

An Asymptotically F-Distributed Chow Test in the Presence of Heteroscedasticity and Autocorrelation*

Yixiao Sun
Department of Economics
UC San Diego, USA

Xuexin Wang
School of Economics and WISE
Xiamen University, China

March 7, 2021

Abstract

This study proposes a simple, trustworthy Chow test in the presence of heteroscedasticity and autocorrelation. The test is based on a series heteroscedasticity and autocorrelation robust variance estimator with judiciously crafted basis functions. Like the Chow test in a classical normal linear regression, the proposed test employs the standard F distribution as the reference distribution, which is justified under fixed-smoothing asymptotics. Monte Carlo simulations show that the null rejection probability of the asymptotic F test is closer to the nominal level than that of the chi-square test.

Keywords: Chow Test, F Distribution, Heteroscedasticity and Autocorrelation, Structural Break.

1 Introduction

For predictive modeling and policy analysis using time series data, it is important to check whether a structural relationship is stable over time. The Chow (1960) test is designed to test whether a break takes place at a given period in an otherwise stable relationship. The test is widely used in empirical applications and has been included in standard econometric textbooks. This paper considers the Chow test in the presence of heteroscedasticity and autocorrelation. There is ample evidence that the Chow test can have very large size distortions if heteroscedasticity and autocorrelation are not accounted for (e.g., Krämer (1989) and Giles and Scott (1992)). Even if we account for them using heteroscedasticity and autocorrelation robust (HAR) variance estimators (e.g., Newey and West (1987) and Andrews (1991)), the test can still over-reject the null hypothesis by a large margin if chi-square critical values are used¹. This is a general problem for any HAR inference, as the chi-square approximation ignores the often substantial finite sample randomness of the HAR variance estimator. To address this problem, the recent literature has developed a new type of asymptotics known as fixed-smoothing asymptotics (see, e.g., Kiefer and Vogelsang (2002a,b, 2005) for early seminal contributions). It is now well known that the

*We would like to thank Esfandiar Maasoumi, the editor, an anonymous associate editor and three anonymous reviewers for their helpful comments. We thank Derrick H. Sun for excellent research assistance.

¹When the Chow test is performed on a single coefficient, normal critical values are typically used on the t statistic. For now, we focus only on the Wald-type Chow test for more than one coefficients so that chi-square critical values are used.

fixed-smoothing asymptotic approximation is more accurate than the chi-square approximation. This has been confirmed by ample simulation evidence and supported by higher-order asymptotic theory in Jansson (2004) and Sun et al. (2008).

In this study, we employ the series HAR variance estimator to implement the Chow test under a known break point in a time series regression where the regressors and regression errors are autocorrelated and the regressor errors may be (conditionally) heteroskedastic. This type of HAR variance estimator is the series counterpart of the kernel HAR variance estimator. Using judiciously crafted orthonormal series in the series HAR variance estimator, we show that the Chow test statistic is asymptotically F-distributed under fixed-smoothing asymptotics. The asymptotic F test is very convenient to use, as the F critical values are readily available from standard statistical tables and programming. This is in contrast to Cho and Vogelsang (2017), where they establish the fixed-b asymptotics for testing structural breaks, employing kernel HAR variance estimators. The fixed-b asymptotic distribution is highly nonstandard, even though it is pivotal. The asymptotic F theory is only possible under series HAR variance estimation.

To establish the asymptotic F theory for the Chow test under fixed-smoothing asymptotics, we have to transform the bases on $L^2[0, 1]$ using a special Gram–Schmidt orthonormalization. The transform ensures that the transformed bases are orthonormal with respect to a special inner product that is built into the problem under consideration. This is in contrast to the existing literature, such as Sun and Kim (2012); Sun (2013); Lazarus et al. (2018); Wang and Sun (2020); Martínez-Iriarte et al. (2020), where the asymptotic F theory is established using the series HAR variance estimator based on orthonormal functions on $L^2[0, 1]$ under the usual inner product. The source of the difference is that the Chow regression involves regressors that are not stationary while the aforementioned papers assume that all regressors are stationary. The regressors in the basic Chow regression are not stationary because they are identically zero before or after the break point. We show that using the orthonormal basis functions on $L^2[0, 1]$ under the usual inner product does not lead to the asymptotic F distribution of the Chow test statistic. For the asymptotic F theory to hold, the special Gram–Schmidt orthonormalization is indispensable.

Monte Carlo simulation experiments show that the F test based on the transformed Fourier bases is as accurate as the nonstandard test based on the usual Fourier bases. The F test and nonstandard test have the same size-adjusted power as the corresponding chi-square tests but much more accurate size. Given its convenience, competitive power, and higher size accuracy, we recommend the F test for practical use.

Our F test theory generalizes the classical Chow test in a linear normal regression where the F distribution is the exact finite sample distribution. The main departures are that we do not make the normality assumption and that we allow for heteroscedasticity and autocorrelation of unknown forms. Without restrictive assumptions such as normality and strict exogeneity, it is in general not possible to obtain the exact finite sample distribution. Instead, we employ the fixed-smoothing asymptotics to show that the Wald statistic is asymptotically F distributed.

This study contributes to the asymptotic F and t test theory in the HAR literature. The structural break regression we consider here can be cast as a linear regression consisting of two special regressors defined in equation (1). One regressor is equal to 1 before the break and follows a stationary process after the break. The other regressor follows a stationary process before the break and is equal to 1 after the break. During the time span of all sample periods, both regressors are then nonstationary. In the nonstationary settings, Sun (2011) and Hwang and Sun (2017) develop the asymptotic F theory in trend regressions and cointegration regressions, respectively. However, neither paper considers the case where the regressors take the special nonstationary

form as we consider here: they are equal to zero during some periods but are stationary during other periods². The paper that is closest to this paper is Liu and Sun (2019), which develops the asymptotic t theory in a simple difference-in-difference (DID) regression where the treatment indicator can be regarded as being subject to a break — it equals 0 before the treatment and 1 after the treatment. In effect, Liu and Sun (2019) has developed the asymptotic t theory in testing whether there is a break in the mean of a time series. Here we consider a more general setting where the effect of a stationary regressor, which includes the constant regressor as a special case, may be subject to a break.

The rest of this paper is organized as follows. Section 2 presents the basic setting and introduces the test statistics. Section 3 establishes the fixed-smoothing asymptotics of the F and t statistics. Section 4 develops asymptotically valid F and t tests. Section 5 extends the basic regression model to include other covariates whose coefficients are known to be stable over time. Section 6 reports the simulation evidence. The last section concludes. Proofs are given in the appendix.

2 Basic Setting and Test Statistics

Given the time series observations $\{X_t \in \mathbb{R}^m, Y_t \in \mathbb{R}\}_{t=1}^T$, we consider the model

$$Y_t = 1 \{t \leq [\lambda T]\} \cdot X_t' \beta_1 + 1 \{t \geq [\lambda T] + 1\} \cdot X_t' \beta_2 + u_t,$$

for $t = 1, 2, \dots, T$ where the unobserved u_t satisfies $E(X_t u_t) = 0$ and $1\{\cdot\}$ is an indicator function. In the above, λ is a known parameter in $(0, 1)$ so that $[\lambda T]$ is the period where the structural break may take place. The effects of X_t on Y_t before and after the break are $\beta_1 \in \mathbb{R}^m$ and $\beta_2 \in \mathbb{R}^m$, respectively. We allow $X_t u_t$ to exhibit autocorrelation and heteroskedasticity of unknown forms. In particular, we allow u_t to be heteroskedastic so that $E(u_t^2 | X_t)$ is a nontrivial function of X_t .

We are interested in testing the null of $H_0: \mathcal{R}\beta_1 = \mathcal{R}\beta_2$ against the alternative $H_1: \mathcal{R}\beta_1 \neq \mathcal{R}\beta_2$ for some $p \times m$ matrix \mathcal{R} . When \mathcal{R} is the $m \times m$ identity matrix, we aim at testing whether β_1 is equal to β_2 . For the moment, we consider the case that all coefficients are subject to a possible break. In Section 5, we consider the case that some of the coefficients are known to be time-invariant.

Let

$$X_{1t} = 1 \{t \leq [\lambda T]\} \cdot X_t, X_{2t} = 1 \{t \geq [\lambda T] + 1\} \cdot X_t. \quad (1)$$

Note that both X_{1t} and X_{2t} are nonstationary. The form of the nonstationarity makes the problem at hand unique. Let $\beta = (\beta_1', \beta_2')'$ and $\tilde{X}_t = (X_{1t}', X_{2t}')'$. Then

$$Y_t = \tilde{X}_t' \beta + u_t,$$

and the hypotheses of interest become $H_0: R\beta = 0$ and $H_1: R\beta \neq 0$ for $R = [\mathcal{R}, -\mathcal{R}] \in \mathbb{R}^{p \times 2m}$.

Denote $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_T)'$, $Y = (Y_1, \dots, Y_T)'$, and $u = (u_1, \dots, u_T)'$. We estimate β by OLS:

$$\hat{\beta} = (\tilde{X}' \tilde{X})^{-1} \tilde{X}' Y.$$

²The structural break regression may also be regarded as a nonlinear regression. When there is only one covariate of interest, it is a nonlinear regression with a stationary covariate. However, our asymptotic theory is based on and tied directly to the perspective of a linear regression with an expanded input space (two separate regressors instead of only one regressor). Depending on how we view the input space and its dimension, a regression can be regarded as linear or nonlinear.

The OLS estimator $\hat{\beta}$ satisfies

$$\sqrt{T}(\hat{\beta} - \beta) = \hat{Q}_{\tilde{X}}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T \tilde{X}_t u_t,$$

where³

$$\hat{Q}_{\tilde{X}} = \frac{\tilde{X}'\tilde{X}}{T} = \begin{pmatrix} T^{-1} \sum_{t=1}^{[T\lambda]} X_t X_t' & O \\ O & T^{-1} \sum_{t=[\lambda]+1}^T X_t X_t' \end{pmatrix}$$

and O is a matrix of zeros. To make inferences on β such as testing whether $R\beta$ is zero, we need to estimate the variance of $T^{-1/2} \sum_{t=1}^T \tilde{X}_t u_t$. To this end, we first construct the residual $\hat{u}_t = Y_t - \tilde{X}_t' \hat{\beta}$, which serves as an estimate for u_t . Given a set of basis functions $\{\phi_j(\cdot)\}_{j=1}^K$, we then construct the series estimator of the variance as

$$\hat{\Omega}_{\tilde{X}} = \frac{1}{K} \sum_{j=1}^K \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j\left(\frac{t}{T}\right) \tilde{X}_t \hat{u}_t \right]^{\otimes 2},$$

where, for a column vector a , $a^{\otimes 2}$ is the outer product of a , that is, $a^{\otimes 2} = aa'$. The asymptotic variance of $R\sqrt{T}(\hat{\beta} - \beta)$ is then estimated by $R\hat{Q}_{\tilde{X}}^{-1}\hat{\Omega}_{\tilde{X}}\hat{Q}_{\tilde{X}}^{-1}R'$.

The Wald statistic for testing $H_0 : R\beta = 0$ against $H_1 : R\beta \neq 0$ is

$$F_T = T \cdot (R\hat{\beta})' \left[R\hat{Q}_{\tilde{X}}^{-1}\hat{\Omega}_{\tilde{X}}\hat{Q}_{\tilde{X}}^{-1}R' \right]^{-1} (R\hat{\beta}).$$

When $p = 1$ and we test $H_0 : R\beta = 0$ against a one-sided alternative, say, $H_1 : R\beta > 0$, we can construct the t statistic:

$$t_T = \frac{\sqrt{T} \cdot R\hat{\beta}}{\left[R\hat{Q}_{\tilde{X}}^{-1}\hat{\Omega}_{\tilde{X}}\hat{Q}_{\tilde{X}}^{-1}R' \right]^{1/2}}.$$

The forms of the F and t statistics are standard. We can also test $H_0 : R\beta = r$ against $H_1 : R\beta \neq r$ for some r . In this case, we can simply replace $R\hat{\beta}$ by $R\hat{\beta} - r$ in the F statistic and the t statistic.

3 Fixed-smoothing Asymptotic Distributions

To establish the asymptotic distributions of F_T and t_T , we maintain the following three assumptions:

Assumption 3.1 $T^{-1} \sum_{t=1}^{[Tr]} X_t X_t' \rightarrow^p Q_{XX} \cdot r$ uniformly over $r \in [0, 1]$ where Q_{XX} is an $m \times m$ invertible matrix.

Assumption 3.2 $T^{-1/2} \sum_{t=1}^{[Tr]} X_t u_t \rightarrow^d \Lambda_X W_m(r)$ for $r \in [0, 1]$ where $\Lambda_X \Lambda_X'$ is the long run variance of $\{X_t u_t\}$ and $W_m(\cdot)$ is an $m \times 1$ standard Brownian process.

Assumption 3.3 The basis functions $\phi_j(\cdot)$, $j = 1, 2, \dots, K$ are piecewise monotonic and piecewise continuously differentiable on $[0, 1]$

³In place of $\hat{Q}_{\tilde{X}}$, we could use $\hat{Q}_{\tilde{X}\tilde{X}}$ but the latter notation is too cumbersome. We opt for a simpler notation when there is no confusion. For the same reason, we use $\hat{\Omega}_{\tilde{X}}$ instead of $\hat{\Omega}_{\tilde{X}U}$ in the next equation.

3.1 and 3.2 are high-level assumptions. Primitive regularity conditions under which these assumptions hold include some moment and mixing conditions; see Phillips and Durlauf (1986); Phillips and Solo (1992); Kiefer and Vogelsang (2002a) for more detailed discussion.

Lemma 3.1 *Let Assumptions 3.1 and 3.2 hold. Then*

$$\sqrt{T}(\hat{\beta} - \beta) := \begin{pmatrix} \sqrt{T}(\hat{\beta}_1 - \beta_1) \\ \sqrt{T}(\hat{\beta}_2 - \beta_2) \end{pmatrix} \rightarrow^d \begin{pmatrix} Q_{XX}^{-1} \Lambda_X \cdot \frac{1}{\lambda} \int_0^\lambda dW_m(\lambda) \\ Q_{XX}^{-1} \Lambda_X \cdot \frac{1}{1-\lambda} \int_\lambda^1 dW_m(\lambda) \end{pmatrix}.$$

If Assumption 3.3 also holds, then

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t \hat{u}_t \rightarrow^d \begin{pmatrix} \Lambda_X \cdot \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_m(r) \\ \Lambda_X \cdot \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_m(r) \end{pmatrix},$$

where

$$\bar{\phi}_{j,1} = \frac{1}{\lambda} \int_0^\lambda \phi_j(s) ds \text{ and } \bar{\phi}_{j,2} = \frac{1}{1-\lambda} \int_\lambda^1 \phi_j(s) ds.$$

Note that $\frac{1}{\lambda} \int_0^\lambda dW_m(\lambda)$ and $\frac{1}{1-\lambda} \int_\lambda^1 dW_m(\lambda)$ are the average changes of the Brownian motion over the intervals $[0, \lambda]$ and $[\lambda, 1]$, respectively. Lemma 3.1 shows that the asymptotic distributions of $\sqrt{T}(\hat{\beta}_1 - \beta_1)$ and $\sqrt{T}(\hat{\beta}_2 - \beta_2)$ are (matrix) proportional to the average changes. Given the independence of the changes of the Brownian motion over any non-overlapping intervals, $\sqrt{T}(\hat{\beta}_1 - \beta_1)$ and $\sqrt{T}(\hat{\beta}_2 - \beta_2)$ are asymptotically independent.

Note that $\bar{\phi}_{j,1}$ can be regarded as an average of $\phi_j(\cdot)$ over the interval $[0, \lambda]$. Similarly, $\bar{\phi}_{j,2}$ can be regarded as an average of $\phi_j(\cdot)$ over the interval $[\lambda, 1]$. So $\phi_j(r) - \bar{\phi}_{j,1}$ and $\phi_j(r) - \bar{\phi}_{j,2}$ are the demeaned versions of $\phi_j(r)$ over the intervals $[0, \lambda]$ and $[\lambda, 1]$, respectively.

Using Lemma 3.1, we can prove our main theorem below.

Theorem 3.1 *Let Assumptions 3.1–3.3 hold. Then, under the null hypothesis,*

$$\begin{aligned} F_T &\rightarrow^d \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right]' \times \left[\frac{1}{K} \sum_{j=1}^K \left\{ \int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \right\}^{\otimes 2} \right]^{-1} \\ &\times \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right] \\ &:= F_\infty, \end{aligned} \tag{2}$$

where

$$\tilde{\phi}_j(r; \lambda) = \frac{1}{\lambda} [\phi_j(r) - \bar{\phi}_{j,1}] \cdot 1\{r \leq \lambda\} - \frac{1}{1-\lambda} [\phi_j(r) - \bar{\phi}_{j,2}] \cdot 1\{r > \lambda\}. \tag{3}$$

When $p = 1$,

$$t_T \rightarrow^d \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right] \times \left[\frac{1}{K} \sum_{j=1}^K \left\{ \int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \right\}^{\otimes 2} \right]^{-1/2} := t_\infty.$$

Like the finite sample distributions, the limiting distributions of F_T and t_T depend on λ and the number and form of the basis functions. This is an attractive feature of the fixed-smoothing approximations, as they capture the effects of all these factors. More importantly, the fixed-smoothing approximations capture the randomness of the HAR variance estimator, which clearly affects the finite sample distributions of F_T and t_T . This is why the fixed-smoothing asymptotic approximations are more accurate than the chi-square or normal approximations.

4 Asymptotic F and t Theory

The limiting distributions F_∞ and t_∞ in Theorem 3.1 are pivotal but nonstandard. We can approximate the nonstandard distributions using a chi-square or t distribution. We can also design a new set of basis functions so that F_∞ and t_∞ become the standard F and t distributions after some multiplicative adjustment.

4.1 Chi-square and normal approximations

Define

$$\tilde{\phi}_0(r; \lambda) = \frac{1}{\lambda} 1\{r \leq \lambda\} - \frac{1}{1-\lambda} 1\{r > \lambda\}.$$

Then

$$\begin{aligned} & \frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \\ &= \int_0^1 \tilde{\phi}_0(r; \lambda) dW_p(r) \sim N\left(0, \int_0^1 [\tilde{\phi}_0(r; \lambda)]^2 dr \cdot I_p\right) \\ &= N\left(0, \frac{1}{\lambda(1-\lambda)} I_p\right), \end{aligned}$$

where I_p is the $p \times p$ identity matrix, and so

$$\begin{aligned} \eta_0 &:= \sqrt{\lambda(1-\lambda)} \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right] \\ &= \sqrt{\lambda(1-\lambda)} \int_0^1 \tilde{\phi}_0(r; \lambda) dW_p(r) \sim N(0, I_p). \end{aligned}$$

As a result,

$$\lambda(1-\lambda) F_T \rightarrow^d \eta'_0 \left(\frac{1}{K} \sum_{j=1}^K \eta_j \eta'_j \right)^{-1} \eta_0 \text{ and } \sqrt{\lambda(1-\lambda)} t_T \rightarrow^d \eta'_0 \left(\frac{1}{K} \sum_{j=1}^K \eta_j \eta'_j \right)^{-1/2},$$

where

$$\eta_j = \int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \text{ for } j = 1, \dots, K.$$

When K is relatively large, it is reasonable to approximate $K^{-1} \sum_{j=1}^K \eta_j \eta'_j$ by its mean:

$$E \left[\frac{1}{K} \sum_{j=1}^K \eta_j \eta'_j \right] = I_p \cdot \frac{1}{K} \sum_{j=1}^K \int_0^1 [\tilde{\phi}_j(r; \lambda)]^2 dr.$$

With such an approximation, we have

$$\begin{aligned} & \lambda(1-\lambda) \cdot \frac{1}{K} \sum_{j=1}^K \int_0^1 [\tilde{\phi}_j(r; \lambda)]^2 dr \cdot F_\infty \sim^a \chi_p^2, \\ & \sqrt{\lambda(1-\lambda)} \left[\frac{1}{K} \sum_{j=1}^K \int_0^1 [\tilde{\phi}_j(r; \lambda)]^2 dr \right]^{1/2} \cdot t_\infty \sim^a N(0, 1), \end{aligned}$$

where ‘ \sim^a ’ signifies distributional approximations. As a result, we can employ the following approximations:

$$F_T^* := \lambda(1-\lambda) \cdot \left\{ \frac{1}{KT} \sum_{j=1}^K \sum_{i=1}^T \left[\tilde{\phi}_{j,T} \left(\frac{i}{T}; \lambda \right) \right]^2 \right\} \cdot F_T \sim^a \chi_p^2, \quad (4)$$

$$t_T^* := \sqrt{\lambda(1-\lambda)} \cdot \left\{ \frac{1}{KT} \sum_{j=1}^K \sum_{i=1}^T \left[\tilde{\phi}_{j,T} \left(\frac{i}{T}; \lambda \right) \right]^2 \right\}^{1/2} \cdot t_T \sim^a N(0, 1), \quad (5)$$

where $\tilde{\phi}_{j,T}(r; \lambda)$ is the finite sample version of $\tilde{\phi}_j(r; \lambda)$ given by

$$\begin{aligned} \tilde{\phi}_{j,T}(r; \lambda) = & \frac{1}{\lambda} \left[\phi_j(r) - \frac{1}{[\lambda T]} \sum_{t=1}^{[\lambda T]} \phi_j \left(\frac{t}{T} \right) \right] \cdot 1\{r \leq \lambda\} \\ & - \frac{1}{1-\lambda} \left[\phi_j(r) - \frac{1}{T - [\lambda T]} \sum_{t=[\lambda T]+1}^T \phi_j \left(\frac{t}{T} \right) \right] \cdot 1\{r > \lambda\}. \end{aligned} \quad (6)$$

It is important to point out that the chi-square and normal approximations are not based on the original Wald and t statistics but rather on their modified versions F_T^* and t_T^* . To a great extent, the chi-square and normal approximations we propose here improve upon the conventional chi-square and normal approximations that are applied directly to the original Wald and t statistics.

Note that the chi-square distribution and standard normal distribution in (4) and (5) are not the asymptotic distributions of F_T^* and t_T^* for a fixed K . The fixed- K asymptotic distributions are given by

$$F_T^* \rightarrow^d \lambda(1-\lambda) \cdot \frac{1}{K} \sum_{j=1}^K \int_0^1 \left[\tilde{\phi}_j(r; \lambda) \right]^2 dr \cdot F_\infty := F_\infty^*, \quad (7)$$

$$t_T^* \rightarrow^d \sqrt{\lambda(1-\lambda)} \cdot \left\{ \frac{1}{K} \sum_{j=1}^K \int_0^1 \left[\tilde{\phi}_j(r; \lambda) \right]^2 dr \right\}^{1/2} \cdot t_\infty := t_\infty^*. \quad (8)$$

These follow directly from Theorem 3.1. The chi-square distribution and standard normal distribution are only approximations to the above nonstandard fixed- K asymptotic distributions.

4.2 Asymptotic F and t Theory

To obtain convenient fixed- K asymptotic approximations, we note that for each $j = 0, 1, \dots, K$, η_j is normal. For each $j \neq 0$, we have

$$\begin{aligned} & \text{cov}(\eta_0, \eta_j) \\ &= \int_0^1 \tilde{\phi}_0(r; \lambda) \tilde{\phi}_j(r; \lambda) dr = \frac{1}{\lambda^2} \int_0^1 [\phi_j(r) - \bar{\phi}_{j,1}] 1\{r \leq \lambda\} dr \\ &+ \frac{1}{(1-\lambda)^2} \int_0^1 [\phi_j(r) - \bar{\phi}_{j,2}] 1\{r > \lambda\} dr = 0. \end{aligned}$$

So η_0 is independent of η_j , $j = 1, \dots, K$. In addition,

$$\text{cov}(\eta_{j_1}, \eta_{j_2}) = \int_0^1 \tilde{\phi}_{j_1}(r; \lambda) \tilde{\phi}_{j_2}(r; \lambda) dr.$$

Therefore, if $\{\tilde{\phi}_j(r; \lambda)\}$ are orthonormal, then η_j for $j = 0, 1, \dots, K$ are independent standard normals. In this case, $\lambda(1 - \lambda)F_\infty$ is a quadratic form in a standard normal vector with an independent weighting matrix. After some adjustment, we can show that $\lambda(1 - \lambda)F_\infty$ is equal to a standard F distribution and that F_T converges to the F distribution. Similarly, $\sqrt{\lambda(1 - \lambda)} \cdot t_T$ converges to Student's t distribution.

Proposition 4.1 *Let Assumptions 3.1–3.3 hold. If $\{\tilde{\phi}_j(r; \lambda)\}$ are orthonormal, then*

$$\tilde{F}_T^* := \frac{K - p + 1}{Kp} \cdot \lambda(1 - \lambda) \cdot F_T \rightarrow^d F_{p, K-p+1},$$

and for $p = 1$,

$$\tilde{t}_T^* := \sqrt{\lambda(1 - \lambda)} \cdot t_T \rightarrow^d t_K$$

where $F_{p, K-p+1}$ is the standard F distribution with the degrees of freedom $(p, K - p + 1)$ and t_K is Student's t distribution with degrees of freedom K .

This is a very convenient result, as the fixed-smoothing asymptotic approximations are standard distributions, and there is no need to simulate critical values.

When $\{\tilde{\phi}_j(r; \lambda)\}$ are orthonormal, we have $K^{-1} \sum_{j=1}^K \int_0^1 [\tilde{\phi}_j(r; \lambda)]^2 dr = 1$. In view of this, we can see that the definitions of \tilde{F}_T^* and \tilde{t}_T^* are similar to those of F_T^* and t_T^* given in (7) and (8). The only difference is that there is an additional degrees-of-freedom adjustment factor in \tilde{F}_T^* when $p > 1$.

4.3 Designing the bases

To design the basis functions such that $\{\tilde{\phi}_j(r; \lambda)\}$ are orthonormal, we need the following lemma.

Lemma 4.1 *Let $\delta(\cdot)$ be the Dirac delta function such that*

$$\int_0^1 \int_0^1 \phi_{j_1}(r) \delta(r - s) \phi_{j_2}(s) dr ds = \int_0^1 \phi_{j_1}(r) \phi_{j_2}(r) dr.$$

Then

$$\int_0^1 \tilde{\phi}_{j_1}(r; \lambda) \tilde{\phi}_{j_2}(r; \lambda) dr = \int_0^1 \int_0^1 C(r, s; \lambda) \phi_{j_1}(r) \phi_{j_2}(s) dr ds,$$

where

$$\begin{aligned} C(r, s; \lambda) = & \left[\delta(r - s) - \frac{1}{\lambda} \right] \frac{1 \{ (r, s) \in [0, \lambda] \times [0, \lambda] \}}{\lambda^2} \\ & + \left[\delta(r - s) - \frac{1}{1 - \lambda} \right] \frac{1 \{ (r, s) \in [\lambda, 1] \times [\lambda, 1] \}}{(1 - \lambda)^2}. \end{aligned}$$

Let

$$W_p(r; \lambda) = \frac{1}{\lambda} \left[W_p(r) - \frac{r}{\lambda} W_p(\lambda) \right] \cdot 1\{0 \leq r \leq \lambda\} \\ - \frac{1}{1-\lambda} \left\{ W_p(r) - W_p(\lambda) - \frac{r-\lambda}{1-\lambda} [W_p(1) - W_p(\lambda)] \right\} \cdot 1\{\lambda < r \leq 1\}$$

be the transformed Brownian motion. Then we have

$$\int_0^1 \tilde{\phi}_{j_1}(r; \lambda) dW_p(r) = \int_0^1 \phi_j(r) dW_p(r; \lambda),$$

and

$$E[dW_p(r; \lambda) dW_p'(s; \lambda)] = I_p \cdot C(r, s; \lambda) dr ds.$$

Therefore, $C(r, s; \lambda)$ can be regarded as the covariance kernel function for the transformed Brownian motion.

To design the basis functions $\{\phi_j(r)\}$ such that $\{\tilde{\phi}_j(r; \lambda)\}$ are orthonormal on $L^2[0, 1]$, we require that $\{\phi_j(r)\}$ be orthonormal with respect to the covariance kernel function $C(r, s; \lambda)$, that is,

$$\int_0^1 \int_0^1 C(r, s; \lambda) \phi_{j_1}(r) \phi_{j_2}(s) dr ds = 1\{j_1 = j_2\}. \quad (9)$$

This can be achieved by applying the Gram-Schmidt orthonormalization to any set of basis functions on $L^2[0, 1]$. The chart below illustrates the procedure:

$\{\phi_j\}$ \downarrow_{GS} $\{\phi_j^*\}$	$\xrightarrow{\text{estimation error}}$ $\xrightarrow{\text{estimation error}}$	$\{\tilde{\phi}_j\}$ (may not be orthonormal on $L^2[0, 1]$) $\{\tilde{\phi}_j^*\}$ (orthonormal on $L^2[0, 1]$)
--	--	---

In the above, $\{\phi_j\}$ is the initial set of basis functions, and $\{\phi_j^*\}$ is the Gram-Schmidt orthonormalized set. “ $\phi_j \rightarrow \tilde{\phi}_j$ ” and “ $\phi_j^* \rightarrow \tilde{\phi}_j^*$ ” reflect the effect of the estimation error in estimating β : had we known β , we would have used the true u_t instead of \hat{u}_t in constructing the variance estimator, and the key elements of the weighting matrix in (2) in Theorem 3.1 would have been $\int_0^1 \phi_j(r) dW_p(r)$ instead of $\int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r)$. The Gram-Schmidt orthonormalization ensures that $\{\phi_j^*\}$ are orthonormal with respect to the covariance kernel $C(r, s; \lambda)$: $\int_0^1 \int_0^1 \phi_{j_1}^*(r) \phi_{j_2}^*(s) C(r, s; \lambda) dr ds = 1\{j_1 = j_2\}$. In view of

$$\int_0^1 \int_0^1 \phi_{j_1}^*(r) \phi_{j_2}^*(s) C(r, s; \lambda) dr ds = \int_0^1 \tilde{\phi}_{j_1}^*(r) \tilde{\phi}_{j_2}^*(r) dr,$$

we have: $\{\tilde{\phi}_j^*\}$ are orthonormal on $L^2[0, 1]$.

If we use $\{\phi_j^*\}$ in constructing the variance estimator, then

$$\lambda(1-\lambda) F_T \rightarrow^d \eta_0' \left(\frac{1}{K} \sum_{j=1}^K \eta_j \eta_j' \right)^{-1} \eta_0$$

for $\eta_j = \int_0^1 \tilde{\phi}_j^*(r; \lambda) dW_p(r) \sim iidN(0, I_p)$ because $\{\tilde{\phi}_j^*\}$ are orthonormal on $L^2[0, 1]$. Moreover, for $j = 1, \dots, K$, η_j is independent of η_0 . Therefore, the asymptotic F theory in Proposition 4.1 holds. Similarly, the asymptotic t theory holds.

Instead of searching for the basis functions that satisfy (9), we search for their discrete versions: the basis vectors. For each basis function $\phi_k(r)$, the corresponding basis vector is defined as

$$\phi_k = \left(\phi_k\left(\frac{1}{T}\right), \phi_k\left(\frac{2}{T}\right), \dots, \phi_k\left(\frac{T}{T}\right) \right)'.$$

Let $\mathbf{C}_T := \mathbf{C}_T(\lambda)$ be the $T \times T$ matrix whose (t_1, t_2) -th element is equal to

$$\begin{aligned} C_T(t_1, t_2; \lambda) &= \left[T \cdot \mathbf{1}\{t_1 = t_2\} - \frac{1}{\lambda} \right] \frac{\mathbf{1}\left\{\left(\frac{t_1}{T}, \frac{t_2}{T}\right) \in [0, \lambda] \times [0, \lambda]\right\}}{\lambda^2} \\ &+ \left[T \cdot \mathbf{1}\{t_1 = t_2\} - \frac{1}{1-\lambda} \right] \frac{\mathbf{1}\left\{\left(\frac{t_1}{T}, \frac{t_2}{T}\right) \in (\lambda, 1] \times (\lambda, 1]\right\}}{(1-\lambda)^2}. \end{aligned}$$

By definition, \mathbf{C}_T is symmetric and positive-definite. It is the discrete version of $C(r, s; \lambda)$. For any two vectors $r_1, r_2 \in \mathbb{R}^T$, we define the inner product

$$\langle r_1, r_2 \rangle = r_1' \mathbf{C}_T r_2 / T^2. \quad (10)$$

Then the discrete analogue of (9) is

$$\langle \phi_{j_1}, \phi_{j_2} \rangle = \mathbf{1}\{j_1 = j_2\} \text{ for } j_1, j_2 = 1, \dots, K. \quad (11)$$

Given any basis vectors ϕ_1, \dots, ϕ_K , we now apply the Gram–Schmidt orthonormalization via the Cholesky decomposition. Let $\phi = (\phi_1, \dots, \phi_K)$ be the $T \times K$ matrix of basis vectors. Let $\mathcal{U}_T \in \mathbb{R}^{K \times K}$ be the upper triangular factor in the Cholesky decomposition of $\phi' \mathbf{C}_T \phi / T^2$ such that $\phi' \mathbf{C}_T \phi / T^2 = \mathcal{U}_T' \mathcal{U}_T$. Define

$$\phi^* = \phi \mathcal{U}_T^{-1} := (\phi_1^*, \dots, \phi_K^*).$$

We then have

$$(\phi^*)' \mathbf{C}_T \phi^* / T^2 = (\mathcal{U}_T')^{-1} \phi' \mathbf{C}_T \phi \mathcal{U}_T^{-1} / T^2 = (\mathcal{U}_T')^{-1} \mathcal{U}_T' \mathcal{U}_T \mathcal{U}_T^{-1} = I_K.$$

That is, the columns of the matrix ϕ^* satisfy the conditions in (11).

Note that the (k_1, k_2) -th element of $\phi' \mathbf{C}_T \phi / T^2$ satisfies

$$\begin{aligned} &\frac{1}{T^2} \sum_{t_2=1}^T \sum_{t_1=1}^T \phi_{k_1}\left(\frac{t_1}{T}\right) C_T(t_1, t_2; \lambda) \phi_{k_2}\left(\frac{t_2}{T}\right) \\ &\rightarrow \int_0^1 \int_0^1 C(r, s; \lambda) \phi_{k_1}(r) \phi_{k_2}(s) dr ds = \int_0^1 \tilde{\phi}_{k_1}(r; \lambda) \tilde{\phi}_{k_2}(r; \lambda) dr \\ &= \text{cov}(\eta_{k_1}, \eta_{k_2}) \text{ as } T \rightarrow \infty. \end{aligned}$$

This implies that \mathcal{U}_T converges to the upper triangular factor of the Cholesky decomposition of $\text{var}(\eta_1, \dots, \eta_K)$. As a result, every transformed basis vector is approximately equal to a linear combination of the original basis vectors. The implied basis functions are thus equal to linear combinations of the original basis functions. Therefore, if Assumption 3.3 holds for the original basis functions, it also holds for the transformed basis functions. It then follows that Proposition

4.1 holds when $\{\phi_1^*, \dots, \phi_K^*\}$ are used as the basis vectors in constructing the asymptotic variance estimator. More specifically, if we estimate Ω by

$$\hat{\Omega}_{\tilde{X}} = \frac{1}{K} \sum_{j=1}^K \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_{j,t}^* \tilde{X}_t \hat{u}_t \right]^{\otimes 2},$$

where $\phi_{j,t}^*$ is the t -th element of the vector ϕ_j^* , then the asymptotic F and t results in Proposition 4.1 hold.

5 The Chow Test in the presence of time-invariant effects

Suppose there is another covariate vector $Z_t \in \mathbb{R}^\ell$ whose effect on Y_t does not change over time so that we have the model:

$$Y_t = 1 \{t \leq [\lambda T]\} \cdot X_t' \beta_1 + 1 \{t \geq [\lambda T] + 1\} \cdot X_t' \beta_2 + Z_t' \gamma + u_t.$$

Let $\beta = (\beta_1', \beta_2')'$ and $\theta = (\beta', \gamma')'$. The OLS estimator of θ is

$$\hat{\theta} = \begin{pmatrix} \tilde{X}' \tilde{X} & \tilde{X}' Z \\ Z' \tilde{X} & Z' Z \end{pmatrix}^{-1} \begin{pmatrix} \tilde{X}' Y \\ Z' Y \end{pmatrix}$$

where $Z = (Z_1, \dots, Z_T)'$.

Let $\hat{U} = (\hat{u}_1, \dots, \hat{u}_T)' = Y - \tilde{X} \hat{\beta} - Z \hat{\gamma} = u - \tilde{X}(\hat{\beta} - \beta) - Z(\hat{\gamma} - \gamma)$. Define

$$\hat{Q}_{[\tilde{X}, Z]} = \frac{1}{T} \begin{pmatrix} \tilde{X}' \tilde{X} & \tilde{X}' Z \\ Z' \tilde{X} & Z' Z \end{pmatrix} \text{ and } \hat{\Omega}_{[\tilde{X}, Z]} = \frac{1}{K} \sum_{j=1}^K \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) [\tilde{X}_t', Z_t']' \hat{u}_t \right]^{\otimes 2}.$$

The Wald statistic for testing $H_0 : R\theta = r$ against $H_1 : R\theta \neq r$ takes the same form

$$F_T = T \cdot (R\hat{\theta} - r)' \left\{ R \hat{Q}_{[\tilde{X}, Z]}^{-1} \hat{\Omega}_{[\tilde{X}, Z]} \hat{Q}_{[\tilde{X}, Z]}^{-1} R' \right\}^{-1} (R\hat{\theta} - r).$$

When $p = 1$, we construct the t statistic:

$$t_T = \frac{\sqrt{T} \cdot (R\hat{\theta} - r)}{\left[R \hat{Q}_{[\tilde{X}, Z]}^{-1} \hat{\Omega}_{[\tilde{X}, Z]} \hat{Q}_{[\tilde{X}, Z]}^{-1} R' \right]^{1/2}}.$$

Note that the dimension of R is different from that in the previous sections.

To establish the asymptotic distributions of F_T and t_T , we maintain the two assumptions below, which are analogous to Assumptions 3.1 and 3.2.

Assumption 5.1 $T^{-1} \sum_{t=1}^{[Tr]} (X_t', Z_t')' (X_t', Z_t') \rightarrow^p Q_{[X, Z]} \cdot r$ uniformly over $r \in [0, 1]$ where $Q_{[X, Z]}$ is an $(\ell + m) \times (\ell + m)$ invertible matrix.⁴

⁴In place of $Q_{[\tilde{X}, Z]}$, we could use $Q_{[\tilde{X}, Z][\tilde{X}, Z]}$ but the latter notation is too cumbersome. We opt for a simpler notation when there is no confusion.

Assumption 5.2 $T^{-1/2} \sum_{t=1}^{[Tr]} (X'_t, Z'_t)' u_t \rightarrow^d \Lambda_{[X,Z]} W_{\ell+m}(r)$ for $r \in [0, 1]$ where $\Lambda_{[X,Z]} \Lambda'_{[X,Z]}$ is the long run variance of the process $\{(X'_t, Z'_t)' u_t\}$ and $W_{\ell+m}(\cdot)$ is an $(\ell + m) \times 1$ standard Brownian process.

We partition $Q_{[X,Z]}$ and $\Lambda_{[X,Z]}$ according to

$$Q_{[X,Z]} = \begin{pmatrix} Q_{XX} & Q_{XZ} \\ Q_{ZX} & Q_{ZZ} \end{pmatrix} \text{ and } \Lambda_{[X,Z]} = \begin{pmatrix} \Lambda_X \\ \Lambda_Z \end{pmatrix},$$

where $Q_{XX} \in \mathbb{R}^{m \times m}$, $Q_{ZZ} \in \mathbb{R}^{\ell \times \ell}$, $\Lambda_X \in \mathbb{R}^{m \times (\ell+m)}$, and $\Lambda_Z \in \mathbb{R}^{\ell \times (\ell+m)}$.

To test for a structural break, we let $R = (\mathcal{R}_\beta, -\mathcal{R}_\beta, O_\gamma)$ for some $p \times m$ full rank matrix \mathcal{R}_β and $p \times \ell$ zero matrix O_γ . The following theorem is an extension of the Chow test in the presence of time-invariant effects.

Theorem 5.1 *Let Assumptions 3.3, 5.1, and 5.2 hold. For $R = (\mathcal{R}_\beta, -\mathcal{R}_\beta, O_\gamma)$, we have*

$$F_T \rightarrow^d \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right]' \times \left[\frac{1}{K} \sum_{j=1}^K \left\{ \int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \right\}^{\otimes 2} \right]^{-1} \\ \times \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right].$$

When $p = 1$,

$$t_T \rightarrow^d \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right] \times \left[\frac{1}{K} \sum_{j=1}^K \left\{ \int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \right\}^{\otimes 2} \right]^{-1/2}.$$

Theorem 5.1 shows that the limiting distributions of the Wald statistic and t statistic are the same as in the case without the extra covariate Z_t . The asymptotic F and t limit theory can be developed in exactly the same way as in Section 4. We present the result formally as a corollary.

Corollary 1 *Let Assumptions 3.3, 5.1, and 5.2 hold. Suppose that the Gram-Schmidt transformed basis vectors $\phi_1^*, \dots, \phi_K^*$ are used in constructing the variance estimator, that is,*

$$\hat{\Omega}_{[\tilde{X}, Z]} = \frac{1}{K} \sum_{j=1}^K \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_{j,t}^* [\tilde{X}'_t, Z'_t]' \hat{u}_t \right]^{\otimes 2}$$

where $\phi_{j,t}^*$ is the t -th element of the vector ϕ_j^* . Then

$$\tilde{F}_T^* := \frac{K-p+1}{Kp} \cdot \lambda(1-\lambda) \cdot F_T \rightarrow^d F_{p, K-p+1},$$

and for $p = 1$,

$$\tilde{t}_T^* := \sqrt{\lambda(1-\lambda)} \cdot t_T \rightarrow^d t_K.$$

Now suppose we allow the coefficients on the covariate X_t to have a break potentially, but we are interested in testing $\mathcal{R}_\gamma \gamma = r$ for some $p \times \ell$ full rank matrix \mathcal{R}_γ . In this case, we can let $R = (O_\beta, O_\beta, \mathcal{R}_\gamma)$ where O_β is the $p \times m$ matrix of zeros. The asymptotic theory is given below.

Theorem 5.2 *Let Assumptions 3.3, 5.1, and 5.2 hold. Assume that $\{\phi_j(r)\}$ are orthonormal on $L^2[0, 1]$ and satisfy $\int_0^1 \phi_j(r) dr = 0$, then for $R = (O_\beta, O_\beta, \mathcal{R}_\gamma)$, we have*

$$\tilde{F}_T^* = \frac{K - p + 1}{Kp} F_T \rightarrow^d F_{p, K-p+1}$$

and for $p = 1$,

$$t_T \rightarrow^d t_K.$$

Theorem 5.2 shows that the asymptotic F and t theory continues to hold if we are interested in the time invariant effects only. In this case, we do not need to use the Gram–Schmidt orthonormalization. All we need is the orthonormality of the basis functions on the usual space $L^2[0, 1]$ and the zero mean condition, namely $\int_0^1 \phi_j(r) dr = 0$. The sine and cosine functions provided latter in (12) satisfy these conditions. The asymptotic F and t theory in Theorem 5.2 is the same as that in a regression with stationary regressors and errors without any structural break (e.g., Sun (2013)). The invariance of the asymptotic results is intriguing, and it exploits the special structure of the Chow regression. It is an open question whether similar results can be established for stationary regressors in a regression with other types of nonstationarity.

Theorems 5.1 and 5.2 consider some particular linear restrictions of the parameters. If we go beyond the Chow test, we may be interested in the effects before and after the break separately. That is, we may be interested in β_1 and β_2 separately. In such a case, the restriction matrix R takes the form $R = (\mathcal{R}_\beta, O_\beta, O_\gamma)$ or $R = (O_\beta, \mathcal{R}_\beta, O_\gamma)$. More generally, we may be interested in a hypothesis that involves parameters from both β and γ so that R has a general form. Proposition 8.1 in the appendix provides the fixed-smoothing asymptotic distributions of the F and t statistics when no restriction is imposed on R . The asymptotic distributions appear to depend on nuisance parameters and cannot be reduced to a standard F or t distribution.

6 Simulation Evidence

In this section, we investigate the finite sample properties of the proposed F test. We consider the linear regression model with $m = 2$ and $X_t = (1, q_t)$. The regressor q_t follows an AR(1) process, and the error u_t follows an independent AR(1) or ARMA(1,1) process. That is,

$$\begin{aligned} q_t &= \rho q_{t-1} + \epsilon_{q,t} \\ u_t &= \rho u_{t-1} + \epsilon_{u,t} + \psi \epsilon_{u,t-1} \end{aligned}$$

where both $\epsilon_{q,t}$ and $\epsilon_{u,t}$ are iid $N(0, 1)$ and $\{\epsilon_{q,t}, t = 1, \dots, T\}$ are independent of $\{\epsilon_{u,t} : t = 1, 2, \dots, T\}$. Note that the AR parameter ρ is the same for q_t and u_t .

We consider the sample sizes $T = 100, 200$, and 500 . We let $\lambda = 0.4$.⁵ Without the loss of generality, we set $\beta_1 = (0, 0)'$ and $\beta_2 = (0, 0)'$ under the null. We consider testing $H_0 : \beta_1 = \beta_2$ against $H_1 : \beta_1 \neq \beta_2$ so that $p = 2$.

We consider two pairs of different tests, both of which are based on the series variance estimators. The first pair uses the (usual) Fourier bases

$$\left\{ \phi_{2j-1}(r) = \sqrt{2} \cos(2j\pi r), \phi_{2j} = \sqrt{2} \sin(2j\pi r), j = 1, \dots, K/2 \right\}. \quad (12)$$

⁵In the supplementary document, we report simulation results when $\lambda = 0.8$ and the errors are (conditionally) heteroskedastic. The results are quite similar to what we report here.

Each test in this pair is based on the same test statistic F_T^* defined in (4) but uses different reference distributions. The first test uses the chi-square approximation (χ_2^2) while the second test uses the nonstandard fixed-smoothing approximation given in (7). We refer to the two tests as “ χ^2 : Fourier Bases” and “ F_∞^* : Fourier Bases”, respectively. The nonstandard critical values are simulated. We approximate the standard Brownian motion in the nonstandard distribution using scaled partial sums of 1000 iid $N(0, 1)$ random variables. To compute the nonstandard critical values, we use 10,000 simulation replications.

The second pair of tests uses the transformed Fourier bases via the Gram–Schmidt orthogonalization given in Section 4.3. Each of the two tests in this pair is based on the same test statistic \tilde{F}_T^* defined in Proposition 4.1. The first test uses the F approximation, and the second test uses the rescaled chi-square distribution $Kp[K - p + 1]^{-1} \chi_2^2$. Equivalently, the second test in this pair employs the test statistic $\tilde{F}_T = \lambda(1 - \lambda) \cdot F_T$ and the chi-square approximation (χ_2^2). We refer to the two tests as “ F : Transformed Bases” and “ χ^2 : Transformed Bases” respectively. The chi-square test in the second pair is used to illustrate the effectiveness of the F approximation in reducing the size distortion.

The nominal level of all tests is 5%. The number of simulation replications is 10,000. Figures 1 and 2 report the null rejection probability for each test for the sample sizes $T = 100$ and $T = 500$ when q_t and u_t follow independent AR(1) processes with the same AR parameter ρ . Several patterns emerge from these two figures:

- Regardless of the bases used, the chi-square tests over-reject the null by a large margin, especially when K is small.
- Regardless of the bases used, the nonstandard test and F test are much more accurate than the chi-square tests.
- For each given value of K , the null rejection probabilities of the nonstandard test and F test are close to each other. This shows that, in terms of size accuracy, using the F approximation (when the transformed Fourier bases are employed) is as good as using the nonstandard approximation (when the Fourier bases are employed). However, the F approximation is more convenient to use and, hence, is preferred.
- For each given value of K , the null rejection probabilities of the two chi-square tests are close to each other, although the one based on the transformed Fourier bases is somewhat more accurate. This shows that the bases do not have a large effect on the quality of the chi-square approximation.
- The nonstandard test and standard F test can still have quite some size distortion if K is large and the regressor and error processes are persistent. The size distortion comes from the bias of the variance estimator. When K is large, we take an average over a frequency window that is too large when the processes are highly persistent, that is, when the spectral density of $\{X_t u_t\}$ is not very flat at the origin. So, it is important to use a data-driven K to obtain an accurate test in practice.
- Comparing the two figures, we see that the size distortion of every test becomes smaller when the sample size is larger.

Figure 3 reports the null rejection probabilities when the sample size T is 200 and when the error process may have an MA component and the AR parameter may be negative. As in Figures 1 and 2, the same patterns emerge.

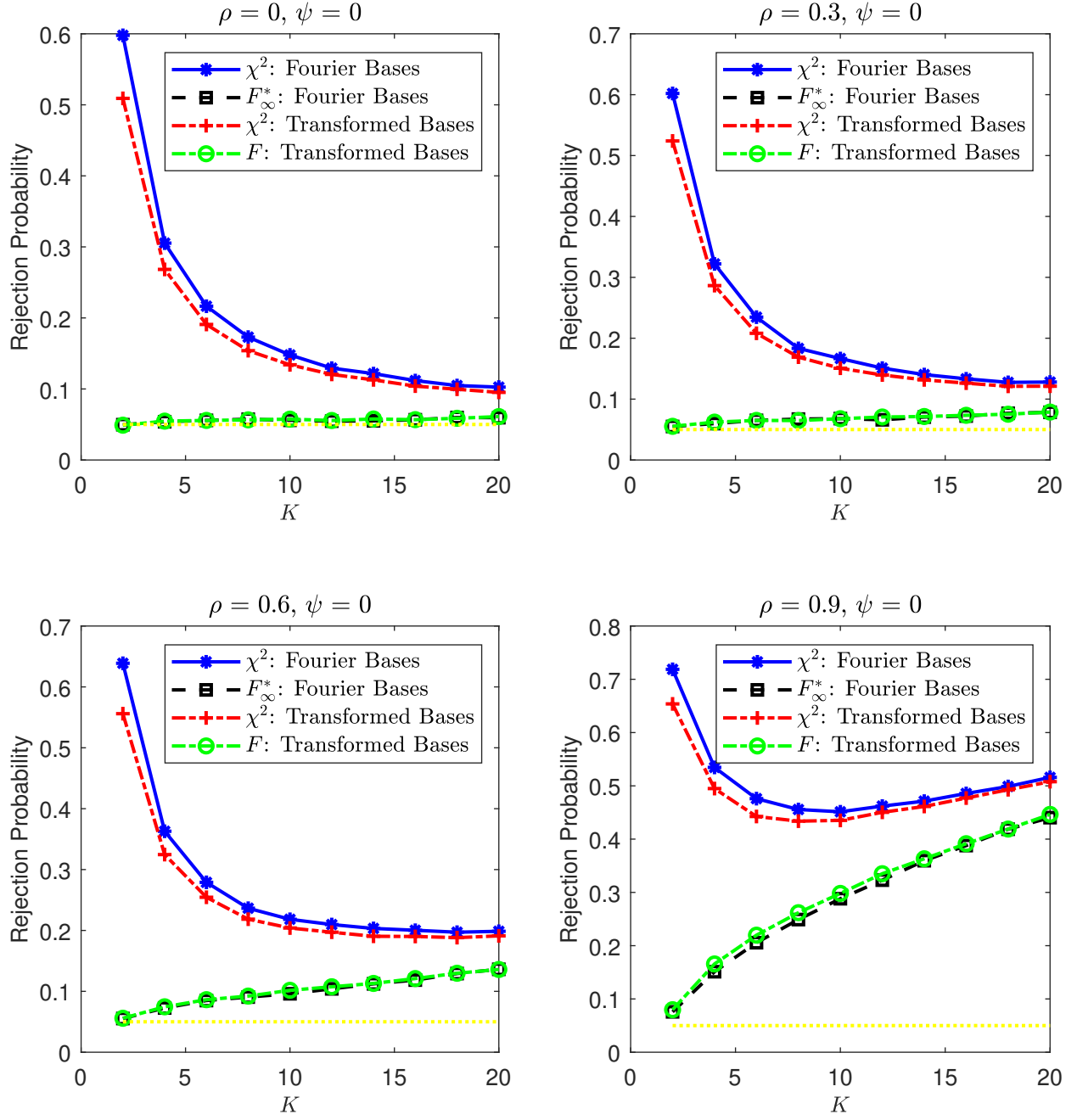


Figure 1: The empirical null rejection probabilities of different 5% tests when $T = 100$ for a range of different K values from 2 to 20 with increment 2.

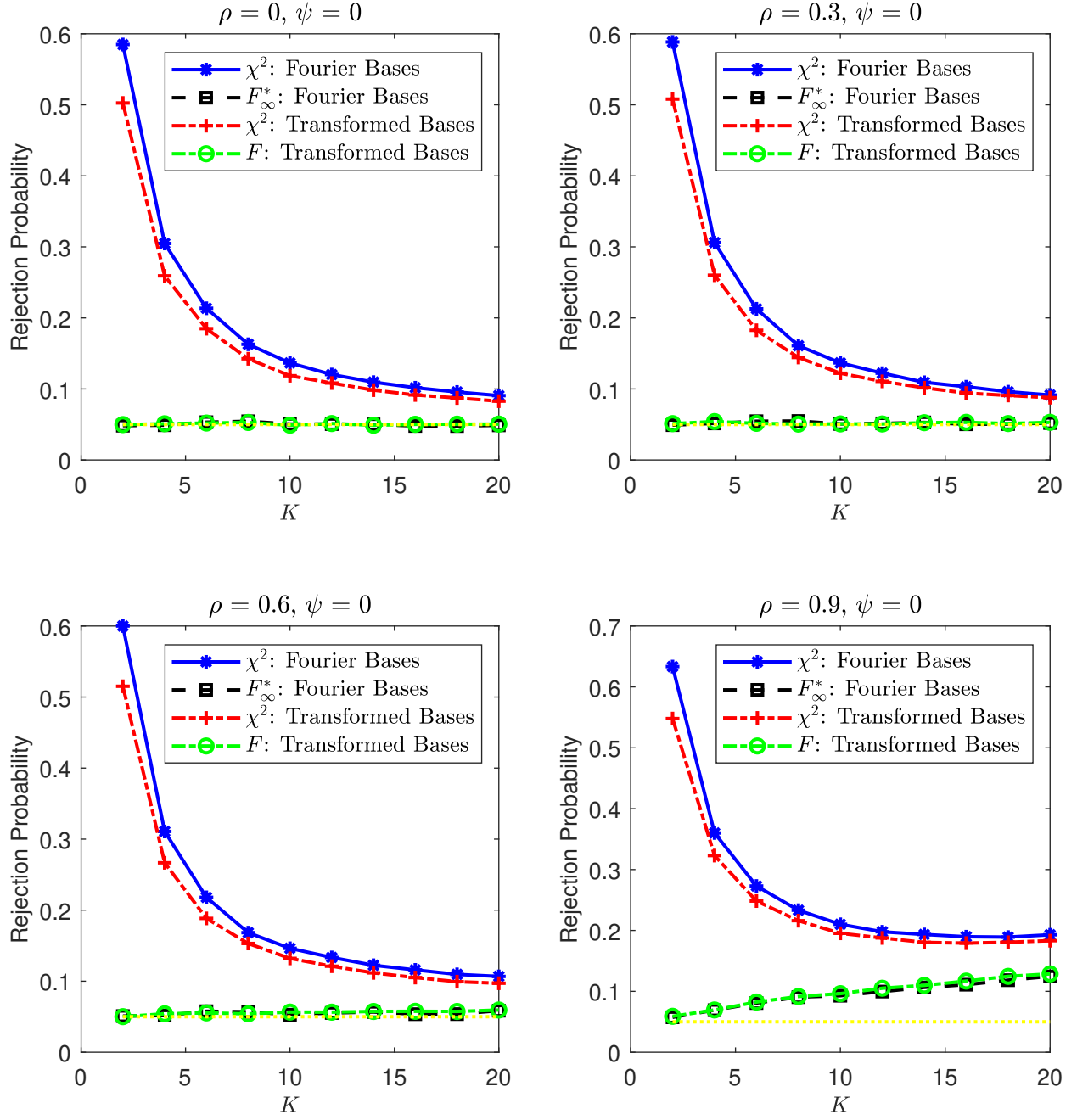


Figure 2: The empirical null rejection probabilities of different 5% tests when $T = 500$ for a range of different K values from 2 to 20 with increment 2.

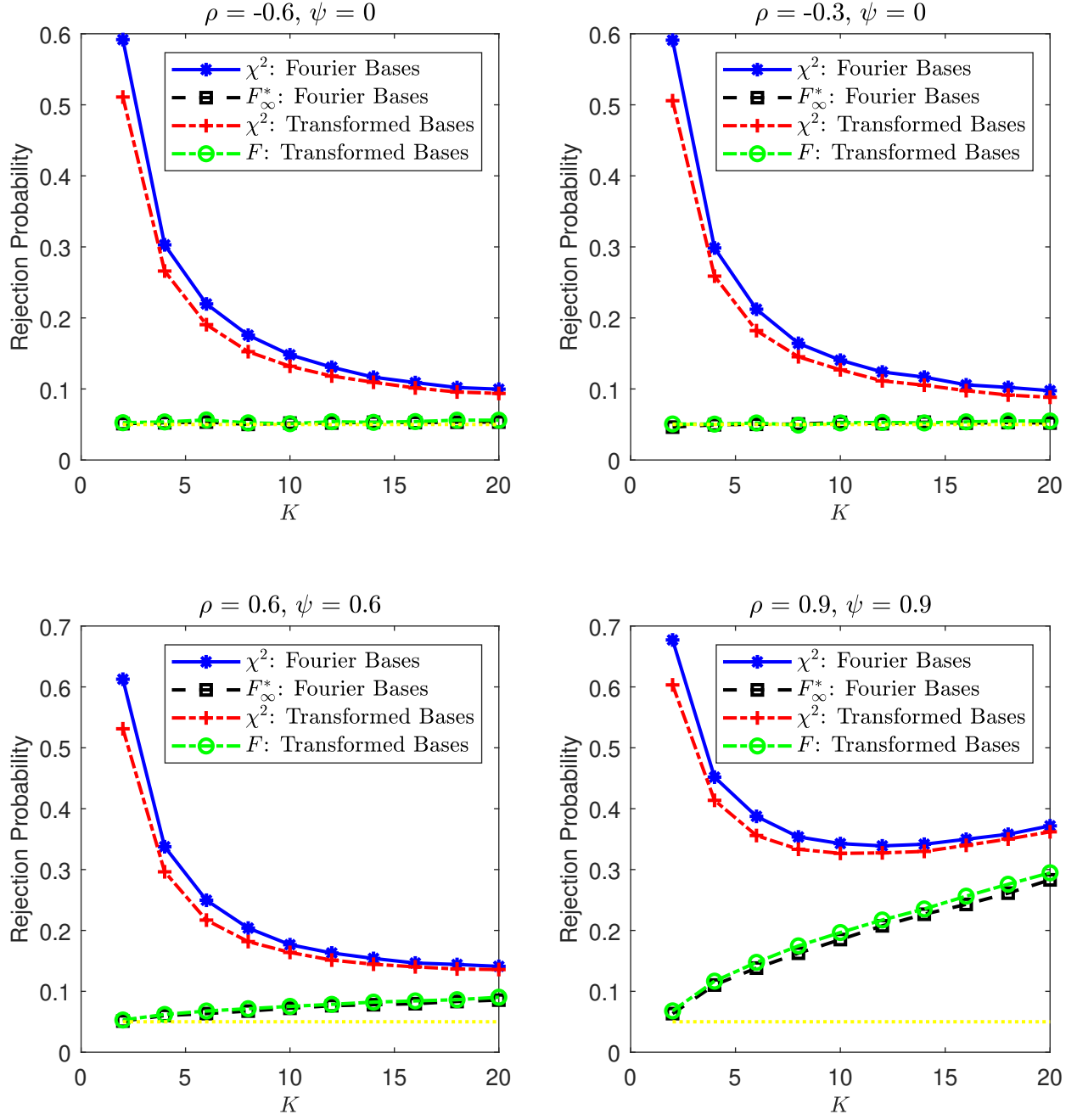


Figure 3: The empirical null rejection probabilities of different 5% tests when $T = 200$ for a range of different K values from 2 to 20 with increment 2.

Next, we consider the size properties of the tests with a data-driven K . Note that

$$R\sqrt{T}(\hat{\beta} - \beta) = \frac{1}{\sqrt{T}} \sum_{t=1}^T R \left(\frac{\tilde{X}'\tilde{X}}{T} \right)^{-1} \tilde{X}_t u_t = \frac{1}{\sqrt{T}} \sum_{t=1}^T v_t + o_p(1)$$

where $v_t = R\hat{Q}_{\tilde{X}}^{-1}\tilde{X}_t u_t$. Then

$$R\hat{Q}_{\tilde{X}}^{-1}\hat{\Omega}_{\tilde{X}}\hat{Q}_{\tilde{X}}^{-1}R' = \frac{1}{K} \sum_{j=1}^K \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \hat{v}_t \right]^{\otimes 2},$$

where $\hat{v}_t = R\hat{Q}_{\tilde{X}}^{-1}\tilde{X}_t u_t$. So $R\hat{Q}_{\tilde{X}}^{-1}\hat{\Omega}_{\tilde{X}}\hat{Q}_{\tilde{X}}^{-1}R'$ can be viewed as the series variance estimator of the long run variance of $\{v_t\}$. We can follow Sun, Phillips and Jin (2008) and Sun (2014) to choose a testing-optimal K that minimizes a weighted average of the type I and type II errors of the proposed test. However, this approach requires high-order expansions. Given that the high-order expansions are highly technical, we opt for a traditional approach here. Following Phillips (2005), we choose K to minimize the mean square error (MSE) of $R\hat{Q}_{\tilde{X}}^{-1}\hat{\Omega}_{\tilde{X}}\hat{Q}_{\tilde{X}}^{-1}R'$. We fit a VAR(1) model to \hat{v}_t and use the fitted model to compute the data-driven MSE-optimal K . Although the MSE-based approach is not best suited for hypothesis testing, Monte Carlo simulations reveal that this choice of K delivers good finite sample performances for the F test.

Table 1 reports the null rejection probabilities and the average values of K used with data-driven choice of K for different sample sizes. The qualitative observations from Figures 1–3 continue to hold with the data-driven K . In particular, the nonstandard test and standard F test are more accurate than the corresponding chi-square tests, especially when the latter have large positive size distortion. The null rejection probabilities of the nonstandard test and the standard F test are close to each other. Similarly, the null rejection probabilities of the two chi-square tests are close to each other. As expected, the average value of K decreases with the persistence of the underlying processes. The higher the persistence, the smaller the average K value, and the more effective the nonstandard test and standard F test in reducing the size distortion.

To simulate the power of the tests, we let $\beta_1 = (0, 0)$ and $\beta_2 = (\delta, \delta)$. Figure 4 presents the size-adjusted power curves as functions of δ when the sample size is 200 and when both q_t and u_t follow AR(1) processes. The figure is representative of other cases. For the two tests in each pair, the size-adjusted powers are the same, as they are based on the same test statistic. Thus, we need only report two power curves: one for the usual Fourier bases and the other for the transformed Fourier bases. The basic message from Figure 4 is that the size-adjusted powers associated with the two sets of bases are very close to each other. This, coupled with its size accuracy and convenience to use, suggests that we use the F test in empirical applications.

7 Conclusion

This study proposes asymptotic F and t tests for structural breaks that are robust to heteroscedasticity and autocorrelation. The tests are based on a special series HAR variance estimator where the basis functions are crafted via the Gram–Schmidt orthonormalization. Monte Carlo simulations show that the F test is much more accurate than the corresponding chi-square test.

This study assumes that there is a single known break point. The asymptotic F and t theory can be extended to the case with multiple but known break points. The theory can also be extended to allow for a linear trend or other deterministic trends, but we need to redesign the

Table 1: The empirical null rejection probabilities of different 5% tests with data-driven choice of K .

	$\rho = 0$ $\psi = 0$	$\rho = 0.3$ $\psi = 0$	$\rho = 0.6$ $\psi = 0$	$\rho = 0.9$ $\psi = 0$	$\rho = -0.6$ $\psi = 0$	$\rho = -0.3$ $\psi = 0$	$\rho = 0.6$ $\psi = 0.6$	$\rho = 0.9$ $\psi = .0.9$
$T = 100$								
χ^2 : Fourier	0.092	0.131	0.227	0.511	0.128	0.091	0.287	0.558
F_{∞}^* : Fourier	0.060	0.076	0.101	0.198	0.067	0.056	0.087	0.170
χ^2 : Transformed	0.089	0.124	0.210	0.473	0.119	0.085	0.259	0.516
F : Transformed	0.064	0.079	0.101	0.209	0.071	0.060	0.088	0.182
Ave(K)	30.00	18.40	9.71	5.29	16.57	26.34	6.14	4.27
$T = 200$								
χ^2 : Fourier	0.069	0.100	0.153	0.396	0.092	0.070	0.197	0.444
F_{∞}^* : Fourier	0.052	0.066	0.079	0.150	0.055	0.050	0.075	0.131
χ^2 : Transformed	0.068	0.094	0.142	0.363	0.088	0.067	0.179	0.406
F : Transformed	0.057	0.069	0.082	0.153	0.058	0.051	0.074	0.135
Ave(K)	70	28.82	14.32	6.10	24.96	46.02	8.56	4.56
$T = 500$								
χ^2 : Fourier	0.055	0.068	0.096	0.222	0.067	0.057	0.119	0.278
F_{∞}^* : Fourier	0.049	0.054	0.062	0.091	0.053	0.048	0.056	0.084
χ^2 : Transformed	0.053	0.064	0.091	0.209	0.064	0.055	0.110	0.253
F : Transformed	0.048	0.053	0.062	0.096	0.051	0.048	0.058	0.086
Ave(K)	144.51	56.47	26.91	9.03	46.58	96.41	15.19	6.05

basis functions. In principle, the tests based on series HAR variance estimation can be extended to accommodate the case with an unknown break point along the line of Cho and Vogelsang (2017). All the basic ingredients have been established in this study. We only need to take the supremum (or other functionals) of the Wald or t statistic over λ as the test statistic. However, the convenient F approximation is lost, as the supremum of the standard distributions is not standard any more. Therefore, it is not clear whether there is still an advantage of using series HAR variance estimators rather than kernel HAR variance estimators.

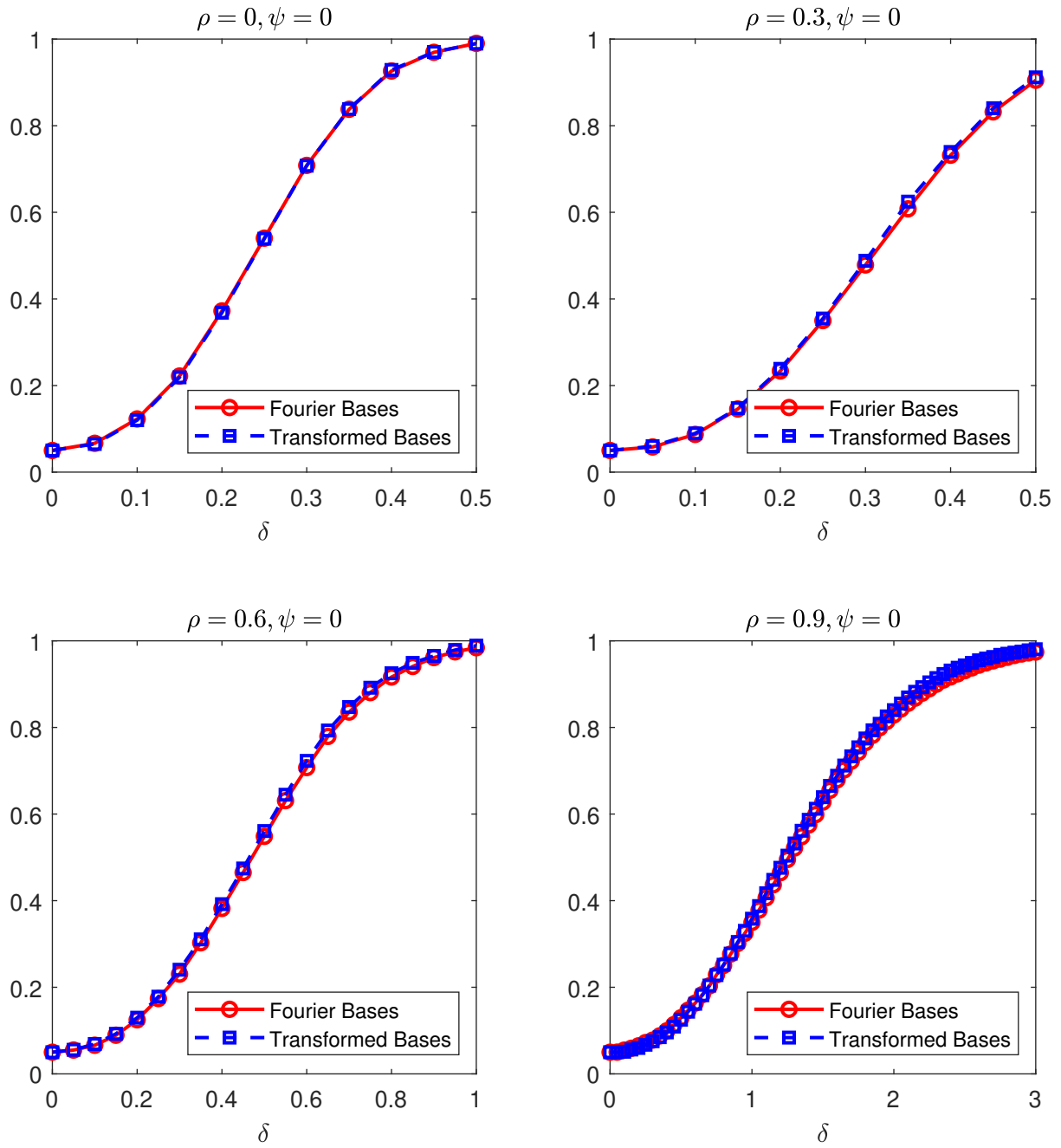


Figure 4: The size-adjusted power curves for different 5% tests when $T = 200$.

8 Appendix of Proofs

Proof of Lemma 3.1. Under Assumption 3.1, we have

$$\frac{\tilde{X}'\tilde{X}}{T} = \begin{pmatrix} T^{-1} \sum_{t=1}^{[T\lambda]} X_t X_t' & O \\ O & T^{-1} \sum_{t=[T\lambda]+1}^T X_t X_t' \end{pmatrix} \rightarrow^p \begin{pmatrix} \lambda Q_{XX} & O \\ O & (1-\lambda) Q_{XX} \end{pmatrix}.$$

Under Assumption 3.2, we have

$$\frac{\tilde{X}'U}{\sqrt{T}} = \begin{pmatrix} \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} X_t u_t \\ \frac{1}{\sqrt{T}} \sum_{t=[T\lambda]+1}^T X_t u_t \end{pmatrix} \rightarrow^d \begin{pmatrix} \Lambda_X W_m(\lambda) \\ \Lambda_X [W_m(1) - W_m(\lambda)] \end{pmatrix}.$$

Hence

$$\begin{aligned} \sqrt{T}(\hat{\beta} - \beta) &\rightarrow^d \begin{pmatrix} \lambda Q_{XX} & O \\ O & (1-\lambda) Q_{XX} \end{pmatrix}^{-1} \begin{pmatrix} \Lambda_X W_m(\lambda) \\ \Lambda_X [W_m(1) - W_m(\lambda)] \end{pmatrix} \\ &= \begin{pmatrix} (\lambda Q_{XX})^{-1} \Lambda_X W_m(\lambda) \\ [(1-\lambda) Q_{XX}]^{-1} \Lambda_X [W_m(1) - W_m(\lambda)] \end{pmatrix} \\ &= \begin{pmatrix} Q_{XX}^{-1} \Lambda_X \cdot \frac{W_m(\lambda)}{\lambda} \\ Q_{XX}^{-1} \Lambda_X \cdot \frac{W_m(1) - W_m(\lambda)}{1-\lambda} \end{pmatrix} = \begin{pmatrix} Q_{XX}^{-1} \Lambda_X \cdot \frac{1}{\lambda} \int_0^\lambda dW_m(\lambda) \\ Q_{XX}^{-1} \Lambda_X \cdot \frac{1}{1-\lambda} \int_\lambda^1 dW_m(\lambda) \end{pmatrix}. \end{aligned}$$

For the second part of the lemma, we have

$$\begin{aligned} &\frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t \hat{u}_t \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t (Y_t - \tilde{X}_t' \hat{\beta}) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t (\tilde{X}_t' \beta + u_t - \tilde{X}_t' \hat{\beta}) \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t u_t - \left[\frac{1}{T} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t \tilde{X}_t' \right] \sqrt{T}(\hat{\beta} - \beta). \end{aligned}$$

Now, it is not hard to show that under Assumption 3.3,

$$\begin{aligned} &\frac{1}{T} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t \tilde{X}_t' \\ &= \begin{pmatrix} \frac{1}{T} \sum_{t=1}^{[T\lambda]} \phi_j \left(\frac{t}{T} \right) X_t X_t' & O \\ O & \frac{1}{T} \sum_{t=[T\lambda]+1}^T \phi_j \left(\frac{t}{T} \right) X_t X_t' \end{pmatrix} \\ &\rightarrow^p \begin{pmatrix} \left[\int_0^\lambda \phi_j(r) dr \right] Q_{XX} & O \\ O & \left[\int_\lambda^1 \phi_j(r) dr \right] Q_{XX} \end{pmatrix}. \end{aligned}$$

Hence,

$$\begin{aligned}
& \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \tilde{X}_t \hat{u}_t \\
& \rightarrow^d \begin{pmatrix} \Lambda_X \int_0^\lambda \phi_j(r) dW_m(r) \\ \Lambda_X \int_\lambda^1 \phi_j(r) dW_m(r) \end{pmatrix} \\
& - \begin{pmatrix} \left[\int_0^\lambda \phi_j(r) dr \right] Q_{XX} & O \\ O & \left[\int_\lambda^1 \phi_j(r) dr \right] Q_{XX} \end{pmatrix} \begin{pmatrix} Q_{XX}^{-1} \Lambda_X \cdot \frac{W_m(\lambda)}{\lambda} \\ Q_{XX}^{-1} \Lambda_X \cdot \frac{W_m(1) - W_m(\lambda)}{(1-\lambda)} \end{pmatrix} \\
& = \begin{pmatrix} \Lambda_X \left\{ \int_0^\lambda \phi_j(r) dW_m(r) - \frac{1}{\lambda} \int_0^\lambda \phi_j(r) dr \cdot W_m(\lambda) \right\} \\ \Lambda_X \left\{ \int_\lambda^1 \phi_j(r) dW_m(r) - \frac{1}{1-\lambda} \int_\lambda^1 \phi_j(r) dr [W_m(1) - W_m(\lambda)] \right\} \end{pmatrix} \\
& = \begin{pmatrix} \Lambda_X \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_m(r) \\ \Lambda_X \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_m(r) \end{pmatrix}.
\end{aligned}$$

■

Proof of Theorem 3.1. We have

$$\begin{aligned}
R\hat{Q}_{\bar{X}}^{-1} & \rightarrow^d (\mathcal{R}, -\mathcal{R}) \begin{pmatrix} \lambda^{-1} Q_{XX}^{-1} & O \\ O & (1-\lambda)^{-1} Q_{XX}^{-1} \end{pmatrix} = \left(\lambda^{-1} \mathcal{R} Q_{XX}^{-1}, -(1-\lambda)^{-1} \mathcal{R} Q_{XX}^{-1} \right), \\
\hat{\Omega}_{\bar{X}} & \rightarrow^d \frac{1}{K} \sum_{j=1}^K \begin{pmatrix} \Lambda_X & O \\ O & \Lambda_X \end{pmatrix} \begin{pmatrix} \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_m(r) \\ \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_m(r) \end{pmatrix}^{\otimes 2} \begin{pmatrix} \Lambda_X & O \\ O & \Lambda_X \end{pmatrix}.
\end{aligned}$$

Hence,

$$\begin{aligned}
& R\hat{Q}_{\bar{X}}^{-1} \hat{\Omega}_{\bar{X}} Q_{\bar{X}}^{-1} R' \\
& \rightarrow^d \left(\lambda^{-1} \mathcal{R} Q_{XX}^{-1}, -(1-\lambda)^{-1} \mathcal{R} Q_{XX}^{-1} \right) \begin{pmatrix} \Lambda_X & O \\ O & \Lambda_X \end{pmatrix} \\
& \times \frac{1}{K} \sum_{j=1}^K \begin{pmatrix} \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_m(r) \\ \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_m(r) \end{pmatrix}^{\otimes 2} \\
& \times \begin{pmatrix} \Lambda_X & O \\ O & \Lambda_X \end{pmatrix} \begin{pmatrix} \lambda^{-1} (\mathcal{R} Q_{XX}^{-1})' \\ (1-\lambda)^{-1} (\mathcal{R} Q_{XX}^{-1})' \end{pmatrix} \\
& = \left(\lambda^{-1} \mathcal{R} Q_{XX}^{-1} \Lambda_X, -(1-\lambda)^{-1} \mathcal{R} Q_{XX}^{-1} \Lambda_X \right) \frac{1}{K} \sum_{j=1}^K \begin{pmatrix} \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_m(r) \\ \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_m(r) \end{pmatrix}^{\otimes 2} \\
& \times \begin{pmatrix} \lambda^{-1} (\mathcal{R} Q_{XX}^{-1} \Lambda_X)', \\ -(1-\lambda)^{-1} (\mathcal{R} Q_{XX}^{-1} \Lambda_X)' \end{pmatrix} \\
& = \mathcal{R} Q_{XX}^{-1} \Lambda_X \frac{1}{K} \sum_{j=1}^K \left\{ \int_0^\lambda \frac{\phi_j(r) - \bar{\phi}_{j,1}}{\lambda} dW_m(r) - \int_\lambda^1 \frac{\phi_j(r) - \bar{\phi}_{j,2}}{1-\lambda} dW_m(r) \right\}^{\otimes 2} (\mathcal{R} Q_{XX}^{-1} \Lambda_X)'.
\end{aligned}$$

Also, under the null, we have

$$\begin{aligned}\sqrt{T} \cdot [R(\hat{\beta} - \beta)] &\rightarrow^d (\mathcal{R}, -\mathcal{R}) \begin{pmatrix} Q_{XX}^{-1} \Lambda_X \cdot \frac{1}{\lambda} \int_0^\lambda dW_m(\lambda) \\ Q_{XX}^{-1} \Lambda_X \cdot \frac{1}{1-\lambda} \int_\lambda^1 dW_m(\lambda) \end{pmatrix} \\ &= \mathcal{R} Q_{XX}^{-1} \Lambda_X \left[\frac{1}{\lambda} \int_0^\lambda dW_m(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_m(\lambda) \right].\end{aligned}$$

Therefore,

$$\begin{aligned}F_T &= T \cdot [R(\hat{\beta} - \beta)]' [R \hat{Q}_{\bar{X}}^{-1} \hat{\Omega}_{\bar{X}} Q_{\bar{X}}^{-1} R']^{-1} [R(\hat{\beta} - \beta)] \\ &\rightarrow^d \left\{ \mathcal{R} Q_{XX}^{-1} \Lambda_X \left[\frac{1}{\lambda} \int_0^\lambda dW_m(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_m(\lambda) \right] \right\}' \\ &\quad \times \left[\mathcal{R} Q_{XX}^{-1} \Lambda_X \frac{1}{K} \sum_{j=1}^K \left\{ \int_0^1 \tilde{\phi}_j(r; \lambda) dW_m(r) \right\}^{\otimes 2} (\mathcal{R} Q_{XX}^{-1} \Lambda_X)' \right]^{-1} \\ &\quad \times \mathcal{R} Q_{XX}^{-1} \Lambda_X \left[\frac{1}{\lambda} \int_0^\lambda dW_m(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_m(\lambda) \right].\end{aligned}$$

Using the distributional equivalence that $\mathcal{R} Q_{XX}^{-1} \Lambda_X W_m =^d A_p W_p$ for a square and invertible matrix A_p , we have

$$\begin{aligned}F_T &\rightarrow^d \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right]' \times \left[\frac{1}{K} \sum_{j=1}^K \left\{ \int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \right\}^{\otimes 2} \right]^{-1} \\ &\quad \times \left[\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right].\end{aligned}$$

The proof for the weak convergence of t_T is similar and is omitted to save space. ■

Proof of Proposition 4.1. We prove the part for the Wald statistic only, as the proof for the t-statistic is similar. Given that $\{\tilde{\phi}_j(r; \lambda)\}$ are orthonormal on $L^2[0, 1]$, we have:

$$\eta_j := \int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \sim iidN(0, I_p).$$

As a consequence,

$$\sum_{j=1}^K \eta_j \eta_j' \sim \mathbb{W}_p(I_p, K),$$

the standard Wishart distribution with degrees of freedom K . So,

$$\frac{K-p+1}{Kp} \lambda(1-\lambda) F_\infty = \frac{K-p+1}{Kp} \cdot \eta_0' \left(\frac{1}{K} \sum_{j=1}^K \eta_j \eta_j' \right)^{-1} \eta_0.$$

Note that $\eta_0, \eta_1, \dots, \eta_K$ are independent standard normal vectors. $\eta'_0 \left(\frac{1}{K} \sum_{j=1}^K \eta_j \eta'_j \right)^{-1} \eta_0$ follows Hotelling's T^2 distribution. Using the relationship between Hotelling's T^2 distribution and the standard F distribution, we have

$$\frac{K-p+1}{Kp} \lambda (1-\lambda) F_\infty \sim F_{p, K-p+1}.$$

It then follows that

$$\tilde{F}_T^* := \frac{K-p+1}{Kp} \lambda (1-\lambda) F_T \rightarrow^d F_{p, K-p+1}.$$

■

Proof of Lemma 4.1. We have

$$\begin{aligned} & \int_0^1 \tilde{\phi}_{j_1}(r; \lambda) \tilde{\phi}_{j_2}(r; \lambda) dr \\ &= \frac{1}{\lambda^2} \int_0^\lambda \left[\phi_{j_1}(r) - \frac{1}{\lambda} \int_0^\lambda \phi_{j_1}(s) ds \right] \left[\phi_{j_2}(r) - \frac{1}{\lambda} \int_0^\lambda \phi_{j_2}(s) ds \right] dr \\ &+ \frac{1}{(1-\lambda)^2} \int_\lambda^1 \left[\phi_{j_1}(r) - \frac{1}{1-\lambda} \int_\lambda^1 \phi_{j_1}(s) ds \right] \left[\phi_{j_2}(r) - \frac{1}{1-\lambda} \int_\lambda^1 \phi_{j_2}(s) ds \right] dr, \end{aligned}$$

where

$$\begin{aligned} & \frac{1}{\lambda^2} \int_0^\lambda \left[\phi_{j_1}(r) - \frac{1}{\lambda} \int_0^\lambda \phi_{j_1}(s) ds \right] \left[\phi_{j_2}(r) - \frac{1}{\lambda} \int_0^\lambda \phi_{j_2}(s) ds \right] dr \\ &= \frac{1}{\lambda^2} \left[\int_0^\lambda \int_0^\lambda \delta(r-s) \phi_{j_1}(r) \phi_{j_2}(s) dr ds - \frac{1}{\lambda} \int_0^\lambda \int_0^\lambda \phi_{j_1}(r) \phi_{j_2}(s) dr ds \right] \\ &= \int_0^1 \int_0^1 \left[\delta(r-s) - \frac{1}{\lambda} \right] \frac{1 \{ (r, s) \in [0, \lambda] \times [0, \lambda] \}}{\lambda^2} \phi_{j_1}(r) \phi_{j_2}(s) dr ds. \end{aligned}$$

Similarly,

$$\begin{aligned} & \frac{1}{(1-\lambda)^2} \int_\lambda^1 \left[\phi_{j_1}(r) - \frac{1}{1-\lambda} \int_\lambda^1 \phi_{j_1}(s) ds \right] \left[\phi_{j_2}(r) - \frac{1}{1-\lambda} \int_\lambda^1 \phi_{j_2}(s) ds \right] dr \\ &= \frac{1}{(1-\lambda)^2} \int_\lambda^1 \int_\lambda^1 \left[\delta(r-s) - \frac{1}{1-\lambda} \right] \phi_{j_1}(r) \phi_{j_2}(s) dr ds \\ &= \int_0^1 \int_0^1 \left[\delta(r-s) - \frac{1}{1-\lambda} \right] \frac{1 \{ (r, s) \in [\lambda, 1] \times [\lambda, 1] \}}{(1-\lambda)^2} \phi_{j_1}(r) \phi_{j_2}(s) dr ds. \end{aligned}$$

Therefore,

$$\int_0^1 \tilde{\phi}_{j_1}(r; \lambda) \tilde{\phi}_{j_2}(r; \lambda) dr = \int_0^1 \int_0^1 C(r, s; \lambda) \phi_{j_1}(r) \phi_{j_2}(s) dr ds.$$

■

The following lemmas facilitate the proof of Theorems 5.1 and 5.2.

Lemma 8.1 *Let Assumptions 5.1 and 5.2 hold. Let*

$$\mathbb{Q} = Q_{XX}^{-1} Q_{XZ} Q_{Z \cdot X}^{-1} Q_{ZX} Q_{XX}^{-1}, Q_{Z \cdot X} = Q_{ZZ} - Q_{ZX} Q_{XX}^{-1} Q_{XZ}.$$

Then

$$\sqrt{T}(\hat{\theta} - \theta) \rightarrow^d Q_{[\tilde{X}, Z]}^{-1} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix}$$

where

$$\begin{aligned} Q_{[\tilde{X}, Z]}^{-1} &= \begin{pmatrix} Q^{\tilde{X}\tilde{X}} & Q^{\tilde{X}Z} \\ Q^{Z\tilde{X}} & Q^{ZZ} \end{pmatrix} \text{ for} \\ Q^{\tilde{X}\tilde{X}} &= \begin{pmatrix} \lambda^{-1} Q_{XX}^{-1} & O \\ O & (1-\lambda)^{-1} Q_{XX}^{-1} \end{pmatrix} + \begin{pmatrix} \mathbb{Q} & \mathbb{Q} \\ \mathbb{Q} & \mathbb{Q} \end{pmatrix} \\ Q^{\tilde{X}Z} &= - \begin{pmatrix} Q_{XX}^{-1} Q_{XZ} Q_{Z \cdot X}^{-1} \\ Q_{XX}^{-1} Q_{XZ} Q_{Z \cdot X}^{-1} \end{pmatrix}, Q^{Z\tilde{X}} = - [Q_{Z \cdot X}^{-1} Q_{XZ} Q_{XX}^{-1} \quad Q_{Z \cdot X}^{-1} Q_{XZ} Q_{XX}^{-1}] \\ Q^{ZZ} &= Q_{Z \cdot X}^{-1}. \end{aligned}$$

Proof of Lemma 8.1. Under Assumptions 5.1 and 5.2, we have

$$\hat{Q}_{[\tilde{X}, Z]} \rightarrow^p Q_{[\tilde{X}, Z]} = \begin{pmatrix} \lambda Q_{XX} & O & \lambda Q_{XZ} \\ O & (1-\lambda) Q_{XX} & (1-\lambda) Q_{XZ} \\ \lambda Q_{ZX} & (1-\lambda) Q_{ZX} & Q_{ZZ} \end{pmatrix},$$

and

$$\frac{1}{\sqrt{T}} \begin{pmatrix} \tilde{X}'U \\ Z'U \end{pmatrix} \rightarrow^d \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix}.$$

Hence

$$\sqrt{T}(\hat{\theta} - \theta) = \hat{Q}_{[\tilde{X}, Z]}^{-1} \frac{1}{\sqrt{T}} \begin{pmatrix} \tilde{X}'U \\ Z'U \end{pmatrix} \rightarrow^d Q_{[\tilde{X}, Z]}^{-1} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix}.$$

Next, we derive the inverse of $Q_{[\tilde{X}, Z]}$. By the matrix partitioned inverse formula, we have

$$Q_{[\tilde{X}, Z]}^{-1} = \begin{pmatrix} \lambda \cdot Q_{XX} & O & \lambda \cdot Q_{XZ} \\ O & (1-\lambda) \cdot Q_{XX} & (1-\lambda) \cdot Q_{XZ} \\ \lambda \cdot Q_{ZX} & (1-\lambda) \cdot Q_{ZX} & Q_{ZZ} \end{pmatrix}^{-1} = \begin{pmatrix} Q^{\tilde{X}\tilde{X}} & Q^{\tilde{X}Z} \\ Q^{Z\tilde{X}} & Q^{ZZ} \end{pmatrix},$$

where

$$\begin{aligned} Q^{\tilde{X}\tilde{X}} &= \left[\begin{pmatrix} \lambda \cdot Q_{XX} & O \\ O & (1-\lambda) \cdot Q_{XX} \end{pmatrix} - \begin{pmatrix} \lambda \cdot Q_{XZ} \\ (1-\lambda) \cdot Q_{XZ} \end{pmatrix} Q_{ZZ}^{-1} [\lambda \cdot Q_{ZX}, (1-\lambda) \cdot Q_{ZX}] \right]^{-1}, \\ Q^{\tilde{X}Z} &= -Q^{\tilde{X}\tilde{X}} \begin{pmatrix} \lambda \cdot Q_{XZ} \\ (1-\lambda) \cdot Q_{XZ} \end{pmatrix} Q_{ZZ}^{-1}, \quad Q^{Z\tilde{X}} = -Q_{ZZ}^{-1} (\lambda \cdot Q_{ZX}, (1-\lambda) \cdot Q_{ZX}) Q^{\tilde{X}\tilde{X}}, \\ Q^{ZZ} &= \left\{ Q_{ZZ} - (\lambda \cdot Q_{ZX} \quad (1-\lambda) \cdot Q_{ZX}) \begin{pmatrix} \lambda \cdot Q_{XX} & O \\ O & (1-\lambda) \cdot Q_{XX} \end{pmatrix}^{-1} \begin{pmatrix} \lambda \cdot Q_{XZ} \\ (1-\lambda) \cdot Q_{XZ} \end{pmatrix} \right\}^{-1} \\ &= [Q_{ZZ} - Q_{ZX} Q_{XX}^{-1} Q_{XZ}]^{-1} = Q_{Z \cdot X}^{-1}. \end{aligned}$$

Using the matrix inverse formula

$$(A - CB^{-1}C')^{-1} = A^{-1} + A^{-1}C(B - C'A^{-1}C)^{-1}C'A^{-1},$$

we have

$$\begin{aligned} Q^{\tilde{X}\tilde{X}} &= \begin{pmatrix} \lambda Q_{XX} & O \\ O & (1-\lambda)Q_{XX} \end{pmatrix}^{-1} \\ &+ \begin{pmatrix} \lambda Q_{XX} & O \\ O & (1-\lambda)Q_{XX} \end{pmatrix}^{-1} \begin{pmatrix} \lambda Q_{XZ} \\ (1-\lambda)Q_{XZ} \end{pmatrix} \\ &\times \left[Q_{ZZ} - \begin{pmatrix} \lambda Q_{XZ} \\ (1-\lambda)Q_{XZ} \end{pmatrix}' \begin{pmatrix} \lambda Q_{XX} & O \\ O & (1-\lambda)Q_{XX} \end{pmatrix}^{-1} \begin{pmatrix} \lambda Q_{XZ} \\ (1-\lambda)Q_{XZ} \end{pmatrix} \right]^{-1} \\ &\times \begin{pmatrix} \lambda Q_{XZ} \\ (1-\lambda)Q_{XZ} \end{pmatrix}' \begin{pmatrix} \lambda Q_{XX} & O \\ O & (1-\lambda)Q_{XX} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \lambda^{-1}Q_{XX}^{-1} & O \\ O & (1-\lambda)^{-1}Q_{XX}^{-1} \end{pmatrix} + \begin{pmatrix} \mathbb{Q} & \mathbb{Q} \\ \mathbb{Q} & \mathbb{Q} \end{pmatrix}. \end{aligned}$$

Using the above, we then have

$$\begin{aligned} Q^{\tilde{X}Z} &= -Q^{\tilde{X}\tilde{X}} \begin{pmatrix} \lambda \cdot Q_{XZ}Q_{ZZ}^{-1} \\ (1-\lambda) \cdot Q_{XZ}Q_{ZZ}^{-1} \end{pmatrix} \\ &= - \left[\begin{pmatrix} \lambda^{-1}Q_{XX}^{-1} & O \\ O & (1-\lambda)^{-1}Q_{XX}^{-1} \end{pmatrix} + \begin{pmatrix} \mathbb{Q} & \mathbb{Q} \\ \mathbb{Q} & \mathbb{Q} \end{pmatrix} \right] \begin{pmatrix} \lambda \cdot Q_{XZ}Q_{ZZ}^{-1} \\ (1-\lambda) \cdot Q_{XZ}Q_{ZZ}^{-1} \end{pmatrix} \\ &= - \left[\begin{pmatrix} Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} \\ Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} \end{pmatrix} + \begin{pmatrix} \mathbb{Q}Q_{XZ}Q_{ZZ}^{-1} \\ \mathbb{Q}Q_{XZ}Q_{ZZ}^{-1} \end{pmatrix} \right] \\ &= - \begin{pmatrix} Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} + \mathbb{Q}Q_{XZ}Q_{ZZ}^{-1} \\ Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} + \mathbb{Q}Q_{XZ}Q_{ZZ}^{-1} \end{pmatrix} = - \begin{pmatrix} Q_{XX}^{-1}Q_{XZ}Q_{Z \cdot X}^{-1} \\ Q_{XX}^{-1}Q_{XZ}Q_{Z \cdot X}^{-1} \end{pmatrix} \end{aligned}$$

and

$$Q^{Z\tilde{X}} = - \begin{pmatrix} Q_{XX}^{-1}Q_{XZ}Q_{Z \cdot X}^{-1} \\ Q_{XX}^{-1}Q_{XZ}Q_{Z \cdot X}^{-1} \end{pmatrix}' = - [Q_{Z \cdot X}^{-1}Q_{XZ}Q_{XX}^{-1}, Q_{Z \cdot X}^{-1}Q_{XZ}Q_{XX}^{-1}],$$

where we have used

$$\begin{aligned} &Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} + \mathbb{Q}Q_{XZ}Q_{ZZ}^{-1} \\ &= Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} + Q_{XX}^{-1}Q_{XZ}Q_{Z \cdot X}^{-1}Q_{ZX}Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} \\ &= Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} + Q_{XX}^{-1}Q_{XZ}Q_{Z \cdot X}^{-1}Q_{ZX}Q_{XX}^{-1}Q_{XZ}Q_{ZZ}^{-1} \\ &= Q_{XX}^{-1}Q_{XZ} [I_\ell + Q_{Z \cdot X}^{-1}Q_{ZX}Q_{XX}^{-1}Q_{XZ}] Q_{ZZ}^{-1} \\ &= Q_{XX}^{-1}Q_{XZ} [I_\ell + Q_{Z \cdot X}^{-1} [Q_{ZX}Q_{XX}^{-1}Q_{XZ} - Q_{ZZ}] + Q_{Z \cdot X}^{-1}Q_{ZZ}] Q_{ZZ}^{-1} \\ &= Q_{XX}^{-1}Q_{XZ}Q_{Z \cdot X}^{-1}. \end{aligned}$$

■

Lemma 8.2 *Let Assumptions 3.3, 5.1, and 5.2 hold. Define*

$$\Lambda_{Z \cdot X} = \Lambda_Z - Q_{ZX} Q_{XX}^{-1} \Lambda_X, \bar{\phi}_{j,0} = \int_0^1 \phi_j(r) dr$$

and

$$\begin{aligned} \bar{\phi}_{j,1}(r, \lambda) &= [\phi_j(r) - \bar{\phi}_{j,1}] \cdot 1\{r \leq \lambda\}, \quad \bar{\phi}_{j,2}(r, \lambda) = [\phi_j(r) - \bar{\phi}_{j,2}] \cdot 1\{r > \lambda\} \\ \bar{\phi}_j^*(r, \lambda) &= (\bar{\phi}_{j,1} - \bar{\phi}_{j,0}) \cdot 1\{r \leq \lambda\} + (\bar{\phi}_{j,2} - \bar{\phi}_{j,0}) \cdot 1\{r > \lambda\}. \end{aligned}$$

Then

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) (\tilde{X}'_t, Z'_t)' \hat{u}_t \\ & \rightarrow^d \begin{pmatrix} \Lambda_X \int_0^1 \bar{\phi}_{j,1}(r, \lambda) dW_{m+\ell}(r) \\ \Lambda_X \int_0^1 \bar{\phi}_{j,2}(r, \lambda) dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 [\bar{\phi}_{j,1}(r, \lambda) + \bar{\phi}_{j,1}(r, \lambda)] dW_{m+\ell}(r) + \Lambda_{Z \cdot X} \int_0^1 \bar{\phi}_j^*(r, \lambda) dW_{m+\ell}(r) \end{pmatrix} \end{aligned}$$

jointly over $j = 1, 2, \dots, K$.

Proof of Lemma 8.2. We have

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) (\tilde{X}'_t, Z'_t)' \hat{u}_t \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) (\tilde{X}'_t, Z'_t)' [Y_t - (\tilde{X}'_t, Z'_t)' \hat{\theta}] \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) (\tilde{X}'_t, Z'_t)' u_t - \left[\frac{1}{T} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) (\tilde{X}'_t, Z'_t)' (\tilde{X}'_t, Z'_t)' \right] \sqrt{T} (\hat{\theta} - \theta). \end{aligned}$$

We proceed to find the weak limit of each term in the above expression. Under Assumptions 3.3 and 5.2, we have

$$\begin{aligned} & \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) (\tilde{X}'_t, Z'_t)' u_t \\ &= \begin{pmatrix} \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) X_t u_t \cdot 1\{t \leq [T\lambda]\} \\ \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) X_t u_t \cdot 1\{t > [T\lambda]\} \\ \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) Z_t u_t \end{pmatrix} \rightarrow^d \begin{pmatrix} \Lambda_X \int_0^\lambda \phi_j(r) dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 \phi_j(r) dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 \phi_j(r) dW_{m+\ell}(r) \end{pmatrix} \end{aligned}$$

jointly over $j = 1, 2, \dots, K$.

To find the weak limit of $T^{-1} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) (\tilde{X}'_t, Z'_t)' (\tilde{X}'_t, Z'_t)'$, we note that under Assumption

5.1,

$$\begin{aligned}
& \frac{1}{T} \sum_{t=1}^{[Tr]} \left(\tilde{X}'_t, Z'_t \right)' \left(\tilde{X}'_t, Z'_t \right) \\
&= \begin{pmatrix} T^{-1} \sum_{t=1}^{[Tr]} X_t X'_t \cdot 1\{t \leq [T\lambda]\} & O & T^{-1} \sum_{t=1}^{[Tr]} X_t Z'_t \cdot 1\{t \leq [T\lambda]\} \\ O & T^{-1} \sum_{t=1}^{[Tr]} X_t X'_t \cdot 1\{t > [T\lambda]\} & T^{-1} \sum_{t=1}^{[Tr]} X_t Z'_t \cdot 1\{t > [T\lambda]\} \\ T^{-1} \sum_{t=1}^{[Tr]} Z'_t X_t \cdot 1\{t \leq [T\lambda]\} & T^{-1} \sum_{t=1}^{[Tr]} Z'_t X'_t \cdot 1\{t > [T\lambda]\} & T^{-1} \sum_{t=1}^{[Tr]} Z'_t Z'_t \end{pmatrix} \\
&\rightarrow^p \begin{pmatrix} rQ_{XX} & O & rQ_{XZ} \\ O & O & O \\ rQ_{ZX} & O & rQ_{ZZ} \end{pmatrix} \cdot 1\{r \leq \lambda\} + \begin{pmatrix} \lambda Q_{XX} & O & \lambda Q_{XZ} \\ O & (r-\lambda)Q_{XX} & (r-\lambda)Q_{XZ} \\ \lambda Q_{ZX} & (r-\lambda)Q_{ZX} & rQ_{ZZ} \end{pmatrix} \cdot 1\{r > \lambda\}.
\end{aligned} \tag{13}$$

Combining this with Assumption 3.3, we have

$$\begin{aligned}
& \frac{1}{T} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \left(\tilde{X}'_t, Z'_t \right)' \left(\tilde{X}'_t, Z'_t \right) \\
&\rightarrow^p \begin{pmatrix} \int_0^\lambda \phi_j(r) dr Q_{XX} & O & \int_0^\lambda \phi_j(r) dr Q_{XZ} \\ O & O & O \\ \int_0^\lambda \phi_j(r) dr Q_{ZX} & O & \int_0^\lambda \phi_j(r) dr Q_{ZZ} \end{pmatrix} + \begin{pmatrix} O & O & O \\ O & \left[\int_\lambda^1 \phi_j(r) dr \right] Q_{XX} & \left[\int_\lambda^1 \phi_j(r) dr \right] Q_{XZ} \\ O & \left[\int_\lambda^1 \phi_j(r) dr \right] Q_{ZX} & \left[\int_\lambda^1 \phi_j(r) dr \right] Q_{ZZ} \end{pmatrix} \\
&= \begin{pmatrix} \int_0^\lambda \phi_j(r) dr \cdot Q_{XX} & O & \int_0^\lambda \phi_j(r) dr \cdot Q_{XZ} \\ O & \int_\lambda^1 \phi_j(r) dr \cdot Q_{XX} & \int_\lambda^1 \phi_j(r) dr \cdot Q_{XZ} \\ \int_0^\lambda \phi_j(r) dr \cdot Q_{ZX} & \int_\lambda^1 \phi_j(r) dr \cdot Q_{ZX} & \int_0^1 \phi_j(r) dr \cdot Q_{ZZ} \end{pmatrix} := Q_{\phi_j, [\tilde{X}, Z]}.
\end{aligned}$$

Hence

$$\begin{aligned}
& \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \left(\tilde{X}'_t, Z'_t \right)' \hat{u}_t \\
&\rightarrow^d \begin{pmatrix} \Lambda_X \int_0^\lambda \phi_j(r) dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 \phi_j(r) dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 \phi_j(r) dW_{m+\ell}(r) \end{pmatrix} - Q_{\phi_j, [\tilde{X}, Z]} Q_{[\tilde{X}, Z]}^{-1} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix}
\end{aligned}$$

jointly over $j = 1, 2, \dots, K$. Some algebraic manipulations show that

$$Q_{\phi_j, [\tilde{X}, Z]} Q_{[\tilde{X}, Z]}^{-1} = \begin{pmatrix} \bar{\phi}_{j,1} \cdot I_m & O & O \\ O & \bar{\phi}_{j,2} \cdot I_m & O \\ (\bar{\phi}_{j,1} - \bar{\phi}_{j,0}) Q_{ZX} Q_{XX}^{-1} & (\bar{\phi}_{j,2} - \bar{\phi}_{j,0}) Q_{ZX} Q_{XX}^{-1} & \bar{\phi}_{j,0} \cdot I_\ell \end{pmatrix}.$$

As a result,

$$\begin{aligned}
& \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) \left(\tilde{X}'_t, Z'_t \right)' \hat{u}_t \\
& \rightarrow^d \begin{pmatrix} \Lambda_X \int_0^\lambda \phi_j(r) dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 \phi_j(r) dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 \phi_j(r) dW_{m+\ell}(r) \end{pmatrix} \\
& - \begin{pmatrix} \bar{\phi}_{j,1} \cdot I_m & O & O \\ O & \bar{\phi}_{j,2} \cdot I_m & O \\ (\bar{\phi}_{j,1} - \bar{\phi}_{j,0}) Q_{ZX} Q_{XX}^{-1} & (\bar{\phi}_{j,2} - \bar{\phi}_{j,0}) Q_{ZX} Q_{XX}^{-1} & \bar{\phi}_{j,0} \cdot I_\ell \end{pmatrix} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix} \\
& = \begin{pmatrix} \Lambda_X \int_0^1 [\phi_{1,j}(r, \lambda)] dW_{m+\ell}(r) \\ \Lambda_X \int_0^1 [\phi_{2,j}(r, \lambda)] dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 [\phi_j(r) - \bar{\phi}_{j,0}] dW_{m+\ell}(r) - Q_{ZX} Q_{XX}^{-1} \Lambda_X \int_0^1 \bar{\phi}_j^*(r, \lambda) dW_{m+\ell}(r) \end{pmatrix} \quad (14)
\end{aligned}$$

jointly over $j = 1, 2, \dots, K$. The last element in (14) can be rewritten as

$$\begin{aligned}
& \Lambda_Z \int_0^1 [\phi_j(r) - \bar{\phi}_{j,0}] dW_{m+\ell}(r) - Q_{ZX} Q_{XX}^{-1} \Lambda_X \int_0^1 \bar{\phi}_j^*(r, \lambda) dW_{m+\ell}(r) \\
& = \Lambda_Z \left[\int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_{m+\ell}(r) + \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_{m+\ell}(r) \right] \\
& + \Lambda_{Z \cdot X} \left[\int_0^\lambda [(\bar{\phi}_{j,1} - \bar{\phi}_{j,0})] dW_{m+\ell}(r) + \int_\lambda^1 [(\bar{\phi}_{j,2} - \bar{\phi}_{j,0})] dW_{m+\ell}(r) \right] \\
& = \Lambda_Z \int_0^1 [\bar{\phi}_{j,1}(r, \lambda) + \bar{\phi}_{j,1}(r, \lambda)] dW_{m+\ell}(r) + \Lambda_{Z \cdot X} \int_0^1 \bar{\phi}_j^*(r, \lambda) dW_{m+\ell}(r).
\end{aligned}$$

■

Proof of Theorem 5.1. For $R = (\mathcal{R}_\beta, -\mathcal{R}_\beta, O_\gamma)$, we have, using the expression for $Q_{[\tilde{X}, Z]}^{-1}$ from Lemma 8.1,

$$\begin{aligned}
RQ_{[\tilde{X}, Z]}^{-1} &= (\mathcal{R}_\beta, -\mathcal{R}_\beta, O_\gamma) \begin{pmatrix} Q^{\tilde{X}\tilde{X}} & Q^{\tilde{X}Z} \\ Q^{Z\tilde{X}} & Q^{ZZ} \end{pmatrix} \\
&= [(\mathcal{R}_\beta, -\mathcal{R}_\beta) Q^{\tilde{X}\tilde{X}}, (\mathcal{R}_\beta, -\mathcal{R}_\beta) Q^{\tilde{X}Z}] \\
&= [(\mathcal{R}_\beta, -\mathcal{R}_\beta) \begin{pmatrix} \lambda^{-1} Q_{XX}^{-1} & O \\ O & (1-\lambda)^{-1} Q_{XX}^{-1} \end{pmatrix}, O] \\
&= [\mathcal{R}_\beta \lambda^{-1} Q_{XX}^{-1}, -\mathcal{R}_\beta (1-\lambda)^{-1} Q_{XX}^{-1}, O]. \quad (15)
\end{aligned}$$

So, by Lemma 8.1, we have

$$\begin{aligned}
R\sqrt{T}(\hat{\theta} - \theta) &\rightarrow^d RQ_{[\tilde{X}, Z]}^{-1} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix} \\
&= \left[\mathcal{R}_\beta \lambda^{-1} Q_{XX}^{-1}, -\mathcal{R}_\beta (1 - \lambda)^{-1} Q_{XX}^{-1}, O \right] \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix} \\
&= \mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X \left(\frac{1}{\lambda} \int_0^\lambda dW_{m+\ell}(\lambda) - \frac{1}{1 - \lambda} \int_\lambda^1 dW_{m+\ell}(\lambda) \right). \tag{16}
\end{aligned}$$

This holds jointly with the weak convergence in Lemma 8.2.

Next, using (15) and Lemma 8.2, we have

$$\begin{aligned}
&R\hat{Q}_{[\tilde{X}, Z]}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) [\tilde{X}'_t, Z'_t]' \hat{u}_t \\
&\rightarrow^d \left[\mathcal{R}_\beta \lambda^{-1} Q_{XX}^{-1}, -\mathcal{R}_\beta (1 - \lambda)^{-1} Q_{XX}^{-1} \right] \begin{pmatrix} \Lambda_X \left[\int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_{m+\ell}(r) \right] \\ \Lambda_X \left[\int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_{m+\ell}(r) \right] \end{pmatrix} \\
&= \mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X \left[\frac{1}{\lambda} \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_{m+\ell}(r) - \frac{1}{1 - \lambda} \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_{m+\ell}(r) \right] \\
&= \mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X \int_0^1 \tilde{\phi}_j(r; \lambda) dW_{m+\ell}(r) \tag{17}
\end{aligned}$$

jointly over $j = 1, \dots, K$.

Combining (16), (17), and the continuous mapping theorem, we obtain

$$\begin{aligned}
F_T &= T \cdot (R\hat{\theta} - r)' \left[R\hat{Q}_{[\tilde{X}, Z]}^{-1} \hat{\Omega}_{[\tilde{X}, Z]} \hat{Q}_{[\tilde{X}, Z]}^{-1} R' \right]^{-1} (R\hat{\theta} - r) \\
&= \left[R\sqrt{T}(\hat{\theta} - \theta) \right]' \left\{ \frac{1}{K} \sum_{j=1}^K \left[R\hat{Q}_{[\tilde{X}, Z]}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) [\tilde{X}'_t, Z'_t]' \hat{u}_t \right]^{\otimes 2} \right\}^{-1} R\sqrt{T}(\hat{\theta} - \theta) \\
&\rightarrow^d \left\{ \mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X \left(\frac{1}{\lambda} \int_0^\lambda dW_{m+\ell}(\lambda) - \frac{1}{1 - \lambda} \int_\lambda^1 dW_{m+\ell}(\lambda) \right) \right\}' \\
&\times \left\{ \frac{1}{K} \sum_{j=1}^K \left[\mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X \int_0^1 \tilde{\phi}_j(r; \lambda) dW_{m+\ell}(r) \right]^{\otimes 2} \right\}^{-1} \\
&\times \left\{ \mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X \left(\frac{1}{\lambda} \int_0^\lambda dW_{m+\ell}(\lambda) - \frac{1}{1 - \lambda} \int_\lambda^1 dW_{m+\ell}(\lambda) \right) \right\}.
\end{aligned}$$

Note that

$$\mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X W_{m+\ell}(\lambda) =^d \mathcal{A}_p W_p(\lambda)$$

for an invertible $p \times p$ matrix \mathcal{A}_p such that $\mathcal{A}'_p \mathcal{A}_p = \mathcal{R}_\beta Q_{XX}^{-1} \Lambda_X \Lambda'_X Q_{XX}^{-1} \mathcal{R}'_\beta$. Using this distribu-

tional equivalence, we then have

$$F_T \rightarrow^d \left(\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right)' \times \left\{ \frac{1}{K} \sum_{j=1}^K \left[\int_0^1 \tilde{\phi}_j(r; \lambda) dW_p(r) \right]^{\otimes 2} \right\}^{-1} \\ \times \left(\frac{1}{\lambda} \int_0^\lambda dW_p(\lambda) - \frac{1}{1-\lambda} \int_\lambda^1 dW_p(\lambda) \right).$$

The result for the t statistic can be proved in the same manner. ■

Proof of Theorem 5.2. Part (a) When $R = (O_\beta, O_\beta, \mathcal{R}_\gamma)$, we have

$$RQ_{[\tilde{X}, Z]}^{-1} = (O_\beta, O_\beta, \mathcal{R}_\gamma) \begin{pmatrix} Q^{\tilde{X}\tilde{X}} & Q^{\tilde{X}Z} \\ Q^{Z\tilde{X}} & Q^{ZZ} \end{pmatrix} = [\mathcal{R}_\gamma Q^{Z\tilde{X}}, \mathcal{R}_\gamma Q^{ZZ}] \\ = \mathcal{R}_\gamma \begin{bmatrix} -Q_{Z \cdot X}^{-1} Q_{XZ} Q_{XX}^{-1}, & -Q_{Z \cdot X}^{-1} Q_{XZ} Q_{XX}^{-1}, & Q_{Z \cdot X} \end{bmatrix}.$$

Using this and Lemma 8.1, we have

$$R\sqrt{T}(\hat{\theta} - \theta) \\ \rightarrow^d \mathcal{R}_\gamma \begin{bmatrix} -Q_{Z \cdot X}^{-1} Q_{XZ} Q_{XX}^{-1}, & -Q_{Z \cdot X}^{-1} Q_{XZ} Q_{XX}^{-1}, & Q_{Z \cdot X} \end{bmatrix} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix} \\ = \mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} \int_0^1 dW_{m+\ell}(r).$$

Next, by Lemma 8.2, we have

$$R\hat{Q}_{[\tilde{X}, Z]}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T \phi_j \left(\frac{t}{T} \right) [\tilde{X}'_t, Z'_t]' \hat{u}_t \\ \rightarrow^d -\mathcal{R}_\gamma Q_{Z \cdot X}^{-1} Q_{ZX} Q_{XX}^{-1} \Lambda_X \left\{ \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_{m+\ell}(r) + \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_{m+\ell}(r) \right\} \\ + \mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_Z \left\{ \int_0^\lambda [\phi_j(r) - \bar{\phi}_{j,1}] dW_{m+\ell}(r) + \int_\lambda^1 [\phi_j(r) - \bar{\phi}_{j,2}] dW_{m+\ell}(r) \right\} \\ + \mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} \left[\int_0^\lambda (\bar{\phi}_{j,1} - \bar{\phi}_{j,0}) dW_{m+\ell}(r) + \int_\lambda^1 (\bar{\phi}_{j,2} - \bar{\phi}_{j,0}) dW_{m+\ell}(r) \right] \\ = \mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} \int_0^1 [\phi_j(r) - \bar{\phi}_{j,0}] dW_{m+\ell}(r)$$

jointly over $j = 1, 2, \dots, K$.

Therefore, using the continuous mapping theorem, we obtain

$$F_T \rightarrow^d \left\{ \mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} \int_0^1 dW_{m+\ell}(r) \right\}' \\ \times \left\{ \frac{1}{K} \sum_{j=1}^K \left[\mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} \int_0^1 [\phi_j(r) - \bar{\phi}_{j,0}] dW_{m+\ell}(r) \right]^{\otimes 2} \right\}^{-1} \\ \times \left\{ \mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} \int_0^1 dW_{m+\ell}(r) \right\}.$$

Note that $\mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} W_{m+\ell}(r) =^d A_p^* W_p(r)$ where $A_p^* = (\mathcal{R}_\gamma Q_{Z \cdot X}^{-1} \Lambda_{Z \cdot X} \Lambda'_{Z \cdot X} Q_{Z \cdot X}^{-1} \mathcal{R}'_\gamma)^{1/2}$ is an invertible square matrix. Plugging this into the above, we obtain

$$F_T \rightarrow^d W_p(1)' \left\{ \frac{1}{K} \sum_{j=1}^K \left[\int_0^1 \phi_j(r) dW_p(r) \right]^{\otimes 2} \right\}^{-1} W_p(1).$$

When $\{\phi_j(r)\}$ are orthonormal on $L^2[0, 1]$ and $\int_0^1 \phi_j(r) dr = 0$, $\left[\int_0^1 \phi_j(r) dW_p(r) \right]^{\otimes 2} \sim^{iid} \text{Wishart}$, and $\int_0^1 \phi_j(r) dW_p(r)$ is independent of $W_p(1)$. So, the above limit distribution has the same form as Hotelling's T^2 distribution. Using the relationship between the T^2 distribution and the F distribution, we obtain that

$$\tilde{F}_T^* = \frac{K - p + 1}{Kp} F_T \rightarrow^d F_{p, K-p+1}.$$

The result for the t statistic can be similarly proved. Details are omitted here. ■

The next proposition gives the fixed-smoothing asymptotic distributions of the F and t statistics when there is no restriction on the matrix R .

Proposition 8.1 *Let Assumptions 3.3, 5.1, and 5.2 hold. Then*

F_T

$$\rightarrow^d \left\{ RQ_{[\tilde{X}, Z]}^{-1} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix} \right\}' \cdot \left[RQ_{[\tilde{X}, Z]}^{-1} \Omega_\infty Q_{[\tilde{X}, Z]}^{-1} R' \right]^{-1} \cdot \left\{ RQ_{[\tilde{X}, Z]}^{-1} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix} \right\}$$

where

$$\Omega_\infty = \frac{1}{K} \sum_{j=1}^K \left\{ \begin{pmatrix} \Lambda_X \int_0^1 \phi_{j,1}(r, \lambda) dW_{m+\ell}(r) \\ \Lambda_X \int_0^1 \phi_{j,2}(r, \lambda) dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 [\phi_{j,1}(r, \lambda) + \phi_{j,2}(r, \lambda)] dW_{m+\ell}(r) + \Lambda_{Z \cdot X} \int_0^1 \bar{\phi}_j^*(r, \lambda) dW_{m+\ell}(r) \end{pmatrix} \right\}^{\otimes 2}.$$

When $p = 1$,

$$t_T \rightarrow^d RQ_{[\tilde{X}, Z]}^{-1} \begin{pmatrix} \Lambda_X \int_0^\lambda dW_{m+\ell}(r) \\ \Lambda_X \int_\lambda^1 dW_{m+\ell}(r) \\ \Lambda_Z \int_0^1 dW_{m+\ell}(r) \end{pmatrix} \cdot \left[RQ_{[\tilde{X}, Z]}^{-1} \Omega_\infty Q_{[\tilde{X}, Z]}^{-1} R' \right]^{-1/2}.$$

Proof of Proposition 8.1. The proposition follows directly from Lemma 8.1, Lemma 8.2, and the continuous mapping theorem. ■

References

- Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59:817–858.
- Cho, C.-K. and Vogelsang, T. J. (2017). Fixed-b inference for testing structural change in a time series regression. *Econometrics*, 5(1).

- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3):591–605.
- Giles, D. and Scott, M. (1992). Some consequences of using the Chow test in the context of autocorrelated disturbances. *Economics Letters*, 38(2):145 – 150.
- Hwang, J. and Sun, Y. (2017). Asymptotic F and t tests in an efficient GMM setting. *Journal of Econometrics*, 198:277–295.
- Jansson, M. (2004). On the error of rejection probability in simple autocorrelation robust tests. *Econometrica*, 72:937–946.
- Kiefer, N. M. and Vogelsang, T. J. (2002a). Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size. *Econometric Theory*, 18:1350–1366.
- Kiefer, N. M. and Vogelsang, T. J. (2002b). Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. *Econometrica*, 70:2093–2095.
- Kiefer, N. M. and Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, 21:1130–1164.
- Krämer, W. (1989). The robustness of the chow test to autocorrelation among disturbances. In Hackl, P., editor, *Statistical Analysis and Forecasting of Economic Structural Change*, pages 45–52. Springer.
- Lazarus, E., Lewis, D. J., Stock, J. H., and Watson, M. W. (2018). HAR inference: Recommendations for practice. *Journal of Business & Economic Statistics*, 36(4):541–559.
- Liu, C. and Sun, Y. (2019). A simple and trustworthy asymptotic t test in difference-in-differences regressions. *Journal of Econometrics*, 210(2):327 – 362.
- Martínez-Iriarte, J., Sun, Y., and Wang, X. (2020). Asymptotic F tests under possibly weak identification. *Journal of Econometrics*, 218(1):140 – 177.
- Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.
- Phillips, P. C. B. (2005). HAC estimation by automated regression. *Econometric Theory*, 21(1):116–142.
- Phillips, P. C. B. and Durlauf, S. N. (1986). Multiple Time Series Regression with Integrated Processes. *Review of Economic Studies*, 53(4):473–495.
- Phillips, P. C. B. and Solo, V. (1992). Asymptotics for linear processes. *Ann. Statist.*, 20(2):971–1001.
- Sun, Y. (2011). Robust trend inference with series variance estimator and testing-optimal smoothing parameter. *Journal of Econometrics*, 164:345–366.
- Sun, Y. (2013). A heteroskedasticity and autocorrelation robust F test using orthonormal series variance estimator. *Econometrics Journal*, 16:1–26.

- Sun, Y. (2014). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. *Journal of Econometrics*, 178:659–677.
- Sun, Y. and Kim, M. S. (2012). Simple and powerful GMM over-identification tests with accurate size. *Journal of Econometrics*, 166:267–281.
- Sun, Y., Philips, P. C. B., and Jin, S. (2008). Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. *Econometrica*, 76:175–194.
- Wang, X. and Sun, Y. (2020). An asymptotic F test for uncorrelatedness in the presence of time series dependence. *Journal of Time Series Analysis*, 41(4):536–550.