Journal of Econometrics 211 (2019) 206–242

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

A model-free consistent test for structural change in regression possibly with endogeneity[☆]

Zhonghao Fu[a,b,1], Yongmiao Hong[c,d,*,2]

[a] Fudan University, China
[b] Shanghai Institute of International Finance and Economics, China
[c] Cornell University, United States
[d] Xiamen University, China

ARTICLE INFO

Article history:
Available online 13 December 2018

JEL classification:
C12
C14

Keywords:
Endogeneity
Fourier transform
Local stationarity
Long-run variance
Model-free test
Structural change

ABSTRACT

Structural instability leads to misleading inference and imprecise prediction of time series models that assume stationarity. We propose a model-free consistent test for structural change in regression by testing the instability of the Fourier transform of data. This novel approach avoids smoothed nonparametric estimation of the unknown regression function and so is free of the "curse of dimensionality" problem. Unlike the existing literature, we allow for endogenous and discrete regressors. By using a proper choice of weighting functions for the transform parameters in the Fourier transform, we avoid numerical integration so that our test statistic is easy to compute. Our test statistic has a convenient asymptotic $N(0, 1)$ distribution under the null hypothesis of no structural change and is consistent against a large class of smooth structural changes as well as abrupt structural breaks with unknown break dates. A Monte Carlo study and an empirical application show that our test performs reasonably well in finite samples.

1. Introduction

Nonlinearity often exists in economic time series data and various nonlinear time series models have been proposed to capture different forms of nonlinearity. Examples include the threshold model (Tong, 1990), the smooth transition model (Teräsvirta, 1994), the Markov regime-switching model (Hamilton, 1989), the artificial neural network model (White, 1989), the functional coefficient model (Cai et al., 2000), and the nonlinear factor model (Bai and Ng, 2008). These nonlinear models can fit a sample reasonably well, but they often perform poorly in out-of-sample prediction (see, e.g., De Gooijer and Kumar, 1992; Clements et al., 2004; González-Rivera and Lee, 2009). Teräsvirta et al. (2010) attribute the failure of nonlinear time series models in out-of-sample prediction to the change of nonlinear features detected in-sample. They argue that only when the nonlinearity documented in-sample also exists in the forecasting period will a nonlinear time series model outperform

☆ This is part of the doctoral dissertation of the first author. We thank three referees for careful and constructive comments that lead to an significant improvement over an earlier version. We also thank Jiti Gao, Nicholas Kiefer, G. Andrew Karolyi, Tae-hwy Lee, Aman Ullah, and participants in New York Camp Econometrics X for their helpful comments and suggestions. We thank Yundong Tu for providing the data in Lee et al. (2014). All remaining errors are solely ours.

* Corresponding author.
E-mail addresses: zhfu@fudan.edu.cn (Z. Fu), yh20@cornell.edu (Y. Hong).

1 Correspondence to: School of Economics, Fudan University; Shanghai Institute of International Finance and Economics.

2 Correspondence to: Department of Economics and Department of Statistical Science, Cornell University; Wang Yanan Institute for Studies in Economics, and Paula and Gregory Chow Center for Economic Research, Xiamen University.

https://doi.org/10.1016/j.jeconom.2018.12.014
0304-4076/

a linear time series model in out-of-sample prediction. However, the nonlinearity detected in-sample could change or even disappear out-of-sample, causing misleading inference and imprecise prediction.

Economic relationships may suffer from structural change due to changing economic environments such as shocks, new policies, preference changes, technological progress, etc. When the time span is long, it is likely that a time series econometric model becomes unstable over time. Many studies have investigated the impact of structural changes on modeling, inference, and prediction (see, e.g., Welch and Goyal, 2008; Stock and Watson, 2009). For any nonlinearity detected in-sample, there exist two possibilities: either the data generating process (DGP) is nonlinear and time-invariant, or the DGP varies over time. Both scenarios can generate similar nonlinear patterns in the data, but their implications on modeling, inference, and prediction are quite different. Therefore it is important for a researcher to detect whether the underlying economic relationship is time-invariant before he or she uses any time series model. This paper proposes a model-free consistent test for structural change in regression so that it can be used to distinguish structural change from nonlinearity and/or model misspecification, among many other things.

Testing instability of parameters in a time series parametric regression model has drawn much attention in the literature. Often a linear time series regression model is considered. Chow (1960), Andrews and Fair (1988), Andrews (1993), Andrews and Ploberger (1994), Bai and Perron (1998, 2003), Hansen (2001), Elliott and Müller (2006), and Qu and Perron (2007) develop tests for parameter constancy against abrupt structural breaks, with known or unknown breakpoints. Lin and Teräsvirta (1994), Cai (2007), Chen and Hong (2012), Kristensen (2012), Zhang and Wu (2012), and Cai et al. (2015) test parameter constancy against smooth structural changes. All of these tests are based on the assumption that the parametric regression model is correctly specified, and regressors are exogenous. When endogenous regressors are present, Hall et al. (2012), Perron and Yamamoto (2014, 2015), and Chen (2015) discuss how to model and test structural changes in a linear time series regression model.

Tests developed under a parametric framework are valid only when a model is correctly specified. However, economic theory usually does not suggest any concrete functional form for the regression function. When a parametric regression model is misspecified, a rejection of stability could be caused by model misspecification, rather than structural changes. To avoid this drawback, we use a nonparametric approach that does not assume any parametric functional form for the regression function. We emphasize that Hidalgo (1995) has pioneered to propose a nonparametric conditional moment test for structural change in regression. Su and Xiao (2008) propose CUSUM-type tests for structural change in a nonparametric time series regression model that allows nonstationary covariates. Su and White (2010) consider testing for structural change in a partially linear regression model. Vogt (2015) also proposes a nonparametric test for structural change in regression by checking whether the shape of a regression function is stable over time, which, to our knowledge, is the only consistent test for structural change in a nonparametric regression model in the existing literature. All of these tests can detect instability of an unknown regression function without model misspecification. However, the existing approaches require smoothed nonparametric estimation of the regression function, so they suffer from the notorious "curse of dimensionality" problem at least in finite samples, especially when the dimension of regressors is high. Also, they restrict regressors to be exogenous and continuous and require to use higher order kernels. When the covariates are endogenous or discrete, no existing test is available to detect structural changes in the unknown regression function.

This paper contributes to the literature on testing for structural change in an unknown regression function in several directions. First, unlike the existing tests, we use a Fourier transform approach that avoids nonparametric estimation of the regression function. As a result, our approach is free of the "curse of dimensionality" problem. This is achieved by testing instability of the Fourier transform of the unknown regression function. The idea of using the Fourier transform to deal with the curse of dimensionality has been used in the literature. For instance, Su and White (2007), and Wang and Hong (2017) propose characteristic function-based tests for conditional independence. Hong et al. (2017) test for strict stationarity via nonparametric smoothing of the joint characteristic function over time. In this paper, we first transform the unknown regression function to the frequency domain and then estimate the time-varying Fourier transform via nonparametric smoothing over a rescaled univariate time index. As a result, our test is asymptotically more efficient than Vogt's (2015) consistent test for structural change under a class of local alternatives. Although our test is asymptotically less powerful than the nonparametric tests of Hidalgo (1995) and Su and Xiao (2008) under certain smooth local alternatives, it is a consistent test under a large class of global alternatives, and it is asymptotically more powerful under a class of nonsmooth local alternatives. Furthermore, since we avoid nonparametric estimation of the unknown regression function, we allow regressors to be continuous, discrete or a mixture of both. In contrast, Hidalgo (1995), Su and Xiao (2008), and Vogt (2015) all require regressors to be continuous and impose certain smoothness conditions on the density functions.

Second, we allow for both exogenous and endogenous covariates. In contrast, all of the existing nonparametric tests require exogenous covariates. Nonparametric regression with endogeneity has drawn increasing attention in the literature (see, e.g., Newey et al., 1999; Ai and Chen, 2003; Newey and Powell, 2003; Hall and Horowitz, 2005; Blundell et al., 2007; Horowitz, 2011; Darolles et al., 2011). By using the Fourier transform of the unknown regression function, we avoid solving the "ill-posed" inverse problem in nonparametric instrumental variable estimation. Our test can serve as a pre-test for instability of the unknown regression function before using any nonparametric instrumental variable estimation. This greatly expands the literature since all of the existing nonparametric structural change tests are only applicable to a regression with exogeneity.

Third, to ensure the consistency of our test against both smooth structural changes and abrupt structural breaks, we have to integrate a transform parameter vector whose dimension is the same as that of covariates or instruments. That is computationally challenging when the dimension of covariates is high. However, the computational burden can be greatly alleviated by using a proper weighing function for the transform parameter. We discuss weighting functions that avoid numerical integration. An example is the joint standard normal density function used by Hong et al. (2017) in testing strict stationarity. Finally, our test statistic is asymptotically pivotal and has a convenient asymptotic null $N(0, 1)$ distribution.

In Section 2, we introduce the hypotheses of interest and a novel testing approach based on the Fourier transform. In Section 3, we propose our test statistic. We derive the asymptotic distribution of our test statistic under the null hypothesis in Section 4 and derive its asymptotic power under global and local alternatives in Section 5. In Section 6, we discuss computation of our test statistic with choices of weighting functions that avoid numerical integration. Section 7 provides a simulation study on the finite sample performance of our test. Section 8 is an empirical application to testing structural changes in predictive regressions for equity returns. We conclude in Section 9. All proofs are relegated to the Appendices. Computer codes to implement our test are available from the authors upon request.

2. Hypotheses of interest and approach

Throughout the paper, i denotes an imaginary number such that $i=\sqrt{-1}$; A^* denotes the complex conjugate of $A \in \mathbb{C}$; Re(A) denotes the real part of A; and C denotes a generic bounded constant. Consider the following time-varying regression model

$$Y_t = g_t(X_t) + v_t,\ t = 1, 2, \ldots, T, \tag{1}$$

where Y_t is the dependent variable, X_t is a d-dimensional vector of covariates, v_t is an unobservable disturbance, $g_t(\cdot)$ is an unknown possibly time-varying function, and T is the sample size. We allow covariates X_t to be either exogenous or endogenous. If X_t is endogenous, then $E(v_t|X_t) \neq 0$. In that case, we assume that there exists a q-dimensional vector of instrumental variables Z_t, where $q \geq d$, such that $E(v_t|Z_t) = 0$, and Z_t and X_t are not independent. We allow both covariates and instrumental variables to be either continuous or discrete.

Our interest is in testing the null hypothesis of no structural change in $g_t(\cdot)$.

$$\mathbb{H}_0 : g_t(\cdot) = g_0(\cdot),\ \text{ for all } t = 1, 2, \ldots, T,$$

for some unknown time-invariant function $g_0(\cdot)$. The alternative hypothesis is

$$\mathbb{H}_A : g_t(\cdot) \neq g_0(\cdot),\ \text{ for some } t = 1, 2, \ldots, T,$$

for any unknown time-invariant function $g_0(\cdot)$. The subscript 0 for $g_0(X_t)$ means the function is constant with respect to time. It does not mean that $g_0(X_t)$ equals the functional form of $g_t(X_t)$ at $t = 0$.

Unlike the existing approaches that rely on nonparametric smoothed estimation of $g_t(X_t)$, we propose a novel approach based on the Fourier transform. Assuming that $\{X_t, Z_t\}_{t=1}^T$ are jointly strictly stationary, we consider the Fourier transform of Y_t using the instrumental variable Z_t:

$$E(Y_t e^{iu'Z_t}) = E\left[g_t(X_t)\, e^{iu'Z_t}\right] + E(v_t e^{iu'Z_t}). \tag{2}$$

Given $E(v_t|Z_t) = 0$, we have $E(v_t e^{iu'Z_t}) = 0$ for all $u \in \mathbb{R}^q$ (Bierens, 1982). Therefore, (2) can be rewritten as

$$\phi_t(u) \equiv E(Y_t e^{iu'Z_t}) = E\left[g_t(X_t)\, e^{iu'Z_t}\right]. \tag{3}$$

(3) reveals the link between the Fourier transform of data and the regression function $g_t(X_t)$. This device separates $g_t(X_t)$ from the error term v_t so that we can examine the instability of $g_t(X_t)$ using its complex-valued projection $\phi_t(u)$. Under the following lemma, we establish that $\phi_t(u)$ is time-invariant for any $u \in \mathbb{R}^q$ if and only if $\mathbb{H}_0$ holds.

Lemma 1. *$Pr[g_t(X_t) = g_0(X_t)] = 1$ for some time-invariant function $g_0(X_t)$ if and only if $\phi_t(u) = \phi_0(u)$ for all $u \in \mathbb{R}^q$ and for some time-invariant complex-valued function $\phi_0(u) \equiv E[g_0(X_t)e^{iu'Z_t}]$.*

Lemma 1 states that $g_t(\cdot)$ is time-invariant if and only if its Fourier transform $\phi_t(u)$ is time-invariant for all u. Therefore, we can rewrite the hypothesis of interest $\mathbb{H}_0$ versus $\mathbb{H}_A$ in the frequency domain

$$\mathbb{H}_0 : \phi_t(u) = \phi_0(u) \text{ for all } u \in \mathbb{R}^q, \tag{4}$$

where $\phi_0(u) \equiv E[g_0(X_t)e^{iu'Z_t}]$. The alternative hypothesis is

$$\mathbb{H}_A : \phi_t(u) \neq \phi_0(u) \text{ for all } u \in \Theta, \tag{5}$$

where $\Theta \subset \mathbb{R}^q$ in some set with positive Lebesgue measure. Intuitively, given that both X_t and Z_t are strictly stationary, the only source of instability of $\phi_t(u)$ will be structural changes in $g_t(X_t)$. Under $\mathbb{H}_0$, the Fourier transform should be a constant function of time for all u in the frequency domain. While under $\mathbb{H}_A$, it is a time-varying function for at least a set with positive Lebesgue measure in the frequency domain. Therefore, to test $\mathbb{H}_0$ against $\mathbb{H}_A$, we can estimate $\phi_t(u)$ and $\phi_0(u)$ consistently by $\hat{\phi}_t(u)$ and $\hat{\phi}_0(u)$ say, and then compare their differences for all $u \in \mathbb{R}^q$. Under $\mathbb{H}_0$, the difference between $\hat{\phi}_t(u)$ and $\hat{\phi}_0(u)$

will be close to 0 for all $u \in \mathbb{R}^q$ as the sample size $T \to \infty$. While under $\mathbb{H}_A$, they will converge to different limits. This idea is similar to Hausman's (1978) test, which compares two finite dimensional parameter estimators that converge to the same limit under the correct model specification but to different limits under model misspecification.

In the context of testing for structural changes, the main advantage of using the Fourier transform is that we avoid nonparametric estimation of unknown regression function $g_t(X_t)$. Even though we still need to estimate $\phi_t(u)$ via nonparametric estimation, it requires only smoothing over the univariate time index. In contrast, nonparametric estimation of $g_t(X_t)$ involves $(d+1)$-dimensional smoothing or even higher if endogeneity exists. The convergence rate of such a nonparametric estimator is rather slow, especially when the dimension of X_t or Z_t is high. It would have an adverse impact on the performance of related tests, at least in finite samples. Indeed, our test is asymptotically more powerful than Vogt's (2015) consistent test against $\mathbb{H}_A$; see Section 5 for more discussion. To ensure the consistency of our test against $\mathbb{H}_A$, we need to check $\mathbb{H}_0$ for all $u \in \mathbb{R}^q$, which requires integration over all $u \in \mathbb{R}^q$. However, we will show that by using certain weighting functions of u, we can avoid computationally expensive numerical integration. This makes the implementation of our test rather simple in practice.

A special case of our framework is when all regressors are exogenous, namely when $E(v_t|X_t) = 0$. All of the existing nonparametric tests for $\mathbb{H}_0$ are valid only under exogeneity. In our framework, we can simply replace Z_t by X_t because $E(v_t|X_t) = 0$ implies $E(v_t e^{iu'X_t}) = 0$ for all $u \in \mathbb{R}^d$. All asymptotic results for our test will remain the same as in the general case with endogeneity.

3. Nonparametric testing

3.1. Consistent estimation for $\phi_t(u)$

Given a time series sample $\{Y_t, X_t, Z_t\}_{t=1}^T$, we first construct consistent estimators for $\phi_t(u)$ and $\phi_0(u)$. Under $\mathbb{H}_0$, $\phi_t(u) = \phi_0(u)$, which is constant over time. We can estimate $\phi_0(u)$ consistently by its sample analog based on the whole sample:

$$\hat{\phi}_0(u) = \frac{1}{T}\sum_{t=1}^{T} Y_t e^{iu'Z_t}.$$

Throughout this paper, we assume

$$g_t(X_t) = g\left(X_t, \frac{t}{T}\right),$$

for some function $g(\cdot,\cdot) : \mathbb{R}^d \times [0,1] \to \mathbb{R}$. We require $g_t(X_t)$ be a function of normalized time ratio $\frac{t}{T}$, rather than t. This implies $\phi_t(u) = \phi(u, \frac{t}{T})$ for some complex-valued function $\phi(\cdot,\cdot)$. We use the local stationarity (e.g., Dahlhaus, 1996) to characterize the instability of the relationship between X_t and Y_t caused by smooth structural changes. The reason that we rescale time t to a finer grid on $[0, 1]$ is to ensure that the amount of data increases around each time point t as the sample size $T \to \infty$. This is necessary for the consistency of $\hat{\phi}_t(u)$. Under $\mathbb{H}_A$, we assume that for each u, $\phi(u, \frac{t}{T})$ is a twice continuously differentiable function of $\frac{t}{T}$ except for a set of a finite number of points on $[0, 1]$. This set of points represents abrupt structural breaks in $g(X_t, \frac{t}{T})$, with unknown break dates. Our test can thus detect instability caused by both abrupt breaks and smooth changes; see Chen and Hong (2012) for detailed discussion.

Next, we construct a consistent nonparametric estimator for $\phi_t(u)$. Consider the following pseudo regression:

$$Y_t e^{iu'Z_t} = \phi_t(u) + \varepsilon_t(u), \tag{6}$$

where $\phi_t(u) = E(Y_t e^{iu'Z_t})$ can be regarded as the mean of the dependent variable $Y_t e^{iu'Z_t}$ at each time t, and $\varepsilon_t(u) = Y_t e^{iu'Z_t} - \phi_t(u)$ is a generalized disturbance with $E[\varepsilon_t(u)] = 0$ for all $u \in \mathbb{R}^q$. From (6), we show that $\phi_t(u)$ can be viewed as a pseudo conditional mean function of $Y_t e^{iu'Z_t}$ on time t. Therefore, we can estimate it consistently by the Nadaraya–Watson (NW) estimator or the local linear (LL) estimator.

The LL estimator has an advantage over the NW estimator in the boundary region under $\mathbb{H}_A$, because the bias of the LL smoothing in the boundary region is of the same order of magnitude as that in the interior region. However, the scales in the boundary and the interior regions are different. Let $K_h(z) = h^{-1}K(\frac{z}{h})$, where $K : [-1, 1] \to \mathbb{R}^+$ is a kernel function and $h \equiv h(T)$ is a bandwidth such that $h \to 0$ and $Th \to \infty$ as $T \to \infty$. As shown in Cai (2007), the asymptotic bias of the LL estimator in the boundary region is proportional to $h^2 b(c)$ where $b(c) = \frac{\mu_{2c}^2 - \mu_{1c}\mu_{2c}}{\mu_{0c}\mu_{2c} - \mu_{1c}^2}$, $\mu_{jc} = \int_{-c}^{-1} \eta^j K(\eta)d\eta$, for $j = 1, 2, 3$, and $c \in [0, 1]$, rather than $h^2 \int_{-1}^{1} \eta^2 K(\eta)d\eta$ in the interior region. Moreover, the asymptotic variance in the boundary region is larger than that in the interior region, which would further complicate the form of our test statistic. We want to avoid trimming the data since the boundary regions $[1, \lfloor Th \rfloor] \cup [T - \lfloor Th \rfloor, T]$ contain abundant information in finite samples, where $\lfloor Th \rfloor$ denotes the integer part of Th. Therefore, we follow Chen and Hong (2012) to use the reflection method which is first introduced by Hall and Wehrly (1991) to construct pseudo data $(Y_t, X_t) = (Y_{-t}, X_{-t})$ for $t \in [1 - \lfloor Th \rfloor, 0]$ and $(Y_t, X_t) = (Y_{2T-t}, X_{2T-t})$ for $t \in [T + 1, T + \lfloor Th \rfloor]$. We then use the union of the original data and the pseudo data to construct our test statistic. This method ensures that the bias in the boundary region is of the same form as that in the interior region because we have symmetric coverage of observations in the boundary regions when the augmented data is used.

Specifically, the NW estimator for $\phi_t(u)$ with the reflection method is defined as

$$\hat{\phi}_t^{NW}(u) = \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} Y_s e^{iu'Z_s} H_{st}^{NW},$$

where $H_{st}^{NW} = \frac{K_h(\frac{s-t}{T})}{\sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} K_h(\frac{s-t}{T})}$.

Alternatively, we can estimate $\phi_t(u)$ using the LL estimator $\hat{\phi}_t^{LL}(u)$. Let $\beta_T(u,t) = [\beta_{T,1}(u,t), \beta_{T,2}(u,t)]'$ be a 2×1 complex-valued vector such that $\beta_{T,1}(u,t) = \phi_t(u)$, and $\beta_{T,2}(u,t) = \frac{\partial \phi_t(u)}{\partial(\frac{t}{T})}$. By the Taylor expansion, for any fixed point $\frac{t}{T} \in [0, 1]$, we have

$$\phi_s(u) = \beta_{T,1}(u,t) + \beta_{T,2}(u,t)\left(\frac{s-t}{T}\right) + O\left[\left(\frac{s-t}{T}\right)^2\right].$$

If $\hat{\beta}_T(u,t)$ is the LL estimator, then

$$\begin{bmatrix} \hat{\beta}_{T,1}(u,t) \\ \hat{\beta}_{T,2}(u,t) \end{bmatrix} = \underset{\beta_T(u,t)\in\mathbb{C}^2}{\operatorname{argmin}} \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \left| Y_s e^{iu'Z_s} - \beta_{T,1}(u,t) - \beta_{T,2}(u,t)\left(\frac{s-t}{T}\right)\right|^2 K_h\left(\frac{s-t}{T}\right). \tag{7}$$

It is straightforward to show that the solution to (7) is given by

$$\begin{bmatrix} \hat{\beta}_{T,1}(u,t) \\ \hat{\beta}_{T,2}(u,t) \end{bmatrix} = \begin{bmatrix} S_{T,0}(u,t) & S_{T,1}(u,t) \\ S_{T,1}(u,t) & S_{T,2}(u,t) \end{bmatrix}^{-1} \begin{bmatrix} \Gamma_{T,0}(u,t) \\ \Gamma_{T,1}(u,t) \end{bmatrix},$$

where $S_{T,j}(u,t) = \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \left(\frac{s-t}{T}\right)^j K_h(\frac{s-t}{T})$, $\Gamma_{T,j}(u,t) = \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \left(\frac{s-t}{T}\right)^j K_h(\frac{s-t}{T}) Y_s e^{iu'Z_s}$. The LL estimator $\hat{\phi}_t^{LL}(u)$ is equal to the intercept estimator $\hat{\beta}_{T,1}(u,t)$ and it has an analytical solution

$$\hat{\phi}_t^{LL}(u) = \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} Y_s e^{iu'Z_s} H_{st}^{LL},$$

where $H_{st}^{LL} = \frac{[S_{T,2}(u,t) - S_{T,1}(u,t)(\frac{s-t}{T})] K_h(\frac{s-t}{T})}{S_{T,0}(u,t) S_{T,2}(u,t) - S_{T,1}(u,t)^2}$.

Both the NW and LL estimators are simply a weighted sum of $Y_s e^{iu'Z_s}$ with different weights. They are essentially asymptotically equivalent to each other since

$$\hat{\phi}_t^{LL}(u) = \frac{1}{T} \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} Y_s e^{iu'Z_s} K_h\left(\frac{s-t}{T}\right)[1 + op(1)],$$

$$\hat{\phi}_t^{NW}(u) = \frac{1}{T} \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} Y_s e^{iu'Z_s} K_h\left(\frac{s-t}{T}\right)[1 + op(1)].$$

Therefore, for notational simplicity we suppress the superscript indicating the type of weighting and write the nonparametric estimator for $\phi_t(u)$ as follows:

$$\hat{\phi}_t(u) = \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} Y_s e^{iu'Z_s} H_{st}. \tag{8}$$

3.2. Test statistic

Under $\mathbb{H}_0$, both $\hat{\phi}_0(u)$ and $\hat{\phi}_t(u)$ can estimate the Fourier transform of data consistently. Therefore their difference will converge to 0 for all u as the sample size grows. On the contrary, if a structural change exists, $\hat{\phi}_t(u)$ still consistently estimates the Fourier transform while $\hat{\phi}_0(u)$ does not. Then they will converge to different limits for some $u \in \Theta$. Therefore, we propose the following test statistic based on the quadratic distance between $\hat{\phi}_t(u)$ and $\hat{\phi}_0(u)$:

$$\hat{Q} = \frac{1}{T} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \hat{\phi}_t(u) - \hat{\phi}_0(u) \right|^2 W(u) du, \tag{9}$$

where $W(\cdot) : \mathbb{R}^q \to \mathbb{R}^+$ is a nonnegative, continuous, and symmetric weighting function for u. Unlike Su and Xiao (2008), the choice of the weighting function does not have any impact on the asymptotic distribution or consistency of our test.

However, the choice of weighting functions will affect the computation of the test statistic since the numerical integration over $\mathbb{R}^q$ is needed. Fortunately, certain classes of $W(\cdot)$ can avoid the numerical integration (see Section 6 for more discussion).

To show the asymptotic distribution of our test statistic $\hat{Q}$, we introduce a standardized version

$$\widehat{SQ} = \left(Th^{1/2}\hat{Q} - \hat{B}\right)/\sqrt{\hat{V}},$$

where

$$\hat{B} = h^{-1/2}\int_{\mathbb{R}^q} |\hat{\Omega}(u,u)| W(u)du \int K^2(\eta)d\eta,$$

and

$$\hat{V} = 2\int_{\mathbb{R}^{2q}} |\hat{\Omega}(u,v)|^2 W(u)W(v)dudv \int \left[\int K(\eta)K(\eta+\lambda)d\eta\right]^2 d\lambda$$

are consistent estimators for the asymptotic mean and asymptotic variance of $Th^{1/2}\hat{Q}$. Here $\hat{\Omega}(u,v)$ is a consistent estimator for the generalized long-run variance

$$\Omega(u,v) = \sum_{j=-\infty}^{\infty} E[\varepsilon_t(u)\varepsilon_{t-j}(v)^*],$$

where $\varepsilon_{t-j}(v)^*$ is the complex conjugate of $\varepsilon_{t-j}(v) \equiv Y_{t-j}e^{iv'Z_{t-j}} - \phi_{t-j}(v)$. We need to consider the generalized long-run variance in both the asymptotic mean and asymptotic variance because the generalized disturbance $\varepsilon_t(u)$ is generally neither an i.i.d. nor a martingale difference sequence. That implies $corr\,[\varepsilon_t(u)\varepsilon_s(u)^*] = E[Y_tY_se^{iu'(Z_t-Z_s)}] - \phi_t(u)\phi_s(u)^* \neq 0$ under both $\mathbb{H}_0$ and $\mathbb{H}_A$. Because the nonparametric smoothing is on the rescaled time t/T, the observations that fall into each small neighborhood via the kernel function are close to each other in terms of time index. Therefore, we need to deal with the possible strong dependence under both $\mathbb{H}_0$ and $\mathbb{H}_A$.

Following Newey and West (1987) and Andrews (1991), we can use the following consistent estimator for $\Omega(u,v)$:

$$\hat{\Omega}(u,v) = \sum_{j=-p_T}^{p_T} k_l\left(\frac{j}{p_T}\right)\hat{\sigma}_j(u,v),$$

where the jth order sample generalized covariance function is

$$\hat{\sigma}_j(u,v) = \begin{cases} \frac{1}{T}\sum_{t=j+1}^{T}\hat{\varepsilon}_t(u)\hat{\varepsilon}_{t-j}(v)^*, & j = 0,1,\ldots,p_T, \\ \frac{1}{T}\sum_{t=1-j}^{T}\hat{\varepsilon}_{t+j}(u)\hat{\varepsilon}_t(v)^*, & j = -1,-2,\ldots,-p_T, \end{cases}$$

and the lag order $p_T = p(T)$ satisfies $p_Th^{1/2} \to \infty$, $p_T/(Th) \to 0$ as $T \to \infty$. $k(\cdot) : [-1,1] \to \mathbb{R}^+$ is a symmetric, bounded, and square-integrable kernel for the lag order j with $k(0) = 1$. $\hat{\varepsilon}_t(u) = Y_te^{iu'Z_t} - \hat{\phi}_t(u)$ is a consistent estimator for $\varepsilon_t(u)$, and $\hat{\varepsilon}_t(u)^*$ is its complex conjugate.

4. Asymptotic distribution

To derive the asymptotic distribution of $\widehat{SQ}$ under $\mathbb{H}_0$, we impose the following conditions.

Assumption 1. (i) $\{X_t', Z_t'\}'$ is a $(d+q)\times 1$ strictly stationary β-mixing process, where the mixing coefficient $\beta(j)$ satisfies $\sum_{j=1}^{\infty} j^2\beta(j)^{\delta/(1+\delta)} < C$ for some $0 < \delta < 1$; (ii) v_t is a weakly stationarity process such that $\sup_t[E(v_t^4)] < C$; and (iii) $\sup_t|g_t(X_t)| < C(X_t)$ with $E[C(X_t)^4] < \infty$.

Assumption 2. The instrument vector Z_t satisfies: (i) $E(v_t|Z_t) = 0$ almost surely, and (ii) $F_{XZ}(x,z) \neq F_X(x)F_Z(z)$, where $F_{XZ}(\cdot)$, $F_X(\cdot)$ and $F_Z(\cdot)$ denote the joint and marginal CDFs for X_t and Z_t.

Assumption 3. $K(\cdot) : [-1,1] \to \mathbb{R}^+$ is a symmetric, bounded and twice continuously differentiable kernel function with $\int_{-1}^{1} K(u)du = 1$, $\int_{-1}^{1} uK(u)du = 0$, and $\int_{-1}^{1} u^2K(u)du = C_k < \infty$.

Assumption 4. $W(\cdot) : \mathbb{R}^q \to \mathbb{R}^+$ is a nonnegative, continuous, symmetric, and integrable weighting function with $\int_{\mathbb{R}^q} |u|_q^4 W(u)du < \infty$.

Assumption 5. $k(\cdot) : [-1,1] \to \mathbb{R}^+$ is a symmetric, bounded, and square-integrable kernel function with $k(0) = 1$.

Assumption 6. (i) The bandwidth $h = c_hT^{-\tau_h}$ for $0 < \tau_h < 2/3$ and $0 < c_h < \infty$; and (ii) the lag order $p_T = c_pT^{\tau_p}$ for $\frac{\tau_h}{2} < \tau_p < 1 - \tau_h$ and $0 < c_p < \infty$.

Assumption 1(i) restricts $\{X_t', Z_t'\}'$ to be jointly strictly stationary. We note that our test is not robust to potential structural changes in the marginal distributions of the regressors or instruments. It is possible to relax this assumption by imposing conditions on structural changes in $\{X_t', Z_t'\}$ such that they change slower than the variation of $g_t(X_t)$ over time. Within a small neighborhood where X_t or Z_t is nearly a stationary process, we can compare a localized weighted estimator and an unweighted estimator to check whether they converge to the same probability limit. That is the same idea as proposed in this paper. Even though our novel Fourier transform approach makes our test less applicable in scope than Su and Xiao (2008) and Vogt (2015), its gains are quite substantial. The most important one is that we can circumvent the "curse of dimensionality" problem. Furthermore, unlike the existing literature, our approach allows for discrete and endogenous covariates. The β-mixing condition is common in time series analysis and is adopted in (e.g.) Hjellvik et al. (1998), Chen and Hong (2012), Wang and Hong (2017). Assumption 1(ii) is a moment condition on v_t and it allows v_t to have time-varying higher moments. Assumption 1(iii) implies that the unknown function $g_t(\cdot)$ is square-integrable, which guarantees the Fourier transform exists.

In Assumption 2, $E(v_t|Z_t) = 0$ characterizes the validity of the instruments. And, the mutual dependence between X_t and Z_t ensures the relevance of the instruments.

Assumption 3 provides conditions on the kernel function $K(\cdot)$ for nonparametric estimation of $\phi(u, \frac{t}{T})$. It includes but does not restrict to the Epanechnikov and Quartic kernels. Unlike the existing approaches, we only require second-order kernels.

Assumption 4 provides regularity conditions on the weighting function $W(\cdot)$. It guarantees the consistency of our test. Although the choice of $W(\cdot)$ does not have an impact on the asymptotic distribution of the test statistic, it does affect the computation in finite samples. Some suitable choices of $W(\cdot)$ can avoid the numerical integration, which is rather appealing in practice, especially when the dimension of Z_t is high.

Assumption 5 provides conditions on the kernel function $k(\cdot)$ for estimating the generalized long-run variance. Kernels such as the Bartlett and Quadratic-Spectral kernels satisfy this condition. We note that our asymptotic result is derived via the generalized error term $\varepsilon_t(u)$ rather than the regression error v_t, where $\varepsilon_t(u) \equiv Y_t e^{iu'Z_t} - \phi_t(u)$ is serially correlated under both $\mathbb{H}_0$ and $\mathbb{H}_A$. So we need to introduce the generalized long-run variance.

Assumption 6 provides conditions on smoothing parameters. It implies that $h \to 0$, $p_T h^{1/2} \to \infty$, and $p_T/(Th) \to 0$, as $T \to \infty$. A data-dependent method to choose an optimal h has been discussed in Robinson (1989). However, such method may affect the size of the test in finite samples because more noise is introduced. As suggested by Chen and Hong (2012), we can use a simple rule-of-thumb to choose h, namely $h = (1/\sqrt{12})T^{-1/5}$, where $1/\sqrt{12}$ is the standard deviation of $U[0, 1]$, because we can approximate the grid points $\{t/T : 1, 2, \ldots, T\}$ by $U[0, 1]$ as $T \to \infty$.

Next we state the asymptotic distribution of $\widehat{SQ}$ under $\mathbb{H}_0$.

Theorem 1. *Suppose Assumptions 1–6 hold. Then under $\mathbb{H}_0$, $\widehat{SQ} \xrightarrow{d} N(0,1)$ as $T \to \infty$.*

The test statistic $\widehat{SQ}$ has a convenient asymptotic null $N(0, 1)$ distribution, which is asymptotically pivotal. Because $\widehat{SQ}$ diverges to $+\infty$ as $T \to \infty$ under $\mathbb{H}_A$ (see Theorem 2), negative values of $\widehat{SQ}$ can occur only under $\mathbb{H}_0$. As a result, one can use one-sided $N(0, 1)$ critical values. For example, the asymptotic critical value at the 5% significance level is 1.645.

Nevertheless, the asymptotic distribution may not provide an accurate approximation in finite samples, especially when the dependence among observation is strong. We can use bootstrap to obtain a better approximation in finite samples. In Section 7, we will examine the finite sample performance of our test using both asymptotic and bootstrapped critical values.

5. Asymptotic power

To study the asymptotic power of the $\hat{Q}$ test, we assume the following condition:

Assumption 7. (i) $g(X_t, \tau) : \mathbb{R}^d \times [0, 1] \to \mathbb{R}$ is a twice continuously differentiable function of τ except for a set of a finite number of points on $[0, 1]$; and (ii) $r(z, \tau) \equiv E[g(X_t, \tau)|Z_t = z]$ is some unknown function which is square-integrable with respect to z and twice continuously differentiable with respect to τ except for a set of a finite number of points on $[0, 1]$, such that $\frac{\partial r(\cdot,\tau)}{\partial \tau} \neq 0$ on a non-zero Borel measurable set.

Assumption 7(i) allows for both smooth structural changes and abrupt structural breaks with unknown breakpoints. For abrupt structural breaks, we require the sizes of breaks are bounded by a stochastic upper bound (see Assumption 1). Assumption 7(ii) is a restriction on the instrumental variable Z_t. It excludes the case where Z_t is orthogonal to the direction where $g_t(X_t)$ has structural changes. For instance, let $g_t(X_t) = X_t^2 + \mathbf{1}(t \leq 0.5T)X_t$ and $E(X_t) = 0$. Then a valid instrument should not only satisfy Assumption 2 but also should be linearly dependent on X_t, i.e., $E(X_t|Z_t) \neq 0$. Because, if $E(X_t|Z_t) = 0$, then the Fourier transform $\phi_t(u)$ will be $\phi_t(u) = E(X_t^2 e^{iu'Z_t}) = \phi_0(u)$, which is a constant function with respect to time. Assumption 7(ii) is necessary for Lemma 1. It requires that the instrumental variable can capture the change of $g_t(X_t)$ over time if a structural change exists.

Theorem 2. *Suppose Assumptions 1–7 hold. Then for any sequence of non-stochastic constants $\{M_T = o(T\sqrt{h})\}$, $Pr(\widehat{SQ} > M_T) \to 1$ under $\mathbb{H}_A$ as $T \to \infty$.*

Theorem 2 implies that $\widehat{SQ}$ is consistent against the alternative hypothesis $\mathbb{H}_A$ at any significance level. Sharing the same merits as Chen and Hong's (2012) test, our test can detect any structural changes subject to Assumption 7, without having to know any prior information about the types of structural changes and the number of breakpoints. Furthermore, we do not restrict the regression function to be a linear model as is considered in Chen and Hong (2012).

To gain additional insight into our test, we consider a class of local smooth structural changes:

$$\mathbb{H}_{A1} : g_t\left(X_t\right) = g_0(X_t) + \kappa_T l_1\left(X_t, \frac{t}{T}\right),$$

where $l_1(X_t, \tau) : \mathbb{R}^d \times [0, 1] \to \mathbb{R}$ is twice continuously differentiable with respect to τ. The nonstochastic factor $\kappa_T \to 0$ as $T \to \infty$ characterizes the speed at which the departure $g_t(X_t) - g_0(X_t) = \kappa_T l_1(X_t, \frac{t}{T})$ converges to 0 as $T \to \infty$. Under $\mathbb{H}_{A1}$, we have

$$|\phi_t\left(u\right) - \phi_0(u)|^2 = \kappa_T^2 \left| E\left[l_1\left(X_t, \frac{t}{T}\right) e^{iu'Z_t}\right]\right|^2.$$

Theorem 3. *Suppose Assumptions 1–7 hold. Then under $\mathbb{H}_{A1}$ with $\kappa_T = T^{-1/2}h^{-1/4}$,*

$$Pr\left[\widehat{SQ} \geq z_\alpha \middle| \mathbb{H}_{A1}\right] \to 1 - \Phi\left(z_\alpha - \frac{\gamma_1}{\sqrt{V}}\right)$$

as $T \to \infty$, where $\Phi(\cdot)$ is the $N(0, 1)$ CDF, z_α is the one-sided critical value of $N(0, 1)$ at significance level α,

$$V = 2\int_{\mathbb{R}^{2q}} |\Omega(u, v)|^2 W(u)W(v)dudv \int \left[\int K(\eta)K(\eta + \lambda)d\eta\right]^2 d\lambda,$$

$\Omega(u, v) = \sum_{j=-\infty}^{\infty} E[\varepsilon_t(u)\varepsilon_{t+j}(u)]$, and

$$\gamma_1 = \int_{\mathbb{R}^q} \left\{ \int_0^1 \left| E\left[l_1(X_t, \tau) e^{iu'Z_t}\right]\right|^2 d\tau - \left| \int_0^1 E\left[l_1(X_t, \tau) e^{iu'Z_t}\right] d\tau \right|^2 \right\} W(u)du < \infty.$$

Theorem 3 implies that $\widehat{SQ}$ has nontrivial power under the class of local alternatives $\mathbb{H}_{A1}$ with a rate of $\kappa_T = T^{-1/2}h^{-1/4}$, which is slightly slower than the parametric rate $T^{-1/2}$. For example, if the bandwidth $h \propto T^{-1/5}$, then the rate $\kappa_T = T^{-9/20}$ is rather close to the parametric rate $T^{-1/2}$. We emphasize that the rate $\kappa_T = T^{-1/2}h^{-1/4}$ is not affected by the dimension of instruments Z_t or covariates X_t, thus free of the "curse of dimensionality" problem. In contrast, Vogt's (2015) consistent test involves smoothing of dimension $d + 1$ and has nontrivial asymptotic power under the class of local alternatives $\mathbb{H}_{A1}$ with a rate of $\kappa_T = T^{-1/2}h^{-(d+1)/4}$, which is slower than the rate of $\kappa_T = T^{-1/2}h^{-1/4}$, especially for a large d. Moreover, Vogt (2015) only considers testing structural changes in regression with exogenous covariates.

We note that the nonparametric tests of Hidalgo (1995) and Su and Xiao (2008) can detect the class of local alternatives $\mathbb{H}_{A1}$ with the parametric rate $\kappa_T = T^{-1/2}$. However, these tests are not consistent for all departures $l_1(X_t, \frac{t}{T})$. For example, Hidalgo's (1995) test has no power against $\mathbb{H}_{A1}$ when $E[\varsigma(X_t)f^2(X_t)l_1(X_t, \frac{t}{T})] = 0$, where $\varsigma(X_t) = \int g(X_t, \tau)d\tau$ and $f(X_t)$ is the density function of X_t. Su and Xiao's (2008) tests have no power against $\mathbb{H}_{A1}$ when $E[\omega(X_t)f(X_t)h_t(X_t)] = 0$, where $\omega(X_t)$ is a weighting function chosen by practitioners. In contrast, our $\widehat{SQ}$ test is consistent for all departures $l_1(X_t, \frac{t}{T})$.

Suppose we consider a class of nonsmooth local alternatives at some given point $\tau_0 \in [0, 1]$:

$$\mathbb{H}_{A2} : g_t\left(X_t\right) = g_0(X_t) + a_T l_2\left(X_t, \frac{t/T - \tau_0}{b_T}\right),$$

where τ_0 is a fixed point in $[0, 1]$, $l_2(X_t, \xi) : \mathbb{R}^{d+1} \to \mathbb{R}$ is twice continuously differentiable with respect to τ. We let $a_T = a(T) \to 0$, and $b_T = b(T) \to 0$ as $T \to \infty$. Under $\mathbb{H}_{A2}$, $g_t(X_t)$ has a nonsmooth spike at point τ_0 as $T \to \infty$. The rate $b_T \to 0$ controls the sharpness of the structural change around τ_0, and $a_T \to 0$ is the speed at which the departure of $g_t(X_t)$ from $g_0(X_t)$ at each point t/T vanishes to 0 as $T \to \infty$. This type of alternatives was first studied by Rosenblatt (1975) in a different context. Chen and Hong (2012) also consider this class of local alternatives in testing parameter instability in a linear time series regression model.

Theorem 4. *Suppose Assumptions 1–7 hold. Then under $\mathbb{H}_{A2}$ with $a_T \to 0$, $b_T \to 0$, $a_T^2 b_T = T^{-1}h^{-1/2}$, $b_T = h^{\tau_b}$, and $1/4 < \tau_b < 1$, we have*

$$Pr\left[\widehat{SQ} \geq z_\alpha \middle| \mathbb{H}_{A2}\right] \to 1 - \Phi\left(z_\alpha - \frac{\gamma_2}{\sqrt{V}}\right)$$

as $T \to \infty$, where $\Phi(\cdot)$ is the $N(0, 1)$ CDF, z_α is the one-sided critical value of $N(0, 1)$ at significance level α,

$$V = 2\int_{\mathbb{R}^{2q}} |\Omega(u, v)|^2 W(u)W(v)dudv \int \left[\int K(\eta)K(\eta + \lambda)d\eta\right]^2 d\lambda,$$

$\Omega(u, v) = \sum_{j=-\infty}^{\infty} E[\varepsilon_t(u)\varepsilon_{t+j}(u)]$, and

$$\gamma_2 = \int_{\mathbb{R}^q}\int_{\mathbb{R}} \left| E\left[l_2(X_t, \xi) e^{iu' Z_t} \right] \right|^2 d\xi W(u) du < \infty.$$

Theorem 4 implies that our test has nontrivial asymptotic power under the class of nonsmooth local alternatives $\mathbb{H}_{A2}$ with $a_T^2 b_T = T^{-1}h^{-1/2}$. In contrast, the nonparametric tests of Hidalgo (1995) and Su and Xiao (2008) may have no power against this class of nonsmooth local alternatives, because it can be shown that these tests can only detect $\mathbb{H}_{A2}$ with $a_T b_T = T^{-1/2}$ or slower. If $a_T = T^{-7/20}(\ln\ln T)^{-\epsilon/2}$, $b_T = T^{-1/5}(\ln\ln T)^{\epsilon}$, $h = T^{-1/5}$, where $\epsilon > 0$ is a constant, then we have $a_T^2 b_T = T^{-1}h^{-1/2}$ but $a_T b_T = o(T^{-1/2})$. In this case, our test has power against $\mathbb{H}_{A2}$ but the tests of Hidalgo (1995) and Su and Xiao (2008) have no power. Also see Chen and Hong (2012, footnote 12, p. 1168) for details.

6. Weighting function

We now discuss the choice of weighting function $W(u)$. Theorem 2 implies that our test is always consistent against $\mathbb{H}_A$ provided a nonnegative, continuous, and integrable weighting function with unbounded support is used. However, the choice of $W(u)$ affects the computation of the test statistic $\widehat{SQ}$, which involves q-dimensional integration. In this section, we will discuss two types of weighting functions that can avoid numerical integration. It is quite attractive in practice, especially when q is large. The first type of weighting function is the joint independent normal probability density function (pdf):

$$W_G(u) = \prod_{k=1}^{q} \frac{1}{\sqrt{2\pi\xi_k^2}} e^{-\frac{u_k^2}{2\xi_k^2}},$$

where ξ_k is the standard deviation for each marginal pdf. A larger ξ_k implies that more weights are given to the values of u that are distant from 0 for dimension k. For each k, we have the following identity:

$$\int_{\mathbb{R}} cos(u_k z_k) \frac{1}{\sqrt{2\pi\xi_k^2}} e^{-\frac{u_k^2}{2\xi_k^2}} du_k = e^{-\frac{z_k^2\xi_k^2}{2}}. \tag{10}$$

Alternatively, we can also use another type of weighting function based on the Laplace pdf:

$$W_L(u) = \prod_{k=1}^{q} \frac{1}{2\lambda_k} e^{-\frac{|u_k|}{\lambda_k}},$$

where the λ_k's are scale parameters for each marginal pdf. For each k, we have the following identity:

$$\int_{\mathbb{R}} cos(u_k z_k) \frac{1}{2\lambda_k} e^{-\frac{|u_k|}{\lambda_k}} du_k = \frac{1}{1+\lambda_k^2 z_k^2}. \tag{11}$$

Let $\mathbb{S}_1 = \{-\lfloor Th \rfloor, -\lfloor Th \rfloor + 1, \ldots, T + \lfloor Th \rfloor\}$, and $\mathbb{S}_2 = \{1, 2, \ldots, T\}$, then by (10) and (11), (9) can be written as

$$\hat{Q} = \frac{1}{T}\sum_{t=1}^{T}\left[\sum_{r,s\in\mathbb{S}_1} A_{sr} H_{st} H_{rt} + \frac{1}{T^2}\sum_{r,s\in\mathbb{S}_2} A_{sr} - \frac{2}{T}\sum_{s\in\mathbb{S}_1, r\in\mathbb{S}_2} A_{sr} H_{st} \right], \tag{12}$$

where

$$A_{sr} = \begin{cases} Y_s Y_r e^{-\frac{\sum_{k=1}^{q}\xi_k^2 (Z_{sk}-Z_{rk})^2}{2}}, & \text{if } W(u) = W_G(u), \\ Y_s Y_r \prod_{k=1}^{q} \frac{1}{1+\lambda_k^2 (Z_{sk}-Z_{rk})^2}, & \text{if } W(u) = W_L(u). \end{cases}$$

Furthermore, the standardized test statistic $\widehat{SQ}$ becomes

$$\widehat{SQ} \equiv \left(Th^{1/2}\hat{Q} - \hat{B} \right) / \sqrt{\hat{V}}, \tag{13}$$

where

$$\hat{B} = h^{-1/2} \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right) \hat{\sigma}(j) \int K^2(\eta) d\eta,$$

$$\hat{V} = 2 \sum_{j,k=-p_T}^{p_T} k\left(\frac{j}{p_T}\right) k\left(\frac{l}{p_T}\right) \hat{\sigma}(j,l) \int \left[\int K(\eta)K(\eta+\lambda) d\eta \right]^2 d\lambda,$$

with

$$\hat{\sigma}(j) = \frac{1}{T}\sum_{s=1+|j|}^{T} \hat{\zeta}(s, s-j),$$

$$\hat{\sigma}(j, l) = \frac{1}{T^2}\sum_{s=1+|j|}^{T}\sum_{r=1+|l|}^{T} \hat{\zeta}(s, r)\hat{\zeta}(s-j, r-l),$$

where $K_{sr} = \frac{1}{Th}K(\frac{s-r}{Th})$, and

$$\hat{\zeta}(j, k) = A_{j,k} - \frac{1}{T}\sum_{l=1}^{T} A_{jl} - \frac{1}{T}\sum_{l=1}^{T} A_{kl} + \frac{1}{T^2}\sum_{l_1,l_2}^{T} A_{l_1 l_2}.$$

Based on the above results, the test statistic can be computed without numerical integration, regardless of the dimension of Z_t.

We note that Székely et al. (2007) propose the following weighting function:

$$W_D(u) = \frac{1}{c_q |u|_q^{q+1}},$$

where $c_q = \frac{\pi^{(1+q)/2}}{\Gamma((1+q)/2)}$, and $\Gamma(\cdot)$ is the complete gamma function. This weighting function can also avoid numerical integration in characteristic function-based tests. However, $W_D(u)$ is not integrable and does not satisfy Assumption 4. So we do not use it in our test.

7. Simulation studies

In this section, we study the finite sample performance of our test in comparison with Su and Xiao's (2008) nonparametric CUSUM-type tests. We consider the cases that covariates are exogenous and endogenous separately.

7.1. The exogeneity case

To study the size performance, we consider the following DGPs:

$$\begin{aligned}
&\text{DGP S.1:}\ \ Y_t = 1 + 1.5X_{1t} + v_t;\\
&\text{DGP S.2:}\ \ Y_t = 1 + 1.5X_{1t} + X_{1t}^2 + v_t\\
&\text{DGP S.3:}\ \ Y_t = 1 + 0.5X_{1t} - 1.5X_{2t} + X_{3t}^2 + v_t;\\
&\text{DGP S.4:}\ \ Y_t = 1 + 0.5X_{1t} + 2X_{4t} + v_t;\\
&\text{DGP S.5:}\ \ Y_t = 1 + 0.5X_{1t}X_{5t} + v_t;
\end{aligned}$$

where

$$\begin{aligned}
X_{1t} &= 0.5X_{1(t-1)} + \nu_{1t};\\
X_{2t} &= 1 - 0.5X_{2(t-1)} + \nu_{2t};\\
X_{3t} &= 0.4X_{1(t-1)} + \nu_{3t};\\
X_{4t} &= \begin{cases} 0 & \text{with } Pr = 0.3,\\ 1 & \text{with } Pr = 0.7;\end{cases}\\
X_{5t} &= \begin{cases} 1 & \text{with } Pr = 0.4,\\ 2 & \text{with } Pr = 0.6;\end{cases}\\
\nu_{jt} &\sim i.i.d.N(0, 1) \quad \text{for } j = 1, 2, 3;\\
v_t &\sim i.i.d.N(0, 1).
\end{aligned}$$

DGPs S.1–S.5 are all time invariant. Under DGP S.1, $g_t(X_t)$ is a linear function. Under DGP S.2, $g_t(X_t)$ is nonlinear. Under DGP S.3, $g_t(X_t)$ is a nonlinear function with a relatively high dimensional regressor ($d = 4$). Under DGPs S.4 and S.5, a discrete covariate exists. It is of an additive form under DGP S.4 and a multiplicative form under DGP S.5. We note that Su and Xiao's (2008) tests are no longer applicable to discrete covariates.

To examine the power performance, we consider the following DGPs:

DGP P.1 [Single Structural Break]:

$$Y_t = \begin{cases} 1.5 + X_{1t} + v_t & \text{if } t \le 0.3T,\\ 1.5 - 2X_{1t} + v_t & \text{otherwise;}\end{cases}$$

DGP P.2 [Multiple Structural Breaks]:

$$Y_t = \begin{cases} 1.5 + X_{1t} + v_t & \text{if } 0.1T < t < 0.3T, \\ 1.2 - X_{1t}^2 + 2sin(X_{1t}) + v_t & \text{if } 0.3T \le t \le 0.7T, \\ -2X_{1t} + v_t & \text{otherwise;} \end{cases}$$

DGP P.3 [Smooth Structural Change]:

$$Y_t = \theta(\tau)(1 + 0.5X_{1t}) + v_t,$$

where $\theta(\tau) = 1.5\tau - e^{-(\tau-0.5)^2}$, and $\tau = t/T$;

DGP P.4 [Smooth Structural Change]:

$$Y_t = \theta_1(\tau)X_{1t} + e^{\tau X_{1t}} - \theta_2(\tau)X_{1t}^2 + v_t,$$

where $\theta_1(\tau) = 0.2e^{-0.7+3.5\tau}$, $\theta_2(\tau) = 2\tau + e^{-16(\tau-0.5)^2} - 1$, and $\tau = t/T$.

Under DGP P.1, there exists a single structural break at $t = 0.3T$. Under DGP P.2, there exist two structural breaks at $t = 0.3T$ and $t = 0.7T$. DGP P.3 is a linear time series model with smooth structural changes in the regression coefficient $\theta(\tau)$. And DGP P.4 is a nonlinear time series model with smooth structural changes.

For each DGP, we simulate 1000 data sets with the sample size $T = 100$ and 200 respectively. We choose the joint $N(0, 1)$ density function for $W(u)$, and the Epanechnikov kernel for $K(\cdot)$. For the choice of bandwidth, we follow Chen and Hong (2012) to use the rule-of-thumb bandwidth $h = (1/\sqrt{12})T^{-1/5}$. Our test statistic involves estimation of a generalized long-run variance, so we need to choose a proper lag order p_T. One choice of p_T is discussed in Lima and Xiao (2010) and Hong et al. (2017):

$$p_T = min\left\{ \left\lfloor \left(\frac{3T}{2}\right)^{1/3} \left(\frac{2\hat{\rho}}{1-\hat{\rho}^2}\right) \right\rfloor, \left\lfloor 8\left(\frac{T}{100}\right)^{1/3} \right\rfloor \right\},$$

where $\hat{\rho}$ is the estimator of the maximum first order autocorrelation of $\{Z_{jt}\}_{j=1}^q$ and Y_t. For $k(\cdot)$, we use the Bartlett kernel.

We compare our test with the nonparametric CUSUM-type tests proposed by Su and Xiao (2008). Their test statistics are based on nonparametric estimation of the unknown regression function. Following Su and Xiao (2008), we consider both types of test statistics:

$$KS_T = \max_{1\le t\le T} \left| \frac{1}{\sqrt{T}\hat{\sigma}} \sum_{j=1}^{t} \hat{v}_j \right|,$$

$$CM_T = \frac{1}{T}\sum_{t=1}^{T} \left(\frac{1}{\sqrt{T}\hat{\sigma}} \sum_{j=1}^{t} \hat{v}_j \right)^2,$$

where $\hat{v}_t = [Y_t - \hat{g}(X_t)]\hat{f}(X_t)\omega(X_t)$. Here $\hat{g}(X_t)$ and $\hat{f}(X_t)$ are the nonparametric kernel estimators for the regression function $g_0(X_t)$ under $\mathbb{H}_0$ and the density function $f(X_t)$ of X_t respectively, and $\omega(X_t)$ is a weighting function. We follow Su and Xiao (2008) to set $\omega(x) = [sin(x) + cos(x)]a_x$, where $a_x = \mathbf{1}[\hat{f}(x) > 0.001/logT]$. We choose a fourth order Epanechnikov kernel. For the bandwidth h, we choose $h = h_0T^{\frac{1}{9}}T^{-\gamma}$ where $\gamma = \frac{1}{7}$, and h_0 is chosen via a least squares cross-validation procedure.

In addition to asymptotic critical values, we also use the following moving block bootstrap (MBB): Step (i), choose a block length $l_T = l(T) \in \mathbb{N}(1 \le l_T \le T)$ such that $l_T \to \infty$ as $T \to \infty$; Step (ii), divide data into $T - l_T + 1$ blocks and generate block data $\{\Xi_t\}_{t=1}^{T-l_T+1}$, where $\Xi_t = \{[Y_t, Z_t], \ldots, [Y_{t+l_T-1}, Z_{t+l_T-1}]\}$; Step (iii), resample $\{\Xi_t\}_{t=1}^{T-l_T+1}$ with replacement to form a bootstrap data set $\{\Xi_t^{\mathcal{B}}\}_{t=1}^{L}$ satisfying $T = \lfloor Ll_T \rfloor$; Step (iv), calculate $\widehat{SQ}^{\mathcal{B}}$ using $\{\Xi_t^{\mathcal{B}}\}_{t=1}^{L}$; Step (v), repeat Steps (iii) to (iv) B times to generate bootstrapped test statistics $\{\widehat{SQ}_b^{\mathcal{B}}\}_{b=1}^{B}$. Then the p-value is given by

$$p^{\mathcal{B}} = \frac{1}{B}\sum_{b=1}^{B} \mathbf{1}(\widehat{SQ} \le \widehat{SQ}_b^{\mathcal{B}}).$$

We reject $\mathbb{H}_0$ when $p^{\mathcal{B}}$ is smaller than the given significance level. Choosing an optimal block length is crucial and many approaches have been proposed (e.g., Lahiri, 1999). In this paper, we adopt Politis and White's (2004) automatic block-length selection procedure.

Table 1 reports the size performance of our test and Su and Xiao's (2008) nonparametric CUSUM-type tests. $\widehat{SQ}^{AS}$ and $\widehat{SQ}^{BS}$ denote the results of $\widehat{SQ}$ using the asymptotic and bootstrapped critical values respectively. For DGPs S.1 and S.2, we find that all tests perform reasonably well in finite samples. For DGP S.3, when the dimension of regressors is high, we find that our test displays some under-rejections using asymptotic critical values, but it improves as the sample size T increases. In both DGP 's S.4 and S.5, where one regressor is discrete, our test performs reasonably well, but Su and Xiao's (2008) test is no longer applicable.

Table 1
Empirical rejection rates in finite samples.

		$\widehat{SQ}^{AS}$		$\widehat{SQ}^{BS}$		KS_T		CM_T	
		5%	10%	5%	10%	5%	10%	5%	10%
DGP S.1	$T = 100$	0.046	0.091	0.053	0.106	0.071	0.123	0.062	0.118
	$T = 200$	0.056	0.106	0.056	0.121	0.077	0.116	0.067	0.109
DGP S.2	$T = 100$	0.058	0.140	0.056	0.136	0.066	0.124	0.061	0.122
	$T = 200$	0.062	0.132	0.052	0.122	0.065	0.128	0.058	0.121
DGP S.3	$T = 100$	0.039	0.089	0.048	0.094	0.071	0.121	0.068	0.119
	$T = 200$	0.048	0.118	0.053	0.122	0.062	0.116	0.058	0.114
DGP S.4	$T = 100$	0.036	0.088	0.042	0.138	–	–	–	–
	$T = 200$	0.049	0.104	0.052	0.114	–	–	–	–
DGP S.5	$T = 100$	0.039	0.094	0.056	0.136	–	–	–	–
	$T = 200$	0.047	0.093	0.052	0.120	–	–	–	–

Notes: (i) Number of replication = 1000; (ii) Number of bootstrapping = 499; (iii) $\widehat{SQ}^{AS}$ and $\widehat{SQ}^{BS}$ denote the result of $\widehat{SQ}$ using asymptotic and bootstrap critical values respectively; (iv) KS_T and CM_T are the Kolmogorov–Smirnov and Cramer–von Mises test statistics of Su and Xiao (2008); (v) The weighting function $W(u)$ is the joint $N(0, 1)$ density function; (vi) For DGP S.4 and DGP S.5, Su and Xiao's (2008) test is no longer applicable.

Table 2
Empirical rejection rates in finite samples.

		$\widehat{SQ}^{AS}$		$\widehat{SQ}^{BS}$		KS_T		CM_T	
		5%	10%	5%	10%	5%	10%	5%	10%
DGP P.1	$T = 100$	0.588	0.794	0.512	0.768	0.818	0.886	0.710	0.836
	$T = 200$	0.614	0.893	0.668	0.881	0.823	0.965	0.791	0.925
DGP P.2	$T = 100$	0.900	0.968	0.776	0.932	0.972	0.990	0.944	0.982
	$T = 200$	0.938	0.989	0.866	0.963	0.992	1.000	0.976	1.000
DGP P.3	$T = 100$	0.226	0.448	0.316	0.530	0.116	0.168	0.086	0.168
	$T = 200$	0.379	0.612	0.462	0.716	0.136	0.197	0.098	0.186
DGP P.4	$T = 100$	0.248	0.394	0.138	0.318	0.166	0.288	0.174	0.294
	$T = 200$	0.301	0.510	0.289	0.512	0.213	0.365	0.201	0.361

Notes: (i) Number of replication = 1000; (ii) Number of bootstrapping = 499; (iii) $\widehat{SQ}^{AS}$ and $\widehat{SQ}^{BS}$ denote the result of $\widehat{SQ}$ using asymptotic and bootstrap critical values respectively; (iv) KS_T and CM_T are the Kolmogorov–Smirnov and Cramer–von Mises test statistics of Su and Xiao (2008); (v) The weighting function $W(u)$ is the joint $N(0, 1)$ density function.

Table 2 reports the power performance of the tests. Not surprisingly, Su and Xiao's (2008) tests outperform our test under DGPs P.1 and P.2 where their tests have nontrivial power. Even so, our test is also very powerful to detect the deviations from $\mathbb{H}_0$ under the two alternatives. Under DGPs P.3 and P.4, Su and Xiao's (2008) tests have lower power. In contrast, our test has reasonable power and is more powerful than Su and Xiao's (2008) tests.

To further demonstrate the power property of our test compared to Su and Xiao's (2008) tests, we plot the finite sample power for different break sizes. Consider the following two DGPs:

DGP P.1' [Single Structural Break]:

$$Y_t = \begin{cases} 1.5 + X_{1t} + v_t & \text{if } t \le 0.3T, \\ 1.5 + (1-c)X_{1t} + v_t & \text{otherwise;} \end{cases}$$

and

DGP P.3' [Smooth Structural Change]:

$$Y_t = \theta(\tau, c)(1 + 0.5X_{1t}) + v_t,$$

where $\theta(\tau, c) = 1 - ce^{-10(\tau-0.5)^2}$ with $\tau = t/T$. Here c measures the magnitude of structural changes. We let c vary from 0 to 2 to investigate its impact on the empirical power of our test and Su and Xiao's (2008) tests. We set the sample size to be $T = 100$, and compute the empirical rejection rates of $\widehat{SQ}$ via the bootstrapped critical values (solid blue lines), the asymptotic critical values (dashed blue lines), and the empirical critical values (star-dotted blue lines). We compute the rejection rates of KS_T (solid black lines) and CM_T (dashed black lines) via the wild bootstrap in Su and Xiao (2008). For each experiment, we simulate 500 random samples and set the number of bootstrap for each replication to be 399. The results for DGP P.1' and DGP P.3' are shown in Figs. 1 and 2.

Fig. 1 presents the comparison of finite sample powers under abrupt structural breaks. It reveals that Su and Xiao's (2008) are more powerful than our tests. It is because Su and Xiao's (2008) tests can detect certain local alternatives at the parametric

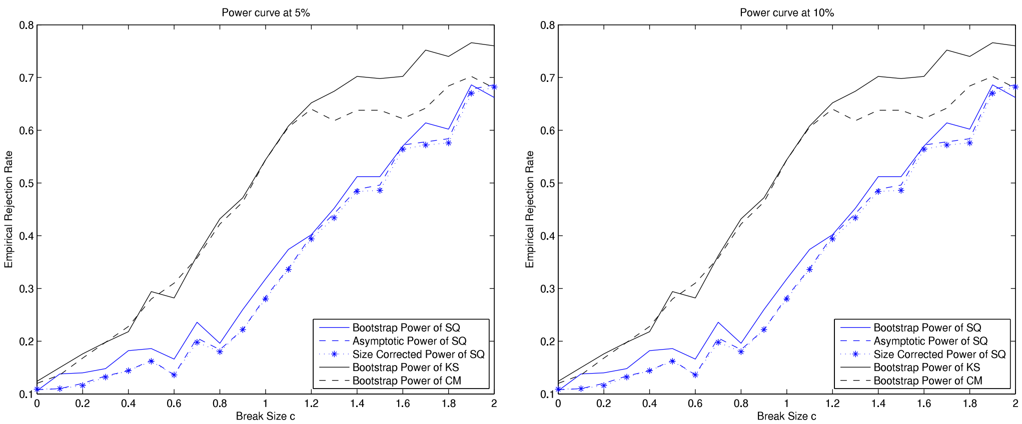

Fig. 1. Empirical rejection rate for DGP P.1'.

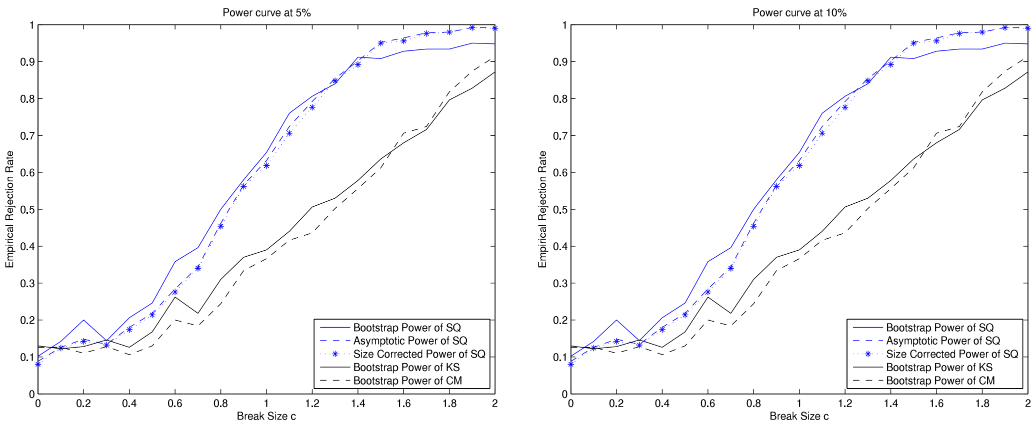

Fig. 2. Empirical rejection rate for DGP P.3'.

rate. However, when it comes to smooth structural changes defined in DGP P.3', our test is more powerful. As shown in Fig. 2, our test can outperform Su and Xiao's (2008) tests under smooth structural changes. Since our test is consistent against all alternatives while Su and Xiao's (2008) tests are not, our test is expected to have all-round power against various alternatives, although it may not have the best power against certain directions.

7.2. The endogeneity case

To study the finite sample performance of our test with endogeneity, we consider the following DGPs.
DGPIV S.1 [No Structural Change]:

$$Y_t = \ln(|X_t - 1| + 1)\mathrm{sgn}(X_t - 1) + v_t,$$
$$X_t = \beta Z_t + w_t,$$

where v_t and w_t are generated by the bivariate normal distribution

$$\begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim i.i.d.N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}\right].$$

Z_t is an instrumental variable and follows an AR(1) process.

$$Z_t = 0.5Z_{t-1} + \xi_t,$$

where $\xi_t \sim$ i.i.d.$N(0, 0.75)$. This DGP is the time series version of the example used by Newey and Powell (2003), where they assume $Z_t \sim$ i.i.d.$N(0, 1)$ with $\beta = 1$. The parameter β represents the strength between the instrumental variable Z_t and the

Table 3
Empirical rejection rates at 5% nominal level.

		$\beta = 0.1$		$\beta = 1$		$\beta = 2$	
		$\widehat{SQ}^{AS}$	$\widehat{SQ}^{BS}$	$\widehat{SQ}^{AS}$	$\widehat{SQ}^{BS}$	$\widehat{SQ}^{AS}$	$\widehat{SQ}^{BS}$
DGPIV S.1	$T = 100$	0.022	0.022	0.016	0.026	0.028	0.046
	$T = 200$	0.012	0.016	0.018	0.020	0.036	0.038
DGPIV P.1	$T = 100$	0.946	0.484	0.948	0.534	0.946	0.566
	$T = 200$	1.000	0.590	1.000	0.650	1.000	0.612
DGPIV P.2	$T = 100$	0.072	0.102	0.074	0.106	0.120	0.150
	$T = 200$	0.160	0.176	0.198	0.208	0.260	0.284

Notes: (i) Number of replication = 500; (ii) Number of bootstrapping = 499; (iii) $\widehat{SQ}^{AS}$ and $\widehat{SQ}^{BS}$ denote the result of $\widehat{SQ}$ using asymptotic and bootstrap critical values respectively; (iv) The weighting function $W(u)$ is the joint $N(0, 1)$ density function.

Table 4
Empirical rejection rates at 10% nominal level.

		$\beta = 0.1$		$\beta = 1$		$\beta = 2$	
		$\widehat{SQ}^{AS}$	$\widehat{SQ}^{BS}$	$\widehat{SQ}^{AS}$	$\widehat{SQ}^{BS}$	$\widehat{SQ}^{AS}$	$\widehat{SQ}^{BS}$
DGPIV S.1	$T = 100$	0.058	0.062	0.062	0.088	0.080	0.112
	$T = 200$	0.048	0.056	0.054	0.072	0.096	0.114
DGPIV P.1	$T = 100$	0.984	0.748	0.988	0.756	0.989	0.748
	$T = 200$	1.000	0.910	1.000	0.932	1.000	0.910
DGPIV P.2	$T = 100$	0.164	0.214	0.168	0.234	0.234	0.300
	$T = 200$	0.274	0.302	0.316	0.360	0.396	0.468

Notes: (i) Number of replication = 500; (ii) Number of bootstrapping = 499; (iii) $\widehat{SQ}^{AS}$ and $\widehat{SQ}^{BS}$ denote the result of $\widehat{SQ}$ using asymptotic and bootstrap critical values respectively; (iv) The weighting function $W(u)$ is the joint $N(0, 1)$ density function.

regressor X_t because their correlation is positively related to β:

$$corr(X_t, Z_t) = \sqrt{\frac{1}{1 + 1/\beta^2}}.$$

We set $\beta = 0.1$, 1, and 2 respectively to investigate the impact of weak and strong instruments on our test.

To examine the power performance, we consider the following DGPs:

DGPIV P.1 [Abrupt Structural Break]:

$$Y_t = \begin{cases} \ln(|X_t - 1| + 1)\text{sgn}(X_t - 1) + v_t & \text{if } t \le 0.3T, \\ \ln(|X_t - 1| + 2) + v_t & \text{otherwise;} \end{cases}$$

DGPIV P.2 [Smooth Structural Change]:

$$Y_t = \theta(\tau)\ln(|X_t - 1| + 1)\text{sgn}(X_t - 1) + v_t,$$

where $\theta(\tau) = 1.5\tau - e^{-(\tau-0.5)^2}$ with $\tau = t/T$.

We only examine the performance of our test because no existing tests are applicable to testing structural changes in nonparametric regression with endogeneity. Tables 3 and 4 report the empirical rejection rates at 10% and 5% significance levels, respectively. Under DGPIV S.1, we find that when the instrumental variable has a weak dependence (e.g., $\beta = 0.1$) with the covariates, $\widehat{SQ}$ is undersized using asymptotic critical values. But it improves as sample size T increases. When the dependence between the instrumental variable and the covariates increases (i.e., $\beta = 2$), the size of our test improves. The empirical rejection rate using bootstrapped critical values is closer to the nominal level than that using asymptotic critical values. Overall, we observe that the strength of correlation between X_t and Z_t affects the size performance of our test. When $\beta = 2$, the finite sample rejection rates of both $\widehat{SQ}^{AS}$ and $\widehat{SQ}^{BS}$ are close to the nominal levels.

Next, we demonstrate the power performance of our test with respect to different magnitudes of structural changes. We made slight modifications to DGPIVs P.1 and P.2:

DGPIV P.1' [Abrupt Structural Break]:

$$Y_t = \begin{cases} \ln(|X_t - 1| + 1)\text{sgn}(X_t - 1) + v_t & \text{if } t \le 0.3T, \\ \ln(|X_t - 1| + 1)\text{sgn}(X_t - 1) + c + v_t & \text{otherwise;} \end{cases}$$

DGPIV P.2' [Smooth Structural Change]:

$$Y_t = \theta(\tau, c)\ln(|X_t - 1| + 1)sgn(X_t - 1) + v_t,$$

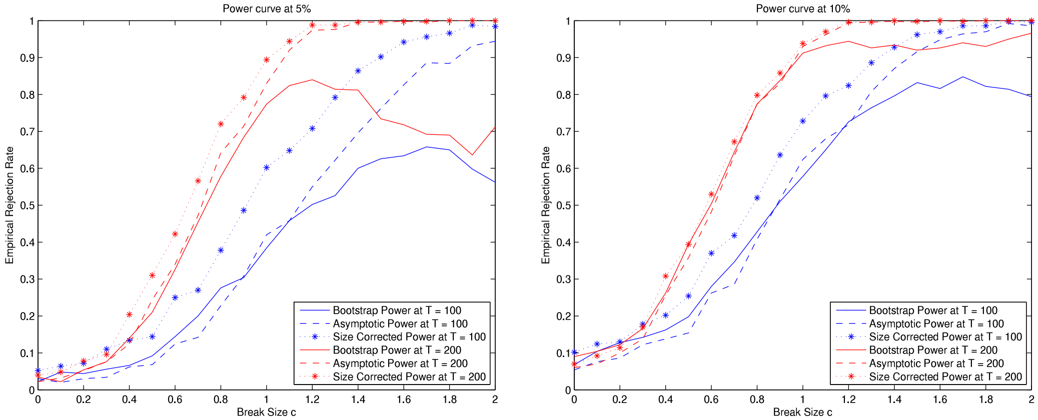

Fig. 3. Empirical rejection rate for DGPIV P.1' with $\beta = 0.1$.

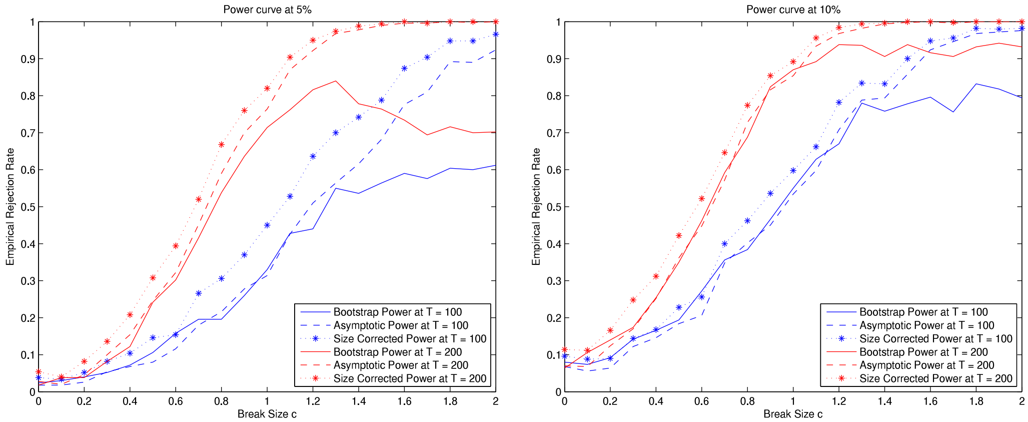

Fig. 4. Empirical rejection rate for DGPIV P.1' with $\beta = 1$.

where $\theta(\tau, c) = 1 - ce^{-10(\tau-0.5)^2}$ with $\tau = t/T$. Here, c measures the strength of signals. We let c increase from 0 to 2 to show its impact on the finite sample power of our test. We compute the empirical rejection rates via the bootstrapped critical values (solid lines), the asymptotic critical values (dashed lines), and the empirical critical values (star-dotted lines). We set the sample size to be 100 (blue lines) and 200 (red lines) to check whether the power increases as the sample size grows. For each experiment, we simulate 500 random samples and set the number of bootstrap for each replication to be 399.

Figs. 3–5 show the empirical power curves under DGPIV P.1' at the 5% and 10% significance levels with $\beta = 0.1$, 1, and 2, respectively. Figs. 6–8 are generated under DGPIV P.2'. In general, we see that our test has nontrivial power for both smooth structural changes and abrupt structural breaks. And the empirical rejection rates increase as the sample size grows.

Furthermore, we have two important observations. One is that our test is slightly undersized using both the bootstrapped and asymptotic critical values, which is characterized by the distance from the star-dotted line (size corrected power) to the solid line (bootstrap power) and the dashed line (asymptotic power). However, the distance becomes smaller as the strength of instrumental variables increases, and the sample size grows. When the sample size is 200, the undersize problem is almost trivial even when the correlation between the instruments and the regressors is relatively small (i.e., $\beta = 0.1$).

The other observation is that the power curve via bootstrapped critical values is not monotonically increasing with the break size c under abrupt structural breaks. Such a problem does not exist when the structural change is smooth. The intuition for that is quite straightforward. Under the case of abrupt structural breaks, the regression function can be viewed as a step function of time. When the break size is relatively small, our smoothed nonparametric estimator can approximate the step function reasonably well such that the asymptotic distribution is still normal. Therefore, we can observe monotonically increasing power as c increases to certain values. However, when the break size is large enough such that the bias of the smoothed nonparametric estimator dominates the behavior of our test statistic, the asymptotic result of our test may not hold anymore. Therefore, the bootstrap cannot capture the asymptotic behavior of our test statistic, leading to the non-monotonic power problem. Such a problem is relatively severe when we approximate the tail distribution (i.e., at the 5% significance level) via bootstrap. Although the bootstrapped power is non-monotonic, we still observe that the empirical

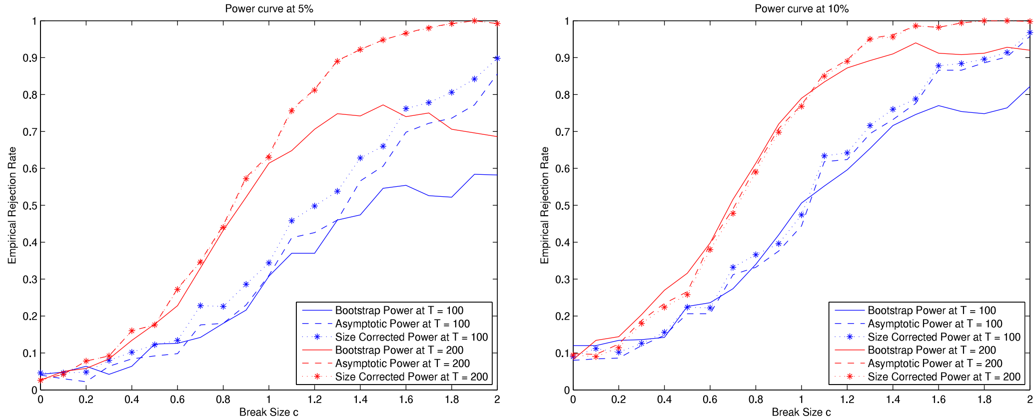

Fig. 5. Empirical rejection rate for DGPIV P.1’ with $\beta = 2$.

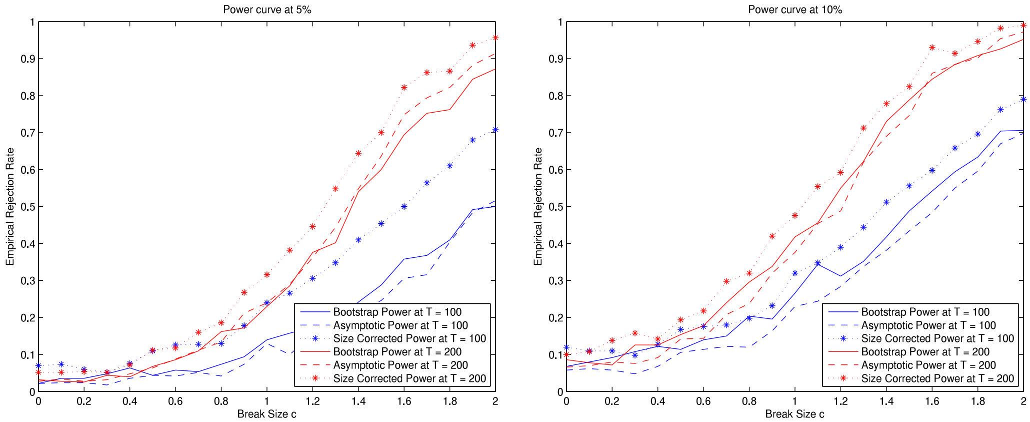

Fig. 6. Empirical rejection rate for DGPIV P.2’ with $\beta = 0.1$.

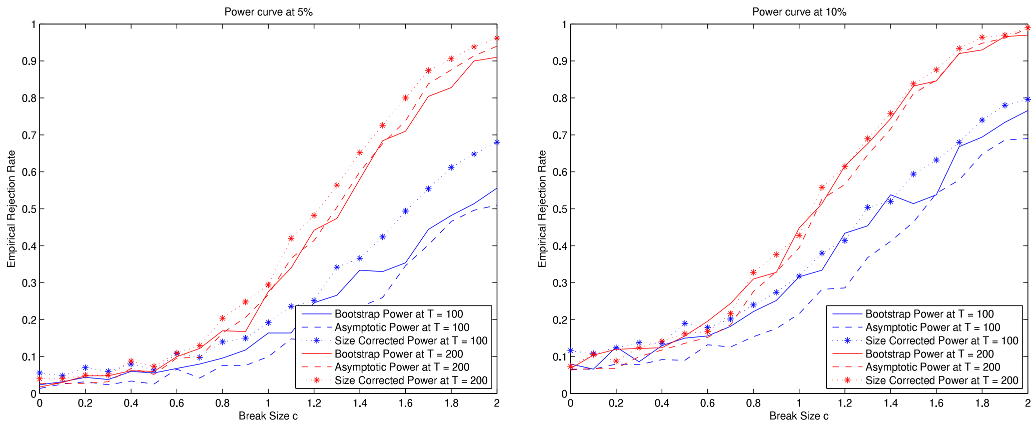

Fig. 7. Empirical rejection rate for DGPIV P.2’ with $\beta = 1$.

rejection rate increases as the sample size grows at each given break size. It implies that our test can still work well in finite samples when there exist abrupt structural breaks. On the contrary, when the structural change is smooth, all three power curves increase as the magnitude of structural changes grows. Because the regression function is a smooth function of time, the smoothed nonparametric estimator can approximate the Fourier transform reasonably well under $\mathbb{H}_A$. Therefore, the asymptotic normality of our test is not affected by the break size c. The finite sample power of our test also increases as the sample size grows. Moreover, when the structural change is smooth, the undersized problem is quite trivial.

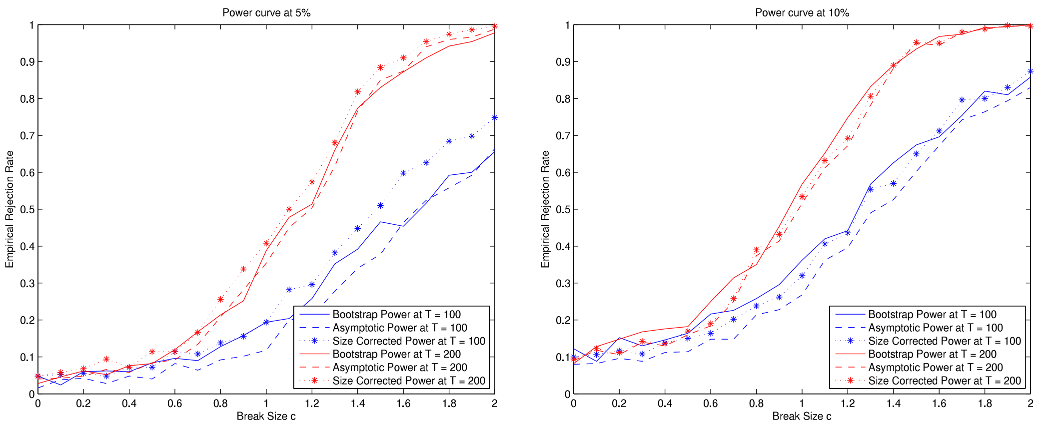

Fig. 8. Empirical rejection rate for DGPIV P.2' with $\beta = 2$.

As pointed out by one referee, our test can detect structural changes that are not explicitly associated with the existing covariates. Consider the following DGP as an example:

$$\begin{aligned} Y_t &= \alpha X_t + \beta_t R_t + \eta_t \\ &= \beta_t E(R_t) + \alpha X_t + \beta_t [R_t - E(R_t)] + \eta_t \\ &= \beta_t \mu_r + \alpha X_t + \beta_t (R_t - \mu_r) + \eta_t, \end{aligned}$$

where X_t is an observable covariate, R_t is unobservable, and $corr(X_t, R_t) \neq 0$. Both X_t and R_t are exogenous such that the true error term η_t satisfies $E(\eta_t | X_t, R_t) = 0$. The parameter $\mu_r \equiv E(R_t)$ is the constant mean of R_t, α is a constant regression coefficient, while β_t changes over time. Here a structural change exists and it is associated with R_t. However, since we do not observe R_t, we may write the model in the following way:

$$Y_t = \beta_t \mu_r + \alpha X_t + v_t,$$

where $v_t \equiv \beta_t(R_t - \mu_r) + \eta_t$. In this model, endogeneity arises given $E(X_t v_t) \neq 0$. According to our approach, we need to find an instrumental variable Z_t such that $E(v_t | Z_t) = 0$ and $E(Z_t X_t) \neq 0$. Our test will detect the structural change in β_t as long as $\mu_r \neq 0$. Therefore, even when we do not have the pertinent covariates that drive the change, we can still know whether a structural change exists. In this sense, our test is more applicable than the existing approaches that require exogeneity.

To sum up, we demonstrate the reasonable finite sample performance of our test in nonparametric regressions with exogeneity and endogeneity. When covariates are all exogenous, our test is more powerful under several alternatives than Su and Xiao's (2008) tests. Moreover, when endogenous or discrete covariates are present, our test also performs reasonably well in finite samples.

8. Empirical application

In this section, we revisit the predictability of equity premium. It has been documented that many financial and macroeconomic variables usually have poor out-of-sample predictive power for equity returns, see (e.g.) Welch and Goyal (2008). Many studies attribute the failure of out-of-sample prediction to the existence of structural changes. Chen and Hong (2012) examine the stability of predictive regressions using 14 financial and economic variables and find substantial evidence against the stability of both univariate and multivariate linear predictive models for equity returns. However, as mentioned by Chen and Hong (2012), the rejection could be due to model misspecification rather than structural changes. Hillebrand et al. (2009) and Campbell and Thompson (2007) find that by imposing sign restrictions, linear predictive models could provide an improved out-of-sample prediction. This implies the failure of out-of-sample prediction could result from model misspecification. Lee et al. (2014) impose monotonicity in both nonparametric and semiparametric predictive regression models, and they find that these models have better predictability for equity premium.

We try to answer the following crucial question: Is the failure of out-of-sample prediction due to the existence of neglected nonlinearity or structural changes? Although Lee et al. (2014) document the superior performance of nonparametric and semiparametric models over linear regression models, they might overlook the instability of the underlying DGP. Our test is applicable since it can detect structural changes without model misspecification.

We follow Lee et al. (2014) and examine 4 predictors: default spread (ds), smoothed earnings–price ratio (se/p), long-term yields on U.S. government bonds (lty), and yields of the 3-Month T-bill on the secondary market (t-bill). The dependent variable is the monthly return of the S&P 500 index (Y_{t+1}), where $Y_{t+1} = \log[(P_{t+1} + D_{t+1})/P_t] - r_t$, P_t is the monthly S&P

Table 5
Stability test for predictive regressions.

	Pre-oil-shock period			Post-oil-shock period		
	$s = 1$	$s = 6$	$s = 12$	$s = 1$	$s = 6$	$s = 12$
ds	0.008	0.0341	0.0661	0.012	0.1603	0.1543
se/p	0.001	0.0342	0.0421	0.008	0.1643	0.1563
lty	0.014	0.0281	0.0862	0.012	0.1784	0.1603
t-bill	0.016	0.0261	0.0601	0.022	0.1423	0.1784

Notes: (i) This table reports the bootstrapped p-values of $\widehat{SQ}$; (ii) Number of bootstrapping = 499; (iii) The predictors are default spread (ds), smoothed earnings–price ratio (se/p), long-term yields on U.S. government bonds (lty), and yields on the 3-Month T-bill on the secondary market (t-bill); (iv) Pre-oil-shock period: January 1950 to December 1975; (v) Post-oil-shock period: January 1976 to December 2005; (vi) The weighting function $W(u)$ is the joint $N(0, 1)$ density function.

500 index, D_t is the dividend paid on the S&P 500 index, and r_t is the 3-month Treasury bill rate. We consider the following predictive regression:

$$Y_{t+s} = g(X_t) + v_{t+s},$$

where s is the number of steps ahead in out-of-sample prediction. Our data are from Lee et al. (2014) and Chen and Hong (2012).

To account for the oil shocks in the 1970s, we divide our sample into two subperiods: the pre-oil-shock sample (January 1950 to December 1975) and the post-oil-shock sample (January 1976 to December 2005). The p-values of our test $\widehat{SQ}$ are obtained by the block-bootstrap as described in Section 7. Table 5 reports the p-values of our test for stability of equity returns prediction. When the forecasting step is one month ahead ($s = 1$), our test rejects the stability of the predictive relationship between equity returns and each predictor during both the pre-oil-shock and post-oil-shock periods. Although the near past predictors should contain relevant information about the outcome one-step ahead, such a relationship is usually unstable due to the rapidly changing economic environment. When $s = 6$, we find that the predictive relationship is unstable for the pre-oil-shock period but is stable for the post-oil-shock period at the 5% significance level. When the forecasting step is one year ahead, only the earnings–price ratio (se/p) has an unstable predictive relationship with the future equity return at the 5% significance level in the pre-oil-shock period. The predictive relationship with each predictor is stable in the post-oil-shock period.

Because we do not assume a parametric functional form for the regression function, it is reasonable that we obtain weaker or less significant results. However, our results, once significant, indicate structural changes which are not contaminated by model misspecification. Therefore, different from the relatively strong evidence of unstable predictive relationships between equity returns and financial/macroeconomic variables detected by Chen and Hong (2012), our findings indicate that the source of rejection in Chen and Hong (2012), who consider a linear predictive model, could come from model misspecification rather than structural changes. However, our test does document strong evidence of structural changes in predictive regressions for $s = 1$ and $s = 6$ in the pre-oil-shock period. It implies that structural change exists and is one important source for poor out-of-sample forecasts, no matter what predictive regression models are used.

9. Conclusion

The failure of out-of-sample forecasting of time series regression models that assume stationarity may be due to the existence of neglected nonlinearity or structural changes. It is important to distinguish structural changes from neglected nonlinearity and/or model misspecification because their implications on modeling, inference, and prediction are quite different. Therefore, it is highly desirable to develop consistent tests for structural changes that are robust to model misspecification.

In this paper, we have proposed a model-free consistent test for structural changes in regression by testing the time-varying property of the Fourier transform of data. It avoids direct nonparametric estimation of the unknown regression function, which is required by the existing nonparametric tests for structural changes. Our test does not suffer from the notorious "curse of dimensionality" problem. As a result, it is asymptotically more powerful than Vogt's (2015) nonparametric test, which is the only consistent test for structural changes in a nonparametric regression model in the existing literature. Although the nonparametric tests of Hidalgo (1995) and Su and Xiao (2008) are asymptotically more powerful than our test against certain smooth local alternatives, they are not consistent tests against a class of fixed alternatives and are asymptotically less powerful than our test against a class of nonsmooth local alternatives. We note that our test is not robust to structural changes in the marginal distributions of the regressors or instruments. Compared to Su and Xiao (2008) and Vogt (2015), we restrict the regressors to be strictly stationary. Although the strict stationarity condition makes our test to be less applicable in scope, we can achieve the efficiency gain. Furthermore, unlike the existing literature, our approach applies to regression models with exogenous and/or endogenous covariates, and we allow the regressors and instruments to be discrete random variables. By using a suitable weighting function, we can avoid the numerical integration over the transform parameters, which is computationally convenient and efficient in practice. Our test statistic follows a convenient asymptotic null $N(0, 1)$ distribution and is consistent against a larger class of smooth structural changes as well as abrupt structural breaks with unknown breakpoints. Simulations show that the test has reasonable size and all-around power against various alternatives of abrupt breaks and smooth changes in both linear and nonlinear time series regression models. An empirical application of our model-free test to predictive regressions suggests that although the neglected nonlinearity is one crucial source of poor out-of-sample predictability for the S&P 500 returns, one still needs to account for potential structural instabilities at certain forecasting horizons.

Acknowledgment

Fu acknowledges financial supports from Shanghai Pujiang Program (No. 18PJC006).

Appendix. Mathematical appendix

Throughout the Appendix, we use the same notations as defined in the paper.

Proof of Lemma 1. First we show the sufficient condition: if $Pr[g_t(X_t) = g_0(X_t)] = 1$ then $\phi_t(u) = \phi_0(u)$ for all $u \in \mathbb{R}^q$.

By definition, $\phi_t(u) = E(Y_t e^{iu'Z_t}) = E[g_t(X_t)e^{iu'Z_t}] = E[g_0(X_t)e^{iu'Z_t}] = \phi_0(u)$. The last equality comes from the definition of $\phi_0(u)$. Since this equality holds for any $u \in \mathbb{R}^q$, we have proved the sufficient condition.

Then we show the necessary condition: if $\phi_t(u) = \phi_0(u)$ for all $u \in \mathbb{R}^q$, then $Pr[g_t(X_t) = g_0(X_t)] = 1$. By Law of Iterated Expectation,

$$
\begin{aligned}
&\phi_t(u) - \phi_0(u) \\
= &\int [g_t(x) - g_0(x)]e^{iu'z} dF_{XZ}(x, z) \\
= &\int E[g_t(X_t) - g_0(X_t)|Z_t = z]e^{iu'z} dF_Z(z),
\end{aligned}
$$

where $F_{XZ}(x, z)$ and $F_Z(z)$ denote the joint CDF of X_t and Z_t and marginal CDF of Z_t. Let $\Delta_t(z) = E[g_t(X_t) - g_0(X_t)|Z_t = z]$. We first show that if $\phi_t(u) = \phi_0(u)$ for all $u \in \mathbb{R}^q$, then $Pr[\Delta_t(Z_t) = 0] = 1$.

Put

$$\delta_t^{(1)}(z) = max\{\Delta_t(z), 0\}, \qquad \delta_t^{(2)}(z) = max\{-\Delta_t(z), 0\},$$

such that $\delta_1(z)$ and $\delta_2(z)$ are non-negative Borel measurable functions on $\mathbb{R}^q$ and

$$\Delta_t(z) = \delta_t^{(1)}(z) - \delta_t^{(2)}(z).$$

Now, let the moments

$$c_t^{(j)} = E[\delta_t^{(j)}(z)] > 0, \quad \text{for } j = 1, 2.$$

Define the new probability measures $\nu_t^{(1)}$ and $\nu_t^{(2)}$ on the Euclidean Borel field $\mathbb{B}$ as

$$\nu_t^{(j)}(B) = \int_B \delta_t^{(j)}(z) dF_Z(z)/c_t^{(j)}, \ \text{for } j = 1, 2,$$

where B is an arbitrary Borel set in $\mathbb{B}$. Now we can rewrite

$$
\begin{aligned}
&\phi_t(u) - \phi_0(u) \\
= &\int \Delta_t(z)e^{iu'z} dF_Z(z) \\
= &\int [\delta_t^{(1)}(z) - \delta_t^{(2)}(z)]e^{iu'z} dF_Z(z) \\
= &\int \delta_t^{(1)}(z)e^{iu'z} dF_Z(z) - \int \delta_t^{(2)}(z)e^{iu'z} dF_Z(z).
\end{aligned}
$$

Given $\nu_t^{(j)}(B) = \int_B \delta_t^{(j)}(z) dF_Z(z)/c_t^{(j)}$ for $j = 1,2$, we have

$$d\nu_t^{(j)}(z)c_t^{(j)} = \delta_t^{(j)}(z) dF_Z(z).$$

Thus

$$
\begin{aligned}
&\phi_t(u) - \phi_0(u) \\
= &\int \delta_t^{(1)}(z)e^{iu'z} dF_Z(z) - \int \delta_t^{(2)}(z)e^{iu'z} dF_Z(z) \\
= &\, c_t^{(1)} \int e^{iu'z} d\nu_t^{(1)}(z) - c_t^{(2)} \int e^{iu'z} d\nu_t^{(2)}(z).
\end{aligned}
$$

Since $\phi_t(u) - \phi_0(u) = 0$ for all $u \in \mathbb{R}^q$ and all t, we have $c_t^{(1)} = c_t^{(2)}$ for all t by letting $u = 0$ and the fact that $\nu_t^{(1)}$ and $\nu_t^{(2)}$ are both probability measures. Therefore, we have

$$\int e^{iu'z} d\nu_t^{(1)}(z) = \int e^{iu'z} d\nu_t^{(2)}(z)$$

for all $u \in \mathbb{R}^q$ and all t. And it is equivalent to say that the two probability measures $\nu_t^{(1)}$ and $\nu_t^{(2)}$ are equal for all t. Then we have

$$\int_B \delta_t^{(1)}(z) dF_Z(z) = \int_B \delta_t^{(2)}(z) dF_Z(z),$$

for every Borel set B and all t. Furthermore, by the definition of $\delta_t^{(1)}(z)$ and $\delta_t^{(2)}(z)$, we have

$$\int_B \Delta_t(z) dF_Z(z) = 0$$

for every Borel set B and all t. Let

$$B_1 = \{z \in \mathbb{R}^q : \Delta_t(z) > 0\},$$

and

$$B_2 = \{z \in \mathbb{R}^q : \Delta_t(z) < 0\},$$

be two Borel sets. Then $B_1 \cup B_2 = \{z \in \mathbb{R}^q : \Delta_t(z) \neq 0\}$ is a null set with respect to $F_Z(z)$. Thus we have proved that if $\phi_t(u) = \phi_0(u)$ for all $u \in \mathbb{R}^q$ and all t, then $Pr[\Delta_t(z) = 0] = 1$ for all t.

Next, we need to show that $Pr[\Delta_t(z) = 0] = 1$ implies $Pr[g_t(X_t) - g_0(X_t) = 0] = 1$. Given $\Delta_t(z) = E[g(X_t, \frac{t}{T})|Z_t = z] - E[g_0(X_t)|Z_t = z]$=0, we have

$$\begin{aligned} \frac{\partial \Delta(z, \frac{t}{T})}{\partial \frac{t}{T}} &= \frac{\partial E[g(X_t, \frac{t}{T})|Z_t = z]}{\partial \frac{t}{T}} \\ &= 0, \end{aligned}$$

for any non-zero Borel sets. However, it contradicts Assumption 7, where $E[g_t(X_t)|Z_t = z] = E[g(X_t, \frac{t}{T})|Z_t = z] = r(z, \frac{t}{T})$ such that $\frac{\partial r(z, \frac{t}{T})}{\partial \frac{t}{T}} \neq 0$ for a non-zero Borel measurable set. Therefore, if $Pr[\Delta_t(z) = 0] = 1$ for all t, then $Pr[g_t(X_t) = g_0(X_t)] = 1$ and we have proved the necessary condition.

Intuitively, the instrument Z_t must capture the time-varying feature of $g_t(X_t)$ after the Fourier transform. Under the case of no endogeneity, this condition is satisfied automatically. ■

Proof of Theorem 1. We decompose $Th^{1/2}\hat{Q}$:

$$\begin{aligned} Th^{1/2}\hat{Q} &= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left|\hat{\phi}_t(u) - \hat{\phi}_0(u)\right|^2 W(u)du \\ &= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left|\hat{\phi}_t(u) - \phi_t(u)\right|^2 W(u)du + h^{1/2} \int_{\mathbb{R}^q} \sum_{t=1}^{T} \left|\hat{\phi}_0(u) - \phi_t(u)\right|^2 W(u)du \\ &\quad - 2h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u) - \phi_t(u)\right]\left[\hat{\phi}_0(u) - \phi_t(u)\right]^*\right\} W(u)du \\ &= \hat{Q}_1 + \hat{Q}_2 - 2\hat{Q}_3. \end{aligned}$$

The proof of Theorem 1 consists of the proofs of Theorems A.1–A.3. The asymptotic distribution is determined by $\hat{Q}_1$. Under $\mathbb{H}_0$, $\hat{Q}_2$ and $\hat{Q}_3$ have no impact on the asymptotic distribution of the test statistic $Th^{1/2}\hat{Q}$. ■

Theorem A.1. *Under the conditions of Theorem 1, $(\hat{Q}_1 - \hat{B})/\sqrt{\hat{V}} \xrightarrow{d} N(0,1)$ as $T \to \infty$, where $\hat{B} = h^{-1/2} \int_{\mathbb{R}^q} |\hat{\Omega}(u, u)| W(u)du \int K^2(\eta)d\eta$ and $\hat{V} = 2\int_{\mathbb{R}^{2q}} |\hat{\Omega}(u, v)|^2 W(u)W(v)dudv \int \left[\int K(\eta)K(\eta + \lambda)d\eta\right]^2 d\lambda$.*

Theorem A.2. *Under the conditions of Theorem 1, $\hat{Q}_2 = op(1)$.*

Theorem A.3. *Under the conditions of Theorem 1, $\hat{Q}_3 = op(1)$.*

Proof of Theorem A.1. To show $(\hat{Q}_1 - \hat{B})/\sqrt{\hat{V}} \xrightarrow{d}$ N(0,1) as $T \to \infty$, it suffices to show the following propositions.

Proposition A.1. *Under the conditions of Theorem 1, $\hat{Q}_1 = B_1 + U + op(1)$, where*

$$B_1 = h^{-1/2} \int_{\mathbb{R}^q} E[|\varepsilon_t(u)|^2] W(u) du \int K^2(\eta) d\eta,$$

$$U = \frac{1}{Th^{1/2}} \sum_{1 \le s \ne r \le T} \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*] W(u) du \int K(\eta) K\left(\eta + \frac{s-r}{Th}\right) d\eta.$$

Proposition A.2. *Under the conditions of Theorem 1, $(U - B_2)/\sqrt{V} \xrightarrow{d} N(0,1)$ as $T \to \infty$, where*

$$B_2 = 2h^{-1/2} \sum_{j=1}^{\infty} \int_{\mathbb{R}^q} E(Re[\varepsilon_t(u)\varepsilon_{t+j}(u)^*]) W(u) du \int K^2(\eta) d\eta,$$

$$V = 2 \int_{\mathbb{R}^{2q}} |\Omega(u, v)|^2 W(u) W(v) du dv \int \left[\int K(\eta) K(\eta + \lambda) d\eta \right]^2 d\lambda.$$

Proposition A.3. *Under the conditions of Theorem 1, $B_1 + B_2 - \hat{B} = op(1)$ and $V - \hat{V} = op(1)$.* ■

Proof of Proposition A.1. Given $Y_t e^{iu'Z_t} = \phi_t(u) + \varepsilon_t(u)$,

$$\begin{aligned}
\hat{Q}_1 &= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \hat{\phi}_t(u) - \phi_t(u) \right|^2 W(u) du \\
&= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} Y_s e^{iu'Z_s} H_{st} - \phi_t(u) \right|^2 W(u) du \\
&= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} [\phi_s(u) + \varepsilon_s(u)] H_{st} - \phi_t(u) \right|^2 W(u) du.
\end{aligned}$$

Under $\mathbb{H}_0$, $\phi_t(u) = \phi_0(u)$ for all t, and by the fact that $\sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} H_{st} = 1$,

$$\begin{aligned}
\hat{Q} &= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \varepsilon_s(u) H_{st} \right|^2 W(u) du \\
&= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \frac{1}{Th} K\left(\frac{s-t}{Th}\right) \varepsilon_s(u) \right|^2 W(u) du + op(1) \\
&= h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=1}^{T} \frac{1}{Th} K\left(\frac{s-t}{Th}\right) \varepsilon_s(u) \right|^2 W(u) du \\
&\quad + h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=1-\lfloor Th \rfloor}^{0} \frac{1}{Th} K\left(\frac{s-t}{Th}\right) \varepsilon_s(u) \right|^2 W(u) du \\
&\quad + h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=T+1}^{T+\lfloor Th \rfloor} \frac{1}{Th} K\left(\frac{s-t}{Th}\right) \varepsilon_s(u) \right|^2 W(u) du \\
&\quad + 2h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \sum_{s=1}^{T} \sum_{r=1-\lfloor Th \rfloor}^{0} \frac{1}{T^2h^2} K\left(\frac{s-t}{Th}\right) K\left(\frac{r-t}{Th}\right) Re[\varepsilon_s(u)\varepsilon_r(u)^*] W(u) du \\
&\quad + 2h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \sum_{s=1}^{T} \sum_{r=T+1}^{T+\lfloor Th \rfloor} \frac{1}{T^2h^2} K\left(\frac{s-t}{Th}\right) K\left(\frac{r-t}{Th}\right) Re[\varepsilon_s(u)\varepsilon_r(u)^*] W(u) du \\
&\quad + 2h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \sum_{s=1-\lfloor Th \rfloor}^{0} \sum_{r=T+1}^{T+\lfloor Th \rfloor} \frac{1}{T^2h^2} K\left(\frac{s-t}{Th}\right) K\left(\frac{r-t}{Th}\right) Re[\varepsilon_s(u)\varepsilon_r(u)^*] W(u) du + op(1) \\
&= A_1 + R_1 + R_2 + R_3 + R_4 + R_5,
\end{aligned}$$

where we use the property that $K(\eta) = 0$ for all $\eta \geq 1$ or $\eta \leq -1$. We decompose A_1 as

$$\begin{aligned}
A_1 &= \frac{1}{T^2h^{3/2}} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left[\sum_{s=1}^{T} |\varepsilon_s(u)|^2 K^2\left(\frac{s-t}{Th}\right) + \sum_{s\neq r}^{T} Re[\varepsilon_s(u)\varepsilon_r(u)^*] K\left(\frac{s-t}{Th}\right) K\left(\frac{r-t}{Th}\right) \right] W(u)du \\
&= \frac{1}{T^2h^{3/2}} K^2(0) \sum_{t=1}^{T} \int_{\mathbb{R}^q} |\varepsilon_t(u)|^2 W(u)du + \frac{1}{T^2h^{3/2}} \sum_{s\neq t}^{T} K^2\left(\frac{s-t}{Th}\right) \int_{\mathbb{R}^q} |\varepsilon_s(u)|^2 W(u)du \\
&\quad + \frac{2}{T^2h^{3/2}} \sum_{s\neq t}^{T} K\left(\frac{s-t}{Th}\right) K(0) \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_t(u)^*] W(u)du \\
&\quad + \frac{1}{T^2h^{3/2}} \sum_{s\neq r\neq t}^{T} K\left(\frac{s-t}{Th}\right) K\left(\frac{r-t}{Th}\right) \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*] W(u)du \\
&= R_6 + A_2 + R_7 + A_3.
\end{aligned}$$

Now we show $A_2 = B_1 + op(1)$:

$$\begin{aligned}
A_2 &= \frac{1}{T^2h^{3/2}} \sum_{s\neq t}^{T} K^2\left(\frac{s-t}{Th}\right) \int_{\mathbb{R}^q} |\varepsilon_s(u)|^2 W(u)du \\
&= \frac{2}{T^2h^{3/2}} \sum_{s=1}^{T-1} \int_{\mathbb{R}^q} |\varepsilon_s(u)|^2 W(u)du \sum_{t=s+1}^{T} K^2\left(\frac{t-s}{Th}\right) \\
&= \frac{2}{T^2h^{3/2}} \sum_{s=1}^{T-1} \int_{\mathbb{R}^q} E[|\varepsilon_s(u)|^2] W(u)du \sum_{j=1}^{T-s} K^2\left(\frac{j}{Th}\right) \\
&\quad + \frac{2}{T^2h^{3/2}} \sum_{s=1}^{T-1} \int_{\mathbb{R}^q} (|\varepsilon_s(u)|^2 - E[|\varepsilon_s(u)|^2]) W(u)du \sum_{j=1}^{T-s} K^2\left(\frac{j}{Th}\right) \\
&= E(A_2) + [A_2 - E(A_2)] .
\end{aligned}$$

For $E(A_2)$,

$$\begin{aligned}
E(A_2) &= \frac{2}{T^2h^{3/2}} \sum_{s=1}^{T-1} \int_{\mathbb{R}^q} E[|\varepsilon_s(u)|^2] W(u)du \sum_{j=1}^{T-s} K^2\left(\frac{j}{Th}\right) \\
&= h^{-1/2} \int_{\mathbb{R}^q} E[|\varepsilon_s(u)|^2] W(u)du \int K^2(\eta)d\eta + o(1), \\
&= B_1 + o(1).
\end{aligned}$$

where the second to last equality comes from

$$\frac{2}{Th} \sum_{j=1}^{T-1} \left(1 - \frac{j}{T}\right) K^2\left(\frac{j}{Th}\right) = \int K^2(\eta)d\eta + o(1).$$

For $A_2 - E(A_2)$, it is straightforward to see that $E[A_2 - E(A_2)] = 0$. Then we just need to check $E[(A_2 - E(A_2))^2] = Var(A_2)$:

$$\begin{aligned}
Var(A_2) &= Var\left[\frac{2}{T^2h^{3/2}} \sum_{s=1}^{T-1} \int_{\mathbb{R}^q} |\varepsilon_s(u)|^2 W(u)du \sum_{j=1}^{T-s} K^2\left(\frac{j}{Th}\right) \right] \\
&= \frac{4}{T^4h^3} \left[\sum_{j=1}^{T-1} (T-j) K^2\left(\frac{j}{Th}\right) \right]^2 \int_{\mathbb{R}^q} \left[Var|\varepsilon_s(u)|^2 + 2\sum_{j=1}^{T-1} (T-j) cov\left|\varepsilon_s(u), \varepsilon_{s+j}(u)^*\right| \right] W(u)du \\
&\leq O(T^{-3}h^{-3}) \\
&= o(1).
\end{aligned}$$

where we use β-mixing conditions. Thus $A_2 - E(A_2) = op(1)$ by Chebyshev's inequality.

Next we show that $A_3 = U + op(1)$:

$$\begin{aligned} A_3 &= \frac{1}{T^2h^{3/2}} \sum_{s\neq r\neq t}^{T} K\left(\frac{s-t}{Th}\right) K\left(\frac{r-t}{Th}\right) \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du \\ &= \frac{1}{Th^{1/2}} \sum_{s\neq r}^{T} \int K(\eta)K\left(\eta + \frac{s-r}{Th}\right) d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du + op(1) \\ &= U + op(1), \end{aligned}$$

where the last equality is by the Riemann approximation of an integral.

Lemma A.1. *Under the conditions of Theorem 1, $R_j = op(1)$ for $j = 1 - 7$.*

Thus we have proved Proposition A.1. The proof of Lemma A.1 is given in the supplementary material. ■

Proof of Proposition A.2. Given that the U-statistic we have here exhibits possible strong dependence on nearby observations, the conventional asymptotic theory may not work out. Thus, we have to first remove the terms with possible strong dependence and consider the asymptotic behavior of the remaining term. Following Hong et al. (2017), we introduce a new tuning parameter p_T that satisfies $p_T \to \infty$, $p_T/Th \to 0$ as $T \to \infty$, and $\sum_{j=p_T}^{\infty} j^2\beta(j) \leq Cp_T^{-1}$. Denote $\mathbb{D}_1 = \{(s, r) : 1 \leq |s-r| \leq p_T, 1 \leq s \neq r \leq T\}$ and $\mathbb{D}_2 = \{(s, r) : p_T < |s-r| < T, 1 \leq s \neq r \leq T\}$. Then we can decompose U as

$$\begin{aligned} U &= \frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_1} \int K(\eta)K\left(\eta + \frac{s-r}{Th}\right) d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du \\ &\quad + \frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_2} \int K(\eta)K\left(\eta + \frac{s-r}{Th}\right) d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du \\ &= U_1 + U_2. \end{aligned}$$

First, we show that $U_1 = B_2 + op(1)$:

$$\begin{aligned} U_1 &= \frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_1} \int K^2(\eta)d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du \\ &\quad + \frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_1} \int K(\eta)\left[K\left(\eta + \frac{s-r}{Th}\right) - K(\eta)\right] d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du \\ &= U_{11} + U_{12}. \end{aligned}$$

Now, we claim that $U_{11} = B_2 + op(1)$ and $U_{12} = op(1)$. For U_{11}, we can decompose it as $U_{11} = E(U_{11}) + U_{11} - E(U_{11})$. Then we show that $E(U_{11}) = B_2 + op(1)$ and $U_{11} - E(U_{11}) = op(1)$.

$$\begin{aligned} E(U_{11}) &= E\left\{\frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_1} \int K^2(\eta)d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du\right\} \\ &= \frac{1}{Th^{1/2}} \sum_{s=1}^{T-1} \sum_{r=s+1}^{\min\{s+p_T,T\}} \int_{\mathbb{R}^q} E\left(Re[\varepsilon_s(u)\varepsilon_r(u)^* + \varepsilon_r(u)\varepsilon_s(u)^*]\right) W(u)du \int K^2(\eta)d\eta + op(1) \\ &= h^{-1/2} \sum_{j=1}^{\min\{p_T,T-s\}} \int_{\mathbb{R}^q} E\left(Re[\varepsilon_t(u)\varepsilon_{t+j}(u)^* + \varepsilon_{t+j}(u)\varepsilon_t(u)^*]\right) W(u)du \int K^2(\eta)d\eta + op(1) \\ &= 2h^{-1/2} \sum_{j=1}^{\infty} \int_{\mathbb{R}^q} E\left(Re[\varepsilon_t(u)\varepsilon_{t+j}(u)^*]\right) W(u)du \int K^2(\eta)d\eta + op(1) \\ &= B_2 + op(1). \end{aligned}$$

To show $U_{11} - E(U_{11}) = op(1)$, we just need to calculate $E\{[U_{11} - E(U_{11})]^2\}$:

$$\begin{aligned} E\{[U_{11} - E(U_{11})]^2\} &= E\left\{\frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_1} \int K^2(\eta)d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du\right. \\ &\quad \left. - E\left[\frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_1} \int K^2(\eta)d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du\right]\right\}^2 \\ &= \frac{1}{T^2h} \sum_{s,r\in\mathbb{D}_1} \sum_{l,m\in\mathbb{D}_1} E\left\{Re\left[\varpi(\Xi_s, \Xi_r)\varpi(\Xi_l, \Xi_m)\right]\right\} \left[\int K^2(\eta)d\eta\right]^2, \end{aligned}$$

where we define

$$\varpi(\Xi_s, \Xi_r) = \int_{\mathbb{R}^q} [\varepsilon_s(u)\varepsilon_r(u)^* - E(\varepsilon_s(u)\varepsilon_r(u)^*)]W(u)du.$$

By Lemma 1 of Yoshihara (1976) and Proposition 4 of Hong et al. (2017), $E\{[U_{11} - E(U_{11})]^2\} = op(T^{-1}h^{-1}p_T) = op(1)$. By Chebyshev's Inequality, we have $U_{11} = E(U_{11}) + op(1)$. Thus, we have $U_{11} = B_2 + op(1)$.

Next, we work on U_{12}: by Lipschitz condition, we have

$$\left| K\left(\eta + \frac{s-r}{Th}\right) - K(\eta) \right| \leq C\frac{|s-r|}{Th}.$$

Thus,

$$\begin{aligned} E(U_{12}) &= \frac{2}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_1} \int K(\eta)\left[K\left(\eta + \frac{s-r}{Th}\right) - K(\eta)\right] d\eta \int_{\mathbb{R}^q} Re\{E[\varepsilon_s(u)\varepsilon_r(u)^*]\}W(u)du \\ &\leq C\frac{1}{Th^{3/2}} \sum_{j=1}^{p_T} j\beta(j)^{\frac{\delta}{1+\delta}} \\ &= o(1). \end{aligned}$$

By similar by tedious derivation as in showing $E\{U_{11} - [E(U_{11})]^2\} = o(1)$, we can show $E\{U_{12} - [E(U_{12})]^2\} = o(p_T^3 T^{-3}h^{-3}) = o(1)$. Thus, $U_{12} = op(1)$ by Chebyshev's inequality.

Now, we work on U_2, which determines the asymptotic distribution of our test statistic:

$$\begin{aligned} U_2 &= \frac{1}{Th^{1/2}} \sum_{s,r\in\mathbb{D}_2} \int K(\eta)K\left(\eta + \frac{s-r}{Th}\right) d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_t(u)^*]W(u)du \\ &= \frac{2}{Th^{1/2}} \sum_{s=1}^{T-p_T-1} \sum_{r=s+p_T+1}^{T} \int K(\eta)K\left(\eta + \frac{s-r}{Th}\right) d\eta \int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_r(u)^*]W(u)du. \end{aligned}$$

First, the mean of U_2 is

$$\begin{aligned} E(U_2) &= \frac{2}{Th^{1/2}} \sum_{j=p_T+1}^{T} (T - j - p_T) \int_{\mathbb{R}^q} E\left\{Re[\varepsilon_s(u)\varepsilon_{s+j}(u)^*]\right\} W(u)du \int K(\eta)K\left(\eta + \frac{j}{Th}\right) d\eta \\ &\leq Ch^{1/2} \sum_{j=p_T+1}^{T} j^2\beta(j) \iint K(\eta)K(\eta+\lambda)d\eta d\lambda \\ &\leq O(h^{1/2}p_T^{-1}) \\ &= o(1). \end{aligned}$$

Given $E(U_2) = o(1)$, we have $Var(U_2) = E(U_2^2) + op(1)$. The variance of U_2 is given by the following lemma.

Lemma A.2. *Under the conditions of Theorem 1, $E(U_2^2) = V + op(1)$, where*

$$V = 2\int_{\mathbb{R}^{2q}} |\Omega(u, v)|^2 W(u)W(v)dudv \int \left[\int K(\eta)K(\eta+\lambda)d\eta\right]^2 d\lambda.$$

The proof of Lemma A.2 is given in the supplementary material. At last, we prove asymptotic normality of $Z \equiv U_2/\sqrt{Var(U_2)} \xrightarrow{d} N(0, 1)$. Here, we make use of the central limit theorem in Hong et al. (2017) and we can show that for all $a \in \mathbb{R}$, the moment generating function of Z is

$$M_Z(a) = E(e^{aZ}) = \sum_{r=1}^{\infty} \frac{1}{r!}\left(\frac{a^2}{2}\right)^r = e^{\frac{a^2}{2}}.$$

By the uniqueness of moment generating function, we can show that $Z \sim N(0, 1)$. ■

Proof of Proposition A.3. We show that the estimators $\hat{B}$ and $\hat{V}$ are consistent estimators for B and V respectively. Given

$$\hat{B} = h^{-1/2}\int_{\mathbb{R}^q} |\hat{\Omega}(u, u)|W(u)du \int K^2(\eta)d\eta,$$

and

$$B = h^{-1/2}\int_{\mathbb{R}^q} |\Omega(u, u)|W(u)du \int K^2(\eta)d\eta,$$

it is sufficient to show that $\hat{\Omega}(u,u) - \Omega(u,u) = op(h^{1/2})$. Let $\bar{\Omega}(u,u) = \sum_{j=-p_T}^{p_T} k(\frac{j}{p_T})\bar{\sigma}_j(u,u)$ where $\bar{\sigma}_j(u,u) = \frac{1}{T-|j|}\sum_{t=1}^{T-|j|}\varepsilon_t(u)\varepsilon_{t+j}(u)$. Then we can decompose

$$\begin{aligned}
&\hat{\Omega}(u,u) - \Omega(u,u)\\
&= \bar{\Omega}(u,u) - \Omega(u,u) + \hat{\Omega}(u,u) - \bar{\Omega}(u,u)\\
&= D_1 + D_2.
\end{aligned}$$

First, we can decompose $D_1 = E(D_1) + D_1 - E(D_1)$.

$$\begin{aligned}
E(D_1) &= E[\bar{\Omega}(u,u) - \Omega(u,u)]\\
&= \sum_{j=-p_T}^{p_T}\left[k\left(\frac{j}{p_T}\right) - 1\right]\sigma_j(u,u) - \sum_{|j|>p_T}\sigma_j(u,u)\\
&= ED_{11} + ED_{12}
\end{aligned}$$

We show that $ED_{11} = O(p_T^{-1})$ by Lipschitz condition. And $ED_{12} = O(p_T^{-1})$ by the condition on p_T: $\sum_{|j|>p_T}\sigma_j(u,u) \le p_T^{-1}$. Therefore, we have $ED_1 = op(h^{1/2})$ since $h^{1/2}p_T \to \infty$. For $D_1 - ED_1$, we can follow Hong et al. (2017) to show that $ED_1^2 = O(T^{-1}) = o(h^{1/2})$. Thus, $D_1 = op(h^{1/2})$ by Chebyshev's inequality.

Now, we will work on D_2:

$$\begin{aligned}
D_2 &= \hat{\Omega}(u,u) - \bar{\Omega}(u,u)\\
&= \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right)\frac{1}{T-|j|}\sum_{t=1}^{T-|j|}[\hat{\varepsilon}_t(u)\hat{\varepsilon}_{t+j}(u) - \varepsilon_t(u)\varepsilon_{t+j}(u)]\\
&= \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right)\frac{1}{T-|j|}\sum_{t=1}^{T-|j|}[\hat{\varepsilon}_t(u)\hat{\varepsilon}_{t+j}(u) - \hat{\varepsilon}_t(u)\varepsilon_{t+j}(u) + \hat{\varepsilon}_t(u)\varepsilon_{t+j}(u) - \varepsilon_t(u)\varepsilon_{t+j}(u)]\\
&= \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right)\frac{1}{T-|j|}\sum_{t=1}^{T-|j|}[\hat{\varepsilon}_t(u)\{\hat{\varepsilon}_{t+j}(u) - \varepsilon_{t+j}(u)\}]\\
&\quad + \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right)\frac{1}{T-|j|}\sum_{t=1}^{T-|j|}[\{\hat{\varepsilon}_t(u) - \varepsilon_t(u)\}\varepsilon_{t+j}(u)]\\
&= \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right)\frac{1}{T-|j|}\sum_{t=1}^{T-|j|}[\hat{\varepsilon}_t(u) - \varepsilon_t(u)][\hat{\varepsilon}_{t+j}(u) - \varepsilon_{t+j}(u)]\\
&\quad + \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right)\frac{1}{T-|j|}\sum_{t=1}^{T-|j|}[\varepsilon_t(u)\{\hat{\varepsilon}_{t+j}(u) - \varepsilon_{t+j}(u)\}]\\
&\quad + \sum_{j=-p_T}^{p_T} k\left(\frac{j}{p_T}\right)\frac{1}{T-|j|}\sum_{t=1}^{T-|j|}[\{\hat{\varepsilon}_t(u) - \varepsilon_t(u)\}\varepsilon_{t+j}(u)]\\
&= D_{21} + D_{22} + D_{23}.
\end{aligned}$$

Under $\mathbb{H}_0$, $\hat{\varepsilon}_t(u) - \varepsilon_t(u) = \phi_0(u) - \hat{\phi}_0(u) = Op(T^{-1/2})$ for all t. Thus, we have $D_{21} = Op(T^{-1}p_T)$, $D_{22} = Op(T^{-1/2}p_T)$, and $D_{22} = Op(T^{-1/2}p_T)$. Then we know that $D_2 = op(h^{1/2})$. Thus, we have shown that $\hat{\Omega}(u,u) - \Omega(u,u) = op(h^{1/2})$ and therefore $\hat{B} - B = op(1)$. Following analogous proof, we can show $\hat{V} - V = op(1)$. ∎

Proof of Theorem A.2. The proof of Theorem A.2 is given in the supplementary material. ∎

Proof of Theorem A.3. The proof of Theorem A.3 is given in the supplementary material. ∎

Proof of Theorem 2. Under $\mathbb{H}_A : \phi_t(u) \neq \phi_0(u)$, we decompose

$$\begin{aligned}
Th^{1/2}\hat{Q} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \hat{\phi}_0(u)\right|^2 W(u)du\\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_0(u)\right|^2 W(u)du
\end{aligned}$$

$$
\begin{aligned}
&+h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_0(u)-\phi_0(u)\right|^2 W(u)du\\
&-2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\hat{\phi}_0(u)-\phi_0(u)\right]^*\right\}W(u)du\\
&=\hat{Q}_4+\hat{Q}_5-\hat{Q}_6.
\end{aligned}
$$

For $\hat{Q}_4$, by the triangle inequality,

$$
\begin{aligned}
\hat{Q}_4 &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u)-\phi_t(u)+\phi_t(u)-\phi_0(u)\right|^2 W(u)du\\
&\le 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u)-\phi_t(u)\right|^2 W(u)du+2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\phi_t(u)-\phi_0(u)\right|^2 W(u)du\\
&=\hat{Q}_{41}+\hat{Q}_{42}.
\end{aligned}
$$

Then we investigate $\hat{Q}_{41}$:

$$
\begin{aligned}
\hat{Q}_{41} &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u)-\phi_t(u)\right|^2 W(u)du\\
&\le 4h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varepsilon_s(u)H_{st}\right|^2 W(u)du+4h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\phi_s(u)H_{st}-\phi_t(u)\right|^2 W(u)du\\
&=\hat{Q}_{411}+\hat{Q}_{412}.
\end{aligned}
$$

It is straightforward to see $\hat{Q}_{411} = 4\hat{Q}_1$. By the results in Propositions A.1 and A.2, $\hat{Q}_{411} = 4B + 4U + op(1)$, where $B = Op(h^{-1/2})$. For $\hat{Q}_{412}$, by the Taylor expansion of $\phi_s(u) \equiv \phi(u, \frac{s}{T})$ around t/T, we have

$$
\phi_s(u)=\phi_t(u)+\phi_t'(u)\left(\frac{s-t}{T}\right)+\frac{1}{2}\phi_t''(u)\left(\frac{s-t}{T}\right)^2+o\left[\left(\frac{s-t}{T}\right)\right]^2.
$$

Therefore,

$$
\begin{aligned}
\hat{Q}_{412} &= 4h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\phi_s(u)H_{st}-\phi_t(u)\right|^2 W(u)du\\
&= 4h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{Th}\sum_{j=-\lfloor Th\rfloor}^{\lfloor Th\rfloor}\left[h\phi_t'(u)\left(\frac{j}{Th}\right)+\frac{h^2}{2}\phi_t''(u)\left(\frac{j}{Th}\right)^2+o\left[\left(\frac{j}{T}\right)\right]^2\right]K\left(\frac{j}{Th}\right)\right|^2 W(u)du\\
&\quad+op(1).
\end{aligned}
$$

Let $\eta = \frac{j}{Th}$, since $\lfloor Th\rfloor \le j \le \lfloor Th\rfloor$, then $-1 \le \eta \le 1$ and by the Riemann approximation of an integral

$$
\begin{aligned}
\hat{Q}_{412} &= 4h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\phi_t'(u)\int_{-1}^{1}\eta hK(\eta)\,d\eta+\frac{1}{2}\phi_t''(u)\int_{-1}^{1}\eta^2h^2K(\eta)\,d\eta+o(h^2)+o(1)\right|^2 W(u)du\\
&\quad+op(1)\\
&= O(Th^{9/2}),
\end{aligned}
$$

where the last equality comes from the assumption on kernel function $K(\cdot)$. Therefore, combining the results on $\hat{Q}_{411}$ and $\hat{Q}_{412}$, we have $\hat{Q}_{41} = Op(h^{-1/2})$. For $\hat{Q}_{42}$:

$$
\begin{aligned}
\lim_{T\to\infty} T^{-1}h^{-1/2}\hat{Q}_{42} &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\phi_t(u)-\phi_0(u)\right|^2 W(u)du\\
&= O(1),
\end{aligned}
$$

where we use Assumption 1(iii) and $\int_{R^q}\|u\|_q^4 W(u)du < \infty$. Thus, $\hat{Q}_{42} = O(Th^{1/2})$. Combining $\hat{Q}_{41}$ and $\hat{Q}_{42}$, we have $\hat{Q}_4 = Op(Th^{1/2})$.

Next, we decompose

$$
\begin{aligned}
\hat{Q}_5 &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_0(u)-\phi_0(u)\right|^2 W(u)du \\
&= Th^{1/2}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right|^2 W(u)du + Th^{1/2}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[\phi_t(u)-\phi_0(u)]\right|^2 W(u)du \\
&\quad + 2Th^{1/2}\int_{\mathbb{R}^q} Re\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\frac{1}{T}\sum_{t=1}^{T}[\phi_t(u)-\phi_0(u)]^*\right]W(u)du \\
&= \hat{Q}_{51}+\hat{Q}_{52}+\hat{Q}_{53}.
\end{aligned}
$$

For $\hat{Q}_{51}$, we decompose

$$
\begin{aligned}
\hat{Q}_{51} &= Th^{1/2}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right|^2 W(u)du \\
&= T^{-1}h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}|\varepsilon_t(u)|^2 W(u)du \\
&\quad + T^{-1}h^{1/2}\sum_{s\neq t}\int_{\mathbb{R}^q} Re[\varepsilon_s(u)\varepsilon_t(u)^*]W(u)du \\
&= \hat{Q}_{511}+\hat{Q}_{512}.
\end{aligned}
$$

It is straightforward to show that $E(\hat{Q}_{511}) = O(h^{1/2})$ and $var(\hat{Q}_{511}) = O(h)$. For $\hat{Q}_{512}$, we show

$$
\begin{aligned}
E(\hat{Q}_{512}) &= T^{-1}h^{1/2}\sum_{s\neq t}\int_{\mathbb{R}^q} Re\{E[\varepsilon_s(u)\varepsilon_t(u)^*]\}W(u)du \\
&= 2h^{1/2}\sum_{j=1}^{T-1}\left(1-\frac{j}{T}\right)\int_{\mathbb{R}^q} Re\{E[\varepsilon_1(u)\varepsilon_{1+j}(u)^*]\}W(u)du \\
&= O(h^{1/2}),
\end{aligned}
$$

by the β-mixing condition. And by Lemma A(ii) of Hjellvik et al. (1998), we can show $E(\hat{Q}_{512}^2) = O(h)$. By Chebyshev's inequality, we have $\hat{Q}_{51} = Op(h^{1/2})$.

Next, it is straightforward to see $\hat{Q}_{52} = Op(Th^{1/2})$ given Assumption 1(iii) and Assumption 4. And by triangle inequality,

$$
\begin{aligned}
\hat{Q}_{53} &\leq Th^{1/2}h^{1/2}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right|^2 W(u)du + Th^{1/2}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[\phi_t(u)-\phi_0(u)]\right|^2 W(u)du \\
&= Op(Th^{1/2}).
\end{aligned}
$$

Therefore, it follows that $\hat{Q}_5 = Op(Th^{1/2})$.

At last, by the triangle inequality,

$$
\begin{aligned}
\hat{Q}_6 &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\hat{\phi}_0(u)-\phi_0(u)\right]^*\right\}W(u)du \\
&\leq h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u)-\phi_0(u)\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_0(u)-\phi_0(u)\right|^2 W(u)du \\
&= Op(Th^{1/2}).
\end{aligned}
$$

Therefore, $Th^{1/2}\hat{Q} = Op(Th^{1/2})$ under $\mathbb{H}_A$. Since we have shown that $\hat{B} = Op(h^{-1/2})$ and $\hat{V} = Op(1)$,

$$
\widehat{SQ} = \frac{Th^{1/2}\hat{Q}-\hat{B}}{\sqrt{\hat{V}}} = Op(Th^{1/2}).
$$

Thus,

$$
Pr(\widehat{SQ} > M_T) \to 1,
$$

as $T\to\infty$ under $\mathbb{H}_A$ for any non-stochastic constants $\{M_T = o(Th^{1/2})\}$. ■

Proof of Theorem 3. Under $\mathbb{H}_{A1}: g_t(X_t) - g_0(X_t) = \kappa_T l_1\left(X_t, \frac{t}{T}\right)$, let

$$\begin{aligned}\phi_t(u) - \phi_0(u) &= \int_{\mathbb{R}^q} [g_t(x) - g_0(x)]\, e^{iu'z} dF_{XZ}(x,z)\\ &= \kappa_T \int_{\mathbb{R}^q} l_1\left(X_t, \frac{t}{T}\right) e^{iu'z} dF_{XZ}(x,z)\\ &\equiv \kappa_T \psi_t(u).\end{aligned}$$

Using the same decomposition as in the proof of Theorem 2, we have

$$\begin{aligned}Th^{1/2}\hat{Q} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \hat{\phi}_0(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_0(u)\right|^2 W(u)du\\ &\quad + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_0(u) - \phi_0(u)\right|^2 W(u)du\\ &\quad - 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u) - \phi_0(u)\right]\left[\hat{\phi}_0(u) - \phi_0(u)\right]^*\right\} W(u)du\\ &= \hat{Q}_4 + \hat{Q}_5 + \hat{Q}_6.\end{aligned}$$

We first decompose $\hat{Q}_4$:

$$\begin{aligned}\hat{Q}_4 &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u) + \phi_t(u) - \phi_0(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u)\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\phi_t(u) - \phi_0(u)\right|^2 W(u)du\\ &\quad + 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u) - \phi_t(u)\right][\phi_t(u) - \phi_0(u)]^*\right\} W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u)\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\kappa_T^2\left|\psi_t(u)\right|^2 W(u)du\\ &\quad + 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u) - \phi_t(u)\right]\kappa_T\psi_t(u)^*\right\} W(u)du\\ &= \hat{Q}_{43} + \hat{Q}_{44} + \hat{Q}_{45}.\end{aligned}$$

For $\hat{Q}_{43}$:

$$\begin{aligned}\hat{Q}_{43} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} Y_s e^{iu'Z_s} H_{st} - \phi_t(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} [\phi_s(u) + \varepsilon_s(u)] H_{st} - \phi_t(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} \varepsilon_s(u) H_{st}\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} \phi_s(u) H_{st} - \phi_t(u)\right|^2 W(u)du\end{aligned}$$

$$+h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varepsilon_s(u)H_{st}\right]\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\phi_s(u)H_{st}-\phi_t(u)\right]^*\right\}W(u)du$$
$$=\hat{Q}_{431}+\hat{Q}_{432}+\hat{Q}_{433}.$$

By Proposition A.1, we have $\hat{Q}_{431}=B+U+op(1)$. Given $l_1\left(X_t,\frac{t}{T}\right)$ is twice continuously differentiable with respect to $\frac{t}{T}$, we have

$$\psi_s(u)=\psi_t(u)+\psi_t'(u)\left(\frac{s-t}{T}\right)+\frac{1}{2}\psi_t''(u)\left(\frac{s-t}{T}\right)^2+o\left[\left(\frac{s-t}{T}\right)\right]^2,$$

by second-order Taylor expansion around $\frac{t}{T}$. Therefore,

$$\begin{aligned}
\hat{Q}_{432} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\phi_s(u)H_{st}-\phi_t(u)\right|^2 W(u)du\\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}[\phi_0(u)+\kappa_T\psi_s(u)]H_{st}-[\phi_0(u)+\kappa_T\psi_t(u)]\right|^2 W(u)du\\
&= h^{1/2}\kappa_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\psi_s(u)H_{st}-\psi_t(u)\right|^2 W(u)du\\
&= h^{1/2}\kappa_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\left[\psi_t'(u)\left(\frac{s-t}{T}\right)+\frac{\psi_t''(u)}{2}\left(\frac{s-t}{T}\right)^2+o\left[\left(\frac{s-t}{T}\right)\right]^2\right]H_{st}\right|^2 W(u)du\\
&= h^{1/2}\kappa_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{Th}\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\left[\psi_t'(u)\left(\frac{s-t}{T}\right)+\frac{\psi_t''(u)}{2}\left(\frac{s-t}{T}\right)^2+o\left[\left(\frac{s-t}{T}\right)\right]^2\right]K\left(\frac{s-t}{Th}\right)\right|^2 W(u)du\\
&\quad+op(1)\\
&= h^{1/2}\kappa_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{Th}\sum_{j=-\lfloor Th\rfloor}^{\lfloor Th\rfloor}\left[h\psi_t'(u)\left(\frac{j}{Th}\right)+\frac{h^2}{2}\psi_t''(u)\left(\frac{j}{Th}\right)^2+o\left[\left(\frac{j}{T}\right)\right]^2\right]K\left(\frac{j}{Th}\right)\right|^2 W(u)du\\
&\quad+op(1)\\
&= O(h^4).
\end{aligned}$$

Based on the Taylor expansion, we show that $\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\phi_s(u)H_{st}-\phi_t(u)=O(\kappa_T h^2)$. Then by Cauchy–Schwarz inequality, we can show that $\hat{Q}_{433}=Op(h^{7/4})$. Therefore, $\hat{Q}_{43}=B+U+op(1)$.

$$\begin{aligned}
\hat{Q}_{44} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\kappa_T^2\left|\psi_t(u)\right|^2 W(u)du\\
&= \frac{1}{T}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\psi_t(u)\right|^2 W(u)du\\
&= \int_{\mathbb{R}^q}\int_0^1|\psi(u,\eta)|^2W(u)d\eta du+op(1).
\end{aligned}$$

Next,

$$\begin{aligned}
\hat{Q}_{45} &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_t(u)\right]\kappa_T\psi_t(u)^*\right\}W(u)du\\
&= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)+\phi_0(u)-\phi_t(u)\right]\kappa_T\psi_t(u)^*\right\}W(u)du\\
&= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\kappa_T\psi_t(u)^*\right\}W(u)du
\end{aligned}$$

$$
\begin{aligned}
& -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{[\phi_t(u)-\phi_0(u)]\,\kappa_T\psi_t(u)^*\right\}W(u)du \\
&= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\kappa_T\psi_t(u)^*\right\}W(u)du \\
& -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} |\kappa_T\psi_t(u)|^2\,W(u)du \\
&= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\kappa_T\psi_t(u)^*\right\}W(u)du \\
& -2\int_{\mathbb{R}^q}\int_0^1 |\psi(u,\eta)|^2 W(u)d\eta du + op(1) \\
&= \hat{Q}_{451}+\hat{Q}_{452}+op(1).
\end{aligned}
$$

We now work on $\hat{Q}_5$ and $\hat{Q}_6$. It follows

$$
\begin{aligned}
\hat{Q}_5 &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_0(u)-\phi_0(u)\right|^2 W(u)du \\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}Y_t e^{iu'Z_t}-\phi_0(u)\right|^2 W(u)du \\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[\phi_t(u)+\varepsilon_t(u)]-\phi_0(u)\right|^2 W(u)du \\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[\phi_t(u)-\phi_0(u)+\varepsilon_t(u)]\right|^2 W(u)du \\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[\kappa_T\psi_t(u)+\varepsilon_t(u)]\right|^2 W(u)du \\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right|^2 W(u)du \\
&\quad + 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du \\
&= \hat{Q}_{51}+\hat{Q}_{52}+\hat{Q}_{53}.
\end{aligned}
$$

Given $\kappa_T = T^{-1/2}h^{-1/4}$, we can show

$$
\begin{aligned}
\hat{Q}_{51} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)\right|^2 W(u)du \\
&= \int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\psi_t(u)\right|^2 W(u)du \\
&= \int_{\mathbb{R}^q}\left|\int_0^1 \psi(u,\tau)d\tau\right|^2 W(u)du + o(1).
\end{aligned}
$$

For $\hat{Q}_{52}$, it is identical to $\hat{Q}_2$ and we have shown that $\hat{Q}_2 = op(1)$. Thus $\hat{Q}_{52} = op(1)$. Given $\psi_t(u)$ is non-stochastic and $\kappa_T = T^{-1/2}h^{-1/4}$,

$$\begin{aligned}\hat{Q}_{53} &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du\\ &= 2T^{-1/2}h^{1/4}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\int_0^1\psi(u,\tau)d\tau\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du + op(1)\\ &= 2T^{1/2}h^{1/4}\int_{\mathbb{R}^q} Re\left\{\left[\int_0^1\psi(u,\tau)d\tau\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du + op(1).\end{aligned}$$

The order of magnitude of $\hat{Q}_{53}$ is solely determined by $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)$. It is trivial to show that $\hat{Q}_{53} = Op(h^{1/4}) = op(1)$. Therefore, we have shown that

$$\hat{Q}_5 = \int_{\mathbb{R}^q}\left|\int_0^1\psi(u,\tau)d\tau\right|^2 W(u)du + op(1).$$

Next, for $\hat{Q}_6$:

$$\begin{aligned}\hat{Q}_6 &= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\hat{\phi}_0(u)-\phi_0(u)\right]^*\right\}W(u)du\\ &= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}Y_te^{iu'Z_t}-\phi_0(u)\right]^*\right\}W(u)du\\ &= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)+\varepsilon_t(u)\right]^*\right\}W(u)du\\ &= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)\right]^*\right\}W(u)du\\ &\quad -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du\\ &\equiv \hat{Q}_{61}+\hat{Q}_{62}.\end{aligned}$$

We claim that $\hat{Q}_{451}+\hat{Q}_{61} = op(1)$:

$$\begin{aligned}&2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\kappa_T\psi_t(u)^*\right\}W(u)du + \hat{Q}_{61}\\ &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}(\phi_s(u)+\varepsilon_s(u)-\phi_0(u))H_{st}\right]\left[\kappa_T\psi_t(u)-\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)\right]^*\right\}W(u)du\\ &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}(\kappa_T\psi_s(u)+\varepsilon_s(u))H_{st}\right]\left[\kappa_T\psi_t(u)-\frac{1}{T}\sum_{t=1}^{T}\kappa_T\psi_t(u)\right]^*\right\}W(u)du\\ &= 2h^{1/2}\kappa_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\psi_s(u)H_{st}\right]\left[\psi_t(u)-\frac{1}{T}\sum_{t=1}^{T}\psi_t(u)\right]^*\right\}W(u)du\\ &\quad + 2h^{1/2}\kappa_T\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varepsilon_s(u)H_{st}\right]\left[\psi_t(u)-\frac{1}{T}\sum_{t=1}^{T}\psi_t(u)\right]^*\right\}W(u)du\\ &\equiv R_8+R_9.\end{aligned}$$

By the weighted Cauchy–Schwarz inequality,

$$\begin{aligned} R_8 &\le 2h^{1/2}\kappa_T^2 \int_{\mathbb{R}^q} \left\{ \left[\sum_{t=1}^{T} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \psi_s(u) H_{st} \right|^2 \right]^{1/2} \left[\sum_{t=1}^{T} \left| \psi_t(u) - \frac{1}{T}\sum_{t=1}^{T}\psi_t(u) \right|^2 \right]^{1/2} \right\} W(u)du \\ &\le 2 \left[h^{1/2}\kappa_T^2 \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \psi_s(u) H_{st} \right|^2 W(u)du \right]^{1/2} \\ &\quad \times \left[h^{1/2}\kappa_T^2 \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \psi_t(u) - \frac{1}{T}\sum_{t=1}^{T}\psi_t(u) \right|^2 W(u)du \right]^{1/2} \\ &= O(1) * O(T^{-1/2}) = o(1), \end{aligned}$$

where the last equality needs $|\psi_s(u)|^2 < C$ for all $u \in \mathbb{R}^q$. And by the triangle inequality

$$\begin{aligned} R_9 &\le h^{1/2}\kappa_T \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \varepsilon_s(u) H_{st} \right|^2 W(u)du \\ &\quad + h^{1/2}\kappa_T \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \psi_t(u) - \frac{1}{T}\sum_{t=1}^{T}\psi_t(u) \right|^2 W(u)du \\ &= Op(T^{-1/2}h^{-3/4}) + O(T^{-1/2}h^{1/4}) \\ &= op(1), \end{aligned}$$

where we have used the result

$$h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} \left| \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \varepsilon_s(u) H_{st} \right|^2 W(u)du = Op(h^{-1/2}),$$

obtained in Proposition A.1. Therefore, we have shown that $R_8 + R_9 = op(1)$.

We leave $\hat{Q}_6$ for now and work on $\hat{Q}_4$:

Lastly, we need to show $\hat{Q}_{62} = op(1)$:

$$\begin{aligned} \hat{Q}_{62} &= -2h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} Re \left\{ \left[\hat{\phi}_t(u) - \phi_0(u) \right] \left[\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t(u) \right]^* \right\} W(u)du \\ &= -2h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} Re \left\{ \left[\sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} (\kappa_T \psi_s(u) + \varepsilon_s(u)) H_{st} \right] \left[\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t(u) \right]^* \right\} W(u)du \\ &= -2h^{1/2}\kappa_T \sum_{t=1}^{T} \int_{\mathbb{R}^q} Re \left\{ \left[\sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \psi_s(u) H_{st} \right] \left[\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t(u) \right]^* \right\} W(u)du \\ &\quad - 2h^{1/2} \sum_{t=1}^{T} \int_{\mathbb{R}^q} Re \left\{ \left[\sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \varepsilon_s(u) H_{st} \right] \left[\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t(u) \right]^* \right\} W(u)du \\ &= \hat{Q}_{621} + \hat{Q}_{622}. \end{aligned}$$

Firstly, $\hat{Q}_{621} = Op(h^{1/4})$ since $\sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} \psi_s(u) H_{st} = O(1)$ and $\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t(u) = Op(T^{-1/2})$. For $\hat{Q}_{622}$, we have shown it is $op(1)$ in the proof of Theorem A.3. Thus, $\hat{Q}_{62} = op(1)$.

Combine all terms in $\hat{Q}_4$, $\hat{Q}_5$, and $\hat{Q}_6$,

$$\begin{aligned} Th^{1/2}\hat{Q} &= \hat{Q}_4 + \hat{Q}_5 + \hat{Q}_6 \\ &= B + U + \int_{\mathbb{R}^q} \left| \int_0^1 \psi(u, \tau) d\tau \right|^2 W(u)du - \int_{\mathbb{R}^q} \int_0^1 |\psi(u, \eta)|^2 W(u) d\eta du + op(1). \quad \blacksquare \end{aligned}$$

Proof of Theorem 4. Under $\mathbb{H}_{A2}: g_t(X_t) - g_0(X_t) = a_T l_2\left(X_t, \frac{t/T-\tau_0}{b_T}\right)$, let

$$\begin{aligned}\phi_t(u) - \phi_0(u) &= \int_{\mathbb{R}^q} [g_t(x) - g_0(x)]\, e^{iu'z} dF_{XZ}(x,z)\\ &= a_T \int_{\mathbb{R}^q} l_2\left(X_t, \frac{t/T-\tau_0}{b_T}\right) e^{iu'z} dF_{XZ}(x,z)\\ &\equiv a_T \vartheta_t(u).\end{aligned}$$

Using the same decomposition as in the proof of Theorem 3, we decompose $Th^{1/2}\hat{Q}$ to $\hat{Q}_4$, $\hat{Q}_5$, and $\hat{Q}_6$. We first investigate $\hat{Q}_4$:

$$\begin{aligned}\hat{Q}_4 &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u) + \phi_t(u) - \phi_0(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u)\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\phi_t(u) - \phi_0(u)\right|^2 W(u)du\\ &\quad + 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u) - \phi_t(u)\right][\phi_t(u) - \phi_0(u)]^*\right\} W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u)\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} a_T^2\left|\varsigma_t(u)\right|^2 W(u)du\\ &\quad + 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u) - \phi_t(u)\right] a_T\varsigma_t(u)^*\right\} W(u)du\\ &= \hat{Q}_{43} + \hat{Q}_{44} + \hat{Q}_{45}.\end{aligned}$$

For $\hat{Q}_{43}$:

$$\begin{aligned}\hat{Q}_{43} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_t(u) - \phi_t(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} Y_s e^{iu'Z_s} H_{st} - \phi_t(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} [\phi_s(u) + \varepsilon_s(u)] H_{st} - \phi_t(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} \varepsilon_s(u) H_{st}\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} \phi_s(u) H_{st} - \phi_t(u)\right|^2 W(u)du\\ &\quad + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} \varepsilon_s(u) H_{st}\right]\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} \phi_s(u) H_{st} - \phi_t(u)\right]^*\right\} W(u)du\\ &= \hat{Q}_{431} + \hat{Q}_{432} + \hat{Q}_{433}.\end{aligned}$$

By Proposition A.1, we have $\hat{Q}_{431} = B + U + op(1)$. Given $l_2\left(X_t, \frac{t/T-\tau_0}{b_T}\right)$ is twice continuously differentiable with respect to $\frac{t}{T}$, we have

$$\varsigma_s(u) = \varsigma_t(u) + \varsigma_t'(u)\left(\frac{s-t}{Tb_T}\right) + \frac{1}{2}\varsigma_t''(u)\left(\frac{s-t}{Tb_T}\right)^2 + o\left[\left(\frac{s-t}{Tb_T}\right)\right]^2,$$

by second-order Taylor expansion around $\frac{t}{T}$. Therefore,

$$\begin{aligned}\hat{Q}_{432} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} \phi_s(u) H_{st} - \phi_t(u)\right|^2 W(u)du\\ &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor} [\phi_0(u) + a_T\varsigma_s(u)] H_{st} - [\phi_0(u) + a_T\varsigma_t(u)]\right|^2 W(u)du\end{aligned}$$

$$
\begin{aligned}
&= h^{1/2}a_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varsigma_s(u)H_{st}-\varsigma_t(u)\right|^2 W(u)du\\
&= h^{1/2}a_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\left[\varsigma_t'(u)\left(\frac{s-t}{Tb_T}\right)+\frac{1}{2}\varsigma_t''(u)\left(\frac{s-t}{Tb_T}\right)^2+o\left[\left(\frac{s-t}{Tb_T}\right)\right]^2\right]H_{st}\right|^2 W(u)du\\
&= h^{1/2}a_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{Th}\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\left[\varsigma_t'(u)\left(\frac{s-t}{Tb_T}\right)+\frac{1}{2}\varsigma_t''(u)\left(\frac{s-t}{Tb_T}\right)^2+o\left[\left(\frac{s-t}{Tb_T}\right)\right]^2\right]K\left(\frac{s-t}{Th}\right)\right|^2 W(u)du\\
&\quad+op(1)\\
&= h^{1/2}a_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{Th}\sum_{j=-\lfloor Th\rfloor}^{\lfloor Th\rfloor}\left[\frac{h}{b_T}\varsigma_t'(u)\left(\frac{j}{Th}\right)+\frac{h^2}{2b_T^2}\varsigma_t''(u)\left(\frac{j}{Th}\right)^2+o\left[\left(\frac{j}{T}\right)\right]^2\right]K\left(\frac{j}{Th}\right)\right|^2 W(u)du\\
&\quad+op(1)\\
&= O(Ta_T^2h^{9/2}b_T^{-4}).
\end{aligned}
$$

By similar arguments as in Theorem 3, $\hat{Q}_{43}=B+U+op(1)$.

$$
\begin{aligned}
\hat{Q}_{44} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}a_T^2\,|\varsigma_t(u)|^2\,W(u)du\\
&= \frac{1}{Tb_T}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\varsigma\left(u,\frac{t/T-\tau_0}{b_T}\right)\right|^2 W(u)du\\
&= \int_{\mathbb{R}^q}\int_{\mathbb{R}}|\varsigma(u,\eta)|^2W(u)d\eta du+op(1).
\end{aligned}
$$

Next,

$$
\begin{aligned}
\hat{Q}_{45} &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}Re\left\{\left[\hat{\phi}_t(u)-\phi_t(u)\right]a_T\varsigma_t(u)^*\right\}W(u)du\\
&= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}[\phi_s(u)+\varepsilon_s(u)]H_{st}-\phi_t(u)\right]a_T\varsigma_t(u)^*\right\}W(u)du\\
&= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}[a_T\varsigma_s(u)+\varepsilon_s(u)]H_{st}-a_T\varsigma_t(u)\right]a_T\varsigma_t(u)^*\right\}W(u)du\\
&= 2h^{1/2}a_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}Re\left\{\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}[\varsigma_s(u)-\varsigma_t(u)]H_{st}\varsigma_t(u)^*\right\}W(u)du\\
&\quad+2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}Re\left\{\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}[\varepsilon_s(u)H_{st}]a_T\varsigma_t(u)^*\right\}W(u)du\\
&= \hat{Q}_{453}+\hat{Q}_{454}.
\end{aligned}
$$

Given the Taylor expansion, we have $\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varsigma_s(u)H_{st}-\varsigma_t(u)=O(h^2b_T^{-2})$. Therefore, it is straightforward to show $\hat{Q}_{453}=Op(a_Th^2b_T^{-2})=op(1)$ via the Riemann approximation of an integral. Similarly, we can show $\hat{Q}_{454}=Op(b_Th^{-1/4})=op(1)$. Therefore $\hat{Q}_{45}=op(1)$.

We now work on $\hat{Q}_5$ and $\hat{Q}_6$. It follows

$$
\begin{aligned}
\hat{Q}_5 &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\hat{\phi}_0(u)-\phi_0(u)\right|^2 W(u)du\\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[\phi_t(u)+\varepsilon_t(u)]-\phi_0(u)\right|^2 W(u)du
\end{aligned}
$$

$$
\begin{aligned}
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[\phi_t(u)-\phi_0(u)+\varepsilon_t(u)]\right|^2 W(u)du \\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}[a_T\varsigma_t(u)+\varepsilon_t(u)]\right|^2 W(u)du \\
&= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}a_T\varsigma_t(u)\right|^2 W(u)du + h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right|^2 W(u)du \\
&\quad + 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\frac{1}{T}\sum_{t=1}^{T}a_T\varsigma_t(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du \\
&= \hat{Q}_{51}+\hat{Q}_{52}+\hat{Q}_{53}.
\end{aligned}
$$

Given $a_T^2 b_T = T^{-1}h^{-1/2}$, we can show

$$
\begin{aligned}
\hat{Q}_{51} &= h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{T}\sum_{t=1}^{T}a_T\varsigma_t(u)\right|^2 W(u)du \\
&= a_T^2 h^{1/2} b_T^2\sum_{t=1}^{T}\int_{\mathbb{R}^q}\left|\frac{1}{Tb_T}\sum_{t=1}^{T}\varsigma_t(u)\right|^2 W(u)du \\
&= b_T\int_{\mathbb{R}^q}\left|\frac{1}{Tb_T}\sum_{t=1}^{T}\varsigma_t(u)\right|^2 W(u)du \\
&= b_T\int_{\mathbb{R}^q}\left|\int_{\mathbb{R}}\varsigma(u,\eta)d\eta\right|^2 W(u)du + o(1) \\
&= O(b_T).
\end{aligned}
$$

For $\hat{Q}_{52}$, it is identical to $\hat{Q}_2$ and we have shown that $\hat{Q}_2 = op(1)$. Thus $\hat{Q}_{52} = op(1)$. Given $\varsigma_t(u)$ is non-stochastic and $a_T^2 b_T = T^{-1}h^{-1/2}$,

$$
\begin{aligned}
\hat{Q}_{53} &= 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\frac{1}{T}\sum_{t=1}^{T}a_T\varsigma_t(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du \\
&= 2h^{1/2}a_T b_T\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\int_{\mathbb{R}}\varsigma(u,\eta)d\eta\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du + op(1) \\
&= 2Th^{1/2}a_T b_T\int_{\mathbb{R}^q} Re\left\{\left[\int_0^1\psi(u,\tau)d\tau\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du + op(1).
\end{aligned}
$$

The order of magnitude of $\hat{Q}_{53}$ is solely determined by $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)$. It is trivial to show that $\hat{Q}_{53} = Op(h^{1/4}b_T^{1/2}) = op(1)$. Therefore, we have shown that $\hat{Q}_5 = op(1)$. Finally, for $\hat{Q}_6$,

$$
\begin{aligned}
\hat{Q}_6 &= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\hat{\phi}_t(u)-\phi_0(u)\right]\left[\hat{\phi}_0(u)-\phi_0(u)\right]^*\right\}W(u)du \\
&= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}Y_s e^{iu'Z_s}H_{st}-\phi_0(u)\right]\left[\frac{1}{T}\sum_{t=1}^{T}Y_t e^{iu'Z_t}-\phi_0(u)\right]^*\right\}W(u)du \\
&= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}a_T\varsigma_s(u)+\varepsilon_s(u)\right]H_{st}\left[\frac{1}{T}\sum_{t=1}^{T}a_T\varsigma_t(u)+\varepsilon_t(u)\right]^*\right\}W(u)du \\
&= -2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}a_T\varsigma_s(u)H_{st}\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du
\end{aligned}
$$

$$
\begin{aligned}
&- 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varepsilon_s(u)H_{st}\right]\left[\frac{1}{T}\sum_{t=1}^{T}a_T\varsigma_t(u)\right]^*\right\}W(u)du \\
&- 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varepsilon_s(u)H_{st}\right]\left[\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t(u)\right]^*\right\}W(u)du \\
&- 2h^{1/2}\sum_{t=1}^{T}\int_{\mathbb{R}^q} Re\left\{\left[\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}a_T\varsigma_s(u)H_{st}\right]\left[\frac{1}{T}\sum_{t=1}^{T}a_T\varsigma_t(u)\right]^*\right\}W(u)du \\
= &\ \hat{Q}_{63}+\hat{Q}_{64}+\hat{Q}_{65}+\hat{Q}_{66}.
\end{aligned}
$$

Given $\frac{1}{T}\sum_{t=1}^{T}\varsigma_t(u) = O(a_T b_T)$, $\sum_{s=t-\lfloor Th\rfloor}^{t+\lfloor Th\rfloor}\varsigma_s(u)H_{st} - \varsigma_t(u) = O(a_T h^2 b_T^{-2})$, we can show $\hat{Q}_{63} = Op(b_T^{1/2}h^{1/4})$, $\hat{Q}_{64} = Op(b_T^{1/2}h^{-1/4})$, and $\hat{Q}_{65} = Op(h^{1/4})$. Combine all terms in $\hat{Q}_4$, $\hat{Q}_5$, and $\hat{Q}_6$,

$$
\begin{aligned}
Th^{1/2}\hat{Q} &= \hat{Q}_4+\hat{Q}_5+\hat{Q}_6 \\
&= B+U+\int_{\mathbb{R}^q}\int_{\mathbb{R}}|\varsigma(u,\eta)|^2W(u)d\eta du+op(1). \qquad \blacksquare
\end{aligned}
$$

Appendix B. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jeconom.2018.12.014.

References

Ai, C., Chen, X., 2003. Efficient estimation of models with conditional moment restrictions containing unknown functions. Econometrica 71 (6), 1795–1843.
Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59 (3), 817–858.
Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61 (4), 821–856.
Andrews, D.W.K., Fair, R.C., 1988. Inference in nonlinear econometric models with structural change. Rev. Econom. Stud. 55 (4), 615–640.
Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62 (6), 1383–1414.
Bai, J., Ng, S., 2008. Forecasting economic time series using targeted predictors. J. Econometrics 146 (2), 304–317.
Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica 66 (1), 47–78.
Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. J. Appl. Econometrics 18 (1), 1–22.
Bierens, H.J., 1982. Consistent model specification tests. J. Econometrics 20 (1), 105–134.
Blundell, R., Chen, X., Kristensen, D., 2007. Semi-Nonparametric iv Estimation of Shape-Invariant Engel Curves. Econometrica 75 (6), 1613–1669.
Cai, Z., 2007. Trending time-varying coefficient time series models with serially correlated errors. J. Econometrics 136 (1), 163–188.
Cai, Z., Fan, J., Yao, Q., 2000. Functional-coefficient regression models for nonlinear time series. J. Amer. Statist. Assoc. 95 (451), 941–956.
Cai, Z., Wang, Y., Wang, Y., 2015. Testing instability in a predictive regression model with nonstationary regressors. Econom. Theory 31 (5), 953–980.
Campbell, J.Y., Thompson, S.B., 2007. Predicting excess stock returns out of sample: can anything beat the historical average? Rev. Financ. Stud. 21 (4), 1509–1531.
Chen, B., 2015. Modeling and testing smooth structural changes with endogenous regressors. J. Econometrics 185 (1), 196–215.
Chen, B., Hong, Y., 2012. Testing for smooth structural changes in time series models via nonparametric regression. Econometrica 80 (3), 1157–1183.
Chow, G.C., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica 28 (3), 591–605.
Clements, M.P., Franses, P.H., Swanson, N.R., 2004. Forecasting economic and financial time-series with non-linear models. Int. J. Forecast. 20 (2), 169–183.
Dahlhaus, R., 1996. On the kullback-leibler information divergence of locally stationary processes. Stochastic Process. Appl. 62 (1), 139–168.
Darolles, S., Fan, Y., Florens, J.P., Renault, E., 2011. Nonparametric instrumental regression. Econometrica 79 (5), 1541–1565.
De Gooijer, J.G., Kumar, K., 1992. Some recent developments in non-linear time series modelling, testing, and forecasting. Int. J. Forecast. 8 (2), 135–156.
Elliott, G., Müller, U.K., 2006. Efficient tests for general persistent time variation in regression coefficients. Rev. Econom. Stud. 73 (4), 907–940.
González-Rivera, G., Lee, T.-H., 2009. Nonlinear Time Series in Financial Forecasting. Springer, New York.
Hall, A.R., Han, S., Boldea, O., 2012. Inference regarding multiple structural changes in linear models with endogenous regressors. J. Econometrics 170 (2), 281–302.
Hall, P., Horowitz, J.L., 2005. Nonparametric methods for inference in the presence of instrumental variables. Ann. Statist. 33 (6), 2904–2929.
Hall, P., Wehrly, T.E., 1991. A geometrical method for removing edge effects from kernel-type nonparametric regression estimators. J. Amer. Statist. Assoc. 86 (415), 665–672.
Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57 (2), 357–384.
Hansen, B.E., 2001. The new econometrics of structural change: dating breaks in U.S. labor productivity. J. Econ. Perspect. 15 (4), 117–128.
Hausman, J.A., 1978. Specification tests in econometrics. Econometrica 46 (6), 1251–1271.
Hidalgo, J., 1995. A nonparametric conditional moment test for structural stability. Econom. Theory 11 (4), 671–698.
Hillebrand, E., Lee, T.-H., Medeiros, M., Bagging constrained forecasts with application to forecasting equity premium, in: JSM Proceedings for Business and Economic Statistics, 2009.
Hjellvik, V., Yao, Q., Tjøstheim, D., 1998. Linearity testing using local polynomial approximation. J. Statist. Plann. Inference 68 (2), 295–321.
Hong, Y., Wang, X., Wang, S., 2017. Testing strict stationarity with applications to macroeconomic time series. Internat. Econom. Rev. (forthcoming).
Horowitz, J.L., 2011. Applied nonparametric instrumental variables estimation. Econometrica 79 (2), 347–394.
Kristensen, D., 2012. Non-parametric detection and estimation of structural change. Econom. J. 15 (3), 420–461.
Lahiri, S.N., 1999. Theoretical comparisons of block bootstrap methods. Ann. Statist. 27 (1), 386–404.
Lee, T.-H., Tu, Y., Ullah, A., 2014. Nonparametric and semiparametric regressions subject to monotonicity constraints: estimation and forecasting. J. Econometrics 182 (1), 196–210.
Lima, L.R., Xiao, Z., 2010. Is there long memory in financial time series? Appl. Financial Econ. 20 (6), 487–500.
Lin, C.-F.J., Teräsvirta, T., 1994. Testing the constancy of regression parameters against continuous structural change. J. Econometrics 62 (2), 211–228.

Newey, W.K., Powell, J.L., 2003. Instrumental variable estimation of nonparametric models. Econometrica 71 (5), 1565–1578.
Newey, W.K., Powell, J.L., Vella, F., 1999. Nonparametric estimation of triangular simultaneous equations models. Econometrica 67 (3), 565–603.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.
Perron, P., Yamamoto, Y., 2014. A note on estimating and testing for multiple structural changes in models with endogenous regressors via 2SLS. Econometric Theory 30 (2), 491–507.
Perron, P., Yamamoto, Y., 2015. Using ols to estimate and test for structural changes in models with endogenous regressors. J. Appl. Econometrics 30 (1), 119–144.
Politis, D.N., White, H., 2004. Automatic block-length selection for the dependent bootstrap. Econometric Rev. 23 (1), 53–70.
Qu, Z., Perron, P., 2007. Estimating and testing structural changes in multivariate regressions. Econometrica 75 (2), 459–502.
Robinson, P.M., 1989. Nonparametric estimation of time-varying parameters. In: Hackl, P. (Ed.), Statistical Analysis and Forecasting of Economic Structural Change. Springer, Berlin, Heidelberg, pp. 253–264.
Rosenblatt, M., 1975. A quadratic measure of deviation of two-dimensional density estimates and a test of independence. Ann. Statist. 3 (1), 1–14.
Stock, J., Watson, M., 2009. Forecasting in Dynamic Factor Models Subject to Structural Instability. Oxford University Press, pp. 1–57.
Su, L., White, H., 2007. A consistent characteristic function-based test for conditional independence. J. Econometrics 141 (2), 807–834.
Su, L., White, H., 2010. Testing structural change in partially linear models. Econom. Theory 26 (06), 1761–1806.
Su, L., Xiao, Z., 2008. Testing structural change in time-series nonparametric regression models. Stat. Interface 1, 347–366.
Székely, G.J., Rizzo, M.L., Bakirov, N.K., 2007. Measuring and testing dependence by correlation of distances. Ann. Statist. 35 (6), 2769–2794.
Teräsvirta, T., 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. J. Amer. Statist. Assoc. 89 (425), 208–218.
Teräsvirta, T., Tjøstheim, D., Granger, C.W.J., 2010. Modelling Nonlinear Economic Time Series. Oxford University Press Oxford.
Tong, H., 1990. Nonlinear Time Series: A Dynamical System Approach. Oxford University Press.
Vogt, M., 2015. Testing for structural change in time-varying nonparametric regression models. Econom. Theory 31 (04), 811–859.
Wang, X., Hong, Y., 2017. Characteristic function based testing for conditional independence: a nonparametric regression approach. Econom. Theory 1–35.
Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financial Stud. 21 (4), 1455–1508.
White, H., 1989. Some asymptotic results for learning in single hidden-layer feedforward network models. J. Amer. Statist. Assoc. 84 (408), 1003–1013.
Yoshihara, K.-i., 1976. Limiting behavior of u-statistics for stationary, absolutely regular processes. Probab. Theory Related Fields 35 (3), 237–252.
Zhang, T., Wu, W.B., 2012. Inference of time-varying regression models. Ann. Statist. 40 (3), 1376–1402.