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**

June 18, 2021

Abstract

This paper studies the problem of testing partial parameter stability in cointegrated regression models. The existing literature considers a variety of models 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). 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. 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 illustrate the usefulness of the procedure in finite samples. In an application to US money demand, we show how the proposed approach can be fruitfully employed to estimate the welfare cost of inflation.

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

JEL Classification: C22

*We thank Robert Taylor (the Editor) and two anonymous referees for their constructive feedback that helped improve the paper in terms of both content and exposition. We are grateful to Luis Filipe Martins for kindly sharing the data used in this paper.

[†]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 paper, 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 an application to US money demand, we show how the proposed approach can be fruitfully employed to estimate the welfare cost of inflation. In particular, we find that the restriction of unitary income elasticity commonly imposed in the literature is not supported by the data with important implications for the trajectory of welfare cost estimates.

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 paper 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 contains the empirical application. Section 6 provides brief concluding remarks. All proofs are provided in Appendix A and additional Monte Carlo simulations are included in Appendix B. As a matter of notation, "$\xrightarrow{p}$" denotes convergence in probability, "$\xrightarrow{d}$" convergence in distribution and "$\Rightarrow$" weak convergence under the Skorohod metric. Further, $O_p(.)$ denotes the stochastic order in its strict sense, i.e., it is not $o_p(.)$.

2 Model and Test Statistics

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

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

for $j = 1, ..., m+1$, ($m+1$ 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: $z_{ft} = z_{f,t-1} + u^f_{zt}$, $z_{bt} = z_{b,t-1} + u^b_{zt}$, for $t = 1, ..., T$, 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 though, when the model contains no break under H_0, as in $H_{0,C}$, the test proposed will be conservative, an issue discussed in more detail in Section 4. 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

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

where $\widetilde{u}_t$ are the residuals from the regression under the null hypothesis and $w(\cdot)$ is a continuous and even function with $|w(.)| \leq 1$, $w(0) = 1$ and $\int_{-\infty}^{\infty} w^2(x)dx < \infty$. KP proposed using the quadratic spectral kernel with the bandwidth chosen via the rule $b_T(\tau) = 1.3221(\hat{a}_2(\tau)T)^{1/5}$ advocated by Andrews (1991), where $\hat{a}_2(\tau) = 4\hat{\rho}(\tau)^2/(1-\hat{\rho}(\tau))^4$, $\hat{\rho}(\tau) = \sum_{t=2}^{T}\hat{u}_t(\tau)\hat{u}_{t-1}(\tau)/\sum_{t=2}^{T}\hat{u}_{t-1}^2(\tau)$, with $\hat{u}_t(\tau)$ the residuals from the regression under the alternative hypothesis. This is a hybrid nonparametric 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 $\xi_t = (u_t, u_{zt}^{f\prime}, u_{zt}^{b\prime})'$, 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): $T^{-1/2}\sum_{t=1}^{[Tr]}\xi_t \Rightarrow B(r)$, with $B(r) = (B_1(r), B_z^f(r)', B_z^b(r)')'$ is a n

vector Brownian motion with symmetric covariance matrix

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

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

•**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 $\gamma_j^0 = (c_j^0, \delta_{fj}^{0\prime}, \delta_{bj}^{0\prime})'$, $j = 1, 2$ and $D_T = diag(1, T^{-1/2} I_{q_f}, T^{-1/2} I_{q_b})$. Then $\gamma_2^0 - \gamma_1^0 = D_T \lambda \upsilon_T$ where $\lambda = (\lambda_c, \lambda_f', \lambda_b')'$ is independent of T and $\upsilon_T > 0$ is a scalar satisfying $\upsilon_T \to 0$ and $T^{1/2}\upsilon_T \to \infty$.

Assumptions A1-A2 are standard in the single equation cointegration literature and the same as in Hansen (1992) and KP. Assumption A2 rules out cointegration among the regressors and implies the presence of a single cointegrating vector between the dependent variable and the regressors. This assumption is standard in the single equation cointegration literature and made in Hansen (1992) and Kejriwal and Perron (2010). It allows us to derive the limit distribution of the structural change tests by ensuring the invertibility of the limiting second moment matrix of the $I(1)$ regressors. Note, however, that our analysis allows both $I(1)$ and $I(0)$ regressors which corresponds to the case where the regressors are cointegrated, albeit trivially so. Monte Carlo evidence in Section 4 illustrates that the proposed two step approach is adequately sized with $I(1)$ and $I(0)$ regressors. Further, in unreported simulations with two $I(1)$ cointegrated regressors, we also confirmed that the two step approach is not subject to size distortions. These results are available upon request.

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 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,* $supF_{T,B}(\tau)$ *is at least* $O_{p}(b_{T}^{-1}(\tau^{0})T)$ *if* $b_{T}(\tau^{0})\nu_{T}^{2}\xrightarrow{p}\infty$*, and at least* $O_{p}(T\nu_{T}^{2})$*, otherwise. (b) If* $\lambda_{f}\neq 0$ *and/or* $\lambda_{b}\neq 0$ *and* $H_{0,C}$ *holds, the same results hold for* $supF_{T,C}(\tau)$*.*

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 $supF_{T,B}(\tau)$ and $supF_{T,C}(\tau)$ are at least $O_{p}(b_{T}^{-1}(\tau^{0})T)$.

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 result stands in stark contrast to that in the standard stationary framework where the limit distribution is invariant to the magnitude of local breaks in parameters not under test (see, e.g., Hsu and Kuan, 2001). The intuition for this invariance is that the omitted break term has the same order of magnitude as the error component and thus does not induce a change in the limit distribution. In our framework, the break magnitude can be fixed or shrink with the sample size at a rate that allows consistent estimation of the break fraction. If the omitted break is on an $I(1)$ regressor, we have the standard spurious regression problem (the effective error is $I(1)$). If the omitted break is on an $I(0)$ regressor with a non-zero mean, then the partial sums of the effective error involve a broken deterministic trend thereby again leading to a spurious regression type problem (see Perron, 1990). In either scenario, it follows from standard results that the sum of squared residuals under both the null and the alternative diverge as does their difference. Since the denominator of the F-statistic is of a lower order of magnitude than the numerator, the test statistic diverges. 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 preceding analysis 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 coefficients 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 $F^{(2)}_{T,i}(\hat{\tau})$ denotes the second step test of the null hypothesis $H_{0,i}$ $[i = B, C]$.

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

Remark 1 *In the first step, we could, in principle, replace the pure structural change test by any partial structural change test, whether or not it involves the regressors whose coefficients are subject to change. We investigate the benefits of this potential modification via simulations in Section 4 and conclude that using the pure structural change test is overall preferable.*

Using Theorems 1 and 2, we can show that the asymptotic size of the two step procedure cannot exceed α, where α is the level of significance used in each step. This is stated in the following corollary.

Corollary 1 *Let $cv_A(\alpha)$ and $cv_i(\alpha)$ denote the level α asymptotic critical values of $\sup F_{T,A}(\tau)$ and $F^{(2)}_{T,i}(\hat{\tau})$, respectively ($i = B, C$). Then, under $H_{0,i}$, $i \in \{B, C\}$, we have*

$$\lim_{T\to\infty} P\left[\{\sup F_{T,A}(\tau) > cv_A(\alpha)\} \cap \left\{F^{(2)}_{T,i}(\hat{\tau}) > cv_i(\alpha)\right\}\right] \leq \alpha$$

Remark 2 *The two step procedure can be applied to a model with multiple breaks where each break affects only a subset of the parameters. Such a model can be represented as a restricted version of the pure structural change model* (1)*. Given the consistency of the first step test as well as the estimated break fractions in the presence of a change in at least one of the parameters, the second step can be used to determine the break(s) that affect a particular parameter of interest by testing the constancy of this parameter across any two adjacent regimes.*

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). For the errors u_t, we consider three different cases: (a) (i.i.d. errors) $u_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$; (b) (AR(1) errors) $u_t = 0.5u_{t-1} + e_t$, $e_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$; (c) (MA(1) errors) $u_t = e_t - 0.5e_{t-1}$, $e_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. The trimming ϵ is set at 15%. Each step of the two step procedure as well as the one step partial KP test uses a 5% nominal level test. The number of replications throughout is 100,000.

In the first set of simulations, the dependent variable y_t is generated by: $y_t = c_t + \delta_t z_t + u_t$; $z_t = z_{t-1} + u_{zt}$, $u_{zt} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. 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 $\delta_t = 1 + \Delta_\delta$. 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. We compare the size and power of the one step 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 (Panels A,B) 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. The power results are size-unadjusted. 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 8%. 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 in Appendix B. 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 (though conservative) in all cases, while the one step KP test is subject to substantial size distortions under DGP-2 (a change in the slope parameter). It is instructive to look at the cases of DGP-1 and DGP-2 in more details. DGP-1 involves no breaks. Hence, the KP test has a 5% asymptotic size and should have highest power, while the two-step procedure is asymptotically conservative. Under the null, for the 5% of the cases in which a rejection does occur, the second step rejects with some probability less than one, given that the estimated break date from the first step is random, so that the limit distribution in Kejriwal and Perron (2008a) does not apply. Still, the power of the two-step procedure remains adequate. By definition, it is less than that of the first step KP test (subject to some simulation errors), but the reduction in power is quite minor; the biggest discrepancy is for AR(1) errors with $T = 120$. Notwithstanding its two step nature, our recommended procedure retains respectable power that increases noticeably with T.

In a second set of simulations, 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 $x_t \overset{i.i.d.}{\sim} \mathcal{N}(1,1)$, $z_t = z_{t-1} + u_{zt}$ and $u_{zt} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. 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 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 final set of simulations considers DGPs with a larger number of regressors. These simulations are motivated by the observation that since the first step of the two step procedure entails applying a structural change test on all model parameters in the first step, its power may be low with a large number of regressors if only a few parameters change.[1] To investigate this possibility, two alternative designs are considered. The first appends DGPs 1-4 with two $I(1)$ regressors generated as independent random walks (and independent of the other variables) with unit coefficients that remain stable throughout the sample. We label these DGPs $1'$-$4'$. The second design appends DGPs 5-8 with one $I(1)$ regressor generated as a random walk and one $I(0)$ regressor generated as *i.i.d.* $N(0, 1)$, the variables being independent of each other as well as the other variables. Each of the two variables have unit coefficients that remain stable throughout the sample. We label these DGPs $5'$-$8'$.

In addition to the proposed two step procedure, we also consider here a modified procedure that replaces the pure structural change test in the first step with a partial structural change test on the coefficient of interest. As shown in Theorem 1, the partial test has unit asymptotic power under the alternative hypothesis that at least one of the parameters change. The first step and final rejection frequencies of the modified procedure are denoted by TS^i_{1st} and TS^i, respectively, while those of the proposed two step procedure are denoted by TS_{1st} and TS, respectively. The objective is to examine the extent of power loss involved by conducting a pure structural change in the first step as opposed to a partial structural change test on the coefficient of interest. A 5% significance level is used in each step of the two procedures.

The results are reported in Table 2. The performance of the proposed and modified

[1] We thank an anonymous referee for pointing this out.

procedures in terms of finite sample size are broadly similar to each other and adequate in both cases. In terms of power, the modified procedure offers discernible improvements only in the case of AR(1) errors when the sample size is small ($T = 120$). In other cases, the improvements are marginal at best. On the other hand, while the proposed approach is simple to implement in practice, the modified approach is computationally costly - with k parameters, the proposed approach only requires $k + 1$ tests while the modified approach requires $2k$ tests. This feature also makes the modified approach more susceptible to multiple testing issues. Thus, the proposed approach can serve as a simple, yet useful addition to the practitioner's toolkit when testing for partial structural change.

Appendix B contains additional Monte Carlo results that explore the behavior of empirical size as a function of sample size/break magnitude. These simulations are motivated by the fact that the size distortions incurred by the two step procedure for DGPs 2b-3b in Table 1 do not decrease as the sample size increases. We show that the proposed procedure remains adequate as the sample size/break magnitude varies and provide a discussion for the observed evolution of the empirical size. See Appendix B for details.

5 Empirical Application

This section applies the proposed procedure to study the stability of US money demand and the associated issue of estimating the welfare cost of inflation. A standard approach to measuring the welfare cost of inflation is due to Bailey (1956), who suggests that such costs can be measured as the area underlying the inverse money demand function that represents the consumer surplus that could be realized from a reduction in the nominal interest rate from a positive level to a near-zero level. The rationale is that if real money balances are treated as a consumption good due to its ability to provide liquidity, inflation can be viewed as a tax on real balances through its effect on nominal interest rates and hence the opportunity cost of holding real balances. In such a framework, the specification of the money demand function naturally plays a crucial role in the estimation of the welfare cost of inflation.

Empirical work in this context has typically relied on two alternative functional forms for the money demand function: the log-log form (Meltzer, 1963) and the semi-log form (Cagan, 1956). The log-log form allowing for time-varying parameters is specified as follows:

$$\ln\left(M/P\right)_t = a_t + c_t \ln\left(Y/P\right)_t + b_t \ln r_t + u_t \tag{2}$$

where M/P denotes real money balances, Y/P denotes real income and r denotes the nominal interest rate. The parameters c_t and b_t measure the time-varying income and interest elasticities of money demand, respectively. The semi-log form is specified as follows:

$$\ln (M/P)_t = a_t + c_t \ln (Y/P)_t + b_t r_t + u_t \tag{3}$$

In addition to (2) and (3), the following specifications that impose a unitary elasticity of money demand are frequently estimated, where m denotes the money-income ratio M/Y:

$$\ln m_t = a_t + b_t \ln r_t + u_t \tag{4}$$

$$\ln m_t = a_t + b_t r_t + u_t \tag{5}$$

The application of Bailey's method to (2) yields the following measure of the welfare cost at a positive interest rate r (see, e.g., Calza and Zaghini, 2010):

$$w(r,t) = \exp(a_t)(Y/P)^{c_t - 1} \frac{-b_t}{1+b_t} r^{1+b_t} \tag{6}$$

while the expression for the semi-log function (3) takes the form:

$$w(r,t) = \frac{\exp(a_t)}{-b_t}(Y/P)^{c_t - 1}[1 - (1 - b_t r)\exp(b_t r)] \tag{7}$$

Equations (6) and (7) are typically evaluated at the average value of (Y/P) over the sample. The corresponding expressions for the restricted specifications (4) and (5) are given as follows:

$$w(r,t) = \exp(a_t) \frac{-b_t}{1+b_t} r^{1+b_t} \tag{8}$$

$$w(r,t) = \frac{\exp(a_t)}{-b_t}[1 - (1 - b_t r)\exp(b_t r)] \tag{9}$$

A wide range of welfare cost estimates is available in the literature depending on whether a log-log or semi-log form is estimated, whether instability is allowed for, as well as whether a unitary income elasticity is imposed. We do not undertake a comprehensive review of this literature here but rather focus on the studies that are more closely related to ours (see Mogliani and Urga, 2018 and Miller et al., 2019, for further references and discussions).

Lucas (2000) argues in favor of a log-log form due to its consistency with inventory-theoretic money demand models, while Ireland (2009) advocates the use of a semi-log specification for post-1980 data due to a shift in the monetary policy regime towards low interest rates. While the former study finds a welfare cost of 10% inflation to be slightly less than

1% of GDP, the latter reports a much lower estimate of about 0.25%. The estimates in both studies are obtained from specifications that assume a unit income elasticity of money demand. Mogliani and Urga (2018) investigate the stability of the log-log form of the money demand function using the joint KP test but again imposes a unitary income elasticity. They find evidence of two breaks (1945 and 1976) and a welfare cost of about 0.1% in the post-1976 period compared to 0.8% over 1945-75. Miller et al. (2019) adopt the time-varying cointegration framework developed in Bierens and Martins (2010) where the coefficients are modeled as smooth functions of time. They find that the unitary income elasticity restriction is rejected for the semi-log form but not for the log-log form. They report that the welfare cost estimate of 10% inflation lies in the range 0.025-0.75% of GDP, with an average of about 0.27% over the sample. In what follows, we will employ the proposed two step approach to test whether the unitary restriction on income elasticity is supported by US data and obtain estimates of the regime-dependent welfare costs accordingly. We do not make an a priori choice between the log-log and semi-log forms and present results for both forms.

Our empirical analysis employs the same dataset as Miller et al. (2019).[2] The data are quarterly and span the period 1959:Q1-2010:Q4. The preliminary unit root analysis in Miller et al. (2019) confirms the presence of a unit root in $\ln(M/P)$, $\ln(Y/P)$ and r. The first step of the two step procedure is implemented using KP's $UDmax$ test that entails taking the maximum of the sup-Wald statistics which allow for one up to five breaks with the trimming level set at 15%. If the $UDmax$ test rejects, the number of breaks is determined using the sequential procedure proposed by KP. Endogeneity of the regressors is accounted for using four leads/lags of the first-differenced regressors while serial correlation is accounted for using a heteroskedasticity and autocorrelation consistent estimate of the long-run variance based on KP's hybrid method using a quadratic spectral kernel with Andrews' (1991) data-dependent bandwidth choice. Given that the first step KP test is consistent against a purely spurious regression, we complement our analysis by testing for the presence of cointegration using Arai and Kurozumi's (2007, AK henceforth) LM-type test for the null hypothesis of cointegration with breaks against the alternative of no cointegration. While AK's test allowed for a single break, its multiple break extension was developed by Kejriwal (2008).

The testing results are presented in Table 3. Panel A reports the findings for the unrestricted models while Panel B reports those for the models that impose a unitary income

[2] The data come from the Federal Reserve Bank of St. Louis' FRED database and consists of a measure of money supply (M1) adjusted for deposit sweep programs, nominal GDP (Y), the three-month US Treasury Bill rate (r) and the GDP deflator (P). See Miller et al. (2019) for further details.

elasticity. The two step results indicate instability in all coefficients for all models except b_t in the unrestricted semi-log model. These results therefore reject the unit elasticity restriction that is often assumed in the literature. The AK test indicates the presence of cointegration regardless of the maintained functional form.

Table 4 reports the results from estimating the models selected by the two step procedure in Table 3. The regime-wise point estimates along with 95% confidence intervals are presented along with the mean welfare cost of 0%, 2% and 10% inflation. The confidence intervals around the welfare cost estimates are obtained using a wild bootstrap procedure using 999 replications. Several features of the results are worth noting. First, the unrestricted model selects a single break in 1993 for both functional forms. In contrast, the restricted models select two or three breaks depending on whether the log-log or semi-log specification is adopted. For the log-log form, Mogliani and Urga (2018) also find evidence of two breaks with the second break located in 1976. Second, the unit income elasticity restriction is rejected in both regimes irrespective of the adopted functional form. Both forms entail a reduction in income elasticity from about 0.5 to -0.3. The pre-break estimate of the income elasticity is consistent with the Baumol-Tobin inventory theoretic approaches to the transactions demand for money as well as empirical findings in Ball (2001). The post-break negative income elasticity, on the other hand, is in conformity with the constant target-threshold monitoring model of Akerlof and Milbourne (1980) in which money is transferred into or out of the account if the cash balance crosses a lower or upper threshold level. The interest elasticity estimate is also lower in more recent decades, consistent with recent findings in Berentsen et al. (2015) and Miller et al. (2019). Third, the imposition of unitary income elasticity leads to upward biased estimates of the welfare cost of inflation, regardless of the choice of functional form. For instance, the mean welfare cost of 10% inflation implied by the unrestricted model is 0.39% and 0.42% for the log-log and semi-log forms, respectively, compared to 0.46% and 0.52% for the restricted model. Fourth, the welfare cost estimates are lower in more recent periods, in accordance with the findings in Berentsen et al. (2015) and Mogliani and Urga (2018). Fifth, compared to the log-log form, the semi-log form tends to produce larger estimates of the welfare cost at the 10% inflation level and smaller estimates at the 0% and 2% levels.

To further investigate the relevance of the unitary elasticity restriction, we re-estimate the restricted model using the break date estimate obtained from the unrestricted model. This allows us to evaluate the impact of imposing the restriction on the estimated welfare costs while controlling for break date misspecification. The results, presented in Table 5, show that the mean restricted welfare cost estimates are now even higher than those reported in Table 4 indicating that the two forms of misspecification have offsetting effects on the magnitude of the welfare cost, with elasticity misspecification inducing a positive effect and break date misspecification inducing a negative effect.

The finding of a break in 1993 in the unrestricted case can be explained by the advancement of information technology with respect to financial products during the post-1993 period and its impact on the demand for money. In particular, the introduction of "sweep technology" in 1993 allowed banks to automatically transfer funds from checking accounts to money market deposit accounts (MMDAs) in which the holder was permitted to make only a few withdrawals every month. This deposit-sweeping software reduced the reserve requirements of banks making more funds available for lending and providing improved access to the money market. Berentsen et al. (2015) construct a microfounded monetary model that can be used to assess the impact of a one-time increase in the access probability to the money market in the early 90s. When calibrated to US data, their model is able to replicate the empirical behavior of the money demand function well suggesting an important role of the sweep technology in explaining the observed changes in money demand.

In summary, the empirical results do not support the typically assumed restriction of unitary income elasticity for both the log-log and semi-log functional forms. Imposing such a restriction leads to overestimating the number of breaks in the money demand relationship as well as the welfare cost of inflation. The unrestricted models point to the prevalence of a single break that can be attributed to technological innovations in the financial sector during the early 90s that affected the demand for money.

6 Conclusion

This paper 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. An application to estimating the welfare cost of US inflation illustrates the relevance of the procedure in practice.

References

Akerlof GA, Milbourne RD. 1980. The short run demand for money. *The Economic Journal* **90**: 885-900.

Andrews DWK. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**: 817-858.

Arai Y, Kurozumi E. 2007. Testing for the null hypothesis of cointegration with a structural break. *Econometric Reviews* **26**: 705-739.

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

Bailey MJ. 1956. The welfare cost of inflationary finance. *Journal of Political Economy* **64**: 93-110.

Ball L. 2001. Another look at long-run money demand. *Journal of Monetary Economics* **47**: 31-44.

Berentsen A, Huber S, Marchesiani A. 2015. Financial innovations, money demand, and the welfare cost of inflation. *Journal of Money, Credit and Banking* **47**: 223-261.

Bierens HJ, Martins LF. 2010. Time-varying cointegration. *Econometric Theory* **26**: 1453-1490.

Cagan P. 1956. The monetary dynamics of hyperinflation. (in M. Friedman ed.) *Studies in the Quantity Theory of Money* Chicago: Chicago University Press.

Calza A, Zaghini A. 2010. Sectoral money demand and the great disinflation in the United States. *Journal of Money, Credit and Banking* **42**: 1663-1678.

Hansen, BE. 1992. Tests for parameter instability in regressions with I (1) processes. *Journal of Business & Economic Statistics* **10**: 321-335.

Hsu CC. 2008. A note on tests of partial parameter stability in the cointegrated system. *Economics Letters* **99**: 500-503.

Hsu CC, Kuan CM. 2001. Distinguishing between trend-break models: Method and empirical evidence. *Econometrics Journal* **4**: 171-190.

Ireland PN. 2009. On the welfare cost of inflation and the recent behavior of money demand. *American Economic Review* **99**: 1040-52.

Kejriwal M. 2008. Cointegration with structural breaks: An application to the Feldstein-Horioka puzzle. *Studies in Nonlinear Dynamics & Econometrics* **12**: 1-37.

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 BS. 1998. Test for partial parameter instability in regressions with I(1) processes. *Journal of Econometrics* **86**: 337-368.

Lucas RE. 2000. Inflation and welfare. *Econometrica* **68**: 247-274.

Meltzer AH. 1963. The demand for money: The evidence from the time series. *Journal of Political Economy* **71**: 219-246.

Miller SM, Martins LF, Gupta R. 2019. A time-varying approach of the US welfare cost of inflation. *Macroeconomic Dynamics* **23**: 775-797.

Mogliani M, Urga G. 2018. On the instability of long-run money demand and the welfare cost of inflation in the United States. *Journal of Money, Credit and Banking* **50**: 1645-1660.

Perron P. 1990. Testing for a unit root in a time series with a changing mean. *Journal of Business & Economic Statistics* **8**: 153-162.

Saikkonen P. 1991. Asymptotically efficient estimation of cointegration regressions. *Econometric Theory* **7**: 1-21.

Table 1: Size and power of the one-step partial KP and two-step (TS) tests (×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
	TS	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
	TS	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
	TS	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
	TS	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
	TS	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
	TS	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
	TS	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
	TS	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
	TS	6.46	99.95	6.07	99.95	5.61	100	5.85	100

Table 2: Size and power of two-step (TS and TS^i) tests ($\times 100$), 5% nominal level, larger number of regressors

		$T=120$				$T=240$			
	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)	TS_{1st}	4.51	93.53	35.25	94.56	4.59	99.97	70.1	100
	TS	2.25	89.83	6.81	90.91	2.14	99.95	7.52	99.99
	TS^i_{1st}	5.29	97.53	30.8	94.97	5.73	99.99	53.15	99.92
	TS^i	2.38	93.17	5.39	91.25	2.38	99.97	5.83	99.91
b (AR(1) errors)	TS_{1st}	2.32	44.94	5.76	47.64	2.83	92.26	13.76	92.39
	TS	1.13	38.32	2.1	39.99	1.22	90.89	3.67	91
	TS^i_{1st}	5.15	72.98	9.97	68.76	5.57	98.16	17.57	96.09
	TS^i	2.18	60.09	3.21	56.62	2.32	96.27	4.23	94.34
c (MA(1) errors)	TS_{1st}	18.36	98.77	62.38	99.04	8.96	100	95	100
	TS	12.09	97.9	11.05	98.21	5.53	100	5.8	100
	TS^i_{1st}	5.72	99.83	45.72	98.34	3.47	100	71.53	100
	TS^i	4.00	98.87	7.36	97.47	2.08	100	4.48	100
Panel B: Testing for a partial change in intercept (c_t) with an $I(1)$ regressor									
a (i.i.d. errors)	TS_{1st}	4.53	93.78	34.83	94.73	5.12	99.99	69.07	100
	TS	1.95	9.64	16.8	29.7	1.95	6.96	37.22	42.98
	TS^i_{1st}	5.01	77.05	54.98	80.06	5.92	93.19	82.79	94.47
	TS^i	1.98	7.64	24.07	24.02	2.11	6.52	41.46	39.99
b (AR(1) errors)	TS_{1st}	2.55	46.15	5.25	46.9	2.78	92.53	12.92	92.37
	TS	1.33	8.26	2.82	11.7	1.29	10.73	6.89	22.35
	TS^i_{1st}	5.53	41.47	15.96	44.63	5.79	68.69	31.76	70.86
	TS^i	2.54	6.55	7.32	10.43	2.35	7.69	13.94	16.34
c (MA(1) errors)	TS_{1st}	19.43	98.53	62.68	98.93	8.74	100	95.3	100
	TS	10.61	8.81	35.11	43.87	4.64	4.62	66.73	66.33
	TS^i_{1st}	3.51	85.15	82.53	87.62	2.14	97.9	98.91	98.1
	TS^i	1.54	7.35	42.6	38.25	0.9	4.48	68.36	64.73
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)	TS_{1st}	81.03	99.83	36.09	99.91	99.44	100	72.09	100
	TS	6.82	99.83	4.66	99.91	6.39	100	6.08	100
	TS^i_{1st}	45.65	98.4	24.4	99.3	79.64	98.78	47.25	99.56
	TS^i	4.45	98.4	3.88	99.3	5.27	98.78	4.39	99.56
b (AR(1) errors)	TS_{1st}	10.06	69.09	2.59	68.6	36.72	99.23	5.75	99.22
	TS	0.93	69.09	0.42	68.6	2.56	99.23	0.95	99.22
	TS^i_{1st}	37.91	97.47	19.99	99.12	72.75	98.93	37.24	99.41
	TS^i	3.86	97.47	3.11	99.12	4.81	98.93	4.03	99.41
c (MA(1) errors)	TS_{1st}	82.55	99.95	40.47	99.92	99.67	100	83.49	100
	TS	6.75	99.95	5.74	99.92	6.15	100	6.61	100
	TS^i_{1st}	42.97	98.32	21.92	99.41	79.3	98.88	43.65	99.51
	TS^i	4.13	98.32	4.45	99.41	5.13	98.88	4.21	99.51

Table 3: Two-step procedure & Arai-Kurozumi (AK) test results, 5% significance level

	Log-Log form	critical values	Semi-Log form	critical values
Panel A: Unrestricted model				
AK Test	0.056	0.115	0.062	0.110
First	423.03	14.47	511.15	14.47
Second : a_t	6.86	3.84	7.33	3.84
Second : c_t	7.03	3.84	7.05	3.84
Second : b_t	5.02	3.84	1.25	3.84
Panel B: Restricted model				
AK Test	0.090	0.147	0.068	0.095
First	613.14	12.25	674.48	12.25
Second : a_t	6.09	5.99	19.03	7.81
Second : b_t	13.83	5.99	13.24	7.81

Table 4: Welfare cost of inflation estimates, with/without restriction, log-log/semi-log form, 95% confidence interval (C.I.)

Regime	Dates	$\hat{a}_t$	$\hat{c}_t$	$\hat{b}_t$	Welfare cost (%)		
					0%	2%	10%
		$C.I.(\hat{a}_t)$	$C.I.(\hat{c}_t)$	$C.I.(\hat{b}_t)$	$C.I.$	$C.I.$	$C.I.$
Panel A1: Unrestricted model, log-log form							
1	1959:Q1-1993:Q3	1.40	.56	-.24	.20	.42	.61
		[.51,2.29]	[.47,.64]	[-.30,-.18]	[.18,.22]	[.39,.46]	[.57,.66]
2	1993:Q4-2010:Q4	9.96	-.33	-.11	.05	.11	.17
		[8.93,10.98]	[-.46,-.21]	[-.15,-.07]	[.04,.05]	[.09,.13]	[.14,.20]
mean	-	5.68	.11	-.18	.12	.27	.39
	-	[5.36,5.99]	[.07,.15]	[-.19,-.16]	[.11,.13]	[.25,.29]	[.36,.42]
Panel A2: Restricted model, log-log form							
1	1959:Q1-1968:Q3	-2.51	-	-.32	.34	.67	.93
		[-2.57,-2.46]	-	[-.33,-.30]	[.30,.39]	[.60,.74]	[.84,1.02]
2	1968:Q4-1976:Q1	-1.97	-	-.11	.08	.19	.29
		[-2.15,-1.79]	-	[-.16,-.06]	[.03,.14]	[.08,.31]	[.13,.46]
3	1976:Q2-2010:Q4	-2.12	-	-.08	.04	.10	.16
		[-2.20,-2.04]	-	[-.11,-.05]	[.04,.05]	[.09,.11]	[.14,.17]
mean	-	-2.20	-	-.17	.15	.32	.46
	-	[-2.27,-2.13]	-	[-.19,-.15]	[.13,.18]	[.27,.37]	[.40,.52]
Panel B1: Unrestricted model, semi-log form							
1	1959:Q1-1993:Q2	2.58	.53	-3.73	.04	.22	.52
		[2.05,3.12]	[.46,.59]	[-4.05,-3.42]	[.03,.04]	[.21,.24]	[.49,.55]
2	1993:Q3-2010:Q4	9.93	-.28	-3.73	.02	.14	.32
		[8.99,10.87]	[-.38,-.18]	[-4.05,-3.42]	[.02,.02]	[.13,.14]	[.30,.34]
mean	-	6.25	.12	-3.73	.03	.18	.42
	-	[5.96,6.56]	[.09,.16]	[-3.98,-3.50]	[.03,.03]	[.17,.19]	[.40,.44]
Panel B2: Restricted model, semi-log form							
1	1959:Q1-1968:Q2	-1.13	-	-8.94	.11	.58	1.17
		[-1.16,-1.09]	-	[-9.72,-8.15]	[.10,.12]	[.53,.64]	[1.09,1.25]
2	1968:Q3-1975:Q4	-1.50	-	-2.36	.02	.15	.36
		[-1.54,-1.46]	-	[-3.28,-1.43]	[.01,.03]	[.08,.22]	[.20,.51]
3	1976:Q1-1983:Q3	-1.71	-	-2.24	.02	.12	.28
		[-1.77,-1.65]	-	[-2.86,-1.63]	[.01,.02]	[.10,.14]	[.24,.33]
4	1983:Q4-2010:Q4	-1.76	-	-2.28	.02	.11	.27
		[-1.78,-1.74]	-	[-2.57,-1.99]	[.02,.02]	[.10,.12]	[.25,.29]
mean	-	-1.53	-	-3.96	.04	.24	.52
	-	[-1.55,-1.50]	-	[-4.36,-3.55]	[.04,.05]	[.22,.26]	[.47,.57]

Table 5: Welfare cost of inflation estimates, restricted model imposed with break dates estimated from the unrestricted model, log-log/semi-log form, 95% confidence interval (C.I.)

Regime	Dates	$\hat{a}_t$	$\hat{c}_t$	$\hat{b}_t$	Welfare cost (%)		
					0%	2%	10%
		$C.I.(\hat{a}_t)$	$C.I.(\hat{c}_t)$	$C.I.(\hat{b}_t)$	$C.I.$	$C.I.$	$C.I.$
Panel A: Restricted model, imposed with break dates estimated from unrestricted model, log-log form							
1	1959:Q1-1993:Q3	-2.94	-	-.42	.49	.88	1.16
		[-3.24,-2.64]	-	[-.54,-.29]	[.41,.60]	[.76,1.02]	[1.03,1.33]
2	1993:Q4-2010:Q4	-2.34	-	-.14	.08	.19	.28
		[-2.70,-1.98]	-	[-.25,-.04]	[.06,.10]	[.14,.23]	[.22,.35]
mean	-	-2.64	-	-.28	.29	.53	.72
	-	[-2.72,-2.56]	-	[-.31,-.26]	[.25,.34]	[.47,.61]	[.65,.81]
Panel B: Restricted model, imposed with break dates estimated from unrestricted model, semi-log form							
1	1959:Q1-1993:Q2	-1.31	-	-6.87	.07	.42	.89
		[-1.61,-1.00]	-	[-10.37,-3.36]	[.06,.09]	[.36,.48]	[.78,.99]
2	1993:Q3-2010:Q4	-1.70	-	-3.88	.03	.19	.43
		[-1.75,-1.65]	-	[-5.33,-2.44]	[.02,.04]	[.15,.22]	[.36,.50]
mean	-	-1.50	-	-5.38	.05	.30	.66
	-	[-1.54,-1.47]	-	[-6.10,-4.70]	[.04,.06]	[.26,.34]	[.59,.73]

Appendix A: Proofs

For any matrix $W_{T\times q} = (w_1, ..., w_T)'$, 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., $\bar{W} = diag(W_1, W_2)$, where $W_i = (w_{T_{i-1}+1}, ..., w_{T_i})'$ $(i = 1, 2)$ with $T_0 = 0$ and $T_2 = T$. Also, let $Y = (y_1, ..., y_T)'$, $\iota_{T\times 1} = (1, ..., 1)'$, $u = (u_1, ..., u_T)'$, $Z_f = (z_{f,1}, ..., z_{f,T})'$, $Z_b = (z_{b,1}, ..., z_{b,T})'$.

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

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

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

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

Let $G_0 = [\iota, Z_f, Z_b] = (G_{0,1}, ..., G_{0,T})'$. We first derive the limiting behavior of the statistic $F_{T,B}(\tau^0)$. The restricted sum of squared residuals is

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

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

$$T^{-1/2}\nu_T^{-1} J_T G_0' \eta = \textstyle\sum_{t=1}^T J_T G_{0,t} T^{-1/2} \nu_T^{-1} \eta_t = T^{-1} \sum_{t=[\tau^0 T]+1}^T \{(T^{1/2} J_T) G_{0,t}\}\{\nu_T^{-1}\eta_t\}$$

$$\Rightarrow (\textstyle\int_{\tau^0}^1 (\lambda_c + \lambda_f' B_z^f(r))dr, \int_{\tau^0}^1 B_z^{f\prime}(r)(\lambda_c + \lambda_f' B_z^f(r))dr, \int_{\tau^0}^1 B_z^{b\prime}(r)(\lambda_c + \lambda_f' B_z^f(r))dr)' = O_p(1).$$

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

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

Let $G_1 = [\iota, Z_f, \bar{Z}_b]$. The unrestricted sum of squared residuals evaluated at τ^0 is

$$SSR_B(\tau^0) = \textstyle\sum_{t=1}^T \hat{u}_t^2 = u'u + 2u'\eta + \eta'\eta - (u' P_{G_1} u + 2u' P_{G_1}\eta + \eta' P_{G_1}\eta) \tag{A.4}$$

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:

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

For the long-run variance estimate $\hat{\sigma}_B^2(\tau^0)$, we have, denoting $b_T^0 = b_T(\tau^0)$,

$$\begin{aligned}
\hat{\sigma}_B^2(\tau^0) &= T^{-1}\textstyle\sum_{t=1}^T \widetilde{u}_t^2 + 2\sum_{j=1}^{T-1} w(j/b_T^0)T^{-1}\sum_{t=j+1}^T \widetilde{u}_t\widetilde{u}_{t-j} && \text{(A.6)} \\
&= [T^{-1}\textstyle\sum_{t=1}^T u_t^2 + O_p(\nu_T^2)] + 2\sum_{j=1}^{T-1} w(j/b_T^0)T^{-1}\sum_{t=j+1}^T u_t u_{t-j} + O_p(b_T^0\nu_T^2) && \text{(A.7)} \\
&= T^{-1}\textstyle\sum_{t=1}^T u_t^2 + 2\sum_{j=1}^{T-1} w(j/b_T^0)T^{-1}\sum_{t=j+1}^T u_t u_{t-j} + o_p(1) + O_p(b_T^0\nu_T^2) && \text{(A.8)} \\
&= \sigma^2 + O_p(b_T^0\nu_T^2)
\end{aligned}$$

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, $\widetilde{u}_{t-j} = (u_{t-j} + \eta_{t-j}) - G'_{0,t-j}(G'_0G_0)^{-1}G'_0(u+\eta)$. It follows that

$$\begin{aligned}
T^{-1}\textstyle\sum_{t=j+1}^T \widetilde{u}_t\widetilde{u}_{t-j} &= T^{-1}\textstyle\sum_{t=j+1}^T [(u_t+\eta_t) - G'_{0,t}(G'_0G_0)^{-1}G'_0(u+\eta)][(u_{t-j}+\eta_{t-j}) \\
&\quad - G'_{0,t-j}(G'_0G_0)^{-1}G'_0(u+\eta)] \\
&= T^{-1}\textstyle\sum_{t=j+1}^T [u_tu_{t-j} + u_t\eta_{t-j} + \eta_t u_{t-j} + \eta_t\eta_{t-j} \\
&\quad - (u_t+\eta_t)G'_{0,t-j}(G'_0G_0)^{-1}G'_0(u+\eta) - (u_{t-j}+\eta_{t-j})G'_{0,t}(G'_0G_0)^{-1}G'_0(u+\eta) \\
&\quad + (u+\eta)'G_0(G'_0G_0)^{-1}G_{0,t}G'_{0,t-j}(G'_0G_0)^{-1}G'_0(u+\eta)] \\
&= T^{-1}\textstyle\sum_{t=j+1}^T u_tu_{t-j} + T^{-1}[O_p(T^{1/2}\nu_T) + O_p(T^{1/2}\nu_T) + O_p(T\nu_T^2) \\
&\quad + O_p(T\nu_T^2) + O_p(T\nu_T^2) + O_p(T\nu_T^2)] \\
&= T^{-1}\textstyle\sum_{t=j+1}^T u_tu_{t-j} + O_p(\nu_T^2), \ \text{uniformly in } j. && \text{(A.9)}
\end{aligned}$$

Using $(b_T^0)^{-1}\sum_{j=1}^{T-1}|w(j/b_T^0)| \to \int_0^{+\infty}|w(x)|dx < \infty$ (e.g. Andrews, 1991), we have from $(A.9)$,

$$\begin{aligned}
& (b_T^0\nu_T^2)^{-1}\{\textstyle\sum_{j=1}^{T-1} w(j/b_T^0)T^{-1}\sum_{t=j+1}^T \widetilde{u}_t\widetilde{u}_{t-j} - \sum_{j=1}^{T-1} w(j/b_T^0)T^{-1}\sum_{t=j+1}^T u_tu_{t-j}\} \\
&\leq (b_T^0)^{-1}\textstyle\sum_{j=1}^{T-1}\left|w(j/b_T^0)\right| \sup_{j\geq 1} |v_T^{-2}T^{-1}\sum_{t=j+1}^T \widetilde{u}_t\widetilde{u}_{t-j} - v_T^{-2}T^{-1}\sum_{t=j+1}^T u_tu_{t-j}| \\
&= [(b_T^0)^{-1}\textstyle\sum_{j=1}^{T-1}|w(j/b_T^0)|]O_p(1) = O_p(1), \quad \text{(A.10)}
\end{aligned}$$

which establishes $(A.8)$.

Combining the results of the numerator (A.5) and the denominator (A.8) of the statistic $F_{T,B}(\tau^0)$, we have: $F_{T,B}(\tau^0) = O_p(T\nu_T^2)/[\sigma^2 + O_p(b_T^0\nu_T^2)] = O_p((b_T^0)^{-1}T)$ if $b_T^0\nu_T^2 \xrightarrow{p} \infty$, and is $O_p(T\nu_T^2)$ otherwise. The result then follows since $supF_{T,B}(\tau) = sup_{\tau\in\Lambda_\epsilon}F_{T,B}(\tau) \geq F_{T,B}(\tau^0)$.

This completes the proof of Theorem 1. ■

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 $X_0(\tau^0) = [\bar{\iota}^0, \bar{Z}_f^0, Z_b]$, where $\bar{\iota}^0$ $[\bar{Z}_f^0]$ is a matrix which diagonally partitions ι $[Z_f]$ at the true break point T_1^0. For the unrestricted regression, we similarly have $X_1(\tau^0) = [\bar{\iota}^0, \bar{Z}_f^0, \bar{Z}_b^0]$. 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 $H = diag(I_{2(1+q_f)}, E_{2q_b \times q_b})$. 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 $F_{T,B}^{(2)}(\tau^0) = [SSR_0(\tau^0) - SSR_1(\tau^0)]/\hat{\sigma}_B^2(\tau^0) \xrightarrow{d} \chi^2(q_b)$.

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

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

where $W_c(.)$ and $W_z^f(.)$ $(q_f \times 1)$ 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 $D_T = diag(1, T^{-1/2}I_{q_f}, T^{-1/2}I_{q_b})$. Under $H_{0,B}$, $\lambda_b = 0_{q_b}$, and

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

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

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

Proof of Corollary 1: We prove the result under the null hypothesis $H_{0,B}$, the proof under $H_{0.C}$ being entirely analogous. Define the following events:

$$\begin{aligned} S_1 &= \{\sup F_{T,A}(\tau) > cv_A(\alpha)\} \\ S_2 &= \left\{F_{T,B}^{(2)}(\hat{\tau}) > cv_B(\alpha)\right\} \end{aligned}$$

Then we have

$$P\left[\{\sup F_{T,A}(\tau) > cv_A(\alpha)\} \cap \left\{F_{T,i}^{(2)}(\hat{\tau}) > cv_i(\alpha)\right\}\right] = P(S_1 \cap S_2) = P(S_1)P(S_2|S_1)$$

We consider the following two cases depending on whether a break exists in the coefficients not of interest under the null.

Case 1: $\lambda_c \neq 0$ and/or $\lambda_f \neq 0$. As $T \to \infty$, we have $P(S_1) \to 1$ (since the first step is consistent), $P(S_2|S_1) \to \alpha$ (by Theorem 2). Thus, $P(S_1 \cap S_2) \to \alpha$.

Case 2: $\lambda_c = \lambda_f = 0$. As $T \to \infty$, we have $P(S_1) \to \alpha$ (by Theorem 1 of Kejriwal and Perron, 2010). Since $P(S_2|S_1) \leq 1$, $\lim_{T\to\infty} P(S_1 \cap S_2) \leq \alpha$. Note that $P(S_2|S_1)$ does not converge to α since the second step test does not have a limiting chi-squared distribution which in turn is due to the fact that the break fraction estimate has a random limit in this case with no break in any of the coefficients.

Combining Cases 1 and 2, the result follows. The proof of Corollary 1 is complete. ■

Appendix B: Additional Monte Carlo Results

This Appendix contains the results of additional Monte Carlo simulations to assess the finite sample performance of the proposed two step procedure. Specifically, this set of simulations examines the impact of the break magnitude and the sample size on test size for DGPs 2b-3b. These simulations are motivated by the observation in Table 1 that the size distortions incurred by the two step procedure for DGPs 2b-3b do not decrease as the sample size increases.

Table B.1 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 empirical size does not monotonically approach the nominal size (5%) 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 the second stage size.

Panel B 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 B.1 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 ($T = 120$).

Table B.1: Rejection rates of the one-step partial KP and two-step (TS) 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
	TS	**5.77**	**9.12**	**10.03**	**9.29**	**8.44**	**7.77**	**7.27**	**6.92**
	TS_{1st}	22.81	65.56	89.91	97.85	99.61	99.94	100	100
	TS_{2nd}	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
	TS	**2.95**	**3.87**	**5.51**	**6.76**	**7.71**	**8.15**	**8.40**	**8.26**
	TS_{1st}	8.54	15.75	29.72	44.54	56.91	66.55	79.63	86.87
	TS_{2nd}	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
	TS	**3.16**	**5.82**	**7.81**	**9.01**	**9.88**	**8.96**	**8.10**	**7.82**
	TS_{1st}	8.67	27.78	48.94	65.42	78.23	98.72	99.96	100
	TS_{2nd}	36.46	20.96	15.95	13.77	12.64	9.08	8.11	7.82
	Δ_c	0.2	0.4	0.6	0.8	1	2	3	4
DGP-3b	KP	3.29	4.27	5.87	8.31	11.21	26.36	33.64	35.26
	TS	**1.53**	**1.86**	**2.37**	**3.10**	**3.91**	**7.06**	**7.62**	**7.36**
	TS_{1st}	2.85	4.18	6.48	10.42	15.73	57.18	85.56	96.32
	TS_{2nd}	53.65	44.38	36.50	29.76	24.87	12.34	8.90	7.64

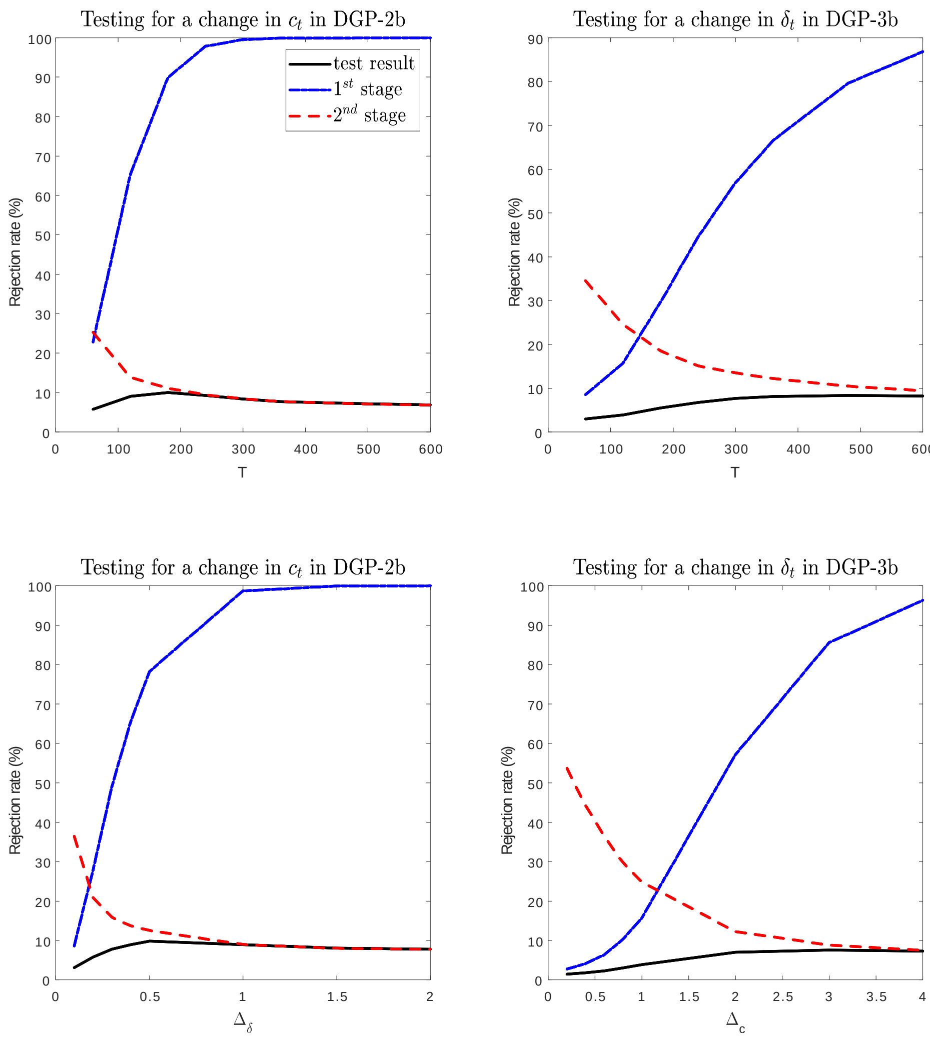

Figure 1: Rejection rates of two-step test for DGP-2b and DGP-3b as a function of sample size/break magnitude.