Supplementary Material for “A General Regression Changepoint Test for Time Series Data”

Here, we begin with illustration of calculations used to derive the form of the quantities $\Delta_k$, $\text{Var}(\Delta_k)$ and (by extension) $\hat{F}_k$ as introduced in the main article. Then, we discuss a pair of generalizations of the model in (1). Specifically, we consider changepoint alternatives that allow the coefficients in the seasonal regressors to change and settings where the stochastic predictors have a changepoint. However, we find that the mathematical simplifications of the methods in the article do not hold in these more general circumstances. The following includes references to notation, equations, lemmas, and theorems from the main article.

Linear Models Theory

Assuming that the changepoint occurs at time $k$, (1) becomes

$$Y_t = \theta'x_t + \Delta'x_tI\{t>k\} + \epsilon_t,$$

where $\theta = \{\alpha', \beta', \gamma', \zeta'\}'$, $x_t = \{x_t', w_t', s_t', v_t'\}'$, and $I_A$ is the indicator of event $A$. In the notation of Section 2.1, we have

$$Y = M_0\theta + \mathcal{M}_k\Delta + \epsilon,$$

where $M_0$ denotes the design matrix under the no changepoint null hypothesis, $\mathcal{M}_k = X_n - X_k$ for $X_t$ in accordance with (13), and $Y$ and $\epsilon$ are response and regression error vectors, respectively. Under the null hypothesis model ($\Delta \equiv 0$), the least squares prediction vector of $Y$ is found by projecting $Y$ onto the column space of $M_0$. That is, $\hat{Y}_0 = P_0Y$, where $P_0$ is the orthogonal projection matrix under $H_0$.

Linear models theory indicates that we can decompose the prediction vector under the
full (i.e., alternative hypothesis) model which allows for changepoints at time $k$ as

$$\hat{Y}_a = \hat{Y}_0 + \hat{Y}_k,$$  \hspace{1cm} (B.1)

which is the sum of two orthogonal vectors with $\hat{Y}_k = P_k Y = (P_a - P_0)Y$, where $P_a$ denotes the projection matrix for the full model. Since the columns of $X_n$ are in the column space of $M_0$, $M_k^\perp := (I - P_0)M_k = -(I - P_0)X_k$. Thus the vector $\hat{Y}_k$, which consists of the projection of $Y$ onto the column space of $M_k^\perp$, satisfies

$$\hat{Y}_k = P_k Y = M_k^\perp [(M_k^\perp)'M_k^\perp]^{-1}(M_k^\perp)' Y$$

$$= (I - P_0)M_k\{(I - P_0)X_k)^{-1}X'_k(I - P_0)Y\}.$$  \hspace{1cm} (B.2)

Since we can re-express (B.1) as

$$\hat{Y}_a = M_0\hat{\theta}_k + M_k\hat{\Delta}_k,$$

where $\hat{\theta}_k$ and $\hat{\Delta}_k$ are the respective (alternative hypothesis) least squares estimates of $\theta$ and $\Delta$, the uniqueness of orthogonal projections and least squares estimates, in conjunction with (B.2), implies that

$$\hat{\Delta}_k = -(X'_k(I - P_0)X_k)^{-1}X'_k(I - P_0)Y,$$

which has variance/covariance matrix

$$\text{Var}(\hat{\Delta}_k) = \tau^2 (X'_k(I - P_0)X_k)^{-1}. \hspace{1cm} (B.3)$$

Derivation of the above uses that $I - P_0$ is symmetric and idempotent. Using (7) and (B.3), the form of $C_k$ in (14) is established.

A Gaussian generalized likelihood ratio test for no change null against the alternative
that a changepoint exist at time \( k \) can be based on statistic
\[
\hat{F}_k = \hat{\Delta}'_k (X_k^T (I - P_0) X_k) \hat{\Delta}_k / \hat{\tau}^2 = \frac{||\hat{Y}_k||^2}{\hat{\tau}^2},
\]
and it follows that
\[
||\hat{Y}_k||^2 = Y'(I - P_0) X_k (X_k' (I - P_0) X_k)^{-1} X_k' (I - P_0) Y = N_k^T C_k^{-1} N_k.
\]

A Changepoint in Coefficients Governing the Seasonal Cycle

We now consider a more general model with seasonal changepoints:
\[
Y_t = x_t'(\alpha + \delta_t) + w_t' \beta + s_t' (\gamma + \delta_t^*) + (s_t^*)' \gamma^* + v_t' \zeta + \epsilon_t,
\] (B.4)
where \( s_t \) is a length-\( p_s \) vector of deterministic time-\( t \) seasonal effects that are allowed to change under \( H_1 \) and \( s_t^* \) is a vector of time-\( t \) seasonal components that are not allowed to change under \( H_1 \). The remaining predictor terms are defined below (1). The shifts are
\[
\delta_t = \begin{cases} 
0, & 1 \leq t \leq c \\
\Delta, & c < t \leq n 
\end{cases},
\]
and
\[
\delta_t^* = \begin{cases} 
0, & 1 \leq t \leq c \\
\Delta^*, & c < t \leq n 
\end{cases}.
\]
Defining \( \Delta = (\Delta', (\Delta^*)')' \), the null hypothesis stipulates that \( \Delta = 0 \), which is to be tested against an alternative that \( \Delta \neq 0 \).
IID Regression Errors

To begin, the errors \( \{\epsilon_t\} \) are assumed IID with \( \text{Var}(\epsilon_t) = \tau^2 \) and are taken as independent of all other stochastic components in the model. Let

\[
X_t = (x_1, \ldots, x_t, 0, \ldots, 0)' \quad \text{and} \quad S_t = (s_1, \ldots, s_t, 0, \ldots, 0)'
\]

for \( t = 1, \ldots, n \), which are matrices of dimension \( n \times p_x \) and \( n \times p_s \), respectively, and whose last \( n - t \) rows contain zeros. Further, let \( \tilde{X}_t = (X_t, S_t) \). We estimate \( \tilde{\Delta} \) with

\[
\tilde{\Delta}_k^* = -C_k^{-1}N_k,
\]

which has variance/covariance matrix \( \text{Var}(\tilde{\Delta}_k^*) = \tau^2 C_k^{-1} \), where

\[
C_k = \tilde{X}_k'(I - P_0)\tilde{X}_k, \tag{B.5}
\]

and

\[
N_k = \tilde{X}_k'(I - P_0)Y, \tag{B.6}
\]

with \( Y = (Y_1, \ldots, Y_n)' \). Note that \( P_0 = M_0'(M_0'M_0)^{-1}M_0' \) is the regression projection matrix under the null hypothesis (with \( M_0 \) denoting the full design matrix under \( H_0 \)). Observe that \( \tilde{N}_k \) is equivalently represented as

\[
\tilde{N}_k = \sum_{t=1}^k \hat{\epsilon}_t \tilde{x}_t,
\]

where \( \tilde{x}_t = (x_t', s_t')' \). As before, our Wald-type test statistic can be expressed as

\[
\tilde{F}_k^* = \frac{\tilde{N}_k' \tilde{C}_k^{-1} \tilde{N}_k}{\tau^2},
\]
where $\hat{\tau}^2$ approximates $\tau^2$ under $H_0$. The limit behavior of

$$\hat{F}^* = \max_{\ell \leq k} \hat{F}^*_k$$

is stated as follows.

**Theorem B.1.** Assume that the data $\{Y_t\}$ obey the model in (B.4) and that the errors $\{\epsilon_t\}$ are IID with variance $\tau^2$. Then

$$\hat{F}^* \xrightarrow{D} \sup_{\ell < z < h} \left\{ \tilde{B}(z) + \tilde{B}^*(z) \right\},$$

for independent Gaussian processes $\{\tilde{B}(z)\}_{z=1}^{\infty}$ and $\{\tilde{B}^*(z)\}_{z=0}^{\infty}$, where $\tilde{B}(z)$ is defined in (20) and where

$$\tilde{B}^*(z) = \frac{\sum_{j=1}^{p_2} B^2_j(z)}{z(1-z)}, \quad z \in (0, 1).$$

Furthermore, $\{B_1(z), \ldots, B_{p_2}(z)\}$ are independent Brownian bridges, each of which can be represented as $B_j(z) = W_j(z) - zW_j(1)$ for a Brownian motion $\{W_j(z)\}$ defined over $z \in [0, 1]$.

**Proof.** As in the proof of Lemma 2.1, let $\tilde{\chi}_t = (x'_t, w'_t)'$ and $\tilde{M}_0 = (\tilde{\chi}_1, \ldots, \tilde{\chi}_n)'$ (this is a version of $M_0$ with only the columns corresponding to the functional form predictors $x_t$ and $w_t$). Furthermore, let $\tilde{\chi}_t^* = (s'_t, v'_t)'$ and $\tilde{M}_0^* = (\tilde{\chi}_1^*, \ldots, \tilde{\chi}_n^*)'$ (so that $M_0 = (\tilde{M}_0, \tilde{M}_0^*)$).

In the proof of Lemma 2.1, it was shown that

$$(n^{-1}M_0'M_0)^{-1} = \begin{pmatrix} (n^{-1}\tilde{M}_0'M_0)^{-1} & 0 \\ 0 & [n^{-1}(\tilde{M}_0^*)'\tilde{M}_0^*]^{-1} \end{pmatrix} + o_p(n^{-1/2}).$$

Likewise,

$$n^{-1}\tilde{X}_k' M_0 = n^{-1} \begin{pmatrix} X'_k \tilde{M}_0 & X'_k \tilde{M}_0^* \\ S'_k \tilde{M}_0 & S'_k \tilde{M}_0^* \end{pmatrix} = \begin{pmatrix} n^{-1}X'_k \tilde{M}_0 & 0 \\ 0 & n^{-1}S'_k \tilde{M}_0^* \end{pmatrix} + o_p(1, k),$$

(B.7)
where $o_p(1, k)$ denotes a random matrix indexed at time $k$ (for example, $A_k$) that satisfies $\max_{\ell \leq n \leq h} \|A_k\|_\infty = o_p(1)$. Lemmas A.1 and A.2 show that the off-diagonal blocks of (B.7) are $o_p(1, k)$. Continuing, we see

$$\tilde{X}_k'M_0(M_0'M_0)^{-1} = \begin{pmatrix} X_k'M_0(\tilde{M}_0'M_0)^{-1} & 0 \\ 0 & S_k'M_0[(\tilde{M}_0')\tilde{M}_0]^*-1 \end{pmatrix} + o_p(1, k).$$

Hence,

$$n^{-1}\tilde{C}_k = \begin{pmatrix} n^{-1}C_k & 0 \\ 0 & n^{-1}C_k^* \end{pmatrix} + o_p(1, k), \quad (B.8)$$

where

$$C_k = X_k'X_k - X_k'\tilde{M}_0(\tilde{M}_0'M_0)^{-1}\tilde{M}_0'X_k$$

with

$$C_k^* = S_k'S_k - S_k'M_0[(\tilde{M}_0')\tilde{M}_0]^*-1(\tilde{M}_0')S_k$$

for $\tilde{C}_k$ as in (B.5). Further,

$$n^{-1}\tilde{N}_k = \begin{pmatrix} n^{-1}N_k \\ n^{-1}N_k^* \end{pmatrix} + o_p(1, k), \quad (B.9)$$

where

$$N_k = X_k'\epsilon - X_k'\tilde{M}_0(\tilde{M}_0'M_0)^{-1}\tilde{M}_0'\epsilon$$

and

$$N_k^* = S_k'\epsilon - S_k'M_0[(\tilde{M}_0')\tilde{M}_0]^*-1(\tilde{M}_0')\epsilon,$$

for $\tilde{N}_k$ as in (B.6) and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)'$. The processes $\{C_k\}$ and $\{N_k\}$ were studied in the proof of Lemma 2.1; our focus now turns to $C_k^*$ and $N_k^*$. 

6
Observe that

$$n^{-1}(\mathbf{M}_0^* | \mathbf{M}_0^*) = \begin{pmatrix} S_n^\prime S_n/n & 0 \\ 0 & (S_n^* | S_n^* / n) \end{pmatrix} \quad \text{and} \quad n^{-1} S_k^\prime \mathbf{M}_0^* = \begin{pmatrix} S_k^\prime S_k/n & 0 \end{pmatrix}$$

in the limit, where $S_n^*$ denotes the columns of $\mathbf{M}_0^*$ not contained in $S_n$. This uses the observations that $n^{-1} \sum_{t=1}^{k} s_t v_t = \alpha_p(1, k)$ and $\sum_{t=1}^{T} s_i \epsilon_{s_j t} = 0$ for $i \neq j$, where we have expressed $s_t = (s_{1, t}, \ldots, s_{p, t})$. The orthogonality formula holds if all seasonal components are harmonic (this property is not needed if all regression coefficients for the seasonal terms are allowed to change). Thus,

$$n^{-1} C_k^* = S_k^\prime S_k/n - (S_k^\prime S_k/n)(S_n^\prime S_n/n)^{-1}(S_n^\prime S_n/n) + \alpha_p(1, k).$$

Likewise,

$$n^{-1/2} N_k^* = S_k^\prime \epsilon / \sqrt{n} - (S_k^\prime S_k/n)(S_n^\prime S_n/n)^{-1}(S_n^\prime \epsilon / \sqrt{n}) + \alpha_p(1, k).$$

Let $D_T = \sum_{j=1}^{T} s_j s_j^\prime$. Without loss of generality, we assume that $k = Td^*$ and $n = Td$, so that if $k = nz$, then $z = d^*/d$. Consequently, $S_k^\prime S_k/d \xrightarrow{p} zD_T$ and therefore

$$\frac{T}{n} C_k^* \xrightarrow{p} zD_T - zD_T D_T^{-1} zD_T = z(1 - z)D_T,$$

for $k/n = z \in (0, 1)$. Let $e_i = \sum_{t=T(i-1)+1}^{T(i+1)} s_t \epsilon_t$ and note that $S_k^\prime \epsilon = \sum_{i=1}^{d^*} e_i$. Notice further that $\text{Var}(e_i) = \tau^2 D_T$. Hence,

$$\frac{T}{n}^{1/2} N_k^* = \frac{1}{\sqrt{d}} \sum_{i=1}^{d^*} e_i - (d^*/d) D_T(D_T)^{-1} \left( \frac{1}{\sqrt{d}} \sum_{i=1}^{d} e_i \right) + \alpha_p(1, k)$$

$$= \frac{1}{\sqrt{d}} \sum_{i=1}^{d^*} e_i - z \left( \frac{1}{\sqrt{d}} \sum_{i=1}^{d} e_i \right) + \alpha_p(1, k).$$
Lastly,
\[ \left( \tau n D_T \right)^{-1/2} N_k^* \Rightarrow W_{p_s}(z) - z W_{p_s}(1), \]  
(B.11)
where \( W_j(z) \) is a vector containing \( j \) independent Brownian motions evaluated at \( z \).

From (B.8) and (B.9), it follows that
\[ \frac{N_k'}{\tau^2} \tilde{C}_k^{-1/2} N_k = \frac{N_k' C_k^{-1} N_k}{\tau^2} + \frac{(N_k^*)'(C_k^*)^{-1} N_k^*}{\tau^2} + o_p(1, k). \]

The limit behavior of the term involving \( N_k \) follows from Theorem 2.1, and the limit behavior of the term involving \( N_k^* \) follows from (B.10) and (B.11). To conclude the proof, we note that the asymptotic independence of these two terms is given by the block-diagonal form of \( \text{Var}(\tilde{N}_k) = \tau^2 \tilde{C}_k \) as \( n \to \infty \), which is evident in (B.8).

Correlated Errors

We next discuss extension of the above to settings involving autocorrelated errors. Versions of Corollary 2.1 and Lemma 2.2 constructed in consideration of tests for changes in coefficients governing seasonal (or stochastic) predictor terms do not hold in general; new approaches are needed to extract that limit behavior of test statistics which are of the flavor of those constructed in the article. Such pursuits indicate that it does not appear these statistics will yield tractable limit distributions in general settings. In the following, we illustrate the complications that arise.

When \( \{\epsilon_t\} \) is IID, the above proof shows that
\[ \text{Var}(N_k^*/\sqrt{n}) \to \tau^2 z(1 - z) D_T/T. \]  
(B.12)
for \( k/n = z \in (0, 1) \). We note that there do not appear to be simplifications of \( \text{Var}(N_k^*) \) in general. For instance, one can not simply replace \( \tau \) with the quantity in (4) (or even a proportional term) to procure a valid result for general \( T \) in settings involving autocorrelated regression errors. Nonetheless, since \( \{n^{-1/2} \sum_{t=1}^{k} s_t \epsilon_t\} \) will obey a multivariate functional
central limit theorem and since \( S_k' S_k / n \xrightarrow{p} z D_T / T \), it might be feasible to develop a statistic that will have the limit distribution in Theorem B.1 in such settings.

To develop a procedure based on \( \{ \hat{Z}_t \} \), the ARMA residuals, one should first examine the process \( \{ n^{-1/2} \sum_{t=1}^k s_t \hat{Z}_t \} \). The behavior of this process cannot be derived directly through \( \{ n^{-1/2} \sum_{t=1}^k s_t \hat{\epsilon}_t \} \) since a version of Lemma 2.2 involving \( s_t \) does not hold in general. There are some circumstances in which \( \{ n^{-1/2} \sum_{t=1}^k s_t \hat{Z}_t \} \) is underlaid by a tractable process and some circumstances in which it is not. We illustrate this in a simplified setting. Specifically, we incorporate a univariate seasonal term in addition to an intercept parameter while enabling autoregressive errors. That is,

\[
Y_t = \mu + \beta s_t + \epsilon_t,
\]

where \( \sum_{t=1}^T s_t = 0 \), \( s_{t+T} = s_t \), and \( \{ \epsilon_t \} \) obeys an AR(1) process. Recall that the ARMA residuals are then calculated via \( \hat{Z}_t = \hat{\epsilon}_t - \hat{\phi} \hat{\epsilon}_{t-1} \), where \( \{ \hat{\epsilon}_t \} \) is the sequence of OLS residuals.

Continuing,

\[
\frac{1}{\sqrt{n}} \sum_{t=1}^k s_t \hat{Z}_t = \frac{1}{\sqrt{n}} \left( \sum_{t=1}^k s_t \hat{\epsilon}_t - \hat{\phi} \sum_{t=1}^k s_t \hat{\epsilon}_{t-1} \right) + o_p(1, k).
\]

The above uses the fact that \( |\hat{\phi} - \phi| = o_p(1) \). Calculations show that

\[
\frac{1}{\sqrt{n}} \sum_{t=1}^k s_t \hat{\epsilon}_t = \frac{1}{\sqrt{n}} \left( \sum_{t=1}^k s_t \epsilon_t - \frac{k}{n} \sum_{t=1}^n s_t \epsilon_t \right) + o_p(1, k)
\]

and that

\[
\frac{1}{\sqrt{n}} \sum_{t=1}^k s_t \hat{\epsilon}_{t-1} = \frac{1}{\sqrt{n}} \left( \sum_{t=1}^k s_t \epsilon_{t-1} - \frac{k}{n} A_T \sum_{t=1}^n s_t \epsilon_t \right) + o_p(1, k),
\]

where \( A_T = \sum_{t=1}^T s_t s_{t-1} / \sum_{t=1}^T s_t^2 \).

If \( T = 2 \), it follows that \( A_T = -1 \) and that \( -s_t = s_{t+1} \). Hence,

\[
\frac{1}{\sqrt{n}} \sum_{t=1}^k s_t \hat{\epsilon}_{t-1} = \frac{1}{\sqrt{n}} \left( \sum_{t=1}^k s_t \epsilon_{t-1} - \frac{k}{n} \sum_{t=1}^n s_t \epsilon_{t-1} \right) + o_p(1, k). \tag{B.13}
\]
Furthermore,

\[
\frac{1}{\sqrt{n}} \sum_{t=1}^{k} s_t \hat{Z}_t = \frac{1}{\sqrt{n}} \left[ \left( \sum_{t=1}^{k} s_t \epsilon_t - \frac{k}{n} \sum_{t=1}^{n} s_t \epsilon_t \right) - \phi \left( \sum_{t=1}^{k} s_t \epsilon_{t-1} - \frac{k}{n} \sum_{t=1}^{n} s_t \epsilon_{t-1} \right) \right] + o_p(1, k)
\]

\[
= \frac{1}{\sqrt{n}} \left( \sum_{t=1}^{k} s_t Z_t - \frac{k}{n} \sum_{t=1}^{n} s_t Z_t \right) + o_p(1, k). \tag{B.14}
\]

The above implies that \( \{n^{-1/2} \sum_{t=1}^{k} s_t \hat{Z}_t \} \) will converge to a Brownian bridge when scaled appropriately.

For general \( T \),

\[
\frac{1}{\sqrt{n}} \sum_{t=1}^{k} s_t \hat{Z}_t = \frac{1}{\sqrt{n}} \left( \sum_{t=1}^{k} s_t Z_t - \frac{k}{n} \sum_{t=1}^{n} s_t Z_t \right) + \phi \frac{1}{\sqrt{n}} \left( \frac{k}{n} \sum_{t=1}^{n} (s_{t+1} - A_T s_t) \epsilon_t \right) + o_p(1, k). \tag{B.15}
\]

We are no longer assured that \( A_T s_t = s_{t+1} \) if \( T > 2 \). Therefore, in general circumstances, \( \{n^{-1/2} \sum_{t=1}^{k} s_t \hat{Z}_t \} \) will observe an unfamiliar limiting process. Specifically, the second term on the right-hand side of (B.15) is \( O_p(1) \), implying that \( \{n^{-1/2} \sum_{t=1}^{k} s_t \hat{Z}_t \} \) is not tied down at \( k = n \).

However, the finding in (B.14) extends to \( T > 2 \) so long as the \( H_0 \) regression model allows each season to have a unique parameter that controls its mean. That is, in the event that \( \{Y_t\} \) has only a seasonal mean sequence (no trend terms or stochastic predictors are used) and when each season is modeled within the regression (i.e., the null hypothesis design matrix is \( T \) dimensional, consisting of one dummy variable for each season), a desirable limit law is extracted. To illustrate the specifics, we assume IID errors and that \( d \) and \( d^* \) satisfy \( n = Td \) and \( k = Td^* \), respectively. The Wald statistic for a change in mean of all seasons at time \( k \) can be written

\[
\hat{F}_k = \frac{1}{\tau^2 d} \sum_{j=1}^{T} \left( \sum_{\ell=0}^{d^*-1} \hat{\epsilon}_{\ell+j} \right)^2 = \frac{1}{\tau^2 d} \sum_{j=1}^{T} \left( \sum_{\ell=0}^{d^*-1} \epsilon_{\ell+j} - d^* \frac{d-1}{d} \sum_{\ell=0}^{d-1} \epsilon_{\ell+j} \right)^2.
\]

The limit of the process \( \{\hat{F}_k\} \) has the form \( \sum_{j=1}^{T} B_j^2(z)/[z(1 - z)] \), where the \( \{B_j(z)\} \)
for \( z \in (0, 1) \) are independent Brownian bridges. If \( \{\epsilon_t\} \) obeys an AR(1) formulation, straightforward algebra shows

\[
\frac{1}{\sqrt{d}} \sum_{\ell=0}^{d^r-1} \hat{Z}_{\ell+j} = \frac{1}{\sqrt{d}} \left( \sum_{\ell=0}^{d^r-1} Z_{\ell+j} - \frac{d^r}{d} \sum_{\ell=0}^{d^r-1} Z_{\ell+j} \right) + o_p(1, k),
\]

where \( \{\hat{Z}_t\} \) is the sequence of ARMA residuals. Hence, the ARMA residuals test becomes

\[
\hat{L}_k = \frac{1}{\sigma^2 d} \sum_{j=1}^{T} \left( \sum_{\ell=0}^{d^r-1} \hat{Z}_{\ell+j} \right)^2 = \frac{1}{\sigma^2 d} \sum_{j=1}^{T} \left( \sum_{\ell=0}^{d^r-1} Z_{\ell+j} - \frac{d^r}{d} \sum_{\ell=0}^{d^r-1} Z_{\ell+j} \right)^2 + o_p(1, k),
\]

which observes the same limit process as did \( \{\hat{F}_k\} \) for IID errors.

**Stochastic Covariates**

A result akin to Theorem B.1 can be developed for testing changepoints in regression coefficients corresponding to stochastic predictors. However, the same complication that arises within (B.15) is encountered when addressing a version of (B.4) which allows the coefficients corresponding to \( v_t \) to change. Specifically, there are not clear avenues in which the methods of the main paper can be extended to test for changes in coefficients on stochastic predictors.

**Regressors with Changepoints**

We now move to circumstances where a univariate stochastic covariate \( \{v_t\} \) contains a changepoint at the known time \( m \), where \( 1 < m \leq n \) but is unaccounted for in the analysis. The objective is to illustrate the complications that may arise in such circumstances.

Suppose the analyst fits the model

\[
Y_t = \mu + \zeta v_t + \epsilon_t, \tag{B.16}
\]

but that the stochastic covariate \( \{v_t\} \) in truth has a changepoint at time \( m \) unbeknown to
the researcher: \( v_t = \mu_v + \Delta_v I_{(t \leq m)} + u_t \) for a stationary sequence \( \{u_t\} \) with \( E[u_t] = 0 \) and \( \text{Var}(u_t) = \sigma_u^2 \).

We wish to detect a shift in \( \mu \) at an unknown time \( c \). The above model contains the regression factors \( f_1(t) = 1 \) and \( v_t \). The null hypothesis model for \( \{Y_t\} \) does not satisfy the assumptions in the main article since \( \{v_t\} \) has a nonstationary mean that is not a linear combination of the other predictors included in the model for \( \{Y_t\} \). Nonetheless, we attempt to quantify the large sample behavior of our Wald statistic in this setting. For the sake of ensuing calculations, we scale for \( \Delta_v \):

\[
v_t = \mu_v^* + I_{(t \leq m)} + u_t^*,
\]

where \( E[u_t^*] = 0 \) and \( \text{Var}(u_t^*) = \nu^2 := (\sigma_v / \Delta_v)^2 \). Let \( \chi_t = (1, v_t)' \). The design matrix under \( H_0 \) is \( M_0 = (\chi_1, \ldots, \chi_n)' \). It follows that

\[
\frac{M'_0 M_0}{n} \Rightarrow G^\dagger := \begin{pmatrix} 1 & \omega \\ \omega & \omega + \nu^2 \end{pmatrix},
\]

where \( \omega = \lim_{n \to \infty} m(n)/n \). The above matrix is similar to \( G(1) \) from Example 5. Here, the Wald statistic is still written \( \hat{F}_k = N_k^{-1} C_k^{-1} N_k / \hat{\tau}^2 \), where \( C_k \) and \( N_k \) are defined in (14) and (15), respectively. Since the off-diagonal elements of \( G^\dagger \) are nonzero, \( \{v_t\} \) will influence the limit distribution of both \( C_k \) and \( N_k \). Specifically, computations show

\[
\frac{C[nz]}{n} \Rightarrow \Omega(z) = \begin{cases} \nu^2 z(1-z) + z(1-\omega)(\omega-z) / \omega(1-\omega) + \nu^2, & z \leq \omega, \\ \nu^2 z(1-z) + \omega(1-z)(z-\omega) / \omega(1-\omega) + \nu^2, & z > \omega, \end{cases}
\]

and

\[
\frac{N[nz]}{\sqrt{n}} \Rightarrow \Lambda(z) = \begin{cases} \nu^2 [W(z) - z W(1)] + \omega(1-\omega) W(z) - z(1-\omega) W(\omega) + \nu V(1), & z \leq \omega, \\ \nu^2 [W(z) - z W(1)] + \omega(1-\omega) W(z) - \omega(z-\omega) W(1) - \omega(1-z) W(\omega) + \nu V(1), & z > \omega, \end{cases}
\]
for $z \in (0, 1)$, where \{\textit{W}(z)\} is the limiting process of $\sum_{t=1}^{k} \epsilon_{t}/(\tau \sqrt{n})$ and \{\textit{V}(z)\} is the limiting process of $\sum_{t=1}^{k} \epsilon_{t} u_{t}/(\tau \nu \sqrt{n})$.

The above illustrates that the limit distribution of a statistic for a change in $\mu$ will depend on characteristics of \{\textit{v}_{t}\}. The most efficient manner of obtaining appropriate critical values in this setting seems to involve a bootstrap or related technique. However, the limit distribution here is a fusion of the limit distributions in Examples 1 and 5. Specifically, as $\nu^{2} \to \infty$ (which occurs when either $\Delta v \to 0$ or $\sigma_{v}^{2} \to \infty$), $\Omega(z)$ and $\Lambda(z)$ are equal to their counterparts in Example 1. Likewise, when $\nu^{2} \to 0$ (which occurs when either $\Delta v \to \infty$ or $\sigma_{v}^{2} \to 0$), $\Omega(z)$ and $\Lambda(z)$ are equal to their counterparts in Example 5. Informal simulations show that the Brownian bridge process in Example 1 provides reasonable critical values if $\nu^{-2}$ is moderately close to zero.