

A Two Step Procedure for Testing Partial Parameter Stability in Cointegrated Regression Models

Mohitosh Kejriwal*

Purdue University

Pierre Perron[†]

Boston University

Xuewen Yu[‡]

Purdue University

January 8, 2020

Abstract

Kejriwal and Perron (2010, KP) provided a comprehensive treatment for the problem of testing multiple structural changes in cointegrated regression models. A variety of models were considered depending on whether all regression coefficients are allowed to change (pure structural change) or a subset of the coefficients is held fixed (partial structural change). In this note, we first show that the limit distributions of the test statistics in the latter case are not invariant to changes in the coefficients not being tested; in fact, they diverge as the sample size increases. To address this issue, we propose a simple two step procedure to test for partial parameter stability. The first entails the application of a joint test of stability for all coefficients as in KP. Upon a rejection, the second conducts a stability test on the subset of coefficients of interest while allowing the other coefficients to change at the estimated breakpoints. Its limit distribution is standard chi-square. The relevant asymptotic theory is provided along with simulations that illustrates the usefulness of the procedure in finite samples.

Keywords: cointegration, partial structural change, break date, sup-Wald tests, joint hypothesis testing.

JEL Classification: C22

*Krannert School of Management, Purdue University, 403 West State Street, West Lafayette IN 47907 (mkejriwa@purdue.edu).

[†]Economics Department, Boston University, 270 Bay State Road, Boston MA 02215 (perron@bu.edu).

[‡]Krannert School of Management, Purdue University, 403 West State Street, West Lafayette IN 47907 (yu656@purdue.edu).

1 Introduction

Kejriwal and Perron (2010, KP henceforth) provided a comprehensive treatment of the problem of testing for multiple structural changes in cointegrated regression models. A number of test statistics were developed, including tests against a prespecified number of breaks, an unknown number of breaks subject to an upper bound and a sequential procedure to estimate the number of breaks. Their framework allows for both nonstationary [$I(1)$] and stationary [$I(0)$] regressors as well as serial correlation and conditional heteroskedasticity in the errors. A variety of models were considered depending on whether all coefficients are allowed to change [pure structural change] or a subset of coefficients is held fixed [partial structural change]. The limiting distributions of the test statistics were shown to be pivotal under the null hypothesis of no structural change and the relevant critical values tabulated. Partial structural change models are useful in that they allow for more powerful testing procedures, as illustrated via simulations by Kuo (1998). In the stationary framework of Bai and Perron (1998), tests of partial parameter stability remain asymptotically valid even in the presence of breaks in coefficients that are not under test. This invariance property facilitates the interpretation of the outcome of these tests and serves to identify the source of instability in the regression model. Such a property, however, no longer holds in the presence of $I(1)$ regressors so that the partial tests of KP can signal the presence of instability as long as *any* of the coefficients are unstable, including those that are not being tested.

In this note, we first show that the limit distributions of the test statistics in the partial structural change models are not invariant to changes in the coefficients not being tested. In fact, the test statistics diverge as the sample size increases. To address this issue, we propose a simple two step procedure to test for partial parameter stability. The first step entails the application of a joint test for the stability of all coefficients as in KP. Upon a rejection, the second step conducts a stability test on the subset of coefficients of interest while allowing the other coefficients to change at the estimated breakpoints. Its limit distribution is standard chi-square. The relevant asymptotic theory is provided along with simulation evidence that illustrates the adequacy of the performance in finite samples.

In a related paper, Hsu and Kuan (2001) studied the problem of distinguishing between intercept and slope breaks in a model with a bounded deterministic trend with a stationary noise component. They showed that the limit distributions of partial break test statistics are non-pivotal and depend on the magnitude of the coefficient break (intercept or slope) not under test. A similar result was demonstrated by Hsu (2008) in the context of cointegrated

regressions. In both studies, however, the asymptotic analysis was conducted in a framework in which the break size shrinks to zero as a function of the sample size at a rate ruling out consistent estimation of the break fractions, thereby invalidating a two-step testing approach. In contrast, our asymptotic framework allows the break fractions to be consistently estimated ensuring the large sample validity of the two step procedure.

This note is structured as follows. Section 2 presents the model and the test statistics. Section 3 details the proposed two step procedure to test for partial parameter stability. Monte Carlo simulation results are reported in Section 4 to assess the performance of the procedure in small samples and Section 5 provides brief concluding remarks. All proofs are provided in Appendix A. As a matter of notation, “ \xrightarrow{p} ” denotes convergence in probability, “ \xrightarrow{d} ” convergence in distribution and “ \Rightarrow ” weak convergence under the Skorohod metric.

2 Model and Test Statistics

The dependent variable y_t is generated according to the linear regression model with m breaks:

$$y_t = c_j + z'_{ft}\delta_{fj} + z'_{bt}\delta_{bj} + u_t, \quad t = T_{j-1} + 1, \dots, T_j \quad (1)$$

for $j = 1, \dots, m + 1$, ($m + 1$ being the number of regimes) where T is the sample size (by convention $T_0 = 0$, $T_{m+1} = T$), z_{ft} and z_{bt} are $(q_f \times 1)$ and $(q_b \times 1)$ vectors of $I(1)$ regressors, defined by: $z_{ft} = z_{f,t-1} + u_{zt}^f$, $z_{bt} = z_{b,t-1} + u_{zt}^b$, for $t = 1, \dots, T$, with z_{f0} and z_{b0} assumed to be fixed constants or $O_p(1)$ random variables. Equation (1), labelled **Model A**, represents a pure structural change model with all regression coefficients including the intercept allowed to change. The null hypothesis of stability is $H_{0,A}$: $c_j = c$, $\delta_{fj} = \delta_f$, $\delta_{bj} = \delta_b$ for all j . We also consider the following two partial structural change models, obtained as special cases of (1), by restricting a subset of the parameters to be fixed across regimes; namely **Model B**: $y_t = c + z'_{ft}\delta_f + z'_{bt}\delta_{bj} + u_t$ and **Model C**: $y_t = c_j + z'_{ft}\delta_f + z'_{bt}\delta_b + u_t$. In Model B, the objective is to test the stability of the coefficients of z_{bt} , i.e., $H_{0,B}$: $\delta_{bj} = \delta_b$ for all j . Similarly, the null hypothesis of interest in model C is the stability of the intercept: $H_{0,C}$: $c_j = c$ for all j . KP considered two additional partial break models: one with the null hypothesis of joint stability of (c_j, δ_{bj}) while holding δ_{fj} fixed across regimes; the other a special case of Model B, which does not include the regressors z_{ft} . They also considered allowing both $I(1)$ and $I(0)$ regressors and a variety of partial break submodels. For brevity, we do not consider these extensions but note that the two step procedure we advocate remains valid in these cases. We also focus on the single break case ($m = 1$) since the extension to multiple breaks is straightforward. KP proposed sup-Wald test statistics for each of

$H_{0,A}$, $H_{0,B}$ and $H_{0,C}$. For a given break fraction $\tau = T_1/T$, the Wald statistic for testing $H_{0,i}$ is $F_{T,i}(\tau) = [SSR_0 - SSR_i(\tau)]/\hat{\sigma}_i^2(\tau)$, where SSR_0 and $SSR_i(\tau)$ [$i = A, B, C$] are the sum of squared residuals under the null hypothesis of stability and that under the alternative of model i , respectively. The scaling factor $\hat{\sigma}_i^2(\tau)$ is an estimate of the long-run variance of u_t . Following KP, it is computed as

$$\hat{\sigma}_i^2(\tau) = T^{-1} \sum_{t=1}^T \tilde{u}_t^2 + 2 \sum_{j=1}^{T-1} w(j/b_T) T^{-1} \sum_{t=j+1}^T \tilde{u}_t \tilde{u}_{t-j}$$

where \tilde{u}_t are the residuals from the regression under the null hypothesis and $w(\cdot)$ is a continuous and even function with $|w(\cdot)| \leq 1$, $w(0) = 1$ and $\int_{-\infty}^{\infty} w^2(x) < \infty$. KP proposed using the quadratic spectral kernel with the bandwidth chosen via the rule $b_T = 1.3221(\hat{a}_2 T)^{1/5}$ advocated by Andrews (1991), where $\hat{a}_2 = 4\hat{\rho}^2/(1 - \hat{\rho})^4$, $\hat{\rho} = \sum_{t=2}^T \hat{u}_t \hat{u}_{t-1} / \sum_{t=2}^T \hat{u}_{t-1}^2$, with \hat{u}_t the residuals from the regression under the alternative hypothesis. This is a hybrid non-parametric estimate that employs residuals under both the null and alternative hypotheses which ensures that the test statistic is adequately sized while bypassing the problem of non-monotonic power that plagues the Lagrange Multiplier type tests (see KP for more details). For some arbitrary small positive number ϵ , define the set $\Lambda_\epsilon = \{\tau : \epsilon \leq \tau \leq 1 - \epsilon\}$. The sup-Wald test is then defined as $\sup F_{T,i}(\tau) = \sup_{\tau \in \Lambda_\epsilon} F_{T,i}(\tau)$. Let $\xi_t = (u_t, u_{zt}^f, u_{zt}^b)'$, a vector of dimension $n = q_f + q_b + 1$. Our analysis is based on the following set of assumptions, where here, and throughout, true values are denoted with a subscript 0:

• **Assumption A1:** The vector ξ_t satisfies the following multivariate Functional Central Limit Theorem (FCLT): $T^{-1/2} \sum_{t=1}^{\lfloor Tr \rfloor} \xi_t \Rightarrow B(r)$, with $B(r) = (B_1(r), B_z^f(r)', B_z^b(r)')'$ is a n vector Brownian motion with symmetric covariance matrix

$$\Omega = \begin{pmatrix} \sigma^2 & \Omega_{1z}^f & \Omega_{1z}^b \\ \Omega_{z1}^f & \Omega_{zz}^{ff} & \Omega_{zz}^{fb} \\ \Omega_{z1}^b & \Omega_{zz}^{bf} & \Omega_{zz}^{bb} \end{pmatrix} \begin{matrix} 1 \\ q_f \\ q_b \end{matrix} = \lim_{T \rightarrow \infty} T^{-1} E(S_T S_T') = \Sigma + \Lambda + \Lambda'$$

where $S_T = \sum_{t=1}^T \xi_t$, $\Sigma = \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T E(\xi_t \xi_t')$ and $\Lambda = \lim_{T \rightarrow \infty} T^{-1} \sum_{j=1}^{T-1} \sum_{t=1}^{T-j} E(\xi_t \xi_{t+j}')$. Also $\sigma^2 > 0$ and $p \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T u_t^2 = \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T E[u_t^2] \equiv \sigma_u^2$.

• **Assumption A2:** The matrix $\begin{pmatrix} \Omega_{zz}^{ff} & \Omega_{zz}^{fb} \\ \Omega_{zz}^{bf} & \Omega_{zz}^{bb} \end{pmatrix}$ is positive definite.

• **Assumption A3:** Let $\gamma_j^0 = (c_j^0, \delta_{fj}^{0f}, \delta_{bj}^{0b})'$, $j = 1, 2$ and $D_T = \text{diag}(1, T^{-1/2} I_{q_f}, T^{-1/2} I_{q_b})$. Then $\gamma_2^0 - \gamma_1^0 = D_T \lambda v_T$ where $\lambda = (\lambda_c, \lambda_f', \lambda_b')$ is independent of T and $v_T > 0$ is a scalar satisfying $v_T \rightarrow 0$ and $T^{1/2} v_T \rightarrow \infty$.

Assumptions A1-A2 are standard and the same as in KP. Assumption A3 adopts a shrinking shifts asymptotic framework whereby the magnitude of the break shrinks to zero as T increases with the coefficients of the $I(1)$ regressors shrinking faster than the intercept break (see Bai, Lumsdaine and Stock, 1998, Kejriwal and Perron, 2008a). The specified rates ensure that the true break fraction $\tau^0 = T_1^0/T$ can be consistently estimated and allows the construction of confidence intervals for the break date. KP derived the limit null distribution of the test statistics for models A, B and C under Assumptions A1-A2 and showed that they are pivotal, allowing the tabulation of critical values to perform the tests. In particular, the limit distributions pertaining to the partial break statistics are derived assuming that *all* parameters are stable under the null hypothesis (i.e., $\lambda = 0$ in Assumption A3), including the subset not under test. The following result shows that the asymptotic size of these test statistics is not invariant to changes in the subset of parameters not being tested.

Theorem 1 *Under Assumptions A1-A3, $\Omega_{1z}^f = \Omega_{1z}^b = 0$ and $\tau^0 \in \Lambda_\epsilon$: a) If $\lambda_c \neq 0$ and/or $\lambda_f \neq 0$ and $H_{0,B}$ holds, $\sup F_{T,B}(\tau)$ is $O_p(b_T^{-1}T)$ if $b_T \nu_T^2 \xrightarrow{p} \infty$, and $O_p(T\nu_T^2)$, otherwise. (b) If $\lambda_f \neq 0$ and/or $\lambda_b \neq 0$ and $H_{0,C}$ holds, the same results hold for $\sup F_{T,C}(\tau)$.*

Theorem 1 shows that the sup-Wald statistics have 100% asymptotic size when the instability comes from the set of parameters not part of the null hypothesis. Hence, the partial break statistics can be expected to suffer from considerable size distortions in finite samples so that a rejection cannot be attributed to a change in the parameters under test. Monte Carlo simulations reported in Section 4 confirm the relevance of this result in finite samples. Note that a similar result holds when the break magnitude is fixed (independent of T), namely $\sup F_{T,B}(\tau)$ and $\sup F_{T,C}(\tau)$ are $O_p(b_T^{-1}T)$.

Theorem 1 is derived under the assumption of strictly exogenous regressors, i.e., $\Omega_{1z}^f = \Omega_{1z}^b = 0$. This is not necessary and is only imposed to simplify the analysis. Endogenous $I(1)$ regressors can be accounted for using the dynamic least squares estimator (DOLS) which entails augmenting the regression with leads and lags of the first-differences of the $I(1)$ regressors (see Saikkonen, 1991) with the number selected using some information criteria (Kejriwal and Perron, 2008b). KP considered a general regression framework which allows for both $I(1)$ and $I(0)$ regressors. It can be shown that the asymptotic size of the partial break KP statistics is again not invariant to changes in a subset of the parameters even when testing the stability of the $I(0)$ coefficients. This size contamination does not occur in the stationary regression framework of Bai and Perron (1998) and is due to the presence of $I(1)$ regressors. In contrast, the two step procedure proposed below remains valid whether

one is interested in testing the stability of the intercept, the $I(1)$ or $I(0)$ coefficients, or any combination of these three sets of parameters.

3 Two Step Procedure

The analysis in the preceding section shows that the partial break KP statistics cannot be used to evaluate the stability of a subset of parameters in the presence of changes in the set of parameters that are not under test. Rather, a rejection by these statistics can only be interpreted as signaling instability in *any* of the model parameters. Thus, if the objective is not only to test for overall model stability but also to determine which particular subset of parameters is unstable, an alternative approach is needed. To achieve this, we propose the following two step procedure: 1) Conduct the test $\sup F_{T,A}(\tau)$ of joint stability of all parameters in regression (1). If the null hypothesis is not rejected at the desired level of significance, stop the procedure and conclude there is no evidence of instability. Otherwise, obtain the break date estimate $\hat{\tau}$ by minimizing the sum of squared residuals from (1) and proceed to the following step; 2) Conduct a F test using chi-squared critical values for the equality of the coefficient across regimes on the subset of coefficients of interest allowing the others to change at the estimated breakpoint. Upon a rejection, conclude in favor of a structural change in the subvector of interest, otherwise the stability cannot be rejected.

The asymptotic validity of the two step procedure follows from (i) the test in the first step is asymptotically pivotal under the null and consistent against alternatives involving a change in at least one parameter and (ii) the break fraction is consistently estimated as long as any of the parameters are subject to a break. The second fact ensures that the F test in the second step converges to a chi-square distribution under the null hypothesis of no structural change in the subvector of interest. This basically follows since the estimate of the break fraction is fast enough to ensure that the limit distribution of the parameter estimate is the same that would prevail if the break date was known. We thus have the following result where $F_{T,i}^{(2)}(\hat{\tau})$ denotes the second step test of the null hypothesis $H_{0,i}$ [$i = B, C$].

Theorem 2 *Suppose Assumptions A1-A3 hold, $\Omega_{1z}^f = \Omega_{1z}^b = 0$ and $\tau^0 \in \Lambda_\epsilon$: Under the conditions of Theorem 1(a), resp, 1(b), a) $F_{T,B}^{(2)}(\hat{\tau}) \xrightarrow{d} \chi^2(q_b)$, resp., b) $F_{T,C}^{(2)}(\hat{\tau}) \xrightarrow{d} \chi^2(1)$.*

4 Monte Carlo Evidence

This section presents the results of Monte Carlo experiments designed to assess the finite sample adequacy of the theoretical results. These will show that (i) the KP partial break

test statistics are subject to substantial over-rejections when the data generating process (DGP) involves a change in the subset of parameters outside those pertaining to the null hypothesis, and (ii) the two step procedure proposed has good size and considerable power in detecting deviations from stability. The design is similar to that in Kuo (1998). The dependent variable y_t is generated by: $y_t = c_t + \delta_t z_t + u_t$; $z_t = z_{t-1} + u_{zt}$, $u_{zt} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$. Four DGPs are considered: DGP-1: $c_t = 1$, $\delta_t = 1$ for all t ; DGP-2: $c_t = 1$ for all t , $\delta_t = 1$ if $t \leq [\tau^0 T]$ and $1 + \Delta_\delta$, otherwise; DGP-3: $c_t = 1$ if $t \leq [\tau^0 T]$ and $1 + \Delta_c$, otherwise, $\delta_t = 1$ for all t ; DGP-4: $c_t = \delta_t = 1$ if $t \leq [\tau^0 T]$, otherwise, $c_t = 1 + \Delta_c$ and $\delta_t = 1 + \Delta_\delta$. For the errors u_t , we consider three different cases: (a) (i.i.d. errors) $u_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$; (b) (AR(1) errors) $u_t = 0.5u_{t-1} + e_t$, $e_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$; (c) (MA(1) errors) $u_t = e_t - 0.5e_{t-1}$, $e_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$. The breakpoint is set at $\tau^0 = 0.5$. The regressor z_t is assumed to be strictly exogenous, i.e., u_{zt} and u_s are independent for all t and s . The trimming ϵ is set at 15%. In all cases, tests with a nominal 5% size are used. The number of replications throughout is 100,000.

In the first set of simulations, we compare the size and power of the partial break KP statistics and the two step procedure for $T \in \{120, 240\}$. The break magnitudes are set at $\Delta_c = 1, \Delta_\delta = 0.4$. Table 1 presents the results. Panel A reports the rejection frequencies when testing for a break in slope (δ) so that DGPs 1 and 3 pertain to size and DGPs 2 and 4 to power. While the partial break KP test has adequate size for DGP-1, size distortions are evident for DGP-3, irrespective of the error structure, which increase with T , consistent with the result in Theorem 1. In contrast, the proposed two step procedure exhibits much better size control across T and error structures, the exact size never exceeding 10%. A seemingly counterintuitive feature of the two step approach is that for DGP-3b [AR(1) errors], the empirical size need not approach the nominal size monotonically as T increases. We investigate this issue in detail later. Panel B reports the rejection frequencies when testing the stability of the intercept c . Here DGPs 1 and 2 correspond to size and DGPs 3 and 4 to power. Similar to the results in Panel A, the two step test has adequate size in all cases, while the one step KP test is subject to substantial size distortions under DGP-2 (a change in the slope parameter). In DGP-2b with positively serially correlated errors, the rejection frequencies of the two step test increase with T , as in Panel A. As expected, there is sometimes a loss in power using the two step test relative to the partial break KP test; e.g., DGP-2 in Panel A, DGP-3 in Panel B, since in these cases the KP test is already tailored to have power against alternatives specified by the DGP. The difference in power is more prominent with AR(1) errors but less so otherwise. Notwithstanding its two step nature, our recommended procedure retains respectable power that increases noticeably with T .

We also consider DGPs involving both $I(1)$ and $I(0)$ regressors given by: $y_t = \mu_t + \beta_t x_t + \delta_t z_t + u_t$, where $x_t \stackrel{i.i.d.}{\sim} \mathcal{N}(1, 1)$, $z_t = z_{t-1} + u_{zt}$ and $u_{zt} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$. Four DGPs are considered: DGP-5: $\mu_t = \beta_t = 1$, $\delta_t = 1$ if $t \leq [\tau^0 T]$, otherwise $\delta_t = 1.4$; DGP-6: $\mu_t = 1$, $\beta_t = \delta_t = 1$, if $t \leq [\tau^0 T]$, otherwise $\beta_t = 3$, $\delta_t = 1.4$; DGP-7: $\beta_t = \delta_t = 1$ and $\mu_t = 1$ if $t \leq [\tau^0 T]$, otherwise $\mu_t = 2$; DGP-8: for $t \leq [\tau^0 T]$, $\mu_t = \beta_t = \delta_t = 1$, and for $t > [\tau^0 T]$, $\mu_t = 2$, $\beta_t = 3$, $\delta_t = 1.4$. For each DGP, we are interested in testing the stability of the $I(0)$ coefficient β_t . Thus DGPs 5 and 7 correspond to size while DGPs 6 and 8 correspond to power. The same three error structures are allowed for u_t as described above. The results are presented in Panel C of Table 1. For DGPs 5 and 7, the size of the two-step procedure is near the nominal 5% level, except when the sample size is small with AR(1) errors, though the distortions in this reduce considerably as T increases. For DGPs 6 and 8, the results show substantial power. The standard KP test is again rejecting far too often, indicating its non-robustness even when testing the stability of $I(0)$ coefficients.

The second set of simulations examines the impact of the break magnitude and T on test size. We focus on DGPs 2b-3b for which we observed an increase in size distortions for the two step procedure as T increased from 120 to 240 for a fixed break magnitude. Table 2 presents the results. Panel A reports the rejection frequencies for T between 60 and 600 with the break magnitude fixed. We include the first step, second step and final rejection frequencies for the two step procedure to investigate the contribution of each to the final test outcome. The following patterns are worth noting. First, the magnitude of the size distortions incurred by the two step approach does not monotonically decrease as T increases, i.e., it initially increases and then decreases. Second, while the increase in the first step rejection frequencies reflects the expected increase in power, the second step rejection frequencies decrease as T increases, reflecting the reduced estimation uncertainty about the break date. The evolution of the final rejection rate (the product of the first and second stage rates) as a function of T thus depends on the rate of increase in first stage power vis-a-vis the rate of reduction in second stage size distortions. This observation explains the pattern of results reported in Table 1 for DGP-2b and DGP-3b. Panel B of Table 2 explores the behavior of test size as a function of the break magnitude when $T = 120$. The results resemble those in Panel A, with a hump-shaped pattern for the final rejection rate, caused by an increase in the first-stage power accompanied by a reduction in sampling uncertainty about the break date, as the magnitude of the break increases. Figure 1 summarizes the results in Table 2 graphically, plotting the first, second and final stage rejection rates as a function of T for a given break size ($\Delta_c = 1, \Delta_\delta = 0.4$) and as a function of break size for a given sample size ($T = 120$).

5 Conclusion

This note dealt with testing for partial parameter stability in cointegrated regression models. Using an asymptotic framework for the break magnitude ensuring consistent estimates of the break fractions, we first showed that existing partial break sup-Wald tests diverge with T when the coefficients not being tested are subject to change. We proposed a simple two step procedure which first tests for joint parameter stability and subsequently conducts a standard chi-squared stability test on the coefficients of interest allowing the other coefficients to change at the breakpoints estimated by minimizing the sum of squared residuals in the pure structural change model. The relevant asymptotic theory is provided and simulations showed the procedure to work well in a variety of scenarios.

References

- Andrews, D. W. K. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica* 59, 817-858.
- Bai, J. (1997), "Estimation of a change point in multiple regression models," *Review of Economics and Statistics* 79, 551-563.
- Bai, J., Perron, P. (1998), "Estimating and testing linear models with multiple structural changes," *Econometrica* 66, 47-78.
- Bai, J., Lumsdaine, R. L., Stock, J. H. (1998), "Testing for and dating common breaks in multivariate time series," *Review of Economic Studies* 65, 395-432.
- Hsu, C. C. (2008), "A note on tests of partial parameter stability in the cointegrated system," *Economics Letters* 99, 500-503.
- Hsu, C. C., Kuan, C. M. (2001), "Distinguishing between trend-break models: Method and empirical evidence," *Econometrics Journal* 4, 171-190.
- Kejriwal, M., Perron, P. (2008a), "The limit distribution of the estimates in cointegrated regression models with multiple structural changes," *Journal of Econometrics* 146, 59-73.
- Kejriwal, M., Perron, P. (2008b), "Data dependent rules for the selection of the number of leads and lags in the dynamic OLS cointegrating regression," *Econometric Theory* 24, 1425-1441.
- Kejriwal, M., Perron, P. (2010), "Testing for multiple structural changes in cointegrated regression models," *Journal of Business & Economic Statistics* 28, 503-522.
- Kuo, B. S. (1998), "Test for partial parameter instability in regressions with I (1) processes," *Journal of Econometrics* 86, 337-368.
- Saikkonen, P. (1991), "Asymptotically efficient estimation of cointegration regressions," *Econometric Theory* 7, 1-21.

Appendix

For any matrix $W_{T \times q} = (w_1, \dots, w_T)'$, define the projection matrices $P_W = W(W'W)^{-1}W'$, $M_W = I_{T \times q} - P_W$ and the matrix \bar{W} that diagonally partitions W at T_1 , i.e., $\bar{W} = \text{diag}(W_1, W_2)$, where $W_i = (w_{T_{i-1}+1}, \dots, w_{T_i})'$ ($i = 1, 2$) with $T_0 = 0$ and $T_2 = T$. Also, let $Y = (y_1, \dots, y_T)'$, $\iota_{T \times 1} = (1, \dots, 1)'$, $u = (u_1, \dots, u_T)'$, $Z_f = (z_{f,1}, \dots, z_{f,T})'$, $Z_b = (z_{b,1}, \dots, z_{b,T})'$.

Proof of Theorem 1: We prove the result for case (a), as the proof of case (b) follows using similar arguments. Throughout, the true values are denoted with a superscript 0. Let

$$\eta = \underbrace{(0, \dots, 0)'}_{1 \times [\tau^0 T]} \underbrace{(\lambda_c \nu_T + z'_{f, [\tau^0 T] + 1} \lambda_f T^{-1/2} \nu_T, \dots, \lambda_c \nu_T + z'_{f, T} \lambda_f T^{-1/2} \nu_T)'}_{1 \times [(1 - \tau^0) T]}$$

and $\bar{\delta}_b^0 = (\delta_{b1}^{0'}, \delta_{b2}^{0'})'$. Under $H_{0,B}$, $\delta_{b1}^0 = \delta_{b2}^0 = \delta_b^0$ so that the DGP is

$$Y = c_1^0 \iota + Z_f \delta_f^0 + Z_b \delta_b^0 + u + \eta = c_1^0 \iota + Z_f \delta_f^0 + \bar{Z}_b \bar{\delta}_b^0 + u + \eta. \quad (\text{A.1})$$

Let $G_0 = [\iota, Z_f, Z_b] = (G_{0,1}, \dots, G_{0,T})'$. The restricted sum of squared residuals is

$$\begin{aligned} SSR_0 &= \sum_{t=1}^T \tilde{u}_t^2 = (u + \eta)' M_{G_0} (u + \eta) = (u + \eta)' (u + \eta) - (u + \eta)' P_{G_0} (u + \eta) \\ &= u'u + 2u'\eta + \eta'\eta - (u' P_{G_0} u + 2u' P_{G_0} \eta + \eta' P_{G_0} \eta). \end{aligned} \quad (\text{A.2})$$

Defining $J_T = \text{diag}(T^{-1/2}, T^{-1} I_{q_f}, T^{-1} I_{q_b})$, we have

$$\begin{aligned} T^{-1/2} \nu_T^{-1} J_T G_0' \eta &= \sum_{t=1}^T J_T G_{0,t} T^{-1/2} \nu_T^{-1} \eta_t = T^{-1} \sum_{t=[\tau^0 T] + 1}^T \{(T^{1/2} J_T) G_{0,t}\} \{\nu_T^{-1} \eta_t\} \\ &\Rightarrow \left(\int_{\tau^0}^1 (\lambda_c + \lambda_f B_z^f(r)) dr, \int_{\tau^0}^1 B_z^{f'}(r) (\lambda_c + \lambda_f B_z^f(r)) dr, \int_{\tau^0}^1 B_z^{b'}(r) (\lambda_c + \lambda_f B_z^f(r)) dr \right)' = O_p(1). \end{aligned}$$

Then it follows that $T^{-1} u'u = O_p(1)$, $T^{-1/2} \nu_T^{-1} u'\eta = T^{-1/2} u'(\nu_T^{-1} \eta) = O_p(1)$

$$\begin{aligned} T^{-1} \nu_T^{-2} \eta'\eta &= T^{-1} (\nu_T^{-1} \eta)' (\nu_T^{-1} \eta) = O_p(1) \\ u' P_{G_0} u &= u' G_0 (G_0' G_0)^{-1} G_0' u = \{u' G_0 J_T\} \{(J_T G_0' G_0 J_T)^{-1}\} \{J_T G_0' u\} = O_p(1) \\ T^{-1/2} \nu_T^{-1} u' P_{G_0} \eta &= \{u' G_0 J_T\} \{(J_T G_0' G_0 J_T)^{-1}\} \{T^{-1/2} \nu_T^{-1} J_T G_0' \eta\} = O_p(1) \\ T^{-1} \nu_T^{-2} \eta' P_{G_0} \eta &= \{T^{-1/2} \nu_T^{-1} \eta' G_0 J_T\} \{(J_T G_0' G_0 J_T)^{-1}\} \{T^{-1/2} \nu_T^{-1} J_T G_0' \eta\} = O_p(1). \end{aligned} \quad (\text{A.3})$$

Let $G_1 = [\iota, Z_f, \bar{Z}_b]$. The unrestricted sum of squared residuals evaluated at $\tau \in \Lambda_\epsilon$ is

$$SSR_B(\tau) = \sum_{t=1}^T \hat{u}_t^2 = u'u + 2u'\eta + \eta'\eta - (u' P_{G_1} u + 2u' P_{G_1} \eta + \eta' P_{G_1} \eta) \quad (\text{A.4})$$

where the orders of $u' P_{G_1} u$, $u' P_{G_1} \eta$ and $\eta' P_{G_1} \eta$ are the same as those of $u' P_{G_0} u$, $u' P_{G_0} \eta$ and $\eta' P_{G_0} \eta$, respectively, stated in (A.3). Combining (A.2) and (A.4), we have:

$$\begin{aligned} &SSR_0 - SSR_B(\tau) \\ &= (u' P_{G_1} u + 2u' P_{G_1} \eta + \eta' P_{G_1} \eta) - (u' P_{G_0} u + 2u' P_{G_0} \eta + \eta' P_{G_0} \eta) \\ &= [O_p(1) + O_p(T^{1/2} \nu_T) + O_p(T \nu_T^2)] - [O_p(1) + O_p(T^{1/2} \nu_T) + O_p(T \nu_T^2)] = O_p(T \nu_T^2) \end{aligned} \quad (\text{A.5})$$

For the long-run variance estimate $\hat{\sigma}_B^2(\tau)$, we have

$$\hat{\sigma}_B^2(\tau) = T^{-1} \sum_{t=1}^T \tilde{u}_t^2 + 2 \sum_{j=1}^{T-1} w(j/b_T) T^{-1} \sum_{t=j+1}^T \tilde{u}_t \tilde{u}_{t-j} \quad (\text{A.6})$$

$$= [T^{-1} \sum_{t=1}^T u_t^2 + O_p(\nu_T^2)] + 2 \sum_{j=1}^{T-1} w(j/b_T) T^{-1} \sum_{t=j+1}^T u_t u_{t-j} + O_p(b_T \nu_T^2) \quad (\text{A.7})$$

$$= T^{-1} \sum_{t=1}^T u_t^2 + 2 \sum_{j=1}^{T-1} w(j/b_T) T^{-1} \sum_{t=j+1}^T u_t u_{t-j} + o_p(1) + O_p(b_T \nu_T^2) \quad (\text{A.8})$$

$$= \sigma^2 + O_p(b_T \nu_T^2)$$

The equality of the first term in (A.6) with the term within square brackets in (A.7) follows from (A.3). For the second term in (A.6), note that, for a given j , $\tilde{u}_{t-j} = (u_{t-j} + \eta_{t-j}) - G'_{0,t-j}(G'_0 G_0)^{-1} G'_0(u + \eta)$. It follows that

$$\begin{aligned} T^{-1} \sum_{t=j+1}^T \tilde{u}_t \tilde{u}_{t-j} &= T^{-1} \sum_{t=j+1}^T [(u_t + \eta_t) - G'_{0,t}(G'_0 G_0)^{-1} G'_0(u + \eta)] [(u_{t-j} + \eta_{t-j}) \\ &\quad - G'_{0,t-j}(G'_0 G_0)^{-1} G'_0(u + \eta)] \\ &= T^{-1} \sum_{t=j+1}^T [u_t u_{t-j} + u_t \eta_{t-j} + \eta_t u_{t-j} + \eta_t \eta_{t-j} \\ &\quad - (u_t + \eta_t) G'_{0,t-j}(G'_0 G_0)^{-1} G'_0(u + \eta) - (u_{t-j} + \eta_{t-j}) G'_{0,t}(G'_0 G_0)^{-1} G'_0(u + \eta) \\ &\quad + (u + \eta)' G_0 (G'_0 G_0)^{-1} G'_{0,t} G'_{0,t-j} (G'_0 G_0)^{-1} G'_0(u + \eta)] \\ &= T^{-1} \sum_{t=j+1}^T u_t u_{t-j} + T^{-1} [O_p(T^{1/2} \nu_T) + O_p(T^{1/2} \nu_T) + O_p(T \nu_T^2) \\ &\quad + O_p(T \nu_T^2) + O_p(T \nu_T^2) + O_p(T \nu_T^2)] \\ &= T^{-1} \sum_{t=j+1}^T u_t u_{t-j} + O_p(\nu_T^2), \quad \text{uniformly in } j. \end{aligned} \quad (\text{A.9})$$

Using $b_T^{-1} \sum_{j=1}^{T-1} |w(j/b_T)| \rightarrow \int_0^{+\infty} |w(x)| dx < \infty$ (e.g. Andrews, 1991), we have from (A.9),

$$\begin{aligned} &(b_T \nu_T^2)^{-1} \left\{ \sum_{j=1}^{T-1} w(j/b_T) T^{-1} \sum_{t=j+1}^T \tilde{u}_t \tilde{u}_{t-j} - \sum_{j=1}^{T-1} w(j/b_T) T^{-1} \sum_{t=j+1}^T u_t u_{t-j} \right\} \\ &\leq b_T^{-1} \sum_{j=1}^{T-1} |w(j/b_T)| \sup_{j \geq 1} |v_T^{-2} T^{-1} \sum_{t=j+1}^T \tilde{u}_t \tilde{u}_{t-j} - v_T^{-2} T^{-1} \sum_{t=j+1}^T u_t u_{t-j}| \\ &= [b_T^{-1} \sum_{j=1}^{T-1} |w(j/b_T)|] O_p(1) = O_p(1), \end{aligned} \quad (\text{A.10})$$

which establishes (A.8).

Combining the results of the numerator (A.5) and the denominator (A.8) of the statistic $F_{T,B}(\tau)$, we have: $F_{T,B}(\tau) = O_p(T \nu_T^2) / [\sigma^2 + O_p(b_T \nu_T^2)] = O_p(b_T^{-1} T)$ if $b_T \nu_T^2 \xrightarrow{p} \infty$, and is $O_p(T \nu_T^2)$ otherwise. This result naturally extends to $\sup F_{T,B}(\tau) = \sup_{\tau \in \Lambda_\epsilon} F_{T,B}(\tau)$, which completes the proof.

Proof of Theorem 2: We only prove (a), as the proof of (b) follows using similar arguments. We first show that $F_{T,B}^{(2)}(\tau^0)$ has a limiting $\chi^2(q_b)$ distribution. For the restricted regression under $H_{0,B}$, we denote the design matrix as $X_0(\tau^0) = [\bar{v}^0, \bar{Z}_f^0, Z_b]$, where \bar{v}^0 $[\bar{Z}_f^0]$ is a matrix which diagonally partitions v $[Z_f]$ at the true break point T_1^0 . For the unrestricted regression, we similarly have $X_1(\tau^0) = [\bar{v}^0, \bar{Z}_f^0, \bar{Z}_b^0]$. For notational simplicity, we simply drop the index τ^0 in $X_0(\tau^0)$ and $X_1(\tau^0)$. First, we note that $Z_b = \bar{Z}_b^0 E$, where $E = (I_{q_b}, I_{q_b})'$. Note that

$X_0 = X_1 H$, where $H = \text{diag}(I_{2(1+q_f)}, E_{2q_b \times q_b})$. It then follows that

$$SSR_0(\tau^0) - SSR_B(\tau^0) = u' M_{X_0} u - u' M_{X_1} u = u'(P_{X_1} - P_{X_0})u.$$

Note that $P_{X_{01}} = P_{X_1} - P_{X_0}$ is an orthogonal projection matrix since the column space of X_0 is included in that of X_1 . Hence, there exists a $(T \times q_b)$ matrix X_{01} with rank q_b that satisfies $P_{X_{01}} = X_{01}(X'_{01}X_{01})^{-1}X'_{01}$. Then, applying a central limit theorem conditional on X_{01} , we have that $[SSR_0(\tau^0) - SSR_B(\tau^0)]/\sigma^2 = u' P_{X_{01}} u/\sigma^2 \xrightarrow{d} \chi^2(q_b)$. Since the limit does not depend on the conditioning matrix X_{01} , it is also the unconditional distribution. Finally, since $\hat{\sigma}_B^2(\tau^0) \xrightarrow{p} \sigma^2$ under $H_{0,B}$, it follows that $F_{T,B}^{(2)}(\tau^0) = [SSR_0(\tau^0) - SSR_B(\tau^0)]/\hat{\sigma}_B^2(\tau^0) \xrightarrow{d} \chi^2(q_b)$.

We next prove that $F_{T,B}^{(2)}(\hat{\tau}) = F_{T,B}^{(2)}(\tau^0) + o_p(1)$. Let \hat{T}_1 be the estimated break date, i.e., $\hat{T}_1 = [T\hat{\tau}]$. From Kejriwal and Perron (2008a, Theorem 2), $\hat{\tau}$ is $T\nu_T^2$ -consistent for τ^0 . Thus $\hat{T}_1 = T_1^0 + O_p(\nu_T^{-2})$. Let $\hat{T}_1 = T_1^0 + [s\nu_T^{-2}]$, $z_{1f,t} = (1, z'_{f,t})'$, $\lambda_{cf} = (\lambda_c, \lambda'_f)'$, $D_{cf,T} = \text{diag}(1, T^{-1/2}I_{q_f})$. Denote \overline{SSR} as the sum of squared residuals from estimating the model without breaks, i.e., $\lambda_c = 0_{1 \times 1}$, $\lambda_b = 0_{q_b \times 1}$ and $\lambda_f = 0_{q_f \times 1}$. Following Bai (1997, Lemma A.5), consider $\hat{T}_1 \leq T_1^0$. Then we can write

$$\begin{aligned} & SSR_0(\tau^0) - SSR_0(\hat{\tau}) \\ &= [\overline{SSR} - SSR_0(\hat{\tau})] - [\overline{SSR} - SSR_0(\tau^0)] \\ &= -\lambda'_{cf} [\nu_T^2 D_{cf,T} (\sum_{t=\hat{T}_1+1}^{T_1^0} z_{1f,t} z'_{1f,t}) D_{cf,T}] \lambda_{cf} + 2\lambda'_{cf} [\nu_T D_{cf,T} \sum_{t=\hat{T}_1+1}^{T_1^0} z_{1f,t} u_t] + o_p(1) \\ &\Rightarrow -|s| \lambda'_{cf} \begin{pmatrix} 1 & W_z^f(\tau^0)' (\Omega_{zz}^{ff})^{1/2} \\ (\Omega_{zz}^{ff})^{1/2} W_z^f(\tau^0) & (\Omega_{zz}^{ff})^{1/2} W_z^f(\tau^0) W_z^f(\tau^0)' (\Omega_{zz}^{ff})^{1/2} \end{pmatrix} \lambda_{cf} \\ &+ 2\lambda'_{cf} \begin{pmatrix} \sigma W_c(-s) \\ \sigma W_c(-s) (\Omega_{zz}^{ff})^{1/2} W_z^f(\tau^0) \end{pmatrix} \equiv L_1(s), \end{aligned} \quad (\text{A.11})$$

where $W_c(\cdot)$ and $W_z^f(\cdot)$ ($q_f \times 1$) are independent Brownian motions on $[0, \infty)$. Let $G_{0,t} = (1, z'_{f,t}, z'_{b,t})' = (z'_{1f,t}, z'_{b,t})'$ and $D_T = \text{diag}(1, T^{-1/2}I_{q_f}, T^{-1/2}I_{q_b})$. Under $H_{0,B}$, $\lambda_b = 0_{q_b}$, and

$$\begin{aligned} SSR_B(\tau^0) - SSR_B(\hat{\tau}) &= -[\lambda'_{cf}, 0'_{q_b}] (\nu_T^2 D_T \{ \sum_{t=\hat{T}_1+1}^{T_1^0} G_{0,t} G'_{0,t} \} D_T) [\lambda'_{cf}, 0'_{q_b}]' \\ &+ 2[\lambda'_{cf}, 0'_{q_b}] (D_T \nu_T \{ \sum_{t=\hat{T}_1+1}^{T_1^0} G_{0,t} u_t \}) + o_p(1) \Rightarrow L_1(s). \end{aligned} \quad (\text{A.12})$$

Thus, from (A.11) and (A.12), we have

$$SSR_0(\hat{\tau}) - SSR_B(\hat{\tau}) = SSR_0(\tau^0) - SSR_B(\tau^0) + o_p(1)$$

which also holds for the case $\hat{T}_1 > T_1^0$ using a symmetric argument. For $\hat{\sigma}_B^2(\hat{\tau})$, following similar arguments as in the proof of Theorem 1, we can decompose $\hat{\sigma}_B^2(\hat{\tau})$ into its variance and covariance components and adapt the technique used in (A.11) to show that each component

converges to the corresponding component of $\hat{\sigma}_B^2(\tau^0)$. The details are omitted. Combining these results, the proof is complete since

$$\begin{aligned} F_{T,B}^{(2)}(\hat{\tau}) &= \frac{SSR_0(\hat{\tau}) - SSR_B(\hat{\tau})}{\hat{\sigma}_B^2(\hat{\tau})} = \frac{SSR_0(\tau^0) - SSR_B(\tau^0) + o_p(1)}{\hat{\sigma}_B^2(\tau^0) + o_p(1)} \\ &= F_{T,B}^{(2)}(\tau^0) + o_p(1) \xrightarrow{d} \chi^2(q_b) \end{aligned}$$

Table 1: Size and power of the KP and two-step tests ($\times 100$), 5% nominal level

		$T = 120$				$T = 240$			
DGP		1	2	3	4	1	2	3	4
Panel A: Testing for a partial change in the coefficient of an $I(1)$ regressor (δ_t)									
a (i.i.d. errors)	KP	3.66	98.83	50.92	93.44	4.28	100	73.93	99.82
	Two-step	2.29	97.06	8.02	97.39	2.63	100	7.05	100
b (AR(1) errors)	KP	2.80	75.90	10.98	66.83	3.67	98.78	27.21	95.20
	Two-step	1.42	59.27	3.96	61.27	1.79	97.10	6.86	97.48
c (MA(1) errors)	KP	2.70	99.92	65.80	96.82	1.88	100	85.26	99.97
	Two-step	3.92	99.70	4.77	99.74	2.02	100	2.97	100
Panel B: Testing for a partial change in intercept (c_t) with an $I(1)$ regressor									
a (i.i.d. errors)	KP	4.43	81.31	84.72	84.88	4.79	92.64	97.12	93.56
	Two-step	2.31	8.40	59.94	66.24	2.64	6.21	81.53	82.19
b (AR(1) errors)	KP	4.27	44.74	28.86	50.09	4.87	71.55	59.32	73.90
	Two-step	1.70	8.95	11.89	27.95	2.12	9.34	33.49	52.35
c (MA(1) errors)	KP	1.10	88.69	97.27	91.44	1.08	97.74	99.94	98.09
	Two-step	2.98	3.88	81.33	82.66	1.63	2.53	94.76	94.20
Panel C: Testing for a partial change in the coefficient of an $I(0)$ regressor (β_t)									
DGP		5	6	7	8	5	6	7	8
a (i.i.d. errors)	KP	47.25	95.36	33.82	97.92	81.10	97.11	64.79	98.78
	Two-step	6.42	99.86	5.07	99.88	6.17	100	6.33	100
b (AR(1) errors)	KP	38.51	93.27	20.31	97.16	72.86	96.8	44.14	98.08
	Two-step	1.80	68.5	0.33	70.24	3.89	99.67	1.61	99.66
c (MA(1) errors)	KP	48.32	95.73	33.7	98.24	82.59	97.53	62.76	98.84
	Two-step	6.46	99.95	6.07	99.95	5.61	100	5.85	100

Table 2: Rejection rates of the KP and two-step tests for DGP-2b, 3b as a function of the sample size and break magnitude, 5% nominal level

Panel A: $\Delta_c = 1$, $\Delta_\delta = 0.4$, with different T .									
T		60	120	180	240	300	360	480	600
DGP-2b	KP	19.77	44.86	61.29	71.20	77.54	81.72	86.32	88.81
	Final	5.77	9.12	10.03	9.29	8.44	7.77	7.27	6.92
	1 st step	22.81	65.56	89.91	97.85	99.61	99.94	100	100
	2 nd step	25.31	13.91	11.16	9.49	8.47	7.77	7.27	6.92
DGP-3b	KP	5.87	11.14	18.85	27.10	34.55	41.29	52.32	60.16
	Final	2.95	3.87	5.51	6.76	7.71	8.15	8.40	8.26
	1 st step	8.54	15.75	29.72	44.54	56.91	66.55	79.63	86.87
	2 nd step	34.58	24.55	18.53	15.17	13.55	12.25	10.55	9.51
Panel B: $T = 120$, with different break magnitude Δ_δ of the parameter not under test.									
Δ_δ		0.1	0.2	0.3	0.4	0.5	1	1.5	2
DGP-2b	KP	12.12	27.72	39.06	44.60	47.84	44.51	39.44	36.58
	Final	3.16	5.82	7.81	9.01	9.88	8.96	8.10	7.82
	1 st step	8.67	27.78	48.94	65.42	78.23	98.72	99.96	100
	2 nd step	36.46	20.96	15.95	13.77	12.64	9.08	8.11	7.82
DGP-3b	KP	3.29	4.27	5.87	8.31	11.21	26.36	33.64	35.26
	Final	1.53	1.86	2.37	3.10	3.91	7.06	7.62	7.36
	1 st step	2.85	4.18	6.48	10.42	15.73	57.18	85.56	96.32
	2 nd step	53.65	44.38	36.50	29.76	24.87	12.34	8.90	7.64

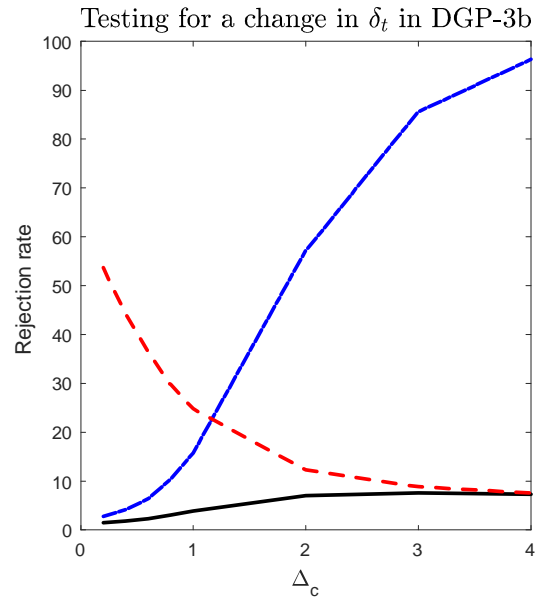
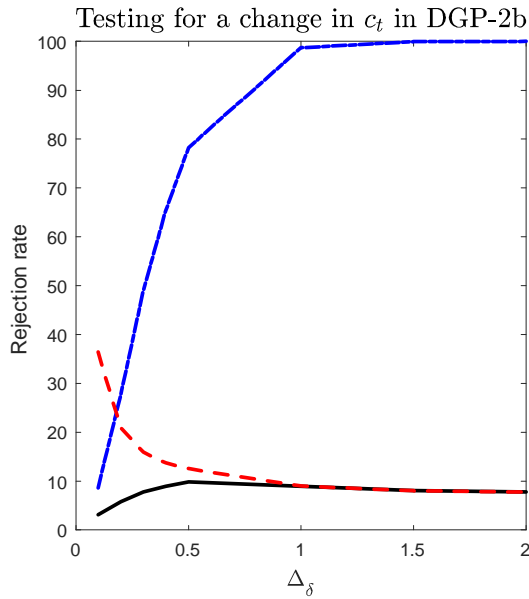
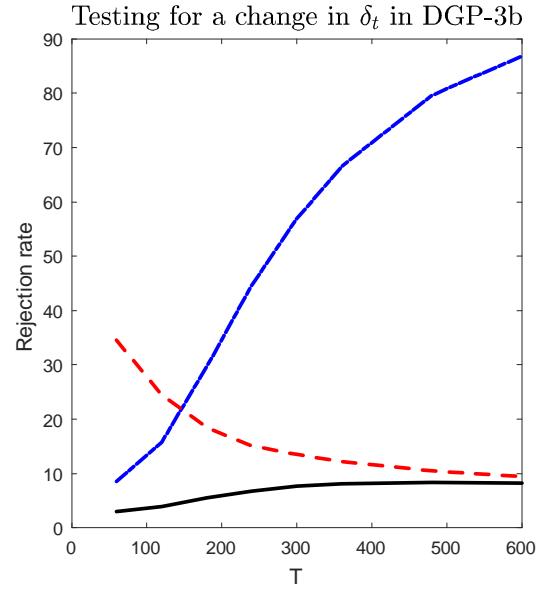
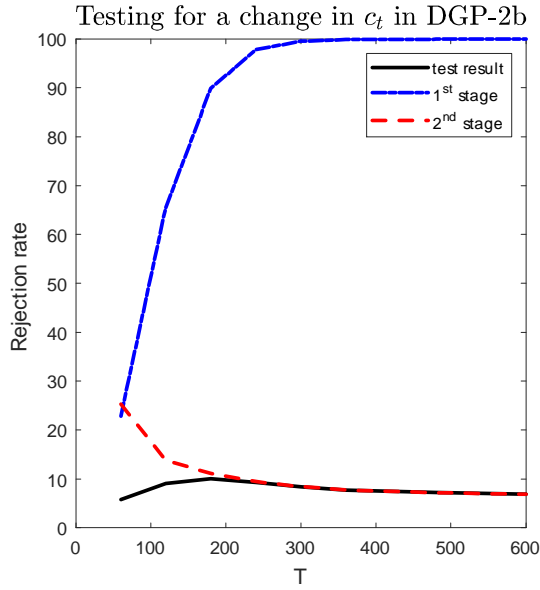


Figure 1: Rejection rates of the two-step test for DGP-2b and DGP-3b as a function of T and the break magnitude.