12-2010

Testing Structural Change in Partially Linear Models

Liangjun SU

Singapore Management University, ljsu@smu.edu.sg

Halbert WHITE

University of California - San Diego

DOI: https://doi.org/10.1017/S0266466609990788

Testing Structural Change in Partially Linear Models *

Liangjun Su[a] and Halbert White[b]

[a]School of Economics, Singapore Management University, Singapore

[b]Department of Economics, UCSD, La Jolla, CA 92093-0508

December 3, 2009

Abstract

We consider two tests of structural change for partially linear time-series models. The first tests for structural change in the parametric component, based on the cumulative sums of gradients from a single semiparametric regression. The second tests for structural change in the parametric and nonparametric components simultaneously, based on the cumulative sums of weighted residuals from the same semiparametric regression. We derive the limiting distributions of both tests under the null hypothesis of no structural change and for sequences of local alternatives. We show that the tests are generally not asymptotically pivotal under the null but may be free of nuisance parameters asymptotically under further asymptotic stationarity conditions. Our tests thus complement the conventional instability tests for parametric models. To improve the finite sample performance of our tests, we also propose a wild bootstrap version of our tests and justify its validity. Finally, we conduct a small set of Monte Carlo simulations to investigate the finite sample properties of the tests.

JEL classifications: C12, C14, C22, C5

*The authors gratefully thank Oliver Linton, and two anonymous referees for their many constructive comments on the previous versions of the paper. They also thank Peter Robinson, Rong Chen, Jiti Gao, Yong Zhou, and the participants of FERM 2007 in Beijing, and the 2008 International Symposium on Nonlinear Time Series in Xiamen, for their valuable comments. The first author gratefully acknowledges financial support from the NSFC under grant numbers 70501001 and 70601001. Address correspondence to: Halbert White, Department of Economics, UCSD, La Jolla, CA 92093-0508, USA. Phone: +1 858 534-3502; e-mail: hwhite@weber.ucsd.edu.

Key Words: CUSUM test, Structural change, Partially linear models, Semiparametric estimation

1 Introduction

Time-series data in economics and finance often have two prominent characteristics, namely, instability and nonlinearity. The first feature has to do with whether the underlying data generating process (DGP) is stable over time, whereas the second has to do with whether the widely used linear model is adequate for modeling the DGP. In principle, many economic and financial factors, such as changes in tastes, technical progress, and economic and financial policies, may lead to an unstable DGP. As the Lucas critique further suggests, changes in economic agents' expectations can induce changes in the reduced-form relationship among economic variables. Even in the absence of instability, applying a linear parametric model to data generated by a nonlinear process can lead to apparent model instability.

Since the seminal work of Page (1955) and Chow (1960), there has developed a large literature on testing for structural change. One procedure that has played a particularly important role in the study of structural change is the CUSUM test proposed by Brown, Durbin, Evans (1975) and extended in a variety of ways by Krämer, Ploberger, and Alt (1988), Ploberger and Krämer (1992), Kuan and Hornik (1995), and Lee and Park (2001), to name just a few. Compared to some other tests in the literature (e.g., Andrews, 1993; Andrews and Ploberger, 1994), CUSUM-type tests are computationally simple and thus easier to implement in practice. On the other hand, all of these conventional procedures assume a parametric regression model, usually linear. If the parametric functional form is misspecified, then the test may not perform as intended.

Linear parametric models provide a parsimonious way to express relationships among variables, but they also impose strong restrictions. Nonparametric models allow for much greater flexibility and thus may have certain advantages in applications. For this reason, Delgado and Hidalgo (2000) advocate conducting nonparametric inference when testing for structural breaks; and Su and Xiao (2008) propose a nonparametric test of structural change in dynamic nonparametric regression models. Nevertheless, as Robinson (1988) has remarked, a correctly specified parametric model affords precise inferences, a badly misspecified one, possibly seriously misleading inference; whereas nonparametric modeling is associated with both greater robustness and lesser precision. An intermediate strategy is to adopt a semiparametric approach. Partially linear models are widely used in this context, motivating the approach we take here.

Since Engle, Granger, Rice, and Weiss (1986), partially linear models have attracted much attention among econometricians. See Robinson (1988), Linton (1995), Fan and Li (1999), Li and Wooldridge (2002), Juhl and Xiao (2005a, 2005b), to mention only a few. To the best of our knowledge, early empirical applications of partially linear models have been focused on either cross-section or conventional panel data. A few exceptions include Engle et al. (1986) and Härdle, Liang, and Gao (2000) who survey empirical applications of partially linear models for some classical time-series data, such as the sunspot, lynx, and the Australian blowfly data.

Recent applications of partially linear models in *economic* time series include four important branches. One is that of partially linear error correction models (e.g., Li and Wooldridge, 2002). See Bachmeier and Li (2002), Lee (2003) and Gaul and Theissen (2006) for empirical data analysis. The second branch generalizes conventional GARCH models to partially linear models. For example, Wu and Xiao (2002) study the relationship between return shocks and conditional volatility, where the impact of return shocks on conditional volatility is specified as a general function and estimated nonparametrically, whereas lagged conditional volatility enters the model linearly. The third branch re-examines certain economic and financial hypotheses by incorporating a nonparametric component. For example, Aneiros-Pérez, Gonzalez-Mánteiga, and Reboredo-Nogueira (2006) propose a new test for the forward premium unbiasedness hypothesis based on a partial linear regression model and find that the forward premium is an unbiased predictor of the spot return when they add a nonparametric component with time as a covariate in the traditional linear regression model. The fourth branch extends the theory and applications of partially linear models from stationary time series data to nonstationary time series data. Using U.S. monthly macroeconomic time series, Juhl and Xiao (2005a) illustrate how using a partially linear model with covariates can lead to a rejection of the unit root null hypothesis when standard unit root tests fail to reject, and Juhl and Xiao (2005b) find that nonparametrically including a stationary covariate in testing a cointegrating relationship may result in conclusions different from those of the standard cointegration test.

In this paper, we thus study tests for structural change using partially linear time-series

DGPs:

$$y_{nt} = x'_{nt}\gamma_{nt} + m_{nt}(z_{nt}) + \varepsilon_{nt},\ t = 1, 2, ..., n, \tag{1.1}$$

where y_{nt} is the dependent variable, x_{nt} is an $\mathbb{R}^p$-valued regressor, z_{nt} is an $\mathbb{R}^q$-valued regressor, γ_{nt} is a $p\times 1$ vector of unknown coefficients, $m_{nt}(\cdot)$ is an unknown but smooth function, and ε_{nt} is a random disturbance term satisfying $E(\varepsilon_{nt}|x_{nt}, z_{nt}) = 0$ a.s. Note that we have written (1.1) using triangular array notation and that both γ_{nt} and $m_{nt}(\cdot)$ may be time-varying. As Hansen (2000a) remarks, this notation facilitates large sample distribution assumptions allowing for a certain degree of non-stationarity in the process $\{(x_{nt},\ z_{nt},\ \varepsilon_{nt}),\ 1 \le t \le n\}$. In this paper, we assume that this triangular array process is a strong (α-) mixing process. We are interested in testing whether (i) the parametric regression coefficient, (ii) the nonparametric component, or (iii) both change over time.

We distinguish two important cases. In the first case, we test the null hypothesis that there is no structural change in the parametric regression coefficient ($H_{0a} : \gamma_{nt} = \gamma_0$ for all $t = 1, ..., n$), allowing the nonparametric component to be unstable over time. In the second case, we test the null hypothesis that there is no structural break in either the parametric regression coefficient or the nonparametric component ($H_{0b} : \gamma_{nt} = \gamma_0$ and $P[m_{nt}(z_{nt}) = m_0(z_{nt})] = 1$ for all $t = 1, ..., n$). Thus, the first test focuses on the stability of the parametric component of the regression function, whereas the second test focuses on the stability of the entire regression relation.

The motivation for the first test is three-fold. First, there are cases where one is only interested in testing the stability of the parametric component of the regression function. For example, such situations arise when one firmly believes that some policy change can only result in the potential change of the relationship between the regressors of the parametric component and the dependent variable, but not between those of the nonparametric component and the dependent variable. Second, we allow the presence of structural breaks in the nonparametric component when we test H_{0a}. As we show, the stability test for the parametric component is robust to instability in the nonparametric component. Third, when one rejects the second null, H_{0b}, it is of interest to know whether the apparent structural change is due to a break in the parametric component or in the nonparametric component. In this case, a test of H_{0a} may provide useful information.

To test H_{0a}, it is desirable to allow potential structural breaks in the regressor process $\{x_{nt}, z_{nt}\}$. When we allow the process $\{x_{nt}, z_{nt}\}$ to be nonstationary under the null, several possibilities arise: (a) the marginal probability density function (PDF) $f_{nt}(\cdot)$ of z_{nt} can be time varying; (b) the conditional expectation function $g_{nt}(\cdot) \equiv E(x_{nt}|z_{nt} = \cdot)$ can be time varying; (c) the nonparametric component $m_{nt}(\cdot)$ in the conditional mean process of y_{nt} can be time varying. We can also consider the marginal distribution of x_{nt}. Nevertheless, because our tests rely directly upon kernel estimation of $f_{nt}(\cdot)$, $g_{nt}(\cdot)$, and $m_{nt}(\cdot)$, the possible breaks in these nonparametric objects will play essential roles. Of course, all the above three types of breaks may be due to the breaks in the joint distribution of $(x_{nt}, z_{nt}, \varepsilon_{nt})$.

To proceed, it is worthwhile to distinguish two categories of breaks, namely, small breaks and fixed breaks. The former means break sizes that shrink to zero as the sample size n tends to ∞, a case that is widely used in the study of local power properties for various tests. The latter means break sizes that do not vanish as $n \to \infty$. For example, if

$$f_{nt}(z) = \begin{cases} f_1(z) & \text{if } t \leq \lceil n\pi_0 \rfloor \\ f_2(z) & \text{if } t \geq \lceil n\pi_0 \rfloor + 1 \end{cases} \quad \text{for some } \pi_0 \in (0,1)$$

where the functions $f_1(\cdot)$ and $f_2(\cdot)$ satisfy $P(f_1(z_{nt}) = f_2(z_{nt})) < 1$ and neither f_1 nor f_2 depends on n, then we say that there is a single fixed break in the nonparametric object $f(\cdot)$ (or f for short). Following the literature, we call π_0 the "break point" of the nonparametric component f. Analogously, one can define fixed breaks in the nonparametric objects $m(\cdot)$ and $g(\cdot)$ (or m and g for short).

In this paper, we will allow for fixed breaks in f but not in m or g. In sharp contrast to pure parametric models (e.g., Andrews, 1993; Bai, 1996; Hansen, 2000a), allowing for fixed breaks in the nonparametric objects m and g when testing structural changes in the finite-dimensional parameters (H_{0a}), is complicated by the need for consistent first-stage nonparametric estimation of these objects. Further, it is much easier to handle nonstationary data in the parametric framework because one has available a variety of applicable weak convergence results. For example, Andrews (1993) assumes that the triangular array of random variables is L^0-NED on a strong mixing process; Hansen (2000a) considers both asymptotically stationary and asymptotically nonstationary processes and allows for structural change in the distribution of the regressors.

Because our test is a nonparametric test for the semiparametric model, we require a preliminary consistent nonparametric estimator in order to consistently estimate the finite-dimensional parameter (γ_0 here). The latter consistency under the null is essential for the derivation of the asymptotic null distribution of our test statistic. Preliminary consistent estimation can be ensured if we consider only small breaks in m and g, corresponding to Hansen's (2000a) asymptotically stationary case. As we show, however, in the case of fixed breaks in both m and g, one generally cannot consistently estimate the finite dimensional parameters using a two-stage kernel method. As a result, deriving the asymptotic null distribution becomes intractable.

Similarly, in the case of the pure nonparametric model $y_{nt} = m_{nt}(z_{nt}) + \varepsilon_{nt}$, $t = 1, 2, ..., n$, one can always estimate the conditional mean object $m_0(z)$ consistently under the null (or the sequence of local alternatives). The test is based on one-step estimation of $m_0(z)$ under the null restriction. There is no preliminary estimate involved that can cause difficulties.

We propose a CUSUM-type test for each hypothesis. The test of H_{0a} is based upon the cumulative sums of gradients from a single semiparametric regression, whereas the test of H_{0b} is based upon the cumulative sums of weighted residuals from the same semiparametric regression. We derive the asymptotic properties of the two tests under their corresponding null and for sequences of local alternatives. We show that the limiting null distributions of the proposed CUSUM tests are generally not asymptotically pivotal if we allow for fixed breaks in the process $\{(x_{nt}, z_{nt}, \varepsilon_{nt})\}$. Nevertheless, under some asymptotic stationarity conditions, these limiting distributions become asymptotically distribution-free under the null hypotheses; each is associated with a vector of independent standard Brownian bridges. We also show that both tests have nontrivial power against $n^{-1/2}$ local alternatives, and we propose a wild bootstrap version of our tests. We demonstrate through simulations that our tests work reasonably well in finite samples.

The paper is organized as follows. In Section 2, we introduce our hypotheses, H_{0a} and H_{0b}, and the corresponding test statistics. We study the asymptotic properties of the CUSUM test of H_{0a} in Section 3 and those of the CUSUM test of H_{0b} in Section 4. In Section 5, we propose a wild bootstrap version of our tests and justify its validity. We provide a small set of Monte Carlo experiments to evaluate the finite sample performance of our tests in Section 6. Section 7 contains concluding remarks. All proofs are relegated to the appendix.

NOTATION: Throughout the paper, B_p denotes a p-dimensional vector of independent standard Brownian bridges on $[0,1]$, $\lceil \cdot \rfloor$ signifies the integer part, $\|\cdot\|$ denotes the Euclidean norm of a matrix (e.g., $\|A\| = [\text{tr}(AA')]^{1/2}$), and $1(\cdot)$ denotes the indicator function of a set. Let $\pi_1 \wedge \pi_2 \equiv \min(\pi_1, \pi_2)$, where $x \equiv y$ indicates that x is defined by y. The operators $\xrightarrow{p}$ and $\xrightarrow{d}$ denote convergence in probability and distribution, respectively. We use $\Rightarrow$ to denote weak convergence in the space $D[0,1]^p$ or $D[0,1]$ of $\overline{p}$-vectors of right-continuous functions with left-hand limits, endowed with the uniform topology (see Pollard (1984)), where $\overline{p} = p$ or 1. We let $\overset{p}{\Rightarrow}$ denote weak convergence in probability as defined by Giné and Zinn (1990); see also Hansen (2000a) and Cavaliere and Taylor (2006).

2 Hypotheses and Test Statistics

2.1 The Hypotheses

Consider the following partially linear data generating process (DGP):

$$y_{nt} = x_{nt}'\gamma_{nt} + m_{nt}(z_{nt}) + \varepsilon_{nt}, \;\; t = 1, 2, ..., n, \tag{2.1}$$

where y_{nt}, x_{nt}, z_{nt}, and ε_{nt} are defined after eq. (1.1). If $m_{nt}(\cdot)$ is absent from the DGP in (2.1), we obtain the conventional time-varying linear regression DGP. If $x_{nt}'\gamma_{nt}$ is absent, however, the DGP in (2.1) becomes time-varying nonparametric (see, e.g., Su and Xiao, 2008).

We consider two scenarios. In the first, allowing (but not requiring) the nonparametric component function $m_{nt}(\cdot)$ to change over time (so $m_{nt} = m_0$ for some m_0 when there is no change), we test whether the coefficient γ_{nt} is stable over time. In this case, the null hypothesis is that for some unknown γ_0, we have

$$H_{0a} : \gamma_{nt} = \gamma_0 \text{ for all } t = 1, ..., n, \tag{2.2}$$

and the alternative hypothesis is the negation of H_{0a}.

In the second case, we consider testing the joint stability of $m_{nt}(\cdot)$ and γ_{nt}. Our null hypothesis here is that for some unknown γ_0 and smooth m_0, we have

$$H_{0b} : \gamma_{nt} = \gamma_0 \text{ and } P[m_{nt}(z_{nt}) = m_0(z_{nt})] = 1 \text{ for all } t = 1, ..., n, \tag{2.3}$$

and the alternative hypothesis is the negation of H_{0b}. When H_{0b} holds, we say that there is no structural change or break in the conditional mean process.

We will not impose restrictions on the conditional variance process $\{E(\varepsilon_{nt}^2|x_{nt}, z_{nt})\}$, or on other aspects of the joint distribution of x_{nt}, z_{nt}, and ε_{nt}. Indeed, following Su and Xiao (2008), we permit time-varying behavior in the conditional variance process and a nonstationary distribution for $\{x_{nt}, z_{nt}, \varepsilon_{nt}\}$ under both the null and alternative hypotheses. Nevertheless, to facilitate the presentation we will assume that some aspects of the process $\{x_{nt}, z_{nt}, \varepsilon_{nt}\}$ are asymptotically stationary in a sense to be defined precisely below.

2.2 Estimation and Test Statistics

We base our tests on estimates of the restricted model

$$y_{nt} = x_{nt}'\gamma + m(z_{nt}) + u_{nt},\ t = 1, 2, ..., n, \tag{2.4}$$

where u_{nt} represents the model residual.

There are several ways to estimate the model of eq. (2.4); one of the more popular methods is the local constant estimator of Robinson (1988). Nevertheless, to handle the random denominator problem, Robinson's estimator requires not only selection of a kernel bandwidth parameter, but also a trimming parameter. To avoid the latter feature, we use density weighted kernel estimation, following Fan and Li (1999).

For this, let $f_{nt}(\cdot)$ be the density function of z_{nt}. We first use kernel methods to estimate $f_{nt} \equiv f_{nt}(z_{nt})$, $E(y_{nt}|z_{nt})$, and $E(x_{nt}|z_{nt})$ as:

$$\widehat{f}_{nt} = \widehat{f}_{nt}(z_{nt}) \equiv n^{-1}\sum_{s\neq t}^{n} K_{hts},\ \widehat{y}_{nt} \equiv n^{-1}\sum_{s\neq t}^{n} y_{ns}K_{hts}/\widehat{f}_{nt},\ \text{and}\ \widehat{x}_{nt} \equiv n^{-1}\sum_{s\neq t}^{n} x_{ns}K_{hts}/\widehat{f}_{nt}, \tag{2.5}$$

where $K_{hts} \equiv h^{-q}K((z_{nt} - z_{ns})/h)$, $K(\cdot)$ is a given kernel function, and $h = h(n)$ is the bandwidth parameter. (We divide by n instead of $n-1$ in eq. (2.5) for notational simplicity.) Then Fan and Li's (1999) density-weighted estimator of γ is given by

$$\widehat{\gamma} \equiv S^{-1}_{(X-\widehat{X})\widehat{f}} S_{(X-\widehat{X})\widehat{f},(Y-\widehat{Y})\widehat{f}}, \tag{2.6}$$

where $(X - \widehat{X})\widehat{f}$ is an $n \times p$ matrix whose tth row is given by $(x_{nt} - \widehat{x}_{nt})'\widehat{f}_{nt}$, $(Y - \widehat{Y})\widehat{f}$ is analogously defined, and, using the notation of Robinson (1988) and Fan and Li (1999), for any two matrices with n rows, A and B, we define $S_{A,B} \equiv n^{-1}\sum_{t=1}^{n} a_t'b_t$ and $S_A \equiv S_{A,A}$, where a_t and b_t are the tth rows of A and B, respectively.

Let $\widehat{\varepsilon}_{nt} \equiv n^{-1}\sum_{s\neq t}^{n} \varepsilon_{ns}K_{hts}/\widehat{f}_{nt}$ and $\widehat{m}_{nt} \equiv n^{-1}\sum_{s\neq t}^{n} m_{ns}(z_{ns})K_{hts}/\widehat{f}_{nt}$. Define $(M - \widehat{M})\widehat{f}$, $\varepsilon\widehat{f}$, and $\widehat{\varepsilon}\widehat{f}$ similarly to $(X-\widehat{X})\widehat{f}$. Then under either null hypothesis we have $\sqrt{n}(\widehat{\gamma} - \gamma_0) = S^{-1}_{(X-\widehat{X})\widehat{f}}\sqrt{n}S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}+\varepsilon\widehat{f}-\widehat{\varepsilon}\widehat{f}}$. Under some regularity conditions, we can show that under either null hypothesis

$$\sqrt{n}(\widehat{\gamma} - \gamma_0) \xrightarrow{d} N\left(0, \Phi^{-1}\Psi\Phi^{-1}\right), \tag{2.7}$$

where

$$\Phi \equiv \lim_{n\to\infty} n^{-1} \sum_{t=1}^{n} E\{[x_{nt} - E(x_{nt}|z_{nt})][x_{nt} - E(x_{nt}|z_{nt})]' f_{nt}^2(z_{nt})\} \text{ and} \tag{2.8}$$

$$\Psi \equiv \lim_{n\to\infty} n^{-1} \sum_{t=1}^{n} E\{[x_{nt} - E(x_{nt}|z_{nt})][x_{nt} - E(x_{nt}|z_{nt})]' \varepsilon_{nt}^2 f_{nt}^4(z_{nt})\}. \tag{2.9}$$

Once we obtain $\widehat{\gamma}$, we can estimate $m(z_t)$ in (2.4) by

$$\widetilde{m}(z_{nt}) \equiv n^{-1} \sum_{s\neq t}^{n} \left(y_{ns} - x_{ns}'\widehat{\gamma}\right) K_{hts}/\widehat{f}_{nt}. \tag{2.10}$$

Let

$$\widetilde{u}_{nt} \equiv y_{nt} - x_{nt}'\widehat{\gamma} - \widetilde{m}(z_{nt}). \tag{2.11}$$

Like Ploberger and Krämer (1992) and Bai (1996), our test statistics are based on these estimated residuals. Under mild conditions,

$$\widehat{\Psi} \equiv n^{-1} \sum_{t=1}^{n} \widetilde{u}_{nt}^2 (x_{nt} - \widehat{x}_{nt})(x_{nt} - \widehat{x}_{nt})' \widehat{f}_{nt}^4 \tag{2.12}$$

consistently estimates Ψ. To test H_{0a}, we thus consider tests based on the stochastic process

$$\Gamma_{na}(\pi) \equiv n^{-1/2} \widehat{\Psi}^{-1/2} \sum_{t=1}^{\lceil n\pi \rfloor} (x_{nt} - \widehat{x}_{nt}) \widetilde{u}_{nt} \widehat{f}_{nt}^2, \ 0 \le \pi \le 1. \tag{2.13}$$

Note that $(x_{nt} - \widehat{x}_{nt}) \widetilde{u}_{nt} \widehat{f}_{nt}^2$ appears as the summand in the first order conditions for the regression of $(y_{nt} - \widehat{y}_{nt}) \widehat{f}_{nt}$ on $(x_{nt} - \widehat{x}_{nt}) \widehat{f}_{nt}$. Therefore, $\Gamma_{na}(\pi)$ is a standardized cumulative sum of the gradients. We will show that under some weak conditions, the process $\Gamma_{na}(\cdot) \equiv \{\Gamma_{na}(\pi) : 0 \le \pi \le 1\}$ converges weakly to a mean-zero Gaussian process $\Gamma_a(\cdot)$.

Note that Γ_{na} will be sensitive to deviations from H_{0a} caused by changes in the parametric regression coefficients. On the other hand, tests based on this process will not have power to detect changes in the nonparametric component. Heuristically, any deviations of $m_{nt}(\cdot)$ from $m_0(\cdot)$ will appear in the residual sequence $\{\widetilde{u}_{nt}\}$, and these are asymptotically orthogonal to $(x_{nt} - \widehat{x}_{nt})$. For this reason, Γ_{na} cannot be used to test H_{0b}.

To test H_{0b}, we propose statistics based on the cumulative sums of weighted residuals:

$$\Gamma_{nb}(\pi) \equiv n^{-1/2} \widehat{\sigma}^{-1} \sum_{t=1}^{\lceil n\pi \rfloor} \widetilde{u}_{nt} \widehat{f}_{nt}, \ 0 \le \pi \le 1, \tag{2.14}$$

where $\widehat{\sigma} \equiv \left\{ n^{-1} \sum_{t=1}^{n} \widetilde{u}_{nt}^2 \widehat{f}_{nt}^2 \right\}^{1/2}$. We will show that $\Gamma_{nb}(\cdot) \equiv \{\Gamma_{nb}(\pi) : 0 \leq \pi \leq 1\}$ converges weakly to a mean-zero Gaussian process $\Gamma_b(\cdot)$.

Let $\mathcal{L}_a(\cdot)$ and $\mathcal{L}_b(\cdot)$ be continuous functionals that measure the fluctuations of $\Gamma_{na}(\cdot)$ and $\Gamma_{nb}(\cdot)$ respectively. By the continuous mapping theorem,

$$\mathcal{L}_a(\Gamma_{na}(\cdot)) \xrightarrow{d} \mathcal{L}_a(\Gamma_a(\cdot)) \text{ and } \mathcal{L}_b(\Gamma_{nb}(\cdot)) \xrightarrow{d} \mathcal{L}_b(\Gamma_b(\cdot)). \tag{2.15}$$

In principle, there is a rich variety of choices for $\mathcal{L}_a$ and $\mathcal{L}_b$. The classical Kolmogorov-Smirnoff measure yields the following CUSUM-type test statistics:

$$KS_{na} \equiv \sup_{0 \leq \pi \leq 1} |\Gamma_{na}(\pi)|_\infty = \max_{1 \leq j \leq n} \left| n^{-1/2} \widehat{\Psi}^{-1/2} \sum_{t=1}^{j} (x_{nt} - \widehat{x}_{nt}) \widetilde{u}_{nt} \widehat{f}_{nt}^2 \right|_\infty, \text{ and} \tag{2.16}$$

$$KS_{nb} \equiv \sup_{0 \leq \pi \leq 1} |\Gamma_{nb}(\pi)| = \max_{1 \leq j \leq n} \left| n^{-1/2} \widehat{\sigma}^{-1} \sum_{t=1}^{j} \widetilde{u}_{nt} \widehat{f}_{nt} \right|, \tag{2.17}$$

where for any p-vector $a_n = (a_{n1}, ..., a_{np})'$, $|a_n|_\infty \equiv \max_{1 \leq i \leq p} |a_{ni}|$. Alternatively, the Cramer-von Mises metric yields the following test statistics:

$$CM_{na} \equiv \int_0^1 \|\Gamma_{na}(\pi)\|^2 ds = \frac{1}{n} \sum_{j=1}^{n} \left\| n^{-1/2} \widehat{\Psi}^{-1/2} \sum_{t=1}^{j} (x_{nt} - \widehat{x}_{nt}) \widetilde{u}_{nt} \widehat{f}_{nt}^2 \right\|^2, \tag{2.18}$$

$$CM_{nb} \equiv \int_0^1 |\Gamma_{nb}(\pi)|^2 ds = \frac{1}{n} \sum_{j=1}^{n} \left(n^{-1/2} \widehat{\sigma}^{-1} \sum_{t=1}^{j} \widetilde{u}_{nt} \widehat{f}_{nt} \right)^2, \tag{2.19}$$

where $\|\cdot\|$ denotes the Euclidean norm. We will study the limiting distributions of KS_{na}, KS_{nb}, CM_{na}, and CM_{nb} below.

3 Asymptotic Properties of $\Gamma_{na}(\cdot)$

In this section, we study the asymptotic properties of $\Gamma_{na}(\cdot)$ under H_{0a} and a sequence of local alternatives. We study $\Gamma_{nb}(\cdot)$ in the next section.

3.1 Assumptions

Let $w_{nt} \equiv (x_{nt}', z_{nt}', \varepsilon_{nt})'$. We will use the mixing coefficients $\alpha_n(j)$, defined by

$$\begin{aligned} \alpha_n(j) = \;& \sup_{1 \leq l \leq n-j} \{P(A \cap B) - P(A)P(B) | A \in \sigma(w_{nt} : 1 \leq t \leq l), \\ & B \in \sigma(w_{nt} : l+j \leq t \leq n)\}, \; j \leq n-1, \\ \alpha_n(j) = \;& 0 \text{ for } j \geq n. \end{aligned}$$

Define the coefficient of strong mixing as $\alpha(j) = \sup_{n\in\mathbb{N}} \alpha_n(j)$ for $j \in \mathbb{N}$ and $\alpha(0) = 1$.

To state the assumptions, let $g_{nt}(z_{nt}) \equiv E(x_{nt}|z_{nt})$ and $v_{nt} \equiv x_{nt} - g_{nt}(z_{nt})$. Let $\sigma^2_{nt}(x,z) \equiv E\left(\varepsilon^2_{nt}|x_{nt} = x, z_{nt} = z\right)$, $\sigma^2_{nt}(z) \equiv E\left(\varepsilon^2_{nt}|z_{nt} = z\right)$, and $\sigma^2_{nt,i}(z) \equiv E\left(v^2_{nt,i}|z_{nt} = z\right)$, where $v_{nt,i}$ is the ith component of v_{nt}, $i = 1, ..., p$. We make the following assumptions on the disturbance term, regressors, kernel function, and bandwidth.

Assumption A1. (i) $\{w_{nt}\}$ is a strong mixing process with mixing coefficients $\alpha(j)$ satisfying $\sup_n \sum_{j=1}^n j^3\alpha(j)^{\eta/(4+\eta)} \leq C < \infty$ for some $\eta > 0$ with $\eta/(4+\eta) \leq 1/2$.

(ii) $E(\varepsilon_{nt}|\mathcal{F}_{n,t-1}) = 0$, where $\mathcal{F}_{n,t-1} \equiv \sigma(x_{nt}, z_{nt}, x_{n,t-1}, z_{n,t-1}, \varepsilon_{n,t-1}, ...)$.

(iii) For all $t \geq 1$, $f_{nt}(\cdot) \in \mathcal{G}^\infty_r$, $m_{nt}(\cdot) \in \mathcal{G}^{4+\eta}_r$, and $g_{nt}(\cdot) \in \mathcal{G}^{4+\eta}_r$ for some integer $r \geq 2$, where $\mathcal{G}^a_r$ is a class of functions defined in Definition C.3 in the Appendix. Also, f_{nt}, m_{nt}, and the elements of g_{nt} each satisfy a global Lipschitz condition: $|\phi(z^*) - \phi(z)| \leq D_\phi(z)\|z^* - z\|$ for all $z^*, z \in \mathbb{R}^q$, where $D_\phi(z_{nt})$ has finite $4+\eta$ moments and $\phi = f_{nt}, m_{nt}$, or an element of g_{nt}.

(iv) $\sup_{n\geq1}\max_{1\leq t\leq n}E(\|\xi_{nt}\|^{4+\eta}) \leq c_{4+\eta} < \infty$ for $\xi_{nt} = \varepsilon_{nt}$ and v_{nt}. For all $t \geq 1$, $(x,z) \to \sigma^2_{nt}(x,z)$, $z \to \sigma^2_{nt}(z)$, and $z \to \sigma^2_{nt,i}(z)$ $(i = 1, ..., p)$ all belong to $\mathcal{G}^2_1$.

(v) With $\xi_{nt} = \varepsilon_{nt}$ or x_{nt}, $\sup_z\sup_{n\geq1}\max_{1\leq t\leq n}E(\|\xi_{nt}\|^{4+\eta}|z_{nt} = z)f_{nt}(z) \leq \overline{b}_1 < \infty$ and for some $\vartheta \geq q$, $\sup_z\sup_{n\geq1}\max_{1\leq t\leq n}\|z\|^\vartheta E(|\xi_{nt}|\ |z_{nt} = z)f_{nt}(z) \leq \overline{b}_2 < \infty$. There is some $t^* < \infty$ such that for all $t \geq t^* > 1$, $\sup_{z,z'}\sup_{n\geq1}\max_{1\leq s,t\leq n}E(|\xi_{ns}\xi_{nt}|\ |z_{ns} = z, z_{nt} = z')$ $f_{n,st}(z,z') \leq \overline{b}_3 < \infty$, where $f_{n,st}$ denotes the joint density of (z_{ns}, z_{nt}).

(vi) For some $\theta \in [1/2, 1)$, we have $\log n/\left(n^\theta h^q\right) = o(1)$, and

$$\frac{q}{\vartheta} + 3 + 2\theta - \frac{1-\theta}{2}\left(\frac{(2\alpha+3)(\eta+2)}{\eta+3} - 2q\right) \leq 0, \tag{3.1}$$

where $\alpha = 4 + 16/\eta$.

(vii) There exists $m_n(\cdot)$ and $g_n(\cdot)$ such that $\max_{1\leq t\leq n}\|m_{nt}(z) - m_n(z)\| \leq \alpha_{mn}c_{mn}(z)$ and $\max_{1\leq t\leq n}\|g_{nt}(z) - g_n(z)\| \leq \alpha_{gn}c_{gn}(z)$ for some functions $c_{mn}(\cdot)$ and $c_{gn}(\cdot)$ and scalar sequences α_{mn} and α_{gn}. In addition, $\sup_{n\geq1}\max_{1\leq t\leq n}E|c_{\xi n}(z_{nt})|^{4+\eta} < \infty$ for $\xi = m$, and g.

(viii) Let $\overline{f}_n(z) \equiv n^{-1}\sum_{t=1}^n f_{nt}(z)$ and $\overline{f}_{nt} \equiv \overline{f}_n(z_{nt})$. $n^{-1}\sum_{t=1}^{\lceil n\pi\rfloor}\overline{f}^2_{nt}v_{nt}v'_{nt} \xrightarrow{p} \Phi(\pi)$ uniformly in π, $n^{-1}\sum_{t=1}^{\lceil n\pi\rfloor}\overline{f}^4_{nt}v_{nt}v'_{nt}\varepsilon^2_{nt} \xrightarrow{p} \Psi(\pi)$ uniformly in π, $n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}^2_{nt}\varepsilon^2_{nt} + o_p(1) \xrightarrow{p} \sigma^2(\pi)$ uniformly in π, and $n^{-1/2}\sum_{t=1}^{\lceil n\cdot\rfloor}\overline{f}^2_{nt}v_{nt}\varepsilon_{nt} \Rightarrow N(\cdot)$, where $\Phi(\pi)$ and $\Psi(\pi)$ are $q\times q$ nonrandom positive definite matrices and $\sigma^2(\pi) > 0$ for any $\pi \in (0,1]$, $\Phi(0) = 0$, $\Psi(0) = 0$, $\sigma^2(0) = 0$, and $N(\cdot)$ is a zero-mean Gaussian process with covariance kernel $E[N(\pi_1)N(\pi_2)'] = \Psi(\pi_1 \wedge \pi_2)$.

Assumption A2. The kernel function $K(\cdot)$ is product kernel of $k(\cdot)$, a symmetric

rth order kernel with compact support $\mathcal{A}$ such that $\int_{\mathbb{R}} a^i k(a)\, da = \delta_{i0}$ $(i = 0, 1, ..., r-1)$, $\sup_{a\in\mathcal{A}} |k(a)| \le \overline{c}_1 < \infty$, and $|k(a) - k(a')| \le \overline{c}_2 |a - a'|$ for any $a, a' \in \mathbb{R}$ and some $\overline{c}_2 < \infty$, where δ_{ij} is Kronecker's delta.

Assumption A3. As $n \to \infty$, $nh^{2q}/(\log n)^2 \to \infty$ and $nh^{4r} \to 0$.

Assumptions A1(i)-(iv) parallel Assumptions (A1)(i)-(iv) in Fan and Li (1999). A noteworthy difference is that Fan and Li (1999) assume a strictly stationary β-mixing (absolutely regular) process in order to use Lemma 1 of Yoshihara (1976). It turns out that we can relax the β-mixing condition to α-mixing by applying Lemma 2.1 of Sun and Chiang (1997) (see also Lemma C.1 in the appendix). Assumption A1(i) implies that $\alpha(j) = o(j^{-(4+16/\eta)})$. The smaller is η, the faster the rate at which $\alpha(j)$ decays to zero. Assumption A1(ii) imposes a martingale difference structure on $\{\varepsilon_{nt}\}$. The smoothness and moment conditions in Assumptions A1(iii)-(iv) are similar to those in Robinson (1988) and Fan and Li (1999). In particular, Assumptions A1(i) and A1(iv) reflect the trade-off between the degree of dependence and the moments of the process $\{x_{nt}, z_{nt}, \varepsilon_{nt}\}$. Assumption A1(v) is used in the proof of Lemma C.5 in the Appendix. It controls the tail behavior of the conditional expectations $E(|\xi_{nt}|^{4+\eta} | z_{nt} = z)$, $E(|\xi_{nt}| \ | z_{nt} = z)$, and $E(|\xi_{ns}\xi_{nt}| \ | z_{ns} = z, z_{nt} = z')$, relative to the marginal density $f_{nt}(z)$ or the joint density $f_{n,st}(z, z')$. Assumption A1(vi) reflects the trade-off between the mixing coefficient, moments of the process $\{\varepsilon_{nt}, x_{nt}, z_{nt}\}$, and the bandwidth h. For fixed $\theta \in [1/2, 1)$ and $\vartheta \ge q$, (3.1) can easily be satisfied by requiring sufficiently small η. Assumption A1(vii) specifies the nonstationary nature of the regressor process $\{x_{nt}, z_{nt}\}$ in terms of $m_{nt}(\cdot)$ and $g_{nt}(\cdot)$. It allows for both fixed and small breaks, as we have not required that the sequences α_{mn} and α_{gn} shrink to zero as $n \to \infty$. (These sequences are not to be confused with the mixing coefficients, subscripted differently.) In the case of fixed breaks for m and g, one can take $\alpha_{mn} = 1$ and $\alpha_{gn} = 1$, respectively. Assumption A1(viii) is a high level assumption. If we only consider small breaks in the process $\{x_{nt}, z_{nt}, \varepsilon_{nt}\}$, one can impose the following linearity assumption on $\Phi(\cdot)$, $\Psi(\cdot)$, and $\sigma^2(\cdot)$:

$$\Phi(\pi) = \pi\Phi, \ \Psi(\pi) = \pi\Psi, \text{ and } \sigma^2(\pi) = \pi\sigma_0^2, \tag{3.2}$$

where Φ and Ψ are defined in (2.8) and (2.9) respectively, and $\sigma_0^2 \equiv \sigma^2(1)$.

Assumption A2 requires the kernel function $K(\cdot)$ to be compactly supported, which can be relaxed at the cost of lengthier arguments (see Hansen, 2008). Assumption A3 is a little bit stronger than the bandwidth condition in Fan and Li (1999), who require $nh^{2q} \to \infty$

and $nh^{4r} \to 0$ as $n \to \infty$. With some lengthier arguments, we conjecture that we can relax $nh^{2q}/(\log n)^2 \to \infty$ to $nh^{2q} \to \infty$.

3.2 Asymptotic Behavior of $\widehat{\gamma}$

Since our test statistics rely heavily upon the asymptotic behavior of $\widehat{\gamma}$, it is worthwhile to study $\widehat{\gamma}$ before we proceed to study the asymptotic properties of our test statistics. Let $\overline{A}_{\xi n}(z) \equiv n^{-1}\sum_{t=1}^{n} f_{nt}(z)\,\xi_{nt}(z)$ for $\xi = m, g,$ or 1. In particular, when $\xi = 1$, we have $\overline{A}_{1n}(z) = \overline{f}_n(z)$. To allow for possible fixed breaks in either m or g, or more generally, the joint distribution of (x_{nt}, z_{nt}), we make the following high level assumptions.

Assumption A1 (vii*)**.** As $n \to \infty$, $n^{-1}\sum_{t=1}^{n}\left[\overline{A}_{gn}(z_{nt}) - \overline{A}_{1n}(z_{nt})\,g_{nt}(z_{nt})\right]\left[\overline{A}_{gn}(z_{nt}) - \overline{A}_{1n}(z_{nt})\,g_{nt}(z_{nt})\right]' \xrightarrow{p} \Phi_{gg}$, and $n^{-1}\sum_{t=1}^{n}[\overline{A}_{gn}(z_{nt}) - \overline{A}_{1n}(z_{nt})\,g_{nt}(z_{nt})][\overline{A}_{mn}(z_{nt}) - \overline{A}_{1n}(z_{nt})\,m_{nt}(z_{nt})] \xrightarrow{p} \Phi_{gm}$, where Φ_{gg} and Φ_{gm} are $q \times q$ and $q \times 1$ nonrandom matrices, respectively.

Clearly, if we have only small breaks in m and g, then $\Phi_{gg} = 0$ and $\Phi_{gm} = 0$. These are generally non-zero if we allow for fixed breaks in m and g. The following theorem characterizes the asymptotic behavior of $\widehat{\gamma}$ under H_{0a}.

Theorem 3.1 *Suppose Assumptions A1-A3 and H_{0a} hold. (a) If Assumption A1(vii*) holds, then*

$$\widehat{\gamma} - \gamma_0 = (\Phi(1) + \Phi_{gg})^{-1}\,\Phi_{gm} + o_p(1).$$

(b) If we have fixed breaks in m but not g and $h^r = o(n^{-1/2})$, then

$$\begin{aligned}\sqrt{n}\,(\widehat{\gamma} - \gamma_0) &= \Phi(1)^{-1}\left\{n^{-5/2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n} K_{hij}K_{hik}\ [m_{nk}(z_{ni}) - m_{ni}(z_{ni})]\,(v_{nj} - v_{ni})\right.\\ &\quad \left. + n^{-1/2}\sum_{i=1}^{n}\overline{f}_n^2(z_{ni})\,v_{ni}\varepsilon_{ni}\right\} + o_p(1).\end{aligned}$$

(c) If we have fixed breaks in g but not m and $h^r = o(n^{-1/2})$ and Assumption A1(vii) holds, then*

$$\begin{aligned}\sqrt{n}\,(\widehat{\gamma} - \gamma_0) &= (\Phi(1) + \Phi_{gg})^{-1}\left\{n^{-5/2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n} K_{hij}K_{hik}\,[g_{nj}(z_{ni}) - g_{ni}(z_{ni})]\,(\varepsilon_{nk} - \varepsilon_{ni})\right.\\ &\quad \left. + n^{-1/2}\sum_{i=1}^{n}\overline{f}_n^2(z_{ni})\,v_{ni}\varepsilon_{ni}\right\} + o_p(1).\end{aligned}$$

(d) If α_{mn} and α_{gn} in Assumption A1(vii) and h in Assumption A3 also satisfy $\alpha_{mn}h^r = o\left(n^{-1/2}\right)$, $\alpha_{gn}h^r = o\left(n^{-1/2}\right)$, and $\alpha_{mn}\alpha_{gn} = o\left(n^{-1/2}\right)$, then

$$\sqrt{n}\left(\widehat{\gamma} - \gamma_0\right) = \Phi\left(1\right)^{-1} n^{-1/2} \sum_{i=1}^{n} \overline{f}_n^2\left(z_{ni}\right) v_{ni}\varepsilon_{ni} + o_p\left(1\right).$$

Remark 1. As mentioned above, if fixed breaks are present in both $m\left(\cdot\right)$ and $g\left(\cdot\right)$, then generally $\Phi_{gm} \neq 0$. This is true even if we have only a one-time simultaneous fixed break in $m\left(\cdot\right)$ and $g\left(\cdot\right)$. As a result, we are unable to estimate γ_0 consistently by $\widehat{\gamma}$ and thus cannot derive an asymptotic null distribution for our test statistics. If we have a fixed break in either m or g but not both, then Theorem 3.1 indicates that $\widehat{\gamma}$ is $\sqrt{n}$-consistent for γ_0 under H_{0a} and some further conditions. Nevertheless, the unknown fixed breaks in m or g contribute to the variance of $\sqrt{n}\left(\widehat{\gamma} - \gamma_0\right)$ and the expansion of our test statistics in a complicated way, which makes characterizing the asymptotic null distribution of our test statistics quite cumbersome. In contrast, if m and g have only small breaks, $\sqrt{n}\left(\widehat{\gamma} - \gamma_0\right)$ will have the same asymptotic distribution under the null as in the purely stationary case. For these reasons, we focus only on this last case in what follows.

3.3 Asymptotic Null Distribution of $\Gamma_{na}\left(\cdot\right)$

The following theorem gives the asymptotic distribution of $\Gamma_{na}\left(\cdot\right)$ under H_{0a}.

Theorem 3.2 *Suppose Assumptions A1-A3 hold. Suppose the conditions in part (d) of Theorem 3.1 hold. Then under H_{0a}, $\Gamma_{na}\left(\cdot\right) \Rightarrow \Gamma_a\left(\cdot\right)$, where $\Gamma_a\left(\pi\right) = \Psi\left(1\right)^{-1/2}\left[N\left(\pi\right) - \Phi\left(\pi\right)\Phi\left(1\right)^{-1} N\left(1\right)\right]$.*

Remark 2. By the continuous mapping theorem, Theorem 3.2 implies that $KS_{na} \xrightarrow{d} \sup_{0\leq\pi\leq1} \left|\Gamma_a\left(\pi\right)\right|_\infty$, and $CM_{na} \xrightarrow{d} \int_0^1 \left\|\Gamma_a\left(\pi\right)\right\|^2 d\pi$. Obviously, the tests KS_{na} and CM_{na} are generally not asymptotically pivotal. The asymptotic null distributions of these test statistics appear to depend on the functions $\Phi\left(\pi\right)$ and $\Psi\left(\pi\right)$ in a complicated way. As there is no way to tabulate the critical values for the tests, we later provide a method to obtain bootstrap p-values.

Remark 3. Nevertheless, when both $\Phi\left(\pi\right)$ and $\Psi\left(\pi\right)$ are linear in π (see (3.2)), the above Γ_{na}-based tests are asymptotically pivotal. In this case, we have $\Gamma_a\left(\pi\right) = \Psi^{-1/2}\left[N\left(\pi\right) - \pi N\left(1\right)\right]$ $= B_p\left(\pi\right)$, where $B_p\left(\cdot\right)$ denotes a vector of p independent standard Brownian bridges defined on $[0,1]$ with zero mean and covariance function $E[B_p(\pi_1) B_p(\pi_2)'] = (\pi_1 \wedge \pi_2 - \pi_1\pi_2) I_p$, and I_p is a $p \times p$ identity matrix. The tests KS_{na} and CM_{na} are then asymptotically distribution free, despite parameter estimation. For this special case, one can easily obtain the critical values for the KS_{na} and CM_{na} test statistics, and reject the null hypothesis H_{0a} for large values of KS_{na} and CM_{na}.

3.4 Local Power of Γ_{na}- based Tests

Now we study the local power properties of the test based on Γ_{na}. We focus on the local alternative

$$H_{1a,n} : \gamma_{nt} = \gamma_0 + n^{-1/2}\delta_1(t/n), \tag{3.3}$$

where $\delta_1(\cdot)$ is an arbitrary non-constant p-dimensional measurable function defined on the $[0,1]$ interval. Following Krämer, Ploberger, and Alt (1988) and Ploberger and Krämer (1992), we only require that $\delta_1(\cdot)$ be expressed as a uniform limit of functions that are constants on intervals. Clearly, if $\delta_1(t/n) = \delta\ 1(t/n \geq \pi_0)$ for some nonzero p-vector δ, then eq. (3.3) includes a one-time shift of the regression coefficient at time $n\pi_0$ as a special case.

Theorem 3.3 *Suppose Assumptions A1-A3 hold. Suppose the conditions in part (d) of Theorem 3.1 hold. Then under $H_{1a,n}$, we have $\Gamma_{na}(\cdot) \Rightarrow \Gamma_a(\cdot) + \Delta_a(\cdot)$, where for $0 \leq \pi \leq 1$,*

$$\Delta_a(\pi) = \Psi(1)^{-1/2}\left\{\int_0^\pi \Phi^{(1)}(s)\,\delta_1(s)\,ds - \Phi(\pi)\,\Phi(1)^{-1}\int_0^1 \Phi^{(1)}(s)\,\delta_1(s)\,ds\right\}, \tag{3.4}$$

and $\Phi^{(1)}(s) = (\partial/\partial s)\Phi(s)$.

Remark 4. Theorem 3.3 implies that the KS_{na} and CM_{na} tests have non-trivial power in detecting $n^{-1/2}$- local alternatives, provided $\Delta_a(\pi) \neq 0$ for π in a set of positive Lebesgue measure. Even a single break at time $t = n\pi_0$, i.e., $\delta_1(t/n) = \delta\ 1(t/n \geq \pi_0)$, affects the right-hand side of (3.4) for all $\pi \in (0,1)$, no matter where the structural change occurs. More importantly, structural changes affect the limiting rejection probabilities only via $\Phi^{(1)}(\cdot)\,\delta_1(\cdot)$; this is a semiparametric analog of the parametric case. In that case, if all structural shifts in the finite dimensional parameters are orthogonal to the mean regressor then the residual-based CUSUM test is not consistent. See Ploberger and Krämer (1992). In addition, if $\Phi(\pi)$ is linear in π (see (3.2)), then the expression for $\Delta_a(\pi)$ reduces to $\Delta_a(\pi) = \Psi(1)^{-1/2}\,\Phi\{\int_0^\pi \delta_1(s)\,ds - \pi\int_0^1 \delta_1(s)\,ds\}$.

Remark 5. As mentioned above, Γ_{na}- based tests have no power to detect structural changes in the nonparametric component $m_0(\cdot)$. This can be seen more clearly from the the right-hand side of (3.4), as $m_{nt}(z_t)-m_0(z_t)$ will necessarily be orthogonal to v_{nt} by the law of iterated expectations: $E[v_{nt}(m_{nt}(z_{nt})-m_0(z_{nt}))]=E[E(v_{nt}|z_{nt})(m_{nt}(z_{nt})-m_0(z_{nt}))]=0$. Thus, if one replaces one of the two v_{nt}'s in the definition of $\Phi(\pi)$ by $m_{nt}(z_t)-m_0(z_t)$, then the matrix $\Phi(\pi)$ becomes zero.

4 Asymptotic Properties of $\Gamma_{nb}(\cdot)$

In this section, we study the asymptotic properties of Γ_{nb} under H_{0b} and a sequence of local alternatives.

4.1 Asymptotic Null Distribution of $\Gamma_{nb}(\cdot)$

Let $S_{11}(\pi_1,\pi_2)\equiv\lim_{n\to\infty}n^{-1}\sum_{i=1}^{\lceil n(\pi_1\wedge\pi_2)\rfloor}E(\overline{f}_{ni}^2\varepsilon_{ni}^2)$, $S_{22}(\pi_1,\pi_2)\equiv\lim_{n\to\infty}n^{-3}\sum_{i=1}^{\lceil n\pi_1\rfloor}\sum_{j=1}^{n}\sum_{k=1}^{\lceil n\pi_2\rfloor}E[f_{ni}(z_{nj})f_{nk}(z_{nj})\varepsilon_{nj}^2]$, $S_{12}(\pi_1,\pi_2)\equiv\lim_{n\to\infty}n^{-2}\sum_{i=1}^{\lceil n\pi_1\rfloor}\sum_{j=1}^{\lceil n\pi_2\rfloor}E\left[\overline{f}_{ni}f_{nj}(z_{ni})\varepsilon_{ni}^2\right]$, and $S_{21}(\pi_1,\,\pi_2)\equiv S_{12}(\pi_2,\pi_1)$. The asymptotic null distribution of Γ_{nb} is given in the next theorem.

Theorem 4.1 *Suppose Assumptions A1-A3 hold. Suppose the conditions in part (d) of Theorem 3.1 hold and $nh^{2r}\to0$ as $n\to\infty$. Then under H_{0b}, $\Gamma_{nb}(\cdot)\Rightarrow\Gamma_b(\cdot)$, where Γ_b is a mean-zero Gaussian process with covariance kernel $E[\Gamma_b(\pi_1)\Gamma_b(\pi_2)]=\sigma_0^{-2}\sum_{i=1}^{2}\sum_{j=1}^{2}(-1)^{i+j}\times S_{ij}(\pi_1,\pi_2)$, and $\sigma_0^2=\lim_{n\to\infty}n^{-1}\sum_{t=1}^{n}E(\overline{f}_{nt}^2\varepsilon_{nt}^2)$.*

Remark 6. Note that we have strengthened the bandwidth condition from $nh^{4r}\to0$ to $nh^{2r}\to0$ in Theorem 4.1. This means that the optimal bandwidth chosen by standard least-squares or generalized cross-validation is not directly applicable to Γ_{nb}-based tests, because such bandwidths converge to zero at the rate $n^{-1/(q+2r)}$. Instead, we require undersmoothing. By the continuous mapping theorem, Theorem 4.1 implies that $KS_{nb}\overset{d}{\to}\sup_{0\leq\pi\leq1}|\Gamma_b(\pi)|$, and $CM_{nb}\overset{d}{\to}\int_0^1|\Gamma_b(\pi)|^2d\pi$. Again, the tests KS_{nb} and CM_{nb} are not asymptotically pivotal in general. We provide a bootstrap method to obtain p-values.

Remark 7. When $\{z_{nt}\}$ is also asymptotically stationary in the sense that $\max_{1\leq t\leq n}|f_{nt}(z)-f_n(z)|\to0\;\forall z$ for some continuous function $f_n(\cdot)$, we can show that under H_{0b}, $\Gamma_{nb}(\pi)=\sigma_0^{-1}\{n^{-1/2}\sum_{t=1}^{\lceil n\pi\rfloor}\overline{f}_{nt}\varepsilon_{nt}-n^{-1/2}\pi\sum_{t=1}^{n}\overline{f}_{nt}\varepsilon_{nt}\}+o_p(1)$. If $\sigma^2(\pi)$ is also linear in π (see (3.2)

), then by Theorem 1 in Herrndorf (1985) we have $\Gamma_{nb}(\cdot) \Rightarrow B_1(\cdot)$ under H_{0b}, where $B_1(\cdot)$ denotes the standard Brownian bridge defined on $[0,1]$. In this special case, the tests KS_{nb} and CM_{nb} are asymptotically distribution free despite parameter estimation. One rejects the null hypothesis H_{0b} for large values of KS_{nb} and CM_{nb}.

4.2 Local Power of Γ_{nb}- based Tests

Now we study the local power of Γ_{nb}-based tests. We focus on the local alternative

$$H_{1b,n} : \gamma_{nt} = \gamma_0 + n^{-1/2}\delta_1(t/n), \; m_{nt}(z_{nt}) = m_0(z_{nt}) + n^{-1/2}\delta_2(z_{nt}, t/n), \tag{4.1}$$

where $\delta_1(\cdot)$ is as defined in (3.3), and $\delta_2(\cdot,\cdot)$ is an arbitrary non-constant measurable function defined on $\mathcal{Z} \times [0,1]$, where $\mathcal{Z}$ is the support of z_{nt}. In addition, we follow Krämer, Ploberger, and Alt (1988) and require that for each z, $\delta_2(z,\cdot)$ can be expressed as a uniform limit of functions that are constants on intervals. Clearly, if $\delta_1(\cdot) \equiv 0$ and $\delta_2(z_{nt}, t/n) = \delta_2(z_{nt}) 1(t/n \geq \pi_0)$ in eq. (4.1), we have the special case of a one-time shift in the nonparametric regression component at time $n\pi_0$.

Theorem 4.2 *Suppose Assumptions A1-A3 hold. Suppose the conditions in part (d) of Theorem 3.1 hold, and $nh^{2r} \to 0$ as $n \to \infty$. Then under $H_{1b,n}$, we have $\Gamma_{nb}(\cdot) \Rightarrow \Gamma_b(\cdot) + \Delta_{b1}(\cdot) + \Delta_{b2}(\cdot)$, where for $0 \leq \pi \leq 1$,*

$$\begin{aligned}
\Delta_{b1}(\pi) &\equiv \sigma_0^{-1} \lim_{n\to\infty} \left\{ n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} E\left(\overline{f}_{ni} x'_{ni}\right) \delta_1(i/n) - n^{-2} \sum_{i=1}^{\lceil n\pi \rfloor} \sum_{j=1}^{n} E[f_{ni}(z_{nj}) x'_{nj}] \delta_1(j/n) \right\}, \\
\Delta_{b2}(\pi) &\equiv \sigma_0^{-1} \lim_{n\to\infty} \left\{ n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} E\left[\overline{f}_{ni} \delta_2(z_{ni}, i/n)\right] - n^{-2} \sum_{i=1}^{\lceil n\pi \rfloor} \sum_{j=1}^{n} E\left[f_{ni}(z_{nj}) \delta_2(z_{nj}, j/n)\right] \right\}.
\end{aligned}$$

Remark 8. The difference between the limiting distribution of Γ_{nb} under $H_{1b,n}$ and its asymptotic null distribution consists of two terms. The first arises from a shift in the parametric component γ_{nt}; the other arises from a shift in the nonparametric component $m_{nt}(\cdot)$. Γ_{nb}-based tests thus have non-trivial power in detecting $n^{-1/2}$-local alternatives whenever these two components do not vanish simultaneously. When we reject H_{0b}, we thus have evidence of structural breaks in either γ_{nt} or $m_{nt}(\cdot)$ or both. To see whether the structural break is caused by a break in the parametric component, we can apply the test of H_{0a} introduced in the previous section. If H_{0a} is not rejected, then the test indicates a structural break in the nonparametric component. Of course, care must be taken to ensure the correct probability of Type I error when conducting such a sequential test.

Remark 9. If $\{z_{nt}\}$ is also asymptotically stationary as in Remark 7, then the proof of Theorem 4.2 can be greatly simplified. In this case, one can readily show that $n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n} E\left[f_{ni}(z_{nj})x_{nj}'\right]\delta_1(j/n) = \pi n^{-1}\sum_{j=1}^{n}E(f_{nj}x_{nj}')\delta_1(j/n)+o(1)$, and $n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j=1}^{n}E[f_{ni}(z_{nj})\delta_2(z_{nj},j/n)] = \pi n^{-1}\sum_{j=1}^{n} E[f_{nj}\delta_2(z_{nj},j/n)]+o(1)$. Now the expressions for $\Delta_{b1}(\pi)$ and $\Delta_{b2}(\pi)$ reduce to $\Delta_{b1}(\pi) = \sigma_0^{-1}(\int_0^\pi R_1^{(s)}(s)\delta_1(s)\,ds - \pi\int_0^1 R_1^{(s)}(s)\delta_1(t)\,dt)$, and $\Delta_{b2}(\pi) = \sigma_0^{-1}[R_2(\pi)-\pi R_2(1)]$, where $R_1(\pi)\equiv\lim_{n\to\infty}n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}E(f_{ni}x_{ni}')$, $R_1^{(1)}(\pi) = (\partial/\partial\pi)R_1(\pi)$, $R_2(\pi)\equiv\int_0^\pi\int f^2(z)\delta_2(z,t)\,dzdt$, and $f(z)\equiv\lim_{n\to\infty}f_n(z)$.

5 Bootstrap Tests

From the previous two sections we see that if we allow fixed breaks in the process $\{x_{nt}, z_{nt}, \varepsilon_{nt}\}$, neither Γ_{na}-based tests nor Γ_{nb}-based tests are asymptotically pivotal in general, preventing tabulation of critical values. To obtain the p-values, we now propose and analyze a wild bootstrap version of our tests.

From the proofs of Theorem 3.2 and Theorem 4.1, we have that under the applicable null hypothesis

$$\Gamma_{na}(\pi) = \widetilde{\Gamma}_{na}(\pi) + o_p(1), \text{ and } \Gamma_{nb}(\pi) = \widetilde{\Gamma}_{nb}(\pi) + o_p(1), \tag{5.1}$$

where

$$\begin{aligned}\widetilde{\Gamma}_{na}(\pi) &= n^{-1/2}\widehat{\Psi}^{-1/2}\left\{\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}^2(x_{nt}-\widehat{x}_{nt})\varepsilon_{nt}\right.\\ &\qquad \left. -\Phi(\pi)\Phi(1)^{-1}\sum_{t=1}^{n}\widehat{f}_{nt}^2(x_{nt}-\widehat{x}_{nt})\varepsilon_{nt}\right\} \qquad (5.2)\\ \widetilde{\Gamma}_{nb}(\pi) &= n^{-1/2}\widehat{\sigma}^{-1}\left\{\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}\varepsilon_{nt} - \sum_{t=1}^{n}\widehat{f}_{\lceil n\pi\rfloor}(z_{nt})\varepsilon_{nt}\right\}, \qquad (5.3)\end{aligned}$$

and $\widehat{f}_{\lceil n\pi\rfloor}(z_{nt})\equiv n^{-1}\sum_{s=1}^{\lceil n\pi\rfloor}K_{hts}$. Even though Theorems 3.2 and 4.1 imply that $\{\widetilde{\Gamma}_{na}(\pi), 0\le\pi\le1\}$ and $\{\widetilde{\Gamma}_{nb}(\pi), 0\le\pi\le1\}$ are not asymptotically pivotal in general under the relevant null, we can mimic the asymptotic distribution of Γ_{na} (resp. Γ_{nb}) by bootstrapping $\widetilde{\Gamma}_{na}$ (resp. $\widetilde{\Gamma}_{nb}$).

To obtain the bootstrap versions of our test statistics, we define the wild bootstrap residuals as

$$u_{nt}^* \equiv \widetilde{u}_{nt}\eta_t, \tag{5.4}$$

where $\{\eta_t\}$ satisfy the conditions stated in Assumption A4(i) below. One can draw such a sequence $\{\eta_t\}$ in a number of ways. In our simulations, we draw $\{\eta_t\}$ independently from a distribution with masses $c = \left(1+\sqrt{5}\right)/\left(2\sqrt{5}\right)$ and $1-c$ at the points $\left(1-\sqrt{5}\right)/2$ and $\left(1+\sqrt{5}\right)/2$, respectively. Consequently, the wild bootstrap draws each u_{nt}^* from a different distribution with mean zero and variance $\widetilde{u}_{nt}^2$, conditional on the data. Our bootstrap processes are then defined by

$$\begin{aligned}\Gamma_{na}^*(\pi) &= n^{-1/2}\widehat{\Psi}^{*-1/2}\left\{\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}^2\left(x_{nt}-\widehat{x}_{nt}\right)u_{nt}^*\right.\\ &\quad\left.-\widehat{\Phi}(\pi)\,\widehat{\Phi}(1)^{-1}\sum_{t=1}^{n}\widehat{f}_{nt}^2\left(x_{nt}-\widehat{x}_{nt}\right)u_{nt}^*\right\} \qquad (5.5)\\ \Gamma_{nb}^*(\pi) &= n^{-1/2}\widehat{\sigma}^{*-1}\left\{\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}u_{nt}^*-\sum_{t=1}^{n}\widehat{f}_{\lceil n\pi\rfloor}(z_{nt})\,u_{nt}^*\right\}, \qquad (5.6)\end{aligned}$$

where $\widehat{\Phi}(\pi) = n^{-1}\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}^2\left(x_{nt}-\widehat{x}_{nt}\right)\left(x_{nt}-\widehat{x}_{nt}\right)'$, $\widehat{\Psi}^* = n^{-1}\sum_{t=1}^{n}\widehat{f}_{nt}^4 u_{nt}^{*2}(x_{nt}-\widehat{x}_{nt})(x_{nt}-\widehat{x}_{nt})$, and $\widehat{\sigma}^{*2} = n^{-1}\sum_{t=1}^{n}\widehat{f}_{nt}^2 u_{nt}^{*2}$. Using Γ_{na}^*, we construct the bootstrap version KS_{na}^* of the statistic KS_{na}. We repeat this procedure B times to obtain the sequence $\left\{KS_{na,j}^*\right\}_{j=1}^{B}$. We reject the null when, for example, $p^* = B^{-1}\sum_{j=1}^{B}1(KS_{na} \leq KS_{na,j}^*)$ is smaller than the desired significance level. The procedure is analogous for CM_{na}, KS_{nb}, and CM_{nb}, where we use Γ_{nb}^* for the latter two statistics. To prove the validity of the above bootstrap procedure, we need the following additional assumption.

Assumption A4. (i) $\{\eta_t\}$ are IID with $E(\eta_t) = 0$, $E\left(\eta_t^2\right) = 1$, and $E(\eta_t^4) < \infty$, and independent of the process $\{y_{nt}, x_{nt}, z_{nt}\}$. (ii) $n^{-1}\sum_{t=1}^{n}\widehat{f}_{\lceil n\pi\rfloor}^2(z_{nt})\left(\widetilde{u}_{nt}-\varepsilon_{nt}\right)^2 = o_p(1)$ for each $\pi \in [0,1]$.

Assumption A4(i) is standard in the literature. Assumption A4(ii) is a high level assumption that parallels to the second condition in Assumption A10 of Delgado and Fiteni (2002). Even though not stated explicitly, the proof of bootstrap consistency in Hansen (2000a) also relies upon similar conditions.

Theorem 5.1 *Suppose Assumptions A1-A4 hold. Then* $\Gamma_{na}^{*}(\cdot) \overset{p}{\Rightarrow} \Gamma_{a}(\cdot)$, *and* $\Gamma_{nb}^{*}(\cdot) \overset{p}{\Rightarrow} \Gamma_{b}(\cdot)$.

Remark 10. Theorem 5.1 shows that each bootstrapped process ($\{\Gamma_{na}^{*}\}$ or $\{\Gamma_{nb}^{*}\}$) converges weakly to the relevant limiting null Gaussian process, thus providing a valid asymptotic basis for approximating the limiting null distribution of test statistics based on $\{\Gamma_{na}\}$ or $\{\Gamma_{nb}\}$. By the properties of the wild bootstrap, our approximation to the limiting null distribution is valid even when the null hypothesis does not hold for the underlying data. This helps ensure reasonable power for the bootstrap test against potential departures from the null hypothesis.

Remark 11. It is worth mentioning that the validity of the above bootstrap procedure relies heavily on Assumption A4(ii), which, unfortunately, we are unable to relax. This assumption can be easily satisfied under either null or local alternatives. This is true no matter whether we have fixed breaks in the marginal PDF $f_{nt}(\cdot)$ of z_{nt} or not. Nevertheless, it may be difficult to ensure this in the case of global alternatives. We note that this is a phenomenon associated with many bootstrap versions of tests for structural change, including those of Hansen (2000a) and Delgado and Fiteni (2002).

In the following, we restrict ourselves to the case where the marginal PDF $f_{nt}(\cdot)$ of z_{nt} has only small breaks and the linearity assumption in (3.2) holds. In this case, we can re-examine the proofs of Theorem 3.2 and Theorem 4.1 and obtain $\Gamma_{na}(\pi) = \widetilde{\Gamma}_{na}(\pi) + o_p(1)$, and $\Gamma_{nb}(\pi) = \widetilde{\Gamma}_{nb}(\pi) + o_p(1)$, where $\widetilde{\Gamma}_{na}(\pi) = n^{-1/2}\widehat{\Psi}^{-1/2}\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}^{2}(x_{nt}-\widehat{x}_{nt})\varepsilon_{nt}$ $-\pi n^{-1/2}\widehat{\Psi}^{-1/2}\sum_{t=1}^{n}\widehat{f}_{nt}^{2}(x_{nt}-\widehat{x}_{nt})\varepsilon_{nt}$, and $\widetilde{\Gamma}_{nb}(\pi) = n^{-1/2}\widehat{\sigma}^{-1}\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}\varepsilon_{nt} - \pi n^{-1/2}\widehat{\sigma}^{-1}\sum_{t=1}^{n}$ $\widehat{f}_{nt}\varepsilon_{nt}$. In this special case, we propose the following bootstrap processes

$$\begin{aligned}
\Gamma_{na}^{*}(\pi) &= n^{-1/2}\widehat{\Psi}^{*-1/2}\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}^{2}(x_{nt}-\widehat{x}_{nt})u_{nt}^{*} - \pi n^{-1/2}\widehat{\Psi}^{*-1/2}\sum_{t=1}^{n}\widehat{f}_{nt}^{2}(x_{nt}-\widehat{x}_{nt})u_{nt}^{*}, \\
\Gamma_{nb}^{*}(\pi) &= n^{-1/2}\widehat{\sigma}^{*-1}\sum_{t=1}^{\lceil n\pi\rfloor}\widehat{f}_{nt}u_{nt}^{*} - \pi n^{-1/2}\widehat{\sigma}^{*-1}\sum_{t=1}^{n}\widehat{f}_{nt}u_{nt}^{*},
\end{aligned}$$

where $\widehat{\Psi}^{*}$and $\widehat{\sigma}^{*2}$ are as defined above. In this case, we have the following corollary.

Corollary 5.2 *Suppose Assumptions A1-A3 and A4(i) hold. Suppose that the linearity condition in (3.2) holds, and* $max_{1\le t\le n}|f_{nt}(z) - f_{n}(z)| \to 0\ \forall z$ *for some continuous function* $f_{n}(\cdot)$. *Then* $\Gamma_{na}^{*}(\cdot) \overset{p}{\Rightarrow} B_{p}(\cdot)$, *and* $\Gamma_{nb}^{*}(\cdot) \overset{p}{\Rightarrow} B_{1}(\cdot)$.

Remark 12. We sketch the proof of the above corollary in the appendix. A crucial point here is that we do not require Assumption A4(ii). Under the stated conditions and the extra condition in part (d) of Theorem 3.1, both Γ_{na}- and Γ_{nb}-based tests are asymptotically pivotal under the relevant null hypothesis. We thus are able to compare the bootstrap versions of the tests with those based on the asymptotic critical values in this case.

6 Monte Carlo Simulations

In this section we present a small set of Monte Carlo experiments to evaluate the finite sample performance of our tests. We consider the following DGP:

$$y_{nt} = \gamma_{nt} x_{nt} + m_{nt}(z_{nt}) + \varepsilon_{nt},\ \varepsilon_{nt} = \sqrt{\vartheta_{nt}}\zeta_{1nt}, \tag{6.1}$$

where $\vartheta_{nt} = 0.05 + 0.90\vartheta_{n,t-1} + 0.05\varepsilon^2_{n,t-1}$; $z_{nt} = 0.5 + 0.8z_{n,t-1} + \zeta_{2nt}$; $x_{nt} = 1 + \cos(z_{nt}) + v_{nt}$; and $\{\zeta_{1nt}\}$, $\{\zeta_{2nt}\}$, and $\{v_{nt}\}$ are each IID $N(0,1)$ sequences, mutually independent of each other. The subscripts for γ_{nt} and $m_{nt}(\cdot)$ indicate that both the parametric and nonparametric components of the regression function may be time-varying. We consider the following specifications of γ_{nt} and $m_{nt}(\cdot)$:

$$\gamma_{nt} = 1 + \delta_1 1(t \geq \lceil n\pi_0 \rfloor),\ \text{and}\ m_{nt}(z_{nt}) = z_{nt} - 0.5z^2_{nt} + \frac{\delta_2 \exp(z_{nt})}{1+\exp(z_{nt})} 1(t \geq \lceil n\pi_0 \rfloor). \tag{6.2}$$

We consider three different break ratios $\pi_0 = 0.25, 0.5, 0.75$, and (δ_1, δ_2) pairs with $\delta_1, \delta_2 = 0$, 0.25, 0.5, and 1.

To construct the test statistics, we choose the fourth order ($r = 4$) Epanechnikov kernel, $K(u) = \frac{3}{4\sqrt{5}}(\frac{15}{8} - \frac{7}{8}u^2)(1 - \frac{1}{5}u^2)1(|u| \leq \sqrt{5})$. To motivate our choice of the bandwidth h, let $X = (x_{n1}, ..., x_{nn})'$, $Y = (y_{n1}, ..., y_{nn})'$, $\widetilde{X} = (\widetilde{x}_{n1}, ..., \widetilde{x}_{nn})'$, $\widetilde{Y} = (\widetilde{y}_{n1}, ..., \widetilde{y}_{nn})'$, and $\mathbb{U} = \widetilde{Y} - \widetilde{X}\widehat{\gamma}$, where $\widetilde{x}_{nt} = (x_{nt} - \widehat{x}_{nt})\widehat{f}_{nt}$, and $\widetilde{y}_{nt} = (y_{nt} - \widehat{y}_{nt})\widehat{f}_{nt}$. The tth element of $\mathbb{U}$ is given by $\widetilde{u}_{nt}\widehat{f}_{nt}$, i.e., the residual from the partially linear regression weighted by the density estimate, $\widehat{f}_{nt}$. Let $\mathbb{K}$ denote the $n \times n$ smoothing matrix whose (s,t)th element is given by $K_{hst}/(n\widehat{f}_{ns})$, and let W =diag$(\widehat{f}_{n1}, ..., \widehat{f}_{nn})$. Then $\widetilde{X} = W(I_n - \mathbb{K})X$, $\widetilde{Y} = W(I_n - \mathbb{K})Y$, and $\widehat{\gamma} = (\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'\widetilde{Y}$. Consequently, $\mathbb{U} = \widetilde{Y} - \widetilde{X}(\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}'\widetilde{Y} = A(h)Y$, where $A(h) = [I_n - \widetilde{X}(\widetilde{X}'\widetilde{X})^{-1}\widetilde{X}']W(I_n - \mathbb{K})$. Following Xu (2006), we propose to choose h to minimize the following generalized cross-validation (GCV) score,

$$GCV_n(h) = n^{-1}Y'A(h)'A(h)Y \ / \ \left(n^{-1}\text{tr}(A(h))\right)^2. \tag{6.3}$$

Let $\widehat{h}$ denote the minimizer of $GCV_n(h)$. Then $\widehat{h}$ is optimal in the sense of Xu (2006), and $\widehat{h} \propto n^{-1/(2r+q)} = n^{-1/9}$. Since undersmoothing is required for the tests based on $\Gamma_{nb}(\cdot)$, we apply a rule of thumb to choose h according to $h = \widehat{h} n^{1/9} n^{-1/\lambda}$. We study the behavior of our tests with different choices $\lambda = 7, 6, 5$ in order to examine the sensitivity of our test to the bandwidth sequence. Robinson (1991, p.448) proposes similar devices. Note that these choices for h and the kernel function meet the requirements for both tests.

In the following, we will report the empirical rejection frequencies of the tests KS_{na}, CM_{na}, KS_{nb} and CM_{nb} for different choices of δ_1 and δ_2. Since the process $\{x_{nt}, z_{nt}, \varepsilon_{nt}\}$ does not exhibit fixed breaks under either the null or local alternatives, the linearity condition in (3.2) holds and Corollary (5.2) applies. Both Γ_{na}- and Γ_{nb}-based tests are asymptotically pivotal under the relevant null hypothesis and they can be conducted based on the asymptotic critical values. To see how well our nonparametric tests based on asymptotic distributions perform in finite samples, we report the rejection frequencies for both the bootstrap versions of the tests, denoted as KS_{na}^b, CM_{na}^b, KS_{nb}^b, CM_{nb}^b, and the tests based on critical values from the pivotal asymptotic distributions. We use 1000 replications for each sample size n and 199 bootstrap resamples for the bootstrap test in each replication. For ease of reference, we refer to the KS_{na}, CM_{na}, KS_{na}^b, and CM_{na}^b tests as a-tests and the KS_{nb}, CM_{nb}, KS_{nb}^b, and CM_{nb}^b tests as b-tests.

6.1 Finite Sample Level

We first examine the finite sample performance of a-tests under H_{0a}. To do so, we set $\delta_1 = 0$ and allow δ_2 to take different values (0, 0.25, 0.5, and 1) in (6.2). Table 1 reports the empirical rejection frequencies of these tests at the nominal level 0.05 and break ratio $\pi_0 = 0.5$. We summarize some important findings from Table 1. (a) Surprisingly, the CM_{na} test based on asymptotic critical values is as good as the bootstrap version of the tests (KS_{na}^b and CM_{na}^b). The KS_{na} test based on asymptotic critical values is undersized for small sample sizes ($n = 100, 200$), but its level improves as n increases. (b) All tests are robust to different choices of bandwidth h (or equivalently λ). (c) The CM_{na}, KS_{na}^b and CM_{na}^b tests behave similarly. In all cases, the empirical levels of these tests are close to the nominal levels. (d) As predicted by our asymptotic theory, the empirical levels of the KS_{na}, CM_{na}, KS_{na}^b and CM_{na}^b tests are robust to the presence of structural changes in the nonparametric component.

To examine the finite sample performance of b-tests under H_{0b}, we set $\delta_1 = \delta_2 = 0$ in (6.2).

Table 1: Finite sample rejection frequencies under H_{0a} (nominal level: 0.05)

δ_2	Test\\n	100			200			400		
		$\lambda = 7$	$\lambda = 6$	$\lambda = 5$	$\lambda = 7$	$\lambda = 6$	$\lambda = 5$	$\lambda = 7$	$\lambda = 6$	$\lambda = 5$
0	KS_{na}	0.023	0.023	0.023	0.028	0.023	0.025	0.043	0.037	0.038
	CM_{na}	0.044	0.044	0.050	0.030	0.029	0.031	0.050	0.047	0.051
	KS^b_{na}	0.044	0.048	0.044	0.045	0.040	0.036	0.051	0.053	0.046
	CM^b_{na}	0.043	0.045	0.047	0.033	0.036	0.036	0.048	0.053	0.050
.25	KS_{na}	0.023	0.023	0.023	0.028	0.023	0.025	0.043	0.041	0.041
	CM_{na}	0.044	0.044	0.050	0.030	0.029	0.031	0.047	0.047	0.053
	KS^b_{na}	0.044	0.048	0.044	0.045	0.040	0.036	0.052	0.054	0.047
	CM^b_{na}	0.043	0.045	0.047	0.033	0.036	0.036	0.047	0.051	0.049
.5	KS_{na}	0.023	0.023	0.023	0.028	0.023	0.025	0.042	0.037	0.041
	CM_{na}	0.044	0.044	0.050	0.030	0.029	0.031	0.052	0.049	0.049
	KS^b_{na}	0.044	0.048	0.044	0.045	0.040	0.036	0.052	0.052	0.049
	CM^b_{na}	0.043	0.045	0.047	0.033	0.036	0.036	0.050	0.052	0.051
1	KS_{na}	0.023	0.023	0.023	0.028	0.023	0.025	0.043	0.040	0.040
	CM_{na}	0.044	0.044	0.050	0.030	0.029	0.031	0.048	0.048	0.053
	KS^b_{na}	0.044	0.048	0.044	0.045	0.040	0.036	0.050	0.055	0.047
	CM^b_{na}	0.043	0.045	0.047	0.033	0.036	0.036	0.046	0.052	0.051

Note. $h = \widehat{h} n^{1/9} n^{-1/\lambda}$, where $\widehat{h}$ is chosen by GCV.

Table 2: Finite sample rejection frequencies under H_{0b} (nominal level: 0.05)

Test\\n	100			200			400		
	$\lambda = 7$	$\lambda = 6$	$\lambda = 5$	$\lambda = 7$	$\lambda = 6$	$\lambda = 5$	$\lambda = 7$	$\lambda = 6$	$\lambda = 5$
KS_{nb}	0.050	0.044	0.040	0.055	0.057	0.053	0.060	0.057	0.055
CM_{nb}	0.059	0.058	0.055	0.062	0.061	0.059	0.061	0.057	0.057
KS^b_{nb}	0.068	0.062	0.054	0.059	0.070	0.064	0.068	0.064	0.056
CM^b_{nb}	0.064	0.057	0.056	0.057	0.060	0.063	0.060	0.056	0.054

Note. $h = \widehat{h} n^{1/9} n^{-1/\lambda}$, where $\widehat{h}$ is chosen by GCV.

Table 2 reports the empirical rejection frequencies of these tests at the nominal level 0.05 and break ratio $\pi_0 = 0.5$. From Table 2 we have findings similar to those in Table 1, except that the KS_{nb} test based on asymptotic critical values performs almost as well as CM_{nb} and the bootstrap version of these two tests (KS^b_{nb} and CM^b_{nb}). As before, all tests are robust to the choice of bandwidth; and all tests have empirical levels close to their nominal levels.

We also conducted tests for other choices of break ratios: $\pi_0 = 0.25, 0.75$ (not tabulated here). We find that the results are similar to the above.

Table 3: Finite sample rejection frequencies under $H_{1a,n}$ (nominal level: 0.05)

δ_1	δ_2	$n=100$						$n=200$					
		$\lambda=7$		$\lambda=6$		$\lambda=5$		$\lambda=7$		$\lambda=6$		$\lambda=5$	
		KS^b	CM^b	KS^b	CM^b	KS^b	CM^b	KS^b	CM^b	KS^b	CM^b	KS^b	CM^b
.25	0	0.12	0.12	0.12	0.12	0.11	0.12	0.23	0.24	0.22	0.23	0.23	0.23
	.25	0.11	0.12	0.11	0.12	0.11	0.12	0.22	0.23	0.22	0.22	0.22	0.22
	.5	0.11	0.11	0.11	0.12	0.11	0.12	0.21	0.21	0.21	0.20	0.21	0.20
	1	0.09	0.09	0.09	0.10	0.10	0.09	0.16	0.17	0.17	0.16	0.17	0.17
.5	0	0.39	0.39	0.37	0.35	0.35	0.35	0.70	0.72	0.70	0.71	0.69	0.69
	.25	0.38	0.36	0.35	0.34	0.33	0.33	0.70	0.71	0.69	0.70	0.68	0.68
	.5	0.35	0.35	0.33	0.33	0.31	0.31	0.68	0.69	0.67	0.67	0.66	0.66
	1	0.29	0.28	0.27	0.27	0.27	0.25	0.60	0.60	0.60	0.59	0.58	0.59
1	0	0.88	0.88	0.86	0.86	0.85	0.86	0.99	0.99	0.99	0.98	0.99	0.99
	.25	0.87	0.87	0.85	0.85	0.84	0.85	0.99	0.99	0.98	0.98	0.99	0.99
	.5	0.85	0.85	0.83	0.83	0.82	0.84	0.99	0.99	0.98	0.98	0.98	0.98
	1	0.80	0.81	0.77	0.78	0.76	0.78	0.99	0.98	0.97	0.97	0.98	0.98

Note. $h=\widehat{h}n^{1/9}n^{-1/\lambda}$, where $\widehat{h}$ is chosen by GCV. KS^b and CM^b refer to KS^b_{na} and CM^b_{na}.

6.2 Finite Sample Power

To examine the power performance of the tests, we first focus on the a-tests. Here and below we conserve space by only reporting results for the bootstrap version of the tests. Table 3 reports the results of these tests based on the bootstrap critical values where the break ratio π_0 is 0.5. Some of the main findings from Table 3 are: (a) As in the level study, the KS^b_{na} and CM^b_{na} tests behave similarly. (b) As the sample size increases, the powers of both tests increase. (c) The choices of the bandwidth sequence have little influence on the power performance of these tests. (d) For fixed δ_2, the powers of both tests increase as the break size δ_1 increases. (e) For fixed δ_1, the power does not increase as δ_2 increases. Instead, there is a general trend suggesting that power may be adversely affected by an increase in δ_2.

We next examine the power performance of the b-tests. Table 4 reports the finite sample performance of these tests based on the bootstrap critical values where the break ratio π_0 is 0.5. We find that: (a) As in the case for size study, the KS^b_{nb} and CM^b_{nb} tests behave similarly in terms of power. (b) As n increases, the powers of both tests increase. (c) The choice of the bandwidth sequence has little influence on the power performance of these tests. (d) When either δ_1 or δ_2 increases, the powers of both tests increase.

Comparing the results in Table 4 for b-tests with those in Table 3 for a-tests, we have two interesting findings. First, when there is a structural break in the parametric component only

Table 4: Finite sample rejection frequencies under $H_{1b,n}$ (nominal level: 0.05)

δ_1	δ_2	$n=100$						$n=200$					
		$\lambda=7$		$\lambda=6$		$\lambda=5$		$\lambda=7$		$\lambda=6$		$\lambda=5$	
		KS^b	CM^b	KS^b	CM^b	KS^b	CM^b	KS^b	CM^b	KS^b	CM^b	KS^b	CM^b
0	.25	0.14	0.14	0.14	0.14	0.14	0.15	0.23	0.24	0.24	0.23	0.23	0.23
	.5	0.38	0.37	0.38	0.38	0.38	0.37	0.68	0.67	0.67	0.67	0.66	067
	1	0.89	0.88	0.88	0.87	0.88	0.88	0.99	0.99	0.99	0.99	0.99	0.99
.25	0	0.12	0.12	0.11	0.11	0.11	0.11	0.15	0.15	0.15	0.14	0.15	0.13
	.25	0.29	0.28	0.29	0.28	0.29	0.28	0.53	0.51	0.51	0.50	0.52	0.50
	.5	0.60	0.58	0.59	0.58	0.59	0.58	0.90	0.90	0.90	0.89	0.89	0.88
	1	0.94	0.93	0.94	0.93	0.94	0.93	0.99	0.99	0.99	0.99	0.99	0.99
.5	0	0.22	0.21	0.21	0.21	0.22	0.21	0.38	0.37	0.37	0.36	0.37	0.38
	.25	0.48	0.47	0.48	0.47	0.47	0.47	0.79	0.78	0.78	0.77	0.78	0.77
	.5	0.75	0.73	0.75	0.73	0.74	0.72	0.97	0.97	0.97	0.96	0.97	0.96
	1	0.97	0.97	0.96	0.96	0.97	0.97	0.99	0.99	0.99	0.99	0.99	0.99
1	0	0.51	0.50	0.51	0.50	0.50	0.49	0.79	0.78	0.78	0.77	0.78	0.77
	.25	0.75	0.72	0.74	0.72	0.73	0.72	0.96	0.95	0.95	0.94	0.96	0.95
	.5	0.89	0.88	0.89	0.87	0.89	0.87	0.99	0.99	0.98	0.98	0.99	0.99
	1	0.99	0.98	0.98	0.98	0.98	0.98	1	1	1	1	1	1

Note. $h = \widehat{h} n^{1/9} n^{-1/\lambda}$, where $\widehat{h}$ is chosen by GCV. KS^b and CM^b refer to KS_{nb}^b and CM_{nb}^b.

(i.e., $\delta_1 \neq 0, \delta_2 = 0$), the a-tests dominate the b-tests in terms of power. Second, except in this case, the b-tests dominate the a-tests in terms of power for the same values of (δ_1, δ_2). This is not surprising, because the a-tests are designed to test for structural changes in the parametric component only. Even though we cannot prove that a-tests are more powerful than the b-tests when we have only breaks in the parametric component, they definitely outperform the b-tests for certain alternatives. On the other hand, if we have breaks in both the parametric and nonparametric components, the b-tests can pick up both types of divergence from the null and are thus expected to be more powerful than the a-tests against certain alternatives. As is well known, no theory can ensure a uniform dominance of one class of such tests over the other class.

6.3 Comparing the a-tests with the Andrews test

It is interesting to compare our a-tests with the Andrews (1993) test. In order to implement the Andrews test, we must specify the conditional mean function parametrically. Suppose that the data are generated according to (6.1) and (6.2), but we pretend that the DGP is

Table 5: Finite sample size of Andrews's test for DGP (6.1) and (6.2) (δ_1=0, nominal level: 0.05)

		Test H_{01}			Test H_{02}		
n	δ_2\test	$SupF_n$	$ExpF_n$	$AveF_n$	$SupF_n$	$ExpF_n$	$AveF_n$
100	0	0.286	0.314	0.256	0.299	0.323	0.243
	0.25	0.286	0.317	0.265	0.331	0.350	0.259
	0.5	0.304	0.336	0.289	0.398	0.425	0.328
	1	0.350	0.378	0.315	0.566	0.602	0.547
200	0	0.308	0.348	0.244	0.379	0.386	0.268
	0.25	0.319	0.339	0.268	0.412	0.409	0.320
	0.5	0.353	0.366	0.306	0.501	0.520	0.434
	1	0.438	0.468	0.405	0.753	0.766	0.720

linear: $y_{nt} = \beta_{0nt} + \beta_{1nt}x_{nt} + \beta_{2nt}z_{nt} + u_{nt}$, and we test the null hypothesis

$$H_{01} : \beta_{1nt} = \beta_1 \text{ for some } \beta_1 \in \mathbb{R} \text{ for all } t \geq 1 \tag{6.4}$$

or

$$H_{02} : \beta_{0nt} = \beta_0 \text{ and } \beta_{1nt} = \beta_1 \text{ for some } (\beta_0, \beta_1) \in \mathbb{R}^2 \text{ for all } t \geq 1. \tag{6.5}$$

We follow Hansen (2000a) and calculate his statistics $SupF_n$, $ExpF_n$, and $AveF_n$. For example, to test H_{01}, we first run the restricted OLS regression $y_{nt} = \beta_0 + \beta_1 x_{nt} + \beta_2 z_{nt} + e_{nt}$, and denote the residuals as $\widehat{e}_t$ and variance estimate as $\widehat{\sigma}^2 = (n-3)^{-1}\sum_{t=1}^n \widehat{e}_t^2$. Then we run the set of unrestricted regressions: $y_{nt} = \beta_0 + \beta_1 x_{nt} + \beta_2 z_{nt} + \vartheta_1 x_{nt} 1(t \geq s) + e_{nt}$. Denote the residuals from the above regression as $\widehat{e}_{ts}$ and the variance estimate as $\widehat{\sigma}_s^2 = (n-4)^{-1}\sum_{t=1}^n \widehat{e}_{ts}^2$. Define $F_s = [(n-3)\widehat{\sigma}^2 - (n-4)\widehat{\sigma}_s^2]/\widehat{\sigma}_s^2$. Then $SupF_n$, $ExpF_n$, and $AveF_n$ are defined as

$$SupF_n = \sup_{s \in (\tau_1, \tau_2)} F_s, \ ExpF_n = \log\left(\int_s \exp(F_s/2)dw(s)\right), \text{ and } AveF_n = \int_s F_s dw(s),$$

where $w(s) = 1/(\tau_2 - \tau_1)$ if $s \in (\tau_1, \tau_2)$ and 0 otherwise. The asymptotic null distributions of these test statistics are given in Andrews (1993) and Andrews and Ploberger (1994). Table 5 reports the finite sample "level" of these tests for the case $\pi_0 = 0.5$ when we choose $\tau_1 = \lceil 0.15n \rfloor$, $\tau_2 = \lceil 0.85n \rfloor$ and the number of replications to be 1000. From Table 5, we see that under this functional misspecification, the level of the Andrews test is highly distorted, and the distortion tends to increase as n or δ_2 increases. In this case, it is inappropriate to compare the power performance of the Andrews test to that of our a-tests. In addition, it is difficult, if possible at all, to calculate the level-adjusted empirical power.

Nevertheless, if we stick to linear DGPs, we can compare the power performance of the

Table 6. Finite sample rejection frequencies under DGP (6.6) (nominal level: 0.05)

		Andrews's tests			Our a-tests					
					$\lambda=7$		$\lambda=6$		$\lambda=5$	
n	Δ_1	$SupF_n$	$ExpF_n$	$AveF_n$	KS_{na}^b	CM_{na}^b	KS_{na}^b	CM_{na}^b	KS_{na}^b	CM_{na}^b
100	0	0.042	0.059	0.054	0.046	0.038	0.050	0.044	0.052	0.049
	0.25	0.157	0.209	0.216	0.146	0.156	0.146	0.152	0.144	0.138
	0.5	0.501	0.588	0.594	0.433	0.428	0.426	0.429	0.412	0.421
	1	0.977	0.986	0.986	0.927	0.914	0.914	0.908	0.900	0.894
200	0	0.042	0.046	0.046	0.040	0.046	0.041	0.039	0.049	0.051
	.25	0.261	0.320	0.335	0.259	0.264	0.267	0.262	0.241	0.243
	0.5	0.812	0.868	0.872	0.741	0.742	0.730	0.726	0.699	0.716
	1	1	1	1	0.993	0.994	0.990	0.992	0.984	0.987

Note. For our nonparametric test, we set $h=\widehat{h}n^{1/9}n^{-1/\lambda}$ where $\widehat{h}$ is chosen by GCV.

two sets of tests. For simplicity, we consider the following linear DGP:

$$y_{nt}=\beta_{1nt}x_{nt}+\beta_{2nt}z_{nt}+\varepsilon_{nt}, \tag{6.6}$$

where we generate $\{x_{nt}\}$ and $\{z_{nt}\}$ as two independent $N(0,1)$ sequences with independent observations and with $\{\varepsilon_{nt}\}$ as in (6.1). We consider testing the null hypothesis H_{01} specified in (6.4). The Andrews test of H_{01} requires no structural change in β_{2nt}, so we now assume that $\beta_{2nt}=1$ for all $t\geq 1$. Table 6 compares the Andrews test of H_{01} with our a-tests when the parameters are generated according to $\beta_{1nt}=1+\Delta_1 1(t\geq\lceil n/2\rfloor)$, and $\beta_{2nt}=1$ for all $t\geq 1$. To save space, for our nonparametric a-tests, we only report the empirical rejection frequencies based upon the bootstrap critical values with 199 bootstrap resamples in each replication. The total number of replications is 1000 for each scenario. When $\Delta_1=0$, Table 6 reports the level behavior of both types of tests. Clearly, the levels of all tests behave reasonably well. When $\Delta_1\neq 0$, Table 6 reports the power behavior of both types of tests. We see that the Andrews parametric tests outperform our nonparametric test in most cases. Nevertheless, the power loss of our a-tests in this case is not severe.

7 Concluding Remarks

In this paper we propose two tests for structural change in partially linear time-series models. One procedure tests for structural change in the parametric component only, and the other tests for structural change in both the parametric and nonparametric components jointly. Our tests complement the conventional procedures for testing for structural change in parametric models and are natural diagnostics for testing for structural change in partially linear regression models. In particular, both tests have non-trivial power to detect deviations from the null at the parametric rate $n^{-1/2}$. The generality of our second test does not come for free, as it requires more stringent assumptions on the bandwidth parameter.

REFERENCES

Andrews, D. W. K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821-856.

Andrews, D. W. K. & W. Ploberger (1994) Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* 62, 1383-1414.

Aneiros-Pérez, G., W. Gonzalez-Mánteiga, & J. C. Reboredo-Nogueira, (2006) A partially linear regression-based test for the forward premium hypothesis. Mimeo, Universidad de Santiago de Compostela.

Bachmeier, L. & Q. Li (2002) Is the term structure nonlinear? A semiparametric investigation. *Applied Economics Letters* 9, 151-153.

Bai, J. (1996) Testing for parameter constancy in linear regressions: an empirical distribution function approach. *Econometrica* 64, 597-622.

Billingsley, P. (1999) *Convergence of Probability Measures*. 2nd ed., John Wiley & Sons: New York.

Bosq, D. (1996) *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*. Springer: New York.

Brown, R. L., J. Durbin, & J.M. Evans (1975) Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society B* 37, 149-163.

Cavaliere, G. & A.M.R. Taylor (2006) Testing for a change in persistence in the presence of a volatility shift. *Oxford Bulletin of Economics and Statistics* 68 (supplement), 761-781.

Chow, G. C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28, 591-605.

Delgado, M. A. & I. Fiteni (2002) External bootstrap tests for parameter stability. *Journal of Econometrics* 109, 275-303.

Delgado, M. A. & J. Hidalgo (2000) Nonparametric inference on structural breaks. *Journal of Econometrics* 96, 113-144.

Engle, R.F., C.W.J. Granger, J. Rice, & A. Weiss (1986) Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association* 81, 310-320.

Fan, Y. & Q. Li (1999) Root-n consistent estimation of partially linear time series models. *Journal of Nonparametric Statistics* 11, 251-269.

Gaul, J. & E. Theissen (2006) A partially linear approach to modelling the dynamics of spot and futures prices. Mimeo. Bonn Graduate School of Economics, University of Bonn.

Giné, E. & J. Zinn (1990) Bootstrapping general empirical measures. *Annals of Probability* 18, 851-869.

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

Hansen, B. E. (2000b) Sample splitting and threshold estimation. *Econometrica* 68, 575-603.

Hansen, B. E. (2008) Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory* 24, 726-748.

Herrndorf, N. (1985) A functional central limit theorem for strongly mixing sequences of random variables. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 69, 541-550.

Härdle, W., H. Liang, & J. Gao (2000) *Partially Linear Models.* Physica-Verlag, Germany.

Juhl, T. & Z. Xiao, (2005a) Partially linear models with unit root. *Econometric Theory* 21, 877-906.

Juhl, T. & Z. Xiao (2005b) Testing for cointegration using partially linear models. *Journal of Econometrics* 124, 363-394.

Krämer, W., W. Ploberger, & R. Alt (1988) Testing for structural change in dynamic models. *Econometrica* 56, 1355-1369.

Kuan, C. M. & K. Hornik (1995) The generalized fluctuation test: a unifying view. *Econometric Reviews* 14, 135-161.

Lee, S. & S. Park (2001) The CUSUM of squares test for scale change in infinite order moving average processes. *Scandinavian Journal of Statistics* 28, 625-644.

Lee, Y. (2003) Effects of introducing five-day work week in Korean labor market: a semiparametric vector error correction approach. Mimeo, Dept. of Economics, Yale University.

Li, Q. (1999) Consistent model specification tests for time series econometric models. *Journal of Econometrics* 92, 101-147.

Li, Q. & J. M. Wooldridge, (2002) Semiparametric estimation of partially linear models for dependent data with generated regressors. *Econometric Theory* 18, 625-645.

Linton, O.B. (1995) Second order approximation in the partially linear regression model. *Econometrica* 63, 1079-1113.

Page, E. S. (1955) A test for change in a parameter occurring at an unknown point. *Biometrika* 42, 523-527.

Ploberger, W. & W. Krämer (1992). The CUSUM test with OLS residuals. *Econometrica* 56, 1355-1369.

Pollard, D. (1984) *Convergence of Stochastic Processes.* Springer-Verlag, New York.

Robinson, P. M. (1988) Root-n-consistent semiparametric regression. *Econometrica* 56, 931-954.

Robinson, P. M. (1991) Consistent nonparametric entropy-based testing. *Review of Economic Studies* 58, 437-453.

Su, L. & Z. Xiao (2008) Testing structural change in time-series nonparametric regression models. *Statistics and Its Interface* 1, 347-366.

Sun, S. & C-Y. Chiang (1997) Limiting behavior of the perturbed empirical distribution functions evaluated at U-statistics for strongly mixing sequences of random variables. *Journal of Applied Mathematics and Stochastic Analysis* 10, 3-20.

Wu, G. & Z. Xiao (2002) A generalized partially linear model of asymmetric volatility. *Journal of Empirical Finance* 9, 287-319.

Xu, W-L. (2006) A note on the optimality of generalized cross-validation bandwidth selection in partially linear models with kernel smoothing estimator. *Acta. Mathematicae Applicatae Sinica* 22, 345-352.

Yoshihara, K. (1976) Limiting behavior of U-statistics for stationary, absolutely regular processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 35, 237-252.

Appendix

A Proof of the Main Results in Sections 3-5

We use C to signify a generic constant whose exact value may vary from case to case. For any random sequence $\{w_i\}$ and function $\phi(w_j, w_i)$, let $E_j[\phi(w_j, w_i)]$ denote expectation with respect to w_j only, e.g., $E_j K_{hij} = \int h^{-d} K((z_{ni} - z)/h) f_{nj}(z) dz$. Let $f_{ni} = f_{ni}(z_{ni})$, $m_{ni} = m_{ni}(z_{ni})$, $g_{ni} = g_{ni}(z_{ni})$, $\overline{f}_{\lceil n\pi \rfloor}(z) = n^{-1}\sum_{j=1}^{\lceil n\pi \rfloor} f_{nj}(z)$, $\overline{f}_n(z) = n^{-1}\sum_{j=1}^{n} f_{nj}(z)$, and $\overline{f}_{ni} = \overline{f}_n(z_{ni})$. Let $f = (f_{n1}, \cdots, f_{nn})'$, $M = (m_{n1}, \cdots, m_{nn})'$, $G = (g'_{n1}, \cdots, g'_{nn})'$, $\varepsilon = (\varepsilon_{n1}, \cdots, \varepsilon_{nn})'$, and $V = (v'_{n1}, \cdots v'_{nn})'$. Similarly, let $\widehat{f} = (\widehat{f}_{n1}, \cdots, \widehat{f}_{nn})'$, and for $\xi = \varepsilon$, V, M, or G, define $\widehat{\xi} = (\widehat{\xi}'_{n1}, \cdots, \widehat{\xi}'_{nn})'$ with $\widehat{\xi}_{ni} \equiv n^{-1}\sum_{j\neq i}^{n} \xi_{nj} K_{hij}/\widehat{f}_{ni}$. We write $A_n \simeq B_n$ to signify that $A_n = B_n(1 + o_p(1))$ as $n \to \infty$. Denote $\nu_n \equiv n^{-1/2}h^{-q/2}\sqrt{\log n} + h^r$.

Proof of Theorem 3.1

Under $H_{0a}: \gamma_{ni} = \gamma_0$, we can write

$$\widehat{\gamma} - \gamma_0 = \widehat{\Phi}^{-1}\left\{S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}} + S_{(X-\widehat{X})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}\right\}. \tag{A.1}$$

We first study the asymptotic behavior of $\widehat{\Phi}$, $S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}}$, and $S_{(X-\widehat{X})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}$. Then we discuss what occurs if we have fixed breaks in both m and g, in either m or g, or in neither. Note that Lemma B.1(i) holds whether we have fixed breaks in m and g or not.

Step 1. We study $\widehat{\Phi}$, $S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}}$, and $S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}}$.

Step 1(i). We show $\widehat{\Phi} \xrightarrow{p} \Phi(1) + \Phi_{gg}$. Write $\widehat{\Phi} = n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2(v_{ni} + g_{ni} - \widehat{x}_{ni})(v_{ni} + g_{ni} - \widehat{x}_{ni})'$ $= \Phi_{n1} + \Phi_{n2} + \Phi_{n3} + \Phi'_{n3}$, where $\Phi_{n1} = n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2 v_{ni}v'_{ni}$, $\Phi_{n2} = n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2(g_{ni} - \widehat{x}_{ni})(g_{ni} - \widehat{x}_{ni})'$, and $\Phi_{n3} = n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2 v_{ni}(g_{ni} - \widehat{x}_{ni})'$. By Lemma B.1(i) and Assumption 1(viii), $\Phi_{n1} =$

$n^{-1}\sum_{i=1}^{n}\overline{f}_n^2(z_{ni})\,v_{ni}v_{ni}'+o_p(1)\overset{p}{\to}\Phi(1)$. By Assumptions A1(i), (iv)-(v), (vii*) and A2-A3, and the repeated use of Lemmas B.1(i), C.1-C.2, and the Chebyshev inequality,

$$\begin{aligned}\Phi_{n2} &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}(x_{nj}-g_{ni})(x_{nk}-g_{ni})\\ &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}(g_{nj}-g_{ni})(g_{nk}-g_{ni})+o_p(1)\\ &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}f_{nj}(z_{ni})\,f_{nk}(z_{ni})\left[g_{nj}(z_{ni})-g_{ni}\right]\left[g_{nk}(z_{ni})-g_{ni}\right]+o_p(1)\\ &= n^{-1}\sum_{i=1}^{n}\left[\overline{A}_{gn}(z_{ni})-\overline{A}_{1n}(z_{ni})\,g_{ni}\right]\left[\overline{A}_{gn}(z_{ni})-\overline{A}_{1n}(z_{ni})\,g_{ni}\right]'+o_p(1)\overset{p}{\to}\Phi_{gg}.\end{aligned}$$

Clearly, if no fixed breaks are present in $g(\cdot)$, then $\Phi_{gg}=0$. It is straightfoward to show that $\Phi_{n3}=o_p(1)$. Hence

$$\widehat{\Phi}\overset{p}{\to}\Phi(1)+\Phi_{gg}. \tag{A.2}$$

Step 1(ii). We analyze $S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}}$. Noting that $x_{nt}=g_{nt}+v_{nt}$, we can write

$$S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}}=S_{(G-\widehat{G})\widehat{f},(M-\widehat{M})\widehat{f}}+S_{(V-\widehat{V})\widehat{f},(M-\widehat{M})\widehat{f}}\equiv S_{n1}+S_{n2},\ \text{say.} \tag{A.3}$$

For S_{n1}, we have

$$\begin{aligned}S_{n1} &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nj}(z_{nj})-m_{nj}(z_{ni})\right]\left[g_{nk}(z_{nk})-g_{nk}(z_{ni})\right]\\ &\quad+n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nj}(z_{nj})-m_{nj}(z_{ni})\right]\left[g_{nk}(z_{ni})-g_{ni}(z_{ni})\right]\\ &\quad+n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nj}(z_{ni})-m_{ni}(z_{ni})\right]\left[g_{nk}(z_{nk})-g_{nk}(z_{ni})\right]\\ &\quad+n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nj}(z_{ni})-m_{ni}(z_{ni})\right]\left[g_{nk}(z_{ni})-g_{ni}(z_{ni})\right]\\ &\equiv S_{n11}+S_{n12}+S_{n13}+S_{n14},\ \text{say.}\end{aligned} \tag{A.4}$$

By using Lemma C.4 repeatedly, it is standard to show that

$$S_{n11}=O_p\left(h^{2r}\right),\ S_{n12}=O_p\left(\alpha_{gn}h^{r}\right),S_{n13}=O_p\left(\alpha_{mn}h^{r}\right),\ \text{and}\ S_{n14}=O_p\left(\alpha_{gn}\alpha_{mn}\right). \tag{A.5}$$

In particular, if we allow fixed breaks in both m and g so that $\alpha_{gn}=\alpha_{mn}=O(1)$, then we have

$$S_{n14}=n^{-1}\sum_{i=1}^{n}\left[\overline{A}_{gn}(z_{ni})-\overline{A}_{1n}(z_{ni})\,g_{ni}\right]\left[\overline{A}_{mn}(z_{ni})-\overline{A}_{1n}(z_{ni})\,m_{ni}\right]+o_p(1)\overset{p}{\to}\Phi_{gm}. \tag{A.6}$$

For S_{n2}, write

$$\begin{aligned} S_{n2} &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left(v_{nj}-v_{ni}\right)\left[m_{nk}\left(z_{nk}\right)-m_{nk}\left(z_{ni}\right)\right] \\ &+n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left(v_{nj}-v_{ni}\right)\left[m_{nk}\left(z_{ni}\right)-m_{ni}\left(z_{ni}\right)\right]\equiv S_{n21}+S_{n22}. \end{aligned} \tag{A.7}$$

By the repeated use of Lemmas C.1-C.2 and the Chebyshev inequality, we can show

$$S_{n21}=o_p(n^{-1/2}) \text{ and } S_{n22}=O_p(\alpha_{mn}n^{-1/2}). \tag{A.8}$$

It follows from (A.3)-(A.8) that

$$\begin{aligned} S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}} &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nj}\left(z_{ni}\right)-m_{ni}\left(z_{ni}\right)\right]\left[g_{nk}\left(z_{ni}\right)-g_{ni}\left(z_{ni}\right)\right] \\ &+n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nk}\left(z_{ni}\right)-m_{ni}\left(z_{ni}\right)\right]\left(v_{nj}-v_{ni}\right) \\ &+O_p\left(h^{2r}+\alpha_{gn}h^{r}+\alpha_{mn}h^{r}\right)+o_p(n^{-1/2}). \end{aligned} \tag{A.9}$$

Step 1(iii). We analyze $S_{(X-\widehat{X})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}$. Write

$$S_{(X-\widehat{X})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}=S_{(G-\widehat{G})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}+S_{(V-\widehat{V})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}\equiv S_{n3}+S_{n4},\ \text{say.} \tag{A.10}$$

For S_{n3}, we have

$$\begin{aligned} S_{n3} &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[g_{nj}\left(z_{nj}\right)-g_{nj}\left(z_{ni}\right)\right]\left(\varepsilon_{nk}-\varepsilon_{ni}\right) \\ &+n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[g_{nj}\left(z_{ni}\right)-g_{ni}\left(z_{ni}\right)\right]\left(\varepsilon_{nk}-\varepsilon_{ni}\right)\equiv S_{n31}+S_{n32}. \end{aligned} \tag{A.11}$$

It is standard to show that

$$S_{n31}=o_p(n^{-1/2}), \text{ and } S_{n32}=O_p(\alpha_{gn}n^{-1/2}). \tag{A.12}$$

Similarly, write

$$\begin{aligned} S_{n4} &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left(v_{nj}-v_{ni}\right)\left(\varepsilon_{nk}-\varepsilon_{ni}\right) \\ &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}v_{nj}\varepsilon_{nk}-n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}v_{nj}\varepsilon_{ni} \\ &\quad -n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}v_{ni}\varepsilon_{nk}+n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}v_{ni}\varepsilon_{ni} \\ &\equiv S_{n41}-S_{n42}-S_{n43}+S_{n44},\ \text{say.} \end{aligned} \tag{A.13}$$

It is standard to show that

$$S_{n4j} = O_p\left(n^{-1}h^{-q}\right) = o_p(n^{-1/2}),\ j = 1,2,3,\ \text{and}\ S_{n44} = n^{-1}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + o_p(n^{-1/2}). \tag{A.14}$$

It follows from (A.10)-(A.14) that

$$\begin{aligned} S_{(X-\widehat{X})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}} &= n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[g_{nj}\left(z_{ni}\right)-g_{ni}\left(z_{ni}\right)\right]\left(\varepsilon_{nk}-\varepsilon_{ni}\right) \\ &\quad + n^{-1}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + o_p(n^{-1/2}). \end{aligned} \tag{A.15}$$

Step 2. We discuss the various cases. Combining (A.1), (A.2), (A.9), and (A.15) yields

$$\begin{aligned} \widehat{\gamma}-\gamma_0 &= \left(\Phi\left(1\right)+\Phi_{gg}\right)^{-1}\left(1+o_p\left(1\right)\right) \\ &\quad \times\Bigg\{n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nj}\left(z_{ni}\right)-m_{ni}\left(z_{ni}\right)\right]\left[g_{nk}\left(z_{ni}\right)-g_{ni}\left(z_{ni}\right)\right] \\ &\quad + n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nk}\left(z_{ni}\right)-m_{ni}\left(z_{ni}\right)\right]\left(v_{nj}-v_{ni}\right) \\ &\quad + n^{-3}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[g_{nj}\left(z_{ni}\right)-g_{ni}\left(z_{ni}\right)\right]\left(\varepsilon_{nk}-\varepsilon_{ni}\right) \\ &\quad + n^{-1}\sum_{i=1}^{n}\overline{f}_{n}^{2}\left(z_{ni}\right)v_{ni}\varepsilon_{ni}\Bigg\} + O_p\left(h^{2r}+\alpha_{gn}h^{r}+\alpha_{mn}h^{r}\right)+o_p(n^{-1/2}). \end{aligned} \tag{A.16}$$

If both m and g have fixed breaks so that one can take $\alpha_{mn} = \alpha_{gn} = 1$, then the first term inside the curly brackets in (A.16) dominates, and by (A.6) we have $\widehat{\gamma}-\gamma_0 = \left(\Phi\left(1\right)+\Phi_{gg}\right)^{-1}\Phi_{gm} + o_p\left(1\right).$

If only m has fixed breaks, so that $\alpha_{mn} = 1$ and $\alpha_{gn} = o\left(1\right)$, then (A.16) in conjunction with (A.5) implies $\sqrt{n}\left(\widehat{\gamma}-\gamma_0\right) = \Phi\left(1\right)^{-1}\Big\{n^{-5/2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[m_{nk}\left(z_{ni}\right)-m_{ni}\left(z_{ni}\right)\right] \times\left(v_{nj}-v_{ni}\right)+n^{-1/2}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni}\Big\} + O_p\left(\sqrt{n}h^{r}\right)+o_p\left(1\right).$ If only g has fixed breaks, so that $\alpha_{gn} = 1$ and $\alpha_{mn} = o\left(1\right)$, then (A.16) in conjunction with (A.7) implies $\sqrt{n}\left(\widehat{\gamma}-\gamma_0\right) = \left(\Phi\left(1\right)+\Phi_{gg}\right)^{-1}\{n^{-5/2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\sum_{k\neq i}^{n}K_{hij}K_{hik}\left[g_{nj}\left(z_{ni}\right)-g_{ni}\left(z_{ni}\right)\right]\left(\varepsilon_{nk}-\varepsilon_{ni}\right)+n^{-1/2}\sum_{i=1}^{n} \overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni}\}+O_p\left(\sqrt{n}h^{r}\right)+o_p\left(1\right).$ If neither m nor g has a fixed break, so that $\alpha_{mn} = o\left(1\right)$ and $\alpha_{gn} = o\left(1\right)$, then combining (A.5), (A.8), and (A.16) yields $\sqrt{n}\left(\widehat{\gamma}-\gamma_0\right) = n^{-1/2}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + O_p\left(\sqrt{n}(\alpha_{mn}h^{r}+\alpha_{gn}h^{r}+\alpha_{mn}\alpha_{gn})\right)+o_p\left(1\right).$ The conclusion then follows under the given extra condition. ■

Proof of Theorems 3.2 and 3.3

The proof of Theorem 3.2 is a special case of that of Theorem 3.3, so we only prove Theorem 3.3. Noting that $\gamma_{ni} = \gamma_0 + n^{-1/2}\delta_1(i/n)$ under $H_{1a,n}$, we have

$$\begin{aligned}
&\sqrt{n}(\widehat{\gamma} - \gamma_0) \\
&= \widehat{\Phi}^{-1}\sqrt{n}\left(S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}} + S_{(X-\widehat{X})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}\right) \\
&+ \widehat{\Phi}^{-1}\left\{n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})x_{ni}'\delta_1(i/n) - n^{-2}\sum_{i=1}^{n}\widehat{f}_{ni}(x_{ni} - \widehat{x}_{ni})\sum_{j\neq i}^{n}K_{hij}x_{nj}'\delta_1(j/n)\right\}.
\end{aligned} \tag{A.17}$$

Under case (d) in Theorem 3.1, we have shown that $\sqrt{n}\left(S_{(X-\widehat{X})\widehat{f},(M-\widehat{M})\widehat{f}} + S_{(X-\widehat{X})\widehat{f},(\varepsilon-\widehat{\varepsilon})\widehat{f}}\right)$ $= n^{-1/2}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + o_p(1)$. By Lemmas B.1(i)-(ii), it is straightforward to show that $n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})x_{ni}'\delta_1(i/n) = n^{-1}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}x_{ni}'\delta_1(i/n) + o_p(1)$ and that the last term inside the curly brackets in (A.17) is $o_p(1)$. In addition, $\widehat{\Phi} = \Phi(1) + o_p(1)$. It follows that

$$\sqrt{n}(\widehat{\gamma} - \gamma_0) = \Phi(1)^{-1}\left\{n^{-1/2}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + n^{-1}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}x_{ni}'\delta_1(i/n)\right\} + o_p(1). \tag{A.18}$$

By (2.1) and (2.11), $\widetilde{u}_{ni} = \varepsilon_{ni} - x_{ni}'(\widehat{\gamma} - \gamma_0) - [\widetilde{m}(z_{ni}) - m_{ni}(z_{ni})] + n^{-1/2}x_{ni}'\delta_1(i/n)$ under $H_{1a,n}$. It follows from (2.13) that

$$\begin{aligned}
&\widehat{\Psi}^{1/2}\Gamma_{na}(\pi) \\
= \;& n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})\varepsilon_{ni} - n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})x_{ni}'(\widehat{\gamma} - \gamma_0) \\
& - n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})[\widetilde{m}(z_{ni}) - m_{ni}(z_{ni})] + n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})x_{ni}'\delta_1(i/n) \\
\equiv \;& A_{n1}(\pi) - A_{n2}(\pi) - A_{n3}(\pi) + A_{n4}(\pi), \text{ say.}
\end{aligned} \tag{A.19}$$

We analyze each of the four terms in the last expression in separate steps.

Step 1. We show that $A_{n1}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + o_p(1)$ uniformly in $\pi \in [0,1]$.

Write

$$A_{n1}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi \rfloor}[v_{ni} - (\widehat{x}_{ni} - g_{ni}(z_{ni}))]\,\widehat{f}_{ni}^{2}\varepsilon_{ni}$$
$$= n^{-1/2}\sum_{i=1}^{\lceil n\pi \rfloor}\Big\{\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + 2\left(\widehat{f}_{ni} - \overline{f}_{ni}\right)\overline{f}_{ni}v_{ni}\varepsilon_{ni} + \left(\widehat{f}_{ni} - \overline{f}_{ni}\right)^{2}v_{ni}\varepsilon_{ni}$$
$$-[\widehat{x}_{ni} - g_{ni}(z_{ni})]\widehat{f}_{ni}\overline{f}_{ni}\varepsilon_{ni} - [\widehat{x}_{ni} - g_{ni}(z_{ni})]\widehat{f}_{ni}\left(\widehat{f}_{ni} - \overline{f}_{ni}\right)\varepsilon_{ni}\Big\}$$
$$\equiv A_{n11}(\pi) + 2A_{n12}(\pi) + A_{n13}(\pi) - A_{n14}(\pi) - A_{n15}(\pi), \text{ say.}$$

It suffices to show that $\sup_{0\leq\pi\leq1}|A_{n1l}(\pi)| = o_p(1)$, $l = 2,3,4,5$. First, write $A_{n12}(\pi) = n^{-3/2}\sum_{i=1}^{\lceil n\pi \rfloor}\sum_{j\neq i}^{n}(K_{hij} - E_jK_{hij})\overline{f}_{ni}v_{ni}\varepsilon_{ni} + n^{-3/2}\sum_{i=1}^{\lceil n\pi \rfloor}\sum_{j\neq i}^{n}[E_jK_{hij} - f_{nj}(z_{ni})]\overline{f}_{ni}v_{ni}\varepsilon_{ni} + n^{-3/2}\sum_{i=1}^{\lceil n\pi \rfloor}\sum_{j\neq i}^{n}[f_{nj}(z_{ni}) - \overline{f}_{ni}]\overline{f}_{ni}v_{ni}\varepsilon_{ni} \equiv A_{n12a}(\pi) + A_{n12b}(\pi) + A_{n12c}(\pi)$, say. By Lemma B.3(i), $\sup_{0\leq\pi\leq1}\|A_{12a}(\pi)\| = o_p(1)$. By the same arguments as in the analysis of $A_{n22}(\pi)$ in the proof of Lemma B.3(ii), we can show that $\sup_{0\leq\pi\leq1}|A_{n12b}(\pi)| = o(1)$. Noting that $\sum_{j\neq i}^{n}(f_{nj}(z_{ni}) - \overline{f}_{ni}) = \overline{f}_{ni} - f_{ni}$, it is straightforward to show that $\sup_{0\leq\pi\leq1}|A_{n12c}(\pi)| = \sup_{0\leq\pi\leq1}|n^{-3/2}\sum_{i=1}^{\lceil n\pi \rfloor}(\overline{f}_{ni} - f_{ni})\overline{f}_{ni}v_{ni}\varepsilon_{ni}| = o_p(1)$. Hence $\sup_{0\leq\pi\leq1}|A_{n12}(\pi)| = o_p(1)$. Next, by Lemma B.1(i) and Assumptions A1 and A3, $\sup_{0\leq\pi\leq1}\|A_{13}(\pi)\| \leq n^{1/2}\sup_{1\leq i\leq n}|\widehat{f}_{ni} - \overline{f}_{ni}|^{2}$ $n^{-1}\sum_{i=1}^{n}\|v_{ni}\varepsilon_{ni}\| = O_p(n^{1/2}\nu_n^2) = o_p(1)$. By Lemma B.3(ii)-(iii), $A_{14}(\pi) = n^{-3/2}\sum_{i=1}^{\lceil n\pi \rfloor}\sum_{j\neq i}^{n}$ $K_{hij}[g_{nj}(z_{nj}) - g_{ni}(z_{ni})]\overline{f}_{ni}\varepsilon_{ni} + n^{-3/2}\sum_{i=1}^{\lceil n\pi \rfloor}\sum_{j\neq i}^{n}K_{hij}v_{nj}\overline{f}_{ni}\varepsilon_{ni} = o_p(1) + o_p(1) = o_p(1)$ uniformly in π. Now, by Lemma B.1(i)-(ii) and Assumptions A1 and A3, $\sup_{0\leq\pi\leq1}\|A_{15}(\pi)\| \leq n^{1/2}\max_{1\leq j\leq n}||(\widehat{x}_{nj} - g_{nj}(z_{nj}))\widehat{f}_{nj}||\max_{1\leq k\leq n}|\widehat{f}_{nk} - \overline{f}_{nk}|n^{-1}\sum_{i=1}^{n}|\varepsilon_{ni}| = O_p(n^{1/2}(\nu_n + \alpha_{gn})\nu_n)$ $= o_p(1)$.

Step 2. We show that $A_{n2}(\pi) = \Phi(\pi)\Phi(1)^{-1}(n^{-1/2}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + n^{-1}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}x_{ni}'$ $\delta_1(i/n)) + o_p(1)$ uniformly in $\pi\in[0,1]$. By Lemma B.2 and eq. (A.18) we can write

$$A_{n2}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})(x_{ni} - \widehat{x}_{ni})'\sqrt{n}(\widehat{\gamma} - \gamma_0) + n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})\widehat{x}_{ni}'\sqrt{n}(\widehat{\gamma} - \gamma_0)$$
$$= \Phi(\pi)\Phi(1)^{-1}\left\{n^{-1/2}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} + n^{-1}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}x_{ni}'\delta_1(i/n)\right\}$$
$$+ \overline{A}_{n2}(\pi)\sqrt{n}(\widehat{\gamma} - \gamma_0) + o_p(1),$$

where $\overline{A}_{n2}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\widehat{f}_{ni}^{2}(x_{ni} - \widehat{x}_{ni})\widehat{x}_{ni}'$. It suffices to show $\sup_{0\leq\pi\leq1}\left|\overline{A}_{n2}(\pi)\right| = o_p(1)$, as $\sqrt{n}\left(\widehat{\gamma}-\gamma_0\right)=O_p\left(1\right)$. Write

$$\begin{aligned}
\overline{A}_{n2}\left(\pi\right) &= n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}\left(x_{ni}-\widehat{x}_{ni}\right)g_{ni}\left(z_{ni}\right)\overline{f}_{ni}+n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}\left(x_{ni}-\widehat{x}_{ni}\right)\left[\widehat{f}_{ni}\widehat{x}_{ni}-g_{ni}\left(z_{ni}\right)\overline{f}_{ni}\right] \\
&= n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\left\{\widehat{f}_{ni}\left[g_{ni}\left(z_{ni}\right)-\widehat{x}_{ni}\right]g_{ni}\left(z_{ni}\right)\overline{f}_{ni}+\left(\widehat{f}_{ni}-\overline{f}_{ni}\right)v_{ni}g_{ni}\left(z_{ni}\right)\overline{f}_{ni}+\overline{f}_{ni}^{2}v_{ni}g_{ni}\left(z_{ni}\right)\right. \\
&\quad \left.+\widehat{f}_{ni}\left(x_{ni}-\widehat{x}_{ni}\right)\left[\widehat{f}_{ni}\widehat{x}_{ni}-g_{ni}\left(z_{ni}\right)\overline{f}_{ni}\right]\right\} \\
&\equiv A_{n21}\left(\pi\right)+A_{n22}\left(\pi\right)+A_{n23}\left(\pi\right)+A_{n24}\left(\pi\right),\ \text{say.}
\end{aligned}$$

By Lemma B.1(i)-(ii), it is straightforward to show that $\sup_{0\leq\pi\leq1}\left\|A_{n21}\left(\pi\right)\right\|=O_p(\nu_n+\alpha_{gn})$, $\sup_{0\leq\pi\leq1}\left\|A_{n22}\left(\pi\right)\right\|=O_p(\nu_n)$, and $\sup_{0\leq\pi\leq1}\left\|A_{n24}\left(\pi\right)\right\|=O_p(\nu_n+\alpha_{gn})$. $\sup_{0\leq\pi\leq1}\left\|A_{n23}\left(\pi\right)\right\|=O_p\left(n^{-1/2}\right)$ by the invariance principle for (heterogeneous) strong mixing processes (e.g., Herrndorf (1985)). It follows that $\sup_{0\leq\pi\leq1}\left\|\overline{A}_{n2}\left(\pi\right)\right\|=o_p\left(1\right)$.

Step 3. We show that $A_{n3}\left(\pi\right)=o_p\left(1\right)$ uniformly in $\pi\in\left[0,1\right]$. Write $A_{n3}\left(\pi\right)=n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^{2}[g_{ni}-\widehat{x}_{ni}][\widetilde{m}\left(z_{ni}\right)-m_{ni}]+n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}(\widehat{f}_{ni}-\overline{f}_{ni})v_{ni}[\widetilde{m}\left(z_{ni}\right)-m_{ni}]+n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}\overline{f}_{ni}v_{ni}[\widetilde{m}\left(z_{ni}\right)-m_{ni}]\equiv A_{n31}\left(\pi\right)+A_{n32}\left(\pi\right)+A_{n33}\left(\pi\right)$, say. It suffices to show that each of these terms is $o_p\left(1\right)$. First, by Lemma B.1(ii)-(iii) and Assumptions A1(vii) and A3, $\sup_{0\leq\pi\leq1}\left\|A_{n31}\left(\pi\right)\right\|=O_p(n^{1/2}(\nu_n+\alpha_{gn})\left(\nu_n+\alpha_{mn}\right))=o_p\left(1\right)$. Similarly, $\sup_{0\leq\pi\leq1}\left\|A_{n32}\left(\pi\right)\right\|=o_p\left(1\right)$. By (2.1) and (2.10),

$$\widetilde{m}\left(z_{ni}\right)\widehat{f}_{ni}=n^{-1}\sum_{j\neq i}^{n}K_{hij}[m_{nj}\left(z_{nj}\right)+\varepsilon_{nj}-x_{nj}'(\widehat{\gamma}-\gamma_{nj})].$$

Under $H_{1a,n}:\gamma_{nj}=\gamma_0+n^{-1/2}\delta_1\left(j/n\right)$, we have

$$\begin{aligned}
A_{n33}\left(\pi\right) &= n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}\overline{f}_{ni}v_{ni}\left[m_{nj}-m_{ni}\right]+n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}\overline{f}_{ni}v_{ni}\varepsilon_{nj} \\
&\quad -n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}\overline{f}_{ni}v_{ni}x_{nj}'\left(\widehat{\gamma}-\gamma_0\right)+n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}\widehat{f}_{ni}\overline{f}_{ni}v_{ni}x_{nj}'\delta_1\left(j/n\right) \\
&\equiv A_{n33a}\left(\pi\right)+A_{n33b}\left(\pi\right)-A_{n33c}\left(\pi\right)+A_{n33d}\left(\pi\right),\ \text{say.}
\end{aligned}$$

By Lemmas B.3(iv)-(v), $\sup_{0\leq\pi\leq1}\left\|A_{n33a}\left(\pi\right)\right\|=o_p\left(1\right)$ and $\sup_{0\leq\pi\leq1}\left\|A_{n33b}\left(\pi\right)\right\|=o_p\left(1\right)$. For $A_{n33c}\left(\pi\right)$, it is easy to show that each element of the $p\times1$ vector $n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}\overline{f}_{ni}v_{ni}x_{nj}$ is $O_p\left(1\right)$, implying that $\sup_{0\leq\pi\leq1}\left\|A_{n33c}\left(\pi\right)\right\|=O_p(\|\widehat{\gamma}-\gamma_0\|)=o_p\left(1\right)$. It is straightforward to show $\sup_{0\leq\pi\leq1}\left\|A_{n33d}\left(\pi\right)\right\|=o_p\left(1\right)$. Hence $\sup_{0\leq\pi\leq1}\left\|A_{n33}\left(\pi\right)\right\|=o_p\left(1\right)$.

Step 4. We show that $A_{n4}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}^{2}v_{ni}x_{ni}'\delta_1(i/n) + o_p(1)$ uniformly in $\pi \in [0,1]$. Write $A_{n4}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}^{2}v_{ni}x_{ni}'\delta_1(i/n) + n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^{2}(g_{ni}(z_{ni}) - \widehat{x}_{ni})x_{ni}'\delta_1(i/n) + n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}(\widehat{f}_{ni}^{2} - \overline{f}_{ni}^{2})v_{ni}x_{ni}'\delta_1(i/n)$. By Lemma B.1, one can show that the last two terms are $o_p(1)$ uniformly in π. The result follows.

Combining (A.19) with the results in Steps 1-4 yields

$$\begin{aligned}\widehat{\Psi}^{1/2}\Gamma_{na}(\pi) &= \left\{n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni} - \Phi(\pi)\Phi(1)^{-1}n^{-1/2}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}\varepsilon_{ni}\right\} \\ &\quad + \left\{n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}^{2}v_{ni}x_{ni}'\delta_1(i/n) - \Phi(\pi)\Phi(1)^{-1}n^{-1}\sum_{i=1}^{n}\overline{f}_{ni}^{2}v_{ni}x_{ni}'\delta_1(i/n)\right\} + o_p(1) \\ &\equiv a_{n0}(\pi) + a_{n1}(\pi) + o_p(1) \text{ uniformly in } \pi.\end{aligned}$$

By Assumption A1(viii), $a_{n0}(\cdot) \Rightarrow N(\cdot) - \Phi(\cdot)\Phi(1)^{-1}N(1)$. By extending Lemma 4 of Krämer, Ploberger and Alt (1988) (see also Bai, 1996), we can show that $a_{n1}(\pi) \xrightarrow{p} \int_0^{\pi}\Phi^{(1)}(s)\delta_1(s)\,ds - \Phi(\pi)\Phi(1)^{-1}\int_0^1\Phi^{(1)}(s)\delta_1(s)\,ds$, where $\Phi^{(1)}(s) = (\partial/\partial s)\Phi(s)$. Under $H_{1a,n}$, we can similarly show that $\widehat{\Psi} = \Psi(1) + o_p(1)$. Consequently, $\Gamma_{na}(\cdot) \Rightarrow \Gamma_a(\cdot) + \Delta_a(\cdot)$ as desired. ■

Proof of Theorems 4.1 and 4.2

By (2.1) and (2.11), $\widetilde{u}_{ni} = \varepsilon_{ni} - x_{ni}'(\widehat{\gamma} - \gamma_{ni}) - [\widetilde{m}(z_{ni}) - m_{ni}(z_{ni})]$. By (2.1) and (2.10), $\widetilde{m}(z_{ni}) = n^{-1}\widehat{f}_{ni}^{-1}\sum_{j\neq i}^{n}K_{hij}[\varepsilon_{nj} - x_{nj}'(\widehat{\gamma} - \gamma_{nj}) + m_{nj}(z_{nj})]$. Under $H_{1b,n}$, we have $\widetilde{u}_{ni} = (\varepsilon_{ni} - \widehat{\varepsilon}_{ni}) - (x_{ni} - \widehat{x}_{ni})'(\widehat{\gamma} - \gamma_0) - \{\widehat{m}_0(z_{ni}) - m_0(z_{ni})\} + \{n^{-1/2}x_{ni}'\delta_1(i/n) - n^{-3/2}\widehat{f}_{ni}^{-1}\sum_{j\neq i}^{n}K_{hij} \times x_{nj}'\delta_1(j/n)\} + \{n^{-1/2}\delta_2(z_{ni}, i/n) - n^{-3/2}\widehat{f}_{ni}^{-1}\sum_{j\neq i}^{n}K_{hij}\delta_2(z_{nj}, j/n)\}$, where $\widehat{m}_0(z_{ni}) = n^{-1}\widehat{f}_{ni}^{-1}\sum_{j\neq i}^{n}K_{hij}m_0(z_{nj})$. It follows that under $H_{1b,n}$,

$$\widehat{\sigma}\Gamma_{nb}(\pi) = B_{n1}(\pi) - B_{n2}(\pi) - B_{n3}(\pi) + B_{n4}(\pi) + B_{n5}(\pi), \tag{A.20}$$

where $B_{n1}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}[\varepsilon_{ni} - \widehat{\varepsilon}_{ni}]$, $B_{n2}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}[x_{ni} - \widehat{x}_{ni}]'[\widehat{\gamma} - \gamma_0]$, $B_{n3}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}[\widehat{m}_0(z_{ni}) - m_0(z_{ni})]$, $B_{n4}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}x_{ni}'\delta_1(i/n) - n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}x_{nj}'\delta_1(j/n)$, and $B_{n5}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}\delta_2(z_{ni}, i/n) - n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}\delta_2(z_{nj}, j/n)$. Note that under H_{0b}, $B_{n4}(\pi)$ and $B_{n5}(\pi)$ vanish in (A.20). The proof of Theorem 4.1 is thus a special case of that of Theorem 4.2.

First, write $B_{n1}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}\varepsilon_{ni} - n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}\varepsilon_{nj} + n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}(\widehat{f}_{ni} - \overline{f}_{ni})\varepsilon_{ni} \equiv B_{n11}(\pi) - B_{n12}(\pi) + B_{n13}(\pi)$, say. By Assumption A1 and analogously to the proof of Lemma B.3(ii), we can show $B_{n12}(\pi) = n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j=1}^{n} f_{ni}(z_{nj})\varepsilon_{nj} + o_p(1)$. Observe that $B_{n13}(\pi) = n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}[K_{hij} - E_j(K_{hij})]\varepsilon_{ni} + n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}[E_j(K_{hij}) - f_{nj}(z_{ni})]\varepsilon_{ni} + n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}[f_{nj}(z_{ni}) - \overline{f}_{ni}]\varepsilon_{ni} \equiv B_{n13a}(\pi) + B_{n13b}(\pi) + B_{n13c}(\pi)$. By arguments similar to the proof of Lemma B.3(i), $\sup_{0\leq\pi\leq1}|B_{n13a}(\pi)| = o_p(1)$. It is easy to show that $\sup_{0\leq\pi\leq1}|B_{n13b}(\pi)| = O_p(h^r) = o_p(1)$. Noting that $\sum_{j\neq i}^{n}(f_{nj}(z_{ni}) - \overline{f}_{ni}) = \overline{f}_{ni} - f_{ni}$, we have $\sup_{0\leq\pi\leq1}|B_{n13c}(\pi)| = \sup_{0\leq\pi\leq1}|n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}(\overline{f}_{ni} - f_{ni})\varepsilon_{ni}| = O_p(n^{-1})$. Consequently

$$B_{n1}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}\varepsilon_{ni} - n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j=1}^{n} f_{ni}(z_{nj})\varepsilon_{nj} + o_p(1) \text{ uniformly in } \pi. \quad \text{(A.21)}$$

Next, write $B_{n2}(\pi) = n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\{\widehat{f}_{ni}[g_{ni} - \widehat{x}_{ni}]'(\widehat{\gamma} - \gamma_0) + (\widehat{f}_{ni} - \overline{f}_{ni})v'_{ni}(\widehat{\gamma} - \gamma_0) + \overline{f}_{ni}v'_{ni}(\widehat{\gamma} - \gamma_0)\} \equiv B_{n21}(\pi) + B_{n22}(\pi) + B_{n23}(\pi)$, say. By Lemma B.1(i)-(ii) and the fact that $\sqrt{n}(\widehat{\gamma} - \gamma_0) = O_p(1)$ under either H_{0b} or $H_{1b,n}$, $\sup_{0\leq\pi\leq1}|B_{n21}(\pi)| = O_p(\nu_n + \alpha_{gn})$ and $\sup_{0\leq\pi\leq1}|B_{n22}(\pi)| = O_p(\nu_n)$. By the invariance principle for $\{n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}v_{ni}\}$ and the fact that $\sqrt{n}(\widehat{\gamma} - \gamma_0) = O_p(1)$, $\sup_{0\leq\pi\leq1}|B_{n23}(\pi)| = O_p\left(n^{-1/2}\right)$. Hence

$$\sup_{0\leq\pi\leq1}|B_{n2}(\pi)| = O_p(\nu_n + \alpha_{gn}) = o_p(1). \quad \text{(A.22)}$$

Using Lemma C.4 we can show that uniformly in z, $|n^{-1}\sum_{j=1}^{n} K_h(z - z_{nj})(m_0(z_{nj}) - m_0(z))| \simeq |n^{-1}\sum_{j=1}^{n} E[K_h(z - z_{nj})(m_0(z_{nj}) - m_0(z))]| \leq h^r D_{m_0}(z)$. Hence, with probability approaching 1 as $n \to \infty$

$$\sup_{0\leq\pi\leq1}|B_{n3}(\pi)| \leq Cn^{-1/2}h^r\sum_{i=1}^{n} D_{m_0}(z_{ni}) = O_p(n^{1/2}h^r) = o_p(1). \quad \text{(A.23)}$$

Next, by Lemma B.1(i), $n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}x'_{ni}\delta_1(i/n) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}x'_{ni}\delta_1(i/n) + o_p(1)$ uniformly in π. One can also show $n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n} K_{hij}x'_{nj}\delta_1(j/n) = n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j=1}^{n} f_{ni}(z_{nj}) \times x'_{nj}\delta_1(j/n) + o_p(1)$. It follows that

$$B_{n4}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}x'_{ni}\delta_1(i/n) - n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j=1}^{n} f_{ni}(z_{nj})x'_{nj}\delta_1(j/n) + o_p(1). \quad \text{(A.24)}$$

Now write $B_{n5}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}\delta_2(z_{ni}, i/n) - n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n} K_{hij}\delta_2(z_{nj}, j/n) \equiv B_{n51}(\pi) - B_{n52}(\pi)$. By Lemma B.1(i), it is easy to show that $B_{n51}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}\delta_2(z_{ni}, i/n) + o_p(1)$ uniformly in π. One can also show that $B_{n52}(\pi) = n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j=1}^{n} f_{ni}(z_{nj})\delta_2(z_{nj}, j/n) + o_p(1)$ uniformly in π. Hence

$$B_{n5}(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}\delta_2(z_{ni}, i/n) - n^{-2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j=1}^{n} f_{ni}(z_{nj})\delta_2(z_{nj}, j/n) + o_p(1). \quad \text{(A.25)}$$

Combining (A.20)-(A.25) yields

$$\begin{aligned}
\widehat{\sigma}\Gamma_{nb}(\pi) &= \left\{ n^{-1/2} \sum_{i=1}^{\lceil n\pi \rfloor} \overline{f}_{ni}\varepsilon_{ni} - n^{-3/2} \sum_{i=1}^{\lceil n\pi \rfloor} \sum_{j=1}^{n} f_{ni}(z_{nj})\varepsilon_{nj} \right\} \\
&+ \left\{ n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} \overline{f}_{ni} x_{ni}^{\prime} \delta_1(i/n) - n^{-2} \sum_{i=1}^{\lceil n\pi \rfloor} \sum_{j=1}^{n} f_{ni}(z_{nj}) x_{nj}^{\prime} \delta_1(j/n) \right\} \\
&+ \left\{ n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} \overline{f}_{ni} \delta_2(z_{ni}, i/n) - n^{-2} \sum_{i=1}^{\lceil n\pi \rfloor} \sum_{j=1}^{n} f_{ni}(z_{nj}) \delta_2(z_{nj}, j/n) \right\} + o_p(1) \\
&\equiv b_{n0}(\pi) + b_{n1}(\pi) + b_{n2}(\pi) + o_p(1) \text{ uniformly in } \pi,
\end{aligned}$$

where $b_{n1}(\pi)$ and $b_{n2}(\pi)$ obviously vanish under H_{0b}. Clearly, $b_{nl}(\pi) \stackrel{p}{\to} \sigma_0 \Delta_{bl}(\pi)$ uniformly in π for $l = 1, 2$, where $\Delta_{b1}(\cdot)$ and $\Delta_{b2}(\cdot)$ are as defined in Theorem 4.2. Under either H_{0b} or $H_{1b,n}$, it is straightforward to show that $\widehat{\sigma}^2 = \sigma_0^2 + o_p(1)$. It remains to show that

$$b_{n0}(\cdot) \Rightarrow \sigma_0 \Gamma_b(\cdot). \tag{A.26}$$

We prove (A.26) in three steps. First, we show convergence of the sample covariance kernel to the specified covariance kernel. Then we establish the convergence of finite dimensional distributions. Finally, we prove the tightness of $\{b_{n0}(\pi)\}$.

First,

$$\begin{aligned}
& E\left[b_{n0}(\pi_1) b_{n0}(\pi_2)\right] \\
= \; & n^{-1} \sum_{i=1}^{\lceil n(\pi_1 \wedge \pi_2) \rfloor} E(\overline{f}_{ni}^2 \varepsilon_{ni}^2) + n^{-3} \sum_{i=1}^{\lceil n\pi_1 \rfloor} \sum_{j=1}^{n} \sum_{k=1}^{\lceil n\pi_2 \rfloor} E\left[f_{ni}(z_{nj}) f_{nk}(z_{nj}) \varepsilon_{nj}^2\right] \\
& - n^{-2} \sum_{i=1}^{\lceil n\pi_1 \rfloor} \sum_{j=1}^{\lceil n\pi_2 \rfloor} E\left[\overline{f}_{ni} f_{nj}(z_{ni}) \varepsilon_{ni}^2\right] - n^{-2} \sum_{i=1}^{\lceil n\pi_2 \rfloor} \sum_{j=1}^{\lceil n\pi_1 \rfloor} E\left[\overline{f}_{ni} f_{nj}(z_{ni}) \varepsilon_{ni}^2\right] \\
\to \; & S_{11}(\pi_1, \pi_2) + S_{22}(\pi_1, \pi_2) - S_{12}(\pi_1, \pi_2) - S_{21}(\pi_1, \pi_2).
\end{aligned}$$

Next, write $b_{n0}(\pi) = n^{-1/2} \sum_{i=1}^{n} \left[\overline{f}_n(z_{ni}) 1(i \le \lceil n\pi \rfloor) - \overline{f}_{\lceil n\pi \rfloor}(z_{ni})\right] \varepsilon_{ni}$. Fix $k \ge 1$, $\omega \equiv (\omega_1, \cdots, \omega_k) \in \mathbb{R}^k$ with $\|\omega\| = 1$, and $(\pi_1, \cdots, \pi_k) \in [0,1]^k$. Let $\varsigma_{ni} = \sum_{j=1}^{k} \omega_j [\overline{f}_n(z_{ni}) 1(i \le \lceil n\pi_j \rfloor)$ $- \overline{f}_{\lceil n\pi_j \rfloor}(z_{ni})]$. By Assumption A1 (iii), the ς_{ni}'s are bounded constants, i.e., $\sup_{n \ge 1} \max_{1 \le i \le n} |\varsigma_{ni}| \le \overline{c} < \infty$. By the Cramér-Wold device, it suffices to show that $\sum_{j=1}^{k} \omega_j b_{n0}(\pi_j) = n^{-1/2} \sum_{i=1}^{n} \varsigma_{ni} \varepsilon_{ni}$ is asymptotically normally distributed. Because the degenerate case is trivial, we assume that $\lim_{n \to \infty} \mathrm{Var}(\sum_{j=1}^{k} \omega_j b_{n0}(\pi_j)) > 0$ if the limit exists. This implies that $n^{-1} s_n^2 \to \underline{c} > 0$ where

$s_n^2 \equiv \sum_{i=1}^n E\left(\varsigma_{ni}^2 \varepsilon_{ni}^2\right)$. In view of the above covariance results, it remains to verify the Lindeberg condition. That is, for each $\epsilon > 0$,

$$L_n(\epsilon) \equiv s_n^{-2} \sum_{i=1}^n E\left[\varsigma_{ni}^2 \varepsilon_{ni}^2 1\left(|\varsigma_{ni}\varepsilon_{ni}| \geq \epsilon s_n\right)\right] \to 0.$$

Since $\sup_{n\geq 1} \max_{1\leq i\leq n} |\varsigma_{ni}| \leq \overline{c}$ and $n/s_n^2 \to 1/\underline{c} < \infty$, we have by the Cauchy-Schwarz and Markov inequalities that

$$\begin{aligned} L_n(\epsilon) &\leq \frac{\overline{c}^2}{s_n^2} \sum_{i=1}^n E\left[\varepsilon_{ni}^2 1\left(|\varepsilon_{ni}| \geq \frac{\epsilon s_n}{\overline{c}}\right)\right] \\ &\leq \frac{\overline{c}^2}{s_n^2} \sum_{i=1}^n [E(\varepsilon_{ni}^4)]^{1/2} \left[P\left(|\varepsilon_{ni}| \geq \frac{\epsilon s_n}{\overline{c}}\right)\right]^{1/2} \leq \frac{\overline{c}^4}{\epsilon^2 s_n^4} \sum_{i=1}^n E\left(\varepsilon_{ni}^4\right) \to 0. \end{aligned}$$

Now, we show the tightness of $\{b_{n0}(\pi)\}$. Let $0 \leq \pi_1 < \pi < \pi_2 \leq 1$. Then by the Cauchy-Schwarz inequality, $E\{[b_{n0}(\pi) - b_{n0}(\pi_1)]^2 [b_{n0}(\pi_2) - b_{n0}(\pi)]^2\} \leq \sum_{l=1}^4 \overline{b}_{nl}$, where

$$\begin{aligned} \overline{b}_{n1} &= 4n^{-2} E\left\{\left[\sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi \rfloor} \overline{f}_{ni}\varepsilon_{ni}\right]^2 \left[\sum_{i=\lceil n\pi \rfloor+1}^{\lceil n\pi_2 \rfloor} \overline{f}_{ni}\varepsilon_{ni}\right]^2\right\}, \\ \overline{b}_{n2} &= 4n^{-6} E\left\{\left[\sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi \rfloor} \sum_{j=1}^n f_{ni}(z_{nj})\varepsilon_{nj}\right]^2 \left[\sum_{i=\lceil n\pi \rfloor+1}^{\lceil n\pi_2 \rfloor} \sum_{j=1}^n f_{ni}(z_{nj})\varepsilon_{nj}\right]^2\right\}, \\ \overline{b}_{n3} &= 4n^{-4} E\left\{\left[\sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi \rfloor} \overline{f}_{ni}\varepsilon_{ni}\right]^2 \left[\sum_{i=\lceil n\pi \rfloor+1}^{\lceil n\pi_2 \rfloor} \sum_{j=1}^n f_{ni}(z_{nj})\varepsilon_{nj}\right]^2\right\}, \text{ and} \\ \overline{b}_{n4} &= 4n^{-4} E\left\{\left[\sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi \rfloor} \sum_{j=1}^n f_{ni}(z_{nj})\varepsilon_{nj}\right]^2 \left[\sum_{i=\lceil n\pi \rfloor+1}^{\lceil n\pi_2 \rfloor} \overline{f}_{ni}\varepsilon_{ni}\right]^2\right\}. \end{aligned}$$

By Assumptions 1(i)-(ii) and Davydov's inequality (e.g., Bosq (1996), p.19),

$$\begin{aligned} \overline{b}_{n1} &= 4n^{-2} E \sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi \rfloor} \sum_{k=\lceil n\pi \rfloor+1}^{\lceil n\pi_2 \rfloor} E\left(\overline{f}_{ni}^2 \varepsilon_{ni}^2 \overline{f}_{nk}^2 \varepsilon_{nk}^2\right) \\ &\quad +8n^{-2} \sum_{\lceil n\pi_1 \rfloor+1 \leq i<j\leq \lceil n\pi \rfloor} \sum_{k=\lceil n\pi \rfloor+1}^{\lceil n\pi_2 \rfloor} \operatorname{Cov}\left(\overline{f}_{ni}\varepsilon_{ni}\overline{f}_{nj}\varepsilon_{nj}, \overline{f}_{nk}^2 \varepsilon_{nk}^2\right) \\ &\leq 4c_1 \frac{\lceil n\pi \rfloor - \lceil n\pi_1 \rfloor}{n} \frac{\lceil n\pi_2 \rfloor - \lceil n\pi \rfloor}{n} + 8c_2 n^{-1} \frac{\lceil n\pi_2 \rfloor - \lceil n\pi \rfloor}{n} \sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi \rfloor} \sum_{\tau=1}^{\lceil n\pi \rfloor - \lceil n\pi_1 \rfloor} \alpha(\tau)^{\eta/(4+\eta)} \\ &\leq C(\pi - \pi_1)(\pi_2 - \pi) \leq C(\pi_2 - \pi_1)^2 \text{ for some large constant } C \end{aligned}$$

where $c_1 \equiv \sup_{n\geq 1}\max_{1\leq i\leq n}\overline{f}_{ni}^4 E(\varepsilon_{ni}^4)$, and $c_2=(2+8/\eta)\,2^{\eta/(4+\eta)}\sup_{n\geq 1}\max_{1\leq i\leq n}\overline{f}_{ni}\left\|\varepsilon_{ni}\right\|_{4+\eta}^4$, where $\left\|\xi\right\|_s \equiv \{E\left|\xi\right|^s\}^{1/s}$ for $s\geq 1$. To find an upper bound for $\overline{b}_{n2}$, we first apply the Cauchy-Schwarz inequality to obtain $\left(\overline{b}_{n2}\right)^2 \leq \overline{b}_{n21}\overline{b}_{n22}$, where $\overline{b}_{n21}=4n^{-6}E[\sum_{i=\lceil n\pi_1\rfloor+1}^{\lceil n\pi\rfloor}\sum_{j=1}^n f_{ni}\left(z_{nj}\right)\varepsilon_{nj}]^4$ and $\overline{b}_{n22}=4n^{-6}E[\sum_{i=\lceil n\pi\rfloor+1}^{\lceil n\pi_2\rfloor}\sum_{j=1}^n f_{ni}\left(z_{nj}\right)\varepsilon_{nj}]^4$. Let $\xi_j=\sum_{i=\lceil n\pi\rfloor+1}^{\lceil n\pi_2\rfloor}f_{ni}\left(z_{nj}\right)\varepsilon_{nj}$, where we suppress the dependence of ξ_j on n, π and π_2. Then by Assumptions A1(i)-(ii), (iv), the Davydov inequality, and the Hölder inequality, we have

$$\begin{aligned}
\overline{b}_{n22} &= 4n^{-6}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}\sum_{j_4=1}^{n}E\left[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}\right] \\
&\leq 96n^{-6}\sum_{1\leq j_1\leq j_2\leq j_3\leq j_4\leq n}^{n}\left|E\left[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}\right]\right| = 96n^{-6}\sum_{1\leq j_1\leq j_2\leq j_3\leq n}^{n}\left|E\left[\xi_{j_1}\xi_{j_2}\xi_{j_3}^2\right]\right| \\
&\leq Cn^{-6}\sum_{1\leq j_1\leq j_2\leq j_3\leq n}^{n}\left\|\xi_{j_1}\right\|_{4+\eta}\left\|\xi_{j_2}\xi_{j_3}^2\right\|_{(4+\eta)/3}\alpha\left(j_2-j_1\right)^{\eta/(4+\eta)} \\
&\leq Cn^{-4}\sup_{n\geq 1}\max_{1\leq j\leq n}\left\|\xi_j\right\|_{4+\eta}^4\sum_{\tau=0}^{\infty}\alpha\left(\tau\right)^{\eta/(4+\eta)} \leq Cn^{-4}\sup_{n\geq 1}\max_{1\leq j\leq n}\left\|\xi_j\right\|_{4+\eta}^4.
\end{aligned}$$

Then by Assumption A1(iii), the definition of ξ_j, and the triangle inequality, we have $\overline{b}_{n22}\leq C\{\sup_{n\geq 1}\max_{1\leq j\leq n}n^{-1}\sum_{i=\lceil n\pi\rfloor+1}^{\lceil n\pi_2\rfloor}||\varepsilon_{nj}||_{4+\eta}\}^4 \leq C\left(\pi_2-\pi\right)^4$. Analogously, we can show that $\overline{b}_{n21}\leq C\left(\pi-\pi_1\right)^4$. Then $\overline{b}_{n2}\leq C\left(\pi_2-\pi_1\right)^4$. Similarly, one can show that $\overline{b}_{nl}\leq C\left(\pi_2-\pi_1\right)^3$, $l=3,4$. It follows that $E\{\left[b_{n0}\left(\pi\right)-b_{n0}\left(\pi_1\right)\right]^2\left[b_{n0}\left(\pi_2\right)-b_{n0}\left(\pi\right)\right]^2\}\leq C\left(\pi_2-\pi_1\right)^2$. By Theorem 13.5 of Billingsley (1999), the weak convergence result follows. ■

Proof of Theorem 5.1

We only prove $\Gamma_{nb}^*\left(\cdot\right)\overset{p}{\Rightarrow}\Gamma_b\left(\cdot\right)$, as the proof for the other case is similar. Let P^* denote the probability conditional on the original sample $\mathcal{W}_n\equiv\{\left(y_{nt},x_{nt},z_{nt}\right)\}_{t=1}^n$ and E^* denote the expectation with respect to P^*. Let $O_{p^*}\left(1\right)$ and $o_{p^*}\left(1\right)$ denote the probability order under the bootstrap, e.g., $b_n=o_{p^*}\left(1\right)$ if for any $\epsilon>0$, $P^*\left(\left\|b_n\right\|>\epsilon\right)=o_p\left(1\right)$. Note that $b_n=o_p\left(1\right)$ implies that $b_n=o_{p^*}\left(1\right)$. We prove the theorem by showing that conditional on $\mathcal{W}_n$, (i) $\widehat{\sigma}^*\Gamma_{nb}^*\left(\cdot\right)\overset{p}{\Rightarrow}\sigma_0 B_1\left(\cdot\right)$, and (ii) $\widehat{\sigma}^{*2}=\sigma_0^2+o_{p^*}\left(1\right)$.

We show (ii) first. By the law of large numbers for independent but non-identically distributed (INID) sequences, $\widehat{\sigma}^{*2}=\overline{\sigma}^{*2}+o_{p^*}\left(1\right)$, where $\overline{\sigma}^{*2}=n^{-1}\sum_{i=1}^n\widehat{f}_{ni}^2\widetilde{u}_{ni}^2$. Now write $\overline{\sigma}^{*2}=n^{-1}\sum_{i=1}^n\widehat{f}_{ni}^2\varepsilon_{ni}^2+n^{-1}\sum_{i=1}^n\widehat{f}_{ni}^2\left(\widetilde{u}_{ni}-\varepsilon_{ni}\right)^2+2n^{-1}\sum_{i=1}^n\widehat{f}_{ni}^2\varepsilon_{ni}\left(\widetilde{u}_{ni}-\varepsilon_{ni}\right)\equiv d_{n1}+d_{n2}+2d_{n3}$, say. By Lemma B.1(i), it is easy to show that $d_{n1}=n^{-1}\sum_{i=1}^n\overline{f}_{ni}^2\varepsilon_{ni}^2+o_p\left(1\right)=\sigma_0^2+o_p\left(1\right)$. By Assumption A4(ii), $d_{n2}=o_p\left(1\right)$. By the Cauchy-Schwarz inequality, $d_{n3}\leq\{d_{n1}\}^{1/2}\{d_{n2}\}^{1/2}=o_p\left(1\right)$. Hence $\widehat{\sigma}^{*2}=\sigma_0^2+o_{p^*}\left(1\right)$.

Now, as in the proof of Theorem 4.2, we prove (i) in three steps. First, we demonstrate that the covariance kernel of $\widehat{\sigma}^* \widetilde{\Gamma}^*_{nb}$ converges to that of $\sigma_0 \Gamma_b$. Then we investigate the finite-dimensional distribution of $\widehat{\sigma}^* \widetilde{\Gamma}^*_{nb}$ conditional on $\mathcal{W}_n$. Finally we show the tightness of $\widehat{\sigma}^* \widetilde{\Gamma}^*_{nb}$. First, by the independence of $\{\eta_i\}$, we have

$$
\begin{aligned}
& E^* \left[\widehat{\sigma}^* \widetilde{\Gamma}^*_{nb} (\pi_1) \widehat{\sigma}^* \widetilde{\Gamma}^*_{nb} (\pi_2) \right] \\
= \; & n^{-1} \sum_{i=1}^{\lceil n(\pi_1 \wedge \pi_2) \rfloor} \widehat{f}^2_{ni} \widetilde{u}^2_{ni} + n^{-3} \sum_{i=1}^{\lceil n\pi_1 \rfloor} \sum_{j=1}^{n} \sum_{k=1}^{\lceil n\pi_2 \rfloor} K_{hij} \widetilde{u}^2_{nj} K_{hkj} \\
& - n^{-2} \sum_{i=1}^{\lceil n\pi_1 \rfloor} \sum_{j=1}^{\lceil n\pi_2 \rfloor} \widehat{f}_{ni} \widetilde{u}^2_{ni} K_{hji} - n^{-2} \sum_{i=1}^{\lceil n\pi_2 \rfloor} \sum_{j=1}^{\lceil n\pi_1 \rfloor} \widehat{f}_{ni} \widetilde{u}^2_{ni} K_{hjk} \\
\equiv \; & S^*_{n11} (\pi_1, \pi_2) + S^*_{n22} (\pi_1, \pi_2) - S^*_{n12} (\pi_1, \pi_2) - S^*_{n21} (\pi_1, \pi_2) .
\end{aligned}
$$

Let $\overline{S}^*_{n11} (\pi) = n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} \widehat{f}^2_{ni} \widetilde{u}^2_{ni}$. Then we can write $\overline{S}^*_{n11} (\pi) = n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} \widehat{f}^2_{ni} \varepsilon^2_{ni} + n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} \widehat{f}^2_{ni}$ $(\widetilde{u}_{ni} - \varepsilon_{ni})^2$ $+ 2n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} \widehat{f}^2_{ni} (\widetilde{u}_{ni} - \varepsilon_{ni}) \varepsilon_{ni} \equiv \overline{S}^*_{n11a} (\pi) + \overline{S}^*_{n11b} (\pi) + 2\overline{S}^*_{n11c} (\pi)$. It is easy to show that $\overline{S}^*_{n11a} (\pi) = n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} \overline{f}^2_{ni} \varepsilon^2_{ni} + o_p (1) \xrightarrow{p} n^{-1} \sum_{i=1}^{\lceil n\pi \rfloor} E(\overline{f}^2_{ni} \varepsilon^2_{ni})$. By Assumption A4(ii),$\overline{S}^*_{n11b} (\pi) = o_p (1)$, and by the Cauchy-Schwarz inequality $\overline{S}^*_{n11c} (\pi) = o_p (1)$. It follows that $S^*_{n11} (\pi_1, \pi_2) \xrightarrow{p} S_{11} (\pi_1, \pi_2)$. Let $\widehat{f}_{\lceil n\pi \rfloor} (z_{ni}) = n^{-1} \sum_{j=1}^{\lceil n\pi \rfloor} K_{hij}$. Then

$$
\begin{aligned}
S^*_{n22} (\pi_1, \pi_2) \; = \; & n^{-1} \sum_{i=1}^{n} \widehat{f}_{\lceil n\pi_1 \rfloor} (z_{ni}) \widehat{f}_{\lceil n\pi_2 \rfloor} (z_{ni}) \varepsilon^2_{ni} + n^{-1} \sum_{i=1}^{n} \widehat{f}_{\lceil n\pi_1 \rfloor} (z_{ni}) \widehat{f}_{\lceil n\pi_2 \rfloor} (z_{ni}) (\widetilde{u}_{ni} - \varepsilon_{ni})^2 \\
& + 2n^{-1} \sum_{i=1}^{n} \widehat{f}_{\lceil n\pi_1 \rfloor} (z_{ni}) \widehat{f}_{\lceil n\pi_2 \rfloor} (z_{ni}) (\widetilde{u}_{ni} - \varepsilon_{ni}) \varepsilon_{ni} \\
\equiv \; & S^*_{n22a} + S^*_{n22b} + 2S^*_{n22c},
\end{aligned}
$$

where we suppress the dependence of $S^{*\prime}_{n22}s$ on π_1 and π_2. Similarly to the proof of Lemma B.1(i), one can show that $\widehat{f}_{\lceil n\pi \rfloor} (z_{ni}) = \overline{f}_{\lceil n\pi \rfloor} (z_{ni}) + O_p (\nu_n)$. With this, it is straightforward to show that $S^*_{n22a} \xrightarrow{p} S_{22} (\pi_1, \pi_2)$. By the Cauchy-Schwarz inequality and Assumption A4(ii), $S^*_{n22b} \leq \{n^{-1} \sum_{i=1}^{n} \widehat{f}_{\lceil n\pi_1 \rfloor} (z_{ni}) (\widetilde{u}_{ni} - \varepsilon_{ni})^2\}^{1/2} \{n^{-1} \sum_{i=1}^{n} \widehat{f}_{\lceil n\pi_2 \rfloor} (z_{ni}) (\widetilde{u}_{ni} - \varepsilon_{ni})^2\}^{1/2} = o_p (1)$, and $S^*_{n22c} \leq \{S^*_{n22a}\}^{1/2} \{S^*_{n22b}\}^{1/2} = o_p (1)$. Hence $S^*_{n22} (\pi_1, \pi_2) \xrightarrow{p} S_{22} (\pi_1, \pi_2)$. Similarly, one can show that $S^*_{n12} (\pi_1, \pi_2) \xrightarrow{p} S_{12} (\pi_1, \pi_2)$. By symmetry, $S^*_{n21} (\pi_1, \pi_2) \xrightarrow{p} S_{21} (\pi_1, \pi_2)$.

We now show the finite dimensional convergence. Write $\widehat{\sigma}^* \widetilde{\Gamma}^*_{nb} (\pi) = n^{-1/2} \sum_{i=1}^{n} [\widehat{f}_n (z_{ni}) 1(i \leq \lceil n\pi \rfloor) - \widehat{f}_{\lceil n\pi \rfloor} (z_{ni})] \widetilde{u}_{ni} \eta_i$. Fix $k \geq 1$, $\omega \equiv (\omega_1, \cdots, \omega_k) \in \mathbb{R}^k$ with $\|\omega\| = 1$, and $(\pi_1, \cdots, \pi_k) \in [0, 1]^k$. Let $\widehat{\varsigma}_{ni} = \sum_{j=1}^{k} \omega_j [\widehat{f}_n (z_{ni}) 1(i \leq \lceil n\pi_j \rfloor) - \widehat{f}_{\lceil n\pi_j \rfloor} (z_{ni})]$. By the Cramér-Wold device, it suffices to show that $F_n \equiv \sum_{j=1}^{k} \omega_j \widehat{\sigma}^* \widetilde{\Gamma}^*_{nb} (\pi_j) = n^{-1/2} \sum_{i=1}^{n} \widehat{\varsigma}_{ni} \widetilde{u}_{ni} \eta_i$ is asymptotically normally distributed given $\mathcal{W}_n$. Write $F_n = n^{-1/2} \sum_{i=1}^{n} \varsigma_{ni} \varepsilon_{ni} \eta_i + n^{-1/2} \sum_{i=1}^{n} (\widehat{\varsigma}_{ni} \widetilde{u}_{ni}$

$-\varsigma_{ni}\varepsilon_{ni}\eta_i) \equiv F_{n1}+F_{n2}$, where $\varsigma_{ni} \equiv \sum_{j=1}^{k}\omega_j[\overline{f}_n(z_{ni})1(i \leq \lceil n\pi_j \rfloor) - \overline{f}_{\lceil n\pi_j \rfloor}(z_{ni})]$. We prove the claim by showing first that conditional on $\mathcal{W}_n$, F_{n1} is asymptotically normally distributed, and then that $F_{n2} = o_{p^*}(1)$. Conditional on $\mathcal{W}_n$, $\{\varsigma_{ni}\varepsilon_{ni}\eta_i\}$ is a mean-zero independent sequence. It remains to verify the Lindeberg or Liapounov condition. The latter holds if $\frac{1}{n}\sum_{i=1}^{n}E^*|\varsigma_{ni}\varepsilon_{ni}\eta_i|^4 \xrightarrow{p} c < \infty$. By the boundedness of ς_{ni}, Assumptions A4(i) and A1(iv), $\frac{1}{n}\sum_{i=1}^{n}E^*|\varsigma_{ni}\varepsilon_{ni}\eta_i|^4 \leq C\frac{1}{n}\sum_{i=1}^{n}\varepsilon_{ni}^4 \xrightarrow{p} C\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E(\varepsilon_{ni}^4) < \infty$. Now $E^*(F_{n2}) = 0$ and $\mathrm{Var}^*(F_{n2}) \leq 2n^{-1}\sum_{i=1}^{n}\widehat{\varsigma}_{ni}^2(\widetilde{u}_{ni}-\varepsilon_{ni})^2 + 2n^{-1}\sum_{i=1}^{n}(\widehat{\varsigma}_{ni}-\varsigma_{ni})^2\varepsilon_{ni}^2$. The first term on the right hand side (r.h.s.) is bounded by $C\max_{1\leq j\leq k}n^{-1}\sum_{i=1}^{n}[\widehat{f}_{ni}^2+\widehat{f}_{\lceil n\pi_j \rfloor}^2(z_{ni})]$ $(\widetilde{u}_{ni}-\varepsilon_{ni})^2 = o_p(1)$ by Assumption A4(ii). The second term is $o_p(1)$ by the consistency of $\widehat{f}_{\lceil n\pi \rfloor}(z_{ni})$ with $\overline{f}_{\lceil n\pi \rfloor}(z_{ni})$ for each π. Hence $\mathrm{Var}^*(F_{n2}) = o_p(1)$ and $F_{n2} = o_{p^*}(1)$ by the conditional Chebyshev inequality.

Finally, the proof of the tightness of $\{\widehat{\sigma}^*\widetilde{\Gamma}_{nb}^*(\pi)\}$ is analogous to that of $\{b_{n0}(\pi)\}$ in Theorem 4.2 so we only sketch some of the differences. For example, now $\overline{b}_{n1}$ in the proof of Theorem 4.2 becomes $\overline{b}_{n1}^* \equiv 4n^{-2}\sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi \rfloor}\sum_{j=\lceil n\pi \rfloor+1}^{\lceil n\pi_2 \rfloor}\widehat{f}_{ni}^2\widetilde{u}_{ni}^2\widehat{f}_{nj}^2\widetilde{u}_{nj}^2E\left(\eta_i^2\eta_j^2\right) \leq 4[n^{-1}\sum_{i=\lceil n\pi_1 \rfloor+1}^{\lceil n\pi_2 \rfloor}\widehat{f}_{ni}^2\widetilde{u}_{ni}^2]^2$ $= 4|H_n(\pi_2)-H_n(\pi_1)|^2$, where $H_n(\pi) \equiv n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\widehat{f}_{ni}^2\widetilde{u}_{ni}^2$. Note that $H_n(\pi) = n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\widehat{f}_{ni}^2\varepsilon_{ni}^2 +$ $n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\widehat{f}_{ni}^2(\widetilde{u}_{ni}-\varepsilon_{ni})^2 + 2n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\widehat{f}_{ni}^2(\widetilde{u}_{ni}-\varepsilon_{ni})\varepsilon_{ni}^2 = n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor}\overline{f}_{ni}^2\varepsilon_{ni}^2 + o_p(1) \xrightarrow{p} \sigma^2(\pi)$ by Assumptions A4(ii) and A1(viii), and $\sigma^2(\cdot)$ is a nondecreasing and continuous function on $[0,1]$. The proof is complete by Theorem 13.5 of Billingsley (1999). ■

Proof of Corollary 5.2

Similarly to the proof of Theorem 5.1, we only prove $\Gamma_{nb}^*(\cdot) \overset{p}{\Rightarrow} B_1(\cdot)$ by showing that conditional on $\mathcal{W}_n$, (i) $\widehat{\sigma}^*\Gamma_{nb}^*(\cdot) \overset{p}{\Rightarrow} \overline{\sigma}B_1(\cdot)$, and (ii) $\widehat{\sigma}^{*2} = \overline{\sigma}^2 + o_{p^*}(1)$, where $\overline{\sigma}^2 = \lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}$ $E\{f_{ni}^2[\varepsilon_{ni}+v_{ni}'(\gamma_{ni}-\overline{\gamma})]^2\}$, $\overline{\gamma} = \Psi^{-1}\Psi_\gamma$ and $\Psi_\gamma = \lim_{n\to\infty}n^{-1}\sum_{i=1}^{n}E(f_{ni}^2v_{ni}v_{ni}')\gamma_{ni}$.

To proceed, we first show that

$$\widehat{\gamma} = S_{(X-\widehat{X})\widehat{f}}^{-1}S_{(X-\widehat{X})\widehat{f},(Y-\widehat{Y})\widehat{f}} = \overline{\gamma} + o_p(1), \text{ and} \tag{A.27}$$

$$\widehat{f}_{ni}\widetilde{u}_{ni} = f_{ni}\left[\varepsilon_{ni}+v_{ni}'(\gamma_{ni}-\overline{\gamma})\right] + o_p(1) \text{ uniformly in } i. \tag{A.28}$$

$S_{(X-\widehat{X})\widehat{f}} = \Psi + o_p(1)$ holds under both the null and alternative hypotheses. Noting that y_{ni} $-E(y_{ni}|z_{ni}) = v_{ni}'\gamma_{ni}+\varepsilon_{ni}$, we have $S_{(X-\widehat{X})\widehat{f},(Y-\widehat{Y})\widehat{f}} = n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2v_{ni}v_{ni}'\gamma_{ni} + n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2v_{ni}\varepsilon_{ni} +$ $n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2v_{ni}(E(y_{ni}|z_{ni})-\widehat{y}_{ni}) + n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2[g_{ni}-\widehat{x}_{ni}][v_{ni}'\gamma_{ni}+\varepsilon_{ni}] + n^{-1}\sum_{i=1}^{n}\widehat{f}_{ni}^2[g_{ni}-\widehat{x}_{ni}]$ $[E(y_{ni}|z_{ni})-\widehat{y}_{ni}] \equiv T_{n1}+T_{n2}+T_{n3}+T_{n4}+T_{n5}$, say. By Lemma B.1(i)-(ii), $\sup_{1\leq i\leq n}|\widehat{f}_{ni} - f_{ni}| = O_p(\nu_n)$, and $\sup_{1\leq i\leq n}||(\widehat{x}_i-g_{ni})\widehat{f}_{ni}|| = O_p(\nu_n+\alpha_{gn})$. Using Lemmas C.4 and C.5, one can also show that $\sup_{1\leq i\leq n}|(\widehat{y}_{ni}-E(y_{ni}|z_{ni}))\widehat{f}_{ni}| = O_p(\nu_n+\alpha_{gn}+\alpha_{mn})$. With these, it is straightforward to show that $T_{n1} = n^{-1}\sum_{i=1}^{n} f_{ni}^2 v_{ni} v_{ni}' \gamma_{ni} + o_p(1) = \Psi_\gamma + o_p(1)$, and $T_{nj} = o_p(1)$ for $j = 2,3,4,5$. Similarly, by (2.10) and (2.11), uniformly in i, $\widehat{f}_{ni}\widetilde{u}_{ni} = \widehat{f}_{ni}[\varepsilon_{ni} + v_{ni}'(\gamma_{ni} - \widehat{\gamma})] + \widehat{f}_{ni}[E(y_{ni}|z_{ni}) - \widehat{y}_{ni}] - \widehat{f}_{ni}[g_{ni} - \widehat{x}_{ni}]'\widehat{\gamma} = f_{ni}[\varepsilon_{ni} + v_{ni}'(\gamma_{ni} - \overline{\gamma})] + O_p(\nu_n + \alpha_{gn} + \alpha_{mn})$.

To show (i), let $M_n^*(\pi) \equiv n^{-1/2}\sum_{i=1}^{\lceil n\pi \rfloor} \widehat{f}_{ni} u_{ni}^*$. Conditionally on $\mathcal{W}$, $M_n^*(\cdot)$ is a mean-zero Gaussian process with independent increments and covariance kernel $E^*[M_n^*(\pi_1) M_n^*(\pi_2)] = n^{-1}\sum_{i=1}^{\lceil n(\pi_1\wedge\pi_2) \rfloor} \widehat{f}_{ni}^2 \widetilde{u}_{ni}^2$ (see, e.g., Cavaliere and Taylor, 2006). Now, by (A.28) and for fixed π, $n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor} \widehat{f}_{ni}^2 \widetilde{u}_{ni}^2 = n^{-1}\sum_{t=1}^{\lceil n\pi \rfloor} f_{ni}^2 [\varepsilon_{ni} + v_{ni}'(\gamma_{ni} - \overline{\gamma})]^2 + o_p(1) \xrightarrow{p} \pi\overline{\sigma}^2$. Since $n^{-1}\sum_{i=1}^{\lceil n\pi \rfloor} \widehat{f}_{ni}^2 \widetilde{u}_{ni}^2$ is monotonically increasing in π and the limit function is continuous in π, the above convergence holds uniformly in π by the proof of Lemma A.10 in Hansen (2000b). Hence, $M_n^*(\cdot) \overset{p}{\Rightarrow} \overline{\sigma} B(\cdot)$, where $B(\cdot)$ is the standard Brownian motion on $[0,1]$. An obvious implication is that $\widehat{\sigma}^* \Gamma_{nb}^*(\cdot) = M_n^*(\cdot) - \cdot M_n^*(1) \overset{p}{\Rightarrow} \overline{\sigma} B_1(\cdot)$. Finally, $\widehat{\sigma}^{*2} = n^{-1}\sum_{i=1}^{n} \widehat{f}_{ni}^2 \widetilde{u}_{ni}^2 + o_{p^*}(1) = \overline{\sigma}^2 + o_{p^*}(1)$, where the first equality follows from the law of large numbers for INID sequences. ■

B Some Technical Lemmas

Recall $f_{ni} = f_{ni}(z_{ni})$, $\overline{f}_{ni} = \overline{f}_n(z_{ni})$, $m_{ni} = m_{ni}(z_{ni})$, $g_{ni} = g_{ni}(z_{ni})$, and $\nu_n \equiv n^{-1/2} h^{-q/2}\sqrt{\log n} + h^r$. We prove the following lemmas under Assumptions A1-A3 without imposing any null hypotheses.

Lemma B.1 *(i)* $\sup_{1\le i\le n} |\widehat{f}_{ni} - \overline{f}_{ni}| = O_p(\nu_n)$; *(ii)* $\sup_{1\le i\le n} ||(\widehat{x}_{ni} - g_{ni}(z_{ni}))\widehat{f}_{ni}|| = O_p(\nu_n + \alpha_{gn})$; *(iii)* $\sup_{1\le i\le n} |(\widetilde{m}(z_{ni}) - m_{ni}(z_{ni}))\widehat{f}_{ni}| = O_p(\nu_n + \alpha_{mn})$.

Proof. By the triangle inequality, $\sup_{1\le i\le n} |\widehat{f}_{ni} - \overline{f}_{ni}| \le \sup_{1\le i\le n} |\frac{1}{n}\sum_{j\ne i}^{n} [K_{hij} - E_j(K_{hij})]| + \sup_{1\le i\le n} |\frac{1}{n}\sum_{j\ne i}^{n} [E_j(K_{hij}) - f_{nj}(z_{ni})]| + \frac{1}{n}\sup_{1\le i\le n} f_{ni}$. By Lemmas C.5 and C.4 the first and second terms are $O_p(n^{-1/2}h^{-q/2}\sqrt{\log n})$ and $O_p(h^r)$, respectively. By Assumption A1(iii), the last term is $O_p(n^{-1})$. Hence (i) follows. Next, write $(\widehat{x}_{ni} - g_{ni}(z_{ni}))\widehat{f}_{ni}(z_{ni}) = n^{-1}\sum_{j\ne i}^{n} K_{hij} [g_{nj}(z_{nj}) - g_{ni}(z_{ni})] + n^{-1}\sum_{j\ne i}^{n} K_{hij} v_{nj} \equiv G_1(z_{ni}) + G_2(z_{ni})$. By the triangle inequality, $\sup_{1\le i\le n} |G_1(z_{ni})| \le \sup_{1\le i\le n} |G_1(z_{ni}) - E[G_1(z_{ni})]| + \sup_{1\le i\le n} |E[G_1(z_{ni})]|$. The first term is $O_p(n^{-1/2}h^{-q/2}\sqrt{\log n})$ by Lemma C.5. Next, by the triangle inequality, Assumption

A1, and Lemma C.4,

$$\begin{aligned}
&\sup_{1\leq i\leq n}\left|E\left[G_1\left(z_{ni}\right)\right]\right| \\
\leq\ & \sup_{1\leq i\leq n} n^{-1}\sum_{j\neq i}^{n}\left|EE_j\left\{K_{hij}\left[g_{nj}\left(z_{nj}\right)-g_{nj}\left(z_{ni}\right)\right]\right\}\right| + \sup_{1\leq i\leq n} n^{-1}\sum_{j\neq i}^{n}\left|E\left\{K_{hij}\left[g_{nj}\left(z_{ni}\right)-g_{ni}\left(z_{ni}\right)\right]\right\}\right| \\
\leq\ & \sup_{1\leq i\leq n} h^r E\left[D_g\left(z_{ni}\right)\right] + \sup_{1\leq i\leq n} 2\alpha_{gn} n^{-1}\sum_{j\neq i}^{n}\left|E\left\{K_h\left(z_{nj}-z_{ni}\right)c_{gn}\left(z_{ni}\right)\right\}\right| = O(h^r)+O(\alpha_{gn}),
\end{aligned}$$

where α_{gn} and $c_{gn}\left(\cdot\right)$ are defined in Assumption A1(vii) and $D_g\left(\cdot\right)$ is as defined in Lemma C.4. Hence $\sup_{1\leq i\leq n}\left|G_1\left(z_{ni}\right)\right| = O_p(\nu_n+\alpha_{gn})$. By Lemma C.5, $\sup_{z\in\mathbb{R}^q}\left|G_2\left(z\right)\right| = O_p(n^{-1/2}h^{-q/2}\sqrt{\log n})$. Hence (ii) follows. Note that $\left(\widetilde{m}\left(z_{ni}\right)-m_{ni}\left(z_{ni}\right)\right)\widehat{f}_{ni} = n^{-1}\sum_{j\neq i}^{n}K_{hij}[m_{nj}\left(z_{nj}\right)-m_{ni}\left(z_{ni}\right)] + n^{-1}\sum_{j\neq i}^{n}K_{hij}\varepsilon_{nj} - n^{-1}\sum_{j\neq i}^{n}K_{hij}x_{nj}'\left(\widehat{\gamma}-\gamma_0\right)$. Uniformly in i, the first term is $O_p\left(\nu_n\right)$ by the same arguments as in the proof of (ii); the second term is $O_p(n^{-1/2}h^{-q/2}\sqrt{\log n})$ by Lemma C.5; and the last term is $O_p\left(n^{-1/2}\right)$, because $\sqrt{n}\left(\widehat{\gamma}-\gamma_0\right) = O_p\left(1\right)$ by Theorem 3.1. Then (iii) follows. ■

Lemma B.2 $n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^2\left(x_{ni}-\widehat{x}_{ni}\right)\left(x_{ni}-\widehat{x}_{ni}\right)' = \Phi\left(\pi\right)+o_p\left(1\right)$ *uniformly in* $\pi\in[0,1]$.

Proof. We only consider the case $p=1$, as the other cases follows from this case and the Cauchy-Schwarz inequality. Noting that $x_{ni}=v_{ni}+g_{ni}$, we have $n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^2\left(x_{ni}-\widehat{x}_{ni}\right)^2 = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^2 v_{ni}^2 + n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^2\left(g_{ni}-\widehat{x}_{ni}\right)^2 + 2n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\widehat{f}_{ni}^2 v_{ni}\left(g_{ni}-\widehat{x}_{ni}\right) \equiv T_{n1}\left(\pi\right)+T_{n2}\left(\pi\right)+2T_{n3}\left(\pi\right)$. Write $T_{n1}\left(\pi\right) = n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}\overline{f}_{ni}^2 v_{ni}^2 + n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}(\widehat{f}_{ni}^2-\overline{f}_{ni}^2)v_{ni}^2$. The first term converges in probability to $\Phi\left(\pi\right)$ uniformly in π by Assumption A1(viii). By Lemma B.1(i) and Assumption A1, the second term is $\sup_{0\leq\pi\leq1}n^{-1}\sum_{i=1}^{\lceil n\pi\rfloor}(\widehat{f}_{ni}^2-\overline{f}_{ni}^2)v_{ni}^2 \leq \max_{1\leq i\leq n}|\widehat{f}_{ni}^2-\overline{f}_{ni}^2|n^{-1}\sum_{i=1}^{n}v_{ni}^2 = o_p\left(1\right)$. Similarly, $\sup_{0\leq\pi\leq1}\left|T_{n2}\left(\pi\right)\right| \leq \max_{1\leq i\leq n}\left|\left(g_{ni}-\widehat{x}_{ni}\right)\widehat{f}_{ni}\right|^2 = o_p\left(1\right)$ by Lemma B.1(ii). By the Cauchy-Schwarz inequality, $\sup_{0\leq\pi\leq1}\left|T_{n3}\left(\pi\right)\right| = o_p\left(1\right)$. ■

Lemma B.3 *(i)* $A_{n1}\left(\pi\right) \equiv n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}\left(K_{hij}-E_jK_{hij}\right)v_{ni}\overline{f}_{ni}\varepsilon_{ni} = o_p\left(1\right)$ *uniformly in* $\pi\in[0,1]$;

(ii) $A_{n2}\left(\pi\right) \equiv n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}(g_{nj}-g_{ni})\overline{f}_{ni}\varepsilon_{ni} = o_p\left(1\right)$ *uniformly in* $\pi\in[0,1]$;

(iii) $A_{n3}\left(\pi\right) \equiv n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}v_{nj}\overline{f}_{ni}\varepsilon_{ni} = o_p\left(1\right)$ *uniformly in* $\pi\in[0,1]$;

(iv) $A_{n4}\left(\pi\right) \equiv n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}(m_{nj}-m_{ni})\overline{f}_{ni}\varepsilon_{ni} = o_p\left(1\right)$ *uniformly in* $\pi\in[0,1]$;

(v) $A_{n5}\left(\pi\right) \equiv n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}K_{hij}v_{ni}\overline{f}_{ni}\varepsilon_{nj} = o_p\left(1\right)$ *uniformly in* $\pi\in[0,1]$.

Proof. We only prove (i) and (ii), as the other cases are similar. To prove (i), let $w_i = \left(\varepsilon_{ni}, v_{ni}', z_{ni}'\right)'$ and $\phi\left(w_j,w_i\right) = \left(K_{hij}-E_jK_{hij}\right)v_{ni}\overline{f}_{ni}\varepsilon_{ni}$. Then $A_{n1}\left(\pi\right) = n^{-3/2}\sum_{1\leq j<i\leq\lceil ns\rfloor}\phi\left(w_j,w_i\right)$

$+n^{-3/2}\sum_{1\le i<j\le n}\phi(w_j,w_i)\ -n^{-3/2}\sum_{\lceil ns\rfloor+1\le i<j\le n}^{n}\phi(w_j,w_i)\ \equiv\ A_{n11}(\pi)+A_{n12}-A_{n13}(\pi)$. It suffices to show $\sup_{0\le\pi\le1}|A_{n11}(\pi)|=o_p(1)$, and $\sup_{0\le\pi\le1}|A_{n13}(\pi)|=o_p(1)$. Write

$$E\left[A_{n11}(\pi)\right]^4=n^{-6}\sum_{1\le i_1<i_2\le\lceil n\pi\rfloor}\sum_{1\le i_3<i_4\le\lceil n\pi\rfloor}\sum_{1\le i_5<i_6\le\lceil n\pi\rfloor}\sum_{1\le i_7<i_8\le\lceil n\pi\rfloor}\phi(w_{i_1},w_{i_2})\phi(w_{i_3},w_{i_4})\times\phi(w_{i_5},w_{i_6})\phi(w_{i_7},w_{i_8}). \tag{B.1}$$

It is easy to show that the dominating terms in the above summation constitute two cases: (a) $i_1,...,i_8$ are distinct integers; (b) $\{i_1,i_2\}$, $\{i_3,i_4\}$, $\{i_5,i_6\}$ and $\{i_7,i_8\}$ form two identical pairs (e.g., $\{i_1,i_2\}=\{i_3,i_4\}$ and $\{i_5,i_6\}=\{i_7,i_8\}$). We will use $EA_{n11(l)}$ to denote these two cases ($l=a,b$).

For case (a), let $i_1,...,i_8$ be distinct integers with $1\le i_j\le\lceil n\pi\rfloor$. Let $1\le k_1<...<k_8\le\lceil n\pi\rfloor$ be the permutation of $i_1,...,i_8$ in ascending order and let d_c be the c-th largest difference among $k_{j+1}-k_j$, $j=1,...,7$. Define $H(k_1,...,k_8)=\phi(w_{i_1},w_{i_2})\phi(w_{i_3},w_{i_4})\phi(w_{i_5},w_{i_6})$ $\phi(w_{i_7},w_{i_8})$. For any $1\le j\le7$, put $P_0^{(8)}\left(E^{(8)}\right)=P\left((w_{i_1},..,w_{i_8})\in E^{(8)}\right)$, and $P_j^{(8)}\left(E^{(j)}\times E^{(8-j)}\right)$ $=P((w_{i_1},..,w_{i_j})\in E^{(j)})P((w_{i_{j+1}},..,w_{i_8})\in E^{(8-j)})$, where $E^{(j)}$ is a Borel set in $\mathbb{R}^{jd}$ and d is the dimension of w_i. It is easy to verify that for any $0\le j\le7$, $\int|H(k_1,...,k_8)|^{1+\eta/4}dP_j^{(8)}\le Ch^{-q\eta}$. By Lemma C.1 with $\widetilde{\eta}=\eta/4$,

$$|E[H(k_1,...,k_8)]|\le\begin{cases}Ch^{-4q\eta/(4+\eta)}\alpha^{\frac{\eta}{4+\eta}}(k_2-k_1) & \text{if } k_2-k_1=d_1\\ Ch^{-4q\eta/(4+\eta)}\alpha^{\frac{\eta}{4+\eta}}(k_8-k_7) & \text{if } k_8-k_7=d_1.\end{cases}$$

Therefore

$$\begin{aligned}&\sum_{\substack{1\le k_1<...<k_8\le n\\k_2-k_1=d_1}}|E[H(k_1,...,k_8)]|\\&\le Ch^{-4q\eta/(4+\eta)}\sum_{k_1=1}^{n-7}\sum_{k_2=k_1+\max_{j\ge3}\{k_j-k_{j-1}\}}^{n-6}\sum_{k_3=k_2+1}^{n-5}\cdots\sum_{k_8=k_7+1}^{n}\alpha^{\frac{\eta}{4+\eta}}(k_2-k_1)\\&\le Ch^{-4q\eta/(4+\eta)}\sum_{k_1=1}^{n-7}\sum_{k_2=k_1+1}^{n-6}(k_2-k_1)^6\alpha^{\frac{\eta}{4+\eta}}(k_2-k_1)\le Cnh^{-4q\eta/(4+\eta)}\sum_{j=1}^{n}j^6\alpha^{\frac{\eta}{4+\eta}}(j).\end{aligned}\tag{B.2}$$

Similarly, we have

$$\sum_{\substack{1\le k_1<...<k_8\le n\\k_8-k_7=d_1}}|E[H(k_1,...,k_8)]|\le Cnh^{-4q\eta/(4+\eta)}\sum_{j=1}^{n}j^6\alpha^{\frac{\eta}{4+\eta}}(j), \tag{B.3}$$

$$\sum_{\substack{1\le k_1<...<k_8\le n\\k_2-k_1=d_2\text{ or }k_8-k_7=d_2}}|E[H(k_1,...,k_8)]|\le Cn^2h^{-4q\eta/(4+\eta)}\sum_{j=1}^{n}j^5\alpha^{\frac{\eta}{4+\eta}}(j), \tag{B.4}$$

$$\sum_{\substack{1\leq k_1<...<k_8\leq n\\ k_2-k_1=d_3 \text{ or } k_8-k_7=d_3}} |E[H(k_1,...,k_8)]| \leq Cn^3h^{-4q\eta/(4+\eta)}\sum_{j=1}^{n} j^4\alpha^{\frac{\eta}{4+\eta}}(j), \quad \text{(B.5)}$$

and for all other subcases ($k_2-k_1=d_c$ and $k_8-k_7=d_{c'}$ for $c,c'\geq 4$) we have

$$\sum_{\substack{1\leq k_1<...<k_8\leq n\\ \text{other subcases}}} |E[H(k_1,...,k_8)]| \leq Cn^4h^{-4q\eta/(4+\eta)}\sum_{j=1}^{n} j^3\alpha^{\frac{\eta}{4+\eta}}(j). \quad \text{(B.6)}$$

By (B.2)-(B.6), Assumption A3, and the fact that $\eta/(4+\eta)<1/2$, we have

$$\begin{aligned} EA_{n11(a)} &\leq n^{-6}\sum_{1\leq k_1<...<k_8\leq n} |E[H(k_1,...,k_8)]| \\ &\leq Cn^{-2}h^{-4q\eta/(4+\eta)}\sum_{j=1}^{n} j^3\alpha^{\frac{\eta}{4+\eta}}(j) = O\left(n^{-2}h^{-4q\eta/(4+\eta)}\right) = o\left(n^{-1}\right). \quad \text{(B.7)}\end{aligned}$$

Now for case (b), some calculations show that

$$EA_{n11(b)} = O\left(n^{-2}h^{-2q}\right) = o\left(n^{-1}\right). \quad \text{(B.8)}$$

Hence $E[A_{n11}(\pi)]^4 = o\left(n^{-1}\right)$ by (B.7)-(B.8) and the remark after (B.1). Let $\epsilon>0$ be arbitrary. Then by the implication rule and the Chebyshev inequality, $P(\sup_{0\leq\pi\leq1}||A_{n11}(\pi)||>\epsilon)\leq\sum_{l=1}^{n}P(A_{n11}(l/n)>\epsilon)\leq\epsilon^{-4}\sum_{l=1}^{n}E|A_{n11}(l/n)|^4=o(1)$. It follows that

$$\sup_{0\leq\pi\leq1}|A_{n11}(\pi)| = o_p(1). \quad \text{(B.9)}$$

Now let $\widetilde{\phi}(w_i,w_j)=\phi(w_j,w_i)$ and $\widetilde{w}_i=w_{n-i+1}$ for $1\leq i,j,\leq n$. Then

$$\sup_{1\leq l\leq n}|\sum_{l\leq i<j\leq n}\phi(w_j,w_i)| = \sup_{1\leq l\leq n}|\sum_{1\leq i<j\leq n-l+1}\phi(w_{n-j+1},w_{n-i+1})| = \sup_{1\leq l\leq n}|\sum_{1\leq i<j\leq l}\widetilde{\phi}(\widetilde{w}_i,\widetilde{w}_j)|.$$

We can thus apply the above method to $\{\widetilde{w}_i\}$ to obtain $\sup_{0\leq\pi\leq1}|A_{n13}(\pi)|=o_p(1)$.

To prove (ii), let $w_i=(\varepsilon_{ni},z_{ni}')'$, $\varphi_0(w_j,w_i)=K_{hij}(g_{nj}-g_{ni})\overline{f}_{ni}\varepsilon_{ni}$, and $\varphi(w_j,w_i)=\varphi_0(w_j,w_i)-E_j[\varphi_0(w_j,w_i)]$. Then $A_{n2}(\pi)=n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}\varphi(w_j,w_i)+n^{-3/2}\sum_{i=1}^{\lceil n\pi\rfloor}\sum_{j\neq i}^{n}E_j\varphi_0(w_j,w_i)\equiv A_{n21}(\pi)+A_{n22}(\pi)$. Analogously to the proof of (i), we can show $\sup_{0\leq\pi\leq1}|A_{n21}(\pi)|=o_p(1)$. For $A_{22}(\pi)$, let $A_{22,j}(\pi)\equiv n^{-1/2}\sum_{i=1}^{\lceil n\pi\rfloor}\{\int K(z)[g_{nj}(z_{ni}+hz)-g_{ni}(z_{ni})]f_{nj}(z_{ni}+hz)dz\}\overline{f}_{ni}\varepsilon_{ni}$. Obviously $E[A_{22,j}(\pi)]=0$, and by Assumptions A1 and A2 we can easily show that $E[A_{22,j}(\pi)]^4=O(h^{4r}+\alpha_{gn}^4)$. By the implication rule, the Chebyshev inequality, and Assumption A3, $\sup_{0\leq\pi\leq1}|A_{22,j}(\pi)|=O_p(n(h^{4r}+\alpha_{gn}^4))=o_p(1)$. It follows that $\sup_{0\leq\pi\leq1}|A_{n22}(\pi)|\leq n^{-1}\sum_{j=1}^{n}\sup_{0\leq\pi\leq1}|A_{22,j}(\pi)|=o_p(1)$. Hence $\sup_{0\leq\pi\leq1}|A_{n2}(\pi)|=o_p(1)$. ■

C Additional Technical Lemmas

This appendix presents some technical lemmas that are used in proving the main results.

Lemma C.1 *Let $\{W_i\}$ be a strong mixing process with mixing coefficient $\alpha(i)$. For any integer $d > 1$ and integers $(i_1, ..., i_d)$ such that $1 \leq i_1 < i_2 < \cdots < i_d$, let θ be a Borel measurable function such that max$\{\int |\theta(w_1, ..., w_d)|^{1+\widetilde{\eta}} dF(w_1, ..., w_d), \int |\theta(w_1, ..., w_p)|^{1+\widetilde{\eta}} dF^{(1)}(w_1, ..., w_j)$ $dF^{(2)}(w_{j+1}, ..., w_d)\} \leq M$ for some $\widetilde{\eta} > 0$ and $M > 0$, where $F = F_{i_1,...,i_d}$, $F^{(1)} = F_{i_1,...,i_j}$, $F^{(2)} = F_{i_j+1,,...,i_d}$ are the distribution functions of $(W_{i_1}, ..., W_{i_d})$, $(W_{i_1}, ..., W_{i_j})$, and $(W_{i_{j+1}}, ..., W_{i_d})$, respectively. Then*

$$\begin{aligned} & \left| \int \theta(w_1, ..., w_d) \, dF(w_1, ..., w_d) - \int \theta(w_1, ..., w_d) \, dF^{(1)}(w_1, ..., w_j) \, dF^{(2)}(w_{j+1}, ..., w_d) \right| \\ \leq \; & 4M^{1/(1+\widetilde{\eta})} \alpha(i_{j+1} - i_j)^{\widetilde{\eta}/(1+\widetilde{\eta})}. \end{aligned}$$

Proof. See Lemma 2.1 of Sun and Chiang (1997). ■

Lemma C.2 *Let $\{W_i\}$ be a strong mixing process with mixing coefficient $\alpha(i)$ and taking values in $\mathbb{R}^a$. Let $\phi(\cdot, \cdot, \cdot)$ be a symmetric Borel measurable function defined on $\mathbb{R}^a \times \mathbb{R}^a \times \mathbb{R}^a$ such that $M_{n12} \equiv \max_{1 \leq i < j < l \leq n} \max \{ \int_{\mathbb{R}^{3a}} |\phi(w_i, w_j, w_l)|^{1+\widetilde{\eta}} dF_i(w_i) \, dF_{jl}(w_j, w_l), \int_{\mathbb{R}^{3a}} |\phi(w_i, w_j, w_l)|^{1+\widetilde{\eta}}$ $dF_i(w_i) \, dF_j(w_j) \, dF_l(w_l)\}$, and $M_{n3} \equiv \max_{1 \leq i < j < l \leq n} \max\{ \int_{\mathbb{R}^{3a}} |\phi(w_i, w_j, w_l)|^{1+\widetilde{\eta}} dF_i(w_i) \, dF_{jl}(w_j, w_l)$, $\int_{\mathbb{R}^{3a}} |\phi(w_i, w_j, w_l)|^{1+\widetilde{\eta}} dF_{ijl}(w_i, w_j, w_l)\}$, where $F_i(\cdot)$, $F_{ij}(\cdot, \cdot)$ and $F_{ijl}(\cdot, \cdot, \cdot)$ are the distributions of W_i, (W_i, W_j), and (W_i, W_j, W_l), respectively. Then*

$$E\left[\sum\sum\sum_{1 \leq i < j < l \leq n} \phi(W_i, W_j, W_l) \right] = O\left(n^3 E\left[\phi(\overrightarrow{W}_i, \overrightarrow{W}_j, \overrightarrow{W}_l)\right]\right) + O\left(n^2 M_{n12}^{1/(1+\widetilde{\eta})}\right) + O\left(n M_{n3}^{1/(1+\widetilde{\eta})}\right),$$

where $\{\overrightarrow{W}_i\}$ denotes an independent process that has the same marginal distribution as the dependent process $\{W_i\}$.

Proof. The proof follows from a modification of that of Lemma B.2 in Fan and Li (1999).

■

The following definition is adopted from Robinson (1988).

Definition C.3 *$\mathcal{G}_\mu^\alpha, \alpha > 0, \mu > 0$, is the class of functions $\vartheta : \mathbb{R}^d \to \mathbb{R}$ satisfying: ϑ is $(m-1)$-times partially differentiable, for $m-1 \leq \mu \leq m$; for some $\rho > 0$, sup$_{y \in \phi_{z\rho}} |\vartheta(y) - \vartheta(z)| / |y - z|^\mu$ $\leq D_\vartheta(z)$ for all z, where $\phi_{z\rho} = \{y : |y - z| < \rho\}$; $Q_\vartheta = 0$ when $m = 1$; Q_ϑ is an $(m-1)$th degree homogeneous polynomial in $y - z$ with coefficients the partial derivatives of ϑ at z of orders 1 through $m - 1$ when $m > 1$; and $\vartheta(z)$, its partial derivatives of order $m - 1$ and less, and $D_\vartheta(z)$ have finite αth moments.*

Lemma C.4 *Suppose K satisfies Assumption A2, $f_{nj} \in \mathcal{G}_r^\alpha$, and $\vartheta_{nj} \in \mathcal{G}_r^\alpha$. Let $z \in \mathbb{R}^q$, and $h \to 0$ as $n \to \infty$. Then (i) $|E[K((z_{nj}-z)/h) - h^q f_{nj}(z)]| \leq h^{q+r} D_{f_{nj}}(z)$ uniformly in z, and (ii) $|E\{[\vartheta_{nj}(z_{nj}) - \vartheta_{nj}(z)] K((z_{nj}-z)/h)\}| \leq h^{q+r} D_{\vartheta_{nj}}(z)$ uniformly in z, where $f_{nj}(\cdot)$ denotes the density function of z_{nj}, and both $D_{f_{nj}}(z_{nj})$ and $D_{\vartheta_{nj}}(z_{nj})$ have finite αth moments.*

Proof. See Lemmas 4-5 of Robinson (1988). ■

To apply Lemma C.4, we will suppress the dependence of $D_{f_{nj}}(\cdot)$ and $D_{\vartheta_{nj}}(\cdot)$ on $j \in \{1, 2, ..., n\}$ by assuming that they are dominated respectively by functions $D_f(\cdot)$ and $D_\vartheta(\cdot)$ that have finite αth moments.

Lemma C.5 *Under Assumptions A1-A3,* $\sup_{z \in \mathbb{R}^q} |\Psi(z) - E\Psi(z)| = O_p(n^{-1/2} h^{-q/2} \sqrt{\log n})$, *where $\Psi(z) = n^{-1} h^{-q} \sum_{i=1}^n \xi_{ni} K((z_{ni}-z)/h)$ and $\xi_{ni} = 1, v_{ni}, \varepsilon_{ni}$ or m_{ni}.*

Proof. The proof follows from Lemma D6 of Su and Xiao (2008). ■