis for trending variables – Ohara (1999) offers critical values for two structural breaks, while Kapetanios (2005) considers up to five structural breaks. Gadea, Montañés, and Reyes (2004) design 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.

Testing the null hypothesis of unit root with multiple structural breaks has been developed in recent years. Nevertheless, none of these papers use GLS detrending procedures to estimate the parameters of the model, though it has been shown that this estimation technique leads to test statistics with better properties – see Elliott, Rothenberg, and Stock (1996) and Ng and Perron (2001). In this paper we extend the approach in Perron and Rodríguez (2003) for the case of multiple structural breaks that affect the slope of the time trend. The paper is only concerned with two test statistics since their limiting distribution can be expressed in a convenient way that allow to compute simplified statistics to cover quite general situations.

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 is devoted to the presentation of the feasible point optimal test with multiple structural breaks, assuming that these break points are known. Furthermore, in this Section we approximate the parameter that is used in the quasi GLS-detrending procedure. Section 4 focuses on the procedure that is applied to estimate the break points when they are unknown. Results in the section shows that the procedure renders T-consistent estimation of the break fraction vector. Section 5 analyses the finite sample performance of the statistics that are proposed in the paper. Finally, Section 6 concludes with some final remarks. All derivations are collected in the companion Appendix.

2 The model

Let y_t be the stochastic process generated according to

$$y_t = d_t + u_t \quad (1)$$

$$u_t = \alpha u_{t-1} + v_t, \; t = 0, \ldots, T, \quad (2)$$

where $\{u_t\}$ is an unobserved stationary 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 stationary disturbance term v_t is defined so that $v_t = \sum_{i=0}^{\infty} \gamma_i \eta_{t-i}$ with $\sum_{i=0}^{\infty} i\left|\gamma_i\right| < \infty$ and where $\{\eta_t\}$ is a martingale difference sequence. We define the long-run and short-run variance as $\sigma^2 = \sigma_\eta^2 \gamma\left(1\right)^2$ and $\sigma_\eta^2 = \lim_{T\to\infty} E\left(\eta_t^2\right)$ respectively. We consider three models, i.e., Model 0 ("level shift" or "crash"'), Model I ("slope change" or "changing growth"), and Model II ("mixed change"). Perron and Rodríguez (2003) consider Models I and II but with only one break. Let $DU_t(T_j^0) = 1$ and $DT_t^*(T_j^0) = \left(t - T_j^0\right)$ for

$t > T_j^0$ and 0 elsewhere, with $T_j^0 = [T\lambda_j^0]$ denoting the j-th date of the break, being $[\cdot]$ the integer part and $\lambda_j^0 \in (0,1)$ the break fraction parameter. As a matter of notation, all true break fractions and break dates are denoted by a superscript, 0. Estimated break fractions and break dates are denoted with a hat. Generic break fractions and break dates are denoted with none of these. We also use the custom that $T_0^0 = 0$ and $T_{m+1}^0 = T$. We collect the m break fraction parameters in the $\lambda^0 = \left(\lambda_1^0, \ldots, \lambda_m^0\right)'$ vector. The deterministic component in (1) $d_t = \psi' z_t(\lambda^0)$ is given by

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

where

$$z_t(\lambda^0) = [z_t'(T_0^0), \ldots, z_t'(T_m^0)]' \text{ and } \psi = (\psi_0', \ldots, \psi_m')'$$

$$z_t(T_0^0) = (1, t),$$

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

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

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.

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 $\eta > 0$. Call the models with this additional assumption as Models 0b and IIb, respectively. 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. Models 0b and IIb have importance in two aspects. The first one is related to the goodness of the finite sample approximation of asymptotic distribution of the unit root tests. In Models 0 and II, the level shifts belong to the class of "slowly evolving trend" defined by Elliott, Rothenberg, and Stock (1996). Hence, ignoring these deterministic components in the unit root procedure has no effect on the asymptotic size and power of the tests but this will clearly worsen the finite sample properties of the associated tests, especially 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" and should not be ignored. This suggests that Models 0b and IIb are better approximations to the process with large level shifts. The second aspect is related to the estimation of unknown break fractions. As shown by Perron and Zhu (2005), the rate of convergence increases when the level shifts are modeled to grow as the sample size grows. This phenomenon has an important implication on the unit root test, which will be shown in the following sections.

GLS detrended unit root test statistics are based on the use of the transformed data $y_t^{\bar{\alpha}}$ and $z_t^{\bar{\alpha}}(\lambda^0)$, where

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

with $\bar{\alpha} = 1 + \bar{c}/T$ and $\bar{c}$ the non-centrality parameter to be defined below. Once the data has been transformed, the deterministic parameters ψ can be estimated through the minimization of the following objective function

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

The minimum of this function is denoted as $S\left(\bar{\alpha}, \lambda^0\right)$.

3 Feasible point optimal test with multiple structural breaks

The definition of the non-centrality parameter $\bar{c}$ is based on the point optimal statistic used in Elliott, Rothenberg, and Stock (1996), which allows us to test the null hypothesis of $\alpha = 1$ in (2) against the alternative hypothesis that $\alpha = \bar{\alpha}$. Using notation in Perron and Rodríguez (2003), the feasible point optimal statistic is given by

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

where $S\left(\bar{\alpha},\lambda^0\right)$ and $S\left(1,\lambda^0\right)$ are the sum of squared residuals (SSR) from a GLS regression with $\alpha=\bar{\alpha}$ and $\alpha=1$, respectively, and $s^2(\lambda^0)$ is the autoregressive estimate of the spectral density at frequency zero of v_t to be defined below. Also, let $W_c(r)$ be the Ornstein-Uhlenbeck process that is the solution to the stochastic differential equation $dW_c(r)=cW_c(r)\,dr+dW(r)$, with $W_c(0)=0$ and $W(r)$ is the standard Brownian motion, where $\Rightarrow$ denotes weak convergence of the associated measure of probability. The limiting distribution of the $P_T^{GLS}\left(c,\bar{c},\lambda^0\right)$ statistic is given in the following Theorem.

Theorem 1 *Let y_t be the stochastic process generated according to (1) and (2) with $\alpha=1+c/T$. Let $P_T^{GLS}\left(c,\bar{c},\lambda^0\right)$ be the statistic defined by (5) with the data obtained from local GLS detrending $(\tilde{y}_t)$ at $\bar{\alpha}=1+\bar{c}/T$. Also, let $s^2(\lambda^0)$ be a consistent estimate of σ^2.*

(i) The limiting distribution of the $P_T^{GLS}\left(c,\bar{c},\lambda^0\right)$ test in Models 0 and 0b is given by

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

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) The limiting distribution of the $P_T^{GLS}\left(c,\bar{c},\lambda^0\right)$ test in Models I, II, and IIb is given by

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

where $M\left(c,0,\lambda^0\right)$ and $M\left(c,\bar{c},\lambda^0\right)$ are as defined in the companion Appendix.

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, Rothenberg, and Stock (1996). Because the timing of breaks are known in the current cases, the test statistic, $P_T^{GLS}\left(c,\bar{c},\lambda^0\right)$ is exactly invariant to the break parameters. Hence, there is no distinction between Models 0 and 0b, and between Models II and IIb. Since Models 0 and I can be viewed as a special case of Model II, no separate proof is provided. The proof for Model II is outlined in the companion Appendix. As can be seen, the limiting distribution of the test statistic for Models I, II, and IIb depends both on the number of structural breaks and on the break fraction vector. From this limiting distribution we can obtain the Gaussian power envelope for different values of $\bar{c}$, and select the $\bar{c}$ parameter so that the asymptotic local power of the test is tangent to the power envelope at 50% power – see Elliott, Rothenberg, and Stock (1996), and Perron and Rodríguez (2003) for further details. For Models 0 and 0b the $\bar{c}$ parameter can be found in Elliott, Rothenberg, and Stock (1996). For Models I, II, and IIb the $\bar{c}$ parameter varies both with the number of structural breaks and with their position. In order to account for this feature, we have approximated the asymptotic $\bar{c}$ parameter for up to $m = 5$ structural break points for all possible combinations of break fraction vectors $\lambda^0 = \left(\lambda_1^0, \cdots, \lambda_m^0\right)'$, $\lambda_i^0 = \{0.1, 0.2, \cdots, 0.9\}$. Instead of providing several tables with the values of the $\bar{c}$ parameter, all the information is summarized through the estimation of one response surface that allows us to approximate the asymptotic $\bar{c}$ parameter for a given vector of break fractions. Visual inspection reveals a u-shapped relationship between the estimated $\bar{c}$ parameters and the different λ^0 vector. Furthermore, the estimated $\bar{c}$ values show some symmetry around $\lambda_i^0 = 0.5$. We have essayed a functional form that accounts for these two features through the specification of powers of λ^0 and $\left|\lambda_i^0 - \lambda_j^0\right|$, $i, j = 1, \cdots, m$, as regressors, i.e. we have estimated

$$\bar{c}\left(\lambda_k^0\right) = \beta_{0,0} + \sum_{l=1}^{4}\sum_{i=1}^{m} \beta_{i,l}(\lambda_{i,k}^0)^l + \sum_{l=1}^{4}\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} \gamma_{i,j,l}\left|\lambda_{i,k}^0 - \lambda_{j,k}^0\right|^l + \varepsilon_k, \tag{6}$$

where we have $k = 1, \cdots, 382$ observations to estimate the parameters of the response surface. The estimated coefficients of (6) are reported in the companion Appendix.

Using the definition of the $\bar{c}$ parameter we can compute other unit root test statistics than the $P_T^{GLS}\left(c, \bar{c}, \lambda^0\right)$ test. Here we follow Perron and Rodríguez (2003) and suggest computing the M-class tests in Ng and Perron (2001) allowing for multiple structural breaks. The tests are based obtained as

$$MZ_\alpha^{GLS}\left(\lambda^0\right) = \left(T^{-1}\tilde{y}_T^2 - s\left(\lambda^0\right)^2\right)\left(2T^{-2}\sum_{t=1}^{T}\tilde{y}_{t-1}^2\right)^{-1} \tag{7}$$

$$MSB^{GLS}\left(\lambda^0\right) = \left(s\left(\lambda^0\right)^{-2}T^{-2}\sum_{t=1}^{T}\tilde{y}_{t-1}^2\right)^{1/2} \tag{8}$$

$$MZ_t^{GLS}\left(\lambda^0\right) = \left(T^{-1}\tilde{y}_T^2 - s\left(\lambda^0\right)^2\right)\left(4s\left(\lambda^0\right)^2T^{-2}\sum_{t=1}^{T}\tilde{y}_{t-1}^2\right)^{-1/2}, \tag{9}$$

with $\tilde{y}_t = y_t - \hat{\psi}' z_t\left(\lambda^0\right)$, where $\hat{\psi}$ minimizes (4). The term $s\left(\lambda^0\right)^2$ denotes the autoregressive estimate of the spectral density at frequency zero of v_t, which is defined as

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

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

$$\Delta \tilde{y}_t = b_0 \tilde{y}_{t-1} + \sum_{j=1}^{k} b_j \Delta \tilde{y}_{t-j} + e_{t,k}, \tag{10}$$

where k is selected using the modified information criteria in Ng and Perron (2001) but computed as in Perron and Qu (2006). Note that we can test the unit root hypothesis using the t-ratio statistic for $b_0 = 0$ in (10), so we get the usual ADF statistic.

As above, we have summarized the 1, 2.5, 5 and 10% percentiles of the previous statistics using response surfaces. In this case, the functional form that has been essayed is given by

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

where as above we have $k = 1, \ldots, 382$ observations to estimate the parameters of the response surface. The estimates of the coefficients of the response surfaces are reported in the companion Appendix.

Theorem 2 *Let y_t be the stochastic process generated according to (1) and (2) with $\alpha = 1 + c/T$. Let $MZ_\alpha^{GLS}(\lambda^0)$, $MSB^{GLS}(\lambda^0)$ and $MZ_t^{GLS}(\lambda^0)$ be the statistics defined by (7)-(9) with the data obtained from local GLS detrending ($\tilde{y}_t$) at $\bar{\alpha} = 1 + \bar{c}/T$. Also, let $s^2(\lambda^0)$ be a consistent estimate of σ^2.*

(i) The limiting distributions of $MZ_\alpha^{GLS}(\lambda^0)$ and $MSB^{GLS}(\lambda^0)$ in Models 0 and 0b are given

by

$$
\begin{aligned}
MZ_{\alpha}^{GLS}(\lambda^0) &\Rightarrow 0.5\left(V_{c,\bar{c}}(1)^2 - 1\right)\left(\int_0^1 V_{c,\bar{c}}(r)^2 dr\right)^{-1} \\
MSB^{GLS}(\lambda^0) &\Rightarrow \left(\int_0^1 V_{c,\bar{c}}(r)^2 dr\right)^{1/2}
\end{aligned}
$$

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) The limiting distributions of $MZ_{\alpha}^{GLS}(\lambda^0)$ *and* $MSB^{GLS}(\lambda^0)$ *in Models I, II and IIb are given by*

$$
\begin{aligned}
MZ_{\alpha}^{GLS}(\lambda^0) &\Rightarrow 0.5\left(V_{c,\bar{c}}(1,\lambda^0)^2 - 1\right)\left(\int_0^1 V_{c,\bar{c}}(r,\lambda^0)^2 dr\right)^{-1} \\
MSB^{GLS}(\lambda^0) &\Rightarrow \left(\int_0^1 V_{c,\bar{c}}(r,\lambda^0)^2 dr\right)^{1/2}
\end{aligned}
$$

where $V_{c,\bar{c}}(r,\lambda^0) = W_c(r) - z_2(r)\, A(\lambda^0)^{-1}\bar{V}\left(\lambda^0\right)$. *The elements* $z_2(r)$, $A\left(\lambda^0\right)$ *and* $\bar{V}\left(\lambda^0\right)$ *are defined in the companion Appendix.*

(iii) The limiting distribution of $MZ_t^{GLS}\left(\lambda^0\right)$ *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 is the same limiting distribution as that for the* $ADF^{GLS}\left(\lambda^0\right)$ *test.*

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). Note, thus, 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 companion Appendix, while the proof for the rest models easily follows from this. We also note that Perron and Rodríguez (2003) showed that each of the M-class tests in Ng and Perron (2001) and the ADF statistic has the same limiting distribution across Models I and II with one break.

4 Unknown break points

Note that previous derivations assume that both the number and position of the structural breaks are known. However, neither are usually known in practice, so that we have to establish a procedure to estimate them. For a given number of structural breaks $m > 0$, we propose to estimate the position of the break points through global minimization of the SSR of the GLS-detrended model:

$$S(\bar{\alpha}, \hat{\lambda}) = \min{}_{\lambda \in \Lambda(\varepsilon)} S\left(\bar{\alpha}, \lambda\right), \tag{11}$$

where the infimum is taken on all possible break fraction vectors defined on the set $\Lambda(\varepsilon)$, with ε being the amount of trimming – some proposals in the literature set $\varepsilon = 0.15$ – and where $\hat{\lambda} = \arg\min{}_{\lambda \in \Lambda(\varepsilon)} S\left(\bar{\alpha}, \lambda\right)$. Note that this approach is different from the one adopted in Perron and Rodríguez (2003), who estimated the location of the break point through the minimization of the SSR under the null and alternative hypotheses. The following Proposition establishes the consistence of the break fraction vector λ^0 as well as its rate of convergence.

Proposition 1 *Let $\{y_t\}_{t=1}^T$ be the stochastic process generated according to (1) and (2) with $\alpha = 1$. Let us assume that $m > 0$ and ψ_j, $j = 1, \ldots, m$, so that there are structural breaks affecting y_t under the null hypothesis. Then, as $T \to \infty$*

(i) in Models I and II,

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

(ii) in Models 0b and IIb,

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

where $\hat{\lambda} = \arg\min{}_{\lambda \in \Lambda(\varepsilon)} S\left(\bar{\alpha}, \lambda\right)$.

The proof is outlined in the companion Appendix. Unfortunately, we also show in the companion Appendix that for the Models I and II the corresponding rate of convergence is not fast enough to warrant that the limit distribution of the statistic is the same as for the known breaks case. However, a simple trimming and reconnection of the data around the estimated break dates can be used to increase the rate of convergence. See Kim and Perron (2006) for more details. The next proposition states that the rate of convergence of the above break fraction estimator is fast enough for the $P_T^{GLS}(c, \bar{c}, \hat{\lambda})$ and M-class tests for Models 0b and IIb, and only for the M-class tests for Models I and II.

Proposition 2 *Let $\{y_t\}_{t=1}^T$ be the stochastic process generated according to (1) and (2) with $\alpha = 1$. Let us assume that $m > 0$ and ψ_j, $j = 1, \cdots, m$. Then, provided that $s(\hat{\lambda})$ is a consistent estimate for σ,*

(i) $P_T^{GLS}(c, \bar{c}, \hat{\lambda})$ has the same limiting distribution as $P_T^{GLS}(c, \bar{c}, \lambda^0)$ in Models 0b and IIb.

(ii) Each of $MZ_\alpha^{GLS}(\hat{\lambda})$, $MSB^{GLS}(\hat{\lambda})$ and $MZ_t^{GLS}(\hat{\lambda})$ has the same limiting distribution as $MZ_\alpha^{GLS}(\lambda^0)$, $MSB^{GLS}(\lambda^0)$ and $MZ_t^{GLS}(\lambda^0)$, respectively, in Models 0b, I, II, and IIb.

The proof is given in the companion Appendix. In practice, the computation of break fractions defined in (11) is somewhat time consuming when a regular grid search is used. Instead, we propose to use a dynamic programming approach. The estimation of the break dates can be done by an iterative procedure similar to that of Perron and Qu (2005). The exact steps are as follows.

1. Compute initial estimated break dates, $\hat{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_m)$ and the associated coefficients, $\hat{\psi} = (\hat{\psi}_0', \hat{\psi}_1', \ldots, \hat{\psi}_m')'$ by OLS using the dynamic algorithm in Bai and Perron (2003) applied to (1).

2. From the given break dates, get an initial value for $\bar{c}(\hat{\lambda})$ using (6).

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 in 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

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

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

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 $\hat{\gamma}_0 = \bar{\alpha}\hat{\mu}_0$. Furthermore,

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

with $\hat{\phi}_1(j) = \left(\hat{\mu}_0 + \hat{\mu}_1 - \hat{\beta}_1 j, \hat{\beta}_0 + \hat{\beta}_1\right)'$ 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 by

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

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

where

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

with $\hat{\phi}_r(j) = \left(\sum_{i=0}^{r} \hat{\mu}_i - \sum_{i=1}^{r-1} \hat{\beta}_i T_i^*(\hat{\psi}, r-1, j) - \beta_r j, \sum_{i=0}^{r} \hat{\beta}_i\right)'$ 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 analogously. Then store the estimated break dates and update $\bar{c}(\hat{\lambda})$ accordingly.

4. Repeat steps 2 and 3 until convergence.

As for the estimation of the number of structural breaks, we can use the BIC information criterion as suggested in Yao (1988).

5 Finite sample performance

5.1 One structural break

The analysis of the performance of the statistic that has been proposed in the paper is considered through the following DGP

$$y_t = d_t + u_t \tag{12}$$

$$d_t = \mu_b DU_t(T_1^0) + \beta_b DT_t^*(T_1^0) \tag{13}$$

$$u_t = \alpha u_{t-1} + v_t, \tag{14}$$

$v_t \sim iid\ N(0,1)$, $u_0 = 0$. We have specified four values for the magnitude of the level shift $\mu_b = \{0, 0.5, 1, 5\}$. For each μ_b value we have analyzed the finite sample performance of the statistics with β_b ranging from -4 to 4 in increments of 0.2. We investigate the sensitivity of the results for 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 1,000 replications are carried out. The empirical size is analyzed with $\alpha = 1$, while the power is essayed with $\alpha = \bar{\alpha} = 1 + \bar{c}/T$, where the $\bar{c}$ parameter depends on the break fraction and is obtained from (6).

Figures 1 to 5 present the results of the empirical size using the asymptotic critical values at the 5% level of significance that are drawn from the estimated response surfaces – note that we use the same scale for the graphs in all cases. In general, the empirical size approaches the nominal one as T increases, which is to be expected since we use the asymptotic critical values. Another common feature is the spike that appears around $\beta_b = 0$. Thus, the statistics show over-rejection distortions for those situations close to the absence of change in the slope of the time series. This is to be expected as well, since the statistics that we have proposed consider the presence of structural breaks affecting the slope of the linear trend under the null hypothesis of unit root. Interestingly, the size distortions diminishes as the magnitude of μ_b increases.

There are, however, some differences in the behaviour of the statistics. Let us first focus on the P_T and MP_T statistics. The theoretical derivations above have revealed that, unless there is a large level shift, the use of the estimated break points that are obtained from the minimization of the SSR does not converge at a fast enough ratio to warrant that the statistics have the same limiting distribution as for the known break case. Simulations reported in the companion Appendix for the P_T statistic show that the size distortions are not so important, although the pattern of these mild distortions depends on the value of λ. Thus, note that under-rejection is observed for $\lambda = 0.3$, whereas mild over-rejection is found for $\lambda = 0.7$. Similar conclusions can be obtained for the MP_T statistic.

The ADF and ZA statistics always show an over-rejection tendency. The MZA and MSB statistics have the right size as T increases for $\lambda = 0.3$ and $\lambda = 0.5$. However, these statistics present mild over-rejection for $\lambda = 0.7$. Finally, the MZT statistic presents under-rejection for $\lambda = 0.3$, although it reaches the right size for the other break fraction values.

The results for the empirical power are presented in the companion Appendix. Since we have set $\alpha = \bar{\alpha}$, the empirical power is expected to be around 0.5. As can be seen, the power of the statistics approaches 0.5 as T gets large, and as λ increases, making the detection of the structural breaks easier. Some of the statistics show power values that are above the expected ones, although this can be thought to be caused by the mild size distortions mentioned previously.

To sum up, the statistics that we have proposed in the paper have good finite sample properties in terms of empirical size and power. The simulations indicate that the statistics have the empirical size close to the nominal one, specially for large T. In those cases where the statistics show size distortions, the amount of distortion is not so important.

5.2 Two structural breaks

Let us consider the DGP given by (12) and (14), with the following deterministic component:

$$d_t = \mu_1 DU_t(T_1^0) + \beta_1 DT_t^*(T_1^0) + \mu_2 DU_t(T_1^0) + \beta_2 DT_t^*(T_2^0),$$

where $\mu_1 = \mu_2 = \mu_b = \{0, 0.5, 1, 5\}$, and β_1 is defined as in the previous section, and β_2 ranges from -5 to 5 in increments of 0.25. Now we consider three vectors of break fractions $\lambda^0 = (0.3, 0.5)$, $(0.3, 0.7)$, and $(0.5, 0.7)$. The other specifications of the Monte Carlo simulation are the same as above.

In the companion Appendix report the empirical size for statistics where the break points have been estimated using the unrestricted algorithm described above for the $\lambda^0 = (0.3, 0.5)$ vector of break fractions. Qualitatively similar conclusions are obtained as for the one break case. The empirical size approaches the nominal one as T, $\beta_1$and β_2 increase. As above, we observe a spike in the pictures when $\beta_1$and β_2 are close to zero and μ_b is small, although the spike disappears for large values of μ_b. Furthermore, except for the ADF statistic, the size distortions that appear when there are no slope shifts are tinny. For values of $\beta_1$and β_2 in the neighborhood of zero, the statistics tend to under-reject, whereas the empirical size approaches the nominal one as $\beta_1$and β_2 get large in absolute value.

6 Conclusion

In this paper we have proposed up to seven test statistics to test the null hypothesis of unit root allowing for the presence of multiple structural breaks that can affect either the level and/or the slope of the deterministic component. We have shown that the minimization of the GLS-detrended-based sum of the squared residuals produces consistent estimates of the break fraction vector. The paper has investigated whether the use of these estimates lead to obtain statistics with the same limiting distribution as when the break points are known. Thus, we have shown that, except for the P_T and MP_T statistics, the rate of convergence of the estimated break fractions suffices, regardless of the magnitude of the level shifts. However, for the P_T and MP_T statistics the rate of convergence of these estimates is not fast enough to warrant that the statistics have the same limiting distribution as for the known break case, unless the time series is affected by large level shifts. The paper has provided response surfaces to approximate the parameter that is used in the GLS estimation, as well as, response surfaces to obtain asymptotic critical values for the percentiles of interest. Furthermore, we have focused on the issue of estimation of multiple break points. Thus, when there is more than one structural break, the simultaneous estimation of all break points is time consuming. The paper has proposed an efficient dynamic algorithm to obtain the estimates when there are more than one structural break. The finite sample performance of the statistics is analyzed using Monte Carlo experiments for the cases of one and two structural breaks. The analysis reveals that the statistics have good finite sample performance.

Figure 1: Empirical size for the ADF test, $\lambda = 0.5$

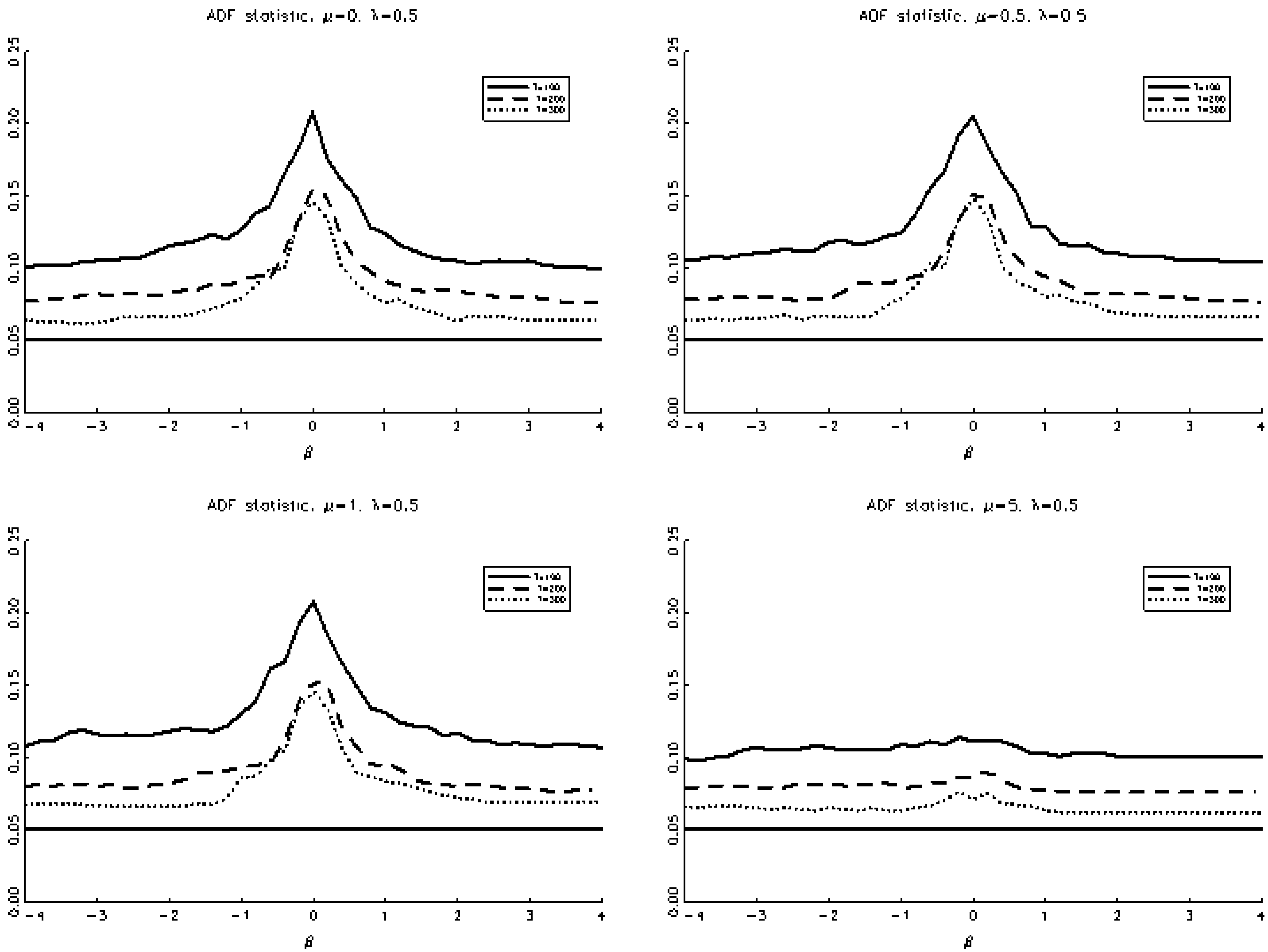


Figure 2: Empirical size for the ZA test, $\lambda = 0.5$

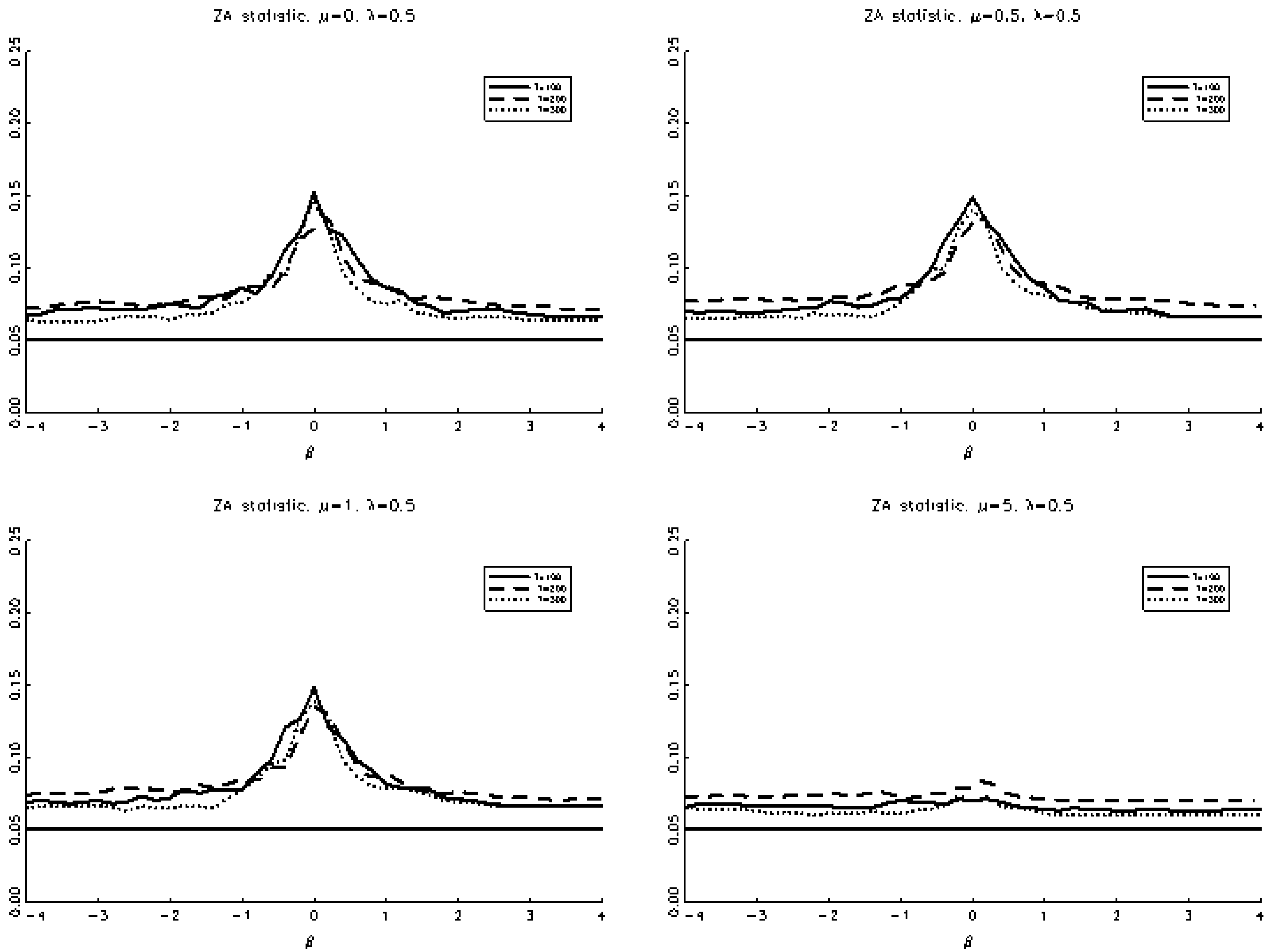


Figure 3: Empirical size for the MZA test, $\lambda = 0.5$

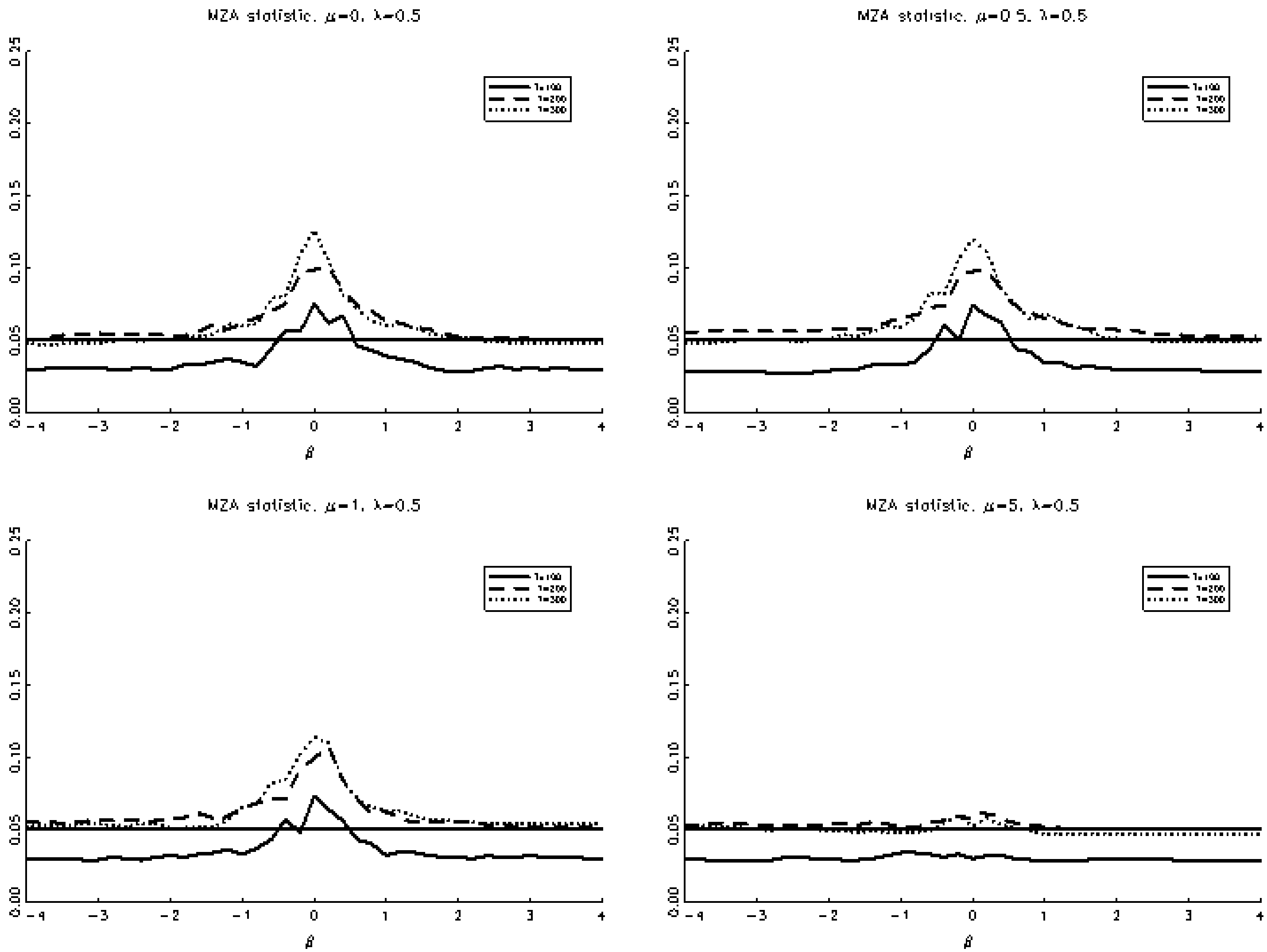


Figure 4: Empirical size for the MSB test, $\lambda = 0.5$

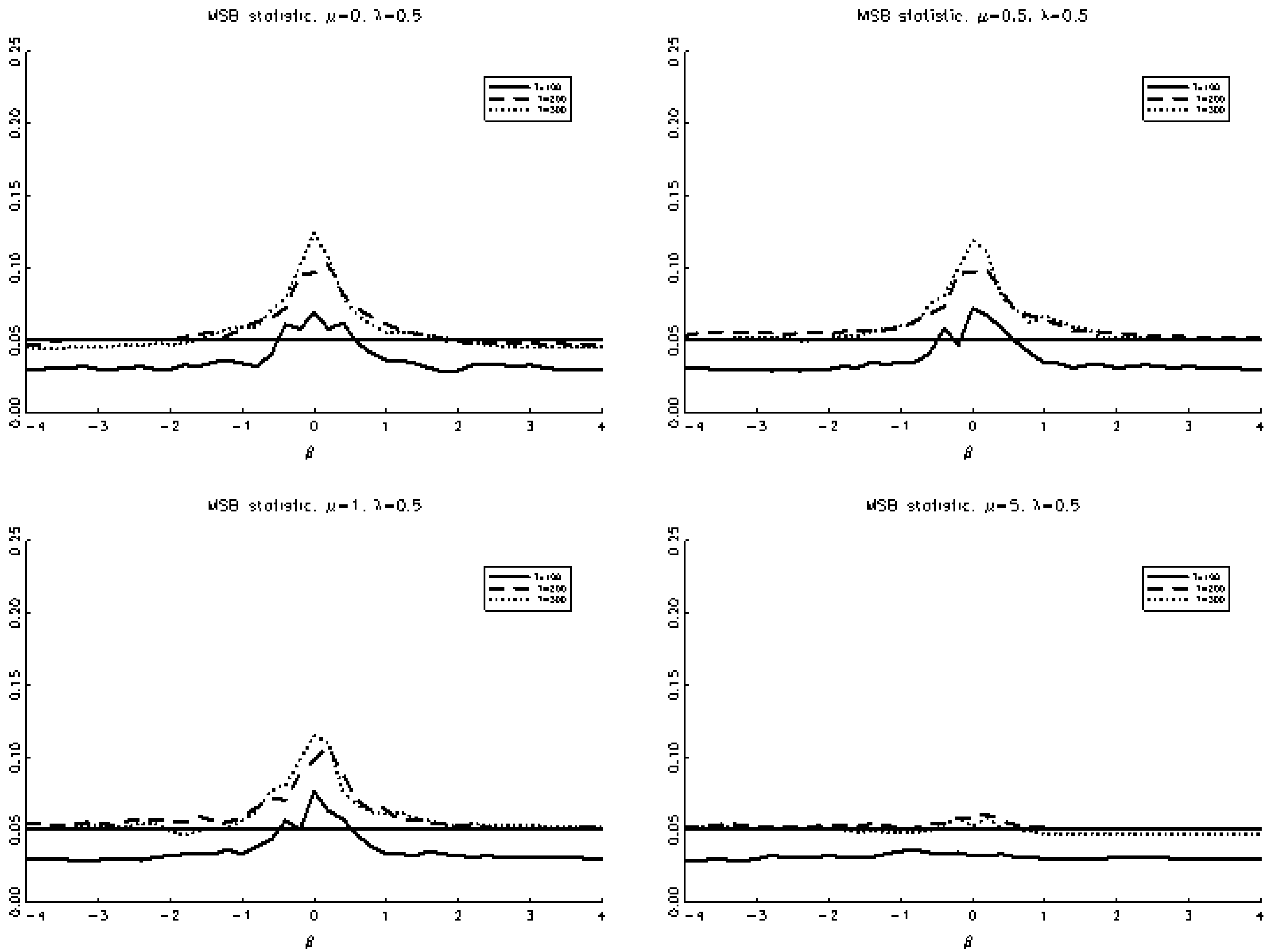
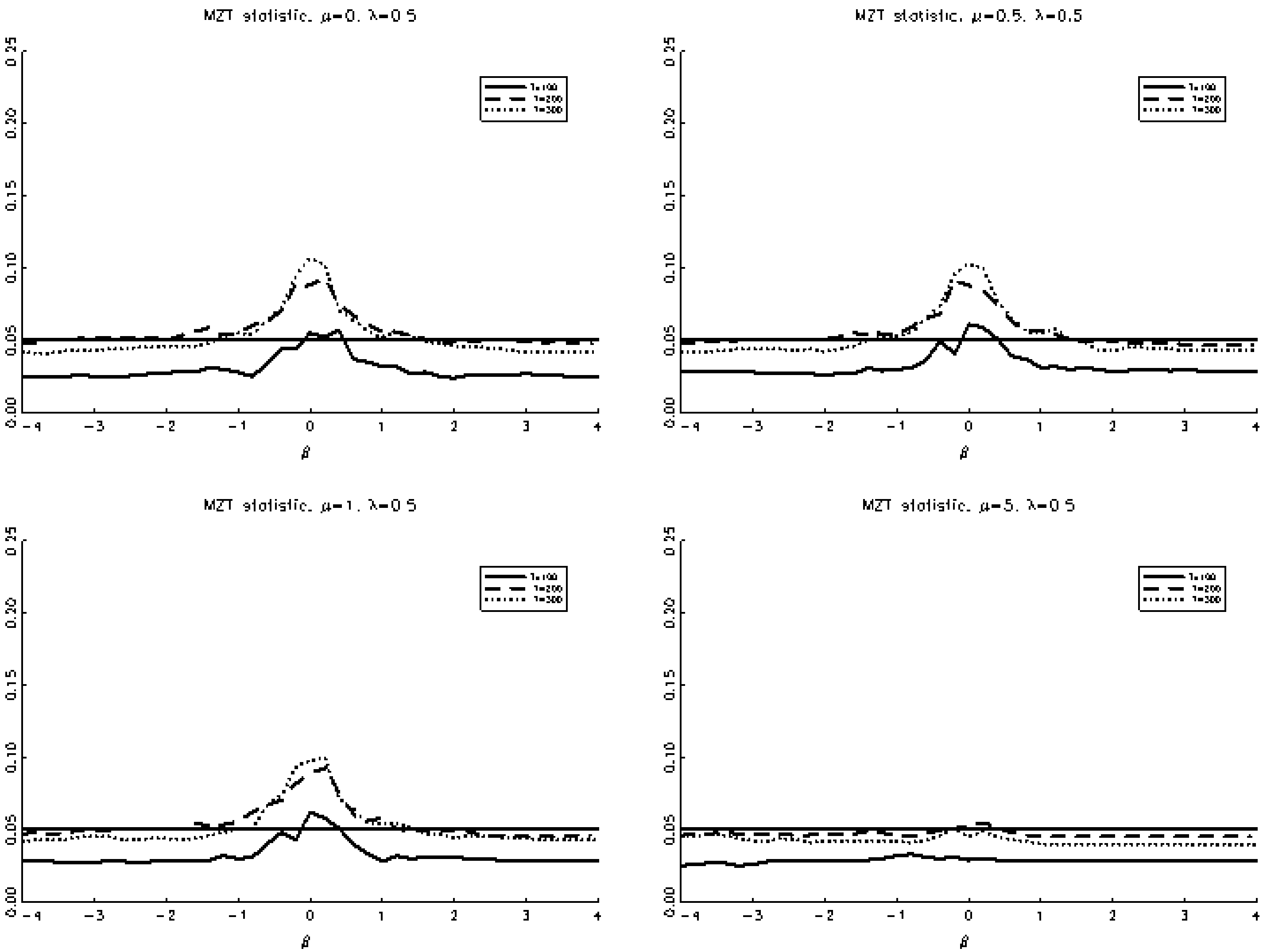


Figure 5: Empirical size for the MZT test, $\lambda = 0.5$



References

Bai, J., and P. Perron (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18, 1–22.

Carrion-i-Silvestre, J. L., A. Sansó, and M. Artís (2004): "Raíces Unitarias Y Cambios Estructurales En Las Macromagnitudes Españolas," *Revista de Economía Aplicada*, p. Forthcoming.

Clemente, J., A. Montañés, and M. Reyes (1998): "Testing for a Unit Root in Variables with a Double Change in the Mean," *Economics Letters*, (59), 175–182.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996): "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

Gadea, M. D., A. Montañés, and M. Reyes (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.

Garcia, R., and P. Perron (1996): "An Analysis of the Real Interest Rate under Regime Shifts," *The Review of Economics and Statistics*, pp. 111–125.

Kapetanios, G. (2005): "Unit Root Testing Against the Alternative Hypothesis of Up to M Structural Breaks," *Journal of Time Series Analysis*, 26, 123–133.

Lee, J. (1996): "Testing for a Unit Root in the Time Series with Trend Breaks," *Journal of Macroeconomics*, 18, 503–519.

Lee, J., C. J. Huang, and Y. Shin (1997): "On Stationary Tests in the Presence of Structural Breaks," *Economics Letters*, (55), 165–172.

Lee, J., and M. Strazicich (2003): "Minimum LM Unit Root Tests with Two Structural Breaks," *Review of Economics and Statistics. Forthcoming.*

Lumsdaine, R. L., and D. H. Papell (1997): "Multiple Trend Breaks and the Unit Root Hypothesis," *Review of Economics and Statistics*, (79), 212–218.

Montañés, A., and M. Reyes (1998): "Effect of a Shift in the Trend Function on Dickey-Fuller Unit Root Tests," *Econometric Theory*, 14, 355–363.

Ng, S., and P. Perron (2001): "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 69(6), 1519–1554.

Ohara, H. I. (1999): "A Unit Root Test with Multiple Trend Breaks: A Theory and an 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, 6), 1361–1401.

Perron, P., and Z. Qu (2006): "A Simple Modification to Improve the Finite Sample Properties of Ng and Perron's Unit Root Tests," *Economics Letters*, forthcoming.

Perron, P., and G. Rodríguez (2003): "GLS Detrending, Efficient Unit Root Tests and Structural Change," *Journal of Econometrics*, 115, 1–27.

Perron, P., and T. Vogelsang (1992): "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity," *Journal of Business & Economic Statistics*, (10, 3), 301–320.

Schmidt, P., and P. C. B. Phillips (1992): "LM Tests for a Unit Root in the Presence of Deterministic Trends," *Oxford Bulletin of Economics and Statistics*, (54), 257–287.

Yao, Y. C. (1988): "Estimating the Number of Change-Points Via Schwarz Criterion," *Statistics and Probability Letters*, 6, 181–189.

Zivot, E., and D. W. K. Andrews (1992): "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit-Root Hypothesis," *Journal of Business & Economic Statistics*, (10, 3), 251–270.