Graduate School of Economics, Hitotsubashi University
Discussion Paper Series No. 2019-01

Pitfalls of Two Step Testing for Changes in the Error Variance and Coefficients of a Linear Regression Model

Pierre Perron Yohei Yamamoto

April 2019

Pitfalls of Two Step Testing for Changes in the Error Variance and Coefficients of a Linear Regression Model*

Pierre Perron† Yohei Yamamoto‡

Boston University Hitotsubashi University

April 20, 2019

Abstract

In empirical applications based on linear regression models, structural changes often occur in both the error variance and regression coefficients possibly at different dates. A commonly applied method is to first test for changes in the coefficients (or in the error variance) and, conditional on the break dates found, test for changes in the variance (or in the coefficients). In this note, we provide evidence that such procedures have poor finite sample properties when the changes in the first step are not correctly accounted for. In doing so, we show that the test for changes in the coefficients (or in the variance) ignoring changes in the variance (or in the coefficients) induces size distortions and loss of power. Our results illustrate a need for a joint approach to test for structural changes in both the coefficients and the variance of the errors. We provide some evidence that the procedures suggested by Perron, Yamamoto and Zhou (2019) provide tests with good size and power.

JEL Classification Numbers: C12, C38

Keywords: structural change; variance shifts, CUSUM of squares tests, hypothesis testing, Sup-LR test.

---

*Some of the material in this paper was part of a paper previously circulated as Zhou, J. and P. Perron (2008): "Testing for Breaks in Coefficients and Error Variance: Simulations and Applications". The latter paper is now part of Perron, Yamamoto and Zhou (2019). This research was carried out in part while Yamamoto was visiting Boston University by financial support from the MEXT Fund for the Promotion of Joint International Research. Yamamoto is grateful to colleagues at Boston University for their hospitality.

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

‡Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo Japan, 186-8601 (yohei.yamamoto@econ.hit-u.ac.jp).

1 Introduction

In a companion paper, Perron, Yamamoto and Zhou (2019) provided a comprehensive treatment of the problem of testing jointly for structural changes in both the regression coefficients and the variance of the errors in a single equation regression model involving stationary regressors, allowing the break dates for the two components to be different or overlap. Their framework is quite general with assumptions no stronger than those in Qu and Perron (2007). The distribution of the errors can be non-normal and conditional heteroskedasticity is permissible. Extensions to the case with serially correlated errors were also covered. They also provided the required tools to address the following testing problems, among others: a) testing for given numbers of changes in regression coefficients and error variance; b) testing for an unknown number of changes less than a pre-specified maximum; c) testing for changes in variance (regression coefficients) allowing for a given number of changes in regression coefficients (variance); d) a sequential procedure to estimate the number of changes present.

Common methods in this context of testing for change in variance (coefficient) when changes in coefficients (variance) are possible are the following. When testing for a change in the error variance, many studies ignore the possibility of changes in coefficients and simply apply standard Sup-Wald type tests (e.g., Andrews, 1993, Bai and Perron, 1998) for changes in the mean of the absolute value of the estimated residuals (e.g., Herrera and Pesavento, 2005, Stock and Watson, 2002). For the problem of testing for a change in variance only (imposing no change in the regression coefficients), a more appropriate test is the CUSUM of squares test of Brown, Durbin and Evans (1975) extended by Deng and Perron (2008a) to allow general conditions on the regressors and the errors (as suggested by Inclán and Tiao, 1994, for normally distributed time series). This test is, however, adequate only with no change in coefficient. Similarly, when testing for a change in coefficients, most work simply use a similar Sup-Wald test applied to the regression coefficient ignoring the possibility of a variance change. It is often the case that changes in both coefficients and variance occur and the break dates need not be the same. Also, a common two step method is to first tests for changes in the regression coefficients and conditioning on the break dates found, then test for changes in variance. As will be shown in this note, all three approaches are clearly inappropriate as they suffer from severe size distortions and/or loss of power. Hence, what is needed is a joint approach when changes are suspected in both the coefficients and variance. This was covered in Perron, Yamamoto and Zhou (2019) and we present simulations results showing that their procedures yield tests with the good size and power.

In this note, we first assess the finite sample properties of structural change tests in coefficients when changes in the error variance are ignored. We then consider the properties of the CUSUM of squares test which tests for changes in the error variance ignoring possible changes in the coefficients. We also consider the two step method to test for a change in the error variance. All methods are shown to suffer from important size distortions and/or power losses. We then present evidence that the joint approach of Perron, Yamamoto and Zhou (2019) provides test with good size and power. Our work is related to that of Pitarakis (2004) and can be viewed as complementary. The tests considered are different. When testing for changes in coefficient, he considers the Sup-LM test (e.g., Andrews, 1993) which is often prone to important conservative size distortions, while we consider the Sup-LR test. When testing for changes in variance, he also considers some LM-type test, while we focus on the CUSUM of squares, shown to be valid under general conditions by Deng and Perron (2008a). He also considers only the issue of size distortions, while we also present results related to power. Finally, we also consider the properties of the commonly used two step method discussed above to detect a change in variance. On the other hand, his work contains theoretical results, while we solely focus on limited simulation experiments.

2 Models and test statistics

The data generating process (DGP) is a sequence of i.i.d. Normal random variable with mean and variance that can change at a single date. We specify

$$y_t = \mu + \delta_2 I(t > T^c) + e_t \tag{1}$$

for $t = 1, ..., T$, where $e_t \sim i.i.d.N(0, 1+\delta_1 I(t > T^v))$ with $I(\cdot)$ the indicator function. To analyze the effect of ignoring a variance break on the size of test for a change in coefficients (here the mean), we consider $\delta_2 = 0$. We consider 3 break dates, $T^v = \{[0.25T], [0.5T], [0.75T]\}$ and variance change δ_1 varying between 0 and 10 in steps of 0.05. The sample size is set to $T = 100$ and 5000 replications are used. The test considered is the standard Sup-LR test (see Andrews, 1993) for a one-time change in μ occurring at some unknown date. To assess the effect on power, we consider $T^v = \{[0.5T], [0.75T]\}; T^c = \{[0.3T]\}$, $T = 100$, $\delta_1 = \{0, 1, 2, 3\}$ and δ_2 varying between 0 and 2.

We consider the effect of a change in mean on the size and power of tests for a change in variance that do not take into account the former change. We consider two testing procedures. One is based on the CUSUM of squares test of Brown, Durbin and Evans (1975) and advocated as a test for a change in variance by Inclán and Tiao (1994), who showed that it is related to the LR test for a change in variance in a sequence of *i.i.d.* Normal random variables (though the equivalence is not exact in finite samples). Deng and Perron (2008a) generalized the conditions under which the test is valid; e.g., allowing for mixing type condition on the errors that permit conditional heteroskedasticity. It is defined by

$$CUSQ = \max_{k+1 \leq r \leq T} \sqrt{T/2} \left| S_T^{(r)} - \frac{r-k}{T-k} \right| \tag{2}$$

where $S_T^{(r)} = (\sum_{t=k+1}^{r} \widetilde{v}_t^2)/(\sum_{t=k+1}^{T} \widetilde{v}_t^2)$ with $\widetilde{v}_t$ the recursive residuals for $t = k+1, ..., T$ and k the number of regressors (here, $k = 1$). Here, its limit distribution under the null hypothesis is the supremum (over $[0, 1]$) of a Brownian Bridge process. To analyze the size of the test, DGP (1) with $\delta_1 = 0$ is used and we set $T^c = \{[.25T], [.5T], [.75T]\}$ with δ_2 varying between 0 and 10. For power, the DGP used is again (1) with δ_1 varying between 0 and 15 and $\delta_2 = \{0, 1, 2, 3\}$. The second procedure we consider is the two step method used by Herrera and Pesavento (2005) and Stock and Watson (2002), among others, which applies a test for a change in the mean of the absolute value of the estimated residuals when the later are obtained allowing for a change in the regression coefficients (here the mean) ignoring the possibility of a break in the error variance. Again, DGP (1) is used to assess the size ($\delta_1 = 0$) and power properties. For size, δ_2 varies between 0 and 10 and we set $T^c = \{[.25T], [.5T], [.75T]\}$, while for power δ_2 varies between 0 and 3 and we consider two sets of break dates, namely $\{T^c = [.5T], T^v = [.3T]\}$ and $\{T^c = [.75T], T^v = [.3T]\}$.

3 Results

The size of the Sup-LR test for a change in μ under DGP (1) is presented in Figure 1. The results show important size distortions unless the break occurs early at $T^v = [0.25T]$, and these are increasing with δ_1. The results for power under DGP (1) are presented in Figure 2, which show that power decreases as the magnitude of δ_1 increases. We next consider the results when testing for variance changes. The size of the CUSQ test (2) ignoring a coefficient change is presented in Figure 3. In all cases the size of the test increases to one rapidly as the magnitude of the change in mean δ_2 increases. This is not surprising in view of the fact that the CUSQ test has power against a change in the regression coefficients as originally argued by Brown, Durbin and Evans (1975); see also Deng and Perron (2008b). The results for power are presented in Figure 4, which show that a change in mean that is unaccounted for can increase the power of the CUSQ test. This result is, however, of little help given the large size distortions. Finally, the results of the two step method are presented in Figures 5 and 6. They show that the test suffers from serious size distortions, which increase as the change in mean increases. For the case of a break in mean at mid-sample, which suffers from conservative size distortions, Figure 6 shows that power decreases as the magnitude of the coefficient break increases.

An online supplement show that the results remain qualitatively the same for the following extended cases: a) models with lagged dependent variables; b) models with multiple structural changes; c) CUSQ tests for a change in variance that correct for potential correlation in the error variance; e.g., conditional heteroskedasticity.

4 Tests allowing for joint changes

Perron, Yamamoto and Zhou (2019) provided a comprehensive treatment for the problem of testing jointly for structural changes in the regression coefficients and the variance of the errors. Here, we consider two versions of their tests to illustrate how they solve the size and power problems. The first version investigates whether a given number (m) of structural changes in the coefficients are present when a given number (n) of structural changes in the error variance are accounted for. The structural change dates for both the coefficients and the variance are unknown and occur at the same or different times. The second version considers whether n structural changes in the error variance are present when m structural changes in the regression coefficients are allowed. Following their labels, we call the former testing problem TP-3 and the latter TP-2. The hypotheses are $H_0 : \{m = 0, n = n_a\}$ versus $H_1 : \{m = m_a, n = n_a\}$ for TP-3 and $H_0 : \{m = m_a, n = 0\}$ versus $H_1 : \{m = m_a, n = n_a\}$ for TP-2, where m_a and n_a are pre-selected values. Denote the break dates for the coefficients by $\{T_1^c, \cdots, T_m^c\}$, those for the error variance by $\{T_1^v, \cdots, T_n^v\}$ and the break fractions by $\{\lambda_1^c, \cdots, \lambda_m^c\}$ and $\{\lambda_1^v, \cdots, \lambda_n^v\}$, respectively. We also let the number of the union of the coefficient and the variance breaks be K.

The test statistics are the quasi-likelihood ratio tests assuming i.i.d. Gaussian disturbances. For TP-3, the log-likelihood function under H_0 is $\log L_T(T_1^v, ..., T_{n_a}^v) = -(T/2)(\log 2\pi + 1) - \sum_{i=1}^{n_a+1}[(T_i^v - T_{i-1}^v)/2)] \log \widetilde{\sigma}_i^2$, where $\widetilde{\sigma}_i^2 = (T_i^v - T_{i-1}^v)^{-1} \sum_{t=T_{i-1}^v+1}^{T_i^v} (y_t - \widetilde{\mu})^2$ for $i = 1, ..., n_a+1$ with $\widetilde{\mu} = T^{-1} \sum_{t=1}^{T} y_t$. Under H_1, $\log L_T(T_1^c, ..., T_{m_a}^c; T_1^v, ..., T_{n_a}^v) = -(T/2)(\log 2\pi + 1) - \sum_{i=1}^{n_a+1}[(T_i^v - T_{i-1}^v)/2)] \log \hat{\sigma}_i^2$, where $\hat{\sigma}_i^2 = (T_i^v - T_{i-1}^v)^{-1} \sum_{t=T_{i-1}^v+1}^{T_i^v} (y_t - \hat{\mu}_{t,j})^2$ for $i = 1, ..., n_a+1$ with $\hat{\mu}_{t,j} = (T_j^c - T_{j-1}^c)^{-1} \sum_{t=T_{j-1}^c+1}^{T_j^c} (y_t/\hat{\sigma}_i)$. Because the break dates are unknown, the supremum type LR test over all the permissible break dates is given by

$$
\begin{aligned}
&\sup LR_{3,T}\left(m_a, n_a, \varepsilon | m=0, n_a\right) \\
&= 2\left[\sup_{\left(\lambda_1^c,\ldots,\lambda_{m_a}^c;\lambda_1^v,\ldots,\lambda_{n_a}^v\right)\in\Lambda_\varepsilon} \log \hat{L}_T(T_1^c,\ldots,T_{m_a}^c;T_1^v,\ldots,T_{n_a}^v) - \sup_{\left(\lambda_1^v,\ldots,\lambda_{n_a}^v\right)\in\Lambda_{v,\varepsilon}} \log \widetilde{L}_T(T_1^v,\ldots,T_{n_a}^v)\right] \\
&= 2\left[\log \hat{L}_T(\widetilde{T}_1^c,\ldots,\widetilde{T}_{m_a}^c;\widetilde{T}_1^v,\ldots,\widetilde{T}_{n_a}^v) - \log \widetilde{L}_T(\hat{T}_1^v,\ldots,\hat{T}_{n_a}^v)\right]
\end{aligned}
$$

where Λ_ε is the union of the set of permissible break fractions for the coefficients and variance and $\Lambda_{v,\varepsilon}$ is a set of the permissible variance break fractions. ε is a small positive trimming value so that

$$
\begin{aligned}
\Lambda_\varepsilon &= \{(\lambda_1^c,\ldots,\lambda_m^c,\lambda_1^v,\ldots,\lambda_n^v)\,;\ \text{for } (\lambda_1,\ldots,\lambda_K) = (\lambda_1^c,\ldots,\lambda_m^c)\cup(\lambda_1^v,\ldots,\lambda_n^v) \\
&\quad |\lambda_{j+1}-\lambda_j| \geq \varepsilon\ (j=1,\ldots,K-1), \lambda_1 \geq \varepsilon, \lambda_K \leq 1-\varepsilon\}
\end{aligned}
$$

and $\Lambda_{v,\varepsilon} = \{(\lambda_1^v,\ldots,\lambda_n^v)\,;\ \left|\lambda_{i+1}^v - \lambda_i^v\right| \geq \varepsilon\ (i=1,\ldots,n_a-1), \lambda_1^v \geq \varepsilon, \lambda_{n_a}^v \leq 1-\varepsilon\}$. For TP-2, the Sup-LR test is

$$
\begin{aligned}
&\sup LR_{2,T}\left(m_a, n_a, \varepsilon | n=0, m_a\right) \\
&= 2\left[\sup_{\left(\lambda_1^c,\ldots,\lambda_{m_a}^c;\lambda_1^v,\ldots,\lambda_{n_a}^v\right)\in\Lambda_\varepsilon} \log \hat{L}_T(T_1^c,\ldots,T_{m_a}^c;T_1^v,\ldots,T_{n_a}^v) - \sup_{\left(\lambda_1^c,\ldots,\lambda_{m_a}^c\right)\in\Lambda_{c,\varepsilon}} \log \widetilde{L}_T(T_1^c,\ldots,T_{m_a}^c)\right] \\
&= 2\left[\log \hat{L}_T(\widetilde{T}_1^c,\ldots,\widetilde{T}_{m_a}^c;\widetilde{T}_1^v,\ldots,\widetilde{T}_{n_a}^v) - \log \widetilde{L}_T(\hat{T}_1^c,\ldots,\hat{T}_{m_a}^c)\right]
\end{aligned}
$$

where $\log \widetilde{L}_T(T_1^c,\ldots,T_{m_a}^c) = -(T/2)(\log 2\pi+1)-(T/2)\log\widetilde{\sigma}^2$, with $\widetilde{\sigma}^2 = T^{-1}\sum_{t=1}^T (y_t-\hat{\mu}_{t,j})^2$, $\hat{\mu}_{t,j} = (T_j^c - T_{j-1}^c)^{-1}\sum_{t=T_{j-1}^c+1}^{T_j^c} y_t$ and $\Lambda_{c,\varepsilon} = \{(\lambda_1^c,\ldots,\lambda_m^c)\,;\ \left|\lambda_{j+1}^c - \lambda_j^c\right| \geq \varepsilon\ (j=1,\ldots,m_a-1), \lambda_1^c \geq \varepsilon, \lambda_{m_a}^c \leq 1-\varepsilon\}$. Perron, Yamamoto and Zhou (2019) showed that the asymptotic distributions of these tests are bounded by limit distributions obtained in Bai and Perron (1998). Hence, somewhat conservative tests are possible using their critical values. They also show that the distortions are very minor via Monte Carlo simulations.

We implemented the *Sup*-$LR_{3,T}$ and *Sup*-$LR_{2,T}$ tests for the same DGP as above. We concentrate on testing for the presence of breaks rather than determining their number. Hence, we use the true values $n_a = 1$ and $m_a = 1$ as applicable when breaks are present. The size of the *Sup*-$LR_{3,T}$ test for a change in μ given one break in the error variance is presented in Figure 7. As explained in Perron, Yamamoto and Zhou (2019), the exact size is slightly smaller than the nominal level; the distortions due to the variance break are, however, minor. The size is more distorted as δ_1 becomes larger for the case of $T^v = [0.25T]$

but there are no evident distortions for the cases with $T^v = [0.5T]$ and $[0.75T]$. The power of the $Sup\text{-}LR_{3,T}$ test is presented in Figure 8 for the cases $T^c = [0.3T]$ and $T^v = [0.5T]$ as well as $T^c = [0.3T]$ and $T^v = [0.75T]$. For the first case, the power function decreases somewhat as the variance break increases. However, and more importantly, all power functions are higher than those presented in Figure 2, which indicates a reliable power performance. For the second case, the magnitude of the variance break has no evident effect on the power function, which remains high.

To test for structural breaks in the error variance, Figure 9 shows the size of the $Sup\text{-}LR_{2,T}$ test which accounts for a change in the mean. Again, the exact size is slightly conservative, as expected, but not affected by the magnitude of δ_2. Figure 10 shows the power functions of the $Sup\text{-}LR_{2,T}$ test for the cases $T^c = [0.5T]$ and $T^v = [0.3T]$ as well as $T^c = [0.75T]$ and $T^v = [0.3T]$. Here, the power functions are not affected by δ_2 for both cases and are higher than those in Figure 6. The results overall illustrate significant improvements of the size and power properties when using the conditional tests.

5 Conclusion

In this note, we provided evidence about the finite sample properties of the following testing procedures: a) applying a Sup-LR test for a change in regression coefficients ignoring the presence of a change in variance; b) applying a CUSUM of squares test for a change in variance ignoring the presence of a change in regression coefficients; c) a two step testing procedure for structural changes in the error variance using a test for a change in the mean of the absolute value of the estimated residuals when the latter are obtained allowing for a change in the regression coefficients ignoring the possibility of a break in the error variance and regression coefficients in a linear regression model possibly at different dates. The results show that all procedures have important size distortion and/or power losses. In an online supplement, the same qualitative results are shown to hold for models with lagged dependent variables, models with multiple structural changes and tests for changes in variance that account for conditional heteroskedasticity. While the setup considered is quite simple, it shows how inference can be misleading when changes in the coefficients and changes in the error variance are not analyzed jointly. To that effect, we presented limited results showing that the tests proposed by Perron, Yamamoto and Zhou (2019) have good size and power in small samples. This paper contains more extensive results about various tests, including sequential methods to estimate the number of breaks in the regression coefficients and error variance, which should be useful in practice.

References

Andrews, D.W.K. (1993): "Tests for parameter instability and structural change with unknown change point," *Econometrica*, 61, 821-856.

Bai, J., and Perron, P. (1998): "Estimating and testing linear models with multiple structural changes," *Econometrica*, 66, 47-78.

Brown, R.L., Durbin, J., and Evans, J.M. (1975): "Techniques for testing the constancy of regression relationships over time," *Journal of the Royal Statistical Society* B, 37, 149-163.

Deng, A., and Perron, P. (2008a): "The limit distribution of the CUSUM of squares test under general mixing conditions," *Econometric Theory*, 24, 809-822.

Deng, A., and Perron, P. (2008b): "A non-local perspective on the power properties of the CUSUM and CUSUM of squares tests for structural change," *Journal of Econometrics*, 142, 212-240.

Herrera, A.M., and Pesavento, E. (2005): "The decline in U.S. output volatility: structural changes and inventory investment," *Journal of Business & Economic Statistics*, 23, 462-472.

Inclán, C., and Tiao, G.C. (1994): "Use of cumulative sums of squares for retrospective detection of changes of variance," *Journal of the American Statistical Association*, 89, 913-923.

Perron, P., Yamamoto, Y., and Zhou, J. (2019): "Testing jointly for structural changes in the error variance and coefficients of a linear regression model," Unpublished manuscript, Boston University.

Pitarakis, J.-Y. (2004): Least-squares estimation and tests of breaks in mean and variance under misspecification," *Econometrics Journal*, 7, 32-54.

Qu, Z., and Perron, P. (2007): "Estimating and testing multiple structural changes in multivariate regressions," *Econometrica,* 75, 459-502.

Stock, J.H., and Watson, M.W. (2002): "Has the business cycle changed and why?" in *NBER Macroeconomics Annual* 17, M. Gertler & K. Rogoff (eds.), MIT press, 159-218.

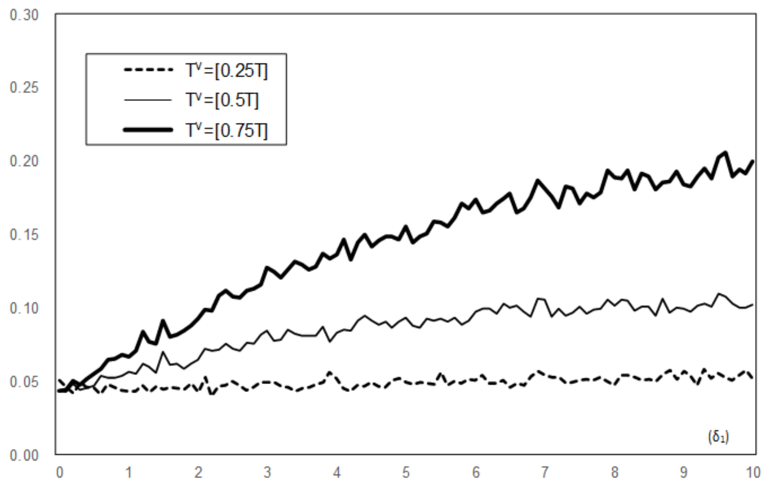

Figure 1: Size of the Sup-LR test for a coefficient change ignoring a variance change

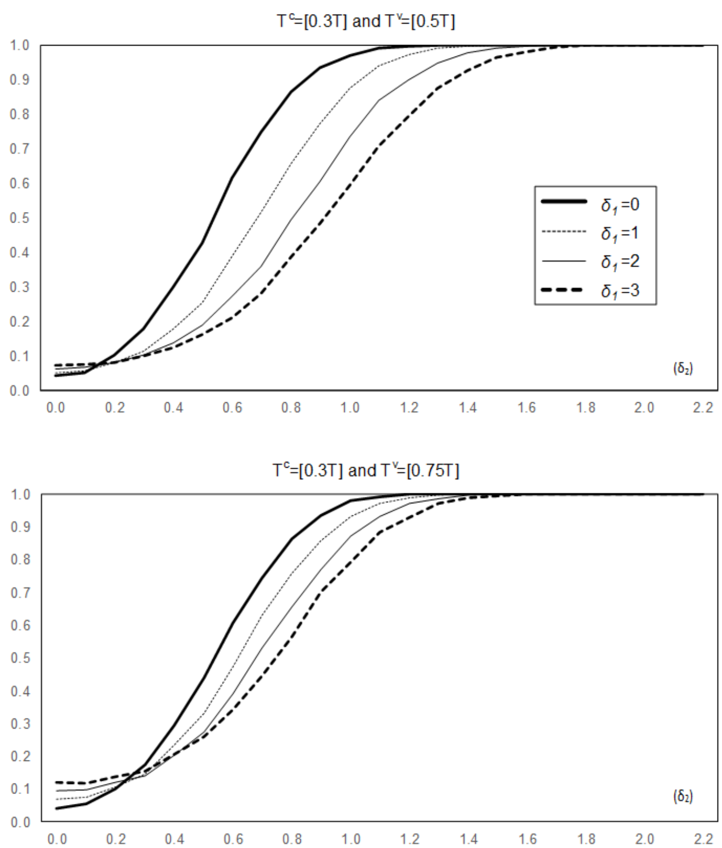

Figure 2: Power of the Sup-LR test for a coefficient change ignoring a variance change

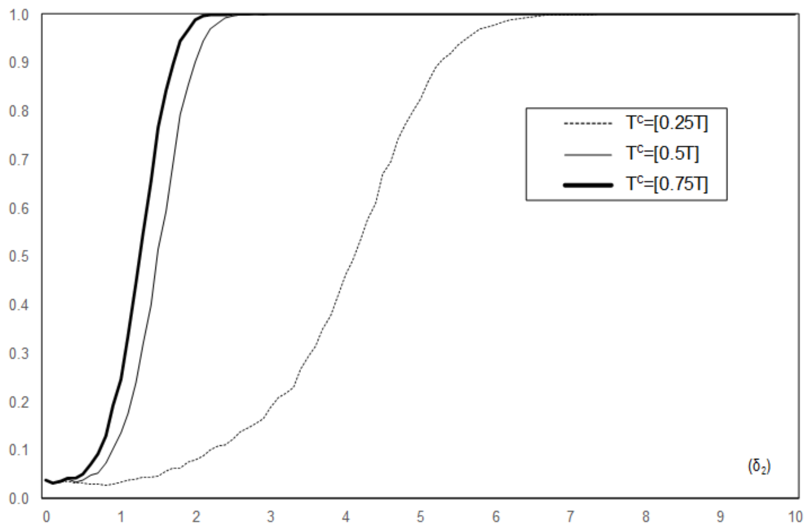

Figure 3: Size of the CUSQ test for a variance change ignoring a coefficient change

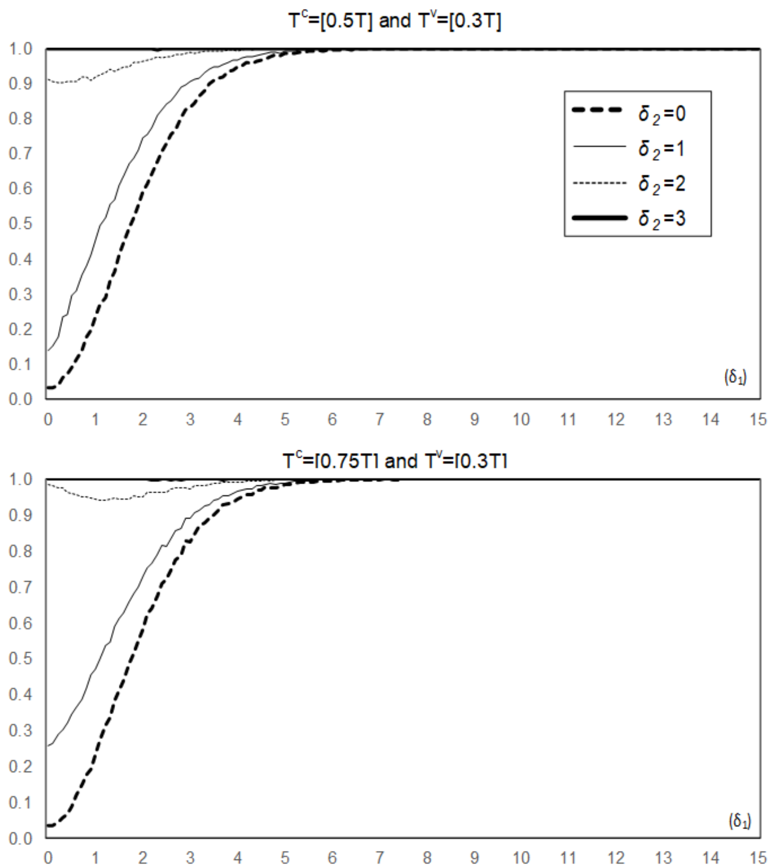

Figure 4: Power of the CUSQ test for a variance change ignoring a coefficient change

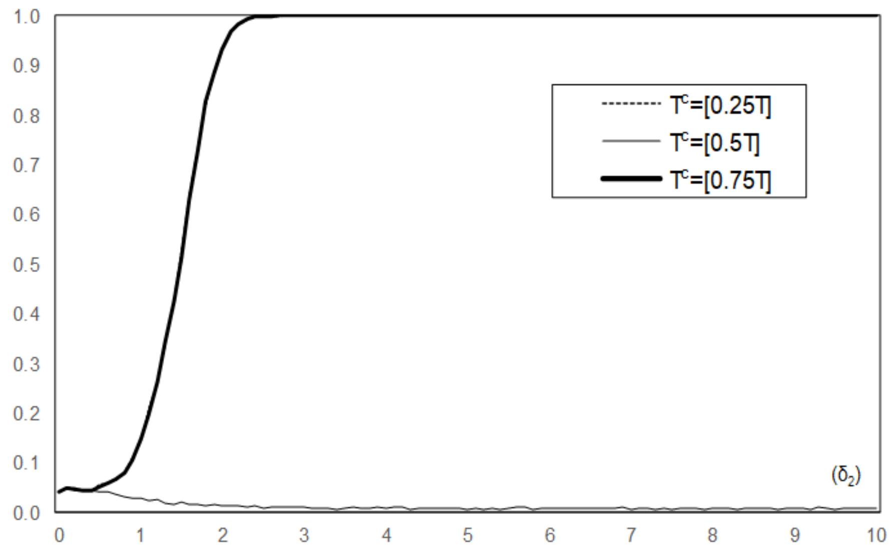

Figure 5: Size of the two step test for a variance change ignoring a coefficient change

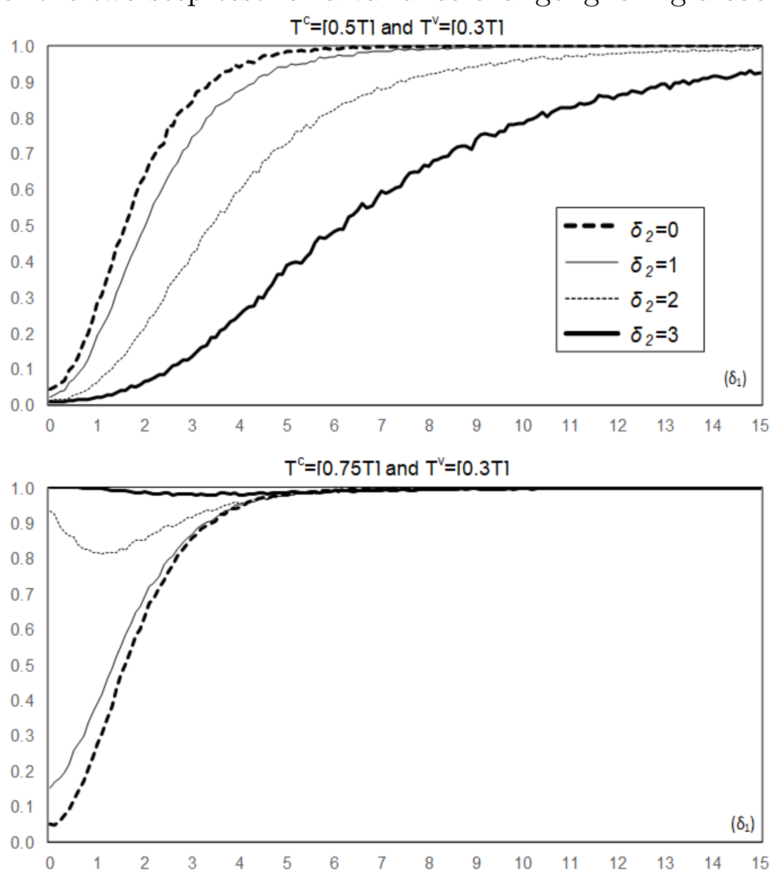

Figure 6: Power of the two step test for a variance change ignoring a coefficient change

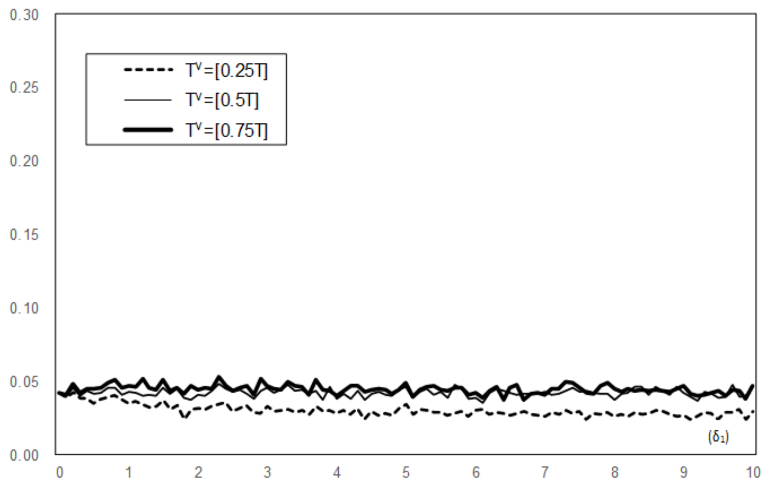

Figure 7: Size of the Sup-$LR_{3,T}$ test for a coefficient change accounting for a variance change

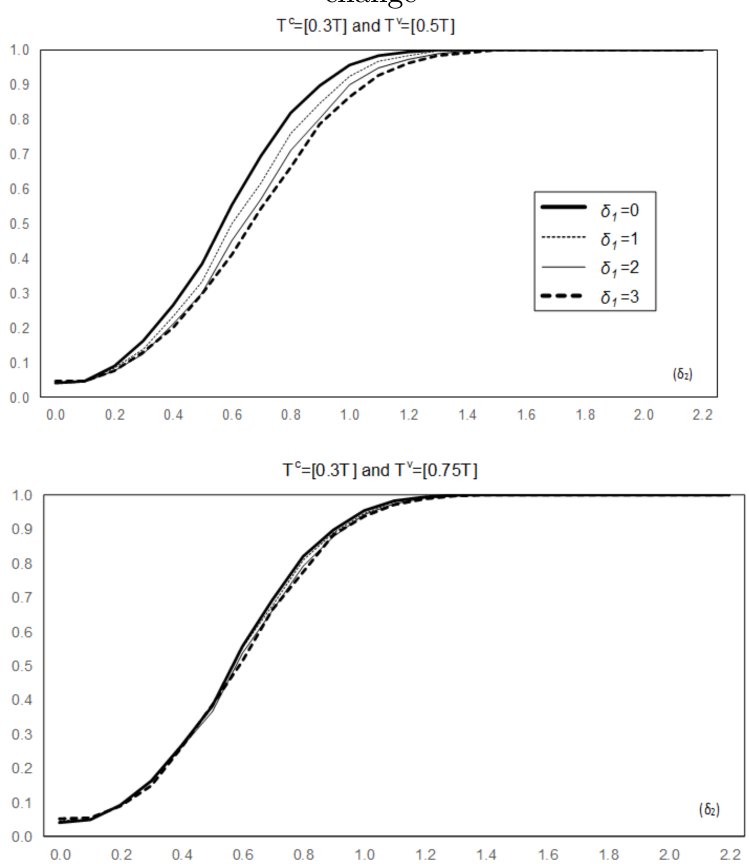

Figure 8: Power of the Sup-$LR_{3,T}$ test for a coefficient change accounting for a variance change

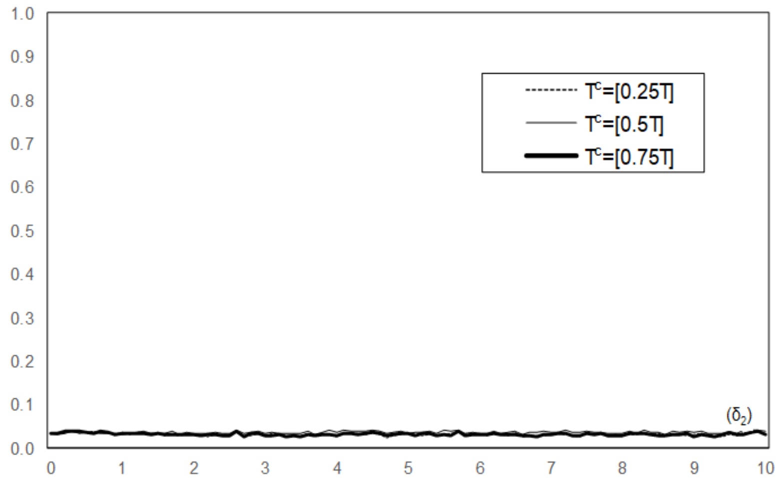

Figure 9: Size of the Sup-$LR_{2,T}$ test for a variance change accounting for a coefficient change

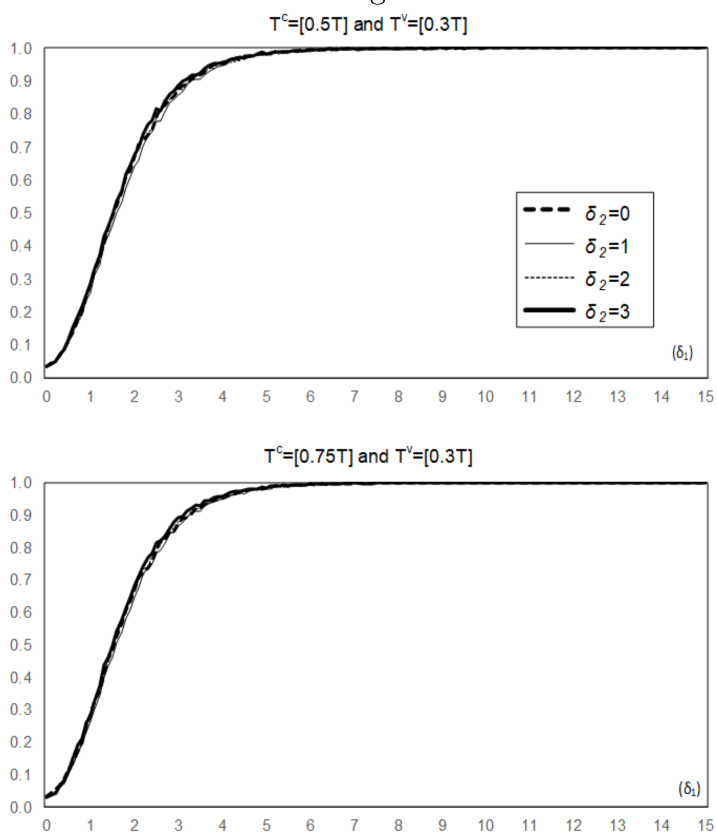

Figure 10: Power of the Sup-$LR_{2,T}$ test for a variance change accounting for a coefficient change