Double Bootstrap Test for a Structural Break when the Disturbance Variance Changes with the Break

Namba, Akio

(Citation)
Kobe University Economic Review,60:33-43

(Issue Date)
2014

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81008767

(URL)
https://hdl.handle.net/20.500.14094/81008767

DOUBLE BOOTSTRAP TEST FOR A STRUCTURAL BREAK WHEN THE DISTURBANCE VARIANCE CHANGES WITH THE BREAK*[1)]

By AKIO NAMBA

In this paper we consider the Wald test statistic proposed by Watt (1979) for testing equality between the sets of regression coefficients in two linear regression models when the disturbance variances may possibly be unequal. This test can be also used as a test for a structural break. As shown by Ohtani and Toyoda (1985) and Honda and Ohtani (1986), the test based on the Wald test statistic suffers from severe size distortion in small samples when the disturbance variances of the two regression models are unequal. To tackle this problem, we apply two kinds of bootstrap methods, i.e., the usual bootstrap and the double bootstrap. Through simulation studies, we show that the size distortion is substantially improved when the bootstrap methods are utilized.

1. Introduction

The test proposed by Chow (1960) has been widely used to test equality between sets of coefficients in two linear regression models, or to test the existence of a structural break in a regression model. However, it is well known that the Chow test suffers from poor performance if the regression model is heteroscedastic, or the disturbance variances of the two linear regression models are unequal [See Toyoda (1974) and Schmidt and Sickles (1977)].

In order to tackle this drawback of the Chow test, several authors proposed alternative testing procedures which are applicable when the disturbance terms are heteroscedastic. Some examples are Watt (1979), Jayatissa (1977), and Rothenberg (1984).[2)] In particular, Watt (1979) proposed the test based on the Wald test statistic. According to Ohtani and Kobayashi (1986) and Thursby (1992), Watt's (1979) test is more powerful than Jayatissa's (1977) test.

Though the test statistic proposed by Watt (1979) is easy to compute, its exact distribution is very complex [See Kobayashi (1986) and Phillips (1986)]. Thus, Watt (1979) proposed to use critical values of a chi-squared distribution based on its asymptotic distribution. However, if we use the critical values of a chi-squared distribution, the test proposed by Watt (1979) suffers from size distortion when the sample size is small. See, e.g., Ohtani and Toyoda (1985) and Honda and Ohtani (1986). In order to avoid this size distortion, Ohtani and Kobayashi (1986) and Kobayashi (1986) proposed the bounds test based on the Wald test statistic. Since their test is based on the upper and lower bounds of the Wald test statistic proposed by Watt (1979), there inherently exists an inconclusive region for the test statistic. Weerahandi (1987) also proposed a test which is exact under the normality assumption of disturbances. Though Weerahandi's (1987) test is exact, it requires a numerical integration when calculating the p-value of the test statistic

1) This work was supported by JSPS KAKENHI Grant Numbers 23243038, 26780136.
2) See also Thursby (1992) for comparisons of several testing methods and their performances.

and thus is not easy to implement.

The above procedures are required since the exact distribution of the Wald test statistic proposed by Watt (1979) is complex when the sample size is small. When the exact distribution of a statistic is complex or unknown, the bootstrap method proposed by Efron (1979) is sometimes useful. Bootstrap is one of the so-called resampling methods where the artificial data are obtained by resampling from the original data. In particular, as shown by Beran (1987, 1988) [see also Hall (1992)], the procedure based on the bootstrap methods yields more accurate results than the conventional asymptotic procedure when the statistic is asymptotically pivotal, i.e., the asymptotic distribution of the statistic does not depend on unknown parameters. Since the Wald test statistic is asymptotically distributed as a chi-squared distribution with known degrees of freedom, it is asymptotically pivotal. Therefore, an improvement is expected if the bootstrap method is applied to the Wald test statistic proposed by Watt (1979).

Beran (1988) also proposed a bootstrap method where another step of bootstrap resampling is executed from the bootstrap sample which is obtained by the usual bootstrap. This is called the double bootstrap method. As discussed in Beran (1988), the performance of the usual bootstrap method can be further improved by the double bootstrap. Several authors investigated the performance of the double bootstrap method. See, e.g., Caers (1998), DiCiccio et al. (1992), Hinkley and Shi (1989), McKnight et al. (2000), Letson and McCullough (1998), McCullough and Vinod (1998), Nankervis (2005), Vinod (1995), and Vinod and McCullough (1995).

Recent literature includes models which permit multiple structural breaks and unknown break points and the methods to investigate them. See, e.g., Bai and Perron (1998), Perron (2006), Boldea et al. (2012), Hall et al. (2012), Perron and Yamamoto (2014) and references therein. However, to examine the validity of the proposed methods and for simplicity, we focus on the model with one possible structural break and a known break point. Thus, in this paper, we apply the usual bootstrap and the double bootstrap methods to the test statistic proposed by Watt (1979). We examine the sizes and the powers of the bootstrap tests by Monte Carlo simulations. The organization of the paper is as follows. In the next section, we introduce the model and the test statistic. Also, the ways to apply the bootstrap methods to the test statistic are explained. It turns out that the bootstrap procedure gets simplified because of the structure of the test statistic. In section 3, we examine the performance of the bootstrap tests by simulations. The simulation results show that the size distortion of Watt's (1979) test is substantially improved by the bootstrap methods. Finally, some concluding remarks are given in section 4.

2. Model, test statistic and the bootstrap methods

Consider two linear regression models

$$y_i = X_i \beta_i + \epsilon_i, \quad i = 1, 2, \tag{1}$$

where y_i is an $n_i \times 1$ vector of observations on a dependent variable, X_i is an $n_i \times k$ matrix of observations on nonstochastic explanatory variables, β_i is a $k \times 1$ vector of coefficients, and ϵ_i is an $n_i \times 1$ vector of error terms and $\epsilon_i \sim \mathrm{N}(0, \sigma_i^2 I_{n_i})$. Also, we assume that X_i is of full column rank.

The task considered in this paper is to test the null hypothesis $H_0 : \beta_1 = \beta_2$ against the alternative $H_1 : \beta_1 \neq \beta_2$. If i denotes the regime, accepting H_0 implies that there is no structural break.

If $\sigma_1^2 = \sigma_2^2$, i.e., the disturbance variances of the two regression models are equal, we can easily test the null hypothesis using the Chow test proposed by Chow (1960). However, as shown by Toyoda (1974) and Schmidt and Sickles (1977), the Chow test has a very poor performance when $\sigma_1^2 \neq \sigma_2^2$. Thus, Watt (1979) proposed the Wald test statistic which takes the heteroscedasticity into consideration:

$$W = (b_1 - b_2)' \left[s_1^2 (X_1' X_1)^{-1} + s_2^2 (X_2' X_2)^{-1} \right]^{-1} (b_1 - b_2), \tag{2}$$

where b_i and s_i^2 are the least squares estimator of β_i and σ_i^2. Though this Wald test statistic is asymptotically valid, as shown by Ohtani and Toyoda (1985) and Honda and Ohtani (1986), the test based on this statistic suffers from severe size distortion in small samples if the critical values of a chi-squared distribution are used. One way of coping with this size distortion is executing the test based on the upper and lower bounds of the Wald test statistic as proposed by Ohtani and Kobayashi (1986) and Kobayashi (1986). However, this testing procedure inherently includes the inconclusive region. Weerahandi (1987) also proposed a test which is exact under normality of the disturbance. However, Weerahandi's (1987) test requires a numerical integration when calculating the p-value of the test statistic and is not easy to implement.

Thus, in this paper, we consider more direct methods, i.e., the bootstrap method proposed by Efron (1979) and the double bootstrap method proposed by Beran (1988). As shown by Beran (1987, 1988), inferences based on asymptotic distributions can be improved by applying the bootstrap method if the statistic considered is asymptotically pivotal, i.e., the asymptotic distribution of the statistic does not depend on unknown parameters. Since the asymptotic distribution of the Wald test statistic given in (2) is a chi-squared distribution with k degrees of freedom, it is asymptotically pivotal. Thus, by applying the bootstrap method to W, a reduction in the size distortion of the test is expected. Beran (1988) also discussed that the performance of the usual bootstrap can be improved by the double bootstrap method. In the following subsections, we will explain the ways to apply the usual and the double bootstrap methods to the Wald statistic given in (2).

2.1 Usual bootstrap method

The application of the usual bootstrap method to W is summarized as follows:

A1. Estimate β_i and σ_i^2 by the ordinary least squares (OLS) method and obtain b_i and s_i^2. Calculate the value of the Wald test statistic W given in (2).

A2. Let $e_i = y_i - X_i b_i$ be the residual vector for $i = 1, 2$. Following Wu (1986), we first rescale the residual vector as $\sqrt{n_i/(n_i - k)}\, e_i$. Drawing a sample of size n_i from the elements of the rescaled residual with replacement and stacking them as an $n_i \times 1$ vector, we obtain a bootstrap sample vector e_i^* for $i = 1, 2$.

A3. Regressing e_i^* on X_i, obtain bootstrap estimates b_i^* and s_i^{2*} for $i = 1, 2$. Using these estimates, calculate the bootstrap version of the Wald test statistic:

$$W_b = (b_1^* - b_2^*)' \left[s_1^{2*}(X_1'X_1)^{-1} + s_2^{2*}(X_2'X_2)^{-1} \right]^{-1} (b_1^* - b_2^*). \tag{3}$$

A4. Repeating the steps 2 and 3 above B times (i.e., $b = 1, 2,...,B$), and calculating the ratio such that W_b exceeds W as

$$p_b = \frac{1}{B} \sum_{i=1}^{B} I(W_b > W), \tag{4}$$

where $I(A)$ is an indicator function such that $I(A) = 1$ when the event A occurs and $I(A) = 0$ otherwise, we obtain the p-value of the test based on the bootstrap method. Thus, if the obtained p-value is less than the assigned significance level α , $H_0 : \beta_1 = \beta_2$ is rejected.

Note that, in step A2 above, we simply regress e_i^* on X_i in order to obtain bootstrap estimates. In the usual bootstrap procedure for a regression model, we usually calculate a bootstrap sample of the dependent variable $y_i^* = X_i\bar{\beta}_i + e_i^*$where $\bar{\beta}_i$ is any estimator of β_i, and obtain bootstrap estimates by regressing y_i^* on X_i. Since, when testing a null hypothesis, a bootstrap sample must be drawn from a model such that the null hypothesis is hold, we need to use an estimator which satisfies $\bar{\beta}_1 = \bar{\beta}_2 = \bar{\beta}$. However, if we regress y_i^* instead of e_i^* under the condition $\bar{\beta}_1 = \bar{\beta}_2 = \bar{\beta}$, we obtain

$$\begin{aligned} b_1^* - b_2^* &= (X_1'X_1)^{-1}X_1'y_1^* - (X_2'X_2)^{-1}X_2'y_2^* \\ &= (X_1'X_1)^{-1}X_1'e_1^* - (X_2'X_2)^{-1}X_2'e_2^*. \end{aligned} \tag{5}$$

This implies that the bootstrap version of the Wald test statistic W_b given in (3) is independent of the choice of $\bar{\beta}$ and that the value of W_b is unchanged whatever estimate $\bar{\beta}$ may be used. Thus, by using the zero vector as $\bar{\beta}$, we can simply regress e_i^* on X_i, and the bootstrap resampling gets simplified because of the structure of the Wald test statistic W.

2.2 Double bootstrap method

The double bootstrap method is the procedure which executes another step of resampling from the bootstrap sample, i.e., the artificial sample obtained by the usual bootstrap resampling. The double bootstrap method is applied to W in the following way.

B1. Estimate β_i and σ_i^2 by the ordinary least squares (OLS) method and obtain b_i and s_i^2 . Calculate the value of the Wald test statistic W.

B2. In a similar way to the steps A2–A3 in the usual bootstrap, obtain a bootstrap sample e_i^* and calculate the bootstrap version of the Wald test statistic W_b.

B3. Let $e_i^\dagger$ be the residual vector obtained by regressing e_i^* on X_i for $i = 1, 2$. As another step of bootstrap, in the similar manner to step A2 above, drawing a sample of size n_i from the elements of $e_i^\dagger$ with replacement,[3)] we obtain a double bootstrap sample vector

3) In this step, we do not rescale the residual since the unconditional expectation of $e_i^\ddagger$ is σ^2 because e_i^* is rescaled.

$e_i^{\ddagger}$. Regressing $e_i^{\ddagger}$ on X_i, and calculating in the similar way to (3), we obtain the double bootstrap version of the Wald statistic W_d .

B4. Repeating the step B3 D times (i.e., d = 1, 2, ...,D) and calculating the ratio such that W_d is greater than or equal to W_b, we can obtain an estimate of the p-value of W_b as

$$\tilde{p}_b = \frac{\sum_{d=1}^{D} I(W_d \geqq W_b)}{D}. \tag{6}$$

B5. Repeating steps B2–B4 B times and calculate the p-value obtained by the double bootstrap as

$$p_D = \frac{\sum_{i=1}^{B} I(\tilde{p}_b < p_b)}{B}, \tag{7}$$

where p_b is defined in (4) above. If $p_D < \alpha$, H_0 is rejected at the 100×α percent significance level.

As we can see above steps, while the usual bootstrap requires B times of iteration, the double bootstrap requires $B \times D$ times of iteration.

3. Simulation results

To investigate the performance of the above mentioned methods, we execute some Monte Carlo simulations in this section. The design of the simulation is as follows:

1. For simplicity, we assume $k = 2$ and $x_{ij} = [1, u_j]$, where x_{ij} is the jth element of X_i and u_j is a random sample drawn from U $[0, 1]$. Thus, the regression model has an intercept and one explanatory variable.
2. Using the number of iteration of resampling in bootstrap $B = 500$ and $D = 300$, and letting $\beta_{11} = \beta_{12} = \sigma_1 = 1$, and $\beta_{21}, \beta_{22}, \sigma_2, n_1, n_2 =$ several values, where β_{ij} is the jth element of β_i, we iterated the procedure explained in the previous section $M =$ 10000 times and test the null hypothesis at $\alpha = 0.10$ (10%), 0.05 (5%) and 0.01 (1%) significance levels. Calculating the ratio when the null hypothesis is rejected out of $M =$ 10000 times, we obtain the empirical power of the test.

Through our simulations, we found that the size distortion of the Wald test which utilized critical values of a chi-squared distribution (i.e., asymptotic test) is severe when the differences between σ_1, n_1 and σ_2, n_2 are large. This coincides with the result in Ohtani and Toyoda (1985). Though we do not show all the results, the results shown here are typical ones obtained by our experiments.

Tables 1 and 2 show the empirical sizes of the tests for $(n_1, n_2) = (10, 50)$ and $(n_1, n_2) = (50, 50)$ obtained through simulations (i.e., $\beta_{21} = \beta_{22} = 1$). To evaluate the correctness of the tests, we test the null hypothesis that the significance level is α by means of the normal approximation of a binomial distribution. $*$, † and ‡ denote that the null hypothesis is rejected at the 10%, 5% and 1% significance levels, respectively. As we can see from Table 1, the size of the test based on the asymptotic distribution of W is not correct at all. The empirical size is far from the nominal size for all values of σ_2 in this case. By the usual bootstrap method, the correctness of the test is much improved. Though the null hypothesis that the significance level is α is rejected in some cases, the empirical size of the usual bootstrap test is much closer to the nominal size than the asymptotic test in all cases. The double bootstrap method yields a more accurate empirical size than the usual bootstrap in some cases, though there are a few cases where the usual bootstrap method has preferable results to the double bootstrap method. From Table 2, we can see the results for the case of $(n_1, n_2) = (50, 50)$. As the difference between n_1 and n_2 gets smaller, the results based on asymptotic distribution of W gets closer to the nominal levels. As mentioned above, this result coincides with the one in Ohtani and Toyoda (1985). However, the empirical size of the asymptotic test is not close enough to the nominal level and the null hypothesis for α is rejected for most of cases considered here. Also, in this case, the usual and double bootstrap methods yield comparable performance. This implies that the improvement over the usual bootstrap by the double bootstrap is hard to obtain under the situation when both n_1 and n_2 are not small, and the usual bootstrap method has already corrected the empirical size to some extent. Though we do not show all the results here, as both n_1 and n_2 get larger, the asymptotic test yields a more accurate empirical size, however, the bootstrap methods generally provide with the better performance than the asymptotic test.

Figures 1 and 2 show the power of the tests for $(n_1, n_2) = (15, 15), (50, 50)$, $\sigma_2 = 2.0$, $\beta_{22} = 1$ and β_{21} = various values, which means that the intercept and the disturbance variance change with the structural break. We can see that the test based on the asymptotic distribution of W has a slightly larger power than the bootstrap tests. However, this is caused by the size distortion of the asymptotic test which can be seen in Tables 1 and 2. Also, the usual bootstrap and double bootstrap tests have almost same powers. Thus, we may assert that the powers of the bootstrap tests are almost comparable to the power of the asymptotic test.

Table 1 Empirical sizes for n_1 = 10 and n_2 = 50.

	Asymptotic Test			Usual Bootstrap			Double Bootstrap		
σ_2	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.1	0:1690‡	0.1116‡	0.0466‡	0.0946*	0.0429‡	0.0085	0.0973	0.0461*	0.0097
0.2	0:1592‡	0.1049‡	0.0444‡	0.0936†	0.0441‡	0.0080†	0.0950*	0.0442‡	0.0068‡
0.3	0:1592‡	0.1057‡	0.0469‡	0.0959	0.0497	0.0115	0.0923†	0.0460*	0.0094
0.4	0:1609‡	0.1066‡	0.0438‡	0.1005	0.0514	0.0117*	0.0971	0.0474	0.0089
0.5	0.1597‡	0.1023‡	0.0419‡	0.0988	0.0492	0.0140	0.0953	0.0442‡	0.0108
0.6	0.1486‡	0.0932‡	0.0377‡	0.0933†	0.0449†	0.0124†	0.0888‡	0.0422‡	0.0099
0.7	0.1494‡	0.0967‡	0.0387‡	0.0992	0.0518	0.0114	0.0938†	0.0465	0.0090
0.8	0.1465‡	0.0945‡	0.0385‡	0.0996	0.0514	0.0149‡	0.0946*	0.0475	0.0128‡
0.9	0.1533‡	0.0974‡	0.0386‡	0.1056*	0.0549†	0.0146‡	0.1019	0.0504	0.0118*
1.0	0.1470‡	0.0911‡	0.0339‡	0.1000	0.0500	0.0117*	0.0944*	0.0460*	0.0102
1.1	0.1426‡	0.0871‡	0.0348‡	0.0987	0.0528	0.0142‡	0.0969	0.0490	0.0130‡
1.2	0.1406‡	0.0922‡	0.0345‡	0.1037	0.0546†	0.0142‡	0.1011	0.0507	0.0120†
1.3	0.1358‡	0.0834‡	0.0310‡	0.0974	0.0478	0.0115	0.0938†	0.0450†	0.0105
1.4	0.1415‡	0.0861‡	0.0298‡	0.1011	0.0532	0.0132‡	0.0978	0.0493	0.0122†
1.5	0.1370‡	0.0858‡	0.0317‡	0.1009	0.0539*	0.0137‡	0.0987	0.0507	0.0117*
1.6	0.1380‡	0.0864‡	0.0298‡	0.1029	0.0527	0.0129‡	0.1002	0.0499	0.0114
1.7	0.1375‡	0.0799‡	0.0267‡	0.0986	0.0487	0.0119*	0.0959	0.0471	0.0101
1.8	0.1377‡	0.0818‡	0.0285‡	0.1025	0.0504	0.0131‡	0.0998	0.0471	0.0114
1.9	0.1390‡	0.0826‡	0.0297‡	0.1038	0.0524	0.0135‡	0.1002	0.0504	0.0115
2.0	0.1365‡	0.0838‡	0.0291‡	0.1022	0.0523	0.0130‡	0.1007	0.0509	0.0109
2.1	0.1407‡	0.0880‡	0.0314‡	0.1093‡	0.0569‡	0.0132‡	0.1068†	0.0532	0.0124†
2.2	0.1363‡	0.0795‡	0.0257‡	0.1010	0.0495	0.0117*	0.0992	0.0479	0.0107
2.3	0.1331‡	0.0803‡	0.0277‡	0.1014	0.0535	0.0120†	0.0990	0.0496	0.0114
2.4	0.1366‡	0.0809‡	0.0286‡	0.1009	0.0536*	0.0135‡	0.0984	0.0502	0.0123†
2.5	0.1342‡	0.0827‡	0.0286‡	0.1041	0.0551†	0.0148‡	0.1010	0.0519	0.0139‡
2.6	0.1306‡	0.0793‡	0.0284‡	0.0996	0.0555†	0.0133‡	0.0991	0.0546†	0.0120†
2.7	0.1337‡	0.0745‡	0.0267‡	0.0988	0.0508	0.0134‡	0.0977	0.0489	0.0105
2.8	0.1366‡	0.0813‡	0.0269‡	0.1058*	0.0530	0.0138‡	0.1028	0.0517	0.0135‡
2.9	0.1346‡	0.0786‡	0.0257‡	0.1022	0.0529	0.0138‡	0.1012	0.0492	0.0129‡
3.0	0.1274‡	0.0775‡	0.0269‡	0.0980	0.0535	0.0141‡	0.0945*	0.0511	0.0128‡
3.1	0.1337‡	0.0763‡	0.0259‡	0.1019	0.0529	0.0144‡	0.1002	0.0506	0.0126‡
3.2	0.1361‡	0.0803‡	0.0257‡	0.1068†	0.0533	0.0145‡	0.1038	0.0517	0.0139‡
3.3	0.1281‡	0.0773‡	0.0261‡	0.1018	0.0528	0.0144‡	0.0991	0.0503	0.0131‡
3.4	0.1325‡	0.0801‡	0.0241‡	0.1053*	0.0541*	0.0133‡	0.1008	0.0519	0.0115
3.5	0.1306‡	0.0726‡	0.0217‡	0.1019	0.0478	0.0114	0.0978	0.0460*	0.0104
3.6	0.1360‡	0.0794‡	0.0254‡	0.1076†	0.0550†	0.0136‡	0.1048	0.0531	0.0128‡
3.7	0.1307‡	0.0783‡	0.0258‡	0.1043	0.0553†	0.0134‡	0.1012	0.0536*	0.0119*
3.8	0.1246‡	0.0698‡	0.0225‡	0.0961	0.0504	0.0115	0.0948*	0.0476	0.0104
3.9	0.1302‡	0.0751‡	0.0228‡	0.1034	0.0537*	0.0128‡	0.1012	0.0517	0.0117*
4.0	0.1361‡	0.0791‡	0.0250‡	0.1084‡	0.0547†	0.0136‡	0.1072†	0.0545†	0.0133‡

Table 2 Empirical sizes for $n_1 = 50$ and $n_2 = 50$.

	Asymptotic Test			Usual Bootstrap			Double Bootstrap		
σ_2	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.1	0.1125‡	0.0592‡	0.0157‡	0.1033	0.0504	0.0130‡	0.1041	0.0512	0.0148‡
0.2	0.1078‡	0.0587‡	0.0132‡	0.0988	0.0508	0.0118*	0.1000	0.0520	0.0131‡
0.3	0.1135‡	0.0587‡	0.0149‡	0.1043	0.0522	0.0130‡	0.1045	0.0530	0.0128‡
0.4	0.1075†	0.0573‡	0.0134‡	0.1006	0.0503	0.0111	0.1005	0.0526	0.0124†
0.5	0.1031	0.0523	0.0112	0.0959	0.0475	0.0090	0.0957	0.0472	0.0093
0.6	0.1009	0.0537*	0.0114	0.0966	0.0495	0.0108	0.0978	0.0499	0.0109
0.7	0.1092‡	0.0584‡	0.0133‡	0.1016	0.0512	0.0122†	0.1010	0.0523	0.0130‡
0.8	0.1032	0.0545†	0.0122†	0.0979	0.0484	0.0107	0.0964	0.0490	0.0099
0.9	0.1142‡	0.0591‡	0.0130‡	0.1069†	0.0535	0.0118*	0.1065†	0.0537*	0.0122†
1.0	0.1115‡	0.0585‡	0.0140‡	0.102	0.0520	0.0128‡	0.1022	0.0525	0.0142‡
1.1	0.1114‡	0.0596‡	0.0159‡	0.1046	0.0548†	0.0135‡	0.1049	0.0554†	0.0143‡
1.2	0.1068†	0.0545†	0.0129‡	0.0992	0.0499	0.0120†	0.1005	0.0518	0.0124†
1.3	0.1032	0.0562‡	0.0146‡	0.0957	0.0492	0.0124†	0.0988	0.0502	0.0136‡
1.4	0.1083‡	0.0588‡	0.0156‡	0.1011	0.0524	0.0136‡	0.1027	0.0535	0.0139‡
1.5	0.1058*	0.0552†	0.0129‡	0.0999	0.0503	0.0122†	0.1003	0.0507	0.0122†
1.6	0.1067†	0.0559‡	0.0123†	0.1002	0.0504	0.0105	0.1018	0.0500	0.0119*
1.7	0.1099	0.0573‡	0.0129‡	0.1008	0.0477	0.0113	0.1007	0.0476	0.0117*
1.8	0.1087‡	0.0568‡	0.0156‡	0.0988	0.0519	0.0133‡	0.1009	0.0524	0.0132‡
1.9	0.1073†	0.0583‡	0.0120†	0.0968	0.0520	0.0122†	0.0968	0.0525	0.0132‡
2.0	0.1134‡	0.0586‡	0.0122†	0.1034	0.0514	0.0109	0.1053*	0.0520	0.0114
2.1	0.1042	0.0582‡	0.0113	0.0986	0.0526	0.0100	0.0991	0.0515	0.0106
2.2	0.1111‡	0.0585‡	0.0147‡	0.1018	0.0520	0.0121†	0.1023	0.0529	0.0132‡
2.3	0.1043	0.0546†	0.0120†	0.0962	0.0474	0.0107	0.0950*	0.0481	0.0117*
2.4	0.1069†	0.0580‡	0.0146‡	0.0992	0.0510	0.0121†	0.1003	0.0515	0.0133‡
2.5	0.1084‡	0.0579‡	0.0127‡	0.1023	0.0513	0.0110	0.1044	0.0535	0.0119*
2.6	0.1158‡	0.0633‡	0.0160‡	0.1063†	0.0577‡	0.0148‡	0.1069†	0.0578‡	0.0148‡
2.7	0.1107‡	0.0590‡	0.0139‡	0.101	0.0537*	0.0116	0.1025	0.0518	0.0137‡
2.8	0.1070†	0.0549†	0.0109	0.0996	0.0483	0.0108	0.1015	0.0488	0.0116
2.9	0.1068†	0.0576‡	0.0124†	0.1008	0.0507	0.0115	0.1020	0.0525	0.0117*
3.0	0.1117‡	0.0583‡	0.0128‡	0.103	0.0519	0.0117*	0.1049	0.0517	0.0122†
3.1	0.1092‡	0.0571‡	0.0124	0.1016	0.0502	0.0097	0.1009	0.0517	0.0110
3.2	0.1070†	0.0527	0.0127‡	0.1002	0.0468	0.0107	0.0999	0.0485	0.0111
3.3	0.1119‡	0.0579‡	0.0139‡	0.1032	0.0516	0.0124†	0.1026	0.0514	0.0124†
3.4	0.1144‡	0.0594‡	0.0142‡	0.1057*	0.0533	0.0134‡	0.1066†	0.0549†	0.0139‡
3.5	0.1066†	0.0550†	0.0128‡	0.0995	0.0491	0.0117*	0.1014	0.0502	0.0118*
3.6	0.1077†	0.0565‡	0.0140‡	0.0996	0.051	0.0137‡	0.1004	0.0511	0.0142‡
3.7	0.1145‡	0.0600‡	0.0147‡	0.1078‡	0.0526	0.0129‡	0.1080‡	0.0556†	0.0139‡
3.8	0.1110‡	0.0580‡	0.0151‡	0.1038	0.0511	0.0124†	0.1046	0.0519	0.0129‡
3.9	0.1098‡	0.0577‡	0.0146‡	0.1014	0.0523	0.0138‡	0.1025	0.0524	0.0145‡
4.0	0.1072†	0.0523	0.0112	0.0985	0.0468	0.0102	0.0998	0.0478	0.0112

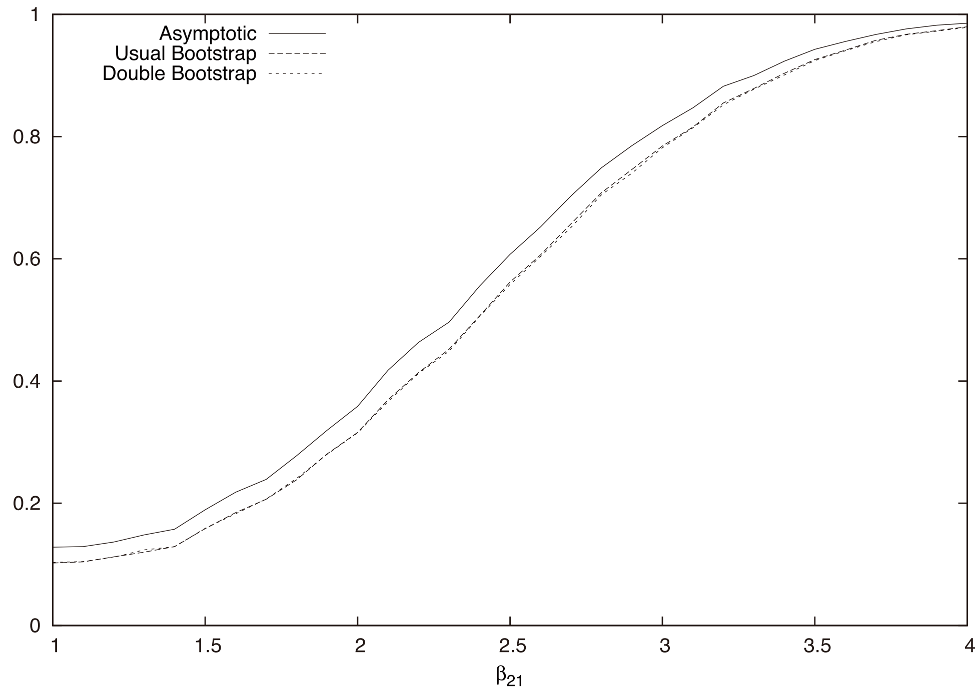

Figure 1 Powers of the tests for $(n_1, n_2) = (20, 20)$ at the 10% significance level.

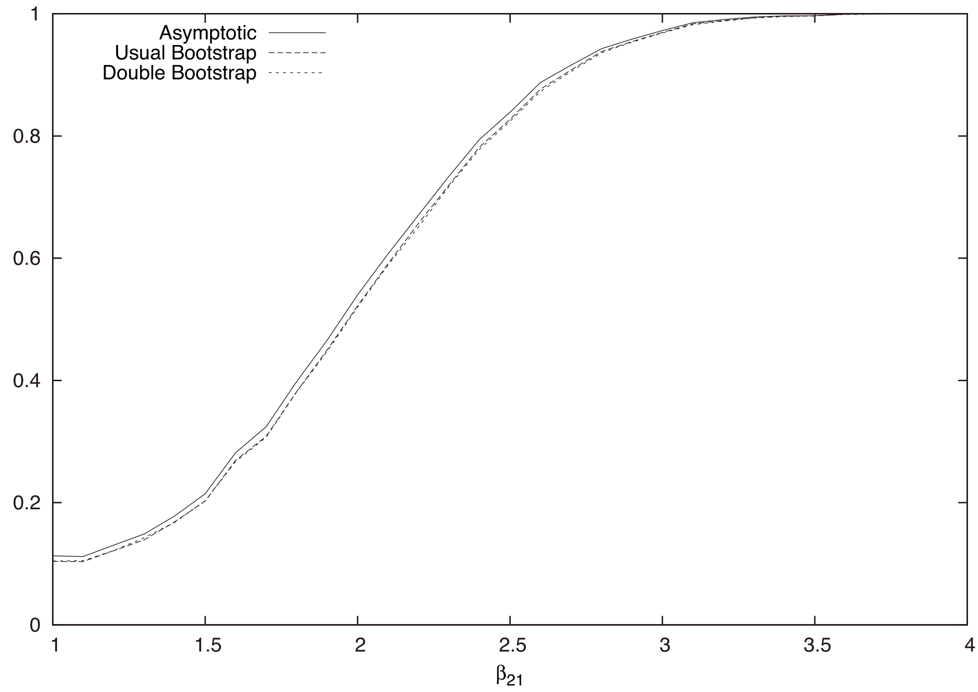

Figure 2 Powers of the tests for $(n_1, n_2) = (50, 50)$ at the 10% significance level.

4. Concluding remarks

In this paper we consider the Wald statistic for a structural break proposed by Watt (1979). We apply the usual bootstrap and double bootstrap methods to the Wald statistic. Our simulation results show that the size distortion of the asymptotic test based on the Wald statistic can be reduced by applying the bootstrap methods. Also, according to our simulation, though the double bootstrap test has a better empirical size in some cases, the superiority of the double bootstrap over the usual bootstrap does not always holds. Since the powers of the bootstrap methods are comparable to the power of the asymptotic test, the superiority of the bootstrap methods over the asymptotic test is obvious. Thus, as a whole, we can see the effectiveness of the bootstrap methods.

In the above sections, in order to examine the validity of the bootstrap methods for the model with structural breaks, we consider a simple model with a possible structural break and a known break point. However, some authors considered models with multiple breaks and unknown break points. In particular, when the break points are unknown, the asymptotic distribution of a test statistic is very complex and the test statistic is not asymptotically pivotal. In such situations, the double bootstrap method may have a better performance than the usual bootstrap method. However, investigating such models are beyond the scope of this paper and a remaining problem for future research.

REFERENCES

Bai, J. & P. Perron (1998), "Estimating and testing linear models with multiple structural changes," Econometrica, 47–78.

Beran, R. (1987), "Prepivoting to reduce level error of confidence sets," Biometrika 74, 457–468.

———— (1988), "Prepivoting test statistics: A bootstrap view of asymptotic refinements," Journal of the American Statistical Association 83, 687–697.

Boldea, O., A. R. Hall, & S. Han (2012), "Asymptotic distribution theory for break point estimators in models estimated via 2SLS," Econometric Reviews 31, 1–33.

Caers, J. & J. V. Dyck (1998), "Nonparametric tail estimation using a double bootstrap method," Computational Statistics & Data Analysis 29, 191–211.

Chow, G. C. (1960), "Tests of equality between sets of coefficients in two linear regressions," Econometrica, 591–605.

DiCiccio, T. J., M. A. Martin, & G. A. Young (1992), "Fast and accurate approximate double bootstrap confidence intervals," Biometrika 79, 285–295.

Efron, B. (1979), "Bootstrap methods: Another look at the jackknife," Annals of Statistics 7, 1–26.

Hall, A. R., S. Han, & O. Boldea (2012), "Inference regarding multiple structural changes in linear models with endogenous regressors," Journal of econometrics 170, 281–302.

Hall, P. (1992), The bootstrap and Edgeworth expansion, Springer.

Hinkley, D. & S. Shi (1989), "Importance sampling and the nested bootstrap," Biometrika 76, 435–446.

Honda, Y. & K. Ohtani (1986), "Modified Wald tests in tests of equality between sets of coefficients in two linear regressions under heteroscedasticity," The Manchester School 54, 208–218.

Jayatissa,W. (1977), "Tests of equality between sets of coefficients in two linear regressions when disturbance variances are unequal," Econometrica 45, 1291–92.

Kobayashi, M. (1986), "A bounds test of equality between sets of coefficients in two linear regressions when disturbance variances are unequal," Journal of the American Statistical Association 81, 510– 513.

Letson, D. & B. McCullough (1998), "Better confidence intervals: the double bootstrap with no pivot," American Journal of Agricultural Economics 80, 552–559.

McCullough, B. & H. D. Vinod (1998), "Implementing the double bootstrap," Computational Economics 12, 79–95.

McKnight, S. D., J. W. McKean, & B. E. Huitema (2000), "A double bootstrap method to analyze linear models with autoregressive error terms," Psychological Methods 5, 87–101.

Nankervis, J. C. (2005), "Computational algorithms for double bootstrap confidence intervals," Computational Statistics & Data Analysis 49, 461–475.

Ohtani, K. & M. Kobayashi (1986), "A bounds test for equality between sets of coefficients in two linear regression models under heteroscedasticity," Econometric Theory, 220–231.

Ohtani, K. & T. Toyoda (1985), "Small sample properties of tests of equality between sets of coefficients in two linear regressions under heteroscedasticity," International Economic Review 26, 37–44.

Perron, P. (2006), "Dealing with structural breaks," Palgrave handbook of econometrics 1, 278–352.

Perron, P. & Y. Yamamoto (2014), "A note on estimating and testing for multiple structural changes in models with endogenous regressors via 2SLS," Econometric Theory 30, 491–507.

Phillips, P. C. (1986), "The exact distribution of the Wald statistic," Econometrica, 881–895.

Rothenberg, T. J. (1984), "Hypothesis testing in linear models when the error covariance matrix is nonscalar," Econometrica, 827–842.

Schmidt, P. & R. Sickles (1977), "Some further evidence on the use of the Chow test under heteroskedasticity," Econometrica, 1293–1298.

Thursby, J. G. (1992), "A comparison of several exact and approximate tests for structural shift under heteroscedasticity," Journal of Econometrics 53, 363–386.

Toyoda, T. (1974), "Use of the Chow test under heteroscedasticity," Econometrica, 601–608.

Vinod, H. & B. McCullough (1995), "Estimating cointegration parameters: an application of the double bootstrap," Journal of Statistical Planning and Inference 43, 147–156.

Vinod, H. D. (1995), "Double bootstrap for shrinkage estimators," Journal of Econometrics 68, 287–302.

Watt, P. A. (1979), "Tests of equality between sets of coefficients in two linear regressions when disturbance variances are unequal: Some small sample properties," The Manchester School 47, 391–396.

Weerahandi, S. (1987), "Testing regression equality with unequal variances," Econometrica, 1211–1215.

Wu, C. F. J. (1986), "Jackknife, bootstrap and other resampling methods in regression analysis," Annals of Statistics 14, 1261–1295.