Hitotsubashi University

Hi-Stat Discussion Paper

Global COE Hi-Stat Discussion Paper Series 227

Research Unit for Statistical and Empirical Analysis in Social Sciences (Hi-Stat)

Testing for Multiple Structural Changes with Non-Homogeneous Regressors

Eiji Kurozumi

February 2012

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

Testing for Multiple Structural Changes with Non-Homogeneous Regressors

Eiji Kurozumi[1]

Department of Economics, Hitotsubashi University

February 21, 2011

Abstract

This paper investigates tests for multiple structural changes with non-homogeneous regressors, such as polynomial trends. We consider exponential-type, supremum-type and average-type tests as well as the corresponding weighted-type tests suggested in the literature. We show that the limiting distributions depend on regressors in general, and we need to tabulate critical values depending on them. Then, we focus on the linear trend case and obtain the critical values of the test statistics. The Mote Carlo simulations are conducted to investigate the finite sample properties of the tests proposed in the paper, and it is found that the specification of the number of breaks is an important factor for the finite sample performance of the tests. Since it is often the case that we cannot prespecify the number of breaks under the alternative but can suppose only the maximum number of breaks, the weighted-type tests are useful in practice.

JEL classification: C12, C22

Key words: multiple breaks, exp-type test, sup-type test, avg-type test, mean-type test

[1]I am grateful to Yuzo Honda and Pierre Perron for helpful comments. All errors are my responsibility. This research was supported by the Ministry of Education, Culture, Sports, Science and Technology under Grants-in-Aid No.18730142 and by the Global COE program of the Research Unit for Statistical and Empirical Analysis in Social Sciences, Hitotsubashi University. Address correspondence to Eiji Kurozumi, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan; e-mail: kurozumi@stat.hit-u.ac.jp

1. Introduction

This paper proposes tests for multiple structural changes with non-homogeneous regressors. In particular, we focus on trending regressors. Tests for structural changes have long been investigated in the econometric and statistical literature, and the most commonly used tests in empirical analysis for a one-time break are the supremum-type (sup-type) test by Andrews (1993) in the GMM framework and the exponential-type (exp-type) and the average-type (avg-type) tests by Andrews and Ploberger (1994) in linear regression models. The latter two tests have an optimal property, which was investigated by Andrews and Ploberger (1994) and Sowell (1996) under the Pitman-type alternative, while Kim and Perron (2009) compared these tests in a framework based on the Bahadur slope.

Although these tests are often used in practice to test for a one-time change, we need to take into account the possibility of multiple structural changes when economic data in long sample periods are available. Bai and Perron (1998) extended the sup-type test for the case of multiple structural changes in univariate regressions, while the multivariate case was considered by Qu and Perron (2007). Andrews, Lee, and Ploberger (1996) investigated the optimality of the exp-type and avg-type tests. Note that these tests are designed for the null hypothesis of no change against the alternative of the fixed number of breaks. On the other hand, Bai and Perron (1998) and Qu and Perron (2007) proposed double maximum tests against the alternative under which only the maximum number of breaks is prespecified, while Bai and Perron (1998), Bai (1999), and Qu and Perron (2007) considered tests for the null hypothesis of ℓ breaks against the alternative of $\ell + 1$ breaks. The multiple structural change tests have an advantage over the single structural change tests in that the former tests are more powerful than the latter when multiple breaks have actually occurred, as shown by Bai and Perron (2006).

The practical difficulty in the multiple structural change tests is that we need to take into account all permissible change points when constructing the test statistics. That is, for the sup-type, the exp-type, and the avg-type tests, we need to construct either the Wald, the likelihood ratio (LR), or the Lagrange multiplier test statistic for all permissible sets of change points, the number of which is proportional to T^m, where m indicates the number of breaks under the alternative. Then, the direct calculation of these test statistics is computationally very expensive when m is large. To overcome this problem, Bai and Perron (2003a) proposed an efficient algorithm for the sup-type test, which requires only the $O(T^2)$ calculations for any number of breaks. Critical values for the sup-type test are tabulated in Andrews (1993) for a one-time break and Bai and Perron (1998, 2003b) for multiple changes, and those for the exp-type and the avg-type tests are given in Andrews and Ploberger (2004) for a one-time change, while asymptotic p-values of these tests can be calculated by the method proposed by Hansen (1997). However, critical values for the exp-type and the avg-type tests with multiple breaks are not yet available.

Most of the above tests assume that regressors are homogeneous in the whole sample period, or at least in each regime under the null hypothesis. However, we sometimes include non-homogeneous regressors, such as trending variables. In this case, most of the above tests are not available in practical analysis. The exception is the LR test, denoted by $\sup F(\ell{+}1|\ell)$, for the null hypothesis of ℓ breaks against the alternative of $\ell + 1$ breaks proposed by Bai (1999). This test allows for polynomial trends, and hence the null hypothesis of no break can be tested using $\sup F(1|0)$. However, as pointed out by Bai and Perron (2006), this test may be less powerful than tests for multiple structural changes when multiple breaks have actually occurred.

In this paper, we develop tests for multiple breaks with non-homogeneous regressors, including trending regressors. We consider sup-type, exp-type, and avg-type tests, as in the literature, and derive the concise expressions of the limiting distributions. It is shown that in general, the limiting distributions depend on non-homogeneous regressors, and then, we need to tabulate critical values depending on the case. For this reason, we focus on the linear trend case and tabulate critical values. Since we need to calculate the Wald test statistics for all permissible break points for the exp-type tests, which is computationally very expensive in the case of more than three breaks, we tabulate the critical values of the exp-type test for at most three breaks, whereas those of the sup-type and avg-type tests are calculated for up to five breaks because they require $O(T^2)$ operations for any given number of breaks under the alternative. Finite sample properties are investigated by Mote Carlo simulations, and it is confirmed that the tests that assume the maximum number of breaks but not the specific number of breaks are useful in practical analysis.

The rest of this paper is organized as follows. Section 2 explains a model and assumptions. The test statistics are given in Section 3, and their limiting distributions are derived. Section 4 discusses the computational problem of the test statistics, and the finite sample properties are investigated in Section 5. Section 6 gives concluding remarks.

2. Model and Assumptions

Let us consider the following regression with m structural changes ($m+1$ regimes):

$$y_t = x_t'\beta_j + \varepsilon_t \quad (j = 1\cdots, m+1 \quad \text{and} \quad t = T_{j-1}+1, \cdots, T_j), \tag{1}$$

where x_t is p-dimensional regressors, including a constant, and ε_t is an error term. We set $T_0 = 0$ and $T_{m+1} = T$ so that the total number of observations is T. The testing problem we consider is given by

$$H_0: \beta_1 = \cdots = \beta_{m+1} \quad \text{vs.} \quad H_1: \beta_i \neq \beta_j \text{ for some } i \neq j,$$

and we then consider the null of no structural change. The following assumptions are made throughout the paper.

Assumption A1 *For some normalizing matrix D_T, (a) $D_T^{-1}\sum_{t=k}^{l} x_t x_t' D_T^{-1}$ is invertible for $l-k > k_0$ for some $0 < k_0 < \infty$. (b) $D_T^{-1}\sum_{t=1}^{[Tr]} x_t x_t' D_T^{-1} \xrightarrow{p} \Omega_r$ uniformly over $0 < r \leq 1$, where Ω_r is a $p \times p$ positive definite matrix for $0 < r \leq 1$ with $\Omega_0 = 0$, $[k]$ signifies the largest integer less than k, and $\xrightarrow{p}$ signifies convergence in probability. (c) $\Omega_s - \Omega_r$ is positive definite for all $0 \leq r < s \leq 1$.*

Assumption A2 *(a) $\{\varepsilon_t\}$ is a martingale difference sequence with respect to $\mathcal{F}_t = \sigma(\varepsilon_t, \varepsilon_{t-1}, \cdots, x_{t+1}, x_t, \cdots)$ with $E[\varepsilon_t^2|\mathcal{F}_{t-1}] = \sigma^2$ for all t. (b) $sup_t E|\varepsilon_t|^{2+\delta} < \infty$ for some $\delta > 0$. (c) For the same normalizing matrix D_T in Assumption A2, $D_T^{-1}\sum_{t=1}^{[Tr]} x_t\varepsilon_t \Rightarrow \sigma G(r)$ for $0 \leq r \leq 1$, where $G(r)$ is a p-dimensional Gaussian process with mean zero and $E[G(r)G'(s)] = \Omega_{r\wedge s}$, and $\Rightarrow$ signifies weak convergence of the associated probability measures.*

Assumption A1(a) is made for the identification of the coefficient. A1(b) allows for non-homogeneous regressors because the second moment of x_t is not asymptotically proportional to the sample fraction r, but possibly depends on r in a complicated way. A1(c) is required for technical reasons. Assumption A2 is standard in linear regressions, but we do not allow for serial correlation in the error term. However, since the lagged dependent variables are allowed as regressors, the dynamic property of the model may be captured through them.

Exactly speaking, Assumptions A1 and A2 are required only under the null hypothesis in order to derive the null limiting distributions of the test statistics and they can be relaxed under the alternative in order for the tests to be consistent. See, for example, Assumptions made in Bai and Perron (1998) for the case of (regime-wise) stationary regressors.

One of the interesting non-homogeneous regressors is a polynomial trend. For example, when x_t is given by

$$x_t' = [1, t, t^2, \cdots, t^d, x_{1t}, \cdots, x_{qt}],$$

where $x_{1t}, \cdots, x_{qt}$ are stationary regressors, we can choose $D_T = diag\{T^{1/2}, T^{3/2}, T^{5/2}, \cdots, T^{d+1/2}, T^{1/2} I_q\}$ and Assumptions A1(b) and A2(c) then become

$$D_T^{-1} \sum_{t=1}^{[Tr]} x_t x_t' D_T^{-1} \xrightarrow{p} \begin{bmatrix} r & \frac{r^2}{2} & \cdots & \frac{r^{d+1}}{d+1} & r\mu_x' \\ \frac{r^2}{2} & \frac{r^3}{3} & \cdots & \frac{r^{d+2}}{d+2} & \frac{r^2}{2}\mu_x' \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{r^{d+1}}{d+1} & \frac{r^{d+2}}{d+2} & \cdots & \frac{r^{2d+1}}{2d+1} & \frac{r^{d+1}}{d+1}\mu_x' \\ r\mu_x & \frac{r^2}{2}\mu_x & \cdots & \frac{r^{d+1}}{d+1} & \Gamma_x \end{bmatrix} \quad \text{and} \quad D_T^{-1} \sum_{t=1}^{[Tr]} x_t \varepsilon_t \Rightarrow \sigma \begin{bmatrix} \int_0^r dW_1(s) \\ \int_0^r s^2 dW_1(s) \\ \vdots \\ \int_0^r s^d dW_1(s) \\ \Gamma_x^{1/2} W_2(r) \end{bmatrix},$$

where μ_x and Γ_x are $q \times 1$ and $q \times q$ and consist of the first and second moments of $x_{1t}, \cdots, x_{qt}$, respectively, $W_1(r)$ and $W_2(r)$ are 1- and q-dimensional standard Brownian motions on $[0, 1]$ and $[0, 1]^q$, respectively, and they are independent of each other. Apparently, the second moment of x_t is not proportional to r, and hence we cannot use the exiting results for multiple structural changes.

3. Tests for Multiple Structural Changes

In this section, we define the test statistics for multiple structural changes and derive their limiting distributions. Let $\hat{\beta} = [\hat{\beta}_1', \cdots, \hat{\beta}_{m+1}]'$ be the least squares estimator of the coefficients for a given number of breaks m with change points $\{T_1, \cdots, T_m\}$, $\hat{\Sigma} = diag\{\hat{\Sigma}_1, \cdots, \hat{\Sigma}_{m+1}\}$ be an $(m+1)p \times (m+1)p$ block-diagonal matrix where $\hat{\Sigma}_j = (\sum_{t=T_{j-1}}^{T_j} x_t x_t')^{-1}$, and $\hat{\sigma}^2$ be a consistent estimator of σ^2. Typically, $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2$, where $\hat{\varepsilon}_t$ is the regression residual. Then, the Wald test statistic for the null hypothesis of H_0 is given by

$$W_T(\Lambda_m) = \left(R\hat{\beta}\right)' \left(\hat{\sigma}^2 R\hat{\Sigma}R'\right)^{-1} \left(R\hat{\beta}\right), \quad \text{where} \quad R = \begin{bmatrix} I_p & -I_p & & & 0 \\ & I_p & -I_p & & \\ & & \ddots & \ddots & \\ & & & I_p & -I_p \end{bmatrix}$$

and $\Lambda_m = \{\lambda_1, \cdots, \lambda_m\}$ with $\lambda_j = T_j/T$ for $j = 1, \cdots, m$ being break fractions. We set $\lambda_0 = 0$ and $\lambda_{m+1} = 1$ for convention because $T_0 = 0$ and $T_{m+1} = T$. Using $W_T(\Lambda_m)$, we construct the exp-type, the sup-type, and the avg-type tests, as in the literature.

$$\begin{aligned} \text{exp-}W_T(m,\epsilon) &= \log\left[\frac{1}{T^*}\sum_{\Lambda_m \in \Lambda_m^\epsilon} \exp\left(\frac{1}{2}W_T(\Lambda_m)\right)\right], & (2) \\ \text{sup-}W_T(m,\epsilon) &= \max_{\Lambda_m \in \Lambda_m^\epsilon} W_T(\Lambda_m), & (3) \\ \text{avg-}W_T(m,\epsilon) &= \frac{1}{T^*}\sum_{\Lambda_m \in \Lambda_m^\epsilon} W_T(\Lambda_m), & (4) \end{aligned}$$

where $\Lambda_m^\epsilon = \{(\lambda_1, \cdots, \lambda_m) : \lambda_j - \lambda_{j-1} \geq \epsilon \;\text{ for }\; j = 1, \cdots, m+1\}$ for a given trimming parameter ϵ and T^* is the number of permissible sets of break fractions included in Λ_m^ϵ. The trimming parameter ϵ should be small; $\epsilon = 0.05$, 0.1, and 0.15 have often been considered in the literature. As discussed in Andrews, Lee, and Ploberger (1996), the exp-type test is optimal against the alternative of the large magnitude of structural changes, whereas the avg-type test is asymptotically most powerful against the alternative of small changes.

The above three tests require the specific number of breaks m under the alternative before constructing the test statistics, but, if we do not want to prespecify the number of breaks, then we may set only the maximum number of breaks given by M and consider the following weighted exp-type and avg-type tests suggested by Andrews, Lee, and Ploberger (1996) as well as the weighted double maximum test proposed by Bai and Perron (1998):

$$
\begin{aligned}
\text{Wexp-}W_T(M,\epsilon) &= \sum_{m=1}^{M} \frac{c_{exp}(p,\alpha,1)}{c_{exp}(p,\alpha,m)} \text{exp-}W_T(m,\epsilon), \\
\text{WDmax-}W_T(M,\epsilon) &= \max_{1\le m\le M} \frac{c_{sup}(p,\alpha,1)}{c_{sup}(p,\alpha,m)} \text{sup-}W_T(m,\epsilon), \\
\text{Wavg -}W_T(M,\epsilon) &= \sum_{m=1}^{M} \frac{c_{avg}(p,\alpha,1)}{c_{avg}(p,\alpha,m)} \text{avg-}W_T(m,\epsilon),
\end{aligned}
$$

where $c_i(p,\alpha,m)$ for $i = exp$, sup, and avg are the critical values of (2)–(4) for a given m with significance level α. These weights are suggested by Bai and Perron (1998).

The limiting distributions of these test statistics are given by the following theorem.

Theorem 1 *Assume that Assumptions A1 and A2 hold. Then, under the null hypothesis,*

$$
\text{exp-}W_T(m,\epsilon) \xrightarrow{d} \log \int_{\Lambda_m \in \Lambda_m^\epsilon} \exp\left(\frac{1}{2} W(\Lambda_m)\right) d\Lambda_m,
$$

$$
\text{sup-}W_T(m,\epsilon) \xrightarrow{d} \sup_{\Lambda_m \in \Lambda_m^\epsilon} W(\Lambda_m), \quad \text{avg-}W_T(m,\epsilon) \xrightarrow{d} \int_{\Lambda_m \in \Lambda_m^\epsilon} W(\Lambda_m) d\Lambda_m,
$$

$$
\text{Wexp-}W_T(M,\epsilon) \xrightarrow{d} \sum_{m=1}^{M} \frac{c_{exp}(p,\alpha,1)}{c_{exp}(p,\alpha,m)} \log \int_{\Lambda_m \in \Lambda_m^\epsilon} \exp\left(\frac{1}{2} W(\Lambda_m)\right) d\Lambda_m,
$$

$$
\text{WDmax-}W_T(M,\epsilon) \xrightarrow{d} \max_{1\le m\le M} \frac{c_{sup}(p,\alpha,1)}{c_{sup}(p,\alpha,m)} \sup_{\Lambda_m \in \Lambda_m^\epsilon} W(\Lambda_m),
$$

$$
\text{Wavg -}W_T(M,\epsilon) \xrightarrow{d} \sum_{m=1}^{M} \frac{c_{avg}(p,\alpha,1)}{c_{avg}(p,\alpha,m)} \int_{\Lambda_m \in \Lambda_m^\epsilon} W(\Lambda_m) d\Lambda_m,
$$

where

$$
\begin{aligned}
W(\Lambda_m) &= Q_{1,m} \qquad (5) \\
&\equiv \sum_{j=1}^{m} \left(\Omega_{\lambda_{j+1}}^{-1} G(\lambda_{j+1}) - \Omega_{\lambda_j}^{-1} G(\lambda_j)\right)' \left(\Omega_{\lambda_j}^{-1} - \Omega_{\lambda_{j+1}}^{-1}\right)^{-1} \left(\Omega_{\lambda_{j+1}}^{-1} G(\lambda_{j+1}) - \Omega_{\lambda_j}^{-1} G(\lambda_j)\right).
\end{aligned}
$$

Remark 1 *Theorem 1 shows that although the Wald test statistic for a given set of break points Λ_m is asymptotically chi-square distributed, $W(\Lambda_{m,1})$ is correlated with $W(\Lambda_{m,2})$ in a complicated way for $\Lambda_{m,1} \neq \Lambda_{m,2}$ $(\Lambda_{m,1}, \Lambda_{m,2} \in \Lambda_m^\epsilon)$, and the test statistics for unknown breaks are then nonstandard.*

Remark 2 *When the regressors are homogeneous with $\Omega_r = r\Omega$, we have $G(r) = B(r)$, where $B(r)$ is a p-dimensional standard Brownian motion on $[0,1]^p$, and it is then not difficult to show that*

$$W(\Lambda_m) = Q_{2,m} \equiv \sum_{j=1}^{m} \frac{\|\lambda_j B(\lambda_{j+1}) - \lambda_{j+1} B(\lambda_j)\|}{\lambda_j \lambda_{j+1}(\lambda_{j+1} - \lambda_j)}, \tag{6}$$

which is given by Bai and Perron (1998) for the case of stationary regressors. Thus, Theorem 1 includes the existing result as a special case.

As we can see from Theorem 1, the limiting distributions of the test statistics depend on the structure of Ω_r, and then, we need to calculate critical values for a given regressor x_t. The dependency of critical values on x_t have sometimes been observed in different situations in the literature. For example, the critical values for unit root tests depend on whether a linear trend is included as a regressor, while the LR tests for cointegrating rank are known to have different distributions depending on the structure of the deterministic term.

In the following, we focus on the case where x_t includes a linear trend, which is widely used in practical analysis. More precisely, let us consider the case where

$$x_t = [1, t, x_{1t}, \cdots, x_{qt}]', \tag{7}$$

with $x_{1t}, \cdots, x_{qt}$ being stationary variables and lagged dependent variables. In this case, we have the following corollary.

Corollary 1 *Assume that Assumptions A1 and A2 hold. Then, under the null hypothesis with x_t given by (7), Theorem 1 holds with $W(\Lambda_m) = Q_{1,m} + Q_{2,m}$, where $Q_{1,m}$ is given by (5) with*

$$\Omega_r^{-1} = \begin{bmatrix} \frac{4}{r} & -\frac{6}{r^2} \\ -\frac{6}{r^2} & \frac{12}{r^3} \end{bmatrix} \quad \text{and} \quad \left(\Omega_r^{-1} - \Omega_s^{-1}\right)^{-1} = \begin{bmatrix} \frac{rs(r^2+rs+s^2)}{(s-r)^3} & \frac{(rs)^2(r+s)}{2(s-r)^3} \\ \frac{(rs)^2(r+s)}{2(s-r)^3} & \frac{(rs)^3}{3(s-r)^3} \end{bmatrix}$$

for $0 < r < s \leq 1$ and $G(r) = [B_1(r), \int_0^r s dB_1(s)]'$, where $B_1(r)$ is a one-dimensional standard Brownian motion on $[0,1]$, while $Q_{2,m}$ is given by (6) with $B(r)$ being a q-dimensional standard Brownian motion independent of $B_1(r)$.

The result in Corollary 1 is similar to that given by Bai (1999) for testing the null hypothesis of ℓ breaks against the alternative of $\ell + 1$ breaks; the limiting distribution is the sum of the two independent distributions corresponding to (constant plus) a linear trend and stationary regressors. We can see that the limiting distribution of Bai's (1999) test with $\ell = 0$ is the same as ours with $m = 0$.

4. Computation of Critical Values

Since the limiting distributions of the test statistics are nonstandard, we obtain the critical values by simulations with $G(r)$ approximated by 1,000 partial sums of the appropriate pseudo random variables. For example, if $G(r)$ is a standard Brownian motion, then we approximate $G(r)$ using the normalized partial sums of $i.i.d.N(0,1)$ pseudo-random variables. However, the computation of the critical values is not necessarily easy for large values of m because the number of permissible sets of breaks is proportional to T^m, so the direct calculation of all permissible Wald test statistics is computationally too expensive when $m \geq 4$. For the sup-type test, Bai and Perron (2003a) gives an efficient algorithm for the computation of the test statistics, which requires operations of order $O(T^2)$ for any given number of breaks; we can use this in our case.

For the avg-type test, we can also calculate the critical values computationally efficiently[2]. Let $Q_1(T_a, T_b)$ be the summand of (5) approximated by the above method with T observations given $T_a = \lambda_a T$ and $T_b = \lambda_b T$. Since the distance between two consecutive break points must be at least $h = \epsilon T$, the permissible ranges of $T_1, T_2, \cdots, T_m$ are $T_1 = h, h+1, \cdots, T - hm$, $T_2 = T_1 + h, T_1 + h + 1, \cdots, T - h(m-1)$, $\cdots$, $T_m = T_{m-1} + h, T_{m-1} + h + 1, \cdots, T - h$; then, the limiting distribution of the avg-type test statistic can be approximated by

$$\frac{1}{T^*} \sum_{T_1=h}^{T-hm} \sum_{T_2=T_1+h}^{T-h(m-1)} \cdots \sum_{T_m=T_{m-1}+h}^{T-h} \sum_{j=1}^{m} Q_1(T_j, T_{j+1}). \tag{8}$$

However, (8) requires the summation operators of order $O(T^m)$, which is computationally expensive as explained above. Instead, we calculate the limiting distributions by noting that each of $Q_1(T_j, T_{j+1})$ appears in (8) many times; if we count them, we can save computational time. For example, $Q_1(T_1, T_2)$ appears as many times as the permissible number of allocations

[2]Pierre Perron pointed out the existence of the efficient calculation of the avg-type test through personal communications.

of $T_3, \cdots, T_m$ in $[T_2,\ T]$. Since, in general, the permissible number of combinations of ℓ breaks in $[T_a, T_b]$ with two consecutive breaks' distance being larger than h is given by

$$k_h(T_a, T_b, \ell) = \frac{1}{\ell!}\prod_{i=1}^{\ell}\{(T_a - T_b + 1) - h(\ell+1) + i\},$$

which is obtained by direct calculations, we can see that $Q_1(T_1, T_2)$ appears $k_h(T_2, T, m-2)$ times in (8). Similarly, we observe $Q_1(T_m, T_{m+1})$ as many times as the number of allocations of $T_1, \cdots, T_{m-1}$ in $[1, T_m]$, which is given by $k_h(1, T_m, m-1)$. For the case of $Q_1(T_j, T_{j+1})$ for $j = 2, \cdots, m-1$, there are $j-1$ and $m-j-1$ breaks allocated in $[1, T_j]$ and $[T_{j+1}, T]$, respectively. Then, we can see that

$$\begin{aligned}(8) \quad = \quad & \frac{1}{k_h(1, T, m)} \sum_{j=1}^{m-1} \sum_{T_a=jh}^{T-h(m-j+1)} \sum_{T_b=T_a+h}^{T-h(m-j)} k_h(1, T_a, j-1) Q_1(T_a, T_b, j) k_h(T_b, T, m-j-1) \\ & + \sum_{T_c=hm}^{T-h} k_h(1, T_c, m-1) Q_1(T_c, T, m), \end{aligned} \tag{9}$$

where $k_h(T_a, T_b, 0) = 1$ for convention. We can see that the number of summation operators on the right hand side of (9) is proportional to $O(T^2)$ for any given number of m.

On the other hand, we cannot find an efficient computational method for the exp-type test. Therefore, we consider the exp-type test only up to $m = 3$.

The critical values in the case of a linear trend are given in Tables 1-3 for $\varepsilon = 0.05$, 0.10, and 0.15 and $q = 0$ to 9, where q is the number of homogeneous regressors. They are obtained by approximating Brownian motions by 1,000 partial sums of $i.i.d.N(0,1)$ pseudo-random variables with 10,000 replications. Because of the above reason, the critical values for the exp-type test is given for only up to $m = 3$ and $M = 3$ whereas those for the sup-type and avg-type tests are obtained for up to $m = 5$ and $M = 3$ and 5. As in the case of homogeneous regressors, the critical values get larger as q and/or m increase.

5. Finite Sample Properties

In this section, we investigate the finite sample properties of the proposed tests via Monte Carlo simulations. We consider two cases where $q = 0$ (DGP0) and $q = 1$ (DGP1). In the case of $q = 0$, the data generating process under the null hypothesis is given by

$$y_t = \beta_1 + \beta_2 t + \varepsilon_t$$

for $t = 1, \cdots, T$ with $\varepsilon_t \sim i.i.d.N(0, 1)$. We set $\beta_1 = \beta_2 = 0$ because all the test statistics are invariant to the true values of β_1 and β_2 under H_0. On the other hand, the DGP1 has an autoregressive (AR) regressor as follows:

$$y_t = \beta_1 + \beta_2 t + \beta_3 x_t + \varepsilon_t, \quad x_t = \phi x_{t-1} + u_t$$

for $t = 1, \cdots, T$, where, again, we set $\beta_1 = \beta_2 = \beta_3 = 0$ without loss of generality. The initial value of x_t is set to $x_0 = 0$ while $\phi = -0.8$, -0.4, 0, 0.4, and 0.8. The sample size T is 120 and 300 and the significance level is set to 0.05. We investigate the case where the maximum number of breaks is three, so that we construct the weighted-type tests with $M = 3$. All computations are conducted using the GAUSS matrix language with 2,000 replications.

Table 4 shows the empirical sizes of the tests. For comparison, we also consider the LR test for the null hypothesis of no break against the alternative of a one-time break proposed by Bai (1999). For DGP0 with $T = 120$, the empirical sizes of all the tests are close to the nominal one except for the exp-type test with $m = 3$, which tends to overly reject the null hypothesis slightly, while the LR test is conservative. However, the size distortions of the tests disappear when the sample size is 300.

On the other hand, when the stationary variable is included as a regressor, almost all the tests reject the null hypothesis more frequently compared to DGP0. As a result, we observe the tendency of the over-rejection for the exponential type test with $m = 2$ and $m = 3$; however, again, this tendency disappears when $T = 300$. As a whole, the empirical sizes of all the tests are not greatly affected by the AR parameter of the stationary regressor as long as ϕ ranges from -0.8 to 0.8; its effect seems marginal.

To see the finite sample powers of the tests, we first consider the following DGP with a one-time break for $q = 0$ and $q = 1$, respectively:

$$\text{DGP0} \quad y_t = \begin{cases} \beta_{1,1} + \beta_{2,1} t + \varepsilon_t & : \quad t = 1, \cdots, T_1, \\ \beta_{1,2} + \beta_{2,1} t + \beta_{2,2}(t - T_1) + \varepsilon_t & : \quad t = T_1 + 1, \cdots, T, \end{cases}$$

$$\text{DGP1} \quad y_t = \begin{cases} \beta_{1,1} + \beta_{2,1} t + \beta_{3,1} x_t + \varepsilon_t & : \quad t = 1, \cdots, T_1, \\ \beta_{1,2} + \beta_{2,1} t + \beta_{2,2}(t - T_1) + \beta_{3,2} x_t + \varepsilon_t & : \quad t = T_1 + 1, \cdots, T, \end{cases}$$

where $T_1 = 0.5T$, $\beta_{1,1} = \beta_{2,1} = \beta_{3,1} = 0$, while $\beta_{1,2} = \beta_{3,2} = 5\gamma$ and $\beta_{2,2} = \gamma$ with γ taking positive values, which are chosen so that the overall shapes of the power functions can be observed.

Figure 1 shows the size-adjusted powers of the tests when $q = 0$ (DGP0). As expected, the test against the correct number of breaks ($m = 1$) is more powerful than the others among the same type of tests. For example, we observe from (i-a) and (i-b) that the exp-type test with $m = 1$ is most powerful, followed by the tests with $m = 2$ and $m = 3$. The effect of the misspecification of the number of breaks is relatively large for the sup-type tests, whereas the avg-type tests are less affected by m. We can also see that the weighted-type tests are the second-best tests compared to tests with a fixed number of breaks. Figures 1(iv-a) and (iv-b) compare the three weighted-type tests and the LR test by Bai (1999). We observe that the weighted avg-type test is most powerful, and the second-best is the weighted exp-type test. The weighted double maximum test and the LR test are inferior to the others in this case, and the former is slightly less powerful than the latter.

Figure 2 corresponds to the case where $q = 1$ (DGP1) and the AR coefficient of x_t is 0. We observe that the relative performance is preserved compared to Figure 1. Regarding the effect of the persistence of x_t, the tests are most powerful when x_t is an *i.i.d.* sequence ($\phi = 0$) and the powers decrease as the absolute values of ϕ get larger. However, the difference between the powers is not substantial, and the effect of the persistence of the regressor is slight for $-0.8 \leq \phi \leq 0.8$ (we do not show the figure to save space).

We next investigate the case where the number of breaks is two. The DGP in this case is given by

$$\text{DGP0} \quad y_t = \begin{cases} \beta_{1,1} + \beta_{2,1}t + \varepsilon_t & : \ t = 1, \cdots, T_1, \\ \beta_{1,2} + \beta_{2,1}t + \beta_{2,2}(t - T_1) + \varepsilon_t & : \ t = T_1 + 1, \cdots, T_2, \\ \beta_{1,3} + \beta_{2,1}t + \beta_{2,2}(t - T_1) + \beta_{2,3}(t - T_2) + \varepsilon_t & : \ t = T_2 + 1, \cdots, T, \end{cases}$$

$$\text{DGP1} \quad y_t = \begin{cases} \beta_{1,1} + \beta_{2,1}t + \beta_{3,1}x_t + \varepsilon_t & : \ t = 1, \cdots, T_1, \\ \beta_{1,2} + \beta_{2,1}t + \beta_{2,2}(t - T_1) + \beta_{3,2}x_t + \varepsilon_t & : \ t = T_1 + 1, \cdots, T_2, \\ \beta_{1,3} + \beta_{2,1}t + \beta_{2,2}(t - T_1) + \beta_{2,3}(t - T_2) + \beta_{3,3}x_t + \varepsilon_t & : \ t = T_2 + 1, \cdots, T, \end{cases}$$

where $T_1 = 0.3T$ and $T_2 = 0.7T$, and we consider two kinds of changes: The first is the case of two successive increases in the coefficients; $\beta_{1,1} = \beta_{2,1} = \beta_{3,1} = 0$ while $\beta_{1,2} = \beta_{3,2} = 5\gamma$ and $\beta_{2,2} = \gamma$ in the second regime and $\beta_{1,3} = \beta_{3,3} = 10\gamma$ and $\beta_{2,3} = \gamma$ in the third regime with

γ taking positive values. The second case is such that the first break occurs in the upwards directions while the dependent variable crashes down by the second break; $\beta_{1,1} = \beta_{2,1} = \beta_{3,1} = 0$ while $\beta_{1,2} = \beta_{3,2} = 5\gamma$ and $\beta_{2,2} = \gamma$ in the second regime and $\beta_{1,3} = \beta_{3,3} = -5\gamma$ and $\beta_{2,3} = -0.5\gamma$ in the third regime.

Figure 3 shows the size-adjusted powers for $q = 0$. We observe that the test with under-misspecification of the number of breaks ($m = 1$) is still most powerful than the others in this case for each type of the tests. As in the one-time break case, the second-best are the weighted-type tests, while the avg-type test is most powerful among the weighted-type tests, followed by the exp-type test, as is observed from (iv-a) and (iv-b). Similar property is observed with a stationary regressor from Figure 4, but as long as the avg-type test is concerned, the specification of the number of breaks does not so much affect the differences in powers. In addition, we do not observe the significant differences in powers among the weighted-type tests. On the other hand, the persistence in the AR regressor affects the finite sample powers very much. The case with $\phi = 0$ is most powerful whereas $\phi = 0.8$ corresponds to the least powerful case. The maximum differences in powers in these two cases with $m = 2$ are 0.490, 0.493, and 0.454 for the exp-type, sup-type and avg-type tests, respectively, when $T = 120$. The similar magnitude of the differences is observed even when $T = 300$ (we do not show the figure to save space).

Figure 5 shows the size-adjusted powers in the second case of the two breaks for $q = 0$. When $T = 120$, the test correctly specifying the number of breaks ($m = 2$) is most powerful in each type of the tests. In particular, the differences in powers are relatively large for the avg-type tests. On the other hand, when $T = 300$, the differences become smaller and as long as the sup-type test is concerned, the test with $m = 1$ becomes most powerful. The differences in powers are more pronounced when a stationary variable is included as a regressor, as is observed from Figure 6. The effect of the persistence of the AR regressor on the powers is mitigated in this case; the largest differences in the powers of the tests with $m = 2$ are 0.169, 0.158, and 0.115 for the exp-type, sup-type and avg-type tests, respectively, when $T = 120$, and the differences becomes marginal when $T = 300$.

In summary, the performance of the tests depends on the DGP and none of the tests dominates the others uniformly, while we also observe that the weighted-type tests are the second-best in most cases and the powers of those tests are close to those of the best tests.

6. Concluding Remarks

In this paper, we have investigated tests for multiple breaks with non-homogeneous regressors. We have derived the limiting distributions of the test statistics in a general case and found that the limiting distributions depend on the regressors. By focusing on the linear trend case, we have obtained the critical values for the sup-type and avg-type tests by computationally efficient methods (although we cannot find such a method for the exp-type test), and have obtained the critical values for the exp-type test only up to $m = 3$. By Monte Carlo simulations, we have showed that the correct specification of the number of breaks is very important in order for the tests to have good power. However, since we often cannot specify the specific number of breaks under the alternative but can only suppose the maximum number of breaks, the weighted-type tests would be useful in practice.

Appendix

Proof of Theorem 1: According to Assumptions A1 and A2, we can see that under the null hypothesis,

$$\begin{aligned} D_T\left(\hat{\beta}_j - \beta\right) &= \left(D_T^{-1}\sum_{t=T_{j-1}+1}^{T_j} x_t x_t' D_T^{-1}\right)^{-1} D_T^{-1}\sum_{t=T_{j-1}+1}^{T_j} x_t u_t \\ &\xrightarrow{d} \sigma\left(\Omega_{\lambda_j} - \Omega_{\lambda_{j-1}}\right)^{-1}\left(G(\lambda_j) - G(\lambda_{j-1})\right) \equiv \sigma\tilde{\Sigma}_{\lambda_j}\tilde{G}(\lambda_j), \end{aligned}$$

where $\tilde{\Sigma}_{\lambda_j} = (\Omega_{\lambda_j} - \Omega_{\lambda_{j-1}})^{-1}$ with $\Omega_{\lambda_0} = 0$ and $\tilde{G}(\lambda_j) = G(\lambda_j) - G(\lambda_{j-1})$ with $G(\lambda_0) = 0$. Then, we can see that

$$W_T(\Lambda_m) \xrightarrow{d} (R\tilde{\Sigma}\tilde{G})'(R\tilde{\Sigma}R')^{-1}(R\tilde{\Sigma}\tilde{G})$$

uniformly over the permissible sets of break fractions, where $\tilde{G} = [\tilde{G}'(\lambda_1), \cdots, \tilde{G}'(\lambda_{m+1})]'$ and $\tilde{\Sigma} = diag\{\tilde{\Sigma}_{\lambda_1}, \cdots, \tilde{\Sigma}_{\lambda_{m+1}}\}$. Then, we need to show that

$$\left(R\tilde{\Sigma}\tilde{G}\right)'\left(R\tilde{\Sigma}R'\right)^{-1}\left(R\tilde{\Sigma}\tilde{G}\right) = Q_{1,m}, \tag{10}$$

where $Q_{1,m}$ is defined in (5). Note that the difficulty is in that the inverse of the variance matrix is given by

$$\left(R\tilde{\Sigma}R'\right)^{-1} = \begin{bmatrix} \tilde{\Sigma}_{\lambda_1}+\tilde{\Sigma}_{\lambda_2} & -\tilde{\Sigma}_{\lambda_2} & 0 & \cdots & 0 \\ -\tilde{\Sigma}_{\lambda_2} & \tilde{\Sigma}_{\lambda_2}+\tilde{\Sigma}_{\lambda_3} & -\tilde{\Sigma}_{\lambda_3} & \ddots & \vdots \\ 0 & -\tilde{\Sigma}_{\lambda_3} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\tilde{\Sigma}_m \\ 0 & \cdots & 0 & -\tilde{\Sigma}_{\lambda_m} & \tilde{\Sigma}_{\lambda_m}+\tilde{\Sigma}_{\lambda_{m+1}} \end{bmatrix}^{-1}, \qquad (11)$$

which does not have a simple closed-form expression.

In order to treat (11), we introduce a nonsingular matrix H and transform the left hand side of (10) using H as $(HR\tilde{\Sigma}\tilde{G})'(HR\tilde{\Sigma}R'H')^{-1}(HR\tilde{\Sigma}\tilde{G})$ and evaluate the transformed expression. More precisely, we define an $mp \times mp$ lower triangular matrix H and decompose the restriction matrix R as

$$H = \begin{bmatrix} -I_p & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ -I_p & \cdots & \cdots & -I_p \end{bmatrix}, \quad R = \left[\begin{array}{c|ccccc} 0 & -I_p & 0 & \cdots & 0 \\ \vdots & I_p & -I_p & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & I_p & -I_p \end{array}\right] + \left[\begin{array}{c|cccc} I_p & 0 & \cdots & \cdots & 0 \\ 0 & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \end{array}\right]$$

$$\text{so that} \quad HR = \left[\begin{array}{c|ccccc} 0 & I_p & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & I_p \end{array}\right] - \left[\begin{array}{c|cccc} I_p & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ I_p & 0 & \cdots & \cdots & 0 \end{array}\right].$$

Let us decompose $\tilde{\Sigma}$ as $\tilde{\Sigma} = \mathrm{diag}\{\tilde{\Sigma}_{\lambda_1}, \tilde{\Sigma}_{2,m+1}\}$, where $\tilde{\Sigma}_{2,m+1} = \mathrm{diag}\{\tilde{\Sigma}_{\lambda_2}, \cdots, \tilde{\Sigma}_{\lambda_{m+1}}\}$. Then, we can see that

$$\begin{aligned} \left(HR\tilde{\Sigma}R'H'\right)^{-1} &= \left(\tilde{\Sigma}_{2,m+1} + F_p\tilde{\Sigma}_{\lambda_1}F_p'\right)^{-1} \\ &= \tilde{\Sigma}_{2,m+1}^{-1} - \tilde{\Sigma}_{2,m+1}^{-1}F_p\left(\tilde{\Sigma}_{\lambda_1}^{-1} + F_p'\tilde{\Sigma}_{2,m+1}^{-1}F_p\right)^{-1}F_p'\tilde{\Sigma}_{2,m+1}^{-1} \\ &= \tilde{\Sigma}_{2,m+1}^{-1} - \tilde{\Sigma}_{2,m+1}^{-1}F_p\left(\tilde{\Sigma}_{\lambda_1}^{-1} + \cdots + \tilde{\Sigma}_{\lambda_{m+1}}^{-1}\right)^{-1}F_p'\tilde{\Sigma}_{2,m+1}^{-1}, \end{aligned} \qquad (12)$$

where $F_p = [I_p, \cdots, I_p]'$ is an $mp \times p$ matrix, while

$$HR\tilde{\Sigma}\tilde{G} = \begin{bmatrix} \tilde{\Sigma}_{\lambda_2}\tilde{G}(\lambda_2) - \tilde{\Sigma}_{\lambda_1}\tilde{G}(\lambda_1) \\ \vdots \\ \tilde{\Sigma}_{\lambda_{m+1}}\tilde{G}(\lambda_{m+1}) - \tilde{\Sigma}_{\lambda_1}\tilde{G}(\lambda_1) \end{bmatrix}. \qquad (13)$$

Then, since $\tilde{\Sigma}_{\lambda_j} = (\Omega_{\lambda_j} - \Omega_{\lambda_{j-1}})^{-1}$ and $\tilde{G}(\lambda_j) = G(\lambda_j) - G(\lambda_{j-1})$, we can see from (13) that

$$\begin{aligned}
(HR\tilde{\Sigma}\tilde{G})'\tilde{\Sigma}_{2,m+1}^{-1}(HR\tilde{\Sigma}\tilde{G}) &= \sum_{j=1}^{m}\tilde{G}'(\lambda_{j+1})\tilde{\Sigma}_{\lambda_{j+1}}\tilde{G}(\lambda_{j+1}) - 2\sum_{j=1}^{m}\tilde{G}'(\lambda_1)\tilde{\Sigma}_{\lambda_1}\tilde{G}(\lambda_{j+1}) \\
&\quad +\tilde{G}'(\lambda_1)\tilde{\Sigma}_{\lambda_1}(\tilde{\Sigma}_{\lambda_2}^{-1} + \cdots + \tilde{\Sigma}_{\lambda_{m+1}}^{-1})\tilde{\Sigma}_{\lambda_1}\tilde{G}(\lambda_1) \\
&= \sum_{j=1}^{m}(G(\lambda_{j+1}) - G(\lambda_j))'(\Omega_{\lambda_{j+1}} - \Omega_{\lambda_j})^{-1}(G(\lambda_{j+1}) - G(\lambda_j)) \\
&\quad -2\left\{G'(\lambda_1)\Omega_{\lambda_1}^{-1}G(\lambda_{m+1}) - G'(\lambda_1)\Omega_{\lambda_1}^{-1}G(\lambda_1)\right\} \\
&\quad +G'(\lambda_1)\Omega_{\lambda_1}^{-1}\Omega_{\lambda_{m+1}}\Omega_{\lambda_1}^{-1}G(\lambda_1) - G'(\lambda_1)\Omega_{\lambda_1}^{-1}G(\lambda_1). \qquad (14)
\end{aligned}$$

Similarly, since $F_p'\tilde{\Sigma}_{2,m+1}^{-1} = [\tilde{\Sigma}_{\lambda_2}^{-1}, \cdots, \tilde{\Sigma}_{\lambda_{m+1}}^{-1}]$, we have

$$\begin{aligned}
F_p'\tilde{\Sigma}_{2,,m+1}^{-1}HR\tilde{\Sigma}\tilde{G} &= \sum_{j=1}^{m}\tilde{G}(\lambda_{j+1}) - (\tilde{\Sigma}_{\lambda_2}^{-1} + \cdots + \tilde{\Sigma}_{\lambda_{m+1}}^{-1})\tilde{\Sigma}_{\lambda_1}\tilde{G}_{\lambda_1} \\
&= G(\lambda_{m+1}) - G(\lambda_1) - (\Omega_{\lambda_{m+1}} - \Omega_{\lambda_1})\Omega_{\lambda_1}^{-1}G(\lambda_1) \\
&= G(\lambda_{m+1}) - \Omega_{\lambda_{m+1}}\Omega_{\lambda_1}^{-1}G(\lambda_1),
\end{aligned}$$

so that

$$\begin{aligned}
&(HR\tilde{\Sigma}\tilde{G})'\tilde{\Sigma}_{2,m+1}^{-1}F_p\left(\tilde{\Sigma}_{\lambda_1}^{-1} + \cdots + \tilde{\Sigma}_{\lambda_{m+1}}^{-1}\right)^{-1}F_p'\tilde{\Sigma}_{2,m+1}^{-1}(HR\tilde{\Sigma}\tilde{G}) \\
&= \left(G(\lambda_{m+1}) - \Omega_{\lambda_{m+1}}\Omega_{\lambda_1}^{-1}G(\lambda_1)\right)'\Omega_{\lambda_{m+1}}^{-1}\left(G(\lambda_{m+1}) - \Omega_{\lambda_{m+1}}\Omega_{\lambda_1}^{-1}G(\lambda_1)\right) \\
&= G'(\lambda_{m+1})\Omega_{\lambda_{m+1}}^{-1}G(\lambda_{m+1}) - 2G'(\lambda_1)\Omega_{\lambda_1}^{-1}G(\lambda_{m+1}) + G'(\lambda_1)\Omega_{\lambda_1}^{-1}\Omega_{\lambda_{m+1}}\Omega_{\lambda_1}^{-1}G(\lambda_1). \qquad (15)
\end{aligned}$$

Then, by combining (12), (14) and (15), we have

$$\begin{aligned}
&(HR\tilde{\Sigma}\tilde{G})'\left(HR\tilde{\Sigma}R'H'\right)^{-1}(HR\tilde{\Sigma}\tilde{G}) \\
&= \sum_{j=1}^{m}(G(\lambda_{j+1}) - G(\lambda_j))'\left(\Omega_{\lambda_{j+1}} - \Omega_{\lambda_j}\right)^{-1}(G(\lambda_{j+1}) - G(\lambda_j)) \\
&\quad +G'(\lambda_1)\Omega_{\lambda_1}^{-1}G(\lambda_1) - G'(\lambda_{m+1})\Omega_{\lambda_{m+1}}^{-1}G(\lambda_{m+1}) \\
&= \sum_{j=1}^{m}\Big\{(G(\lambda_{j+1}) - G(\lambda_j))'\left(\Omega_{\lambda_{j+1}} - \Omega_{\lambda_j}\right)^{-1}(G(\lambda_{j+1}) - G(\lambda_j)) \\
&\quad +G'(\lambda_j)\Omega_{\lambda_j}^{-1}G(\lambda_j) - G'(\lambda_{j+1})\Omega_{\lambda_{j+1}}^{-1}G(\lambda_{j+1})\Big\} \\
&= \sum_{j=1}^{m}\left(\Omega_{\lambda_{j+1}}^{-1}G(\lambda_{j+1}) - \Omega_{\lambda_j}^{-1}G(\lambda_j)\right)'\left(\Omega_{\lambda_{j+1}}^{-1} - \Omega_{\lambda_j}^{-1}\right)^{-1}\left(\Omega_{\lambda_{j+1}}^{-1}G(\lambda_{j+1}) - \Omega_{\lambda_j}^{-1}G(\lambda_j)\right),
\end{aligned}$$

where the last equality is obtained by using the following three equivalent expressions:

$$\begin{aligned}
\left(\Omega_{\lambda_{j+1}}-\Omega_{\lambda_j}\right)^{-1} &= \Omega_{\lambda_{j+1}}^{-1}+\Omega_{\lambda_{j+1}}(\Omega_{\lambda_j}^{-1}-\Omega_{\lambda_{j+1}}^{-1})^{-1}\Omega_{\lambda_{j+1}}^{-1} \\
&= -\Omega_{\lambda_j}^{-1}+\Omega_{\lambda_j}^{-1}(\Omega_{\lambda_j}^{-1}-\Omega_{\lambda_{j+1}}^{-1})^{-1}\Omega_{\lambda_j}^{-1} \\
&= \Omega_{\lambda_j}^{-1}(\Omega_{\lambda_j}^{-1}-\Omega_{\lambda_{j+1}}^{-1})^{-1}\Omega_{\lambda_{j+1}}^{-1}.\blacksquare
\end{aligned}$$

Proof of Corollary 1: Since a constant term is included as a regressor, the stationary variables can be assumed to be mean zero without loss of generality. Similarly, because the lagged dependent variables can be decomposed into a constant, a linear trend, and stationary components, we can treat the expectation of the lagged dependent variables to be zero because 1 and t are included as regressors. Then, we can see that Ω_r becomes a block diagonal matrix with the first 2-by-2 diagonal block given by

$$\begin{bmatrix} r & \frac{r^2}{2} \\ \frac{r^2}{2} & \frac{r^3}{3} \end{bmatrix}$$

and the last q-by-q block given by $r\Omega_2$, where Ω_2 consists of the second moments of the stationary regressors. The result immediately follows because of the diagonality of Ω_r.$\blacksquare$

References

[1] Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica* 61, 821-856 (Erratum, 71, 395-397).

[2] Andrews, D. W. K., I. Lee and W. Ploberger (1996). Optimal Changepoint Tests for Normal Linear Regression. *Journal of Econometrics* 70, 9-38.

[3] Andrews, D. W. K. and W. Ploberger (1994). Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative, *Econometrica* 62, 1383-1414.

[4] Bai, J. and P. Perron (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* 66, 47-78.

[5] Bai, J. (1999). Likelihood Ratio Tests for Multiple Structural Changes. *Journal of Econometrics* 91, 299-323.

[6] Bai, J. and P. Perron (2003a). Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics* 18, 1-22.

[7] Bai, J. and P. Perron (2003b). Critical Values for Multiple Structural Change Tests. *Econometrics Journal* 6, 72-78.

[8] Bai, J. and P. Perron (2006). Estimating and Testing Linear Models with Multiple Structural Changes, D. Corbae, S. N. Durlauf and B. E. Hansen ed., *Econometric Theory and Practice*. Cambridge University Press, Cambridge.

[9] Hansen, B. E. (1997). Approximate Asymptotic P Values for Structural-Change Tests. *Journal of Business and Economic Statistics* 15, 60-67.

[10] Kim, D. and P. Perron (2009). Assessing the Relative Power of Structural Break Tests Using a Framework Based on the Approximate Bahadur Slope. *Journal of Econometrics* 149, 26-51.

[11] Qu, Z. and P. Perron (2007). Estimating and Testing Structural Changes in Multivariate Regressions. *Econometrica* 75, 459-502.

[12] Sowell, B. (1996) Optimal Tests for Parameter Instability in the Generalized Method of Moments Framework. *Econometrica* 64, 1085-1107.

Table 1: Asymptotic critical values of the exp-type test with level α

		q									
m	α	0	1	2	3	4	5	6	7	8	9
exp-$W_T(m, 0.05)$ ($\varepsilon = 0.05$)											
	.10	2.63	3.60	4.49	5.38	6.18	6.98	7.86	8.59	9.36	10.14
1	.05	3.26	4.32	5.28	6.21	7.10	7.99	8.81	9.64	10.42	11.17
	.01	4.80	5.83	7.10	8.00	9.00	9.93	10.79	11.70	12.61	13.43
	.10	4.55	6.32	7.97	9.61	11.16	12.70	14.23	15.71	17.20	18.61
2	.05	5.35	7.18	8.97	10.66	12.26	13.88	15.43	16.99	18.51	19.96
	.01	7.20	9.28	10.98	12.93	14.77	16.42	18.12	19.66	21.35	22.97
	.10	6.29	8.80	11.18	13.49	15.75	17.98	20.20	22.41	24.57	26.60
3	.05	7.21	9.82	12.37	14.79	17.12	19.38	21.62	23.90	26.08	28.27
	.01	9.37	12.11	14.65	17.28	19.69	22.40	24.67	27.10	29.29	31.68
Wexp-$W_T(3, 0.05)$ ($\varepsilon = 0.05$, $M = 3$)											
	.10	7.81	10.65	13.33	15.99	18.30	20.77	23.30	25.51	27.84	30.25
	.05	9.66	12.79	15.60	18.29	21.02	23.59	26.20	28.55	30.78	33.08
	.01	14.14	17.35	20.81	23.75	26.46	29.21	31.79	34.42	37.02	39.87
exp-$W_T(m, 0.10)$ ($\varepsilon = 0.10$)											
	.10	2.59	3.56	4.45	5.33	6.11	6.90	7.74	8.48	9.24	10.00
1	.05	3.25	4.27	5.25	6.16	6.98	7.92	8.72	9.52	10.32	11.08
	.01	4.82	5.90	7.04	8.00	8.98	9.82	10.67	11.60	12.50	13.34
	.10	4.43	6.18	7.80	9.38	10.84	12.35	13.89	15.32	16.75	18.14
2	.05	5.28	7.06	8.80	10.41	11.97	13.54	15.14	16.70	18.13	19.57
	.01	7.10	9.20	10.89	12.65	14.40	16.27	17.77	19.32	20.85	22.56
	.10	6.07	8.56	10.84	13.05	15.21	17.39	19.56	21.65	23.70	25.75
3	.05	7.06	9.60	12.02	14.36	16.59	18.83	20.98	23.24	25.40	27.41
	.01	9.15	12.00	14.31	16.80	19.30	21.78	23.97	26.20	28.52	30.91
Wexp-$W_T(3, 0.10)$ ($\varepsilon = 0.10$, $M = 3$)											
	.10	7.70	10.55	13.17	15.81	18.10	20.48	22.97	25.13	27.44	29.78
	.05	9.63	12.63	15.44	18.11	20.61	23.36	25.79	28.20	30.56	32.86
	.01	14.18	17.25	20.69	23.48	26.26	28.81	31.47	34.19	36.78	39.25
exp-$W_T(m, 0.15)$ ($\varepsilon = 0.15$)											
	.10	2.56	3.51	4.42	5.25	6.05	6.83	7.61	8.40	9.13	9.86
1	.05	3.21	4.25	5.20	6.10	6.94	7.84	8.60	9.38	10.23	10.99
	.01	4.84	5.90	6.92	7.93	8.97	9.79	10.63	11.56	12.34	13.26
	.10	4.33	6.06	7.68	9.19	10.62	12.12	13.62	14.93	16.37	17.77
2	.05	5.21	6.99	8.60	10.18	11.75	13.31	14.83	16.28	17.84	19.24
	.01	7.07	9.14	10.67	12.59	14.07	15.87	17.50	19.08	20.54	22.05
	.10	5.92	8.34	10.47	12.63	14.77	16.89	18.96	20.96	22.97	24.99
3	.05	6.87	9.35	11.69	13.90	16.12	18.37	20.45	22.57	24.59	26.63
	.01	8.96	11.78	14.17	16.57	18.91	21.25	23.38	25.65	27.97	30.07
Wexp-$W_T(3, 0.15)$ ($\varepsilon = 0.15$, $M = 3$)											
	.10	7.56	10.34	13.09	15.56	17.87	20.22	22.52	24.86	27.00	29.22
	.05	9.50	12.60	15.28	17.91	20.41	23.04	25.44	27.78	30.34	32.57
	.01	14.18	17.27	20.33	23.19	26.35	28.65	31.13	33.85	36.18	38.80

Table 2: Asymptotic critical values of the sup-type test with level α

		q									
m	α	0	1	2	3	4	5	6	7	8	9
sup-$W_T(m, 0.05)$ ($\varepsilon = 0.05$)											
	.10	12.10	14.25	16.10	17.94	19.79	21.60	23.21	24.86	26.60	28.16
1	.05	13.70	16.03	17.95	20.00	21.84	23.69	25.48	27.18	28.88	30.56
	.01	17.50	20.34	22.48	24.58	26.38	28.19	30.44	32.28	33.74	35.65
	.10	22.41	26.62	30.15	33.64	36.92	40.39	43.62	46.80	49.82	52.85
2	.05	24.78	28.87	32.69	36.18	39.57	43.07	46.31	49.66	52.68	55.69
	.01	28.96	33.67	38.03	41.47	45.18	48.70	52.20	55.87	59.03	62.18
	.10	31.02	36.92	42.31	47.35	52.25	57.16	61.86	66.62	71.06	75.49
3	.05	33.62	39.65	45.20	50.35	55.36	60.21	65.27	69.79	74.57	79.06
	.01	38.72	45.22	50.93	56.43	61.67	66.86	71.76	77.15	81.92	86.45
	.10	38.89	46.48	53.58	60.33	66.66	72.94	79.16	85.22	91.34	97.06
4	.05	41.57	49.40	56.67	63.51	69.96	76.38	82.90	89.11	94.90	100.96
	.01	47.60	56.17	63.02	70.45	77.15	83.76	90.96	97.12	103.79	109.31
	.10	45.90	55.12	63.92	72.13	79.96	87.85	95.34	102.75	110.20	117.24
5	.05	48.72	58.51	67.47	75.77	83.77	91.71	99.60	107.03	114.51	121.77
	.01	55.08	65.82	74.31	83.09	91.16	99.74	108.76	116.48	124.60	131.19
WDmax-$W_T(3, 0.05)$ ($\varepsilon = 0.05$, $M = 3$)											
	.10	13.13	15.31	17.24	19.16	21.00	22.92	24.57	26.25	28.01	29.60
	.05	14.68	17.19	19.19	21.29	23.20	25.04	26.87	28.61	30.41	31.92
	.01	18.77	21.64	23.80	25.99	27.78	29.76	31.96	33.98	35.30	37.40
WDmax-$W_T(5, 0.05)$ ($\varepsilon = 0.05$, $M = 5$)											
	.10	13.26	15.48	17.43	19.36	21.22	23.09	24.78	26.51	28.21	29.82
	.05	14.93	17.38	19.35	21.47	23.40	25.24	27.04	28.87	30.66	32.33
	.01	19.04	21.83	24.01	26.17	28.08	30.04	32.19	34.14	35.53	37.52
sup-$W_T(m, 0.10)$ ($\varepsilon = 0.10$)											
	.10	11.54	13.54	15.52	17.41	19.02	20.71	22.50	24.00	25.61	27.31
1	.05	13.32	15.51	17.43	19.31	21.10	22.97	24.73	26.51	28.11	29.55
	.01	16.96	19.41	21.61	23.80	25.75	27.55	29.27	31.18	32.96	34.62
	.10	20.62	24.33	27.86	31.36	34.61	37.90	41.25	44.25	47.32	50.30
2	.05	22.66	26.67	30.48	34.00	37.29	40.68	44.05	47.26	50.45	53.37
	.01	27.30	31.99	35.54	39.27	42.85	46.62	49.79	53.20	56.51	59.75
	.10	27.70	33.21	38.46	43.38	48.28	53.09	57.60	62.13	66.51	70.83
3	.05	30.07	36.10	41.48	46.63	51.36	56.22	60.75	65.57	70.15	74.54
	.01	35.16	41.76	47.25	52.22	57.53	62.85	67.49	72.46	77.27	82.54
	.10	33.70	40.75	47.42	53.92	60.04	66.34	72.18	78.18	83.90	89.32
4	.05	36.35	43.93	50.79	57.30	63.64	69.94	76.21	82.23	88.04	93.74
	.01	42.27	50.28	57.46	64.29	70.72	77.03	83.19	89.61	96.12	101.48
	.10	38.69	47.19	55.36	63.19	70.58	78.13	85.22	92.66	99.46	106.41
5	.05	41.64	50.56	58.80	66.86	74.55	82.08	89.39	96.87	104.25	111.19
	.01	47.39	57.01	66.05	74.41	82.18	90.19	97.62	105.40	113.23	120.14
WDmax-$W_T(3, 0.10)$ ($\varepsilon = 0.10$, $M = 3$)											
	.10	12.53	14.71	16.72	18.63	20.22	22.01	23.76	25.33	27.08	28.73
	.05	14.47	16.67	18.68	20.50	22.42	24.31	26.17	27.99	29.52	31.14
	.01	18.19	20.71	22.78	25.22	27.27	29.07	30.80	32.68	34.62	36.45
WDmax-$W_T(5, 0.10)$ ($\varepsilon = 0.10$, $M = 5$)											
	.10	12.75	14.95	16.92	18.85	20.51	22.25	24.11	25.63	27.35	29.02
	.05	14.68	16.93	18.93	20.75	22.75	24.69	26.45	28.26	29.79	31.35
	.01	18.48	20.87	23.09	25.54	27.55	29.34	31.14	33.10	34.87	36.65

Table 2: (Continued)

m	α	q = 0	1	2	3	4	5	6	7	8	9
sup-$W_T(m, 0.15)$ $(\varepsilon = 0.15)$											
	.10	11.07	13.03	14.97	16.81	18.48	20.16	21.83	23.34	24.92	26.65
1	.05	12.80	14.95	16.91	18.70	20.42	22.35	23.95	25.75	27.51	29.01
	.01	16.54	18.87	20.81	23.03	25.07	26.88	28.70	30.65	32.28	33.97
	.10	18.92	22.71	26.21	29.58	32.78	35.97	39.10	42.04	45.16	47.93
2	.05	21.05	25.03	28.73	32.02	35.33	38.73	42.05	45.15	48.25	51.17
	.01	25.62	30.32	33.59	37.55	40.88	44.58	47.79	50.84	54.05	57.56
	.10	24.66	30.05	34.93	39.72	44.36	49.06	53.52	57.83	62.22	66.33
3	.05	27.02	32.68	37.98	42.84	47.61	52.63	56.94	61.33	65.77	70.13
	.01	32.29	38.60	43.66	49.08	53.92	59.02	63.64	68.15	73.02	77.19
	.10	28.50	35.30	41.37	47.56	53.46	59.31	64.99	70.59	76.16	81.62
4	.05	31.28	38.23	44.75	50.91	56.70	63.22	68.78	74.44	80.19	85.77
	.01	36.62	44.41	51.11	57.56	64.34	70.51	76.35	82.29	88.67	94.30
	.10	28.82	36.30	43.61	50.63	57.41	64.02	70.49	77.03	83.70	89.74
5	.05	31.39	39.52	46.95	54.35	61.10	67.97	74.56	81.42	88.16	94.58
	.01	36.78	45.68	53.60	61.06	68.40	76.30	82.94	89.34	96.70	103.25
WDmax-$W_T(3, 0.15)$ $(\varepsilon = 0.15,\ M = 3)$											
	.10	12.12	14.23	16.20	18.07	19.74	21.55	23.19	24.75	26.41	28.18
	.05	13.85	16.17	18.16	19.98	21.75	23.58	25.41	27.21	28.93	30.55
	.01	17.74	20.23	22.30	24.11	26.50	28.22	30.29	32.08	33.87	35.82
WDmax-$W_T(5, 0.15)$ $(\varepsilon = 0.15,\ M = 5)$											
	.10	12.39	14.52	16.51	18.40	20.11	21.93	23.61	25.24	26.87	28.64
	.05	14.27	16.50	18.52	20.36	22.22	24.10	25.85	27.66	29.41	31.04
	.01	18.12	20.83	22.62	24.71	26.95	28.65	30.67	32.81	34.45	36.48

Table 3: Asymptotic critical values of the avg-type test with level α

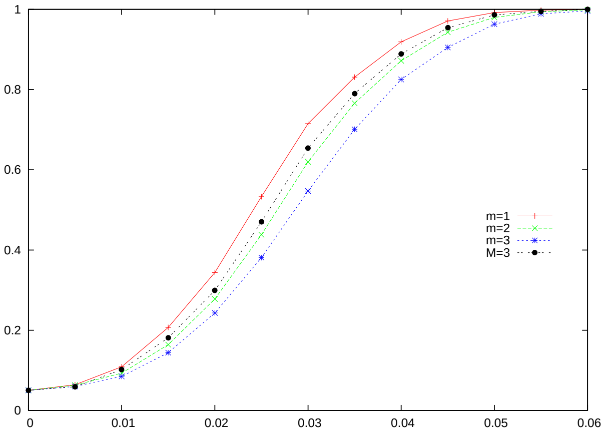
m	α	q = 0	1	2	3	4	5	6	7	8	9
avg-$W_T(m, 0.05)$ ($\varepsilon = 0.05$)											
	.10	3.51	4.86	6.24	7.44	8.72	9.93	11.10	12.34	13.52	14.63
1	.05	4.32	5.77	7.08	8.49	9.84	11.17	12.38	13.63	14.84	16.04
	.01	6.32	7.80	9.20	10.51	12.02	13.44	14.90	16.33	17.56	18.76
	.10	6.36	8.94	11.39	13.82	16.16	18.52	20.82	23.10	25.44	27.63
2	.05	7.49	10.19	12.78	15.25	17.71	20.21	22.66	25.01	27.42	29.70
	.01	10.15	12.97	15.63	18.22	20.95	23.64	26.21	28.82	30.97	33.48
	.10	9.00	12.78	16.35	19.89	23.27	26.71	30.12	33.48	36.85	40.19
3	.05	10.35	14.29	18.02	21.60	25.11	28.79	32.26	35.88	39.31	42.80
	.01	13.46	17.55	21.24	25.13	29.07	32.87	36.41	40.29	43.91	47.48
	.10	11.50	16.42	21.15	25.78	30.27	34.84	39.25	43.71	48.19	52.57
4	.05	13.14	18.26	23.10	27.82	32.39	37.13	41.67	46.35	50.95	55.40
	.01	16.67	21.70	26.69	31.82	36.73	41.87	46.28	51.53	56.34	60.94
	.10	14.09	20.11	25.91	31.58	37.18	42.86	48.27	53.86	59.46	64.87
5	.05	15.80	22.11	28.07	33.89	39.64	45.44	50.92	56.71	62.32	68.00
	.01	19.70	25.90	32.09	38.29	44.45	50.54	56.45	62.38	68.48	74.21
Wavg-$W_T(3, 0.05)$ ($\varepsilon = 0.05$, $M = 3$)											
	.10	10.48	14.54	18.59	22.20	26.00	29.70	33.16	36.81	40.38	43.77
	.05	12.94	17.24	21.10	25.29	29.42	33.36	37.08	40.76	44.34	47.90
	.01	18.66	23.18	27.59	31.42	35.85	40.21	44.85	48.85	52.16	55.89
Wavg-$W_T(5, 0.05)$ ($\varepsilon = 0.05$, $M = 5$)											
	.10	17.35	24.19	30.91	36.92	43.27	49.33	55.16	61.24	67.18	72.71
	.05	21.24	28.53	35.01	41.96	48.70	55.31	61.60	67.57	73.90	79.67
	.01	30.94	38.59	45.79	51.96	59.72	66.38	73.57	80.71	86.74	92.61
avg-$W_T(m, 0.10)$ ($\varepsilon = 0.10$)											
	.10	3.58	4.98	6.40	7.60	8.88	10.11	11.34	12.52	13.77	14.88
1	.05	4.45	5.97	7.32	8.70	10.10	11.48	12.68	13.92	15.22	16.39
	.01	6.64	8.15	9.49	10.93	12.36	13.91	15.42	16.79	18.06	19.35
	.10	6.50	9.19	11.67	14.12	16.48	18.86	21.19	23.47	25.85	28.09
2	.05	7.74	10.47	13.09	15.63	18.07	20.68	23.08	25.48	27.95	30.19
	.01	10.67	13.39	16.19	18.76	21.45	24.40	26.89	29.43	31.82	34.33
	.10	9.21	13.09	16.71	20.33	23.69	27.22	30.65	33.93	37.46	40.84
3	.05	10.67	14.68	18.48	22.14	25.71	29.39	32.93	36.64	40.03	43.39
	.01	14.16	18.28	21.96	25.98	30.04	33.72	37.40	41.17	44.99	48.72
	.10	11.82	16.91	21.58	26.24	30.88	35.52	39.87	44.37	48.96	53.33
4	.05	13.64	18.78	23.77	28.50	33.18	38.04	42.60	47.43	52.02	56.42
	.01	17.28	22.72	27.69	32.90	37.79	42.84	47.76	52.73	57.84	62.53
	.10	14.47	20.66	26.48	32.14	37.88	43.78	49.12	54.65	60.44	65.82
5	.05	16.35	22.84	28.87	34.83	40.53	46.48	52.09	57.96	63.60	69.29
	.01	20.44	27.22	33.37	39.73	45.63	51.88	57.75	64.26	70.31	76.11
Wavg-$W_T(3, 0.10)$ ($\varepsilon = 0.10$, $M = 3$)											
	.10	10.66	14.87	19.01	22.65	26.52	30.19	33.89	37.39	41.10	44.44
	.05	13.22	17.72	21.75	25.88	30.09	34.25	37.88	41.48	45.42	48.90
	.01	19.31	24.06	28.36	32.32	36.83	41.53	45.73	50.08	53.43	57.61
Wavg-$W_T(5, 0.10)$ ($\varepsilon = 0.10$, $M = 5$)											
	.10	17.68	24.64	31.56	37.61	43.95	49.98	56.21	62.08	68.22	73.94
	.05	21.81	29.27	35.96	42.96	49.75	56.61	62.74	69.00	75.49	81.26
	.01	32.36	39.74	46.82	53.46	61.04	68.48	76.05	82.83	88.92	95.59

Table 3: (Continued)

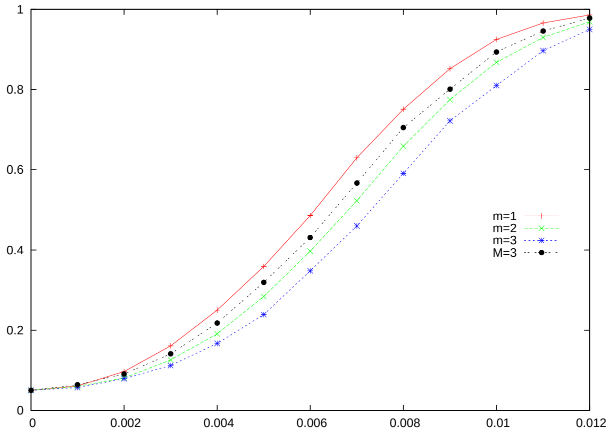
m	α	q 0	1	2	3	4	5	6	7	8	9
avg-$W_T(m, 0.15)$ ($\varepsilon = 0.15$)											
	.10	3.66	5.10	6.53	7.78	9.07	10.31	11.59	12.70	13.99	15.12
1	.05	4.61	6.13	7.54	8.96	10.35	11.72	12.94	14.24	15.50	16.74
	.01	6.89	8.52	9.87	11.37	12.87	14.30	15.92	17.30	18.60	19.99
	.10	6.65	9.36	11.94	14.35	16.78	19.18	21.58	23.85	26.17	28.49
2	.05	7.92	10.74	13.45	15.95	18.45	21.01	23.55	25.95	28.42	30.72
	.01	10.99	13.80	16.54	19.43	22.14	24.85	27.29	30.07	32.60	35.17
	.10	9.41	13.40	17.10	20.66	24.12	27.72	31.09	34.55	37.98	41.44
3	.05	10.99	15.16	18.98	22.55	26.30	30.00	33.66	37.27	40.88	44.22
	.01	14.63	18.78	22.70	26.87	30.78	34.52	38.21	42.30	46.21	49.68
	.10	12.16	17.37	22.11	26.85	31.49	36.15	40.68	45.28	49.82	54.19
4	.05	14.02	19.32	24.36	29.14	33.96	38.86	43.48	48.28	52.87	57.59
	.01	17.79	23.53	28.92	34.27	39.16	44.09	49.14	54.31	59.61	64.31
	.10	15.09	21.38	27.47	33.43	39.16	44.86	50.55	56.14	61.93	67.48
5	.05	17.01	23.80	30.13	36.22	42.17	48.22	54.08	60.00	65.63	71.42
	.01	21.31	28.98	35.54	41.93	48.24	54.26	60.74	67.02	73.35	79.35
Wavg-$W_T(3, 0.15)$ ($\varepsilon = 0.15$, $M = 3$)											
	.10	10.89	15.18	19.42	23.12	27.02	30.67	34.56	37.90	41.69	45.04
	.05	13.62	18.14	22.31	26.61	30.76	34.98	38.49	42.36	46.14	49.76
	.01	20.07	24.95	29.01	33.45	38.20	42.72	47.18	51.23	54.94	59.15
Wavg-$W_T(5, 0.15)$ ($\varepsilon = 0.15$, $M = 5$)											
	.10	17.91	25.01	32.01	38.30	44.64	50.68	57.04	62.46	69.05	74.68
	.05	22.39	29.83	36.81	43.74	50.63	57.41	63.58	69.96	76.41	82.33
	.01	33.59	41.02	47.67	55.21	62.69	70.18	77.68	84.14	90.41	97.53

Table 4: Empirical sizes of the tests

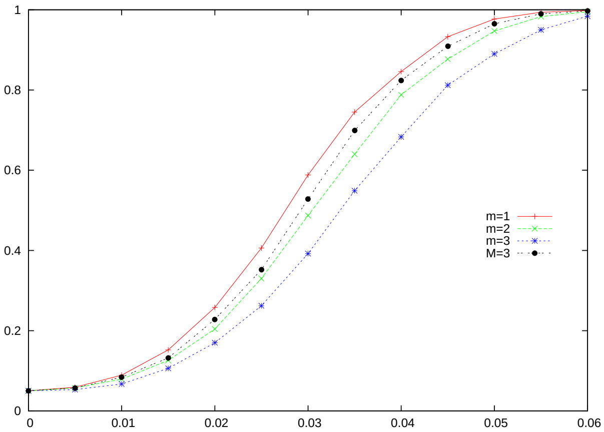
	DGP0		DGP1									
T	120	300			120					300		
ϕ	-	-	-0.8	-0.4	0.0	0.4	0.8	-0.8	-0.4	0.0	0.4	0.8
exp(1)	0.058	0.049	0.064	0.062	0.059	0.058	0.060	0.048	0.048	0.050	0.043	0.050
exp(2)	0.064	0.056	0.085	0.080	0.079	0.076	0.077	0.053	0.062	0.060	0.057	0.063
exp(3)	0.071	0.057	0.096	0.095	0.089	0.094	0.092	0.058	0.069	0.064	0.059	0.066
Wexp	0.066	0.053	0.076	0.072	0.071	0.073	0.076	0.051	0.057	0.060	0.053	0.059
sup(1)	0.041	0.044	0.046	0.043	0.041	0.038	0.041	0.035	0.040	0.036	0.035	0.040
sup(2)	0.043	0.047	0.057	0.053	0.050	0.049	0.050	0.041	0.053	0.043	0.043	0.048
sup(3)	0.053	0.052	0.065	0.062	0.058	0.064	0.067	0.044	0.050	0.050	0.049	0.054
WDmax	0.045	0.042	0.056	0.056	0.050	0.055	0.053	0.041	0.045	0.044	0.041	0.050
avg(1)	0.047	0.051	0.055	0.052	0.049	0.051	0.052	0.041	0.043	0.048	0.044	0.047
avg(2)	0.052	0.055	0.058	0.057	0.059	0.053	0.053	0.048	0.052	0.054	0.050	0.057
avg(3)	0.054	0.053	0.060	0.054	0.059	0.052	0.053	0.049	0.055	0.055	0.048	0.056
Wavg	0.051	0.054	0.056	0.056	0.059	0.051	0.051	0.047	0.049	0.053	0.047	0.053
LR	0.037	0.041	0.041	0.040	0.039	0.036	0.037	0.035	0.038	0.034	0.034	0.039



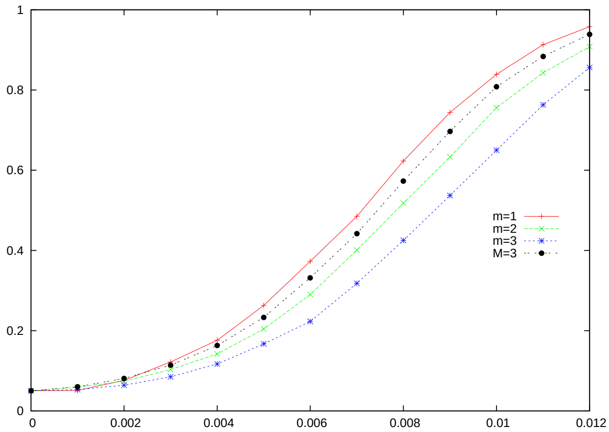
(i-a) exp-type tests ($T = 120$)



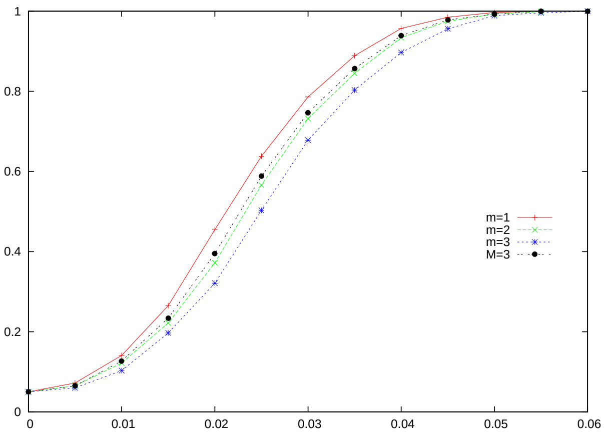
(i-b) exp-type tests ($T = 300$)



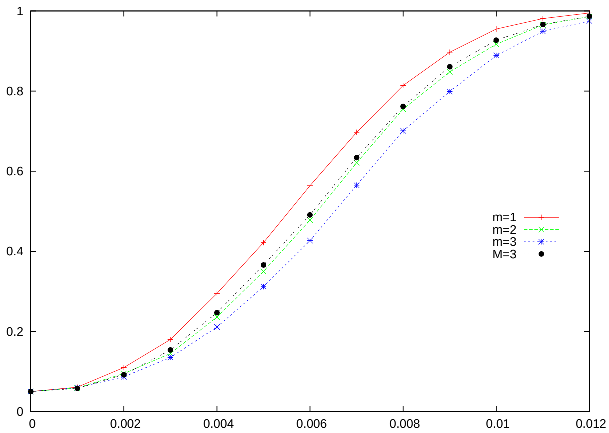
(ii-a) sup-type tests ($T = 120$)



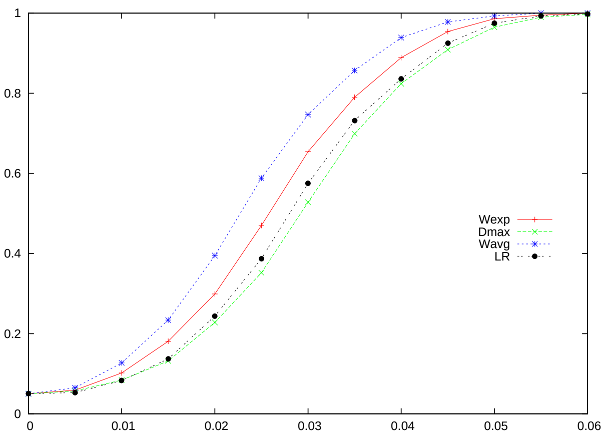
(ii-b) sup-type tests ($T = 300$)



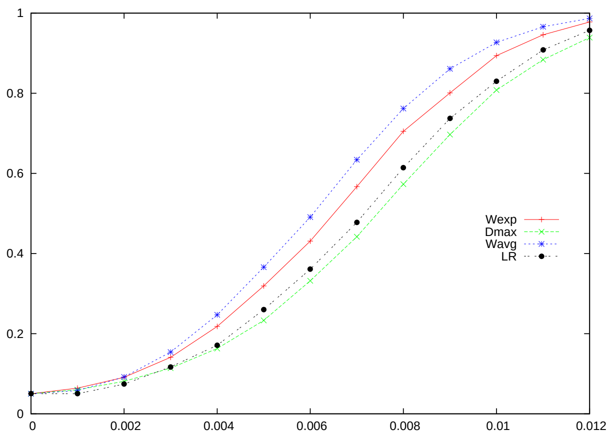
(iii-a) avg-type tests ($T = 120$)



(iii-b) avg-type tests ($T = 300$)

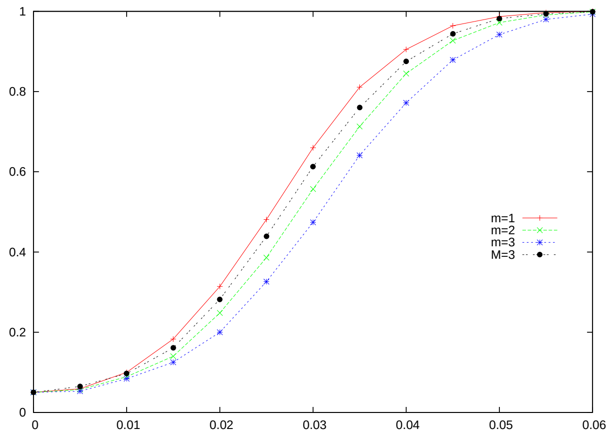


(iv-a) weighted-type tests ($T = 120$)

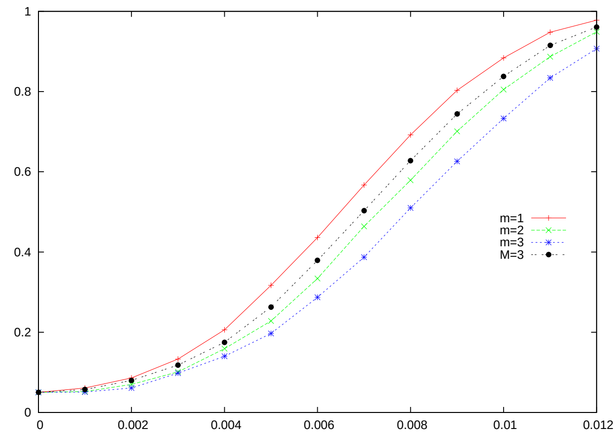


(iv-b) weighted-type tests ($T = 300$)

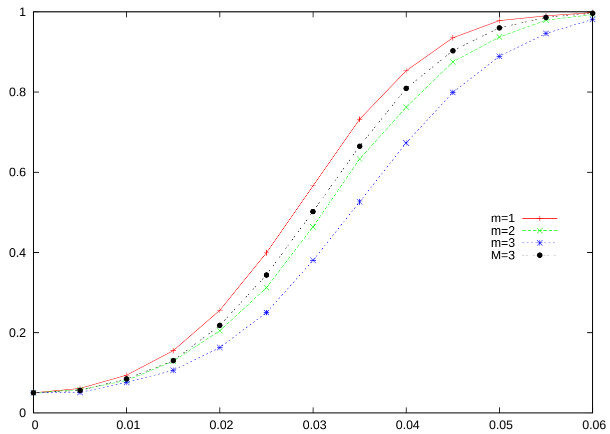
Figure 1: Size adjusted power ($q = 0$, $m = 1$)



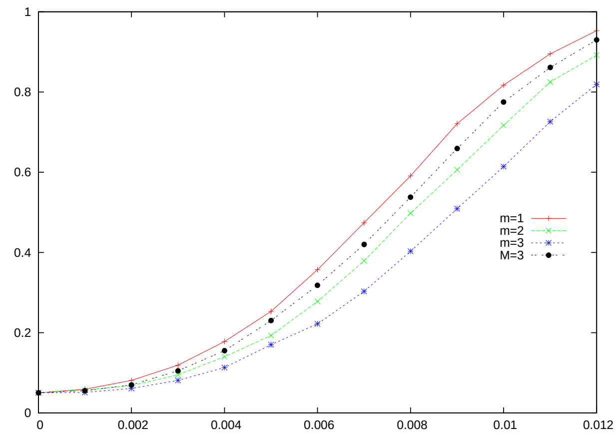
(i-a) exp-type tests ($T = 120$)



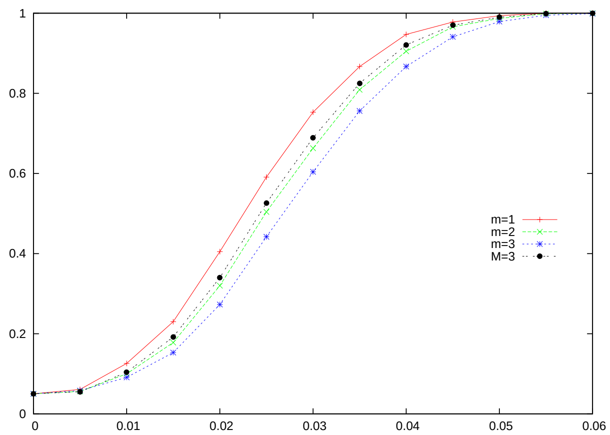
(i-b) exp-type tests ($T = 300$)



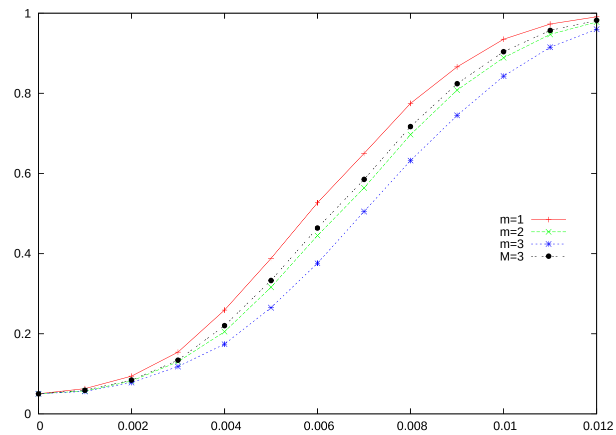
(ii-a) sup-type tests ($T = 120$)



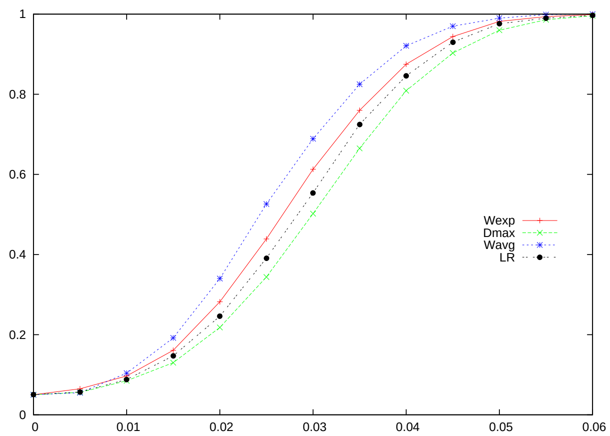
(ii-b) sup-type tests ($T = 300$)



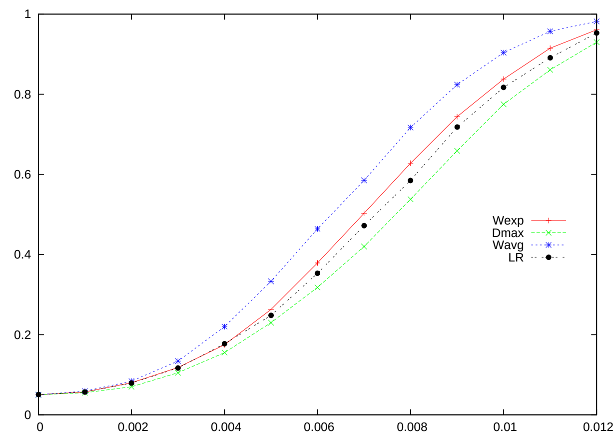
(iii-a) avg-type tests ($T = 120$)



(iii-b) avg-type tests ($T = 300$)

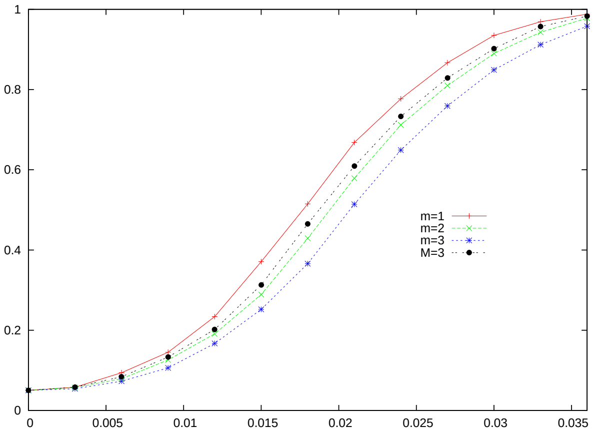


(iv-a) weighted-type tests ($T = 120$)

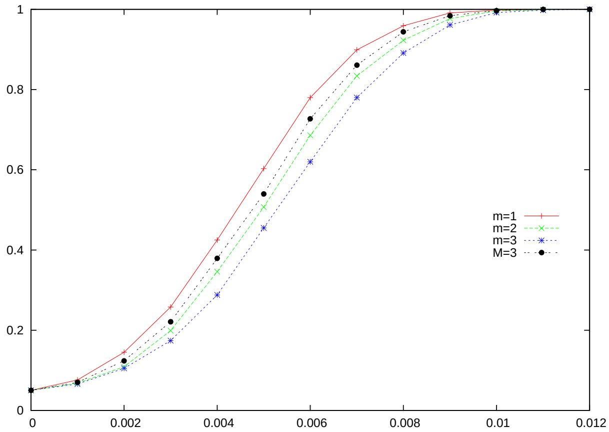


(iv-b) weighted-type tests ($T = 300$)

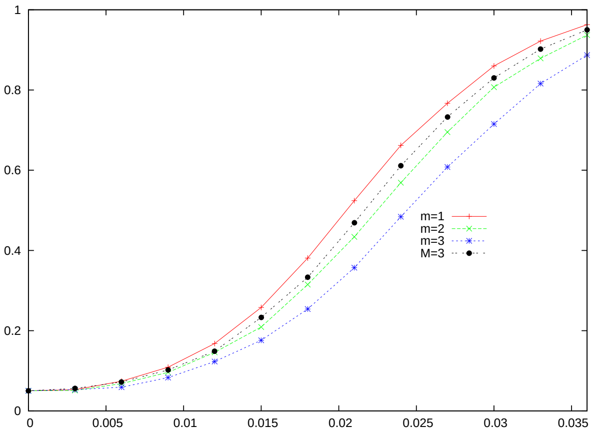
Figure 2: Size adjusted power ($q = 1$, $m = 1$, $\phi = 0$)



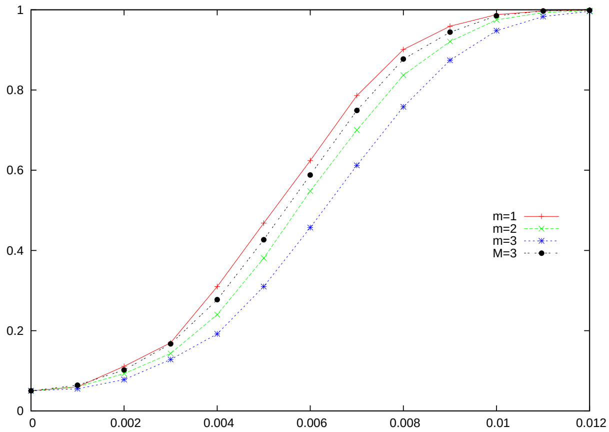
(i-a) exp-type tests ($T = 120$)



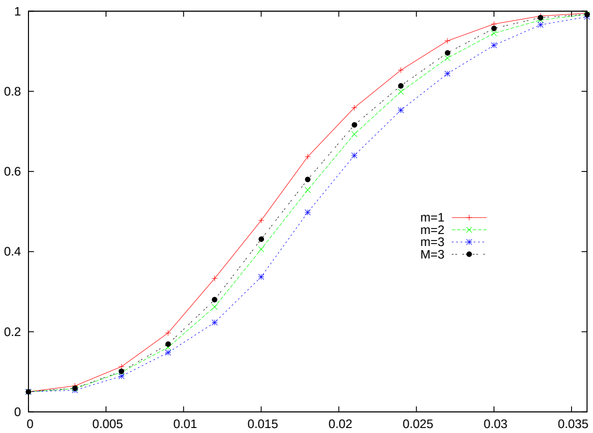
(i-b) exp-type tests ($T = 300$)



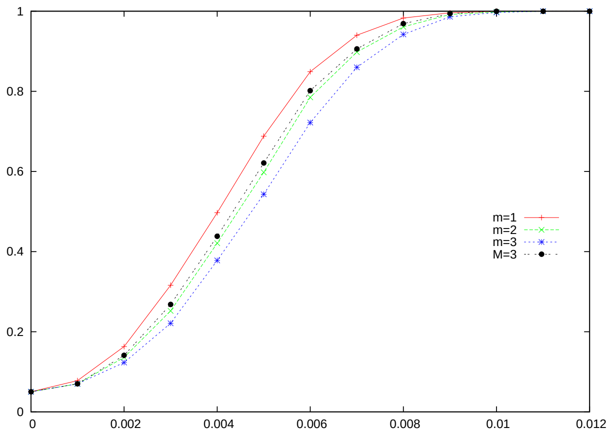
(ii-a) sup-type tests ($T = 120$)



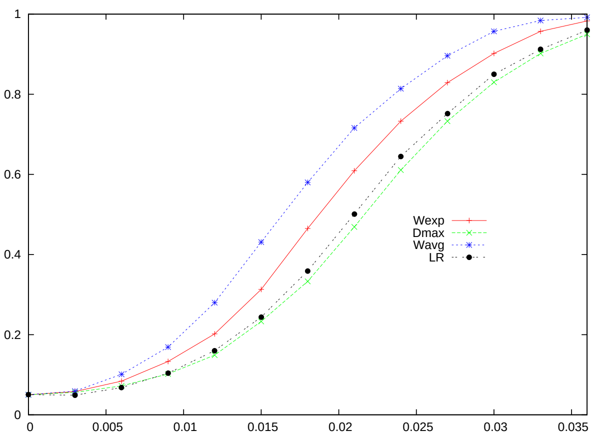
(ii-b) sup-type tests ($T = 300$)



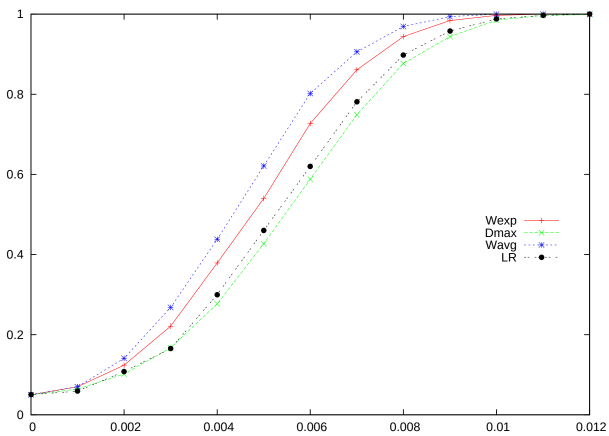
(iii-a) avg-type tests ($T = 120$)



(iii-b) avg-type tests ($T = 300$)

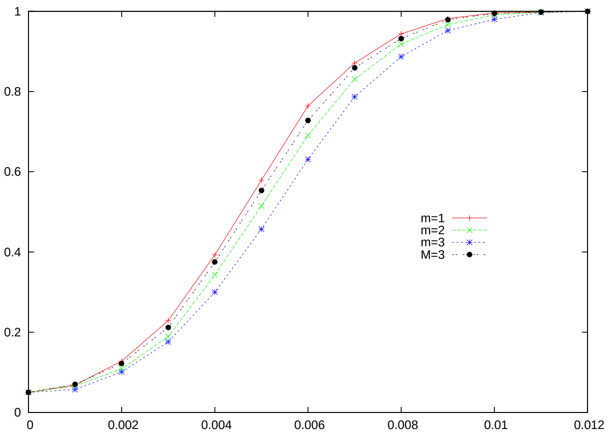


(iv-a) weighted-type tests ($T = 120$)

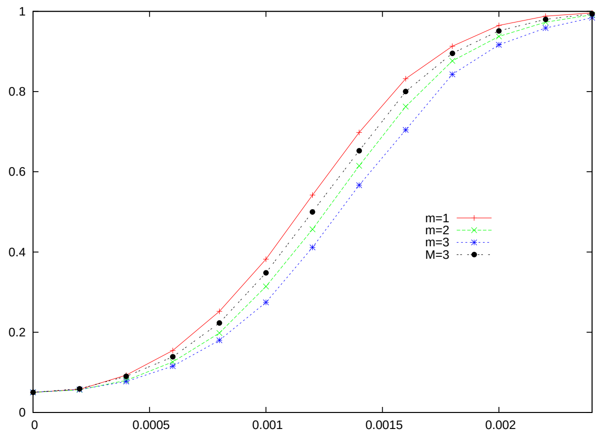


(iv-b) weighted-type tests ($T = 300$)

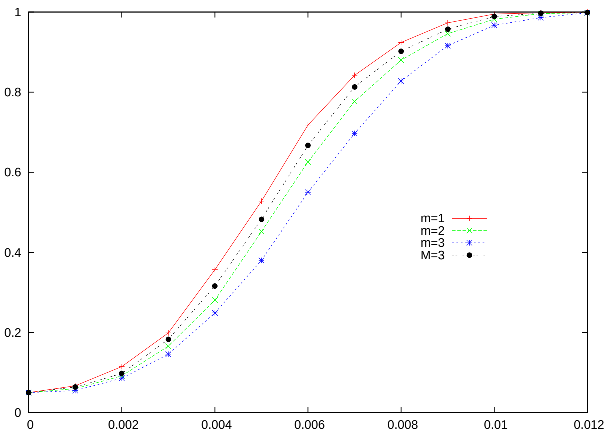
Figure 3: Size adjusted power ($q = 0$, $m = 2$, case a)



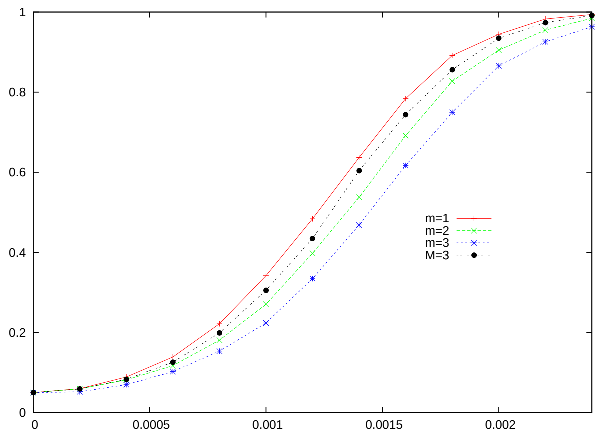
(i-a) exp-type tests ($T = 120$)



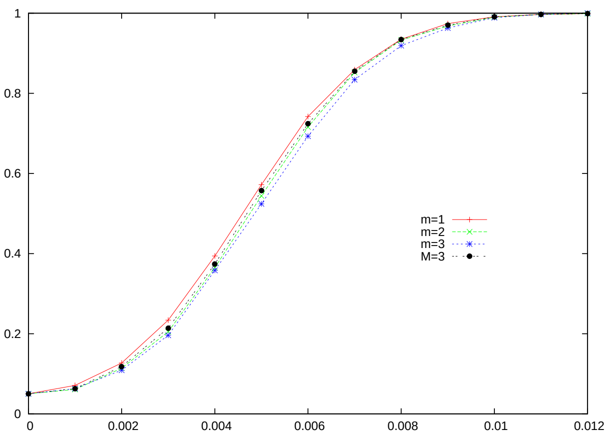
(i-b) exp-type tests ($T = 300$)



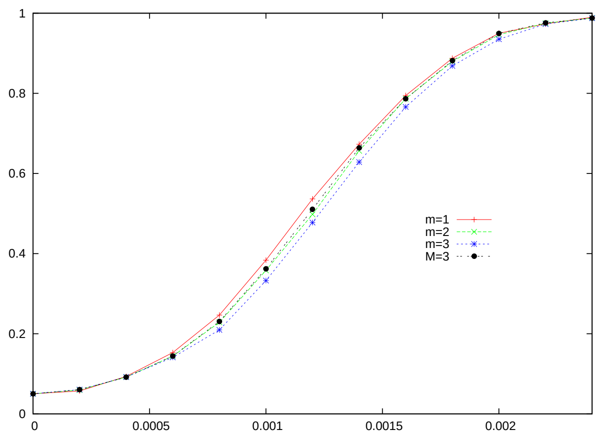
(ii-a) sup-type tests ($T = 120$)



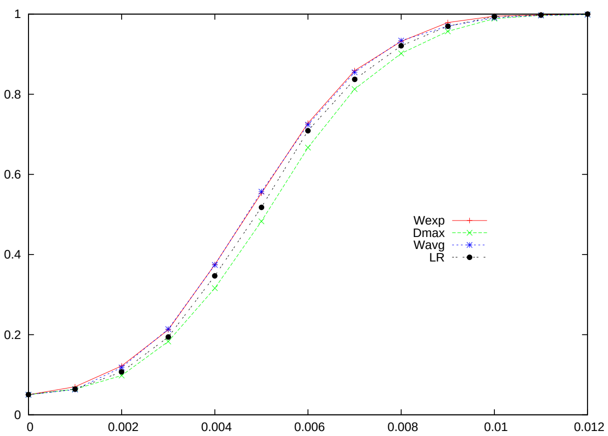
(ii-b) sup-type tests ($T = 300$)



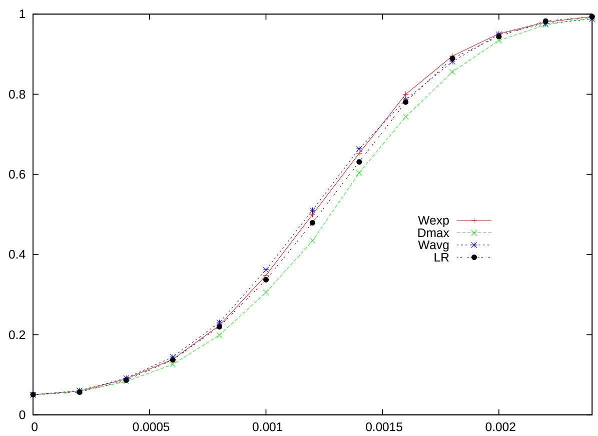
(iii-a) avg-type tests ($T = 120$)



(iii-b) avg-type tests ($T = 300$)

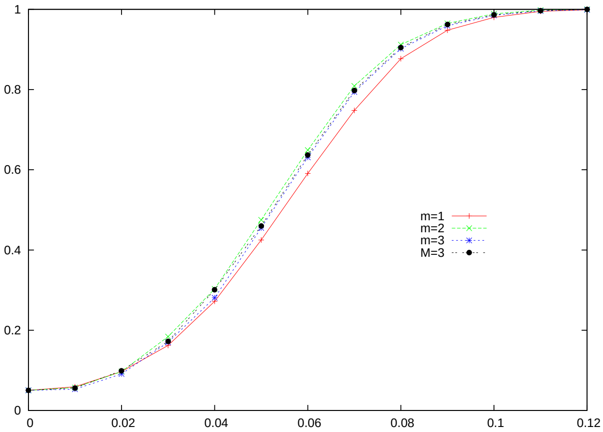


(iv-a) weighted-type tests ($T = 120$)

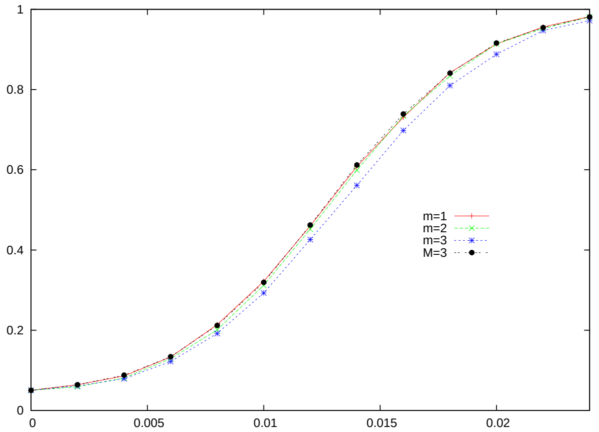


(iv-b) weighted-type tests ($T = 300$)

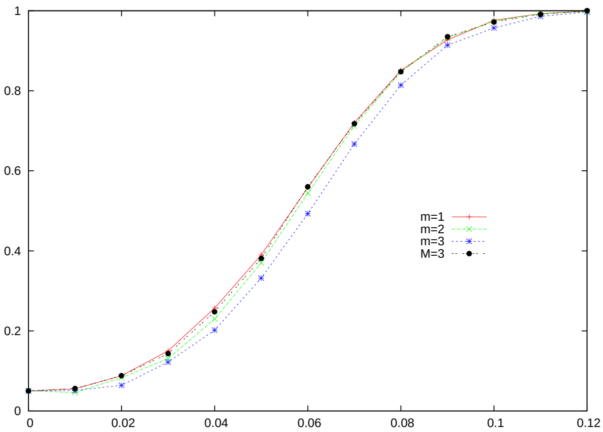
Figure 4: Size adjusted power ($q = 1$, $m = 2$, $\phi = 0$, case a)



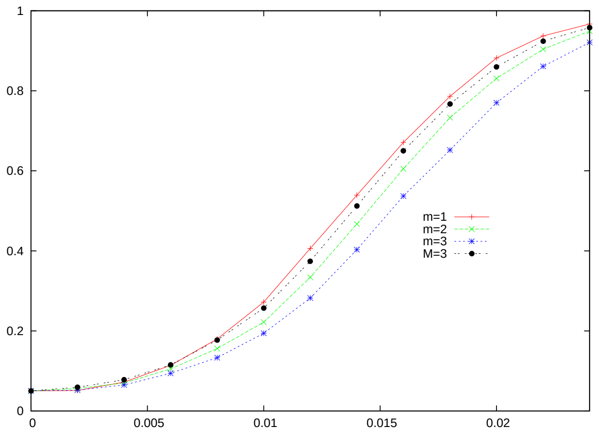
(i-a) exp-type tests ($T = 120$)



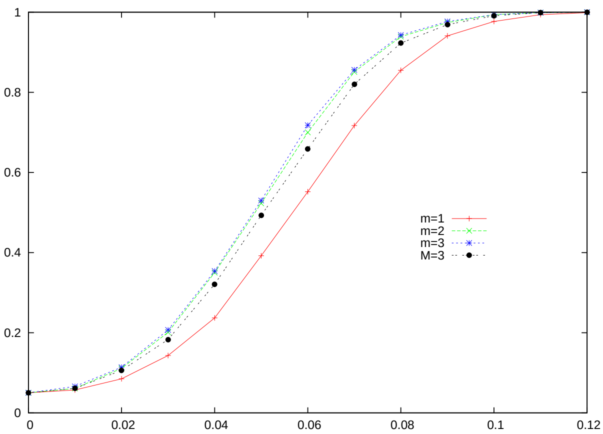
(i-b) exp-type tests ($T = 300$)



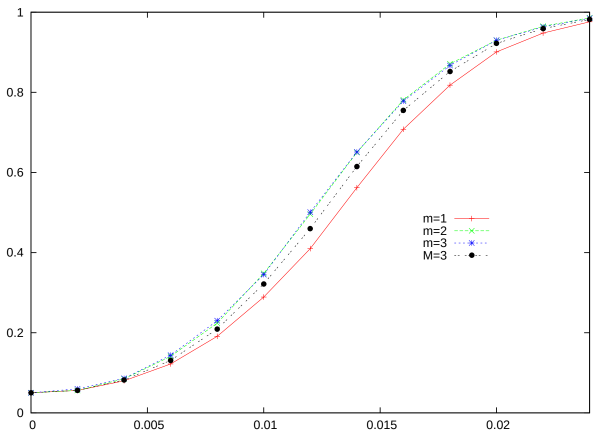
(ii-a) sup-type tests ($T = 120$)



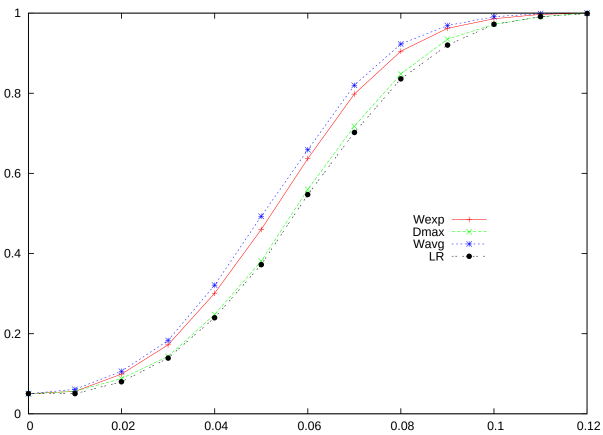
(ii-b) sup-type tests ($T = 300$)



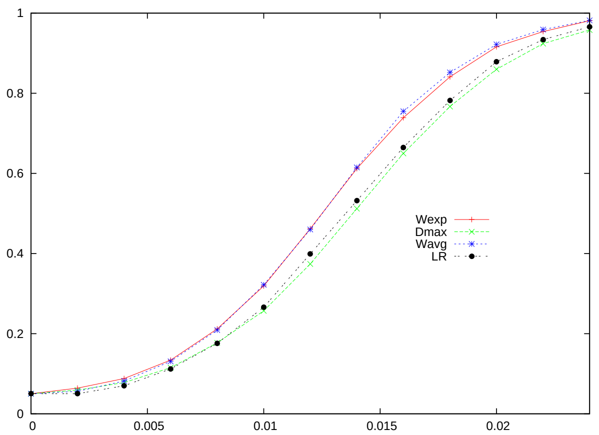
(iii-a) avg-type tests ($T = 120$)



(iii-b) avg-type tests ($T = 300$)

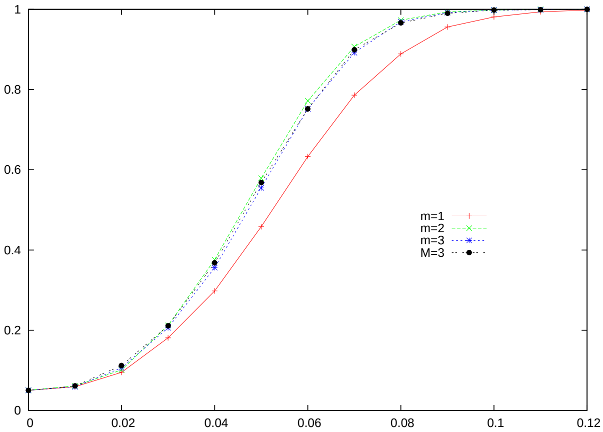


(iv-a) weighted-type tests ($T = 120$)

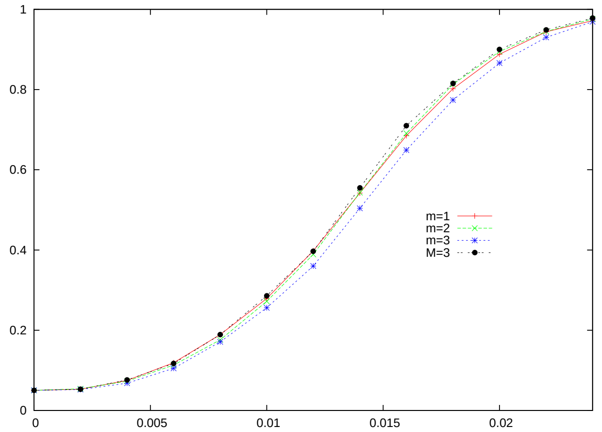


(iv-b) weighted-type tests ($T = 300$)

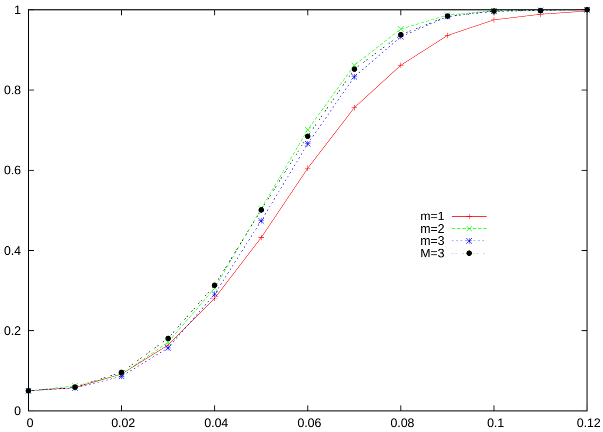
Figure 5: Size adjusted power ($q = 0$, $m = 2$, case b)



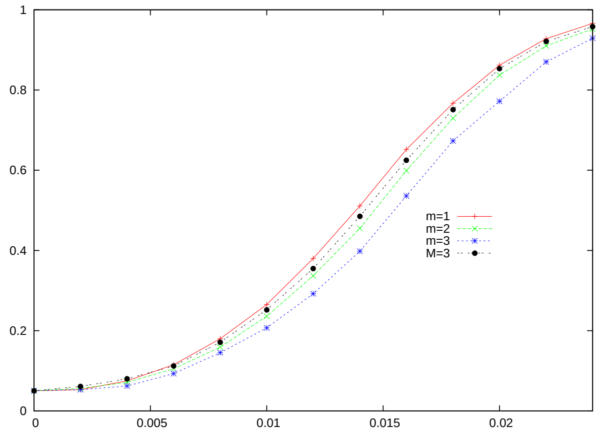
(i-a) exp-type tests ($T = 120$)



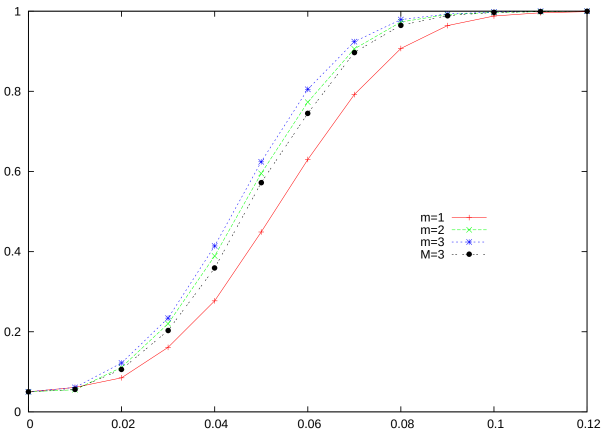
(i-b) exp-type tests ($T = 300$)



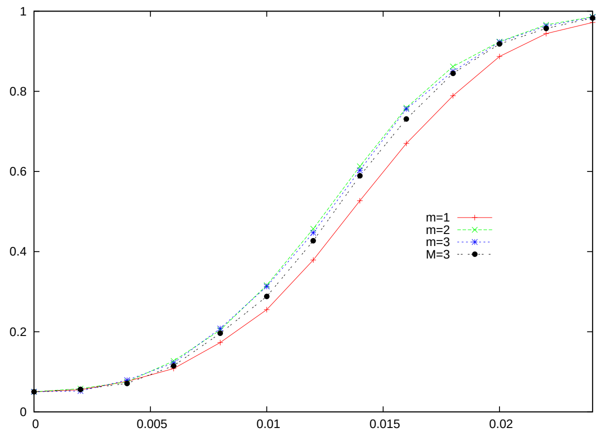
(ii-a) sup-type tests ($T = 120$)



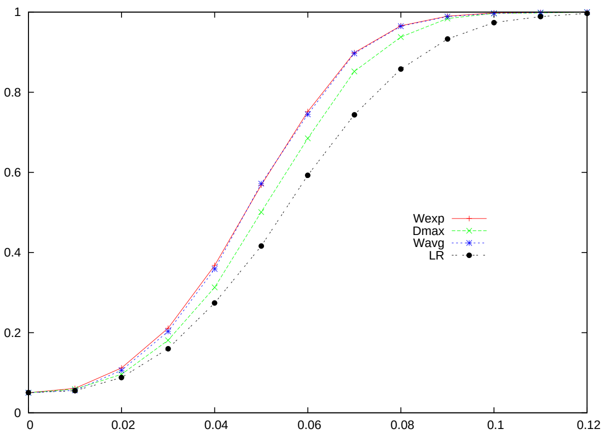
(ii-b) sup-type tests ($T = 300$)



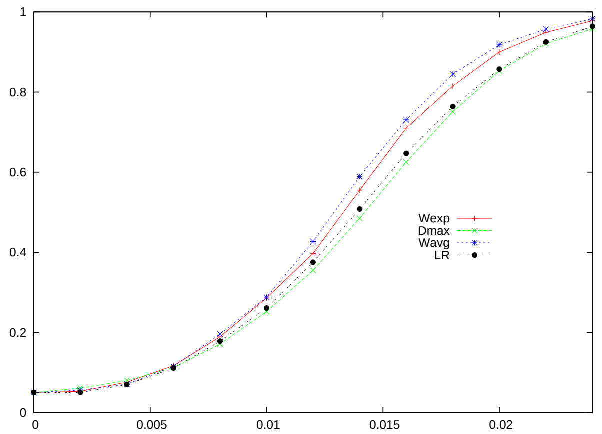
(iii-a) avg-type tests ($T = 120$)



(iii-b) avg-type tests ($T = 300$)



(iv-a) weighted-type tests ($T = 120$)



(iv-b) weighted-type tests ($T = 300$)

Figure 6: Size adjusted power ($q = 1$, $m = 2$, $\phi = 0$, case b)