Consistent Testing for Structural Changes in Time Series Models via Discrete Fourier Transform*

Zhonghao Fu[a], Yongmiao Hong[b], Xia Wang[c]

[a] Fudan University
[b] Cornell University and Xiamen University
[c] Sun Yat-Sen University

May 4, 2019

Abstract

We propose Kolmogorov-Smirnov and Cramér-von Mises type tests for structural changes in linear time series models via the Discrete Fourier Transform (DFT). If we ignore the potential structural changes in coefficients, then the OLS or 2SLS estimators for linear time series models are inconsistent. The residuals will contain information about parameters' instability. Therefore, the basic idea of the proposed tests is to take DFT of residuals and to capture such information in the frequency domain. The proposed tests are not only powerful against smooth structural changes and abrupt structural breaks, but also could detect a class of local alternatives at the parametric rate, which is asymptotically more efficient than the existing consistent tests. Furthermore, our tests are robust to unknown structural changes in explanatory variables and instrumental variables, which makes our test widely applicable, especially in macroeconomic models. Simulation studies demonstrate excellent finite sample performance of our tests. In an application to Taylor rules, we find significant evidence of structural changes during the post-1979 period, which is treated as a stable period by Clarida et al. (2000).

Keywords: Discrete Fourier transform, Endogeneity, Structural changes, Time series model, Kolmogorov-Smirnov test, Cramér-von Mises test.

JEL classification: C12; C14

*Fu is from School of Economics, Fudan University, and Shanghai Institute of International Finance and Economics (SIIFE), Email: *zhfu@fudan.edu.cn*; Hong is from Department of Economics and Department of Statistics and Data Science, Cornell University, and Wang Yanan Institute for Studies in Economics (WISE) and MOE Key Laboratory of Econometrics, Xiamen University, Email: *yh20@cornell.edu*; Wang is from Lingnan College, Sun Yat-sen University, Email: *wangxia25@mail.sysu.edu.cn*. We thank Haiqiang Chen, Liyuan Cui, Jiti Gao, Kristoffer Nimark, and seminar participants in the econometrics Brown-bag seminar at Xiamen Univerisity, the econometrics seminar at Cornell University, and the CUHK workshop on econometrics for their helpful comments and suggestions. Address correspondence to Yongmiao Hong, e-mail: *yh20@cornell.edu*; All remaining errors are solely ours.

1 Introduction

Testing for structural change in a time series regression model has drawn much attention in the literature, since Chow (1960) and Quant (1960). A time series dynamics usually suffers from abrupt structural breaks or smooth structural changes due to varying economic environments, e.g., shocks, policy shifts, technology progress, preference changes, etc. Many empirical studies have confirmed the prevalence of structural instability in financial and macroeconomic time series. For example, Welch and Goyal (2008) show that the in-sample significant predictability of most financial and macroeconomic variables fails to yield better out-of-sample forecasts of the U.S. equity premium than the simple historical mean of equity returns. One possible reason is the instability of model parameters and it is confirmed by Chen and Hong (2012). In labor economics, Hansen (2001) finds strong evidence of a structural break in labor productivity between 1992 and 1996, and weaker evidence of a structural break in the 1960s and the early 1980s. In macroeconomics, Stock and Watson (1996) find substantial instability in 76 representative US monthly post-war macroeconomic time series. Zhang et al. (2008) show that the instability of parameters in the New Keynesian Phillips Curve results in conflicting conclusions about the key determinant of the short-run inflation dynamics. In the current big data era, the abundance of data is not necessarily equivalent to sufficiency of relevant information. If the underlying DGP changes over time, big data may also result in misleading conclusions. Therefore, detecting the existence of structural change is crucial for econometric modeling and inference.

Although most existing studies focus on dealing with abrupt structural breaks (e.g., see Perron, 2006), estimation and testing with smooth structural changes have drawn increasing attention. Intuitively, it is quite likely that economic agents digest and react to shocks (e.g., policies shifts or unexpected income) in a gradual manner. Even when the change is abrupt at the individual level, it is likely to behave as a smooth change at the aggregate level. As a result, the parameters in a time series model usually change smoothly over time rather than shift abruptly. Smooth structural changes can be modeled parametrically. One example is the Smooth Transition Regression (STR) model developed by Lin and Teräsvirta (1994), where they choose a particular parametric function to model the time-varying parameters. However, economic theories usually do not indicate how parameters evolve over time. To avoid misspecification, a nonparametric time-varying parameter model was introduced by Robinson (1989, 1991) and further studied by Orbe et al. (2000, 2005), Cai (2007), Chen and Hong (2012), Kristensen (2012), Zhang and Wu (2012), Cai et al. (2015), and Xu (2015). Among many others, Chen and Hong (2012) propose a generalized Hausman's test for both smooth structural changes and abrupt structural breaks in a linear time series regression model. Kristensen (2012) considers estimation and testing in both mean and variance in a time series dynamics. Cai et al. (2015) test the instability of model parameter when the covariates follow a unit root process. Xu (2015) constructs a CUSUM-type test for smooth structural changes in regression coefficient when the variance is time-varying. Although the existing tests can detect the instability of parameters of unknown forms in a linear time series model, they are restricted in a conditional mean framework. When endogenous covariates are present, their tests are not applicable.

Several recent works have considered estimation and testing in a linear time series model with endogenous covariates. Hall et al. (2012) extend Bai and Perron's (1998) approach to estimation and testing for abrupt breaks of linear models with endogeneity using the Two-Stage Least Squares (2SLS). However, their approach requires firstly identifying the structural breaks in the reduced form. If there exist too many breaks or smooth structural changes in the reduced form, then their approach is not applicable. Perron and Yamamoto (2014) consider estimation and testing based on the 2SLS by extending Perron and Qu (2006). Perron and Yamamoto (2015) propose a testing method for abrupt structural breaks using OLS rather than 2SLS. They show the OLS-based test is more powerful than tests based on the 2SLS in most cases. However, the inference is restricted to the dates and magnitudes of breaks in the reduced form, and their test is not consistent against local alternatives of certain directions. Unlike the studies that focus on abrupt structural breaks, Chen (2015) proposes a Two-Stage Local Linear (2SLL) method for estimation and testing when the unknown parameters change smoothly over time. However, the test statistic is computed via smoothed nonparametric estimation which requires choosing bandwidths for both the structural function and the first stage reduced form. Furthermore, certain smoothness condition is needed to ensure the consistency of smoothed nonparametric estimation, which is restrictive since we usually have no prior knowledge about the property of unknown parameters over time. Moreover, the rate of alternatives that Chen (2015) can detect is $T^{-1/2}h^{-1/4}$, where $h \to 0$ is a bandwidth. This is slower than the parametric rate $T^{-1/2}$.

In this paper, we propose a novel test for structural change in a time series model via DFT. Unlike the existing tests that focus on time domain analysis, we investigate structural stability in the frequency domain. The intuition is straightforward: if structural change exists, then estimation methods based on the whole sample will miss it because the OLS or 2SLS estimator cannot capture the local behavior of the parameters at each time point. Consequently, the estimated residuals will contain such information. By projecting the estimated residuals on the frequency domain via DFT, we can infer the existence of structural change by examining the corresponding spectrum at each frequency. The merits of our frequency domain based approach are as follows. First, our test is consistent against both abrupt structural breaks and smooth structural changes because the Fourier transform can capture all the time series property of the unknown parameter. Second, compared to the consistent tests for smooth structural changes (e.g., Chen and Hong, 2012; Zhang and Wu, 2012; Cai et al., 2015; Chen, 2015), our test avoids smoothed nonparametric estimation of the unknown parameter. As a result, our test is tuning parameter-free and can detect a class of local alternatives at a faster rate. Third, we allow for endogenous covariates, and our test can be viewed as a unified framework. Moreover, our test is robust to possible structural changes in both the covariates and instruments. This is an improvement over Hall et al. (2012) and Perron and Yamamoto (2014), where the instability in the first stage reduced form has a nontrivial impact on testing and estimation. In addition, compared to the existing tests for abrupt structural breaks, e.g., Andrews' (1993) supremum test and Bai and Perron's (1998) double maximum test, we do not require trimming of the boundary region. Furthermore, unlike the existing literature that cannot distinguish structural change from model misspecification, our novel testing via the DFT is robust to model misspecification in certain cases. Therefore, the only source of rejection for our test will be the structural change.

The rest of this paper is organized as follows. In Section 2, we propose the test based on the DFT in a time series regression model with exogenous covariates. In Section 3, we provide the asymptotic theory. In Section 4, we extend our test by adding endogenous covariates. In Section 5, we discuss the robustness of our tests under misspecified models. In Section 6, we show the finite sample performance of our tests via a simulation study. In Section 7, we examine the stability of the U.S. Taylor rule. We then provide a concluding remark and discuss the possible extensions of our tests to other frameworks in Section 8.

2 Framework and Approach

In this paper, we consider testing for structural change of unknown form in the following linear time series regression model:

$$Y_t = X_t'\beta_t + \varepsilon_t, \tag{2.1}$$

where $\{X_t, Y_t\}_{t=1}^T$ is a $\mathbb{R}^d \times \mathbb{R}$-valued observable random sample, $\beta_t \in \Theta$ is a $\mathbb{R}^d$-valued unknown parameter vector that may vary with time, Θ denotes the parameter space which is bounded, and ε_t is an unobservable disturbance such that $E(\varepsilon_t|X_t) = 0$.

The hypothesis of interest is

$$\mathbb{H}_0 : \beta_t = \beta_0 \text{ for some unknown } \beta_0 \in \Theta \text{ and for all } t = 1, 2, ..., T,$$

against

$$\mathbb{H}_A : \beta_t \text{ is a time-varying parameter.}$$

Under $\mathbb{H}_0$, the true model parameter is equal to β_0, which implies there is no structural change. While under $\mathbb{H}_A$, there exist abrupt structural breaks if β_t is a step function of time; and there exist smooth structural changes if $\beta_t \equiv \beta(\frac{t}{T})$ is a smooth function of the rescaled time $\frac{t}{T} \in (0, 1]$.

A straightforward way of testing $\mathbb{H}_0$ is via the generalized Hausman's test (Hausman, 1978), which is comparing the consistent estimates for β_t under the null and alternative hypothesis, e.g., Chen and Hong (2012), Kristensen (2012), Zhang and Wu (2012), and Cai et al. (2015). Because the two estimators converge to the same probability limit only under $\mathbb{H}_0$. However, estimating the unrestricted model requires tuning parameters, i.e., a bandwidth. Even though the choice of a bandwidth has a trivial impact asymptotically, it is crucial in a finite sample. It is possible that two practitioners may obtain conflicting results by using two different bandwidths. To avoid the undesirable features of smoothed nonparametric estimation, we propose a novel test based on the DFT.

Throughout this paper, $\mathbf{i}$ denotes the imaginary number such that $\mathbf{i} = \sqrt{-1}$, A^* denotes the complex conjugate of A, $\text{Re}(A)$ means the real part of A, and $\|\cdot\|$ denotes the Euclidean norm. We further let $X_t(u) \equiv X_t e^{\mathbf{i}u2\pi t/T}$ be the product of a random variable X_t and the Fourier basis function of time, where $X_t(0) = X_t$. Furthermore, we let $\hat{Q}_{xx}(u) \equiv \frac{1}{T}\sum_{t=1}^T X_t(u)X_t'$ be a complex-valued random matrix of u that depends on sample size T. When $u = 0$, we have $\hat{Q}_{xx}(0) \equiv \hat{Q}_{xx} = \frac{1}{T}\sum_{t=1}^{T} X_t X_t'$. And we let $Q_{xx}(u) \equiv \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E(X_t X_t') e^{\mathbf{i}u2\pi t/T}$ be the probability limit for all $u \in \mathbb{R}$. In particular, when X_t is weakly stationary, $Q_{xx}(u) = E(X_t X_t') \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau$. Here we allow X_t to be nonstationary such that it can change smoothly or shift abruptly. The nonstationarity here is similar to the asymptotical mse-stationarity in Hansen (2000).

Definition 2.1 (Hansen, 2000) *An array $\{X_t\}_{t=1}^{T}$ is asymptotically mse-stationary if*

$$\frac{1}{T}\sum_{t=1}^{[\tau T]} X_t X_t' \xrightarrow{p} \tau Q_{xx},$$

for $\tau \in (0, 1]$, where $Q_{xx} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E(X_t X_t')$.

The asymptotical mse-stationarity allows for certain deterministic cyclic process, seasonal dummies, and 'stationary' forms of heteroscedasticity. In this paper, our definition of nonstationarity is weaker than the asymptotical mse-stationarity since we do not need the convergence to hold for all $\tau \in [0, 1]$. Our analysis does not involve any sample splitting, and we only require it to hold in the whole sample, i.e., $\tau = 1$. Therefore, we can allow X_t to have structural breaks of unknown forms, and that is an improvement over the literature.

The idea of our test is to capture the time-varying feature of β_t without estimating it directly. However, β_t is incorporated in Y_t through X_t. Given we do not restrict X_t to be stationary, structural changes in Y_t could result from the time-varying marginal distribution of X_t. So we need first to purge that from Y_t via the OLS.

Consider the OLS estimator $\hat{\beta}$:

$$\begin{aligned}
\hat{\beta} &= \hat{Q}_{xx}^{-1} \frac{1}{T}\sum_{t=1}^{T} X_t Y_t \\
&= \hat{Q}_{xx}^{-1} \frac{1}{T}\sum_{t=1}^{T} X_t X_t' \beta_t + \hat{Q}_{xx}^{-1} \frac{1}{T}\sum_{t=1}^{T} X_t \varepsilon_t \\
&\xrightarrow{p} \tilde{\beta},
\end{aligned}$$

where

$$\tilde{\beta} \equiv Q_{xx}^{-1} \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E(X_t X_t')\beta_t$$

is a weighted average of β_t. In particular, when X_t is weakly stationary, $\tilde{\beta} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \beta_t$. Since the probability limit of the OLS estimator $\hat{\beta}$ is a constant vector, it cannot capture the time-varying feature of β_t at each time t. Therefore, the estimated residuals contain the behavior of β_t over time. Consider the following decomposition of $\hat{\varepsilon}_t$:

$$\begin{aligned}
\hat{\varepsilon}_t &= Y_t - X_t'\hat{\beta} \\
&= X_t'(\beta_t - \hat{\beta}) + \varepsilon_t
\end{aligned}$$

$$= \quad \varepsilon_t - X_t'\hat{Q}_{xx}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} X_t\varepsilon_t\right) + X_t'\left[\beta_t - \hat{Q}_{xx}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} X_tX_t'\beta_t\right)\right],$$

where $\hat{\varepsilon}_t$ is decomposed into three components: the disturbance ε_t, the estimation uncertainty, and the time-varying behavior of β_t which is equal to 0 under $\mathbb{H}_0$. Under $\mathbb{H}_A$, we can extract β_t's local feature over time via DFT without estimating it directly.

Consider the following complex-valued empirical process:

$$\hat{A}(u) = \frac{1}{T}\sum_{t=1}^{T} X_t\hat{\varepsilon}_t e^{\mathbf{i}u2\pi t/T}, \tag{2.2}$$

which is the DFT of $X_t\hat{\varepsilon}_t$. By the definition of $\hat{\varepsilon}_t$,

$$\begin{aligned}
\hat{A}(u) &= \frac{1}{T}\sum_{t=1}^{T} X_t\hat{\varepsilon}_t e^{\mathbf{i}u2\pi t/T} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left[X_t(u) - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}X_t\right](X_t'\beta_t + \varepsilon_t) \\
&= \frac{1}{T}\sum_{t=1}^{T}\hat{M}_t(u)X_t'\beta_t + \frac{1}{T}\sum_{t=1}^{T}\hat{M}_t(u)\varepsilon_t \\
&\equiv \hat{A}_1(u) + \hat{A}_2(u),
\end{aligned}$$

where we let

$$\hat{M}_t(u) = X_t e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}X_t$$

be a $d \times 1$ complex-valued process of u and it has the following property:

$$\frac{1}{T}\sum_{t=1}^{T}\hat{M}_t(u)X_t' = 0 \tag{2.3}$$

for all $u \in \mathbb{R}$. Intuitively, $\hat{M}_t(u)$ can be viewed as a projection of X_t to a frequency space which is orthogonal to X_t under $\mathbb{H}_0$. And (3) holds even when the time series dynamics of X_t has structural change.

Under $\mathbb{H}_0$, $\hat{A}_1(u)$ is identically 0 for all u. Then the DFT $\hat{A}(u)$ is dominated by $\hat{A}_2(u)$ which is purely a stochastic process that converges to 0 in frequency domain (zero spectrum). While under $\mathbb{H}_A$, $\hat{M}_t(u)$ is no longer orthogonal to $X_t'\beta_t$. Then $\hat{A}_1(u)$ will capture the time-varying property of both X_t and β_t. As a result, the DFT is dominated by $\hat{A}_1(u)$, and it will converge to a nonzero spectrum in the frequency domain. In particular, if X_t is weakly stationary, and $\beta_t \equiv \beta(\frac{t}{T})$ is a smooth function of the rescaled time $\frac{t}{T} \in (0, 1]$, then the probability limit of $\hat{A}_1(u)$ is proportional to the pseudo-covariance of β_t and $e^{\mathbf{i}u2\pi t/T}$ in the sense that $\frac{t}{T}$ follows the $U[0, 1]$ distribution. Therefore, we can test for structural change by checking the behaviors of $\hat{A}(u)$ at each frequency. Intuitively, the DFT is equivalent to running a nonparametric regression of Y_t on time because the Fourier basis function is an infinite sum of polynomials. Under $\mathbb{H}_0$, the DFT only captures the noise, and the impact of X_t has been purged. However, under $\mathbb{H}_A$, the DFT can capture the time-varying feature of both X_t and β_t and thus will converge to a nonzero spectrum.

The DFT also has an alternative theoretical interpretation that it can be viewed as a generalized Hausman's test in the frequency domain. Since it is proportional to the difference between two estimators which converge to the same limit only under $\mathbb{H}_0$.

$$\begin{aligned}
\hat{A}(u) &= \frac{1}{T}\sum_{t=1}^{T} X_t \hat{\varepsilon}_t e^{\mathbf{i}u2\pi t/T} \\
&= \frac{1}{T}\sum_{t=1}^{T} X_t(u) Y_t - \hat{Q}_{xx}(u)\hat{\beta} \\
&= \hat{Q}_{xx}(u)\left[\hat{Q}_{xx}^{-}(u)\frac{1}{T}\sum_{t=1}^{T} X_t(u)Y_t - \hat{Q}_{xx}^{-1}\frac{1}{T}\sum_{t=1}^{T} X_t Y_t\right] \\
&\equiv \hat{Q}_{xx}(u)[\hat{\beta}(u) - \hat{\beta}],
\end{aligned}$$

where $\hat{\beta}(u) \equiv \hat{Q}_{xx}(u)^{-}\frac{1}{T}\sum_{t=1}^{T} X_t(u)Y_t$ is a complex-valued random vector function of u, and $\hat{Q}_{xx}^{-}(u)$ denotes the generalized inverse[1] of $\hat{Q}_{xx}(u)$ at each u. For each $u \in \mathbb{R} \setminus \mathbb{Z}$, $\hat{\beta}(u)$ is the solution to

$$\min_{\beta \in \mathbb{R}^d} WSSR(\beta, u),$$

where $WSSR(\beta, u) \equiv \sum_{t=1}^{T}(Y_t - X_t'\beta)^2 e^{\mathbf{i}u2\pi t/T}$ is a weighted sum of squared residuals. In matrix notation, let $\mathcal{X} = [X_1', X_2', ..., X_T']'$, $\mathcal{Y} = [Y_1, Y_2, ..., Y_T]'$, and $\mathcal{W}(u)$ be a diagonal matrix with the diagonal elements being $[e^{\mathbf{i}u2\pi/T}, e^{\mathbf{i}u2\pi 2/T}, ..., e^{\mathbf{i}u2\pi T/T}]$, then $\hat{\beta}(u) = [\mathcal{X}'\mathcal{W}(u)\mathcal{X}]^{-}\mathcal{X}\mathcal{W}(u)\mathcal{Y}$ is a WLS (Weighted Least Squares) estimator. At every fixed u, the weighting matrix assigns weights to each time t and the weights are the Fourier basis functions of time t. The objective function $WSSR(\beta, u)$ can be viewed as a DFT of the squared residuals for each $\beta \in \Theta$. Intuitively, when the unknown parameter is constant with respect to time, the weights $e^{\mathbf{i}u\pi t/T}$ will have trivial impact on the minimization problem. And the optimal solution to minimizing $WSSR(\beta, u)$ at each u will converge to the same limit. Therefore, under $\mathbb{H}_0 : \beta_t = \beta_0$, $\hat{\beta}(u)$ is consistent for β_0:

$$\begin{aligned}
\hat{\beta}(u) &= \hat{Q}_{xx}(u)^{-1}\frac{1}{T}\sum_{t=1}^{T} X_t(u)(X_t'\beta_0 + \varepsilon_t) \\
&= \beta_0 + \hat{Q}_{xx}(u)^{-1}\frac{1}{T}\sum_{t=1}^{T} X_t(u)\varepsilon_t \\
&\xrightarrow{p} \beta_0,
\end{aligned}$$

for each $u \in \mathbb{R} \setminus \mathbb{Z}$ given $E(\varepsilon_t|X_t) = 0$. Although $\hat{\beta}(u)$ is a complex-valued function of the nuisance parameter u, the entire complex-valued term vanishes to 0 as $T \to \infty$ under $\mathbb{H}_0$. Therefore the

[1]Here we use the generalized inverse because $\sum_{t=1}^{T} e^{\mathbf{i}u2\pi t/T} \to 0$ when $u \in \mathbb{Z}$.

probability limit of the WLS estimator at each $u \in \mathbb{R} \setminus \mathbb{Z}$ will always be identical to each other. However, if there exists structural change in the unknown parameter, the weights will capture that information since various u delivers different weights to each time point t. Thus, under $\mathbb{H}_A : \beta_t \neq \beta_0$, for each fixed $u \in \mathbb{R} \setminus \mathbb{Z}$,

$$
\begin{aligned}
\hat{\beta}(u) &= \hat{Q}_{xx}^{-1}(u) \frac{1}{T} \sum_{t=1}^{T} X_t(u) Y_t \\
&= \hat{Q}_{xx}^{-1}(u) \frac{1}{T} \sum_{t=1}^{T} X_t(u) X_t' \beta_t + \hat{Q}_{xx}^{-1}(u) \frac{1}{T} \sum_{t=1}^{T} X_t(u) \varepsilon_t \\
&\xrightarrow{p} Q_{xx}^{-1}(u) \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E(X_t X_t') e^{\mathbf{i} u 2\pi t/T} \beta_t,
\end{aligned}
$$

which is different from the probability limit of $\hat{\beta}$. In this sense, our test is a generalized Hausman's test in the frequency domain. Unlike the existing generalized Hausman's tests that compare the difference between two estimators at each time point t, our comparison is at each frequency u.

In fact, there has been a long history of spectral analysis in time series econometrics, e.g., see Granger and Hatanaka (1964), Hannan (1965, 1967), Engle (1974), Granger and Watson (1984), Choi and Phillips (1993), and Corbae et al. (2002). Among many others, Engle (1974) proposes the 'Band Spectrum Regression' for static time series models. The idea is first to transform X_t and Y_t from the time domain to frequency domain using the DFT and then run an OLS via the corresponding DFTs. The sample is no longer indexed by time but by frequency. Then we can get the same estimator as the conventional OLS in time domain. By running regressions at different frequencies, we can get insights on whether the model can capture various aspects of the data. Unlike the literature that mainly focuses on spectrum analysis in stationary time series, this paper applies frequency domain analysis to nonstationary time series, namely analysis of structural change.

The DFT projects the time-varying feature of β_t to the frequency domain without loss of information. By this device, we can avoid the smoothed nonparametric estimation of β_t and can test $\mathbb{H}_0$ consistently by checking each frequency. Our test does not need prior information about the type of structural change. If structural change exists, the DFT will capture it that $\hat{A}_1(u)$ will converge to a nonzero spectrum.

3 Asymptotic Theory

To examine whether the DFT is significantly different from a zero spectrum, we need to establish the asymptotic theory. Let "$\xrightarrow{p}$", "$\xrightarrow{d}$", and "$\Rightarrow$"denote convergence in probability, convergence in distribution, and weak convergence respectively. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For any two sigma-fields $\mathcal{A}$ and $\mathcal{B}$, define the following absolute regular mixing coefficient

$$\alpha(\mathcal{A},\mathcal{B}) = \frac{1}{2}\sup\sum_{j=1}^{J}\sum_{k=1}^{K}\left|P(A_j\cap B_k)-P(A_j)P(B_k)\right|,$$

where $A_i \in \mathcal{A}$, $B_j \in \mathcal{B}$, and the supremum is taken over all pairs of finite partitions $\{A_j\}_{j\in J}$ and $\{B_k\}_{k\in K}$ of the sample space that are $\mathcal{A}$ and $\mathcal{B}$ measurable respectively. The absolute regularity condition is necessary for us to establish the uniform convergence from the sample moments to the population moments of the complex-valued empirical process. However, since we do not restrict $\{X_t', \varepsilon_t\}'$ to be stationary, we need a stronger condition to restrict the temporal dependence to be uniformly weak. Let $-\infty \leq J_1 \leq J_2 \leq \infty$, and $\mathcal{F}_{J_1}^{J_2} \equiv \sigma(X_t, \varepsilon_t)$ such that $J_1 \leq t \leq J_2$, then for each $m \geq 1$, define the following dependence coefficient:

$$\alpha(m) = \sup_{t\in\mathbb{Z}} \alpha(\mathcal{F}_{-\infty}^{t}, \mathcal{F}_{t+m}^{\infty}).$$

We define mixing coefficient $\alpha(m)$ as in Bradley (2005) for nonstationary random sequences. Because the distribution at each t could be different, we have to take supremum over all $t \in \mathbb{Z}$ to restrict the largest possible dependence. When $\{X_t', \varepsilon_t\}'$ is strictly stationary, one simply has

$$\alpha(m) = \alpha(\mathcal{F}_{-\infty}^{0}, \mathcal{F}_{m}^{\infty}).$$

Now we provide the following regularity conditions.

Assumption 3.1 *$\{X_t', \varepsilon_t\}_{t=1}^{\prime T}$ is a $(d+1)\times 1$ absolutely regular process uniformly over $t \in \mathbb{Z}$ with the mixing coefficient such that $\sum_{j=1}^{\infty} \alpha(j)^{\frac{\delta-1}{\delta}} < C < \infty$ for some $\delta > 1$.*

Assumption 3.2 *$\{\varepsilon_t\}$ is a MDS process such that $E(\varepsilon_t|I_{t-1}) = 0$ a.s. for all t, where I_{t-1} is a sigma-field generated by $\{X_{t-1}X_{t-2}, ..., \varepsilon_{t-1}, \varepsilon_{t-2}, ...\}$.*

Assumption 3.3 *$E(\varepsilon_t|X_t) = 0$ almost surely for all t.*

Assumption 3.4 *(i) $V_{xx}(u_1, u_2) \equiv \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} Re\left\{E[X_t(u_1)X_t(u_2)^*\varepsilon_t^2]\right\}$ is finite and positive definite for all $(u_1, u_2) \in \mathbb{R}^2$; (ii) $Q_{xx} = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E(X_tX_t')$ is nonsingular and finite.*

Assumption 3.5 *$E(X_{jt}^{4\delta}) \leq C < \infty$ for all $j = 1, ..., d$ and all t; $E(\varepsilon_t^{4\delta}) \leq C < \infty$ for all t.*

Assumption 3.6 *$\beta_t \in \Theta$, where Θ is a bounded parameter space in $\mathbb{R}^d$ such that $\frac{1}{T}\sum_{t=1}^{T}\|\beta_t\|^2 < C < \infty$.*

Assumption 3.1 allows the structure of X_t to change over time but restricts its temporal dependence to be weak. Under this condition, our test is applicable to dynamic time series regression models. The structural change in X_t can be either smooth or abrupt. That means our test is robust to unknown structural change in the covariates. The mixing condition restricts the temporal dependence to be weak uniformly over time t. This is crucial for us to apply the law of large numbers and the central limit theorem by Bradley and Tone (2015). We note that Assumption

3.1 excludes the case when X_t is unit root or near unit root. Such cases will be investigated in subsequent studies.

Assumption 3.2 restricts the model disturbance ε_t to be serially uncorrelated. This is only for the simplicity of delivering our main result. We can allow ε_t to exhibit serial correlation and use the long-run variance-covariance (e.g., Newey and West, 1987). Assumption 3.3 excludes the existence of endogenous covariates. In Section 4, we will consider endogenous covariates. Assumption 3.3 implies the mean of ε_t is 0 and does not vary with time. Assumption 3.4(i) allows ε_t to be conditional heteroscedastic which is common in time series data. Assumption 3.4(ii) ensures the existence of the OLS estimator. Assumption 3.5 is a regular moment condition on ε_t and X_t. Assumption 3.5 allows the variance of ε_t to be time-varying. If both X_t and ε_t are weakly stationary, then Assumption 3.4 (i) can be implied by Assumption 3.5. Compared to the existing literature, e.g., Chen and Hong (2012), Kristensen (2012), Cai et al. (2015), we do not impose any smoothness assumption on the unknown parameter β_t. We only need it to be square-summable, which is stated in Assumption 3.6.

Theorem 3.1 *Suppose Assumptions 3.1-3.5 hold, under* $\mathbb{H}_0$*,*

$$\sqrt{T}\hat{A}(u) \Rightarrow \mathcal{G}(u),$$

where $\mathcal{G}(u)$ *is a zero-mean complex-valued Gaussian process with covariance kernel*

$$\begin{aligned}\mathcal{K}(u_1, u_2) &\equiv cov[\mathcal{G}(u_1), \mathcal{G}(u_2)^*] \\ &= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left[M_t(u_1) M_t(u_2)'^* \varepsilon_t^2\right],\end{aligned}$$

where $M_t(u) = X_t e^{\mathbf{i}u2\pi t/T} - Q_{xx}(u) Q_{xx}^{-1} X_t$.

Theorem 3.1 provides the asymptotic null distribution of the DFT $\hat{A}(u)$ scaled by $\sqrt{T}$. $M_t(u)$ is the limit $\hat{M}_t(u)$. At each frequency u, it converges to a normal distribution with mean 0 and variance $\mathcal{K}(u, u)$ as $T \to \infty$. If both X_t and ε_t are weakly stationary, then

$$M_t(u) = X_t \left(e^{\mathbf{i}u2\pi t/T} - \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau \right),$$

and

$$\mathcal{K}(u_1, u_2) = E(X_t X_t' \varepsilon_t^2) \widetilde{\mathrm{cov}}(e^{\mathbf{i}u_1 2\pi\tau}, e^{-\mathbf{i}u_2 2\pi\tau}),$$

where $\widetilde{\mathrm{cov}}(e^{\mathbf{i}u_1 2\pi\tau}, e^{-\mathbf{i}u_2 2\pi\tau}) = \left[\int_0^1 e^{\mathbf{i}2\pi(u_1-u_2)\tau} d\tau - \int_0^1 e^{\mathbf{i}2\pi u_1\tau} d\tau \int_0^1 e^{-\mathbf{i}2\pi u_2\tau} d\tau\right]$ is a pseudo-covariance in the sense that τ follows the $U[0, 1]$ distribution. Intuitively, the covariance kernel contains two components. $E(X_t X_t \varepsilon_t^2)$ is the asymptotic variance of $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_t \varepsilon_t$, which is the uncertainty introduced by the OLS estimation. The other component can be viewed as the noise introduced by the Fourier transform. When X_t and ε_t are nonstationary, the two components will intertwine with each other.

Next, we want to show that the DFT $\hat{A}(u)$ converge to a non-zero spectrum under $\mathbb{H}_A$ in the frequency domain, which is essential for the consistency of our test.

Theorem 3.2 *Suppose Assumptions 3.1 to 3.6 hold, under $\mathbb{H}_A$, as $T \to \infty$*

$$\sup_{u \in \mathbb{R}} \left\| \hat{A}(u) - \tilde{A}(u) \right\| \xrightarrow{p} 0,$$

where $\tilde{A}(u) \equiv \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E[M_t(u) X_t'] \beta_t$.

Theorem 3.2 shows the limit of $\hat{A}(u)$ is a nonzero spectrum in the frequency domain. If we let X_t be weakly stationary, and $\beta_t \equiv \beta(\frac{t}{T})$ be a smooth function of rescaled time $t/T \in (0, 1]$, then

$$\begin{aligned} \tilde{A}(u) &= E(X_t X_t') \lim_{T\to\infty} \left[\frac{1}{T} \sum_{t=1}^{T} \beta_t e^{\mathbf{i}u2\pi t/T} - \frac{1}{T} \sum_{t=1}^{T} \beta_t \frac{1}{T} \sum_{t=1}^{T} e^{\mathbf{i}u2\pi t/T} \right] \\ &\equiv E(X_t X_t') \widetilde{\text{cov}}[\beta(\tau), e^{\mathbf{i}u2\pi\tau}], \end{aligned}$$

where $\widetilde{\text{cov}}[\beta(\tau), e^{\mathbf{i}u2\pi\tau}] = \int_0^1 \beta(\tau) e^{\mathbf{i}u2\pi\tau} d\tau - \int_0^1 \beta(\tau) d\tau \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau$ is a pseudo-covariance in the sense that τ follows the $U[0, 1]$ distribution. As long as $\beta_t \equiv \beta(\frac{t}{T})$ is a non-constant function of time t, $\widetilde{\text{cov}}[\beta(\tau), e^{\mathbf{i}u2\pi\tau}]$ will be different from 0 for $u \neq 0$. Theorem 3.2 shows that the DFT $\hat{A}(u)$ is equivalent to the sample analog of a pseudo-covariance between β_t and $e^{\mathbf{i}u2\pi t/T}$ under $\mathbb{H}_A$. The result also explains why we construct $\hat{A}(u)$ by considering the DFT of $X_t \hat{\varepsilon}_t$ rather than $\hat{\varepsilon}_t$. When X_t is weakly stationary such that $E(X_t) = 0$, then our test will have no power if we consider the Fourier transform of $\hat{\varepsilon}_t$. For this reason, we construct $\hat{A}(u)$ using the Fourier transform of $X_t \hat{\varepsilon}_t$ to ensure that it converges to a nonzero spectrum under $\mathbb{H}_A$.

One way to examine the behavior the DFT in the frequency domain is to compose a test for each fixed u. However, the power of a test using $\hat{A}(u)$ depends on u under $\mathbb{H}_A$. Intuitively, $\hat{A}(u)$ can be viewed as a series approximation of a weighted average of β_t using the Fourier basis functions. Frequency u represents how volatile the $\sin(\cdot)$ and $\cos(\cdot)$ series are. If the variation of β_t over time is very slow, then we expect the Fourier basis functions with a lower frequency can approximate β_t better than those with a higher frequency. Thus, we should test $\mathbb{H}_0$ by assigning large weights to small frequencies, and *vice versa*. Unfortunately, the information on β_t is usually unknown *a prior*. To ensure the DFT test is consistent against all alternatives, we consider the following Kolmogorov-Smirnov (KS) and Cramér-von Mises (CM) type test statistics:

$$\hat{K} = T \sup_{u \in \mathbb{R}} \|\hat{A}(u)\|^2,$$

and

$$\hat{C} = T \int_{\mathbb{R}} \|\hat{A}(u)\|^2 W(u) du,$$

where $W(\cdot)$ is a weighting function that assigns weights to each frequency, and it satisfies

$$\int_R |u|^2 W(u) du < \infty. \tag{3.1}$$

$\|\hat{A}(u)\|^2$ is called the 'periodogram' and it reveals the magnitude of the DFT at each frequency. The periodogram is also proportional to the quadratic difference between the WLS and the OLS estimators at each frequency.

The computation of $\hat{C}$ involves integration over all u. We note that proper choices of the weighting function $W(u)$ can avoid numerical integration of u, and $\hat{C}$ will have a closed-form expression. One example is the zero-mean normal density function

$$W(u) = \frac{1}{\sqrt{2\pi\xi^2}} \exp\left(-\frac{u^2}{2\xi^2}\right).$$

By the following identity

$$\int_{\mathbb{R}} \cos(uz) \frac{1}{\sqrt{2\pi\xi^2}} \exp\left(-\frac{u^2}{2\xi^2}\right) du = \exp\left(-\frac{z^2\xi^2}{2}\right),$$

we have

$$\hat{C} \quad = \quad \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} X_s' X_t \hat{\varepsilon}_s \hat{\varepsilon}_t \exp\left(-\frac{2\pi^2(s-t)^2\xi^2}{T^2}\right),$$

where ξ is the standard deviation that measures the dispersion of weights assigned around 0. In general, the results are insensitive to the choice of ξ. Another example is the Laplace density weighting

$$W(u) = \frac{\lambda}{2} e^{-\lambda|u|}.$$

By the identity that

$$\int_{\mathbb{R}} \cos(uz) \frac{\lambda}{2} e^{-\lambda|u|} = \frac{\lambda^2}{\lambda^2 + z^2},$$

we can show

$$\hat{C} \quad = \quad \frac{1}{T} \sum_{s=1}^{T} \sum_{t=1}^{T} X_s' X_t \hat{\varepsilon}_s \hat{\varepsilon}_t \left(\frac{\lambda^2 T^2}{\lambda^2 T^2 + 4\pi^2(s-t)^2}\right).$$

Here λ determines the dispersion of weights assigned around 0. As λ increases, more weights are assigned to higher frequencies.

After providing the test statistics $\hat{K}$ and $\hat{C}$ based on the DFT, we now show the asymptotic null distributions.

Theorem 3.3 *Suppose Assumptions 3.1-3.5 hold, under* $\mathbb{H}_0$*, by the continuous mapping theorem,*

$$\hat{K} \xrightarrow{d} \sup_{u \in \mathbb{R}} \|\mathcal{G}(u)\|^2,$$

and

$$\hat{C} \xrightarrow{d} \int_{\mathbb{R}} \|\mathcal{G}(u)\|^2 W(u) du,$$

as $T \to \infty$.

Theorem 3.3 provides the asymptotic distributions of $\hat{K}$ and $\hat{C}$ under $\mathbb{H}_0$. They are not pivotal because they depend on the unknown data generating process (DGP). We follow Hansen (1996) and propose the following multiplier bootstrap to obtain critical values.

- Step(i). Use the sample $\{Y_t, X_t'\}_{t=1}^T$ to estimate the model via the OLS, and compute $\hat{K}$ and $\hat{C}$;
- Step (ii). Generate $i.i.d.N(0,1)$ random variables $\{v_{bt}\}_{t=1}^T$ and compute compute $\hat{K}_b$ and $\hat{C}_b$ using $\hat{A}_b(u) = \frac{1}{T}\sum_{t=1}^T \hat{M}_t(u)\hat{\varepsilon}_t v_{bt}$, where

$$\hat{M}_t(u) = X_t e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1} X_t,$$

  and $\hat{\varepsilon}_t$ is the estimated residuals from OLS;
- Step (iii). Repeat Step (ii) for a total of B times to obtain B bootstrap test statistics $\{\hat{K}_b, \hat{C}_b\}_{b=1}^B$;
- Step (iv). Compute the bootstrap p-values for $\hat{K}$ and $\hat{C}$ respectively, with $p_{B,T}^K = \frac{1}{B}\sum_{b=1}^B \mathbf{1}(\hat{K}_b > \hat{K})$ and $p_{B,T}^C = \frac{1}{B}\sum_{b=1}^B \mathbf{1}(\hat{C}_b > \hat{C})$, where $\mathbf{1}(\cdot)$ is the indicator function.

$p_{B,T}^K$ and $p_{B,T}^C$ are the conditional p-values for $\hat{K}$ and $\hat{C}$ via resampling for B times. Next, we show they are consistent for the true p-values. Let $F^K(\cdot)$ and $F^C(\cdot)$ denote the distribution function of $\sup_{u\in\mathbb{R}} \|\mathcal{G}(u)\|^2$ and $\int_{\mathbb{R}} \|\mathcal{G}(u)\|^2 W(u)du$ respectively. Define $p_T^K = 1 - F^K(\hat{K})$ and $p_T^C = 1 - F^C(\hat{C})$ be the asymptotic p-values, such that

$$\lim_{T\to\infty} Pr\{p_T^j \leq \alpha | \mathbb{H}_0\} = \alpha,$$

for $j = K, C$. Conditional on the sample $\{Y_t, X_t\}_{t=1}^T$, let $F_T^K(\cdot)$ and $F_T^C(\cdot)$ be the conditional distribution functions of $\hat{K}$ and $\hat{C}$, respectively. Define $\tilde{p}_T^K = 1 - F_T^K(\hat{K})$ and $\tilde{p}_T^C = 1 - F_T^C(\hat{C})$ to be the conditional p-values, where $\hat{F}_T^K(\cdot)$ and $\hat{F}_T^K(\cdot)$ are generated by replacing $\hat{A}(u)$ with

$$\hat{A}(u) = \frac{1}{T}\sum_{t=1}^T \hat{M}_t(u)\hat{\varepsilon}_t v_t,$$

where $\{v_t\}_{t=1}^T$ is $i.i.d.N(0,1)$. Then we can show the consistency of the resampling method by the following Corollary.

Corollary 3.1 *Under the regularity conditions in Theorem 3.1,* $p^j_{B,T} \xrightarrow{p} \tilde{p}^j_T$ *as* $B \to \infty$*, and*

$$\tilde{p}^j_T \Rightarrow p^j,$$

where p^j *is the true p-value for* $j = K, C$.

Next, we show the asymptotic power of our test statistics.

Theorem 3.4 *Under Assumptions 3.1-3.6 and* $\mathbb{H}_A$*, for any sequence of non-stochastic constants* $\{c_T = o(T)\}$*, as* $T \to \infty$*,*

$$\begin{aligned} P(\hat{K} > c_T) &\to 1, \\ P(\hat{C} > c_T) &\to 1. \end{aligned}$$

This result shows the power of $\hat{K}$ and $\hat{C}$ against fixed alternatives approaches to 1 as $T \to \infty$. Therefore, the tests $\hat{K}$ and $\hat{C}$ are consistent against both abrupt structural breaks and smooth structural changes.

We now investigate the local power property of our tests and compare with existing consistent tests in the literature. Consider the following local alternatives under $\mathbb{H}_{A1}$:

$$\beta_t = \beta_0 + \Delta_T \phi_t,$$

where $\phi_t \in \mathbb{R}^d$ is a nonrandom function of time t such that $0 < \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \|\phi_t\|^2 < \infty$. ϕ_t can be either a smooth function of $\frac{t}{T}$ (smooth structural changes) or a step function of t (abrupt structural breaks).

Theorem 3.5 *Suppose Assumptions 3.1-3.6 and* $\mathbb{H}_{A1}$ *with* $\Delta_T = T^{-1/2}$ *hold. Then as* $T \to \infty$*,*

$$\begin{aligned} \hat{K} &\xrightarrow{d} \sup_{u\in\mathbb{R}} \|\xi(u) + G(u)\|^2, \\ \hat{C} &\xrightarrow{d} \int_{\mathbb{R}} \|\xi(u) + G(u)\|^2 W(u) du, \end{aligned}$$

where

$$\xi(u) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[M_t(u) X_t']\phi_t.$$

Theorem 3.5 shows that our tests $\hat{K}$ and $\hat{C}$ have nontrivial power against a class of local alternatives with the parametric rate $\Delta_T = T^{-1/2}$. In contrast, Chen and Hong (2012), Zhang and Wu (2012), and Cai et al. (2015) can only detect a class of local alternatives at a rate of $\Delta_T = T^{-1/2}h^{-1/4}$, where the bandwidth $h \to 0$ as $T \to \infty$. This is the advantage of using the Fourier transform by avoiding smoothed nonparametric estimation. The price we have to pay is that the asymptotic distributions of our tests $\hat{K}$ and $\hat{C}$ are not pivotal. Moreover, we are free from boundary problem which is common in smoothed nonparametric testing. Chen and Hong (2012) use boundary reflection to increase the convergence rate of boundary estimation. Our test does not require such a delicate treatment. Compared to the existing tests that achieve parametric rates such as Andrews' (1993) supremum and Bai and Perron's (1998) double maximum tests, our test does not require trimming of the data. Their tests do not have uniform power for all $t/T \in (0, 1]$.

4 Extension to Endogenous Covariates

In this section, we extend our tests to a time series regression model with endogenous covariates, i.e., $E(\varepsilon_t|X_t) \neq 0$. Suppose there exists a set of instruments $Z_t \in \mathbb{R}^l$ such that $E(\varepsilon_t|Z_t) = 0$ a.s., and $E(X_t Z_t') \neq 0$, where $l \geq d$.

To test $\mathbb{H}_0$, we now consider the following complex-valued empirical process:

$$\hat{A}^{IV}(u) \quad = \quad \frac{1}{T}\sum_{t=1}^{T} \hat{X}_t \hat{\varepsilon}_t e^{\mathbf{i}u2\pi t/T},$$

where $\hat{\varepsilon}_t = Y_t - X_t'\hat{\beta}_{2sls}$ is the estimated residual from the Two-Stage Least Squares (2SLS) estimation, and $\hat{X}_t \in \mathbb{R}^d$ is the fitted value from the first stage regression of X_t on instrumental variables Z_t. $\hat{A}^{IV}(u)$ can be viewed as the DFT of $\hat{X}_t\hat{\varepsilon}_t$. The 2SLS estimator cannot capture the time-varying feature of the unknown parameter. So we focus on the estimated residual $\hat{\varepsilon}_t$ from the 2SLS and consider the following decomposition:

$$\begin{aligned}
\hat{\varepsilon}_t &= Y_t - X_t'\hat{\beta}_{2sls} \\
&= X_t'(\beta_t - \hat{\beta}_{2sls}) + \varepsilon_t \\
&= X_t'\left[\beta_t - \left(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx}\right)^{-1}\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zy}\right] + \varepsilon_t \\
&= \varepsilon_t - X_t'\left(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx}\right)^{-1}\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} Z_t\varepsilon_t\right) \\
&\quad + X_t'\left[\beta_t - \left(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx}\right)^{-1}\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} Z_t X_t'\beta_t\right)\right].
\end{aligned}$$

Like in the case where all covariates are exogenous, the estimated residuals from the 2SLS can also be decomposed into three components: the disturbance ε_t, the estimation uncertainty, and the local feature of β_t over time. Under $\mathbb{H}_0 : \beta_t = \beta_0$ for all t, $\hat{\beta}_{2sls}$ is consistent for β_0. Then the third component will be identically 0 and the DFT will converge to 0 in probability at every frequency. While under $\mathbb{H}_A$, $\hat{\beta}_{2sls}$ fails to capture the time-varying feature of β_t, and such information will be contained in the estimated residuals. Therefore, the DFT will converge to a nonzero spectrum in the frequency domain. In particular, if both X_t and Z_t are weakly stationary, and $\beta_t \equiv \beta(\frac{t}{T})$ is a smooth function of the rescaled time $\frac{t}{T} \in (0, 1]$, then the probability limit of the DFT under $\mathbb{H}_A$ is proportional to the pseudo-covariance of β_t and $e^{\mathbf{i}u2\pi t/T}$ in the sense that $\frac{t}{T}$ follows the $U[0, 1]$ distribution.

For the first stage regression, we simply estimate it by the OLS:

$$\hat{X}_t = \hat{\gamma}' Z_t,$$

where $\hat{\gamma} = (\sum_{t=1}^T Z_t Z_t')^{-1} \sum_{t=1}^T Z_t X_t' \equiv \hat{Q}_{zz}^{-1}\hat{Q}_{zx}$ is the OLS estimator. As $T \to \infty$, $\hat{X}_t$ will converge in probability to $\tilde{X}_t \equiv \gamma' Z_t$. It is possible that the true relationship between X_t and Z_t is unstable such that γ is not constant with respect to time. If that is the case, then $\hat{\gamma}$ will not be consistent for γ. However, that has a trivial impact on our analysis since our goal is not to estimate γ consistently but rather to use $\tilde{X}_t$ as a proxy to obtain the 2SLS estimator. So $\tilde{X}_t$ can be viewed as a linear projection of X_t on Z_t. As long as X_t and Z_t are not orthogonal to each other, we can achieve the consistency of the 2SLS estimators for β_0 under $\mathbb{H}_0$. Moreover, we do not restrict either X_t or Z_t to be weakly stationary. Thus $\tilde{X}_t$ can also be nonstationary such that its distribution may change smoothly or break abruptly. The fitted value $\hat{X}_t$ from the first stage regression serves as the regressor in the second stage. It has the same role as X_t in the case where all covariates are exogenous. Therefore its instability will have no impact on testing for structural change in the structural equation. The tests by Hall et al. (2012) and Perron and Yamamoto (2014) rely on comparing consistent estimates for β_t using the 2SLS in subsamples. Therefore they need first to investigate the stability of the reduce form. This is a drawback of testing in the time domain. However, our frequency domain based approach can avoid this issue and is robust to instability in both covariates and instruments.

Like the DFT in Section 2, $\hat{A}^{IV}(u)$ can also be interpreted as a generalized Hausman's test in the frequency domain. Let $\hat{X}_t(u) \equiv \hat{X}_t e^{\mathbf{i}u2\pi t/T} = \hat{Q}_{xz}\hat{Q}_{zz}^{-1} Z_t e^{\mathbf{i}u2\pi t/T}$, we can rewrite the DFT as

$$\begin{aligned}
\hat{A}^{IV}(u) &= \frac{1}{T}\sum_{t=1}^{T} \hat{X}_t(u)\hat{\varepsilon}_t \\
&= \hat{Q}_{\hat{x}x}(u)\left[\hat{\beta}^{IV}(u) - \hat{\beta}_{2sls}\right] \\
&= \hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx}(u)\left[\hat{\beta}^{IV}(u) - \hat{\beta}_{2sls}\right],
\end{aligned}$$

where we define the following linear IV estimator for β_0:

$$\begin{aligned}
\hat{\beta}^{IV}(u) &\equiv \left[\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx}(u)\right]^{-} \hat{Q}_{xz}\hat{Q}_{zz}^{-1}\frac{1}{T}\sum_{t=1}^{T} Z_t Y_t e^{\mathbf{i}u2\pi t/T} \\
&= [\mathcal{X}'\mathcal{P}_Z\mathcal{W}(u)\mathcal{X}]^{-}\mathcal{X}'\mathcal{P}_Z\mathcal{W}(u)\mathcal{Y},
\end{aligned}$$

where $\mathcal{P}_Z = \mathcal{Z}(\mathcal{Z}'\mathcal{Z})^{-1}\mathcal{Z}'$ is a projection matrix. Compared to the WLS defined in Section 2, $\hat{\beta}^{IV}$ contains an additional projection to the space spanned by Z_t. In particular, when $u = 0$, $\hat{\beta}^{IV}(0) = [\mathcal{X}'\mathcal{P}_Z\mathcal{X}]^{-1}\mathcal{X}'\mathcal{P}_Z\mathcal{Y}$ is exactly the 2SLS estimator. Under $\mathbb{H}_0$,

$$\hat{\beta}^{IV}(u) = \left[\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx}(u)\right]^{-1} \hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zy}(u)$$

$$
\begin{aligned}
&= \beta_0 + \left[\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx}(u)\right]^{-1}\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\frac{1}{T}\sum_{t=1}^{T} Z_t\varepsilon_t e^{\mathbf{i}u2\pi t/T} \\
&\overset{p}{\to} \beta_0,
\end{aligned}
$$

for any $u \in \mathbb{R} \setminus \mathbb{Z}$. It implies $\hat{\beta}^{IV}(u)$ is consistent for β_0 under $\mathbb{H}_0$, and it can be viewed as a general class of estimators that contain the 2SLS estimator as a special case. Following analogous arguments, $\hat{\beta}^{IV}(u)$ will converge to the same probability limit for all $u \in \mathbb{R}$ under $\mathbb{H}_0 : \beta_t = \beta_0$. However, $\hat{\beta}^{IV}(u)$ will converge to a nonzero spectrum under $\mathbb{H}_A$.

Now we provide the asymptotic theory for the test based on $\hat{A}^{IV}(u)$.

Assumption 4.1 *$\{X_t', Z_t', \varepsilon_t\}_{t=1}^{\prime T}$ is a $(d+l+1)\times 1$ absolutely regular process uniformly over $t \in \mathbb{Z}$ with the mixing coefficient such that $\sum_{j=1}^{\infty}\alpha(j)^{\frac{\delta-1}{\delta}} < C < \infty$ for some $\delta > 1$.*

Assumption 4.2 *$\{\varepsilon_t\}$ is a MDS process such that $E(\varepsilon_t|I_{t-1}) = 0$ a.s., where I_{t-1} is a Sigma-algebra generated by $\{Z_{t-1}, Z_{t-2}, ..., \varepsilon_{t-1}, \varepsilon_{t-2}, ...\}$.*

Assumption 4.3 *$E(\varepsilon_t|X_t) \neq 0$ for some t and $E(Z_t\varepsilon_t) = 0$ a.s. for all t.*

Assumption 4.4 *(i) $V_{zz}(u_1, u_2) \equiv \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E[Z_t(u_1)Z_t(u_2)^*\varepsilon_t^2]$ is finite and positive definite for all $(u_1, u_2) \in \mathbb{R}^2$;*
(ii) $Q_{zz} = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E(Z_tZ_t')$ is nonsingular and finite;
(iii) $Q_{xz} = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E(X_tZ_t')$ is finite and of full rank.

Assumption 4.5 *$E(Z_{jt}^{4\delta}) \leq C < \infty$ and $E(X_{kt}^{4\delta}) \leq C < \infty$ for some $\delta > 1$, all $j = 1, ..., l$, $k = 1, ..., d$, and all t; $E(\varepsilon_t^{4\delta}) \leq C < \infty$ for all t.*

Like Assumption 3.1, Assumption 4.1 allows both X_t and Z_t to have smooth structural changes or abrupt structural breaks. The temporal dependence has been controlled by uniform mixing coefficient for possibly nonstationary time series. Assumption 4.2 restricts the disturbance to be serially uncorrelated. Assumption 4.3 states the existence of endogeneity and the validity of instrumental variable Z_t. Assumptions 4.4 and 4.5 are regular moment conditions on X_t, Z_t, and ε_t which are rather mild.

Theorem 4.1 *Suppose Assumptions 4.1-4.5 hold, under $\mathbb{H}_0$,*

$$
\sqrt{T}\hat{A}^{IV}(u) \Rightarrow \mathcal{G}^{IV}(u)
$$

where $\mathcal{G}^{IV}(u)$ is a zero-mean complex-valued Gaussian process with covariance kernel

$$
\begin{aligned}
\mathcal{K}^{IV}(u_1, u_2) &\equiv cov\left[\mathcal{G}^{IV}(u_1), \mathcal{G}^{IV}(u_2)^*\right] \\
&= \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T} E\left[M_t^{IV}(u_1)M_t^{IV}(u_2)'^*\varepsilon_t^2\right],
\end{aligned}
$$

where $M_t^{IV}(u) = \tilde{X}_t e^{\mathbf{i}u2\pi t/T} - Q_{\tilde{x}x}(u)Q_{\tilde{x}\tilde{x}}^{-1}\tilde{X}_t$.

The asymptotic null distribution for the DFT $\sqrt{T}\hat{A}^{IV}(u)$ is similar to Theorem 3.1 expect that the covariance kernels are different. When endogenous covariates are present, we need to first project X_t on Z_t, which introduces additional estimation uncertainty to the variance of the DFT. When both X_t and Z_t are weakly stationary, we can show

$$\mathcal{K}^{IV}(u_1, u_2) = \gamma' E\left(Z_t Z_t' \varepsilon_t^2\right) \gamma \widetilde{\text{cov}}(e^{\mathbf{i}u_1 2\pi\tau}, e^{-\mathbf{i}u_2 2\pi\tau}).$$

The covariance kernel for the DFT $\sqrt{T}\hat{A}^{IV}(u)$ is equal to the variance of score function in the IV regression multiplied by a pseudo-covariance introduced by the Fourier transform.

Theorem 4.2 *Suppose Assumption 4.1 to 4.5 and 3.6 hold, under $\mathbb{H}_A$, as $T \to \infty$,*

$$\sup_{u\in\mathbb{R}} \left\| \hat{A}^{IV}(u) - \tilde{A}^{IV}(u) \right\| \xrightarrow{p} 0,$$

where $\tilde{A}^{IV}(u) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[M_t^{IV}(u)X_t']\beta_t$.

Theorem 4.2 shows the limit of the DFT is a nonzero spectrum in the frequency domain, and it is a weighted average of the unknown parameter β_t. If we let X_t and Z_t be weakly stationary, and $\beta_t \equiv \beta(\frac{t}{T})$ be a smooth function of rescaled time $t/T \in (0, 1]$, then the limit of $\hat{A}^{IV}(u)$ is

$$\begin{aligned} \tilde{A}^{IV}(u) &= \gamma' E(Z_t X_t') \lim_{T\to\infty} \left[\frac{1}{T}\sum_{t=1}^{T} \beta_t e^{\mathbf{i}u2\pi t/T} - \frac{1}{T}\sum_{t=1}^{T} \beta_t \frac{1}{T}\sum_{t=1}^{T} e^{\mathbf{i}u2\pi t/T} \right] \\ &\equiv \gamma' E(Z_t X_t') \widetilde{\text{cov}}[\beta(\tau), e^{\mathbf{i}u2\pi\tau}], \end{aligned}$$

where $\widetilde{\text{cov}}[\beta(\tau), e^{\mathbf{i}u2\pi\tau}]$ is defined as in Section 3. As long as β_t is not constant with respect to time, the pseudo-covariance will be different from 0 for $u \neq 0$. It is crucial to ensure the consistency of our test based on $\hat{A}^{IV}(u)$. To examine the behavior of the DFT $\hat{A}^{IV}(u)$ at each frequency u, we consider the following KS and CM type test statistics:

$$\hat{K}^{IV} = T \sup_{u\in\mathbb{R}} \left\| \hat{A}^{IV}(u) \right\|^2,$$

and

$$\hat{C}^{IV} = T \int_{\mathbb{R}} \left\| \hat{A}^{IV}(u) \right\|^2 W(u) du,$$

where $W(\cdot)$ is a weighting function that satisfies (3.1).

Theorem 4.3 *Suppose Assumptions 4.1 to 4.5 hold, under $\mathbb{H}_0$, by the continuous mapping theorem,*

$$\hat{K}^{IV} \xrightarrow{d} \sup_{u\in\mathbb{R}} \left\| \mathcal{G}^{IV}(u) \right\|^2,$$

and

$$\hat{C}^{IV} \xrightarrow{d} \int_{\mathbb{R}} \left\|\mathcal{G}^{IV}(u)\right\|^2 W(u)du.$$

Theorem 4.3 provides the asymptotic distributions of $\hat{K}^{IV}$ and $\hat{C}^{IV}$ under $\mathbb{H}_0$. Our test is tuning parameter-free, but we need resampling methods to obtain critical values. Like in the OLS case, we follow Hansen (1996) and propose the following resampling procedures.

- Step(i). Use the sample $\{Y_t, X_t', Z_t'\}_{t=1}^T$ to estimate the model via the 2SLS, and compute $\hat{K}^{IV}$ and $\hat{C}^{IV}$;
- Step (ii). Generate $i.i.d.N(0,1)$ random variables $\{v_{bt}\}_{t=1}^T$ and compute $\hat{K}_b^{IV}$ and $\hat{C}_b^{IV}$ using $\hat{A}_b(u) = \frac{1}{T}\sum_{t=1}^T \hat{M}_t^{IV}(u)\hat{\varepsilon}_t v_{bt}$, where

$$\hat{M}_t^{IV}(u) = \hat{X}_t e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{\hat{x}x}(u)\hat{Q}_{\hat{x}\hat{x}}^{-1}\hat{X}_t,$$

  and $\hat{\varepsilon}_t$ is the estimated residuals from the 2SLS;
- Step (iii). Repeat Step (ii) for a total of B times to obtain B bootstrap test statistics $\{\hat{K}_b^{IV}, \hat{C}_b^{IV}\}_{b=1}^B$;
- Step (iv). Compute the bootstrap p-values for $\hat{K}^{IV}$ and $\hat{C}^{IV}$ respectively, with $p_{B,T}^{K,IV} = \frac{1}{B}\sum_{b=1}^B \mathbf{1}(\hat{K}_b^{IV} > \hat{K}^{IV})$ and $p_{B,T}^{C,IV} = \frac{1}{B}\sum_{b=1}^B \mathbf{1}(\hat{C}_b^{IV} > \hat{C}^{IV})$, where $\mathbf{1}(\cdot)$ is the indicator function.

$p_{B,T}^{K,IV}$ and $p_{B,T}^{C,IV}$ are the conditional p-values for $\hat{K}^{IV}$ and $\hat{C}^{IV}$ via resampling for B times. By similar proof as in Corollary 3.1, we can show they are consistent for the true p-values.

Next, we state the asymptotic power of $\hat{K}^{IV}$ and $\hat{C}^{IV}$ against fixed alternatives.

Theorem 4.4 *Suppose Assumptions 4.1-4.5 and 3.6 hold. Then under $\mathbb{H}_A$, for any sequence of non-stochastic constants $\{c_T = o(T)\}$, as $T \to \infty$,*

$$\begin{aligned} P(\hat{K}^{IV} > c_T) &\to 1, \\ P(\hat{C}^{IV} > c_T) &\to 1. \end{aligned}$$

Theorem 4.4 shows that our test statistics $\hat{K}^{IV}$ and $\hat{C}^{IV}$ diverge to infinity at a speed of T when $\mathbb{H}_0$ is failed. Unlike Perron and Yamamoto (2015), $\hat{K}^{IV}$ and $\hat{C}^{IV}$ are consistent against both abrupt structural breaks and smooth structural changes. Compared to Chen's (2015) nonparametric consistent test, $\hat{K}^{IV}$ and $\hat{C}^{IV}$ avoid smoothed nonparametric estimation.

To investigate the local power property of $\hat{K}^{IV}$ and $\hat{C}^{IV}$, we consider $\mathbb{H}_{A1}$ defined in Section 3.

Theorem 4.5 *Suppose Assumptions 4.1-4.5 and $\mathbb{H}_{A1}$ with $\Delta_T = T^{-1/2}$ hold. Then as $T \to \infty$,*

$$\begin{aligned} \hat{K}^{IV} &\xrightarrow{d} \sup_{u\in\mathbb{R}} \left\|\xi^{IV}(u) + \mathcal{G}^{IV}(u)\right\|^2, \\ \hat{C}^{IV} &\xrightarrow{d} \int_{\mathbb{R}} \left\|\xi^{IV}(u) + \mathcal{G}^{IV}(u)\right\|^2 W(u)du, \end{aligned}$$

where

$$\xi(u)^{IV} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E\left[M_t^{IV}(u)X_t'\right]\phi_t.$$

Theorem 4.5 implies that our tests $\hat{K}^{IV}$ and $\hat{C}^{IV}$ have nontrivial power against a class of local alternatives with rate $\Delta_T = T^{-1/2}$. This rate is faster than the nonparametric rate $T^{-1/2}h^{-1/4}$ in Chen (2015) where the bandwidth $h \to 0$ as $T \to \infty$. Thus, our test is asymptotically more powerful.

5 Model Misspecification

Many existing tests for structural change are based on a correctly specified conditional mean model. If the model is misspecified, then the rejection of a structural change test may come from model misspecification rather than the instability of unknown parameters (Chen and Hong, 2012). Unlike the existing literature that cannot distinguish the source of rejection from structural change and model misspecification, our frequency-based approach will be free of model misspecification under certain scenarios.

Suppose the true DGP is

$$Y_t = X_t'\beta(\xi_t) + \varepsilon_t,$$

where we assume $\beta(\xi_t) : \mathbb{R}^p \to \mathbb{R}^d$ is some unknown function of random variable $\xi_t \in \mathbb{R}^p$. When $\xi_t = X_t$, then the DGP can be viewed as a nonlinear function of X_t. We assume $E(\varepsilon_t|X_t) = 0$ a.s., and ξ_t is weakly stationary.

Consider the the DFT $\hat{A}(u)$ defined in Section 3:

$$\begin{aligned}\hat{A}(u) &= \frac{1}{T}\sum_{t=1}^{T} X_t\hat{\varepsilon}_t e^{\mathbf{i}u2\pi t/T} \\ &= \frac{1}{T}\sum_{t=1}^{T} M_t(u)X_t\beta(\xi_t) + \frac{1}{T}\sum_{t=1}^{T} M_t(u)\varepsilon_t \\ &= \frac{1}{T}\sum_{t=1}^{T} X_tX_t'\beta(\xi_t)e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}\frac{1}{T}\sum_{t=1}^{T} X_tX_t'\beta(\xi_t) \\ &\quad +\hat{A}_2(u).\end{aligned}$$

The DFT can also be decomposed to two parts. The second component $\hat{A}_2(u)$ will converge to a zero spectrum in frequency domain due to the orthogonality condition between X_t and ε_t. And it determines the asymptotic distribution of the DFT. Under the null hypothesis of no structural change, the first component will converge to 0 in the following two scenarios:

- Scenario 1: X_t is nonstationary but is independent of ξ_t.

$$\begin{aligned} & \frac{1}{T}\sum_{t=1}^{T} X_t X_t' \beta(\xi_t) e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}\frac{1}{T}\sum_{t=1}^{T} X_t X_t' \beta(\xi_t) \\ \xrightarrow{p} \; & Q_{xx}(u)E[\beta(\xi_t)] - Q_{xx}(u)Q_{xx}^{-1}Q_{xx}E[\beta(\xi_t)] \\ = \; & 0, \end{aligned}$$

as $T \to \infty$. Intuitively, when ξ_t is weakly stationary and independent of X_t, the first component of the DFT will behave like a projection of X_t using $\hat{M}_t(u)$ defined in Section 3. It will converge to a zero spectrum because $\sum_{t=1}^{T} M_t(u)X_t = 0$ for all u. Therefore, our test is robust to model misspecification when the true DGP is a functional coefficient model (e.g., Cai et al., 2000) with the dependent variable of regression coefficients being independent of X_t. In this case, the only source of rejection for our tests will be structural change.

- Scenario 2: X_t is weakly stationary.

$$\begin{aligned} & \frac{1}{T}\sum_{t=1}^{T} X_t X_t' \beta(\xi_t) e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}\frac{1}{T}\sum_{t=1}^{T} X_t X_t' \beta(\xi_t) \\ \xrightarrow{p} \; & E[X_t X_t' \beta(\xi_t)] \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau - Q_{xx}\int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau Q_{xx}^{-1} E[X_t X_t' \beta(\xi_t)] \\ = \; & 0, \end{aligned}$$

as $T \to \infty$. Intuitively, when both X_t and ξ_t are stationary, our test cannot capture any information about the unknown parameter as a function of time, so the DFT will always converge to 0 for all u.

The derivation above shows a salient feature of testing based on frequency domain analysis. The DFT we proposed in this paper is designed particularly to detect if the unknown parameter is a deterministic function of time. Even when the model is misspecified, as long as the unknown parameter contains no information about the deterministic time trend, our test will be robust to model misspecification. In particular, when both X_t and ξ_t are weakly stationary, the only source of rejection for our test is structural instability. Compared with the existing tests that cannot distinguish the structural change from model misspecification, the advantage of our tests is obvious.

6 Monte Carlo Simulations

In this section, we study the finite sample performance of the proposed tests through Monte Carlo simulations. We consider the time series regressions with exogenous covariates and endogenous covariates respectively. For the case of exogenous covariates, we compare our test with Andrews' (1993) supremum LM test, Bai and Perron's (1998) double maximum test and Chen and Hong's (2012) generalized Hausman test. For the case of endogenous covariates, we compare our test with

Chen's (2015) test.

6.1 Exogenous Covariate

In this subsection, we check the performance of our tests for exogenous covariates. To examine the size and power, we consider the following regressions:
DGP.S1: (No structural change)

$$Y_t = 1 + 0.5X_t + \varepsilon_t$$

DGP.P1: (Single structural break)

$$Y_t = \begin{cases} 1 + 0.5X_t + \varepsilon_t, & \text{if } t \leq 0.3T, \\ 1.2 + X_t + \varepsilon_t, & \text{otherwise;} \end{cases}$$

DGP.P2: (Multiple structural breaks)

$$Y_t = \begin{cases} 1 + 0.1X_t + \varepsilon_t, & \text{if } 0.1T \leq t \leq 0.3T \\ 1 + X_t + \varepsilon_t, & \text{if } 0.7T \leq t \leq 0.9T, \\ 1 + 0.5X_t + \varepsilon_t, & \text{otherwise.} \end{cases}$$

DGP.P3: (Monotonic smooth structural change)

$$Y_t = 1.5 + \beta(t/T)X_t + \varepsilon_t$$

where $\beta(u) = 0.5 + 0.5\{1 + \exp[-20(u - 0.5)]\}^{-1}$.
DGP.P4: (Non-monotonic smooth structural change)

$$Y_t = 1 + \theta(t/T)X_t + \varepsilon_t$$

where $\theta(\tau) = 0.5 + 1.5\exp[-3(\tau - 0.5)^2]$;

To examine the robustness of tests, we follow Chen and Hong (2012) to consider three cases for the error term $\{\varepsilon_t\}$: (i) i.i.d. case $\varepsilon_t \sim i.i.d.N(0,1)$; (ii) ARCH case $\varepsilon_t = \sqrt{h_t}\nu_t$, $h_t = 0.2 + 0.5\varepsilon_{t-1}^2$, $\nu_t \sim i.i.d.N(0,1)$;(iii) Heteroscedasticity case $\varepsilon_t = \sqrt{h_t}\nu_t$, $h_t = 0.2 + 0.5X_t^2$, $\nu_t \sim i.i.d.N(0,1)$. In addition, we show that our tests are robust to structural changes in X_t. To check the finite sample performance of our tests, we also consider three cases for the covariate X_t:
Case 1: no structural change in X_t

$$X_t = 0.5X_{t-1} + \eta_t;$$

Case 2: abrupt structural break in X_t

$$X_t = \begin{cases} 1.5 + 0.5X_{t-1} + \eta_t, & \text{if } t \leq T/4, \\ -0.3X_{t-1} + \eta_t, & \text{otherwise.} \end{cases}$$

Case 3: smooth structural change in X_t

$$X_t = 1 + \alpha(t/T)X_{t-1} + \eta_t$$
$$\alpha(\tau) = 1.5 - 1.5\exp[-3(\tau - 0.5)^2];$$

where the innovation $\eta_t \sim i.i.d.N(0,1)$ is independent of the error term ε_t.

DGP.S1 satisfies the null hypothesis of structural stability and is used to study the sizes of our tests. Specifically, we examine the performance of our tests under various combinations of error terms and covariate, which allow for i.i.d., conditional heteroscedasticity and heteroscedasticity error terms and stable, abrupt structural break and smooth structural change in X_t. DGP.P1-P4 describe various kinds of structural changes, including the single structural break, multiple structural breaks, monotonic and non-monotonic smooth structural change. We check the power property of our tests by using DGP.P1-P4 with various combinations of ε_t and X_t.

For each DGP, we generate 500 data sets of the random sample $\{X_t, Y_t\}_{t=1}^T$ for $T = 100$, 200 and 500. The computation of $\hat{K}$ involves searching over all possible $u \in \mathbb{R}$ which is computationally infeasible. Thanks to the periodicity and symmetry of the trigonometric function, we can compute $\hat{K}$ using a finite interval $\mathcal{U}$ that covers certain full periods of the sine and cosine waves. In this simulation study, we choose $\mathcal{U} = [0.01, 0.02, ..., 1]$. One can choose different grids based on specific problems. Our simulation studies show that the choice of grids has little impact on the performance of our $\hat{K}$ test. We compute $\hat{K}$ as follows:

$$\begin{aligned}\hat{K} &= \max_{u\in\mathcal{U}} T\hat{A}(u)'\hat{A}(u)^* \\ &= \max_{u\in\mathcal{U}} \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T} X_s'X_t\hat{\varepsilon}_s\hat{\varepsilon}_t\cos\left[\frac{2\pi u(s-t)}{T}\right].\end{aligned}$$

For $\hat{C}$, we use the standard normal density weighting function $W(u) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{u^2}{2})$ and compute it as follows:

$$\begin{aligned}\hat{C} &= T\int_{\mathbb{R}} \hat{A}(u)'\hat{A}(u)^*W(u)du \\ &= \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T} X_s'X_t\hat{\varepsilon}_s\hat{\varepsilon}_t\exp\left[-\frac{2\pi^2(s-t)^2}{T^2}\right].\end{aligned}$$

The critical values of $\hat{K}$ and $\hat{C}$ are obtained by the resampling method described in Section 3, and we set the number of resampling in each replication to be $B = 200$. For Chen and Hong's (2012) generalized Hausman test, we follow their paper to use the rule of thumb bandwidth $h = 1/\sqrt{12}T^{-1/5}$ and the uniform kernel. The critical values of Chen and Hong's (2012) test is computed by the wild bootstrap method provided in their paper. Following Andrews (1993), we choose the trimming parameter $\pi = 0.15$ for the tests of Andrews (1993) and Bai and Perron (1998). For Bai and Perron's (1998) test, we follow Chen and Hong (2012) to set the upper bound of the number of breaks as 5. We consider the heteroscedasticity-robust versions of all tests.

Table 1 reports the empirical sizes of various tests at 5% and 10% significant levels when the sample size $T = 100$, 200 and 500. We can see the proposed tests $\hat{K}$ and $\hat{C}$ are robust to unknown structural change in X_t and unknown conditional volatility dynamics of ε_t. The empirical rejection rates of both $\hat{K}$ and $\hat{C}$ are close to the nominal levels in both small and large samples. For other tests, Chen and Hong's (2012) generalized Hausman test also delivers reasonable size. In contract, Andrews' (1993) sup-LM test tends to under-rejection, while Bai and Perron's (1998) test displays strong over-rejection, especially for the cases of $N = 100$ and $N = 200$.

Table 1: Empirical size for DGP.S1 with exogenous covariate

		X_t has no change						X_t has abrupt break						X_t has smooth change					
		N=100		N=200		N=500		N=100		N=200		N=500		N=100		N=200		N=500	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
I.I.D.	$\hat{K}$	4.2	10.2	4.8	11.0	5.0	11.8	5.4	11.6	6.2	11.8	4.4	9.0	6.4	12.6	4.4	7.8	6.2	11.8
	$\hat{C}$	4.4	10.2	5.6	11.0	5.0	11.2	5.4	10.4	6.0	10.0	4.8	8.8	6.6	12.8	4.0	8.0	7.0	11.4
	$\hat{H}_h$	4.4	9.2	4.4	9.6	5.6	11.8	6.4	13.8	7.2	14.2	4.0	9.2	6.4	11.8	4.0	8.8	4.2	8.4
	sup-LM	2.0	5.0	2.2	5.0	4.2	10.2	2.2	5.0	3.2	7.0	4.0	8.6	1.8	5.8	3.0	6.8	4.0	9.0
	UDMax	14.6	23.2	8.0	16.4	7.0	14.2	9.4	15.2	7.8	11.4	2.2	4.2	11.6	17.4	6.2	10.0	5.6	9.2
ARCH	$\hat{K}$	4.6	9.6	3.8	9.4	7.2	13.2	7.0	12.8	6.6	11.4	4.6	8.6	5.4	13.0	6.0	11.2	5.0	10.0
	$\hat{C}$	5.6	9.8	5.4	11.0	6.8	11.8	6.4	11.6	6.4	10.4	4.4	8.2	5.4	11.2	6.0	10.6	5.0	10.0
	$\hat{H}_h$	8.0	13.2	5.8	9.4	7.6	13.2	8.0	13.2	9.0	14.4	5.0	10.4	6.8	13.8	6.0	11.0	6.4	10.8
	sup-LM	2.0	6.2	2.2	5.4	4.2	9.6	1.2	4.8	3.0	6.4	4.6	8.2	1.8	4.4	2.0	5.8	3.4	7.2
	UDMax	14.2	22.0	8.2	13.6	7.6	13.2	10.6	15.2	6.6	10.4	3.0	4.8	11.8	18.8	6.8	11.2	4.4	8.4
Heter.	$\hat{K}$	5.2	13.0	4.8	12.2	3.8	8.6	6.6	13.0	5.8	11.8	4.0	9.6	7.2	13.6	8.0	12.2	5.8	11.4
	$\hat{C}$	5.2	10.8	4.8	10.8	3.2	9.2	5.2	9.4	5.2	11.0	4.4	9.0	8.2	12.4	7.6	12.2	6.0	10.6
	$\hat{H}_h$	5.2	9.8	5.2	9.4	2.6	8.0	4.8	9.2	7.6	13.2	4.4	9.8	3.8	9.2	5.4	10.2	6.8	10.8
	sup-LM	1.4	2.4	1.6	5.0	3.6	7.2	1.0	2.6	2.4	6.2	3.4	7.8	1.2	4.2	1.2	3.6	2.6	6.8
	UDMax	20.4	32.4	15.8	23.6	7.8	14.6	14.6	21.0	8.2	12.8	2.0	4.6	12.6	18.2	9.4	13.0	6.8	10.2

Notes: (i) $\hat{K}$ denotes this paper's KS type test computed using a grid of $\mathcal{U} = [0.01, 0.02, ..., 1]$; (ii) $\hat{C}$ is this paper's CM type test computed using the standard normal weighting function; (iii) $\hat{H}$ denotes Chen and Hong's (2012) generalized Hausman test with bandwidths $h = (1/\sqrt{12})T^{-1/5}$; (iv) sup-LM denotes Andrews' (1993) supremum LM test; (v) UDMax denotes Bai and Perron's (1998) double maximum test; (vi) Number of replication = 500, Number of resampling = 200. The main entries report the average percentage of rejection.

Table 2 provides the empirical powers of various tests for DGP.P1-P4 at the 5% and 10% significant levels when the sample size $T = 100$, 200 and 500. We first consider the single structural break process given by DGP.P1. All tests are powerful against this DGP and Bai and Perron's (1998) test is most powerful for $T = 100$ and 200. As shown in Table 1, Bai and Perron's (1998) test tends to over-reject seriously even under the null hypothesis. Therefore, it is not surprise to see that Bai and Perron's (1998) test has the highest reject rate. The proposed tests $\hat{K}$ and $\hat{C}$ perform better than Andrews' (1993) test and Chen and Hong's (2012) test. Andrews' (1993) test is the worst one when the sample size is small, which may due to the under-rejection of this test for the small sample. Next, we consider the multiple structural breaks process given by DGP.P2. Under DGP.P2, the proposed tests $\hat{K}$ and $\hat{C}$ dominate all other tests. Chen and Hong's (2012) test performs better than Andrews' (1993) test and Bai and Perron's (1998) test. Then, we consider the monotonic smooth structural changes given by DGP.P3. The $\hat{K}$ and $\hat{C}$ tests are most powerful except the case of $T = 100$ when X_t is stable. As we mentioned before, the high rejection rate of Bai and Perron's (1998) test may due to the severe over-rejection when the sample size is small. When the sample size increases, our tests outperform Bai and Perron's (1998) test. Finally, we consider the non-monotonic smooth structural changes. Our tests outperform other tests and the improvement is significant. Andrews' (1993) test almost has no power when the sample size is small. To sum up, our test is powerful to detect bothy smooth and abrupt structural changes. They outperform Chen and Hong's (2012) generalized Hausman test for any cases, which is consistent with our analysis on the relative efficiency between our tests and Chen and Hong's (2012) test.

Table 2: Empirical power for DGP.P1-P4 with exogenous covariate (i.i.d. error term)

		X_t has no change						X_t has abrupt break						X_t has smooth change					
		N=100		N=200		N=500		N=100		N=200		N=500		N=100		N=200		N=500	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
DGP.P1	$\hat{K}$	44.6	57.0	74.8	84.0	99.2	99.6	92.2	96.0	99.8	99.8	100	100	98.6	100	100	100	100	100
	$\hat{C}$	47.8	59.6	79.2	86.4	99.6	99.6	92.0	95.4	99.8	100	100	100	99.4	100	100	100	100	100
	$\hat{H}_h$	38.4	51.8	71.2	80.6	99.0	99.4	76.0	83.8	98.6	99.6	100	100	95.2	97.0	100	100	100	100
	sup-LM	23.2	39.4	69.8	79.8	99.2	99.6	72.0	84.0	99.0	99.8	100	100	93.2	96.8	100	100	100	100
	UDMax	57.8	69.0	81.4	86.6	99.4	99.8	48.6	59.4	74.2	84.4	99.8	100	97.4	98.4	100	100	100	100
DGP.P2	$\hat{K}$	53.4	65.8	82.0	87.0	99.6	99.8	84.0	90.2	98.2	99.2	100	100	91.4	96.2	99.8	100	100	100
	$\hat{C}$	53.8	65.6	81.8	87.2	99.4	99.8	84.6	91.4	98.8	99.4	100	100	92.2	96.2	99.6	100	100	100
	$\hat{H}_h$	37.6	52.4	73.6	84.6	99.4	99.6	63.0	75.4	96.6	98.8	100	100	86.8	91.8	99.6	100	100	100
	sup-LM	22.0	36.0	60.2	73.6	98.4	99.0	37.8	54.2	83.0	92.0	100	100	65.0	75.0	96.0	98.2	100	100
	UDMax	49.6	61.6	72.6	81.2	97.8	98.8	32.2	40.6	64.6	73.8	88.4	93.4	62.4	73.4	89.4	94.0	100	100
DGP.P3	$\hat{K}$	51.4	61.8	83.8	90.4	99.8	99.8	74.8	82.8	97.2	99.0	100	100	99.2	99.8	100	100	100	100
	$\hat{C}$	52.4	61.0	84.0	90.6	99.8	99.8	73.0	79.0	96.6	99.0	100	100	99.2	100	100	100	100	100
	$\hat{H}_h$	35.4	48.0	72.0	81.6	98.6	99.4	54.4	66.8	88.4	93.2	99.8	100	89.4	96.0	100	100	100	100
	sup-LM	27.4	43.8	71.2	82.4	99.6	99.8	40.0	57.0	88.4	94.8	100	100	89.2	95.2	100	100	100	100
	UDMax	53.2	64.2	83.8	88.0	99.6	99.8	32.8	41.2	40.8	55.2	93.4	97.4	96.0	97.2	99.8	100	100	100
DGP.P4	$\hat{K}$	52.8	64.2	86.6	91.6	100	100	72.4	83.8	98.8	99.8	100	100	91.4	96.0	100	100	100	100
	$\hat{C}$	48.0	58.4	83.4	88.6	100	100	65.4	77.4	98.0	99.2	100	100	89.2	93.6	100	100	100	100
	$\hat{H}_h$	36.0	51.0	76.4	84.2	99.6	99.8	49.4	63.4	95.0	97.6	100	100	82.8	89.6	99.8	99.8	100	100
	sup-LM	6.6	16.0	26.4	44.6	91.2	97.6	22.2	35.6	75.2	87.4	100	100	24.6	41.0	76.2	87.8	99.8	100
	UDMax	49.2	59.8	74.6	84.6	99.4	99.8	42.4	53.6	71.6	79.6	99.4	99.6	63.2	74.6	87.0	92.2	99.6	99.6

Notes: See notes in Table 1.

6.2 Endogenous Covariate

In this subsection, we examine the finite sample performance of our tests $\hat{K}^{IV}$ and $\hat{C}^{IV}$ for endogenous covariates. We still use DGP.S1 and DGP.P1-P4 in the last subsection to examine the size and power of our tests. To allow for misspecification of the relationship between endogenous covariate X_t and instrumental variable Z_t, we consider the follow three cases:

Case 1: no structural change

$$X_t = 0.5Z_t + \eta_t$$

Case 2: abrupt structural change

$$X_t = \begin{cases} 0.8Z_t + \eta_t, & \text{if } t \leq T/4, \\ 0.4Z_t + \eta_t, & \text{otherwise.} \end{cases}$$

Case 3: smooth structural change

$$\begin{aligned} X_t &= \gamma(t/T)Z_t + \eta_t \\ \gamma(\tau) &= 3.5\exp[-(4u-1)^2] + 3.5\exp[-(4u-3)^2] - 1.5 \end{aligned}$$

The third case has been considered by Chen (2015). Following Chen (2015), we generate the instrument variable Z_t as follows:

$$Z_t = 1 + 0.5Z_{t-1} + u_t$$

where $u_t \sim i.i.d.N(0,1)$ is independent of ε_t and η_t. To check the robustness of our test, we also consider three cases for the error term ε_t: (i) i.i.d. case: $\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim i.i.d.N\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$; (ii) ARCH case $\varepsilon_t = \sqrt{h_t}\nu_t$, $h_t = 0.2 + 0.5\varepsilon_{t-1}^2$, $\begin{pmatrix} \nu_t \\ \eta_t \end{pmatrix} \sim i.i.d.N\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$;(iii) Heteroscedasticity case $\varepsilon_t = \sqrt{h_t}\nu_t$, $h_t = 0.2 + 0.5X_t^2$, $\begin{pmatrix} \nu_t \\ \eta_t \end{pmatrix} \sim i.i.d.N\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Here, the parameter ρ measures the degree of endogeneity. We set $\rho = 0.6$ in this subsection. We also redo the simulations by setting $\rho = 0.2$ and $\rho = 0.8$ respectively. Although the reject rates of all tests are affected by this parameter, the conclusion is the same as those reported here.

By using a grid $\mathcal{U} = [0.01, 0.02, ..., 1]$ for $\hat{K}^{IV}$ and the standard normal density weighting function $W(u) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{u^2}{2})$ for $\hat{C}^{IV}$, we have

$$\begin{aligned} \hat{K}^{IV} &= \max_{u\in\mathcal{U}} T\hat{A}^{IV}(u)'\hat{A}^{IV}(u)^* \\ &= \max_{u\in\mathcal{U}} \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\hat{X}_s'\hat{X}_t\hat{\varepsilon}_s\hat{\varepsilon}_t\cos\left[\frac{2\pi u(s-t)}{T}\right], \end{aligned}$$

and

$$
\begin{aligned}
\hat{C}^{IV} &= T\int_{\mathbb{R}} \hat{A}^{IV}(u)'\hat{A}^{IV}(u)^{*}W(u)du \\
&= \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\hat{X}_s'\hat{X}_t\hat{\varepsilon}_s\hat{\varepsilon}_t\exp\left[-\frac{2\pi^2(s-t)^2}{T^2}\right],
\end{aligned}
$$

respectively. Compared to the case with exogenous covariates, X_t and $\hat{\varepsilon}_t = Y_t - X_t\hat{\beta}$ are replaced by $\hat{X}_t$ and $\hat{\varepsilon}_t = Y_t - \hat{X}_t'\hat{\beta}_{2sls}$, respectively.

We compare the finite sample power of $\hat{K}^{IV}$ and $\hat{C}^{IV}$ with the consistent test $\hat{H}^{IV}$ proposed by Chen (2015). The critical values of $\hat{K}^{IV}$ and $\hat{C}^{IV}$ are obtained by the resampling method described in Section 4, and we set the number of resampling in each replication to be $B = 200$. Chen's (2015) test involves two different bandwidths, which are used in two different stages. We follow Chen's (2015) Cross-validation method to choose these two bandwidths. For each DGPs, we generate 500 data sets of the random sample $\{X_t, Z_t, Y_t\}_{t=1}^{T}$ for $T = 100$, 200 and 500 respectively.

Table 3: Empirical size for DGP.S1 with endogenous covariate

		X_t has no change						X_t has abrupt break						X_t has smooth change					
		N=100		N=200		N=500		N=100		N=200		N=500		N=100		N=200		N=500	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
I.I.D.	$\hat{K}^{IV}$	5.0	8.8	4.4	8.6	3.6	7.8	4.8	9.2	5.8	9.0	5.0	9.4	6.2	11.4	4.2	8.6	3.4	10.0
	$\hat{C}^{IV}$	4.2	8.4	4.2	8.0	3.6	7.0	5.4	9.2	6.0	9.2	5.0	9.0	6.0	10.0	4.2	8.2	4.0	8.2
	$\hat{H}^{IV}$	3.4	8.6	4.2	8.0	3.8	6.8	3.4	7.8	2.0	3.8	4.2	11.0	3.0	6.8	1.6	3.6	0.6	1.8
ARCH	$\hat{K}^{IV}$	4.2	9.0	5.4	8.8	3.4	8.0	4.6	9.2	4.8	9.4	3.6	9.2	5.6	10.6	4.4	9.6	3.6	9.0
	$\hat{C}^{IV}$	4.0	7.6	4.8	8.2	3.2	7.2	4.8	9.4	5.0	9.6	4.0	8.2	5.4	9.6	4.4	8.6	3.8	8.2
	$\hat{H}^{IV}$	4.4	7.6	6.6	10.0	4.8	9.8	3.4	7.8	0.8	3.0	1.0	2.4	4.0	7.6	0.6	3.4	4.0	6.8
Heter.	$\hat{K}^{IV}$	4.8	9.6	3.8	7.8	4.0	7.8	4.4	9.6	5.0	9.4	3.4	8.0	6.2	11.8	3.8	10.0	3.8	7.6
	$\hat{C}^{IV}$	3.8	8.0	3.2	6.8	4.0	7.8	3.8	9.4	4.0	10.0	3.6	7.6	5.8	11.2	3.6	9.8	3.6	7.2
	$\hat{H}^{IV}$	3.8	6.6	4.0	8.6	4.0	7.6	3.6	9.0	2.6	5.6	2.8	5.0	6.0	10.2	3.0	6.8	3.8	6.2

Notes: (i) $\hat{K}^{IV}$ denotes this paper's KS type test for endogenous regressors computed using a grid of $\mathcal{U} = [0.01, 0.02, ..., 1]$; (ii) $\hat{C}^{IV}$ is this paper's CM type test for endogenous regressors computed using the standard normal weighting function; (iii) $\hat{H}^{IV}$ denotes Chen's (2015) test; (iv) Number of replication = 500, Number of resampling = 200. The main entries report the average percentage of rejection.

Table 3 reports the empirical sizes of these tests at both 5% and 10% significance level when the sample size $T = 100$, 200 and 500. As shown in the table, the proposed tests $\hat{K}^{IV}$ and $\hat{C}^{IV}$ are robust to unknown structural changes in covariate. Our tests deliver reasonable sizes for most cases. They tend to under-reject the null hypothesis slightly for some cases but are still acceptable. Chen's (2015) test also has reasonable size for the cases of X_t has no structural change, but it shows more under-rejection when X_t suffers from structural changes.

Table 4: Empirical power for DGP.P1-P4 with endogenous covariate

		X_t has no change						X_t has abrupt break						X_t has smooth change					
		N=100		N=200		N=500		N=100		N=200		N=500		N=100		N=200		N=500	
		5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
DGP.P1	$\hat{K}^{IV}$	82.0	88.2	99.2	99.8	100	100	75.4	83.2	96.6	98.0	100	100	100	100	100	100	100	100
	$\hat{C}^{IV}$	85.2	91.6	99.6	99.6	100	100	79.8	87.6	98.0	98.8	100	100	100	100	100	100	100	100
	$\hat{H}^{IV}$	35.6	48.8	70.8	78.0	100	100	53.8	65.8	89.0	92.0	100	100	100	100	100	100	100	100
DGP.P2	$\hat{K}^{IV}$	64.6	73.2	90.4	94.0	100	100	56.0	69.0	80.2	87.0	99.0	99.6	99.8	99.8	100	100	100	100
	$\hat{C}^{IV}$	64.8	73.8	90.2	93.8	100	100	57.2	69.0	80.6	87.6	99.6	99.6	99.8	100	100	100	100	100
	$\hat{H}^{IV}$	31.0	41.4	58.6	67.4	95.8	97.8	48.6	62.8	73.8	81.2	97.8	98.8	99.6	99.8	100	100	100	100
DGP.P3	$\hat{K}^{IV}$	67.6	78.4	94.4	97.0	100	100	60.6	69.6	83.2	87.8	99.4	99.6	94.8	96.6	100	100	100	100
	$\hat{C}^{IV}$	68.2	76.6	93.8	97.2	100	100	61.8	69.4	82.8	88.2	99.4	99.6	97.2	98.6	100	100	100	100
	$\hat{H}^{IV}$	30.4	41.6	65.0	69.4	98.4	98.6	47.6	59.4	72.6	78.8	98.2	99.0	94.0	96.0	100	100	100	100
DGP.P4	$\hat{K}^{IV}$	70.4	81.6	95.2	97.4	100	100	65.0	77.2	95.2	97.4	100	100	96.8	98.0	100	100	100	100
	$\hat{C}^{IV}$	64.2	75.0	92.6	95.6	100	100	58.2	71.4	92.2	96.2	100	100	96.8	98.4	100	100	100	100
	$\hat{H}^{IV}$	8.4	14.4	22.2	31.0	55.2	60.6	6.8	13.0	19.4	25.6	58.6	62.6	95.0	97.6	100	100	100	100

Notes: See notes in Table 3.

Table 4 reports the empirical powers of our tests and Chen's (2015) test at the 5% and 10% significance levels when the sample size $T = 100$, 200 and 500. From the table, we could get the following important findings. First, our $\hat{K}^{IV}$ and $\hat{C}^{IV}$ tests are powerful in detecting all forms of structural changes given by DGP.P1-P4 and the simulation results are consistent with our theoretical conclusion that our test is able to detect both sudden structural breaks and smooth structural changes. Second, our tests are more powerful than Chen's (2015) test in all DGPs. In fact, the nonparametric test proposed by Chen (2015) relies on nonparametric estimations for both the structural equation and the first stage reduced form. While our tests are based on the 2SLS using the whole sample. Third, the unknown structural change in the covariates and instruments may help to increase the power of our tests.

7 Empirical Application to Taylor Rule

In this section, we apply our tests to check whether the U.S. Taylor rule suffer from structural changes. Taylor rule, firstly proposed by Taylor (1993), describes central banks' behavior through a linear function of interest rate to inflation gap and output gap. It has been a benchmark of central banks when conducting monetary policy to stabilize price and smooth output fluctuations. After Taylor (1993), Clarida et al. (1998, 2000) extend the rule to the forward-looking version by incorporating central banks' forward-looking behavior and tendency to smooth changes in interest rates. However, some recent studies argue that the reaction function of the interest rate to inflation and output gap may suffer from structural changes and try to model the structural changes by employing nonlinear models. For example, Kim and Nelson (2006) and Boivin (2006) adopt time varying parameter models to fit the U.S. data during the past 50 years. Brüggemann and Riedel (2011) use logistic smooth transition regression (LSTR) models with time varying parameters to estimate the U.K. monetary policy reaction function and Zheng et al. (2012) introduce a regime-switching model to examine the reaction of China's monetary policy. However, all these papers model the non-constancy of parameters directly without formally testing the significance of instability behavior of the parameters. In fact, most of the 90% confidence bands of model parameters given in those papers cover a horizontal line through the whole period. Based on that, the existence of structural change is not evident.

Following Clarida et al. (1998, 2000) and Kim and Nelson (2006), the forward-looking Taylor rule with smooth interest rate adjustment could be written as follows:

$$r_t = (1-\rho)[c + \beta\pi_{t+p} + \gamma y_{t+q}] + \rho r_{t-1} + \varepsilon_t, \tag{7.1}$$

where r_t is the nominal short-term interest rate at time t, π_{t+p} is the inflation at time $t+p$, and y_{t+q} is the output gap at time $t+q$. The error $\varepsilon_t = -(1-\rho)\{\beta[\pi_{t+p} - E_t(\pi_{t+p})] + \gamma[y_{t+q} - E_t(y_{t+q})]\} + m_t$, where E_t is the conditional expectation operator. Conditional on information available to the monetary authority at time t, $E_t(\pi_{t+p})$ is the p-th periods ahead forecast for inflation at time t, and m_t is a random error term caused by the central bank's control of interest rate. $c = \bar{r} - (\beta-1)\pi^*$, where $\bar{r}$ is the equilibrium real interest rate, and π^* is the target value of inflation. Suppose we define $\beta_0 = (1-\rho)c$, $\beta_1 = (1-\rho)\beta$, $\beta_2 = (1-\rho)\gamma$ and $\beta_3 = \rho$, then Eq. (7.1) could be written as:

$$r_t = \beta_0 + \beta_1\pi_{t+p} + \beta_2 y_{t+q} + \beta_3 r_{t-1} + \varepsilon_t \tag{7.2}$$

If $p < 0$ and $q < 0$, then Eq. (7.2) is the backward-looking Taylor rule, and $\varepsilon_t = m_t$. In this case, the parameters in (7.2) could be consistently estimated via OLS. If $p = 0$ and $q = 0$, then (7.2) is contemporary Taylor rule. If $p > 0$ and $q > 0$, then (7.2) is the forward-looking Taylor rule. As shown by Kim and Nelson (2006), the contemporary and forward-looking Taylor rules suffer from endogenous problem since the error term ε_t contains forecast errors, and should be estimated by IV method. As in Clarida et al. (1998, 2000) and Kim and Nelson (2006), we consider $p = q = 1$ for the forward-looking Taylor rule.

Using the U.S. quarterly data from 1960Q1 to 2018Q1, we investigate the stability of the backward-looking, contemporary and forward-looking monetary reaction functions by the proposed frequency domain based tests. We apply the frequency domain based tests proposed in Section 3, Andrews' (1993) supremum LM test and Chen and Hong's (2012) generalized Hausman test to check the stability of backward-looking Taylor rule corresponding to $p = -1, q = -1$ and apply the tests proposed in Section 4 and Chen's (2015) generalized Hausman test to check the stability of contemporary and forward-looking Taylor rules corresponding to $p = 0, q = 0$ and $p = 1, q = 1$ respectively. As in Clarida et al. (1998, 2000) and Kim and Nelson (2006), the interest rate is measured by the average Federal Funds rate in the first month of each quarter, expressed in annal rates, the inflation measure is the annualized rate of change of the GDP deflator between two subsequent quarters, and the output gap is the series constructed by the Congressional Budget Office. The set of instruments includes four lags of each of the following variables: the federal funds rate, output gap, inflation, commodity price inflation, M2 growth, and the spread between the long-term bond rate and the three-month Treasury Bill rate.

Table 5: Testing for structural changes in Taylor rules

	$\hat{C}$	$\hat{K}$	$\sup LM$	$\hat{H}$
	1960Q1-2018Q1: the whole sample			
$p=-1, q=-1$	0.003	0.019	0.041	0.004
$p=0, q=0$	0.000	0.002	–	0.872
$p=1, q=1$	0.000	0.000	–	0.908
	1960Q1-1996Q4: Clarida et al. (2000)			
$p=-1, q=-1$	0.002	0.014	0.002	0.012
$p=0, q=0$	0.000	0.001	–	0.824
$p=1, q=1$	0.000	0.004	–	0.918
	1960Q1-2001Q2: Kim and Nelson (2006)			
$p=-1, q=-1$	0.000	0.006	0.001	0.008
$p=0, q=0$	0.000	0.000	–	0.876
$p=1, q=1$	0.000	0.002	–	0.884
	1960Q1-1979Q2: Clarida et al. (2000)'s pre-1979			
$p=-1, q=-1$	0.028	0.039	0.000	0.002
$p=0, q=0$	0.091	0.127	–	0.872
$p=1, q=1$	0.534	0.393	–	0.964
	1979Q3-1996Q4: Clarida et al. (2000)'s post-1979			
$p=-1, q=-1$	0.043	0.043	0.001	0.006
$p=0, q=0$	0.008	0.006	–	0.364
$p=1, q=1$	0.046	0.039	–	0.482

Notes: (i) numbers in the main entries are p-values; (ii) $\hat{C}$ and $\hat{K}$ denote the proposed frequency domain based tests, which are calculated by the OLS based tests given in Section 3 for the $p=-1, q=-1$, and are calculated by the IV based tests given in section 4 for $p=0, q=0$ and $p=1, q=1$; (iii) supLM is Andrews' (1993) supremum LM test; (iv) $\hat{H}$ is the generalized Hausman test, which is calculated according to Chen and Hong (2012) for the case of $p=-1, q=-1$, and is calculated according to Chen (2015) for the other cases. (v) the p-values of $\hat{C}$, $\hat{K}$ and Chen and Hong's (2012) test are based on 2000 bootstrap iterations, while the p-values of Chen's (2015) test are based on 500 bootstrap iterations due to the time consuming CV bandwidth selection. (vi) $\hat{K}$ is calculated using a grid of $\mathcal{U} = [0.01, 0.02, \cdots, 1]$ with 100 grid point in the set $[0, 1]$.

Table 5 reports the results of the proposed tests $\hat{C}$ and $\hat{K}$, Andrews' (1993) supremum LM test as well as Chen and Hong's (2012) or Chen's (2015) generalized Hausman test. We first check the stability of Taylor rules for the whole sample from 1960Q1 to 2018Q1. The p-values of the proposed tests based on bootstrapped distributions all less than 5% for any cases. This results document the existence of structural changes in various Taylor rules. We also check the stability of Taylor rules for the sub-samples 1960Q1-1996Q4 and 1960Q1-2001Q2, which are the samples considered by Clarida et al. (2000) and Kim and Nelson (2006) respectively. The p-values of the proposed tests reported in the second and third parts of Table 5 also less than 5% for all cases, show that there exists substantial evidence of structural instabilities of the Taylor rules during these periods. Clarida et al. (2000) point out that there is a significant difference in the way of monetary policy was conducted pre- and post-1979, the year Paul Volcker was appointed the Chairman of the Board of Governors of the Federal Reserve System. Therefore, they divide the sample into two sub-periods: the pre-1979 period from 1960Q1 to 1979Q2 and the post-1979 period from 1979Q3 to 1996Q4 and estimate the stable forward-looking Taylor rule for each period. The fourth and fifth parts of Table 5 report the test results for these two periods. According to the results of the proposed tests, we find no significant evidence against the stability of the contemporary and forward-looking Taylor rules for the first period, which encompasses the tenures of William M. Martin, Arthur Burns and G. William Miller as Federal Reserve chairmen. However, the results reveal the existence of structural changes for Taylor rules during the second period, which is supposed to be stable by Clarida et al. (2000).

Table 5 also reports the results of Chen and Hong's (2012) generalized Hausman test and Andrews' (1993) supremum LM test for the backward-looking Taylor rule and the results of Chen's (2015) generalized Hausman test for the contemporary and forward-looking Taylor rules. Our experience shows that the results of Andrews' (1993) test are affected by the choice of trimming parameter. For example, if we set the trimming parameter $\pi_0 = 0.15$, the p-value for the whole sample is 0.041, while it changes to 0.063 if we set $\pi_0 = 0.05$, which will fail to reject the null of no structural changes at the 5% level. Table 5 reports the results by setting $\pi_0 = 0.15$. For the backward-looking Taylor rule, the results of Chen and Hong's (2012) generalized Hausman test and Andrews' (1993) supremum LM test are consistent with our frequency domain based tests for exogenous variable. However, the results of Chen's (2015) test can not detect the possible structural changes in contemporary and forward-looking Taylor rules for any samples. This is consistent with our theoretical results and simulation studies that our tests are asymptotically more powerful than Chen's (2015) test.

To sum up, our results reveal the existence of structural changes in Taylor rules. In particular, we find that the Taylor rules in the post-1979 period from 1979Q3 to 1996Q4 suffer from structural changes, which is treated as stable by Clarida et al. (2000). This empirical findings are not only useful in understanding the central banks' monetary policy, but also justify the use of nonlinear models for Taylor rules.

8 Conclusion

This paper proposes a novel DFT-based approach to testing for structural change in a linear time series regression model, which is consistent against both abrupt structural breaks and smooth structural changes. It avoids smoothed nonparametric estimation of the unknown model parameter. Therefore it is tuning parameter-free and can detect a class of local alternatives at the parametric rate, which is asymptotically more efficient than the existing smoothed nonparametric tests. Our approach applies to linear time series regression models with both exogenous and endogenous variables. In particular, it is robust to structural changes of unknown type in both regressors and instruments. Furthermore, we show that our tests are robust to certain model misspecification such that the only source of rejection is the structural change.

Our tests can also be extended to many other frameworks where the constancy of unknown parameters is investigated. The only input we need is sample moment condition which is identical to 0 implied by the estimation method. To extract the local information contained in the unknown parameter, we can use the DFT as a filter to capture it. The testing method we employed in the linear time series model is based on the first order condition. That implies we can test the constancy of parameter characterized by moment conditions for multiple equations (e.g., DSGE models). We can also apply our testing method to test structural change in volatility dynamics, where the only input we need is the score function. Compared to the test based on the local QMLE in Chen and Hong (2016), our DFT-based approach is much easier to implement and will be asymptotically more powerful. Moreover, based on the discussion in Section 5, our tests can also examine the constancy of a functional coefficient model by replacing the Fourier basis functions of time with the Fourier basis functions of a random variable. In this way, nonparametric smoothed estimation can be avoided.

Mathematical Appendix

Throughout the appendix, "$\xrightarrow{p}$", "$\xrightarrow{d}$", and "$\Rightarrow$"denote convergence in probability, convergence in distribution, and weak convergence respectively. We use the same notations as in the paper and we denote C to be a generic bounded constant, A^* to be the conjugate transpose of $A \in \mathbb{C}$, and $\mathrm{Re}(A)$ to be the real part of $A \in \mathbb{C}$.

A Technical Lemmas

Lemma A.1 *Suppose Assumptions 3.1 to 3.5 hold,*

$$\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]e^{iu2\pi t/T},$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\varepsilon_t e^{iu2\pi t/T},$$

are stochastically equicontinuous.

Proof of Lemma A.1. We first show that $\frac{1}{T}\sum_{t=1}^{T}\left[X_tX_t' - E(X_tX_t')\right]e^{\mathbf{i}u2\pi t/T}$ is stochastically equicontinuous, i.e., we need to show that, for any $\epsilon > 0$ and $\kappa > 0$, there exists $\delta > 0$ such that

$$\lim_{T\to\infty} P\left[\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta}\left\|\frac{1}{T}\sum_{t=1}^{T}\left[X_tX_t' - E(X_tX_t')\right]\left(e^{\mathbf{i}u_1 2\pi t/T} - e^{\mathbf{i}u_2 2\pi t/T}\right)\right\| > \kappa\right] < \epsilon.$$

Let $\bar{u} = au_1 + (1-a)u_2$ for some $a \in (0,1)$ and

$$e^{\mathbf{i}u_1 2\pi t/T} = e^{\mathbf{i}u_2 2\pi t/T} + \mathbf{i}2\pi t/Te^{\mathbf{i}\bar{u}2\pi t/T}(u_1 - u_2),$$

then

$$\begin{aligned}
& \lim_{T\to\infty} P\left[\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta}\left\|\frac{1}{T}\sum_{t=1}^{T}\left[X_tX_t' - E(X_tX_t')\right]\left(e^{\mathbf{i}u_1 2\pi t/T} - e^{\mathbf{i}u_2 2\pi t/T}\right)\right\| > \kappa\right] \\
= & \lim_{T\to\infty} P\left[\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta}\left\|\frac{1}{T}\sum_{t=1}^{T}\left[X_tX_t' - E(X_tX_t')\right]\left(\mathbf{i}2\pi t/Te^{\mathbf{i}\bar{u}2\pi t/T}\right)(u_1-u_2)\right\| > \kappa\right] \\
\leq & \lim_{T\to\infty} P\left[\sup_{\bar{u}\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left[X_tX_t' - E(X_tX_t')\right]\left(\mathbf{i}2\pi t/Te^{\mathbf{i}\bar{u}2\pi t/T}\right)\right\| > \kappa/\delta\right] \\
\leq & \lim_{T\to\infty} P\left[\sup_{\bar{u}\in\mathbb{R}}\frac{1}{T}\sum_{t=1}^{T}\left\|\left[X_tX_t' - E(X_tX_t')\right]\left(\mathbf{i}2\pi t/Te^{\mathbf{i}\bar{u}2\pi t/T}\right)\right\| > \kappa/\delta\right] \\
\leq & \lim_{T\to\infty} P\left[\sqrt{\frac{1}{T}\sum_{t=1}^{T}\|X_tX_t' - E(X_tX_t')\|^2}\sup_{\bar{u}\in\mathbb{R}}\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left\|\left(\mathbf{i}2\pi t/Te^{\mathbf{i}\bar{u}2\pi t/T}\right)\right\|^2} > \kappa/\delta\right] \\
= & \lim_{T\to\infty} P\left[\sqrt{\frac{1}{T}\sum_{t=1}^{T}\|X_tX_t' - E(X_tX_t')\|^2}\sqrt{\frac{1}{T}\sum_{t=1}^{T}4\pi t^2/T^2} > \kappa/\delta\right]
\end{aligned}$$

$$< \quad \epsilon,$$

where the third to last inequality is by triangle inequality and the second to last is by Cauchy-Swartz inequality. As is shown in Andrews (1994), the last inequality holds since we always find a $\delta > 0$ small enough given $\frac{1}{T}\sum_{t=1}^{T}\left\|[X_tX_t' - E(X_tX_t')]\right\|^2$ is $O_P(1)$, and $\frac{1}{T}\sum_{t=1}^{T} 4\pi t^2/T^2$ is $O(1)$. By analogous argument, we can show $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t\varepsilon_t e^{\mathbf{i}u2\pi t/T}$ is also stochastically equicontinuous. ■

Lemma A.2 *Suppose Assumptions 3.1 to 3.5 hold, then as $T \to \infty$,*

$$\sup_{u\in\mathbb{R}}\left\|\hat{Q}_{xx}(u) - Q_{xx}(u)\right\| \xrightarrow{p} 0.$$

Proof of Lemma A.2.

$$\begin{aligned}
& \sup_{u\in\mathbb{R}}\left\|\hat{Q}_{xx}(u) - Q_{xx}(u)\right\| \\
= & \sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\,e^{\mathbf{i}u2\pi t/T}\right\| \\
\le & \sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\cos(u2\pi t/T)\right\| + \sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\sin(u2\pi t/T)\right\| \\
\equiv & R_1 + R_2, \text{ say.}
\end{aligned}$$

Next, we show that $R_1 = o_P(1)$. Let $\mathbb{D}$ be the space of all functions $\theta : [0,1] \to [-1,1]$, where $\theta(\tau) = \cos(u2\pi\tau)$ for $\tau = t/T$, $t = 1,2,...,T$, and $\theta(\tau) = 0$ otherwise. Therefore,

$$\begin{aligned}
\lim_{T\to\infty} P(R_1 > 2\kappa) \le & \lim_{T\to\infty} P\left(\sup_{\theta\in\mathbb{D}}\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\,\theta(\tau)\right\| > 2\kappa\right) \\
\le & \lim_{T\to\infty} P\Bigg(\max_{j\le J}\sup_{\tilde{\theta}\in B(\theta_j,\delta)}\Bigg\{\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\left[\tilde{\theta}(\tau) - \theta_j(\tau)\right]\right\| \\
& + \left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\,\theta_j(\tau)\right\|\Bigg\} > 2\kappa\Bigg) \\
\le & \lim_{T\to\infty} P\left(\sup_{\theta\in\mathbb{D}}\sup_{\tilde{\theta}\in B(\theta_j,\delta)}\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\left[\tilde{\theta}(\tau) - \theta_j(\tau)\right]\right\| > \kappa\right) \\
& + \lim_{T\to\infty} P\left(\max_{j\le J}\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\,\theta_j(\tau)\right\|\Bigg\} > \kappa\right) \\
< & \kappa,
\end{aligned}$$

where we let $\{B(\theta_j,\delta) : j = 1,2,...,J\}$ be a finite cover of $\mathbb{D}$ such that $\theta \in B(\theta_j,\delta)$ if and only if $d(\theta,\theta_j) \equiv \sqrt{\int_0^1 |\theta(\tau) - \theta_j(\tau)|^2 d\tau} \le \delta$. To let the last inequality hold, we need to show: (i). $\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\,\theta(\tau)$ is stochastically equicontinuous; and (ii). $\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\,\theta(\tau) = o_P(1)$ for any $\theta \in \mathbb{D}$.

For (i):

$$\lim_{T\to\infty} P\left[\sup_{\theta_1,\theta_2\in\mathbb{D}:d(\theta_1,\theta_2)<\delta}\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\,(\theta_1(\tau) - \theta_2(\tau))\right\| > \kappa\right]$$

$$\leq \lim_{T\to\infty} P\left[\left\|\frac{1}{T}\sum_{t=1}^{T}[X_tX_t' - E(X_tX_t')]\right\| > \kappa/\delta\right] < \epsilon,$$

by analogous arguments in the proof of Lemma A.1. The point-wise convergence (ii) is easy to verify given Assumptions 3.1, 3.2, 3.5, and the boundedness condition of $\mathbb{D}$. By analogous proof, we can show that $R_2 = o_P(1)$. Thus, we have proved the result. ■

B Proof of Theorems

Proof of Theorem 3.1. Under $\mathbb{H}_0$, $\hat{A}_1(u) = 0$, thus

$$\begin{aligned}
\sqrt{T}\hat{A}(u) &= \sqrt{T}\hat{A}_2(u) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{M}_t(u)\varepsilon_t \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}M_t(u)\varepsilon_t \\
&\quad +\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\hat{M}_t(u) - M_t(u)\right]\varepsilon_t.
\end{aligned}$$

We first show that the second term converge to 0 uniformly over all u:

$$\begin{aligned}
&\sup_{u\in\mathbb{R}}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\hat{M}_t(u) - M_t(u)\right]\varepsilon_t\right\| \\
=\ &\sup_{u\in\mathbb{R}}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1} - Q_{xx}(u)Q_{xx}^{-1}\right]X_t\varepsilon_t\right\| \\
=\ &\sup_{u\in\mathbb{R}}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1} - Q_{xx}(u)Q_{xx}^{-1}\right]X_t\varepsilon_t\right\| \\
=\ &\sup_{u\in\mathbb{R}}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1} - \hat{Q}_{xx}(u)Q_{xx}^{-1} + \hat{Q}_{xx}(u)Q_{xx}^{-1} - Q_{xx}(u)Q_{xx}^{-1}\right]X_t\varepsilon_t\right\| \\
\leq\ &\sup_{u\in\mathbb{R}}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{Q}_{xx}(u)\left[\hat{Q}_{xx}^{-1} - Q_{xx}^{-1}\right]X_t\varepsilon_t\right\| + \sup_{u\in\mathbb{R}}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\hat{Q}_{xx}(u) - Q_{xx}(u)\right]Q_{xx}^{-1}X_t\varepsilon_t\right\| \\
\leq\ &\sup_{u\in\mathbb{R}}\left\|\hat{Q}_{xx}(u)\right\|\left\|\hat{Q}_{xx}^{-1} - Q_{xx}^{-1}\right\|\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\varepsilon_t\right\| + \sup_{u\in\mathbb{R}}\left\|\hat{Q}_{xx}(u) - Q_{xx}(u)\right\|\left\|Q_{xx}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\varepsilon_t\right\| \\
=\ &O_P(1)\cdot o_P(1)\cdot O_P(1) + o_P(1)\cdot O_P(1) = o_P(1),
\end{aligned}$$

where the last equality is by Lemma A.2, and Assumptions 3.1-3.5. Define

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}M_t(u)\varepsilon_t \equiv S_T(u),$$

then $\sqrt{T}\hat{A}(u) = S_T(u) + o_P(1)$. We want to first show that $S_T(u)$ converges in distribution to a normal distribution for each fixed $u \in \mathbb{R}$. Then we show that $S_T(u)$ is stochastically equicontinuous and establish the weak convergence result.

Let

$$
\begin{aligned}
U_T(u) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \Lambda' M_t(u) \varepsilon_t \\
&= \sum_{t=1}^{T} \left[\frac{1}{\sqrt{T}} \sum_{j=1}^{d} \Lambda_j M_{jt}(u) \varepsilon_t \right],
\end{aligned}
$$

where $\Lambda \in \mathbb{R}^d$ is any nonzero vector such that $\Lambda'\Lambda = 1$, and Λ_j and $M_{jt}(u)$ represent the jth entry of $d \times 1$ column vectors Λ and $M_t(u)$ respectively.

Furthermore we define

$$
\begin{aligned}
\zeta_T^2(u) &\equiv E[U_T(u)^2] \\
&= \frac{1}{T} \sum_{s,t=1}^{T} \sum_{j,k=1}^{d} E\left[\Lambda_j M_{jt}(u) \varepsilon_t \Lambda_k M_{ks}(u)^* \varepsilon_s\right] \\
&= \frac{1}{T} \sum_{j,k=1}^{d} \Lambda_j \Lambda_k \left[\sum_{s,t=1}^{T} E[M_{jt}(u) \varepsilon_t M_{ks}(u)^* \varepsilon_s] \right] \\
&= \frac{1}{T} \sum_{j,k=1}^{d} \Lambda_j \Lambda_k \left[\sum_{t=1}^{T} E[M_{jt}(u) M_{kt}(u)^* \varepsilon_t^2] \right] \\
&= O(1).
\end{aligned}
$$

for all $u \in \mathbb{R}$, where the second to last equality comes from the fact that $E(\varepsilon_t | I_{t-1}) = 0$ almost surely, and last equality is due to the fact that $E|X_{tj}|^{4\delta} < \infty$, $E|\varepsilon_t|^{4\delta} < \infty$ and $\Lambda_j^2 < 1$ for all $j = 1, 2, ..., d$.

By Cramer-Wold device, to show joint normality, we just need to show

$$
\zeta_T(u)^{-1} U_T(u) \xrightarrow{d} N(0, 1),
$$

for each fixed $u \in \mathbb{R}$ and all Λ.

Let

$$
\Psi_t(u) \equiv \frac{1}{\sqrt{T}} \sum_{j=1}^{d} \Lambda_j X_{jt}(u) \varepsilon_t,
$$

then $U_T(u) = \sum_{t=1}^{T} \Psi_t(u)$. By Theorem 1.1 in Bradley and Tone (2015), we need to show the following:

(i) $\alpha'(\{\Psi_t(u)\}_{t=1}^{T}, m) \to 0$, as $m \to \infty$;

(ii) $\rho'(\{\Psi_t(u)\}_{t=1}^{T}, 1) < 1$;

(iii) $\zeta_T^2(u) > 0$;

(iv) $E(\Psi_t(u)^2) = \frac{1}{T} \sum_{j,k=1}^{d} \Lambda_j \Lambda_k E[M_{jt}(u) M_{kt}(u)^* \varepsilon_t^2] < \infty$;

(v) The Lindeberg condition holds:

$$\lim_{T\to\infty} \frac{1}{\zeta_T^2(u)} \sum_{t=1}^{T} E\left[\Psi_t(u)^2 \mathbf{1}\left(|\Psi_t(u)| > \zeta_T(u)\epsilon\right)\right] = 0,\ \forall \epsilon > 0.$$

First, Condition (i) can be implied by Assumption 3.1.

$$\begin{aligned}
\alpha'(\{\Psi_t(u)\}_{t=1}^T, m) &= \sup_{1\le t\le T-m} \alpha\left[\sigma\left(\sum_{j=1}^{d} \Lambda_j M_{jt}(u)\varepsilon_t\right), \sigma\left(\sum_{j=1}^{d} \Lambda_j M_{jt+m}(u)\varepsilon_{t+m}\right)\right] \\
&= \sup_{1\le t\le T-m} \alpha[\sigma(X_t\varepsilon_t), \sigma(X_{t+m})\varepsilon_{t+m}] \\
&\le Cm^{-\nu},
\end{aligned}$$

for all u. Thus $\alpha'(\{\Psi_t(u)\}_{t=1}^T, m) \to 0$ as $m \to \infty$.

For condition (ii), notice that $\Psi_t(u)$ is martingale difference sequence process:

$$E(\Psi_t(u)|I_{t-1}) = E\left(\frac{1}{\sqrt{T}} \sum_{j=1}^{d} \Lambda_j M_{jt}(u)\varepsilon_t \middle| I_{t-1}\right) = 0,$$

almost surely.

Thus

$$\rho'(\{\Psi_t(u)\}_{t=1}^T, 1) \quad = \quad 0,$$

where the maximal coefficient of correlation

$$\rho(\mathcal{A}, \mathcal{B}) \equiv \sup_{f\in L^2(\mathcal{A}), g\in L^2(\mathcal{B})} |corr(f, g)|,$$

and $L^2(\mathcal{A})$ consists any $\mathcal{A}$-measurable random variables with finite second moments.

Condition (iii) has been shown by $\zeta_T^2(u) = O(1)$ and thus is satisfied.

Condition (iv) is satisfied due to $E|X_{tj}|^{4\delta} < \infty$, $E|\varepsilon_t|^{4\delta} < \infty$ and $\tau_j^2 < 1$ for all $j = 1, 2, ..., d$.

Lastly, we verify the Lindeberg condition: $\forall \epsilon > 0$,

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{\zeta_T^2(u)} \sum_{t=1}^{T} E\left[\Psi_t(u)^2 \mathbf{1}\left(|\Psi_t(u)| > \zeta_T(u)\epsilon\right)\right] \\
\le\ &\lim_{T\to\infty} \frac{1}{\zeta_T^2(u)} \sum_{j=1}^{d} \left(E\left[\Psi_t(u)^4\right]\right)^{1/2} E\left[\mathbf{1}\left(|\Psi_t(u)| > \zeta_T(u)\epsilon\right)\right]^{1/2} \\
\le\ &\lim_{T\to\infty} \frac{1}{\zeta_T^2(u)} \sum_{j=1}^{d} \left(E\left[\Psi_t(u)^4\right]\right)^{1/2} \left(\frac{1}{\zeta_T^2(u)\epsilon^2} E\left[\Psi_t(u)^2\right]\right)^{1/2} \\
=\ &\lim_{T\to\infty} \frac{1}{\zeta_T^3(u)\epsilon} \sum_{j=1}^{d} \left(E\left[\Psi_t(u)^4\right]\right)^{1/2} \left(E\left[\Psi_t(u)^2\right]\right)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{T\to\infty} \frac{1}{\zeta_T^3(u)\epsilon} \sum_{j=1}^{d} \left(\frac{1}{T^2} E\left[\sum_{j=1}^{d} \Lambda_j X_{jt}(u)\varepsilon_t \right]^4 \right)^{1/2} \left(\frac{1}{T} E\left[\sum_{j=1}^{d} \Lambda_j X_{jt}(u)\varepsilon_t \right]^2 \right)^{1/2} \\
&= O(T^{-3/2}) = o(1).
\end{aligned}
$$

where the first inequality is due to the Cauchy-Schwartz inequality given $E[\Psi_t(u)^2] = E[|\Psi_t(u)^2|]$, $E[\mathbf{1}\left(|\Psi_t(u)| > \zeta_T(u)\epsilon\right)] = E[|\mathbf{1}\left(|\Psi_t(u)| > \zeta_T(u)\epsilon\right)|]$ due to the non-negativity of indicator function, and $E[\mathbf{1}\left(|\Psi_j(T,u)| > \zeta_T(u)\epsilon\right)^2] = E[\mathbf{1}\left(|\Psi_j(T,u)| > \zeta_T(u)\epsilon\right)]$. The second inequality is due to the Chebyshev's Inequality. Thus, we have verified condition (iv).

By Theorem 1.1 in Bradley and Tone (2015), for each $u \in \mathbb{R}$,

$$
\zeta_T(u)^{-1} \sum_{t=1}^{T} \Psi_t(u) \xrightarrow{d} N(0,1).
$$

Since this result holds for all $\Lambda'\Lambda = 1$, by Cramer-Wold Device,

$$
S_T(u) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} M_t(u)\varepsilon_t \xrightarrow{d} N[0, \mathcal{K}(u,u)],
$$

for each fixed $u \in \mathbb{R}$, where

$$
\mathcal{K}(u,u) \quad = \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E[M_t(u) M_t(u)^{*'} \varepsilon_t^2].
$$

Next, we show $S_T(u)$ is stochastically equicontinuous over $u \in \mathbb{R}$, i.e., we need to show that, for any $\epsilon > 0$ and $\kappa > 0$, there exists $\delta > 0$ such that

$$
\lim_{T\to\infty} P\left[\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta} \|S_T(u_1) - S_T(u_2)\| > \kappa \right] < \epsilon.
$$

Given

$$
\begin{aligned}
& \sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta} \|S_T(u_1) - S_T(u_2)\| \\
\leq \quad & \sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[X_t(u_1) - X_t(u_2)\right] \varepsilon_t \right\| \\
& + \sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta} \|Q_{xx}(u_1) - Q_{xx}(u_2)\| \left\| Q_{xx}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_t\varepsilon_t \right\|,
\end{aligned}
$$

$S_T(u)$ is stochastically equicontinuous by Lemma A.1 and the fact that $|\left\| Q_{xx}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_t\varepsilon_t \right\| = O_P(1)$. Therefore, we have

$$
S_T(u) \Rightarrow \mathcal{G}(u)
$$

where $\mathcal{G}(u)$ is a zero-mean complex-valued Gaussian process with covariance kernel

$$\mathcal{K}(u_1,u_2) \quad = \quad \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E[M_t(u_1)M_t(u_2)^{*'}\varepsilon_t^2].$$

Furthermore, when X_t is weakly stationary

$$\begin{aligned} E[M_t(u_1)M_t(u_2)^{*'}\varepsilon_t^2] &= E\left\{[X_t(u_1)-Q_{xx}(u_1)Q_{xx}^{-1}X_t][X_t(u_2)-Q_{xx}(u_2)Q_{xx}^{-1}X_t]^{*'}\varepsilon_t^2\right\} \\ &= E\left(X_tX_t'\varepsilon_t^2\right)\left(e^{\mathbf{i}u_1 2\pi t/T}-\int_0^1 e^{\mathbf{i}u_1 2\pi\tau}d\tau\right)\left(e^{\mathbf{i}u_2 2\pi t/T}-\int_0^1 e^{\mathbf{i}u_2 2\pi\tau}d\tau\right)^{*}. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathcal{K}(u_1,u_2) &= E\left(X_tX_t'\varepsilon_t^2\right)\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left(e^{\mathbf{i}u_1 2\pi t/T}-\int_0^1 e^{\mathbf{i}u_1 2\pi\tau}d\tau\right)\left(e^{\mathbf{i}u_2 2\pi t/T}-\int_0^1 e^{\mathbf{i}u_2 2\pi\tau}d\tau\right)^{*} \\ &= E\left(X_tX_t'\varepsilon_t^2\right)\left(\int_0^1 e^{\mathbf{i}2\pi(u_1-u_2)\tau}d\tau-\int_0^1 e^{\mathbf{i}2\pi u_1\tau}d\tau\int_0^1 e^{-\mathbf{i}2\pi u_2\tau}d\tau\right). \end{aligned}$$

■

Proof of Theorem 3.2. Under $\mathbb{H}_A$,

$$\begin{aligned} \hat{A}(u) &= \hat{A}_1(u)+\hat{A}_2(u) \\ &= \hat{A}_1(u)+o_P(1). \end{aligned}$$

Let

$$\tilde{A}(u)=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\left[M_t(u)X_t'\right]\beta_t,$$

we have

$$\begin{aligned} &\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\hat{A}_1(u)-\tilde{A}(u)\right\| \\ \leq\ &\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\hat{M}_t(u)X_t'-M_t(u)X_t'\right\}\beta_t\right\|+\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{M_t(u)X_t'-E\left[M_t(u)X_t'\right]\right\}\beta_t\right\| \\ =\ &\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\left[Q_{xx}(u)Q_{xx}^{-1}-\hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}\right]\frac{1}{T}\sum_{t=1}^{T}X_tX_t'\beta_t\right\| \\ &+\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{X_t(u)X_t'-E[X_t(u)X_t']\right\}\beta_t-Q_{xx}(u)Q_{xx}^{-1}\frac{1}{T}\sum_{t=1}^{T}[X_tX_t'-E(X_tX_t')]\beta_t\right\| \\ \leq\ &\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|Q_{xx}(u)Q_{xx}^{-1}-\hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}\right\|\left\|\frac{1}{T}\sum_{t=1}^{T}X_tX_t'\beta_t\right\| \\ &+\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{X_t(u)X_t'-E[X_t(u)X_t']\right\}\beta_t\right\| \\ &+\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\|Q_{xx}(u)\|\left\|Q_{xx}^{-1}\frac{1}{T}\sum_{t=1}^{T}[X_tX_t'-E(X_tX_t')]\beta_t\right\| \end{aligned}$$

$$= \quad R_1 + R_2 + R_3, \text{ say.}$$

By Proof of Theorem 3.1, and Assumption 3.4-3.6, $R_1 = 0$. We can show $R_2 = 0$ by analogous proof of Lemma A.2 and Assumption 3.6. $R_3 = 0$ is by the moment conditions in Assumption 3.5, point-wise convergence, and Assumption 3.6.

Thus, we have shown

$$\sup_{u \in \mathbb{R}} \left\| \hat{A}_1(u) - \tilde{A}(u) \right\| = o_P(1).$$

Furthermore, when X_t is weakly stationary,

$$\begin{aligned}
\tilde{A}(u) &= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E\left[M_t(u) X_t'\right] \beta_t \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left[E\left(X_t X_t'\right) e^{\mathbf{i}u2\pi t/T} \beta_t - \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau E(X_t X_t') \beta_t \right] \\
&= E(X_t X_t') \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \left[e^{\mathbf{i}u2\pi t/T} \beta_t - \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau \beta_t \right].
\end{aligned}$$

■

Proof of Theorem 3.3.

Theorem 3.1 established the weak convergence of $\sqrt{T}\hat{A}(u)$ to a complex-valued Gaussian process $\mathcal{G}(u)$. By continuous mapping theorem, we have

$$\hat{K} \xrightarrow{d} \sup_{u \in \mathbb{R}} \|\mathcal{G}(u)\|^2,$$

and

$$\hat{C} \xrightarrow{d} \int_{\mathbb{R}} \|\mathcal{G}(u)\|^2 W(u) du.$$

■

Proof of Corollary 3.1. We refer to Theorem 2 in Hansen (1996) by showing that Assumption 1 and Assumption 2 are satisfied.

First, $\{X_t, Y_t\}_{t=1}^T$ is an absolutely regular process under Assumption 3.1. The stationarity assumption in Hansen (1996) is equivalent to the uniform boundedness of $E(|X_{jt}|)$ for all j and all t. The corresponding regression score in this paper is

$$s_T(u) \equiv \hat{M}_t(u)\varepsilon_t,$$

we need to show

$$\inf_{u \in \mathbb{R}} \left\{ \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E[\hat{M}_t(u)\hat{M}_t(u)^*] \right\} > 0.$$

Given $\hat{M}_t(u) = X_t e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1} X_t$,

$$\frac{1}{T} \sum_{t=1}^{T} E[\hat{M}_t(u)\hat{M}_t(u)^*]$$

$$
\begin{aligned}
&= \frac{1}{T}\sum_{t=1}^{T} E\left\{\left[X_t e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}X_t\right]\left[X_t e^{\mathbf{i}u2\pi t/T} - \hat{Q}_{xx}(u)^{*}\hat{Q}_{xx}^{-1}X_t\right]'\right\} \\
&= \frac{1}{T}\sum_{t=1}^{T} E\left\{X_tX_t' - \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}X_tX_t'e^{-\mathbf{i}u2\pi t/T} - X_tX_t'e^{\mathbf{i}u2\pi t/T}\hat{Q}_{xx}^{-1}\hat{Q}_{xx}(u)^{*} + \hat{Q}_{xx}(u)\hat{Q}_{xx}^{-1}X_tX_t'\hat{Q}_{xx}^{-1}\hat{Q}_{xx}(u)^{*}\right\} \\
&\xrightarrow{p} Q_{xx} - Q_{xx}(u)Q_{xx}^{-1}Q_{xx}(u)^{*} \sim \text{p.s.d.}
\end{aligned}
$$

Next, we need to show that $s_T(u)$ satisfies the Lipschitz condition. Then it is equivalent to show that

$$
\left\|X_t\varepsilon_t e^{\mathbf{i}u_1 2\pi t/T} - X_t\varepsilon_t e^{\mathbf{i}u_2 2\pi t/T}\right\| \le C|u_1 - u_2|,
$$

for some $C < \infty$. Observe that by Taylor expansion,

$$
\begin{aligned}
& X_t\varepsilon_t e^{\mathbf{i}u_1 2\pi t/T} - X_t\varepsilon_t e^{\mathbf{i}u_2 2\pi t/T} \\
= \;& X_t\varepsilon_t\left[e^{\mathbf{i}\bar{u}2\pi t/T}2\pi\frac{t}{T}(u_1 - u_2)\right].
\end{aligned}
$$

We can choose C to be $2\pi E(\|X_t\|^2)E(\varepsilon_t^2)$ and the Lipschitz condition is satisfied.

Then by Theorem 2 of Hansen (1996), the consistency of the resampling method is established.

■

Proof of Theorem 3.4. Under $\mathbb{H}_A$, by Theorem 3.2,

$$
\begin{aligned}
\hat{A}(u) &= \hat{A}_1(u) + \hat{A}_2(u) \\
&= \tilde{A}(u) + o_P(1).
\end{aligned}
$$

Given Lemma A.1, it is easy to show that $\tilde{A}(u)$ is stochastically equicontinuous. Therefore,

$$
\hat{K} = T\sup_{u\in\mathbb{R}}\|\hat{A}(u)\|^2 = O_P(T),
$$

and

$$
\hat{C} = T\int_{\mathbb{R}}\|\hat{A}(u)\|^2 W(u)du = O_P(T).
$$

■

Proof of Theorem 3.5. Under $\mathbb{H}_{A1} : \beta_t = \beta_0 + \Delta_T\phi_t$,

$$
\begin{aligned}
\hat{A}(u) &= \frac{1}{T}\sum_{t=1}^{T}\hat{M}_t(u)X_t'(\beta_0 + \Delta_T\phi_t) + \hat{A}_2(u) \\
&= \frac{\Delta_T}{T}\sum_{t=1}^{T}\hat{M}_t(u)X_t'\phi_t + \hat{A}_2(u),
\end{aligned}
$$

where the last equality is by

$$
\frac{1}{T}\sum_{t=1}^{T}\hat{M}_t(u)X_t' = 0.
$$

Given $\Delta_T = \frac{1}{\sqrt{T}}$,

$$\sqrt{T}\hat{A}(u) = \frac{1}{T}\sum_{t=1}^{T}\hat{M}_t(u)X_t'\phi_t + \sqrt{T}\hat{A}_2(u).$$

By analogous derivations in Proof of Theorem 3.2, we can show

$$\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{M}_t(u)X_t'\phi_t - \xi(u)\right\| \xrightarrow{p} 0$$

where

$$\xi(u) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E[M_t(u)X_t']\phi_t.$$

By Theorem 3.1

$$\sqrt{T}\hat{A}_2(u) \Rightarrow \mathcal{G}(u),$$

where $\mathcal{G}(u)$ is a complex-valued Gaussian process defined in Theorem 3.1.

Therefore, under $\mathbb{H}_{A1}$,

$$\sqrt{T}\hat{A}(u) \Rightarrow \xi(u) + \mathcal{G}(u).$$

By continuous mapping theorem,

$$\begin{aligned}
\hat{K} &\xrightarrow{d} \sup_{u\in\mathbb{R}}\|\xi(u)+\mathcal{G}(u)\|, \\
\hat{C} &\xrightarrow{d} \int_{\mathbb{R}}\|\xi(u)+\mathcal{G}(u)\|^2 W(u)du.
\end{aligned}$$

■

Proof of Theorem 4.1. Under $\mathbb{H}_0 : \beta_t = \beta_0$,

$$\begin{aligned}
\hat{A}^{IV}(u) &= \frac{1}{T}\sum_{t=1}^{T}\hat{X}_t\hat{\varepsilon}_t e^{\mathbf{i}u2\pi t/T} \\
&= \frac{1}{T}\sum_{t=1}^{T}\hat{X}_t(Y_t - X_t'\hat{\beta}_{2sls})e^{\mathbf{i}u2\pi t/T} \\
&= \frac{1}{T}\sum_{t=1}^{T}M_t^{IV}(u)\varepsilon_t + \frac{1}{T}\sum_{t=1}^{T}\left[\hat{M}_t^{IV}(u) - M^{IV}(u)\right]\varepsilon_t,
\end{aligned}$$

where

$$\begin{aligned}
\hat{M}_t^{IV}(u) &= \hat{X}_t(u) - \hat{Q}_{\hat{x}x}(u)\hat{Q}_{\hat{x}\hat{x}}^{-1}\hat{X}_t \\
&= \hat{\gamma}'Z_t(u) - \hat{\gamma}'\hat{Q}_{zx}(u)(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}'Z_t,
\end{aligned}$$

and

$$M_t^{IV}(u) = \tilde{X}_t(u) - Q_{\tilde{x}x}(u)Q_{\tilde{x}\tilde{x}}^{-1}\tilde{X}_t$$

$$= \quad \gamma' Z_t(u) - \gamma' Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma' Z_t.$$

Following Proof of Theorem 3.1, we first show that $\sqrt{T}\hat{A}^{IV}(u)$ converges to a normal distribution pointwise, and then establish the weak convergence result via stochastic equicontinuity. By the algebra that

$$\hat{a}\hat{b}\hat{c} - abc = \hat{a}(\hat{b} - b)\hat{c} + \hat{a}b(\hat{c} - c) + (\hat{a} - a)bc,$$

$$\begin{aligned}
& \sup_{u\in\mathbb{R}} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\hat{M}_t^{IV}(u) - M_t^{IV}(u) \right] \varepsilon_t \right\| \\
= \;& \sup_{u\in\mathbb{R}} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\hat{\gamma}' Z_t(u) - \gamma' Z_t(u) - \hat{\gamma}' \hat{Q}_{zx}(u)(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}' Z_t + \gamma' Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma' Z_t \right] \varepsilon_t \right\| \\
\leq \;& \sup_{u\in\mathbb{R}} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} (\hat{\gamma} - \gamma)' Z_t(u)\varepsilon_t \right\| \\
& + \sup_{u\in\mathbb{R}} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\hat{\gamma}' \hat{Q}_{zx}(u)(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}' - \gamma' Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma' \right] Z_t\varepsilon_t \right\| \\
\leq \;& \sup_{u\in\mathbb{R}} \left\| (\hat{\gamma} - \gamma)' \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t(u)\varepsilon_t \right\| \\
& + \sup_{u\in\mathbb{R}} \left\| \left[\hat{\gamma}' [\hat{Q}_{zx}(u) - Q_{zx}(u)](\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}' \right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t\varepsilon_t \right\| \\
& + \sup_{u\in\mathbb{R}} \left\| \left[\hat{\gamma}' Q_{zx}(u)[(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}' - (Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'] \right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t\varepsilon_t \right\| \\
& + \sup_{u\in\mathbb{R}} \left\| [(\hat{\gamma} - \gamma)' Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'] \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t\varepsilon_t \right\|, \\
\leq \;& \left\| (\hat{\gamma} - \gamma)' \right\| \sup_{u\in\mathbb{R}} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t(u)\varepsilon_t \right\| \\
& + \left\| \hat{\gamma}' \right\| \sup_{u\in\mathbb{R}} \left\| \hat{Q}_{zx}(u) - Q_{zx}(u) \right\| \left\| (\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}' \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t\varepsilon_t \right\| \\
& + \|\hat{\gamma}'\| \sup_{u\in\mathbb{R}} \|Q_{zx}(u)\| \left\| (\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}' - (Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma' \right\| \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t\varepsilon_t \right\| \\
& + \|\hat{\gamma}' - \gamma\| \sup_{u\in\mathbb{R}} \|Q_{zx}(u)\| \left\| (Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma' \right\| \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Z_t\varepsilon_t \right\| \\
= \;& o_P(1)\cdot O_P(1) + O_P(1)\cdot o_P(1)\cdot O_P(1) + O_P(1)\cdot O_P(1)\cdot o_P(1) + o_P(1)\cdot O_P(1)\cdot O_P(1) \\
= \;& o_P(1),
\end{aligned}$$

where the second to last equality is by the following results:
(i). $\hat{\gamma} = \hat{Q}_{zz}^{-1}Q_{zx} = O_P(1)$;
(ii). $\hat{\gamma} - \gamma = o_P(1)$;
(iii). $(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1})\hat{\gamma} - (Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma = o_P(1)$;
(iv). $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z_t\varepsilon_t = O_P(1)$;

(v). $\sup_{u\in\mathbb{R}}\left\|\hat{Q}_{zx}(u)-Q_{zx}(u)\right\|=o_P(1)$;
where (v). is implied by Lemma A.2.

Next, we follow the similar arguments in Proof of Theorem 3.1, and we can show

$$S_T^{IV}(u)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}M_t^{IV}(u)\varepsilon_t\xrightarrow{d}N[0,\mathcal{K}^{IV}(u,u)],$$

for each fixed $u\in\mathbb{R}$, where

$$\mathcal{K}^{IV}(u,u)\quad=\quad\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E[M_t^{IV}(u)M_t^{IV}(u)^{*'}\varepsilon_t^2].$$

Next, we show $S_T^{IV}(u)$ is stochastically equicontinuous over $u\in\mathbb{R}$, i.e., we need to show that, for any $\epsilon>0$ and $\kappa>0$, there exists $\delta>0$ such that

$$\lim_{T\to\infty}P\left[\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta}\left\|S_T^{IV}(u_1)-S_T^{IV}(u_2)\right\|>\kappa\right]<\epsilon.$$

Given

$$\begin{aligned}&\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta}\left\|S_T^{IV}(u_1)-S_T^{IV}(u_2)\right\|\\ \leq\;&\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\gamma'\left[Z_t(u_1)-Z_t(u_2)\right]\varepsilon_t\right\|\\ &+\sup_{u_1,u_2\in\mathbb{R}:|u_1-u_2|<\delta}\|\gamma'[Q_{zx}(u_1)-Q_{zx}(u_2)]\|\left\|(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Z_t\varepsilon_t\right\|,\end{aligned}$$

and we can show that $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Z_t(u)\varepsilon_t$ and $Q_{zx}(u)$ are stochastically equicontinuous by similar arguments as in Proof of Lemma A.1. Therefore $S_T^{IV}(u)$ is stochastically equicontinuous and

$$S_T^{IV}(u)\Rightarrow\mathcal{G}^{IV}(u)$$

where $\mathcal{G}^{IV}(u)$ is a zero-mean complex-valued Gaussian process with covariance kernel

$$\mathcal{K}^{IV}(u_1,u_2)\quad=\quad\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E[M_t^{IV}(u_1)M_t^{IV}(u_2)^{*'}\varepsilon_t^2].$$

Furthermore, when $\{X_t,Z_t\}$ is weakly stationary

$$\begin{aligned}E[M_t^{IV}(u_1)M_t^{IV}(u_2)^{*'}\varepsilon_t^2]&=E\left\{\left[\gamma'Z_t(u_1)-\gamma'Z_t\int_0^1e^{\mathbf{i}u_12\pi\tau}d\tau\right]\left[\gamma'Z_t(u_2)-\gamma'Z_t\int_0^1e^{\mathbf{i}u_22\pi\tau}d\tau\right]^{*'}\varepsilon_t^2\right\}\\&=\gamma'E\left(Z_tZ_t'\varepsilon_t^2\right)\gamma\left(e^{\mathbf{i}u_12\pi t/T}-\int_0^1e^{\mathbf{i}u_12\pi\tau}d\tau\right)\left(e^{\mathbf{i}u_22\pi t/T}-\int_0^1e^{\mathbf{i}u_22\pi\tau}d\tau\right)^*.\end{aligned}$$

Therefore,

$$
\begin{aligned}
\mathcal{K}^{IV}(u_1,u_2) &= \gamma' E\left(Z_t Z_t' \varepsilon_t^2\right)\gamma \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\left(e^{\mathbf{i}u_1 2\pi t/T}-\int_0^1 e^{\mathbf{i}u_1 2\pi\tau}d\tau\right)\left(e^{\mathbf{i}u_2 2\pi t/T}-\int_0^1 e^{\mathbf{i}u_2 2\pi\tau}d\tau\right)^* \\
&= \gamma' E\left(Z_t Z_t' \varepsilon_t^2\right)\gamma\left(\int_0^1 e^{\mathbf{i}2\pi(u_1-u_2)\tau}d\tau-\int_0^1 e^{\mathbf{i}2\pi u_1\tau}d\tau\int_0^1 e^{-\mathbf{i}2\pi u_2\tau}d\tau\right).
\end{aligned}
$$

■

Proof of Theorem 4.2. Under $\mathbb{H}_A$,

$$
\begin{aligned}
\hat{A}^{IV}(u) &= \frac{1}{T}\sum_{t=1}^{T}\hat{M}_t^{IV}(u)X_t'\beta+\frac{1}{T}\sum_{t=1}^{T}\hat{M}_t^{IV}(u)\varepsilon_t \\
&= \hat{A}_1^{IV}(u)+\hat{A}_2^{IV}(u).
\end{aligned}
$$

By proof of Theorem 4.1, we have $\hat{A}_2^{IV}(u)=O_P(T^{-1/2})=o_P(1)$. Let

$$
\tilde{A}^{IV}(u)=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\left[M_t^{IV}(u)X_t'\right]\beta_t,
$$

it follows

$$
\begin{aligned}
&\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\hat{A}_1^{IV}(u)-\tilde{A}^{IV}(u)\right\| \\
\le\ &\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{\hat{M}_t^{IV}(u)X_t'-M_t^{IV}(u)X_t'\right\}\beta_t\right\|+\lim_{T\to\infty}\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{M_t^{IV}(u)X_t'-E\left[M_t^{IV}(u)X_t'\right]\right\}\beta_t\right\| \\
=\ &R_1+R_2,\ \text{say.}
\end{aligned}
$$

Firstly, we show $R_1=0$.

$$
\begin{aligned}
&\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left[\hat{M}_t^{IV}(u)-M_t^{IV}(u)\right]X_t'\beta_t\right\| \\
=\ &\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left[\hat{\gamma}'Z_t(u)-\gamma'Z_t(u)-\hat{\gamma}'\hat{Q}_{zx}(u)(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}'Z_t+\gamma'Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'Z_t\right]X_t'\beta_t\right\| \\
\le\ &\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}(\hat{\gamma}-\gamma)'Z_t(u)X_t'\beta_t\right\| \\
&+\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left[\hat{\gamma}'\hat{Q}_{zx}(u)(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}'-\gamma'Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'\right]Z_tX_t'\beta_t\right\| \\
\le\ &\sup_{u\in\mathbb{R}}\left\|(\hat{\gamma}-\gamma)'\frac{1}{T}\sum_{t=1}^{T}Z_t(u)X_t'\beta_t\right\| \\
&+\sup_{u\in\mathbb{R}}\left\|\left[\hat{\gamma}'[\hat{Q}_{zx}(u)-Q_{zx}(u)](\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}'\right]\frac{1}{T}\sum_{t=1}^{T}Z_tX_t'\beta_t\right\| \\
&+\sup_{u\in\mathbb{R}}\left\|\left[\hat{\gamma}'Q_{zx}(u)[(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}'-(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma']\right]\frac{1}{T}\sum_{t=1}^{T}Z_tX_t'\beta_t\right\|
\end{aligned}
$$

$$
\begin{aligned}
& +\sup_{u\in\mathbb{R}}\left\|\left[(\hat{\gamma}-\gamma)'Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'\right]\frac{1}{T}\sum_{t=1}^{T}Z_tX_t'\beta_t\right\|, \\
\leq\ & \left\|(\hat{\gamma}-\gamma)'\right\|\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}Z_t(u)X_t'\beta_t\right\| \\
& +\left\|\hat{\gamma}'\right\|\sup_{u\in\mathbb{R}}\left\|\hat{Q}_{zx}(u)-Q_{zx}(u)\right\|\left\|(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}'\frac{1}{T}\sum_{t=1}^{T}Z_tX_t'\beta_t\right\| \\
& +\|\hat{\gamma}'\|\sup_{u\in\mathbb{R}}\|Q_{zx}(u)\|\left\|(\hat{Q}_{xz}\hat{Q}_{zz}^{-1}\hat{Q}_{zx})^{-1}\hat{\gamma}'-(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'\right\|\left\|\frac{1}{T}\sum_{t=1}^{T}Z_tX_t'\beta_t\right\| \\
& +\|\hat{\gamma}'-\gamma\|\sup_{u\in\mathbb{R}}\|Q_{zx}(u)\|\left\|(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'\right\|\left\|\frac{1}{T}\sum_{t=1}^{T}Z_tX_t'\beta_t\right\| \\
=\ & o_P(1)\cdot O_P(1)+O_P(1)\cdot o_P(1)\cdot O_P(1)+O_P(1)\cdot O_P(1)\cdot o_P(1)+o_P(1)\cdot O_P(1)\cdot O_P(1) \\
=\ & o_P(1).
\end{aligned}
$$

Comparing to the proof that shows

$$
\sup_{u\in\mathbb{R}}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left[\hat{M}_t^{IV}(u)-M_t^{IV}(u)\right]\varepsilon_t\right\|=o_P(1),
$$

we use an additional fact that $\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}Z_tX_t'\beta_t\right\|=O_P(1)$, which is implied by Assumption 4.5 and 3.6.

Next, we show $R_2=0$.

$$
\begin{aligned}
& \sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\left\{M_t^{IV}(u)X_t'-E\left[M_t^{IV}(u)X_t'\right]\right\}\beta_t\right\| \\
=\ & \sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\gamma'\left\{Z_t(u)X_t'-E[Z_t(u)X_t']\right\}\beta_t-\gamma'Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'\frac{1}{T}\sum_{t=1}^{T}[Z_tX_t'-E(Z_tX_t')]\beta_t\right\| \\
\leq\ & \sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\gamma'\left\{Z_t(u)X_t'-E[Z_t(u)X_t']\right\}\beta_t\right\|+\sup_{u\in\mathbb{R}}\left\|\gamma'Q_{zx}(u)(Q_{xz}Q_{zz}^{-1}Q_{zx})^{-1}\gamma'\frac{1}{T}\sum_{t=1}^{T}[Z_tX_t'-E(Z_tX_t')]\beta_t\right\| \\
=\ & o_P(1),
\end{aligned}
$$

by similar arguments as in Proof of Theorem 3.2.

Thus, we have shown

$$
\sup_{u\in\mathbb{R}}\left\|\hat{A}_1^{IV}(u)-\tilde{A}^{IV}(u)\right\|=o_P(1).
$$

Furthermore, when $\{X_t,Z\}$ is weakly stationary,

$$
\begin{aligned}
\tilde{A}(u) &= \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\left[M_t^{IV}(u)X_t'\right]\beta_t \\
&= \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E\left\{\left[\gamma'Z_t(u_1)-\gamma'Z_t\int_0^1 e^{\mathbf{i}u_1 2\pi\tau}d\tau\right]X_t'\right\}\beta_t
\end{aligned}
$$

$$
\begin{aligned}
&= \lim_{T\to\infty} \gamma' \frac{1}{T}\sum_{t=1}^{T}\left[\gamma' E\left(Z_t X_t'\right) e^{\mathbf{i}u2\pi t/T}\beta_t - \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau \gamma' E(Z_t X_t')\beta_t\right]\gamma \\
&= \gamma' E(Z_t X_t') \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\left[e^{\mathbf{i}u2\pi t/T}\beta_t - \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau \beta_t\right] \\
&= Q_{xz}Q_{zz}^{-1}Q_{zx} \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}\left[e^{\mathbf{i}u2\pi t/T}\beta_t - \int_0^1 e^{\mathbf{i}u2\pi\tau} d\tau \beta_t\right].
\end{aligned}
$$

■

Proof of Theorem 4.3.

Theorem 4.1 established the weak convergence of $\sqrt{T}\hat{A}^{IV}(u)$ to a complex-valued Gaussian process $\mathcal{G}^{IV}(u)$. By continuous mapping theorem, we have

$$
\hat{K}^{IV} \xrightarrow{d} \sup_{u\in\mathbb{R}} \|\mathcal{G}^{IV}(u)\|^2,
$$

and

$$
\hat{C}^{IV} \xrightarrow{d} \int_{\mathbb{R}} \|\mathcal{G}^{IV}(u)\|^2 W(u) du.
$$

■

Proof of Theorem 4.4. Under $\mathbb{H}_A$, by Theorem 4.2,

$$
\begin{aligned}
\hat{A}^{IV}(u) &= \hat{A}_1^{IV}(u) + \hat{A}_2^{IV}(u) \\
&= \tilde{A}^{IV}(u) + o_P(1).
\end{aligned}
$$

Given Lemma A.1, it is easy to show that $\tilde{A}^{IV}(u)$ is stochastically equicontinuous. Therefore,

$$
\hat{K}^{IV} = T \sup_{u\in\mathbb{R}} \|\hat{A}^{IV}(u)\|^2 = O_P(T),
$$

and

$$
\hat{C^{IV}} = T\int_{\mathbb{R}} \|\hat{A}^{IV}(u)\|^2 W(u) du = O_P(T).
$$

■

Proof of Theorem 4.5. Under $\mathbb{H}_{A1} : \beta_t = \beta_0 + \Delta_T \phi_t$,

$$
\begin{aligned}
\hat{A}^{IV}(u) &= \frac{1}{T}\sum_{t=1}^{T} \hat{M}_t^{IV}(u) X_t'(\beta_0 + \Delta_T\phi_t) + \hat{A}_2^{IV}(u) \\
&= \frac{\Delta_T}{T}\sum_{t=1}^{T} \hat{M}_t^{IV}(u) X_t'\phi_t + \hat{A}_2^{IV}(u),
\end{aligned}
$$

where the last equality is by

$$
\frac{1}{T}\sum_{t=1}^{T} \hat{M}_t^{IV}(u) X_t' = 0.
$$

Given $\Delta_T = \frac{1}{\sqrt{T}}$,

$$\sqrt{T}\hat{A}^{IV}(u) = \frac{1}{T}\sum_{t=1}^{T}\hat{M}_t^{IV}(u)X_t'\phi_t + \sqrt{T}\hat{A}_2^{IV}(u).$$

By analogous derivations in Proof of Theorem 3.2, we can show

$$\sup_{u\in\mathbb{R}}\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{M}_t^{IV}(u)X_t'\phi_t - \xi^{IV}(u)\right\| \xrightarrow{p} 0$$

where

$$\xi^{IV}(u) = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}E[M_t^{IV}(u)X_t']\phi_t.$$

By Theorem 4.1

$$\sqrt{T}\hat{A}_2^{IV}(u) \Rightarrow \mathcal{G}^{IV}(u),$$

where $\mathcal{G}^{IV}(u)$ is a complex-valued Gaussian process defined in Theorem 4.1.

Therefore, under $\mathbb{H}_{A1}$,

$$\sqrt{T}\hat{A}^{IV}(u) \Rightarrow \xi^{IV}(u) + \mathcal{G}^{IV}(u).$$

By continuous mapping theorem,

$$\hat{K}^{IV} \xrightarrow{d} \sup_{u\in\mathbb{R}}\|\xi^{IV}(u) + \mathcal{G}^{IV}(u)\|,$$

$$\hat{C}^{IV} \xrightarrow{d} \int_{\mathbb{R}}\|\xi^{IV}(u) + \mathcal{G}^{IV}(u)\|^2 W(u)du.$$

■

References

Andrews, D. W., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica: Journal of the Econometric Society, 821–856.

Andrews, D. W., Fair, R. C., 1988. Inference in nonlinear econometric models with structural change. The Review of Economic Studies 55 (4), 615–640.

Andrews, D. W., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica: Journal of the Econometric Society, 1383–1414.

Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. Econometrica, 47–78.

Bierens, H. J., 1982. Consistent model specification tests. Journal of Econometrics 20 (1), 105–134.

Boivin, J., 2006. Has us monetary policy changed? evidence from drifting coefficients and real-time data. Journal of Money, Credit and Banking 38 (5), 1149–1173.

Bradley, R. C., Tone, C., 2015. A central limit theorem for non-stationary strongly mixing random fields. Journal of Theoretical Probability, 1–20.

Bradley, R. C., et al., 2005. Basic properties of strong mixing conditions. a survey and some open questions. Probability surveys 2 (2), 107–144.

Brüggemann, R., Riedel, J., 2011. Nonlinear interest rate reaction functions for the uk. Economic Modelling 28 (3), 1174–1185.

Cai, Z., 2007. Trending time-varying coefficient time series models with serially correlated errors. Journal of Econometrics 136 (1), 163–188.

Cai, Z., Fan, J., Yao, Q., 2000. Functional-coefficient regression models for nonlinear time series. Journal of the American Statistical Association 95 (451), 941–956.

Cai, Z., Wang, Y., Wang, Y., 2015. Testing instability in a predictive regression model with non-stationary regressors. Econometric Theory 31 (05), 953–980.

Calvo, G. A., 1983. Staggered prices in a utility-maximizing framework. Journal of monetary Economics 12 (3), 383–398.

Campbell, J. Y., Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21 (4), 1509–1531.

Chen, B., 2015. Modeling and testing smooth structural changes with endogenous regressors. Journal of 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.

Chen, B., Hong, Y., 2016. Detecting for smooth structural changes in garch models. Econometric Theory 32 (03), 740–791.

Choi, I., Phillips, P. C., 1993. Testing for a unit root by frequency domain regression. Journal of Econometrics 59 (3), 263–286.

Chow, G. C., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica: Journal of the Econometric Society, 591–605.

Clarida, R., Galı, J., Gertler, M., 1998. Monetary policy rules in practice: Some international evidence. European Economic Review 42 (6), 1033–1067.

Clarida, R., Gali, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. The Quarterly Journal of Economics 115 (1), 147–180.

Corbae, D., Ouliaris, S., Phillips, P. C., 2002. Band spectral regression with trending data. Econometrica 70 (3), 1067–1109.

Dahlhaus, R., 1996. On the kullback-leibler information divergence of locally stationary processes. Stochastic Processes and their Applications 62 (1), 139–168.

Engle, R. F., 1974. Band spectrum regression. International Economic Review, 1–11.

Estrella, A., Fuhrer, J. C., 2002. Dynamic inconsistencies: Counterfactual implications of a class of rational-expectations models. The American Economic Review 92 (4), 1013–1028.

Estrella, A., Fuhrer, J. C., 2003. Monetary policy shifts and the stability of monetary policy models. Review of Economics and Statistics 85 (1), 94–104.

Fuhrer, J., Moore, G., 1995. Inflation persistence. The Quarterly Journal of Economics, 127–159.

Fuhrer, J. C., 1997. The (un) importance of forward-looking behavior in price specifications. Journal of Money, Credit, and Banking, 338–350.

Gali, J., Gertler, M., 1999. Inflation dynamics: A structural econometric analysis. Journal of monetary Economics 44 (2), 195–222.

Gali, J., Gertler, M., Lopez-Salido, J. D., 2005. Robustness of the estimates of the hybrid new keynesian phillips curve. Journal of Monetary Economics 52 (6), 1107–1118.

Granger, C., Hatanaka, M., 1964. Spectral analysis of economic time series.

Granger, C. W., Watson, M. W., 1984. Time series and spectral methods in econometrics. Handbook of econometrics 2, 979–1022.

Hall, A. R., Han, S., Boldea, O., 2012. Inference regarding multiple structural changes in linear models with endogenous regressors. Journal of econometrics 170 (2), 281–302.

Hannan, E. J., 1965. The estimation of relationships involving distributed lags. Econometrica: Journal of the Econometric Society, 206–224.

Hannan, E. J., 1967. The estimation of a lagged regression relation. Biometrika 54 (3-4), 409–418.

Hansen, B. E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica: Journal of the econometric society, 413–430.

Hansen, B. E., 2000. Testing for structural change in conditional models. Journal of Econometrics 97 (1), 93–115.

Hansen, B. E., 2001. The new econometrics of structural change: Dating breaks in us labor productivity. The Journal of Economic Perspectives 15 (4), 117–128.

Hausman, J. A., 1978. Specification tests in econometrics. Econometrica: Journal of the Econometric Society, 1251–1271.

Hong, Y., Wang, S., Wang, X., 2015. Testing strict stationarity with applications to macroeconomic and financial time series. Working paper.

Kim, C.-J., Nelson, C. R., 2006. Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex post data. Journal of Monetary Economics 53 (8), 1949–1966.

Kristensen, D., 2012. Non-parametric detection and estimation of structural change. The Econometrics Journal 15 (3), 420–461.

Lin, C.-F. J., Teräsvirta, T., 1994. Testing the constancy of regression parameters against continuous structural change. Journal of Econometrics 62 (2), 211–228.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708.

Orbe, S., Ferreira, E., Rodríguez-póo, J., 2000. A nonparametric method to estimate time varying coefficients under seasonal constraints. Journal of nonparametric statistics 12 (6), 779–806.

Orbe, S., Ferreira, E., Rodríguez-póo, J., 2005. Nonparametric estimation of time varying parameters under shape restrictions. Journal of Econometrics 126 (1), 53–77.

Perron, P., 2006. Dealing with structural breaks. Palgrave handbook of econometrics 1, 278–352.

Perron, P., Qu, Z., 2006. Estimating restricted structural change models. Journal of Econometrics 134 (2), 373–399.

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 (02), 491–507.

Perron, P., Yamamoto, Y., 2015. Using ols to estimate and test for structural changes in models with endogenous regressors. Journal of Applied Econometrics 30 (1), 119–144.

Quandt, R. E., 1960. Tests of the hypothesis that a linear regression system obeys two separate regimes. Journal of the American statistical Association 55 (290), 324–330.

Robinson, P. M., 1989. Nonparametric estimation of time-varying parameters. In: Statistical Analysis and Forecasting of Economic Structural Change. Springer, pp. 253–264.

Robinson, P. M., 1991. Time-varying nonlinear regression. Springer.

Rudebusch, G. D., 2002. Assessing nominal income rules for monetary policy with model and data uncertainty. The Economic Journal 112 (479), 402–432.

Sbordone, A. M., 2002. Prices and unit labor costs: a new test of price stickiness. Journal of Monetary economics 49 (2), 265–292.

Sbordone, A. M., 2005. Do expected future marginal costs drive inflation dynamics? Journal of Monetary Economics 52 (6), 1183–1197.

Stock, J. H., Watson, M. W., 1996. Evidence on structural instability in macroeconomic time series relations. Journal of Business & Economic Statistics 14 (1), 11–30.

Taylor, J. B., 1993. Discretion versus policy rules in practice. In: Carnegie-Rochester conference series on public policy. Vol. 39. Elsevier, pp. 195–214.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21 (4), 1455–1508.

Xu, K.-L., 2015. Testing for structural change under non-stationary variances. The Econometrics Journal 18 (2), 274–305.

Zhang, C., Osborn, D. R., Kim, D. H., 2008. The new keynesian phillips curve: from sticky inflation to sticky prices. Journal of Money, Credit and Banking 40 (4), 667–699.

Zhang, T., Wu, W. B., 2012. Inference of time-varying regression models. The Annals of Statistics 40 (3), 1376–1402.

Zheng, T., Xia, W., Huiming, G., 2012. Estimating forward-looking rules for china's monetary policy: A regime-switching perspective. China Economic Review 23 (1), 47–59.