SUPPLEMENT: A FULLY FLEXIBLE CHANGEPOINT TEST FOR REGRESSION MODELS WITH STATIONARY ERRORS

Michael W. Robbins

RAND Corporation

Appendix A: Technical Assumptions

The following assumptions, which are largely borrowed from work such as Robbins *et al.* (2016), are imposed on the predictor sequences $\{\widetilde{\mathbf{x}}_t\}$, $\{\widetilde{\mathbf{s}}_t\}$, and, $\{\widetilde{\mathbf{v}}_t\}$ and error sequence $\{\epsilon_t\}$:

Assumption 1. *The sequence $\{\widetilde{\mathbf{v}}_t\}$ satisfies a functional central limit theorem.*

Assumption 2. *If $q_x > 0$, the functions f_1 through f_{q_x} are continuous and differentiable over the set K of admissible changepoints. It is also imposed that $f_j^2 > 0$ over the set K for $j = 1, \ldots, p_x$.*

Assumption 3. *Let $\widetilde{\boldsymbol{\chi}}_t = (\widetilde{\mathbf{x}}_t', \widetilde{\mathbf{s}}_t', \widetilde{\mathbf{v}}_t')'$. The matrix $n^{-1}\sum_{t=1}^n \widetilde{\boldsymbol{\chi}}_t \widetilde{\boldsymbol{\chi}}_t'$ is invertible for each $n \geq p_x + p_s + p_v$ with probability 1 in that it has a probability limit with a minimum eigenvalue that is bounded away from zero.*

Assumption 4. *The regression errors $\{\epsilon_t\}$ are independent of the process $\{\widetilde{\mathbf{v}}_t\}$ and satisfy*

$$\epsilon_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \tag{A.1}$$

where $\{Z_t\}$ is a sequence of mean zero independent and identically distributed (IID) random innovations that have a variance denoted by σ^2 and a finite $(2+\eta)^{\text{th}}$ moment for some

$\eta > 0$. *Also, the causal coefficients* $\{\psi_j\}$ *have a geometrically decaying structure that obeys* $|\psi_j| \leq \omega r^{-j}$ *for all* $j \geq 0$ *and some finite* ω *and* $r > 1$.

Appendix B: The Limit of $\mathbf{N}_{\mathrm{x},k}$

Define strict functional form versions of $\widetilde{\mathbf{x}}_t$ and $\mathbf{x}_t$ as $\widetilde{\boldsymbol{f}}(z) = (f_1(z), \ldots, f_{p_x}(z))'$ and $\boldsymbol{f}(z) = (f_1(z), \ldots, f_{q_x}(z))'$, respectively, for $z \in [0, 1]$. Also, let

$$\mathbf{G}(z) = \int_0^z \boldsymbol{f}(u)\boldsymbol{f}(u)'du, \qquad \mathbf{G}^*(z) = \int_0^z \widetilde{\boldsymbol{f}}(u)\boldsymbol{f}(u)'du$$

$$\text{and} \qquad \widetilde{\mathbf{G}}(z) = \int_0^z \widetilde{\boldsymbol{f}}(u)\widetilde{\boldsymbol{f}}(u)'du.$$

Likewise, let

$$\boldsymbol{\Gamma}(z) = \int_0^z \boldsymbol{f}(u)dW(u) \qquad \text{and} \qquad \widetilde{\boldsymbol{\Gamma}}(z) = \int_0^z \widetilde{\boldsymbol{f}}(u)dW(u),$$

where $\{W(z)\}_{z\in[0,1]}$ is a Wiener process. Define

$$\boldsymbol{\Omega}(z) = \mathbf{G}(z) - \mathbf{G}^*(z)'\widetilde{\mathbf{G}}(1)^{-1}\mathbf{G}^*(z) \qquad \text{and} \qquad \boldsymbol{\Lambda}(z) = \boldsymbol{\Gamma}(z) - \mathbf{G}^*(z)'\widetilde{\mathbf{G}}(1)^{-1}\widetilde{\boldsymbol{\Gamma}}(1).$$

Robbins *et al.* (2016) prove that

$$(n\tau^2)^{-1}\widehat{\mathrm{Var}}(\mathbf{N}_{\mathrm{x},\lfloor nz\rfloor}) \Rightarrow \boldsymbol{\Omega}(z) \qquad \text{and} \qquad (n\tau^2)^{-1/2}\mathbf{N}_{\mathrm{x},\lfloor nz\rfloor} \Rightarrow \boldsymbol{\Lambda}(z), \tag{A.2}$$

as $n \to \infty$ for $z \in K$. The result in (9) follows directly from the above.

Appendix C: Specific Representations for $\mathbf{s}_t$

The form of the ARMA residuals-based statistic $\widehat{L}^*_{\mathrm{s},k}$ can be simplified if seasonal component $\mathbf{s}_t$ obeys one of a pair of commonly used representations. First, consider that $\mathbf{s}_t$ takes the harmonic form

$$\mathbf{s}_t = (\mathbf{s}'_{j_1,t}, \ldots, \mathbf{s}'_{j_\rho,t})' \quad \text{where} \quad \mathbf{s}_{j,t} = (\cos(2\pi jt/T), \sin(2\pi jt/T))' \tag{A.3}$$

for $j \in (j_1, \ldots, j_\rho) \subseteq (1, \ldots, T/2)$ and $\rho \leq p_s/2$. Further, assume that any terms contained within $\mathbf{s}^*_t$ are among those in $(\mathbf{s}'_{1,t}, \ldots, \mathbf{s}'_{T/2,t})'$ which are not in $\mathbf{s}_t$. If $\{\mathbf{s}_t\}$ and $\{\mathbf{s}^*_t\}$ follow this representation, the matrix $\mathbf{D}_T$ has a diagonal form; specifically, $\mathbf{D}_T = \mathbf{I}_{q_s}/2$ where $\mathbf{I}_d$ is an identity matrix in d dimensions. Consider a second situation where seasonality is modeled exhaustively (i.e., each season is allocated its own mean term through the use of dummy variables). In this case, it holds that $q_s = T - 1$ so long as $\mathbf{x}_t$ contains an intercept term. Further, write $\mathbf{s}_t = (s_{1,t}, \ldots, s_{T-1,t})'$ and let

$$s_{j,t} = \begin{cases} 1 - T^{-1}, & \text{if } (t-j)/T \text{ an integer}, \\ -T^{-1}, & \text{otherwise.} \end{cases} \tag{A.4}$$

This equation defines an indicator variable that has been centered so as to satisfy the requirement that $\sum_{t=1}^{T} \mathbf{s}_t = \mathbf{0}$. Since $\mathbf{s}_t$ exhaustively models the periodicity, $\mathbf{s}^*_t$ is empty. In this case, $\mathbf{D}_T = \mathbf{I}_{q_s} - T^{-1}\mathbf{J}$, where $\mathbf{J}$ is a $q_s \times q_s$ matrix of ones. Under either of the above formulations for $\mathbf{s}_t$, the quantity $\widehat{L}_k$ can be further simplified by replacing $\mathbf{R}^*_{\mathrm{s},k}$ with the process $\mathbf{R}_{\mathrm{s},k}$ due to the following (a proof of which is provided in the supplement.).

Corollary A.1. *Given the conditions of Theorem 3, assume that* $\{\mathbf{s}_t\}$ *obeys (A.3) or (A.4).*

Let $\mathbf{R}_{\mathrm{s},k} = \sum_{t=1}^{k} \mathbf{s}_t \hat{Z}_t$ *and*

$$\widehat{L}_{\mathrm{s},k} = \frac{\mathbf{R}'_{\mathrm{s},k}(\mathbf{D}_T)^{-1}\mathbf{R}_{\mathrm{s},k}}{\hat{\sigma}^2 k(1-\frac{k}{n})}.$$

It follows that

$$\widehat{L}_{\mathrm{s},k} - \widehat{L}^*_{\mathrm{s},k} = o_p(1,k).$$

Appendix D: Nonstationary Stochastic Covariates

In the main text, we assumed that $\{\widetilde{\mathbf{v}}_t\}$ is stationary with zero mean. Now, we generalize to circumstances where $\{\widetilde{\mathbf{v}}_t\}$ has nonzero mean; specifically, consider that $\{\widetilde{\mathbf{v}}_t\}$ is generated via

$$\widetilde{\mathbf{v}}_t = \boldsymbol{\xi}'\boldsymbol{a}_t + \widetilde{\mathbf{u}}_t,$$

where $\{\widetilde{\mathbf{u}}_t\}$ (which is decomposed as $\widetilde{\mathbf{u}}_t = (\mathbf{u}'_t, (\mathbf{u}^*_t)')'$ in the same manner as the other regressor vectors introduced in Sections 1 and 2) is stationary with zero mean, $\{\boldsymbol{a}_t\}$ is a vector of known deterministic design points, and $\boldsymbol{\xi}$ is a matrix of constants.

Assume that predictor sequence given by $\{\boldsymbol{a}_t\}$ is contained within the predictors in $\{(\mathbf{x}'_t, \mathbf{s}'_t)'\}$. It follows that the OLS residuals take on the same values when $\{\widetilde{\mathbf{v}}_t\}$ is used as a predictor as they do when $\{\widetilde{\mathbf{u}}_t\}$ is used in its place when fitting the regression (this is the case for residuals calculated under both the null and alternative hypotheses). Therefore, the sequences $\{\widehat{F}_k\}$ and $\{\widehat{F}^*_k\}$, defined in (6) and (21), respectively, are unchanged if $\{\widetilde{\mathbf{u}}_t\}$ were used in place of $\{\widetilde{\mathbf{v}}_t\}$, and the limit laws given in Theorems 1 and 2 still hold.

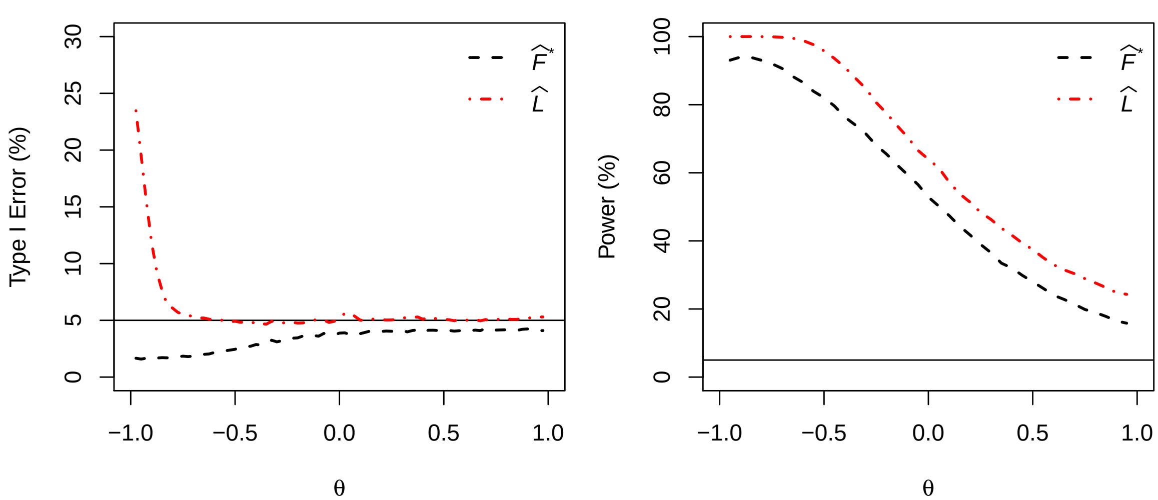
However, one must filter a nonstationary mean sequence out of $\{\mathbf{v}_t\}$ prior to calculating $\{\mathbf{R}^*_{\mathrm{v},k}\}$, as defined in (24). Let $\{\widehat{\mathbf{u}}_t\}$ denote residuals from a regression of $\{\mathbf{v}_t\}$ on $\{\boldsymbol{a}_t\}$ and let $\widehat{\boldsymbol{\xi}}_q$ denote a $\sqrt{n}$-consistent estimator of $\boldsymbol{\xi}_q$, where $\boldsymbol{\xi}_q$ gives the first q_v rows of $\boldsymbol{\xi}$. Define

$$\mathbf{R}^{\dagger}_{\mathrm{v},k} = \sum_{t=1}^{k} \widehat{\mathbf{u}}_t \hat{Z}_t - \frac{k}{n} \sum_{t=1}^{n} \widehat{\mathbf{u}}_t \hat{Z}_t.$$

If $\mathbf{R}^{*}_{\mathrm{v},k}$ is replaced with $\mathbf{R}^{\dagger}_{\mathrm{v},k}$ in the calculation of $\widehat{L}$, the convergence illustrated in Theorem 3 will hold. If $\boldsymbol{a}_t$ contains terms exogenous to $(\mathbf{x}_t', \mathbf{s}_t')'$, we recommend homogenizing $\{\widetilde{\mathbf{v}}_t\}$ prior fitting any regressions (and therefore prior to calculating test statistics). In this case, the limit theory outlined above applies; formal proof of these claims is omitted for brevity.

Appendix E: Additional Simulations The motivation for use of the $\widehat{F}^*$ statistic is that it does not impose a parametric model on the error structure. Therefore, we examine the performance of the $\widehat{L}$ statistic when the serial correlation in $\{\epsilon_t\}$ is not correctly modeled. Specifically, we generate $\{\epsilon_t\}$ using various values of θ while fixing $\phi = 0$ (this implies the er-

Figure A.1: Simulated size (left) and power (right) of the $\widehat{F}^*$ and *misspecified* $\widehat{L}$ tests for a nominal significance level of 0.05 when alternative model H1a is considered and when the error sequence $\{\epsilon_t\}$ is generated from an MA(1) model with parameter θ with $n = 1000$. Results are shown for various choices of θ, where $\delta_x = \delta_s = \delta_v = 0.107$ for power comparisons. Results for size are based on 100,000 independently simulated datasets for each value of θ, whereas 25,000 datasets are generated for power calculations



rors are sampled from a MA(1) model). Then, when the $\widehat{L}$ statistic is calculated, an AR(p_{ar}) model, with p_{ar} selected using the AIC criterion, is fit to the regression residuals. The size of the $\widehat{F}^*$ and the misspecified $\widehat{L}$ tests are approximated under alternative H1a, and then the power for this alternative is calculated while fixing $\delta_x = \delta_s = \delta_v = 0.107$. Results are shown in Figure A.1. The findings indicate that the $\widehat{L}$ statistic still outperforms the $\widehat{F}^*$ statistic (with regards to both size and power), even when the error model is incorrectly specified. It is expected that power for both tests will decrease as θ increases (Robbins *et al.*, 2011a).

Appendix F: Proofs

Theorem 1. As is stipulated by the conditions of Theorem 1, this proof assumes IID regression errors (i.e., $\epsilon_t = Z_t$ and thus $\tau^2 = \sigma^2$). To begin, let

$$\mathbf{X}_t = (\mathbf{x}_1, \ldots, \mathbf{x}_t, \mathbf{0}, \ldots, \mathbf{0})' \quad \text{and} \quad \widetilde{\mathbf{X}}_t = (\widetilde{\mathbf{x}}_1, \ldots, \widetilde{\mathbf{x}}_t, \mathbf{0}, \ldots, \mathbf{0})',$$

which are matrices of dimension $n \times q_x$ and $n \times p_x$, respectively, the last $n - t$ rows of which contain zeros. Similarly, define

$$\mathbf{S}_t = (\mathbf{s}_1, \ldots, \mathbf{s}_t, \mathbf{0}, \ldots, \mathbf{0})' \quad \text{and} \quad \widetilde{\mathbf{S}}_t = (\widetilde{\mathbf{s}}_1, \ldots, \widetilde{\mathbf{s}}_t, \mathbf{0}, \ldots, \mathbf{0})',$$

and

$$\mathbf{V}_t = (\mathbf{v}_1, \ldots, \mathbf{v}_t, \mathbf{0}, \ldots, \mathbf{0})' \quad \text{and} \quad \widetilde{\mathbf{V}}_t = (\widetilde{\mathbf{v}}_1, \ldots, \widetilde{\mathbf{v}}_t, \mathbf{0}, \ldots, \mathbf{0})'.$$

Further, let $\mathbf{M}_t = (\mathbf{X}_t, \mathbf{S}_t, \mathbf{V}_t)$ and $\widetilde{\mathbf{M}}_t = (\widetilde{\mathbf{X}}_t, \widetilde{\mathbf{S}}_t, \widetilde{\mathbf{V}}_t)$. Note that $\mathbf{M}_n$ is the full design matrix under $\mathcal{H}_0$. The null hypothesis OLS estimator of $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_x', \boldsymbol{\Delta}_s', \boldsymbol{\Delta}_v')'$, when a changepoint is assumed to occur at time k with $k/n \in K$, is $\widehat{\mathbf{\Delta}}_k = -\mathbf{C}_k^{-1}\mathbf{N}_k$ where

$$\mathbf{C}_k = \mathbf{M}_k'(\mathbf{I} - \mathbf{P}_n)\mathbf{M}_k, \tag{A.5}$$

and

$$\mathbf{N}_k = \mathbf{M}_k'(\mathbf{I} - \mathbf{P}_n)\mathbf{Y}, \tag{A.6}$$

with $\mathbf{Y} = (Y_1, \ldots, Y_n)'$. In the above, $\mathbf{P}_n = \widetilde{\mathbf{M}}_n(\widetilde{\mathbf{M}}_n'\widetilde{\mathbf{M}}_n)^{-1}\widetilde{\mathbf{M}}_n'$ is the projection matrix under the null hypothesis. Conditional on $\widetilde{\mathbf{V}}_n$, it holds that $\mathrm{Var}(\widehat{\mathbf{\Delta}}_k) = \tau^2\mathbf{C}_k^{-1}$.

Note that

$$n^{-1}\mathbf{M}_k'\mathbf{M}_k = \begin{pmatrix} n^{-1}\mathbf{X}_k'\mathbf{X}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & n^{-1}\mathbf{S}_k'\mathbf{S}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & n^{-1}\mathbf{V}_k'\mathbf{V}_k \end{pmatrix} + o_p(1, k).$$

Lemmas A.1 and A.2 of Robbins *et al.* (2016) are used to show that the off-diagonal blocks of the above matrix are zero asymptotically. Similarly,

$$(n^{-1}\widetilde{\mathbf{M}}_n'\widetilde{\mathbf{M}}_n)^{-1} = \begin{pmatrix} (n^{-1}\widetilde{\mathbf{X}}_n'\widetilde{\mathbf{X}}_n)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (n^{-1}\widetilde{\mathbf{S}}_n'\widetilde{\mathbf{S}}_n)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (n^{-1}\widetilde{\mathbf{V}}_n'\widetilde{\mathbf{V}}_n)^{-1} \end{pmatrix} + \mathcal{O}_p(n^{-1/2}).$$

Likewise, we now see

$$
\begin{aligned}
n^{-1}\mathbf{M}_k'\widetilde{\mathbf{M}}_n &= n^{-1}\begin{pmatrix} \mathbf{X}_k'\widetilde{\mathbf{X}}_n & \mathbf{X}_k'\widetilde{\mathbf{S}}_n & \mathbf{X}_k'\widetilde{\mathbf{V}}_n \\ \mathbf{S}_k'\widetilde{\mathbf{X}}_n & \mathbf{S}_k'\widetilde{\mathbf{S}}_n & \mathbf{S}_k'\widetilde{\mathbf{V}}_n \\ \mathbf{V}_k'\widetilde{\mathbf{X}}_n & \mathbf{V}_k'\widetilde{\mathbf{S}}_n & \mathbf{V}_k'\widetilde{\mathbf{V}}_n \end{pmatrix} \\
&= \begin{pmatrix} n^{-1}\mathbf{X}_k'\widetilde{\mathbf{X}}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & n^{-1}\mathbf{S}_k'\widetilde{\mathbf{S}}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & n^{-1}\mathbf{V}_k'\widetilde{\mathbf{V}}_n \end{pmatrix} + o_p(1,k).
\end{aligned}
$$

Continuing, we see

$$
\mathbf{M}_k'\widetilde{\mathbf{M}}_n(\widetilde{\mathbf{M}}_n'\widetilde{\mathbf{M}}_n)^{-1} = \begin{pmatrix} \mathbf{X}_k'\widetilde{\mathbf{X}}_n(\widetilde{\mathbf{X}}_n'\widetilde{\mathbf{X}}_n)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_k'\widetilde{\mathbf{S}}_n(\widetilde{\mathbf{S}}_n'\widetilde{\mathbf{S}}_n)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_k'\widetilde{\mathbf{V}}_n(\widetilde{\mathbf{V}}_n'\widetilde{\mathbf{V}}_n)^{-1} \end{pmatrix} + o_p(1,k). \quad \text{(A.7)}
$$

Hence,

$$
n^{-1}\mathbf{C}_k = \begin{pmatrix} n^{-1}\mathbf{C}_{\mathrm{x},k} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & n^{-1}\mathbf{C}_{\mathrm{s},k} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & n^{-1}\mathbf{C}_{\mathrm{v},k} \end{pmatrix} + o_p(1,k), \quad \text{(A.8)}
$$

where

$$
\mathbf{C}_{\mathrm{x},k} = \mathbf{X}_k'\mathbf{X}_k - \mathbf{X}_k'\widetilde{\mathbf{X}}_n(\widetilde{\mathbf{X}}_n'\widetilde{\mathbf{X}}_n)^{-1}\widetilde{\mathbf{X}}_n'\mathbf{X}_k, \quad \text{(A.9)}
$$

with

$$
\mathbf{C}_{\mathrm{s},k} = \mathbf{S}_k'\mathbf{S}_k - \mathbf{S}_k'\widetilde{\mathbf{S}}_n(\widetilde{\mathbf{S}}_n'\widetilde{\mathbf{S}}_n)^{-1}\widetilde{\mathbf{S}}_n'\mathbf{S}_k, \quad \text{(A.10)}
$$

and

$$\mathbf{C}_{\mathrm{v},k} = \mathbf{V}_k'\mathbf{V}_k - \mathbf{V}_k'\widetilde{\mathbf{V}}_n(\widetilde{\mathbf{V}}_n'\widetilde{\mathbf{V}}_n)^{-1}\widetilde{\mathbf{V}}_n'\mathbf{V}_k,$$

for $\mathbf{C}_k$ is defined in (A.5).

Shifting the focus to the process $\{\mathbf{N}_k\}$, we first note that

$$\mathbf{N}_k = \begin{pmatrix} \mathbf{N}_{\mathrm{x},k} \\ \mathbf{N}_{\mathrm{s},k} \\ \mathbf{N}_{\mathrm{v},k} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_k'\boldsymbol{\epsilon} - \mathbf{X}_k'\widetilde{\mathbf{M}}_n(\widetilde{\mathbf{M}}_n'\widetilde{\mathbf{M}}_n)^{-1}\widetilde{\mathbf{M}}_n'\boldsymbol{\epsilon} \\ \mathbf{S}_k'\boldsymbol{\epsilon} - \mathbf{S}_k'\widetilde{\mathbf{M}}_n(\widetilde{\mathbf{M}}_n'\widetilde{\mathbf{M}}_n)^{-1}\widetilde{\mathbf{M}}_n'\boldsymbol{\epsilon} \\ \mathbf{V}_k'\boldsymbol{\epsilon} - \mathbf{V}_k'\widetilde{\mathbf{M}}_n(\widetilde{\mathbf{M}}_n'\widetilde{\mathbf{M}}_n)^{-1}\widetilde{\mathbf{M}}_n'\boldsymbol{\epsilon} \end{pmatrix},$$

for $\mathbf{N}_k$ as defined in (A.6) and for $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$. Recall that $\mathbf{N}_{\mathrm{x},k}$, $\mathbf{N}_{\mathrm{s},k}$, and $\mathbf{N}_{\mathrm{v},k}$ were defined in (8). Using $\mathrm{Var}(\mathbf{N}_k) = \tau^2\mathbf{C}_k$ and (A.8), it holds that these three processes are asymptotically uncorrelated. Applying the result of (A.7), we see

$$n^{-1/2}\begin{pmatrix} \mathbf{N}_{\mathrm{x},k} \\ \mathbf{N}_{\mathrm{s},k} \\ \mathbf{N}_{\mathrm{v},k} \end{pmatrix} = n^{-1/2}\begin{pmatrix} \mathbf{X}_k'\boldsymbol{\epsilon} - \mathbf{X}_k'\widetilde{\mathbf{X}}_n(\widetilde{\mathbf{X}}_n'\widetilde{\mathbf{X}}_n)^{-1}\widetilde{\mathbf{X}}_n'\boldsymbol{\epsilon} \\ \mathbf{S}_k'\boldsymbol{\epsilon} - \mathbf{S}_k'\widetilde{\mathbf{S}}_n(\widetilde{\mathbf{S}}_n'\widetilde{\mathbf{S}}_n)^{-1}\widetilde{\mathbf{S}}_n'\boldsymbol{\epsilon} \\ \mathbf{V}_k'\boldsymbol{\epsilon} - \mathbf{V}_k'\widetilde{\mathbf{V}}_n(\widetilde{\mathbf{V}}_n'\widetilde{\mathbf{V}}_n)^{-1}\widetilde{\mathbf{V}}_n'\boldsymbol{\epsilon} \end{pmatrix} + o_p(1,k). \tag{A.11}$$

The processes $\{\mathbf{C}_{\mathrm{x},k}\}$ and $\{\mathbf{N}_{\mathrm{x},k}\}$ were studied in the proof of Lemma 2.1 in Robbins *et al.* (2016); the focus now turns to $\{\mathbf{C}_{\mathrm{s},k}\}$ and $\{\mathbf{N}_{\mathrm{s},k}\}$.

Let $\mathbf{D}_T = \sum_{j=1}^T \mathbf{s}_j\mathbf{s}_j'/T$ with $\mathbf{D}_T^* = \sum_{j=1}^T \widetilde{\mathbf{s}}_j\mathbf{s}_j'/T$ and $\widetilde{\mathbf{D}}_T = \sum_{j=1}^T \widetilde{\mathbf{s}}_j\widetilde{\mathbf{s}}_j'/T$. Consequentially,

$$n^{-1}\widetilde{\mathbf{S}}_n'\widetilde{\mathbf{S}}_n = \widetilde{\mathbf{D}}_T + \mathcal{O}(n^{-1}),$$

and

$$n^{-1}\mathbf{S}_k'\widetilde{\mathbf{S}}_n = (k/n)\mathbf{D}_T^* + \mathcal{O}(n^{-1},k), \quad \text{with} \quad n^{-1}\mathbf{S}_k'\mathbf{S}_k = (k/n)\mathbf{D}_T + \mathcal{O}(n^{-1},k).$$

To derive (10), note that

$$\begin{aligned} n^{-1}\mathbf{C}_{\mathrm{s},\lfloor nz\rfloor} &\Rightarrow z\mathbf{D}_T - z^2(\mathbf{D}_T^*)'(\widetilde{\mathbf{D}}_T)^{-1}\mathbf{D}_T^*, \\ &= z\mathbf{D}_T - z^2\begin{pmatrix} \mathbf{I}_{p_s} & \mathbf{0} \end{pmatrix}\mathbf{D}_T^*, \\ &= z(1-z)\mathbf{D}_T. \end{aligned}$$

The second line follows from the fact that $\mathbf{D}_T^*$ equals the first q_s columns of $\widetilde{\mathbf{D}}_T$, and similarly the third line uses the observation that the first q_s rows of $\mathbf{D}_T^*$ equal $\mathbf{D}_T$. Likewise,

$$\begin{aligned} \mathbf{N}_{\mathrm{s},k} &= \mathbf{S}_k'\boldsymbol{\epsilon} - (k/n)(\widetilde{\mathbf{D}}_T^*)'(\widetilde{\mathbf{D}}_T)^{-1}\widetilde{\mathbf{S}}_n'\boldsymbol{\epsilon} + o_p(\sqrt{n},k) \\ &= \mathbf{S}_k'\boldsymbol{\epsilon} - (k/n)\begin{pmatrix} \mathbf{I}_{q_s} & \mathbf{0} \end{pmatrix}^{-1}\widetilde{\mathbf{S}}_n'\boldsymbol{\epsilon} + o_p(\sqrt{n},k) \\ &= \mathbf{S}_k'\boldsymbol{\epsilon} - (k/n)\mathbf{S}_n'\boldsymbol{\epsilon} + o_p(\sqrt{n},k), \end{aligned}$$

which illustrates (12).

Recall that $\boldsymbol{e}_i = \sum_{t=T(i-1)+1}^{iT}\mathbf{s}_t\epsilon_t$ and note that $\mathbf{S}_k'\boldsymbol{\epsilon} = \sum_{i=1}^{m^*}\boldsymbol{e}_i$, where it is assumed that $n = Tm$ and $k = Tm^*$. Note further that $\mathrm{Var}(\boldsymbol{e}_i) = \tau^2 T\mathbf{D}_T$. It follows that

$$(T/n)^{1/2}\mathbf{N}_{\mathrm{s},k} = \frac{1}{\sqrt{m}}\sum_{i=1}^{m^*}\boldsymbol{e}_i - \frac{k}{n}\left(\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\boldsymbol{e}_i\right) + o_p(1,k),$$

This formula, in combination with (10), yields (14).

Similar approaches are taken to extract the limit behavior of $\mathbf{C}_{\mathrm{v},k}$ and $\mathbf{N}_{\mathrm{v},k}$. Define

$\widetilde{\boldsymbol{\Sigma}}_v = \mathrm{E}\left[\widetilde{\mathbf{v}}_t\widetilde{\mathbf{v}}_t'\right]$, let $\boldsymbol{\Sigma}_v^*$ denote the first q_v columns of $\widetilde{\boldsymbol{\Sigma}}_v$ and let $\boldsymbol{\Sigma}_v$ denote the first q_v rows of $\boldsymbol{\Sigma}_v^*$. Furthermore,

$$n^{-1}\widetilde{\mathbf{V}}_n'\widetilde{\mathbf{V}}_n \to \widetilde{\boldsymbol{\Sigma}}_v, \quad n^{-1}\mathbf{V}_{\lfloor nz\rfloor}'\widetilde{\mathbf{V}}_n \Rightarrow z\boldsymbol{\Sigma}_v^*, \quad \text{and} \quad n^{-1}\mathbf{V}_{\lfloor nz\rfloor}'\mathbf{V}_{\lfloor nz\rfloor} \Rightarrow z\boldsymbol{\Sigma}_v.$$

Formulas (11) and (13) are derived using arguments akin to those that provide (10) and (12). Specifically,

$$n^{-1}\mathbf{C}_{\mathrm{v},\lfloor nz\rfloor} \Rightarrow z(1-z)\boldsymbol{\Sigma}_v$$

and

$$\mathbf{N}_{\mathrm{v},k} = \mathbf{V}_k'\boldsymbol{\epsilon} - (k/n)\mathbf{V}_n'\boldsymbol{\epsilon} + o_p(\sqrt{n}, k).$$

Note that the sequence $\{\mathbf{v}_t\epsilon_t\}$ is devoid of autocorrelation and observes $\mathrm{Var}(\mathbf{v}_t\epsilon_t) = \tau^2\boldsymbol{\Sigma}_v$. Therefore, the identity in (15) is now evident.

From (A.8) and (A.11), it follows that

$$\widehat{F}_k = \frac{\mathbf{N}_{\mathrm{x},k}'\mathbf{C}_{\mathrm{x},k}^{-1}\mathbf{N}_{\mathrm{x},k}}{\hat{\tau}^2} + \frac{\mathbf{N}_{\mathrm{s},k}'(\mathbf{D}_T)^{-1}\mathbf{N}_{\mathrm{s},k}}{\hat{\tau}^2 k(1-\frac{k}{n})} + \frac{\mathbf{N}_{\mathrm{v},k}'(\widehat{\boldsymbol{\Sigma}}_v)^{-1}\mathbf{N}_{\mathrm{v},k}}{\hat{\tau}^2 k(1-\frac{k}{n})} + o_p(1,k),$$

where $\widehat{F}_k$ is defined in (6). The limit behavior of the term involving $\mathbf{N}_{\mathrm{x},k}$ follows from (A.2), and the limit behavior of the term involving $\mathbf{N}_{\mathrm{s},k}$ follows from (10) and (12). Likewise, the limit distribution of the term involving $\mathbf{N}_{\mathrm{v},k}$ follows from (11) and (13). The block-diagonal form of $\mathrm{Var}(\mathbf{N}_k) = \tau^2\mathbf{C}_k$ as $n \to \infty$, which is evident in (A.8), implies pairwise asymptotic independence of $\mathbf{N}_{\mathrm{x},k}$, $\mathbf{N}_{\mathrm{s},k}$ and $\mathbf{N}_{\mathrm{v},k}$. To establish (asymptotic) process independence, calculations similar to those which yield the form of $\mathbf{C}_k$ can be used to establish that $\mathbf{N}_{\mathrm{x},k}$ and $\mathbf{N}_{\mathrm{s},k'}$, for example, are asymptotically uncorrelated for $k \neq k'$. □

Lemma 1. Let $\{\mathbf{b}_t\}$ and $\{\widetilde{\mathbf{b}}_t\}$ be sequences of vectors that satisfy

$$\sum_{t=1}^{k} \mathbf{b}_t \hat{\epsilon}_{t-i} = \sum_{t=1}^{k} \mathbf{b}_t \epsilon_{t-i} - \sum_{t=1}^{k} \mathbf{b}_{t+i} \widetilde{\mathbf{b}}_t \left(\sum_{t=1}^{n} \widetilde{\mathbf{b}}_t \widetilde{\mathbf{b}}_t' \right)^{-1} \sum_{t=1}^{n} \widetilde{\mathbf{b}}_t \epsilon_t + o_p(\sqrt{n}, k), \tag{A.12}$$

where $\{\hat{\epsilon}_t\}$ is the sequence of OLS residuals generated using (5) and where $\{\epsilon_t\}$ is the sequence of regression errors generated from the white noise ARMA errors $\{Z_t\}$ in accordance with (22). Assume further that

$$\frac{1}{n} \sum_{t=1}^{\lfloor nz \rfloor} \mathbf{b}_{t+i} \widetilde{\mathbf{b}}_t' \Rightarrow z \mathbf{\Gamma}_b(i) \quad \text{and} \quad \frac{1}{n} \sum_{t=1}^{n} \widetilde{\mathbf{b}}_t \widetilde{\mathbf{b}}_t' \rightarrow \widetilde{\mathbf{\Gamma}}_b(0) \tag{A.13}$$

for some sequence of deterministic matrices $\{\mathbf{\Gamma}_b(i)\}$ with arbitrary $i \geq 0$ and for some matrix $\widetilde{\mathbf{\Gamma}}_b(0)$. It follows that

$$\begin{aligned} \sum_{t=1}^{k} \mathbf{b}_t \hat{\epsilon}_{t-i} - \frac{k}{n} \sum_{t=1}^{n} \mathbf{b}_t \hat{\epsilon}_{t-i} &= \sum_{t=1}^{k} \mathbf{b}_t \epsilon_{t-i} - (k/n) \mathbf{\Gamma}_b(i) [\widetilde{\mathbf{\Gamma}}_b(0)]^{-1} \sum_{t=1}^{n} \widetilde{\mathbf{b}}_t \epsilon_t \\ &\quad - \frac{k}{n} \left(\sum_{t=1}^{n} \mathbf{b}_t \epsilon_{t-i} - \mathbf{\Gamma}_b(i) [\widetilde{\mathbf{\Gamma}}_b(0)]^{-1} \sum_{t=1}^{n} \widetilde{\mathbf{b}}_t \epsilon_t \right) + o_p(\sqrt{n}, k) \\ &= \sum_{t=1}^{k} \mathbf{b}_t \epsilon_{t-i} - \frac{k}{n} \sum_{t=1}^{n} \mathbf{b}_t \epsilon_{t-i} + o_p(\sqrt{n}, k). \end{aligned} \tag{A.14}$$

Let the sequence $\{\pi_j\}_{j=0}^{\infty}$ denote the coefficients from the expansion of $(1 - \phi_1 z - \cdots - \phi_p z^p)/(1 + \theta_1 z + \cdots + \theta_q z^q)$, and let $\{\hat{\pi}_j\}_{j=0}^{\infty}$ represent the versions of these coefficients when calculated using the $\hat{\phi}_j$ and $\hat{\theta}_j$. Hence,

$$Z_t = \sum_{j=0}^{\infty} \pi_j \epsilon_{t-j} \quad \text{and} \quad \hat{Z}_t = \sum_{j=0}^{\infty} \hat{\pi}_j \hat{\epsilon}_{t-j}, \tag{A.15}$$

and

$$\begin{aligned}\sum_{t=1}^{k}\mathbf{b}_t\hat{Z}_t-\frac{k}{n}\sum_{t=1}^{n}\mathbf{b}_t\hat{Z}_t &= \sum_{j=0}^{\infty}\hat{\pi}_j\left(\sum_{t=1}^{k}\mathbf{b}_t\hat{\epsilon}_{t-j}-\frac{k}{n}\sum_{t=1}^{n}\mathbf{b}_t\hat{\epsilon}_{t-j}\right)\\ &= \sum_{j=0}^{\infty}\pi_j\left(\sum_{t=1}^{k}\mathbf{b}_t\epsilon_{t-j}-\frac{k}{n}\sum_{t=1}^{n}\mathbf{b}_t\epsilon_{t-j}\right)+o_p(\sqrt{n},k). \quad \text{(A.16)}\end{aligned}$$

The last line in the above uses (A.14) and the facts that the elements of $\{\pi_t\}$ decay at an exponential rate while $\{\hat{\pi}_t\}$ converges to $\{\pi_t\}$ at an even quicker rate. It follows that

$$\sum_{t=1}^{k}\mathbf{b}_t\hat{Z}_t-\frac{k}{n}\sum_{t=1}^{n}\mathbf{b}_t\hat{Z}_t=\sum_{t=1}^{k}\mathbf{b}_tZ_t-\frac{k}{n}\sum_{t=1}^{n}\mathbf{b}_tZ_t+o_p(\sqrt{n},k).$$

Note that (A.7) and, therefore, (A.11) hold in the event that $\mathbf{M}_k$ is substituted with an analogous version that has the row of $\mathbf{M}_k$ that corresponds to $(\mathbf{x}_t', \mathbf{s}_t', \mathbf{v}_t')$ replaced with $(\mathbf{x}_{t+i}', \mathbf{s}_{t+i}', \mathbf{v}_{t+i}')$ for $t=1,\dots,k$ and for $i \geq 0$. Using this observation as well as the fact that $\{n^{-1/2}\sum_{t=1}^{k}\mathbf{b}_t\hat{\epsilon}_{t-i}\}$ and $\{n^{-1/2}\sum_{t=1}^{k}\mathbf{b}_{t+i}\hat{\epsilon}_t\}$ are asymptotically equivalent, we see that $\{\mathbf{s}_t\}$ and $\{\widetilde{\mathbf{s}}_t\}$ obey (A.12), as do $\{\mathbf{v}_t\}$ and $\{\widetilde{\mathbf{v}}_t\}$. Lastly, it holds that these predictor sequences obey (A.13), which yields the lemma's main result. □

Theorem 3. It is first illustrated that the processes $\{\mathbf{R}_{\mathrm{x},k}\}$, $\{\mathbf{R}_{\mathrm{s},k}\}$ and $\{\mathbf{R}_{\mathrm{v},k}\}$ are asymptotically uncorrelated. In light of (23), (A.11), and Lemma 1, it is sufficient to show that the following three processes are asymptotically uncorrelated:

$$\left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{k}\mathbf{x}_t\epsilon_t\right\}, \quad \left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{k}\mathbf{s}_tZ_t\right\}, \quad \text{and} \quad \left\{\frac{1}{\sqrt{n}}\sum_{t=1}^{k}\mathbf{v}_tZ_t\right\},$$

for $1 \leq k \leq n$, where $\{\epsilon_t\}$ is generated from $\{Z_t\}$ via the causal representation in (A.1). Calculations show that $\sum_{t=1}^{k}\mathbf{x}_t\epsilon_t=\sum_{t=1}^{k}\mathbf{y}_tZ_t$ where $\mathbf{y}_t=\sum_{j=t}^{k}\psi_{j-t}\mathbf{x}_j$.

The proof of Theorem 1 illustrates that the latter two processes of the three processes above are uncorrelated in large samples. Further calculations show that the remaining pairwise covariances of these processes are diven by $n^{-1}\sum_{t=1}^{k}\mathbf{y}_t\mathbf{s}_t'$ and $n^{-1}\sum_{t=1}^{k}\mathbf{y}_t\mathbf{v}_t'$. Furthermore,

$$\sum_{t=1}^{k}\mathbf{y}_t\mathbf{s}_t' = \sum_{j=0}^{k-1}\psi_j\sum_{t=1}^{k}\mathbf{x}_{t+j}\mathbf{s}_t' \quad \text{and} \quad \sum_{t=1}^{k}\mathbf{y}_t\mathbf{v}_t' = \sum_{j=0}^{k-1}\psi_j\sum_{t=1}^{k}\mathbf{x}_{t+j}\mathbf{v}_t'.$$

Using the geometrically decaying structure of $\{\psi_j\}$ in addition to Lemmas A.1 and A.2 of Robbins *et al.* (2016), it holds that $n^{-1}\sum_{t=1}^{k}\mathbf{y}_t\mathbf{s}_t' = \mathcal{O}(n^{-1},k)$ and $n^{-1}\sum_{t=1}^{k}\mathbf{y}_t\mathbf{v}_t' = \mathcal{O}_p(n^{-1/2},k)$. This, in combination with (23), Lemma 1 and (25), illustrates the result in the theorem. □

Corollary 1. Let $\mathbf{b}_t$ denote one of $\mathbf{x}_t$, $\mathbf{s}_t$ or $\mathbf{v}_t$ and correspondingly let $\boldsymbol{\Delta}_b$ denote either $\boldsymbol{\Delta}_x$, $\boldsymbol{\Delta}_s$ or $\boldsymbol{\Delta}_v$. In the event that $\boldsymbol{\Delta}_b \neq \mathbf{0}$ and that the changepoint occurs at time c, we see

$$\begin{aligned}\sum_{t=1}^{k}\mathbf{b}_t\hat{\epsilon}_{t-i} &= \left[\sum_{t=1}^{\min\{k,c\}}\mathbf{b}_{t+i}\mathbf{b}_t' - \sum_{t=1}^{k}\mathbf{b}_{t+i}\widetilde{\mathbf{b}}_t\left(\sum_{t=1}^{n}\widetilde{\mathbf{b}}_t\widetilde{\mathbf{b}}_t'\right)^{-1}\sum_{t=1}^{c}\mathbf{b}_t\mathbf{b}_t'\right]\boldsymbol{\Delta}_b \\ &\quad + \sum_{t=1}^{k}\mathbf{b}_t\epsilon_{t-i} - \sum_{t=1}^{k}\mathbf{b}_{t+i}\widetilde{\mathbf{b}}_t\left(\sum_{t=1}^{n}\widetilde{\mathbf{b}}_t\widetilde{\mathbf{b}}_t'\right)^{-1}\sum_{t=1}^{n}\widetilde{\mathbf{b}}_t\epsilon_t + o_p(\sqrt{n},k), \quad \text{(A.17)}\end{aligned}$$

which expands upon (A.12).

Our focus now turns to $\mathbf{b}_t = \mathbf{x}_t$. Let

$$\mathcal{A}_k = \sum_{i=0}^{p_{\text{ar}}}\hat{\phi}_i^*\sum_{t=1}^{k}f\left(\frac{t}{n}\right)\hat{\epsilon}_t - \sum_{j=0}^{q_{\text{ma}}}\hat{\theta}_j^*\sum_{t=1}^{k}f\left(\frac{t}{n}\right)\hat{Z}_t,$$

where $f(t/n)$ denotes an arbitrary element of $\mathbf{x}_t$ and where $\hat{\phi}_i^* = -\hat{\phi}_i$ and $\hat{\theta}_j^* = \hat{\theta}_j$ for $i = 1,\ldots,p_{\text{ar}}$ and $j = 1,\ldots,q_{\text{ma}}$ with $\hat{\phi}_0^* = \hat{\theta}_0^* = 1$. Following the proof of Lemma 2.2 of

Robbins *et al.* (2016), it holds that

$$\mathcal{A}_k + \frac{1}{n}\sum_{i=0}^{p_{\rm ar}} i\hat{\phi}_i^* \sum_{t=1}^{k} \dot{f}(\xi_{ti})\hat{\epsilon}_t - \frac{1}{n}\sum_{j=0}^{q_{\rm ma}} j\hat{\theta}_j^* \sum_{t=1}^{k} \dot{f}(\xi_{tj})\hat{Z}_t = o_p(n^{1/\nu}, k),$$

for some $\nu \geq 2$ where $\cdot f(z)$ is the first derivative of $f(z)$. Let

$$\mathcal{B}_c(k) = \left[\sum_{t=1}^{\min\{k,c\}} \mathbf{x}_t\mathbf{x}_t' - \sum_{t=1}^{k} \mathbf{x}_t\widetilde{\mathbf{x}}_t \left(\sum_{t=1}^{n} \widetilde{\mathbf{x}}_t\widetilde{\mathbf{x}}_t'\right)^{-1} \sum_{t=1}^{c} \mathbf{x}_t\mathbf{x}_t'\right] \mathbf{\Delta}_x,$$

and note that $n^{-1}\mathcal{B}_c(k) = \mathcal{O}_p(1,k)$. Using this and calculations that illustrate (A.17), we can show that $n^{-1}\sum_{t=1}^{k} \dot{f}(\xi_{ti})\hat{\epsilon}_t = \mathcal{O}_p(1,k)$ when $\mathbf{\Delta}_x \neq \mathbf{0}$. Similarly, $n^{-1}\sum_{t=1}^{k} \dot{f}(\xi_{tj})\hat{Z}_t = \mathcal{O}_p(1,k)$, which follows from the application of (A.15). Consequentially,

$$\mathcal{A}_k = o_p(n^{1/\nu}, k). \tag{A.18}$$

Using (A.17), (A.18), and the fact that $n^{-1}\mathcal{B}_c(k) = \mathcal{O}_p(1,k)$, we see that $n^{-1/2}\mathbf{R}_{\mathrm{x},k}$ diverges at rate of $n^{1/2}$ if $\mathbf{\Delta}_x \neq \mathbf{0}$, which proves that $\lim_{n\to\infty} P(\widehat{L}_{\mathrm{x},k} > c_\alpha) = 1$ for any constant c_α.

To illustrate consistency of $\arg\max_k \widehat{L}_{\mathrm{x},k}$ as an estimator of the changepoint time, note that $(\mathcal{B}_c(k))'\mathbf{C}_{\mathrm{x},k}^{-1}\mathcal{B}_c(k)$ is maximized when $k = c$ by Lemma A.2 of Bai (1997). This, (A.18), and Lemma A.4 of Bai (1997) show that $\arg\max_k \widehat{L}_{\mathrm{x},k} \xrightarrow{\mathcal{P}} \kappa$, where $c/n \to \kappa$.

Next, we focus on the case where $\mathbf{b}_t = \mathbf{s}_t$ or $\mathbf{b}_t = \mathbf{v}_t$. Formula (A.17) and calculations akin to those which provide (A.14) imply

$$\begin{aligned}\sum_{t=1}^{k} \mathbf{b}_t\hat{\epsilon}_{t-i} - \frac{k}{n}\sum_{t=1}^{n} \mathbf{b}_t\hat{\epsilon}_{t-i} \quad &= \quad n\min\left\{\frac{k}{n}, \frac{c}{n}\right\}\left(1 - \max\left\{\frac{k}{n}, \frac{c}{n}\right\}\right)\mathbf{\Gamma}_b(i)\mathbf{\Delta}_b \\ &\sum_{t=1}^{k} \mathbf{b}_t\epsilon_{t-i} - \frac{k}{n}\sum_{t=1}^{n} \mathbf{b}_t\epsilon_{t-i} + o_p(\sqrt{n}, k).\end{aligned}$$

So long as $\{\hat{\pi}_j\}$ are reasonable approximations under the alternative hypothesis, we mimic (A.16) to yield

$$\begin{aligned}\sum_{t=1}^{k}\mathbf{b}_t\hat{Z}_t - \frac{k}{n}\sum_{t=1}^{n}\mathbf{b}_t\hat{Z}_t &= n\min\left\{\frac{k}{n},\frac{c}{n}\right\}\left(1-\max\left\{\frac{k}{n},\frac{c}{n}\right\}\right)\sum_{j=0}^{\infty}\pi_j\mathbf{\Gamma}_b(j)\mathbf{\Delta}_b \\ &\quad + \sum_{t=1}^{k}\mathbf{b}_t Z_t - \frac{k}{n}\sum_{t=1}^{n}\mathbf{b}_t Z_t + o_p(\sqrt{n},k).\end{aligned}$$

Therefore, $n^{-1/2}\mathbf{R}^*_{\mathrm{s},k}$ diverges at rate of $n^{1/2}$ if $\mathbf{\Delta}_s \neq \mathbf{0}$. Note that $\min\{k,c\}(n-\max\{k,c\})$ is maximized when $k = c$. Furthermore, from Lemma A.4 of Bai (1997), it holds that $\arg\max_k \widehat{L}_{\mathrm{s},k} \xrightarrow{\mathcal{P}} \kappa$, where $c/n \to \kappa$. Analogous results hold for $\mathbf{R}^*_{\mathrm{v},k}$ and $\arg\max_k \widehat{L}_{\mathrm{v},k}$.

□

Corollary A.1. We first assume that the seasonal terms obey the harmonic representation in (A.3). Basic trigonometric identities can be used to establish that $\mathbf{D}_T = \mathbf{I}_{q_s}/2$. This result in combination with the observation that $\sum_{t=1}^{n}\mathbf{s}_t\hat{Z}_t = \mathcal{O}_p(1)$ establishes the finding of Corollary A.1. To illustrate the latter formula, we establish that $\sum_{t=1}^{n}\mathbf{s}_t\hat{\epsilon}_{t-i} = \mathcal{O}_p(1)$ for all $i \geq 0$; the invertibility expansions used in (A.16) can then be applied in order to show that $\sum_{t=1}^{n}\mathbf{s}_t\hat{Z}_t = \mathcal{O}_p(1)$.

Define

$$\mathbf{H}_{j,i} = \begin{pmatrix} \cos(2\pi ji/T) & -\sin(2\pi ji/T) \\ \sin(2\pi ji/T) & \cos(2\pi ji/T) \end{pmatrix},$$

and let $\mathbf{H}_i$ denote a block-diagonal matrix of dimension $p_s \times p_s$ written as

$$\mathbf{H}_i = \begin{pmatrix} \mathbf{H}_{j_1,i} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{j_2,i} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H}_{j_J,i} \end{pmatrix}.$$

The matrix $\mathbf{H}_i$ can be used to establish a recursion between the $\mathbf{s}_t$. Specifically,

$$\mathbf{H}_i \mathbf{s}_t = \mathbf{s}_{t+i}. \tag{A.19}$$

Further, note that $\mathbf{H}_i = 2\sum_{t=1}^{T} \mathbf{s}_{t+i}\mathbf{s}_t'/T$. Define

$$\begin{aligned} \mathbf{N}_{\mathrm{s},n}(i) := \sum_{t=1}^{n} \mathbf{s}_{t+i}\hat{\epsilon}_t &= \mathbf{S}_{n,i}'\boldsymbol{\epsilon} - \mathbf{S}_{n,i}'\widetilde{\mathbf{M}}_n(\widetilde{\mathbf{M}}_n'\widetilde{\mathbf{M}}_n)^{-1}\widetilde{\mathbf{M}}_n\boldsymbol{\epsilon} \\ &= \mathbf{S}_{n,i}'\boldsymbol{\epsilon} - \mathbf{S}_{n,i}'\widetilde{\mathbf{S}}_n(\widetilde{\mathbf{S}}_n'\widetilde{\mathbf{S}}_n)^{-1}\widetilde{\mathbf{S}}_n\boldsymbol{\epsilon} + \mathcal{O}_p(1), \end{aligned}$$

where $\mathbf{S}_{n,i} = (\mathbf{s}_{1+i}, \ldots, \mathbf{s}_{n+i})'$, for $t = 1, \ldots, n$.

It follows that

$$\widetilde{\mathbf{S}}_n'\widetilde{\mathbf{S}}_n/n = \mathbf{I}_{p_s}/2 + \mathcal{O}(n^{-1}), \quad \text{and} \quad \mathbf{S}_{n,i}'\widetilde{\mathbf{S}}_n/n = \begin{pmatrix} \mathbf{H}_i/2 & \mathbf{0} \end{pmatrix} + \mathcal{O}(n^{-1}).$$

Continuing,

$$\begin{aligned} \mathbf{N}_{\mathrm{s},n}(i) &= \sum_{t=1}^{n} \mathbf{s}_{t+i}\epsilon_t - \mathbf{H}_i \sum_{t=1}^{n} \mathbf{s}_t\epsilon_t + \mathcal{O}_p(1) \\ &= \mathbf{0} + \mathcal{O}_p(1), \end{aligned} \tag{A.20}$$

which is derived using the recursion in (A.19). The above also implies that $\sum_{t=1}^{n} \mathbf{s}_t \hat{\epsilon}_{t-i} = \mathcal{O}_p(1)$.

Next, assume that $\mathbf{s}_t$ satisfies the formulation in (A.4). Then, it holds that

$$\mathbf{S}'_{n,i}\mathbf{S}_n(\mathbf{S}'_n\mathbf{S}_n)^{-1} = \mathbf{H}^*_i + \mathcal{O}(n^{-1}),$$

where $\mathbf{H}^*_i$ is defined as follows. First,

$$\mathbf{H}^*_1 = \begin{pmatrix} -\mathbf{1}' & -1 \\ \mathbf{I}_{T-2} & \mathbf{0} \end{pmatrix},$$

where $\mathbf{1}$ is a vector of ones. Next, $\mathbf{H}^*_2$ is defined by permuting the rows of $\mathbf{H}_1$ so that the bottom row of $\mathbf{H}^*_2$ equals the top row of $\mathbf{H}^*_1$ and the remaining rows of $\mathbf{H}^*_1$ are each shifted down. In general, define $\mathbf{H}^*_i$ by permuting the rows of $\mathbf{H}^*_{i-1}$ in a similar manner. A recursion analogous to (A.19) can be established: $\mathbf{H}^*_i\mathbf{s}_t = \mathbf{s}_{t+i}$. Therefore, (A.20) is again satisfied, and the proof is completed. □

Appendix G: Parameter Estimates for Data Examples

OLS parameter estimates under both the null and alternative are provided in Table A.1 for the Mauna Loa data example and in Table A.2 for the Barrow, AK data example. (Notation in the tables is in line with the notation provided in the article. That is, α's govern trend, β's govern seasonality, γ's govern covariates, and Δ's quantify changes. For example, $\hat{\Delta}_{s,1,1}$ is the post-changepoint change in the parameter $\hat{\beta}_{1,1}$.)

Table A.1: Parameter estimates for the Mauna Loa data analysis

	Coef.	Estimate
H0	$\hat{\alpha}_1$	315.086
	$\hat{\alpha}_2$	34.833
	$\hat{\alpha}_3$	62.580
	$\hat{\alpha}_4$	1.763
	$\hat{\alpha}_5$	1.664
	$\hat{\alpha}_6$	-15.103
	$\hat{\beta}_{1,1}$	2.552
	$\hat{\beta}_{2,1}$	-0.653
	$\hat{\beta}_{1,2}$	0.022
	$\hat{\beta}_{2,2}$	-0.057
	$\hat{\beta}_{1,3}$	1.120
	$\hat{\beta}_{2,3}$	-0.413
	$\hat{\beta}_{1,4}$	-0.086
	$\hat{\beta}_{2,4}$	0.044
	$\hat{\gamma}_1$	-0.015
H1c*	$\hat{\alpha}_1$	315.088
	$\hat{\alpha}_2$	34.819
	$\hat{\alpha}_3$	62.586
	$\hat{\alpha}_4$	1.793
	$\hat{\alpha}_5$	1.628
	$\hat{\alpha}_6$	-15.098
	$\hat{\beta}_{1,1}$	2.424
	$\hat{\beta}_{2,1}$	-0.614
	$\hat{\beta}_{1,2}$	0.018
	$\hat{\beta}_{2,2}$	-0.049
	$\hat{\beta}_{1,3}$	0.890
	$\hat{\beta}_{2,3}$	-0.335
	$\hat{\beta}_{1,4}$	-0.096
	$\hat{\beta}_{2,4}$	0.086
	$\hat{\gamma}_1$	-0.014
	$\widehat{\Delta}_{s,1,1}$	0.186
	$\widehat{\Delta}_{s,2,1}$	-0.058
	$\widehat{\Delta}_{s,1,2}$	0.006
	$\widehat{\Delta}_{s,2,2}$	-0.012
	$\widehat{\Delta}_{s,1,3}$	0.336
	$\widehat{\Delta}_{s,2,3}$	-0.113
	$\widehat{\Delta}_{s,1,4}$	0.014
	$\widehat{\Delta}_{s,2,4}$	-0.061

Table A.2: Parameter estimates for the Barrow, AK data analysis

	Coef.	Estimate
H0	$\hat{\alpha}_1$	-13.003
	$\hat{\alpha}_2$	2.072
	$\hat{\beta}_1$	-8.151
	$\hat{\beta}_2$	-17.113
	$\hat{\beta}_3$	-22.994
	$\hat{\beta}_4$	-25.149
	$\hat{\beta}_5$	-26.663
	$\hat{\beta}_6$	-25.111
	$\hat{\beta}_7$	-17.225
	$\hat{\beta}_8$	-6.285
	$\hat{\beta}_9$	1.796
	$\hat{\beta}_{10}$	4.885
	$\hat{\beta}_{11}$	4.095
	$\hat{\gamma}_1$	-0.017
H1b	$\hat{\alpha}_1$	-12.263
	$\hat{\alpha}_2$	-0.169
	$\hat{\beta}_1$	-7.731
	$\hat{\beta}_2$	-16.982
	$\hat{\beta}_3$	-23.250
	$\hat{\beta}_4$	-25.322
	$\hat{\beta}_5$	-27.109
	$\hat{\beta}_6$	-25.276
	$\hat{\beta}_7$	-17.045
	$\hat{\beta}_8$	-6.445
	$\hat{\beta}_9$	1.806
	$\hat{\beta}_{10}$	5.030
	$\hat{\beta}_{11}$	4.340
	$\hat{\gamma}_1$	-0.017
	$\widehat{\Delta}_{x,1}$	-4.616
	$\widehat{\Delta}_{x,2}$	7.312
	$\widehat{\Delta}_{s,1}$	-0.787
	$\widehat{\Delta}_{s,2}$	-0.259
	$\widehat{\Delta}_{s,3}$	0.442
	$\widehat{\Delta}_{s,4}$	0.302
	$\widehat{\Delta}_{s,5}$	0.788
	$\widehat{\Delta}_{s,6}$	0.282
	$\widehat{\Delta}_{s,7}$	-0.354
	$\widehat{\Delta}_{s,8}$	0.266
	$\widehat{\Delta}_{s,9}$	-0.050
	$\widehat{\Delta}_{s,10}$	-0.256
	$\widehat{\Delta}_{s,11}$	-0.443
	$\widehat{\Delta}_{v,1}$	-0.007

Supplement References

Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics* **79**, 551–563.

Robbins, M., Gallagher, C., Lund, R. and Aue, A. (2011). Mean shift testing in correlated data. *Journal of Time Series Analysis* **32**, 498–511.

Robbins, M. W., Gallagher, C. M. and Lund, R. B. (2016). A general regression changepoint test for time series data. *Journal of the American Statistical Association* **111**, 670–683.