Estimation of Average Derivatives of Latent Regressors: With an Application to Inference on Buffer-Stock Saving *

Hao Dong † Yuya Sasaki ‡

September, 2022

Abstract

This paper proposes a density-weighted average derivative estimator based on two noisy measures of a latent regressor. Both measures have classical errors with possibly asymmetric distributions. We show that the proposed estimator achieves the root-n rate of convergence, and derive its asymptotic normal distribution for statistical inference. Simulation studies demonstrate excellent small-sample performance supporting the root-n asymptotic normality. Based on the proposed estimator, we construct a formal test on the sub-unity of the marginal propensity to consume out of permanent income (MPCP) under a nonparametric consumption model and a permanent-transitory model of income dynamics with nonparametric distribution. Applying the test to four recent waves of U.S. Panel Study of Income Dynamics (PSID), we reject the null hypothesis of the unit MPCP in favor of a sub-unit MPCP, supporting the buffer-stock model of saving.

Keywords: Average derivative, latent variables, income dynamics, consumption

JEL Codes: C14, C23, D31

---

*Financial supports from Bush Institute-SMU Economic Growth Initiative Research Grant and SMU University Research Council Grant are gratefully acknowledged.

†Hao Dong. Department of Economics, Southern Methodist University, 3300 Dyer Street, Dallas, TX 75275, US. Email: haod@smu.edu. Phone: (214) 768-3288.

‡Yuya Sasaki. Department of Economics, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235-1819, US. Email: yuya.sasaki@vanderbilt.edu. Phone: (615) 343-3016.

1 Introduction

Rational forward-looking agents should have a unit marginal propensity to consume out of permanent shocks in income (MPCP). Carroll (2009) demonstrates that the MPCP is strictly less than one in the context of the buffer-stock model where inpatient consumers have a standard precautionary saving motive, and further shows through simulations across a wide range of structural assumptions that the MPCP ranges from 0.75 to 0.92. Thus, differentiating between the standard model and Carroll's buffer-stock model can be accomplished by accessing whether the MPCP is strictly less than one.

When we take this testable implication of the buffer-stock model to statistical inference based on empirical data, we encounter a fundamental issue. Namely, we do not observe the permanent income in data. If we could observe the permanent income shock X^*, as well as the consumption growth Y, then a statistical inference about the theory of Carroll boils down to inference about the nonparametric regression function $g(\cdot) = E[Y|X^* = \cdot]$. Specifically, rejecting the null hypothesis that a (weighted) average derivative of g is greater than or equal to one against the alternative that it is strictly less than one will provide a statistical support for the theory of Carroll. Under the unobservability of the permanent income shock X^* in data, however, the existing econometric methods of estimation and inference for (weighted) average derivatives do not apply. To overcome this, we develop a novel method and theory of estimation and inference for weighted average derivatives when the latent regressor X^* is unobserved, but two noisy measures of X^* are available in data, as is the case with the standard permanent-transitory models of earnings and income dynamics.

This paper, in terms of its technical aspects, belongs to the vast literature on measurement error models and deconvolution. See books by Carroll, Ruppert, Stefanski and Crainiceanu (2006), Meister (2009) and Horowitz (2009) and surveys by Chen, Hong and Nekipelov (2011), Schennach (2016) and Schennach (2021) for reviews. The literature on deconvolution started out with the deconvolution kernel density methods under known error distributions (Carroll and Hall, 1988; Stefanski and Carroll, 1990; Fan, 1991a,b; Bissantz, Dümbgen, Holzmann and Munk, 2007; Bissantz and Holzmann, 2008; van Es and Gugushvili, 2008; Lounici and Nickl, 2011; Schmidt-Hieber, Munk and Dümbgen, 2013), followed by those under unknown error distributions (Diggle and Hall, 1993; Horowitz and Markatou, 1996; Neumann and Hössjer, 1997; Efromovich, 1997; Li and Vuong, 1998; Delaigle, Hall and Meister, 2008; Johannes, 2009; Comte and Lacour, 2011; Kato and Sasaki, 2018; Kato, Sasaki and Ura, 2021). Among the latter set of papers, Horowitz and Markatou (1996) and Delaigle, Hall and Meister (2008) use repeated measurements with symmetrically and identically distributed errors, while Li and Vuong (1998) propose an alternative estimator based on Kotlarski's lemma (Kotlarski, 1967) that does not require known error distribution – also see Bohnomme and Robin (2010) and Comte and Kappus (2015). Also related is Adusumilli, Kurisu, Otsu and Whang (2020) who studies distribution function instead of density function.

These deconvolution kernel density methods extend to methods for nonparametric errors-in-variables regression. Fan and Truong (1993) and Fan and Masry (1992) study Nadaraya-Watson estimator under known error distribution, followed by extensions by Delaigle and Meister (2007), Delaigle, Fan and Carroll (2009) and Delaigle, Hall and Jamshidi (2015). Delaigle, Hall and Meister (2008), Adusumilli and Otsu (2018) and Kato and Sasaki (2019) consider cases of unknown error distribution with symmetrically and identically distributed errors, while Li (2002), Schennach (2004), Schennach, White and Chalak (2012), Schennach and Hu (2013) and Hu and Sasaki (2015) consider cases of unknown error distribution with repeated measurements. Fan (1995) and Dong, Otsu and Taylor (2021) study average derivatives of nonparametric errors-in-variables regression under known error distribution and unknown symmetric error distribution, respectively. The current paper is closely related to the last two references in that we are also interested in $\sqrt{n}$-consistent estimation and inference for average derivatives for the purpose of the aforementioned statistical inference about the hypothesis of buffer-stock saving. However, unlike Fan (1995) or Dong, Otsu and Taylor (2021), we allow for the unknown error distribution to be non-symmetric, in light of the recent empirical reports that components of earnings and income have skewed distributions (e.g., Bonhomme and Robin, 2010; Guvenen, Ozkan and Song, 2014; Hu, Moffitt and Sasaki, 2019; Guvenen, Karahan, Ozkan and Song, 2021).

Despite the extensive econometric and statistical literature on deconvolution as summarized in the prior two paragraphs, none of the existing papers to the best of our knowledge can conduct statistical inference for the MPCP under an unknown and possibly asymmetric error distribution, even though the skewness has been reported to be very likely by recent empirical studies on earnings and income dynamics. Motivated by the aforementioned economic question concerning the MPCP, therefore, this paper fills this important void in the deconvolution literature by proposing novel methods of estimation and inference for average derivatives of latent regressors whose error distribution can be both unknown and non-symmetric.

Regarding the application to the inference on buffer-stock saving motivated by Carroll (2009), there are a couple of related existing papers. In particular, Blundell, Pistaferri and Preston (2008) and Arellano, Blundell and Bonhomme (2017) investigate MPCP using empirical data. Blundell, Pistaferri and Preston (2008) use a linear parametric consumption model and conduct inference on the parameter that represents a constant MPCP. In contrast, we use a nonparametric consumption model with a possibly non-constant MPCP. Arellano, Blundell and Bonhomme (2017) use a nonlinear model of income dynamics, although their identification and estimation approach *per se* does not lead to statistical inference on MPCP. In contrast, at the cost of linear model of income dynamics, our proposed approach allows for statistical inference on MPCP under a nonparametric consumption model, and thus nonparametrically enables hypothesis testing about buffer-stock saving. In this way, the novel method proposed in this paper complements the existing literature on the empirical analysis of MPCP.

2 Methodology

2.1 Motivation and overview

Consider the permanent-transitory model of income dynamics

$$\iota_{jt} = \pi_{jt} + \tau_{jt}$$

$$\pi_{jt} = \pi_{jt-1} + \eta_{jt}$$

where ι_{jt}, π_{jt}, τ_{jt} and η_{jt} denote observed log income, latent log permanent income, latent log transitory income and latent permanent income shock, respectively, of individual j in year t. Under this setup,

$$\underbrace{\iota_{jt} - \iota_{jt-2}}_{=:X_j} = \underbrace{\eta_{jt}}_{=:X_j^*} + \underbrace{\eta_{jt-1} + \tau_{jt} - \tau_{jt-2}}_{=:\epsilon_j} \qquad \text{and} \tag{2.1}$$

$$\underbrace{\iota_{jt+1} - \iota_{jt-1}}_{=:W_j} = \underbrace{\eta_{jt}}_{=:X_j^*} + \underbrace{\eta_{jt+1} + \tau_{jt+1} - \tau_{jt-1}}_{=:\nu_j} \tag{2.2}$$

hold, where (X_j, W_j) is observed and $(X_j^*, \epsilon_j, \nu_j)$ is unobserved by a researcher. Although not necessary, if we assume that the shocks, η_{jt-1}, η_{jt}, η_{jt+1}, τ_{jt-2}, τ_{jt-1}, τ_{jt}, and τ_{jt+1}, are mutually independent, then the latent components X_j^*, ϵ_j, and ν_j are also mutually independent.[1]

Let Y_j denote consumption growth $C_{jt} - C_{jt-1}$ from year $t-1$ to year t. Consider the nonparametric regression function $g(\cdot) = E[Y_j | X_j^* = \cdot]$. The derivative g' of this function quantifies the MPCP introduced in Section 1. The theory of Carroll (2009) is that the buffer-stock model that arises under inpatient consumers with a standard precautionary saving motive implies $g' < 1$ rather than $g' = 1$. This testable implication leads to the null and alternative hypotheses

$$H_0 : \theta_1 \geq 0 \qquad \text{and} \qquad H_1 : \theta_1 < 0, \tag{2.3}$$

where $\theta_1 = E[\{g'(X^*) - 1\} f(X^*)]$ and f is the density of X^*. Rejection of H_0 in favor of H_1 implies that there is at least some location x^* in the support of X^* such that $g'(x^*) < 1$, thus providing statistical support for the theory of Carroll (2009).

As we formally present in Section 2.2 ahead, we propose to estimate θ_1 by

$$\hat{\theta}_1 = -\frac{2}{n^2 b_n^3} \sum_{j=1}^{n} \sum_{k=1}^{n} (Y_j - W_j) \int \hat{\mathbb{K}}\left(\frac{x - X_j}{b_n}\right) \hat{\mathbb{K}}'\left(\frac{x - X_k}{b_n}\right) dx, \tag{2.4}$$

where $\hat{\mathbb{K}}(u) = \frac{1}{2\pi} \int e^{-\mathrm{i}tu} \frac{K^{\mathrm{ft}}(t)}{\hat{f}_\epsilon^{\mathrm{ft}}(t/b_n)} dt$, $\mathrm{i} = \sqrt{-1}$, K^{ft} is the Fourier transform of a kernel function K, b_n is a bandwidth parameter, and $\hat{f}_\epsilon^{\mathrm{ft}}(t) = \frac{\hat{f}_X^{\mathrm{ft}}(t)}{\hat{f}^{\mathrm{ft}}(t)}$ with $\hat{f}_X^{\mathrm{ft}}(t) = \frac{1}{n} \sum_{j=1}^{n} e^{\mathrm{i}tX_j}$ and $\hat{f}^{\mathrm{ft}}(t) = \exp\left(\int_0^t \frac{\mathrm{i}\sum_{j=1}^{n} X_j e^{\mathrm{i}sW_j}}{\sum_{j=1}^{n} e^{\mathrm{i}sW_j}} ds\right)$. In Section 3, we further show that $\sqrt{n}(\hat{\theta}_1 - \theta_1)$

[1] Even with more periods of data, if we wish errors like ϵ_j and ν_j to keep sharing no common component, due to the structure of the permanent-transitory model of income dynamics, we can have no more than two noisy measures of X_j^*. Hence, the repeated measurements setting here is fundamental.

converges to a normal distribution, which facilitates statistical inference for θ_1. Given the estimator $\hat{\theta}_1$ along with its estimated standard error, we may conduct a formal statistical test of the implication (2.3) of the theory of Carroll (2009).

2.2 Average derivative estimator

To understand $\hat{\theta}_1$ defined in (2.4) in a general framework, we consider the estimation of $\theta_c = E[\{g'(X^*) - c\}f(X^*)]$ for a constant c. We set $c = 1$ for the test of (2.3), whereas one can set $c = 0$ if inference for the average derivative is the objective *per se*. Suppose that f is continuously differentiable and $\{g(x) - cx\}f^2(x) \to 0$ as $|x| \to \infty$. By integration by parts, θ_c can be expressed by

$$\theta_c = -2E[(Y - cX^*)f'(X^*)].$$

If X^* were directly observed, θ_c could be estimated by $\tilde{\theta}_c = -\frac{2}{n}\sum_{j=1}^{n}(Y_j - cX_j^*)\tilde{f}'(X_j^*)$, where $\tilde{f}(x)$ denotes the kernel density estimator of $f(x)$, and $\tilde{\theta}_c$ could be understood as the density-weighted average derivative estimator by Powell, Stock and Stoker (1989) with dependent variable $Y - cX^*$. In the case when X^* is unobserved, however, $\tilde{\theta}_c$ is infeasible.

Motivated by (2.1)–(2.2), suppose that we can observe two noisy measurements of X^*, denoted by X and W, which are generated by

$$X = X^* + \epsilon, \quad W = X^* + \nu, \tag{2.5}$$

where ϵ and ν are measurement errors associated with X and W, respectively. In particular, ϵ and ν have zero mean and are classical; that is, ϵ and ν are independent of X^*.[2] To construct an estimator of θ_c in this case, note that

$$\theta_c = -2E[(Y - cW)f'(X^*)] = -2\int h_c(x)f'(x)dx, \tag{2.6}$$

where $h_c(x) = \{g(x) - cx\}f(x)$. Let f_A denote the density of a random variable A, $a^{\mathrm{ft}}(t) = \int e^{\mathrm{i}tx}a(x)dx$ denote the Fourier transform of a function a, and $\{Y_j, X_j, W_j\}_{j=1}^{n}$ be

[2] We will invoke a weaker assumption than the full statistical independence formally in Section 3.

an i.i.d. sample of (Y, X, W).

If f_ϵ were known, h_c and f could be estimated using the deconvolution techniques by

$$\check{h}_c(x) = \frac{1}{nb_n}\sum_{j=1}^{n}\mathbb{K}\left(\frac{x - X_j}{b_n}\right)(Y_j - cW_j),$$

$$\check{f}(x) = \frac{1}{nb_n}\sum_{j=1}^{n}\mathbb{K}\left(\frac{x - X_j}{b_n}\right),$$

where $\mathbb{K}(u) = \frac{1}{2\pi}\int e^{-\mathrm{i}tu}\frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)}dt$ is the deconvolution kernel function based on the characteristic function f_ϵ^{ft} of the true measurement error ϵ. Hence, it would be natural to estimate θ_c by the following plug-in estimator

$$\begin{aligned}\check{\theta}_c &= -2\int \check{h}_c(x)\check{f}'(x)dx \\ &= -\frac{2}{n^2b_n^3}\sum_{j=1}^{n}\sum_{k=1}^{n}(Y_j - cW_j)\int \mathbb{K}\left(\frac{x - X_j}{b_n}\right)\mathbb{K}'\left(\frac{x - X_k}{b_n}\right)dx.\end{aligned}$$

Now, suppose that f_ϵ is unknown. Assume that f^{ft}, f_ϵ^{ft} and f_ν^{ft} do not vanish anywhere, according to Kotlarski's (1967) identity, $f^{\mathrm{ft}}(t) = \exp\left(\int_0^t \frac{\mathrm{i}E[Xe^{\mathrm{i}sW}]}{E[e^{\mathrm{i}sW}]}ds\right)$, which together with $f_\epsilon^{\mathrm{ft}}(t) = \frac{E[e^{\mathrm{i}tX}]}{f^{\mathrm{ft}}(t)}$ implies that $\hat{f}_\epsilon^{\mathrm{ft}}$ defined in Section 2.1 is a plug-in estimator of f_ϵ^{ft} based on the sample analogs of $E[e^{\mathrm{i}sX}]$ and $E[e^{\mathrm{i}tW}]$, and $E[Xe^{\mathrm{i}sW}]$. Therefore, to estimate θ_c when f_ϵ is unknown, it is natural to replace f_ϵ^{ft} in $\check{\theta}_c$ by $\hat{f}_\epsilon^{\mathrm{ft}}$, which gives

$$\hat{\theta}_c = -\frac{2}{n^2b_n^3}\sum_{j=1}^{n}\sum_{k=1}^{n}(Y_j - cW_j)\int \hat{\mathbb{K}}\left(\frac{x - X_j}{b_n}\right)\hat{\mathbb{K}}'\left(\frac{x - X_k}{b_n}\right)dx,$$

where $\hat{\mathbb{K}}(u) = \frac{1}{2\pi}\int e^{-\mathrm{i}tu}\frac{K^{\mathrm{ft}}(t)}{\hat{f}_\epsilon^{\mathrm{ft}}(t/b_n)}dt$, defined in Section 2.1, is the deconvolution kernel function based on the estimated measurement error characteristic function $\hat{f}_\epsilon^{\mathrm{ft}}$.

3 Main result

In this section, we presents a formal theory that provides the asymptotic validity of the test procedure proposed in Section 2. Specifically, we derive the asymptotic distribution for $\hat{\theta}_c$ and propose an estimator for its asymptotic variance. To this end, we make the following assumptions.

Assumption.

(1) *$\{Y_j, X_j, W_j\}_{j=1}^n$ is an i.i.d. sample of (Y, X, W), where (X, W) satisfies (2.5), $E[|X^*|^{2+\eta}] < \infty$ for some $\eta > 0$, and $E[Y^2] < \infty$. Measurement errors (ϵ, ν) are independent from X^* and satisfy $E[Y|X^*, \epsilon] = E[Y|X^*]$, $E[\epsilon|\nu] = 0$, $E[\nu|\epsilon] = 0$, $E[|\epsilon|^{2+\eta}] < \infty$, and $E[|\nu|^{2+\eta}] < \infty$. Characteristic functions f^{ft}, f_ϵ^{ft} and f_ν^{ft} do not vanish anywhere.*

(2) *h_c and f have α continuous, bounded and integrable derivatives, and satisfy*

$$|f^{(\alpha)}(x + \Delta x) - f^{(\alpha)}(x)| < m(x)|\Delta x|, \quad |h_c^{(\alpha)}(x + \Delta x) - h_c^{(\alpha)}(x)| < m(x)|\Delta x|$$

for some bounded and integrable function $m(x)$ with $E[|m(X)|^2(1+|Y-cW|)^2] < \infty$.

(3) *K is symmetric, differentiable, and$\int K(u)du = 1$, $\int u^l K(u)du = 0$ for $1 \leq l < \alpha$, and $\int u^\alpha K(u)du \neq 0$. Also K^{ft} is compactly supported on $[-1, 1]$ and bounded.*

(4) $\frac{n^{-1/2} b_n^{-2} \log(1/b_n)^2 \{\inf_{|t| \leq b_n^{-1}} |f_\epsilon^{\mathrm{ft}}(t)|\}^{-2} \{\inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|\}^{-2}}{\min\left\{ \{\inf_{|t| \leq b_n^{-1}} |f_\epsilon^{\mathrm{ft}}(t)|\}^2, \{\inf_{|t| \leq b_n^{-1}} |f_\nu^{\mathrm{ft}}(t)|\}^4 \{\inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|\}^2 b_n^2 \right\}} \to 0$ *and* $n^{1/2} b_n^\alpha \to 0$ *as* $n \to \infty$.

(5) *$Var[\xi_c(Y, X, W)] < \infty$ where*

$$\xi_c(y, x, w) = \frac{1}{\pi} \int \left\{ \begin{array}{c} \{\{h_c'\}^{\mathrm{ft}}(-t) - (y - cw)\{f'\}^{\mathrm{ft}}(-t)\} \frac{e^{itx}}{f_\epsilon^{\mathrm{ft}}(t)} \\ + \left\{ \begin{array}{c} \{f^{\mathrm{ft}}(t)\{h_c'\}^{\mathrm{ft}}(-t) - \{f'\}^{\mathrm{ft}}(-t) h_c^{\mathrm{ft}}(t)\} \\ \times \left\{ -\frac{e^{itx}}{f^{\mathrm{ft}}(t) f_\epsilon^{\mathrm{ft}}(t)} + \int_0^t \left(-\frac{\{f^{\mathrm{ft}}\}'(s)}{f^{\mathrm{ft}}(s)} + \mathrm{i}x \right) \frac{e^{isw}}{f^{\mathrm{ft}}(s) f_\nu^{\mathrm{ft}}(s)} ds \right\} \end{array} \right\} \end{array} \right\} dt.$$

Assumption (1) requires random sampling and imposes conditions on the distribution of (Y, X^*) and measurement errors (ϵ, ν). In particular, under the classical measurement error assumptions, $E[Y|X^*, \epsilon] = E[Y|X^*]$ and $E[\nu|\epsilon] = E[\nu]$ are imposed for the identification of $E[Y - cW|X^*]$[3], $E[\nu] = 0$ is imposed for $\theta_c = -2E[(Y - cW)f'(X^*)]$, and the

[3]The full independence between the measurement error in regressor and dependent variable has been commonly adopted in the literature of nonparametric regression with errors-in-variables, (e.g., Fan and Truong, 1993; Delaigle and Meister, 2007; Meister, 2009). In our setting, however, it is sufficient to require the mean independence. In particular, observe that the identification of $E[Y - cW|X^*]$ hinges on

$$\{E[Y - cW|X] f_X\}(x) = \int \{E[Y - cW|X^*] f\}(x - e) f_\epsilon(e) de,$$

which holds if $E[Y - cW|X^*, \epsilon] = E[Y - cW|X^*]$, or equivalently $E[Y|X^*, \epsilon] = E[Y|X^*]$ and $E[\nu|\epsilon] = E[\nu]$, under the classical measurement error assumption.

non-vanishing characteristic functions and $E[\epsilon|\nu] = 0$ is imposed for the Kotlarski's identity[4]. $E[|X^*|^{2+\eta}] < \infty$ and $E[|\epsilon|^{2+\eta}] < \infty$ are regularity conditions required by Lemma 1, which is used to characterize the uniform convergence rate of the empirical characteristic function of (X, W) and their first-order derivatives over an expanding region. We remark in the context of (2.1)–(2.2) that our conditions, $E[\epsilon|\nu] = 0$ and $E[\nu|\epsilon] = 0$, do not rule out the typical assumptions about the permanent-transitory models of income and earnings dynamics in which the permanent shocks $\{\eta_{jt}\}_t$ and the transitory components $\{\tau_{jt}\}_t$ are white noise processes (e.g., Bonhomme and Robin, 2010).

Assumption (2) constitutes mild assumptions on the smoothness of the density function f and the regression function g, which is equivalent to Assumptions 3 and 5 in Powell, Stock and Stoker (1989). Assumption (3) concerns the kernel function K. Specifically, we require K to be a symmetric α-th order kernel, which together with Assumption (2) can be used to control the magnitude of the estimation bias. In addition, we also require K^{ft} to be compactly supported, which is to regularize the deconvolution problem that is well-known to be ill-posed.

Assumption (4) gives two conditions on the bandwidth b_n. In particular, the first condition is needed to control the estimation variance, and the second condition requires g and f to be smooth enough so that the estimation bias is asymptotically negligible. Here, we maintain a general expression without specifying the decay rates of the tails of f^{ft}, f_ϵ^{ft} and f_ν^{ft} as is typical in the deconvolution literature. By doing so, we can apply our result to a larger set of measurement error distributions, including both ordinary smooth distributions and supersmooth distributions. Assumption (5) is a high-level assumption on the boundedness of the asymptotic variance of $\hat{\theta}_c$; an analogous assumption is made in Fan (1995) and Dong, Otsu and Taylor (2021).

Theorem. *Under Assumptions (1) - (5), we have*

$$\sqrt{n}\{\hat{\theta}_c - \theta_c\} \xrightarrow{d} N(0, Var[\xi_c(Y, X, W)]).$$

This is the main result of the paper. To understand it, we can decompose ξ_c into two

[4]The full independence between ϵ and ν has been commonly adopted in the literature on identification and estimation based on Kotlarski's identity (e.g., Li and Vuong, 1998; Kato, Sasaki and Ura, 2021; Dong, Otsu and Taylor, 2022), but it can be relaxed to the mean independence; See Schennach(2004).

parts as $\xi_c(Y,X,W)=\xi_{c,1}(Y,X,W)+\xi_{c,2}(Y,X,W)$, where

$$\xi_{c,1}(Y,X,W)=\frac{1}{\pi}\int\left\{\{h_c'\}^{\mathrm{ft}}(-t)-(Y-cW)\{f'\}^{\mathrm{ft}}(-t)\right\}\frac{e^{\mathrm{i}tX}}{f_\epsilon^{\mathrm{ft}}(t)}dt,$$

$$\xi_{c,2}(Y,X,W)=\frac{1}{\pi}\int\left\{\begin{array}{c}\{\{h_c'\}^{\mathrm{ft}}(-t)f^{\mathrm{ft}}(t)-h_c^{\mathrm{ft}}(t)\{f'\}^{\mathrm{ft}}(-t)\}\\ \times\left\{-\frac{e^{\mathrm{i}tX}}{f^{\mathrm{ft}}(t)f_\epsilon^{\mathrm{ft}}(t)}+\int_0^t\left(-\frac{\{f^{\mathrm{ft}}\}'(s)}{f^{\mathrm{ft}}(s)}+\mathrm{i}X\right)\frac{e^{\mathrm{i}sW}}{f^{\mathrm{ft}}(s)f_\nu^{\mathrm{ft}}(s)}ds\right\}\end{array}\right\}dt,$$

and consider, for ease of illustration, a special case in which f_ϵ^{ft} is of the form

$$f_\epsilon^{\mathrm{ft}}(t)=\frac{1}{c_0+c_1t+\cdots+c_\beta t^\beta},$$

where β is a positive integer, and $c_0=1,c_1,\ldots,c_\beta$ are complex numbers, which includes the Laplace distribution as a special case when $c_1=0$ and $\beta=2$.

In such cases, using $\{a^{(k)}\}^{\mathrm{ft}}(t)=(-\mathrm{i}t)^k a^{\mathrm{ft}}(t)$ for a positive integer k, we obtain that

$$\begin{aligned}\xi_{c,1}(Y,X,W)=&\frac{1}{\pi}\int\left\{\{h_c'\}^{\mathrm{ft}}(-t)-(Y-cW)\{f'\}^{\mathrm{ft}}(-t)\right\}\left\{c_0+c_1t+\cdots+c_\beta t^\beta\right\}e^{\mathrm{i}tX}dt\\ =&\sum_{k=0}^{\beta}\frac{c_k}{\pi\mathrm{i}^k}\int\left\{\underbrace{(\mathrm{i}t)^{k+1}h_c^{\mathrm{ft}}(-t)}_{\{h_c^{(k+1)}\}^{\mathrm{ft}}(-t)}-(Y-cW)\underbrace{(\mathrm{i}t)^k\{f'\}^{\mathrm{ft}}(-t)}_{\{f^{(k+1)}\}^{\mathrm{ft}}(-t)}\right\}e^{\mathrm{i}tX}dt\\ =&\sum_{k=0}^{\beta}(-\mathrm{i})^k2c_k\left\{(Y-cW)f^{(k+1)}(X)-h_c^{(k+1)}(X)\right\},\end{aligned}$$

which when $c=0$ coincides with $2r(X,Y)$ defined as in equation (25) of Fan (1995). Thus, $\xi_{c,1}(Y,X,W)$ characterizes the randomness of $\check{\theta}_c$, which is the estimator of θ_c when f_ϵ is known. Furthermore, compared to $2r(X,Y)$ in Fan (1995), $\xi_{c,1}(Y,X,W)$ allows non-zero value of c and can cover a larger set of measurement error distributions.

Since $\xi_{c,1}(Y,X,W)$ characterizes the randomness in the estimation of θ_c when f_ϵ is known, the additional randomness introduced by using $\hat{f}_\epsilon^{\mathrm{ft}}$ in the place of f_ϵ^{ft} is reflected by $\xi_{c,2}(Y,X,W)$. It is worthy to note that the structure of $\xi_{c,2}(Y,X,W)$ is similar to that of $\xi_{c,1}(Y,X,W)$, but is more complicated. In general, it is difficult to simplify $\xi_{c,2}(Y,X,W)$ as we did for $\xi_{c,1}(Y,X,W)$ even when f^{ft}, f_ϵ^{ft} and f_ν^{ft} are all specified. However, there are special cases in which $\xi_{c,2}(Y,X,W)$ can be completely ignored. To see a case in point, note that $\{h_c'\}^{\mathrm{ft}}(-t)f^{\mathrm{ft}}(t)-h_c^{\mathrm{ft}}(t)\{f'\}^{\mathrm{ft}}(-t)=\mathrm{i}t\{h_c^{\mathrm{ft}}(-t)f^{\mathrm{ft}}(t)-h_c^{\mathrm{ft}}(t)f^{\mathrm{ft}}(-t)\}$, which implies that if $h_c^{\mathrm{ft}}(-t)=h_c^{\mathrm{ft}}(t)$ and $f^{\mathrm{ft}}(-t)=f^{\mathrm{ft}}(t)$, for example when both h_c and f are symmetric around zero, $\xi_{c,2}(Y,X,W)=0$, i.e. the estimation error brought by using $\hat{f}_\epsilon^{\mathrm{ft}}$ in the place of f_ϵ^{ft} is exactly zero and $\hat{\theta}_c$ has exactly the same asymptotic distribution as that of $\check{\theta}_c$.

To test hypotheses on θ_c like H_0, besides the asymptotic distribution of $\hat{\theta}_c$, we also need to estimate $s_c^2 = Var[\xi_c(Y,X,W)]$. Observe that we can rewrite

$$
\begin{aligned}
\xi_c(y,x,w) &= \frac{1}{\pi}\int \mathrm{i}t \left\{ \begin{array}{c} \{h_c^{\mathrm{ft}}(-t)-(y-cw)f^{\mathrm{ft}}(-t)\}\frac{e^{\mathrm{i}tx}}{f_\epsilon^{\mathrm{ft}}(t)} \\ +\left\{ \begin{array}{c} \{f^{\mathrm{ft}}(t)h_c^{\mathrm{ft}}(-t)-f^{\mathrm{ft}}(-t)h_c^{\mathrm{ft}}(t)\} \\ \times\left\{-\frac{e^{\mathrm{i}tx}}{f^{\mathrm{ft}}(t)f_\epsilon^{\mathrm{ft}}(t)}+\int_0^t\left(-\frac{\{f^{\mathrm{ft}}\}'(s)}{f^{\mathrm{ft}}(s)}+\mathrm{i}x\right)\frac{e^{\mathrm{i}sw}}{f^{\mathrm{ft}}(s)f_\nu^{\mathrm{ft}}(s)}ds\right\} \end{array}\right\} \end{array}\right\} dt \\
&= \frac{1}{\pi}\int \mathrm{i}t \left\{ \begin{array}{c} \{h_c^{\mathrm{ft}}(-t)-(y-cw)f^{\mathrm{ft}}(-t)\}\frac{e^{\mathrm{i}tx}}{f_\epsilon^{\mathrm{ft}}(t)} \\ +\left\{ \begin{array}{c} \{f^{\mathrm{ft}}(t)h_c^{\mathrm{ft}}(-t)-f^{\mathrm{ft}}(-t)h_c^{\mathrm{ft}}(t)\} \\ \times\left\{-\frac{e^{\mathrm{i}tx}}{f_X^{\mathrm{ft}}(t)}+\int_0^t\left(-\frac{\mathrm{i}E[Xe^{\mathrm{i}sW}]}{f_W^{\mathrm{ft}}(s)}+\mathrm{i}x\right)\frac{e^{\mathrm{i}sw}}{f_W^{\mathrm{ft}}(s)}ds\right\} \end{array}\right\} \end{array}\right\} dt,
\end{aligned}
$$

where the first step uses $\{a^{(k)}\}^{\mathrm{ft}}(t)=(-\mathrm{i}t)^k a^{\mathrm{ft}}(t)$ and the second step follows from $f_X^{\mathrm{ft}} = f^{\mathrm{ft}}f_\epsilon^{\mathrm{ft}}$, $f_W^{\mathrm{ft}} = f^{\mathrm{ft}}f_\nu^{\mathrm{ft}}$ and $f^{\mathrm{ft}}(t)=\exp\left(\int_0^t \frac{\mathrm{i}E[Xe^{\mathrm{i}sW}]}{E[e^{\mathrm{i}sW}]}ds\right)$, which implies $E[\xi_c(Y,X,W)]=0$ and $s_c^2 = E[\xi_c^2(Y,X,W)]$. In practice,$f^{\mathrm{ft}}$, f_ϵ^{ft}, f_X^{ft}, f_W^{ft}, $E[Xe^{\mathrm{i}tW}]$ and h_c^{ft} are all unknown, and we have to estimate them. f^{ft}, f_ϵ^{ft} and f_X^{ft} can be estimated by $\hat{f}^{\mathrm{ft}}$, $\hat{f}_\epsilon^{\mathrm{ft}}$, and f_X^{ft} defined in Section 2, f_W^{ft} can be estimated by $\hat{f}_W^{\mathrm{ft}}(t)=\frac{1}{n}\sum_{j=1}^n e^{\mathrm{i}tW_j}$, $E[Xe^{\mathrm{i}tW}]$ can be estimated by $\hat{E}[Xe^{\mathrm{i}tW}]=\frac{1}{n}\sum_{j=1}^n X_j e^{\mathrm{i}sW_j}$, and h_c^{ft} can be estimated by $\hat{h}_c^{\mathrm{ft}}(t)=\frac{\frac{1}{n}\sum_{j=1}^n (Y_j-cW_j)e^{\mathrm{i}tX_j}}{\hat{f}_\epsilon^{\mathrm{ft}}(t)}$. Therefore, we can estimate s_c^2 by $\hat{s}_c^2=\frac{1}{n}\sum_{j=1}^n \hat{\xi}_c^2(Y_j,X_j,W_j)$ with

$$
\hat{\xi}_c(y,x,w) = \frac{1}{\pi}\int \mathrm{i}t \left\{ \begin{array}{c} \{\hat{h}_c^{\mathrm{ft}}(-t)-(y-cw)\hat{f}^{\mathrm{ft}}(-t)\}\frac{e^{\mathrm{i}tx}}{\hat{f}_\epsilon^{\mathrm{ft}}(t)} \\ +\left\{ \begin{array}{c} \{\hat{f}^{\mathrm{ft}}(t)\hat{h}_c^{\mathrm{ft}}(-t)-\hat{f}^{\mathrm{ft}}(-t)\hat{h}_c^{\mathrm{ft}}(t)\} \\ \times\left\{-\frac{e^{\mathrm{i}tx}}{\hat{f}_X^{\mathrm{ft}}(t)}+\int_0^t\left(-\frac{\mathrm{i}\hat{E}[Xe^{\mathrm{i}sW}]}{\hat{f}_W^{\mathrm{ft}}(s)}+\mathrm{i}x\right)\frac{e^{\mathrm{i}sw}}{\hat{f}_W^{\mathrm{ft}}(s)}ds\right\} \end{array}\right\} \end{array}\right\} K^{\mathrm{ft}}(tb_n)dt,
$$

where $K^{\mathrm{ft}}(tb_n)$ is introduced to regularize the integration.

4 Simulation

This section presents simulation studies to analyze the finite sample performance of the proposed method of inference about θ_c. We generate N independent copies of the observed variables (Y, X, U) via the structural equations

$$
\begin{aligned}
Y &= f(X^*) + U = X^* - \delta X^* + U, \\
X &= X^* + \epsilon, \qquad \text{and} \\
W &= X^* + \nu,
\end{aligned}
$$

where the latent variables (X^*, U, ϵ, ν) are in turn generated independently from the standard normal distribution. Note in this setting that the null hypothesis $H_0 : \theta_1 \geq 0$ is true if and only if $\delta \leq 0$. Furthermore, positive values of the design parameter δ in this data generating process measures deviations from the null hypothesis. We run sets of simulations across combinations of the values of $\delta \in [0.0, 0.5]$ and $N \in \{250, 500\}$, where each set of simulations consists of 2,500 Monte Carlo iterations.

[FIGURE 1 HERE]

Figure 1 plots the Monte Carlo frequencies of rejecting the null hypothesis $H_0 : \theta_1 \geq 0$ against the alternative $H_1 : \theta_1 < 0$ based on the one-sided test with our estimator $\hat{\theta}_1$ and its standard error estimator $\hat{s}_1$. The nominal size of the test is set to 0.05. The horizontal axis of the figure measures the deviation $\delta \in [0.0, 0.5]$ away from the null hypothesis H_0. The dashed (respectively, solid) curve indicates the results with $N = 250$ (respectively, 500). Observe that the rejection frequency at $\delta = 0$ is close to the nominal size, 0.05. As δ becomes larger, on the other hand, the rejection frequencies increase. For a given value of $\delta > 0$, the rejection frequency is larger for the larger sample size, demonstrating the power of the test as well as the size control.

We ran additional simulations with alternative data generating designs, only to find very similar simulation results to the baseline results presented above. Overall, the simulation outcomes demonstrate excellent small-sample performance of the estimation and inference methods. Our observations that the asymptotic approximations are already accurate even at such small sample sizes N as 250 demonstrate the practical merit of our root-n consistent test under the highly sophisticated problem of errors-in-variables nonparametric regressions.

5 Application

This section revisits the analysis of the MPCP introduced in Sections 1–2. Recall from (2.1) and (2.2) that we require four time periods of panel data $\left\{\{\iota_{j\tau}\}_{\tau=t-2}^{t+1}\right\}_{j=1}^{n}$ on observed income to construct the variables $\{(X_j, W_j)\}_{j=1}^{n}$ that we use as inputs for our proposed method of inference. Using the U.S. Panel Study of Income Dynamics (PSID) for the four most recent survey years 2013, 2015, 2017, and 2019, we aim to test the null hypothesis $H_0 : \theta_1 \geq 0$ of (super-) unit MPCP against the alternative hypothesis $H_1 : \theta_1 < 0$ of sub-unit MPCP. Rejecting the null hypothesis H_0 supports the buffer-stock model that arises with inpatient consumers having a standard precautionary saving motive.

The income variable ι_{jt} is defined by the log total family income of unit j reported in year t. The consumption variable C_{jt} is similarly defined by the log family expenditures of unit j reported in year t, where the categories of consumption consist of food, housing, telephone/internet, transportation, vehicle, education, child care, health care, household repairs, household furnishing, clothing, and recreation. The first three columns of Table 1 display summary statistics of this data set. Displayed values are the sample means. Parentheses enclose sample standard deviations.

[TABLE 1 HERE]

Adapting (2.1) and (2.2) to this panel data set, we construct $X_j = \iota_{j2017} - \iota_{j2013}$ and $W_j = \iota_{j2019} - \iota_{j2015}$. Likewise, we construct the outcome variable by $Y_j = C_{j2017} - C_{j2015}$. We drop units that experience a missing value or an infinite value for X_j, W_j or Y_j. Consequently, we obtain a balanced panel of 5976 household units. The last three columns of Table 1 display summary statistics of the constructed variables $\{(X_j, W_j, Y_j)\}_{j=1}^{n}$. Again, displayed values are the sample means, and parentheses enclose sample standard deviations.

Applying our proposed method of estimation and inference, we obtain the point estimate of $\hat{\theta}_1 = -0.0607$ with the estimated standard error of $\hat{s}_1 = 0.0052$. Using our asymptotic normality results along with these estimates, we formally reject the one-sided test of the null hypothesis $H_0 : \theta_1 \geq 0$ of (super-) unit MPCP in favor of the alternative hypothesis $H_1 : \theta_1 < 0$ of sub-unit MPCP. Our test result supports the buffer-stock model that arises with inpatient consumers having a standard precautionary saving motive. Even though a number of prior studies have calibrated or estimated the MPCP under various models of income dynamics, to our best knowledge, our result is the first formal statistical inference result about the MPCP using flexible nonparametric distribution in the permanent-transitory model of income processes.

6 Conclusion

In this paper, we propose a density-weighted average derivative estimator when two noisy measures of a latent regressor is available. Both measures have classical errors, and the error distributions are possibly asymmetric. We show that this estimator achieves the root-n rate of convergence and is asymptotically normal. Simulation studies demonstrate excellent small-sample performance, and support the merit of the root-n asymptotic normality. Based on the proposed estimator, we construct a test on the sub-unity of MPCP under a nonparametric consumption model. In particular, under a permanent-transitory model of income dynamics, we construct two noisy measures of a permanent income shock using four periods data. With an application using recent waves of the U.S. PSID, we reject the null hypothesis of unit MPCP in favor or a sub-unit MPCP, supporting the buffer-stock model of saving.

A Proof of the theorem

Let $\hat{\mu}_\iota(t) = \frac{1}{n}\sum_{l=1}^n \mu_{\iota,l}(t)$ and $\mu_\iota(t) = E[\mu_{\iota,1}(t)]$ for $\iota = 1,2,3$ with $\mu_{1,l}(t) = e^{\mathrm{i}tX_l}$, $\mu_{2,l}(t) = e^{\mathrm{i}tW_l}$, and $\mu_{3,l}(t) = X_l e^{\mathrm{i}tW_l}$. Then, $\hat{f}_\epsilon^{\mathrm{ft}}(t) = \hat{\mu}_1(t)\exp\left(-\int_0^t \frac{\mathrm{i}\hat{\mu}_3(s)}{\hat{\mu}_2(s)}ds\right)$ and $f_\epsilon(t) = \mu_1(t)\exp\left(-\int_0^t \frac{\mathrm{i}\mu_3(s)}{\mu_2(s)}ds\right)$. By expanding $(\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3)$ around (μ_1, μ_2, μ_3), we obtain

$$\hat{\mathbb{K}}(u) = \mathbb{K}(u) + \mathbb{A}(u) + \mathbb{R}(u), \tag{A.1}$$

where $\mathbb{A}(u) = \frac{1}{2\pi}\int e^{-\mathrm{i}tu}\frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)}\hat{\Pi}(t/b_n)dt$ and $\mathbb{R}(u) = \frac{1}{2\pi}\int e^{-\mathrm{i}tu}\frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)}\hat{\Pi}^{\mathrm{res}}(t/b_n)dt$ with

$$\hat{\Pi}(t) = \frac{1}{n}\sum_{l=1}^n \Pi_l(t), \quad \Pi_l(t) = -\frac{\delta_{1,l}(t)}{\mu_1(t)} + \mathrm{i}\int_0^t \left\{-\frac{\mu_3(s)\delta_{2,l}(s)}{\mu_2^2(s)} + \frac{\delta_{3,l}(s)}{\mu_2(s)}\right\}ds,$$

$$\hat{\Pi}^{\mathrm{res}}(t) = \frac{\hat{\delta}_1^2(t)}{\mu_1(t)+\hat{\delta}_1(t)} - \int_0^t \mathrm{i}\left\{-\frac{\mu_3(s)\hat{\delta}_2(s)}{\mu_2^2(s)} + \frac{\hat{\delta}_3(s)}{\mu_2(s)}\right\}\frac{\hat{\delta}_2(s)}{\mu_2(s)+\hat{\delta}_2(s)}ds$$

$$+\int_0^t \mathrm{i}\left\{-\frac{\mu_3(s)\hat{\delta}_2(s)}{\mu_2^2(s)} + \frac{\hat{\delta}_3(s)}{\mu_2(s)}\right\}\left\{1 - \frac{\hat{\delta}_2(s)}{\mu_2(s)+\hat{\delta}_2(s)}\right\}ds\left\{-\frac{\hat{\delta}_1(t)}{\mu_1(t)} + \frac{\hat{\delta}_1^2(t)}{\mu_1(t)+\hat{\delta}_1(t)}\right\}$$

$$-\frac{1}{2}e^{\bar{\phi}(t)}\left(\int_0^t \left\{-\frac{\mu_3(s)\hat{\delta}_2(s)}{\mu_2^2(s)} + \frac{\hat{\delta}_3(s)}{\mu_2(s)}\right\}\left\{1 - \frac{\hat{\delta}_2(s)}{\mu_2(s)+\hat{\delta}_2(s)}\right\}ds\right)^2\left\{1 - \frac{\hat{\delta}_1(t)}{\mu_1(t)} + \frac{\hat{\delta}_1^2(t)}{\mu_1(t)+\hat{\delta}_1(t)}\right\},$$

for some $|\bar{\phi}(t)| \le \left|\int_0^t \left\{-\frac{\mu_3(s)\hat{\delta}_2(s)}{\mu_2^2(s)} + \frac{\hat{\delta}_3(s)}{\mu_2(s)}\right\}\left\{1 - \frac{\hat{\delta}_2(s)}{\mu_2(s)+\hat{\delta}_2(s)}\right\}ds\right|$, where $\hat{\delta}_\iota(t) = \frac{1}{n}\sum_{l=1}^n \delta_{\iota,l}(t)$ with $\delta_{\iota,l}(t) = \mu_{\iota,l}(t) - \mu_\iota(t)$ for $\iota = 1,2,3$. Here, $\mathbb{A}$ denotes the Fréchet derivative of $\hat{\mathbb{K}}$ as a functional of $(\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3)$ evaluated at (μ_1, μ_2, μ_3) in the direction of $(\hat{\delta}_1, \hat{\delta}_2, \hat{\delta}_3)$, and $\mathbb{R}$ contains the remainders. Observe that (A.1) implies

$$\hat{\theta}_c = \frac{2}{n^2}\sum_{j=1}^n\sum_{k=1}^n(-1)b_n^{-3}(Y_j - cW_j)\int \hat{\mathbb{K}}\left(\frac{x-X_j}{b_n}\right)\hat{\mathbb{K}}'\left(\frac{x-X_k}{b_n}\right)dx$$

$$= \underbrace{\frac{2}{n^2}\sum_{j=1}^n\sum_{k=1}^n(-1)b_n^{-3}(Y_j - cW_j)\int \left\{\begin{array}{c} \mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}'\left(\frac{x-X_k}{b_n}\right) \\ +\mathbb{A}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}'\left(\frac{x-X_k}{b_n}\right) + \mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{A}'\left(\frac{x-X_k}{b_n}\right)\end{array}\right\}dx}_{=:S}$$

$$+ \underbrace{\frac{2}{n^2}\sum_{j=1}^n\sum_{k=1}^n(-1)b_n^{-3}(Y_j - cW_j)\int \left\{\begin{array}{c} \mathbb{A}\left(\frac{x-X_j}{b_n}\right)\mathbb{A}'\left(\frac{x-X_k}{b_n}\right) \\ +\mathbb{R}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}'\left(\frac{x-X_k}{b_n}\right) + \mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{R}'\left(\frac{x-X_k}{b_n}\right) \\ +\mathbb{R}\left(\frac{x-X_j}{b_n}\right)\mathbb{A}'\left(\frac{x-X_k}{b_n}\right) + \mathbb{A}\left(\frac{x-X_j}{b_n}\right)\mathbb{R}'\left(\frac{x-X_k}{b_n}\right) \\ +\mathbb{R}\left(\frac{x-X_j}{b_n}\right)\mathbb{R}'\left(\frac{x-X_k}{b_n}\right)\end{array}\right\}dx}_{=:T}$$

First, we are going to show

$$T = o_p(n^{-1/2}). \tag{A.2}$$

To show (A.2), decompose $T = T_1 + T_2 + T_3 + T_4$, where

$$\begin{aligned}
T_1 &= \frac{-2}{n^2 b_n^3}\sum_{j=1}^n\sum_{k=1}^n (Y_j - cW_j)\int \mathbb{A}\left(\frac{x-X_j}{b_n}\right)\mathbb{A}'\left(\frac{x-X_k}{b_n}\right)dx,\\
T_2 &= \frac{-2}{n^2 b_n^3}\sum_{j=1}^n\sum_{k=1}^n (Y_j - cW_j)\int \left\{\mathbb{R}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}'\left(\frac{x-X_k}{b_n}\right) + \mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{R}'\left(\frac{x-X_k}{b_n}\right)\right\}dx,\\
T_3 &= \frac{-2}{n^2 b_n^3}\sum_{j=1}^n\sum_{k=1}^n (Y_j - cW_j)\int \left\{\mathbb{R}\left(\frac{x-X_j}{b_n}\right)\mathbb{A}'\left(\frac{x-X_k}{b_n}\right) + \mathbb{A}\left(\frac{x-X_j}{b_n}\right)\mathbb{R}'\left(\frac{x-X_k}{b_n}\right)\right\}dx,\\
T_4 &= \frac{-2}{n^2 b_n^3}\sum_{j=1}^n\sum_{k=1}^n (Y_j - cW_j)\int \mathbb{R}\left(\frac{x-X_j}{b_n}\right)\mathbb{R}'\left(\frac{x-X_k}{b_n}\right)dx.
\end{aligned}$$

For T_1, we have

$$\begin{aligned}
|T_1| &= \frac{1}{\pi n^2 b_n^3}\left|\sum_{j=1}^n\sum_{k=1}^n (Y_k - cW_k)\iint\left\{\begin{array}{c}\frac{1}{2\pi}\int e^{-\mathrm{i}(t_1+t_2)x/b_n}dx\, t_1 e^{\mathrm{i}\left(\frac{t_1X_j+t_2X_k}{b_n}\right)}\\ \times\frac{K^{\mathrm{ft}}(t_1)K^{\mathrm{ft}}(t_2)}{f_\epsilon^{\mathrm{ft}}(t_1/b_n)f_\epsilon^{\mathrm{ft}}(t_2/b_n)}\hat{\Pi}(t_1/b_n)\hat{\Pi}(t_2/b_n)\end{array}\right\}dt_1dt_2\right|\\
&= \frac{1}{\pi n^2 b_n^2}\left|\sum_{j=1}^n\sum_{k=1}^n (Y_k - cW_k)\iint\left\{\begin{array}{c}\frac{1}{2\pi}\int e^{-\mathrm{i}(t_1+t_2)\tilde{x}}d\tilde{x}\, t_1 e^{\mathrm{i}\left(\frac{t_1X_j+t_2X_k}{b_n}\right)}\\ \times\frac{K^{\mathrm{ft}}(t_1)K^{\mathrm{ft}}(t_2)}{f_\epsilon^{\mathrm{ft}}(t_1/b_n)f_\epsilon^{\mathrm{ft}}(t_2/b_n)}\hat{\Pi}(t_1/b_n)\hat{\Pi}(t_2/b_n)\end{array}\right\}dt_1dt_2\right|\\
&= \frac{1}{\pi n^2 b_n^2}\left|\sum_{j=1}^n\sum_{k=1}^n (Y_k - cW_k)\int t e^{\mathrm{i}t\left(\frac{X_j-X_k}{b_n}\right)}\frac{|K^{\mathrm{ft}}(t)|^2}{|f_\epsilon^{\mathrm{ft}}(t/b_n)|^2}|\hat{\Pi}(t/b_n)|^2dt\right|\\
&= O_p\left(\frac{\{\sup_{|t|\le b_n^{-1}}|\hat{\Pi}(t)|\}^2}{b_n^2\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2}\right),
\end{aligned}$$

where the second step follows from the change of variables $\tilde{x} = x/b_n$, the third step follows by $\int \delta(t-s)f(t)dt = f(s)$ with Dirac delta function $\delta(t) = \frac{1}{2\pi}\int e^{-\mathrm{i}tx}dx$, and the last step uses the implication that K^{ft} is supported on $[-1,1]$ under Assumption (3). Similar arguments show

$$\begin{aligned}
|T_2| &= O_p\left(\frac{\sup_{|t|\le b_n^{-1}}|\hat{\Pi}^{\mathrm{res}}(t)|}{b_n^2\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2}\right),\\
|T_3| &= O_p\left(\frac{\sup_{|t|\le b_n^{-1}}|\hat{\Pi}(t)|\sup_{|t|\le b_n^{-1}}|\hat{\Pi}^{\mathrm{res}}(t)|}{b_n^2\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2}\right),\\
|T_4| &= O_p\left(\frac{\{\sup_{|t|\le b_n^{-1}}|\hat{\Pi}^{\mathrm{res}}(t)|\}^2}{b_n^2\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2}\right),
\end{aligned}$$

and thus (A.2) follows by Lemma 2 and Assumption (4).

Hence, for the asymptotic distribution of $\hat{\theta}_c$, it is sufficient to focus on S. Let $a_j = (Y_j, X_j, W_j)$ and let $\mathrm{Sym}(\mathcal{I})$ denote the collection of all permutations of an ordered set $\mathcal{I}$. Observe that

$$
\begin{aligned}
S &= \frac{2}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}(-1)b_n^{-3}(Y_j - W_j)\int\left\{\begin{array}{c}\mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}'\left(\frac{x-X_k}{b_n}\right)\\ +\mathbb{A}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}'\left(\frac{x-X_k}{b_n}\right)+\mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{A}'\left(\frac{x-X_k}{b_n}\right)\end{array}\right\}dx\\
&= \frac{2}{n^3}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\underbrace{(-1)b_n^{-3}(Y_j - W_j)\int\left\{\begin{array}{c}\mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}'\left(\frac{x-X_k}{b_n}\right)\\ +\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)}\frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)}\Pi_l(t/b_n)dt\right\}\mathbb{K}'\left(\frac{x-X_k}{b_n}\right)\\ +\mathbb{K}\left(\frac{x-X_j}{b_n}\right)\left\{\frac{-\mathrm{i}}{2\pi}\int te^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)}\frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)}\Pi_l(t/b_n)dt\right\}\end{array}\right\}dx}_{=:q_n(a_j,a_k,a_l)}\\
&= \frac{(n-1)(n-2)}{n^2}\underbrace{\binom{n}{3}^{-1}\sum_{j=1}^{n-2}\sum_{k=j+1}^{n-1}\sum_{l=k+1}^{n}\overbrace{\sum_{(j',k',l')\in\mathrm{Sym}((j,k,l))}q_n(a_j,a_k,a_l)/3}^{=:p_n(a_j,a_k,a_l)}-E[p_n(a_j,a_k,a_l)]}_{=:U_n}\\
&\quad+\underbrace{\frac{(n-1)(n-2)}{n^2}E[p_n(a_1,a_2,a_3)]}_{=:B_n}\\
&\quad+\underbrace{\frac{2}{n^3}\left\{\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}\left\{\begin{array}{c}q_n(a_j,a_j,a_k)+q_n(a_k,a_k,a_j)\\ +q_n(a_j,a_k,a_j)+q_n(a_k,a_j,a_k)\\ +q_n(a_j,a_k,a_k)+q_n(a_k,a_j,a_j)\end{array}\right\}+\sum_{j=1}^{n}q_n(a_j,a_j,a_j)\right\}}_{=:R_n},
\end{aligned}
$$

where U_n is a 3rd order U-statistic with symmetric kernel p_n, and its Hájek projection is given by

$$
U_n^* = \frac{3}{n}\sum_{j=1}^{n}\{r_n(a_j) - E[r_n(a_j)]\},
$$

where $r_n(a_j) = E_j[p_n(a_j, a_k, a_l)]$ and $E_j[\cdot] = E[\cdot|a_j]$. Then, we can write

$$
S - \theta_c = U_n^*(1-1/n)(1-2/n) + \{U_n - U_n^*\}(1-1/n)(1-2/n) + \{B_n - \theta_c\} + R_n. \quad \text{(A.3)}
$$

First, we are going to show

$$
R_n = o_p(n^{-1/2}). \quad \text{(A.4)}
$$

To show (A.4), decompose $R_n = R_{n,1} + R_{n,2} + R_{n,3}$, where

$$
\begin{aligned}
R_{n,1} &= \frac{2}{n^3}\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}\{q_n(a_j,a_j,a_k)+q_n(a_k,a_k,a_j)+q_n(a_j,a_k,a_k)+q_n(a_k,a_j,a_j)\}\\
R_{n,2} &= \frac{2}{n^3}\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}\{q_n(a_j,a_k,a_j)+q_n(a_k,a_j,a_k)\},\\
R_{n,3} &= \frac{2}{n^3}\sum_{j=1}^{n} q_n(a_j,a_j,a_j).
\end{aligned}
$$

Observe that

$$
\begin{aligned}
q_n(d_j,d_k,d_l) &= \frac{\mathrm{i}(Y_j - cW_j)}{2\pi b_n^3}\iint\left\{\begin{array}{c}\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}(t_1+t_2)x/b_n}dx\right\}t_2 e^{\mathrm{i}\left(\frac{t_1X_j+t_2X_k}{b_n}\right)}\\ \times\frac{K^{\mathrm{ft}}(t_1)K^{\mathrm{ft}}(t_2)}{f_\epsilon^{\mathrm{ft}}(t_1/b_n)f_\epsilon^{\mathrm{ft}}(t_2/b_n)}\{1+\Pi_l(t_1/b_n)+\Pi_l(t_2/b_n)\}\end{array}\right\}dt_1dt_2\\
&= \frac{\mathrm{i}(Y_j - cW_j)}{2\pi b_n^2}\iint\left\{\begin{array}{c}\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}(t_1+t_2)\tilde{x}}d\tilde{x}\right\}t_2 e^{\mathrm{i}\left(\frac{t_1X_j+t_2X_k}{b_n}\right)}\\ \times\frac{K^{\mathrm{ft}}(t_1)K^{\mathrm{ft}}(t_2)}{f_\epsilon^{\mathrm{ft}}(t_1/b_n)f_\epsilon^{\mathrm{ft}}(t_2/b_n)}\{1+\Pi_l(t_1/b_n)+\Pi_l(t_2/b_n)\}\end{array}\right\}dt_1dt_2\\
&= \frac{\mathrm{i}(cW_j - Y_j)}{2\pi b_n^2}\int te^{\mathrm{i}t\left(\frac{X_j-X_k}{b_n}\right)}\frac{|K^{\mathrm{ft}}(t)|^2}{|f_\epsilon^{\mathrm{ft}}(t/b_n)|^2}\{1+\Pi_l(t/b_n)+\Pi_l(-t/b_n)\}\,dt,
\end{aligned}
$$

where the third step follows from the change of variable $\tilde{x} = x/b_n$ and the last step follows from the property of Dirac delta function. Using the fact that K^{ft} is supported on $[-1,1]$ under Assumption (3), this implies

$$
\begin{aligned}
E[|R_{n,1}|] &\le \frac{2(n-1)}{n^2}\{E[|q_n(a_1,a_1,a_2)|]+E[|q_n(a_1,a_2,a_2)|]\} = O\left(\frac{\max\{1,\sup_{|t|\le b_n^{-1}}E[|\Pi_1(t)|]\}}{nb_n^2\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2}\right),\\
E[|R_{n,2}|] &\le \frac{2(n-1)}{n^2}E[|q_n(a_1,a_2,a_1)|] = O\left(\frac{\max\{1,\sup_{|t|\le b_n^{-1}}E[|(Y_1-cW_1)\Pi_1(t)|]\}}{nb_n^2\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2}\right),\\
E[|R_{n,3}|] &\le \frac{2}{n^2}E[|q_n(a_1,a_1,a_1)|] = O\left(\frac{\max\{1,\sup_{|t|\le b_n^{-1}}E[|(Y_1-cW_1)\Pi_1(t)|]\}}{n^2b_n^2\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2}\right),
\end{aligned}
$$

and (A.4) follows by Lemma 3 and Assumption (4).

Second, under Assumption (4), we have

$$
B_n - \theta_c = o(n^{-1/2}), \tag{A.5}
$$

which follows from

$$
\begin{aligned}
&E[p_n(a_1,a_2,a_3)] = 2E[q_n(a_1,a_2,a_3)]\\
=&-\frac{2}{b_n^3}\int E\left[(Y-cW)\mathbb{K}\left(\frac{x-X}{b_n}\right)\right]E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right]dx\\
=&-2\int\left\{\frac{1}{b_n}E\left[(Y-cW)K\left(\frac{x-X^*}{b_n}\right)\right]\right\}\left\{\frac{1}{b_n^2}E\left[K'\left(\frac{x-X^*}{b_n}\right)\right]\right\}dx\\
=&\underbrace{-2\int h_c(x)f'(x)dx}_{\theta_c}-2\int\left\{\begin{array}{c}
h(x)\frac{b_n^{\alpha-1}}{(\alpha-1)!}\int K(u)\{\overbrace{f^{(\alpha)}(x+b_n\bar{u}_f)-f^{(\alpha)}(x)}^{\leq m(x)b_n|u|}\}u^{\alpha-1}du\\
+f'(x)\frac{b_n^{\alpha}}{\alpha!}\int K(u)\{\underbrace{h_c^{(\alpha)}(x+b_n\bar{u}_h)-h_c^{(\alpha)}(x)}_{\leq m(x)b_n|u|}\}u^{\alpha}du\\
+\left\{\begin{array}{c}
\frac{b_n^{\alpha-1}}{(\alpha-1)!}\int K(u)\{f^{(\alpha)}(x+b_n\bar{u}_f)-f^{(\alpha)}(x)\}u^{\alpha-1}du\\
\times\frac{b_n^{\alpha}}{\alpha!}\int K(u)\{h_c^{(\alpha)}(x+b_n\bar{u}_h)-h_c^{(\alpha)}(x)\}u^{\alpha}du
\end{array}\right\}
\end{array}\right\}dx\\
=&\theta_c+O(b_n^{\alpha})
\end{aligned}
$$

for some $\bar{u}_h$ and $\bar{u}_f$ such that $\max\{|\bar{u}_h|,|\bar{u}_f|\}\leq|u|$, where the second step follows from $E[\Pi_l(t/b_n)]=0$, the third step follows from Lemma 4, the fourth step follows from Lemma 5, and the last step follows from the Lipschitz conditions on $h_c^{(\alpha)}$ and $f^{(\alpha)}$ under Assumption (2).

Also, note by Lemma A.3 of Ahn and Powell (1993) that

$$
U_n-U_n^*=o_p(n^{-1/2}) \tag{A.6}
$$

if $E[|p_n(a_j,a_k,a_l)|^2]=o(n)$, which follows from Assumption (4), Lemma 3, and

$$
\begin{aligned}
&E[|p_n(a_j,a_k,a_l)|^2]\leq 4E[|q_n(a_j,a_k,a_l)|^2]\\
=&4E\left[\left|\frac{\mathrm{i}(cW_j-Y_j)}{2\pi b_n^2}\int te^{\mathrm{i}t\left(\frac{X_j-X_k}{b_n}\right)}\frac{|K^{\mathrm{ft}}(t)|^2}{|f_\epsilon^{\mathrm{ft}}(t/b_n)|^2}\{1+\Pi_l(t/b_n)+\Pi_l(-t/b_n)\}dt\right|^2\right]\\
\leq&\frac{E[|Y_j-cW_j|^2]}{b_n^4}\iint\left\{\begin{array}{c}
|t_1t_2|\frac{|K^{\mathrm{ft}}(t_1)|^2|K^{\mathrm{ft}}(t_2)|^2}{|f_\epsilon^{\mathrm{ft}}(t_1/b_n)|^2|f_\epsilon^{\mathrm{ft}}(t_2/b_n)|^2}\\
\times E\left[\begin{array}{c}
\{1+|\Pi_l(t_1/b_n)|+|\Pi_l(-t_1/b_n)|\}\\
\times\{1+|\Pi_l(t_2/b_n)|+|\Pi_l(-t_2/b_n)|\}
\end{array}\right]
\end{array}\right\}dt_1dt_2\\
=&O\left(\frac{\max\{1,\sup_{|t|\leq b_n^{-1}}E[|\Pi_1(t)|^2]\}}{b_n^4\{\inf_{|t|\leq b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^4}\right),
\end{aligned}
$$

where the second step follows by $k \neq l$ and the last step follows by the fact that K^{ft} is supported on $[-1,1]$ as in Assumption (3) and the Cauchy-Schwarz inequality.

Finally, observe that

$$3r_n(a_j) = E_j \begin{bmatrix} q_n(a_j,a_k,a_l) + q_n(a_j,a_l,a_k) \\ +q_n(a_k,a_j,a_l) + q_n(a_l,a_j,a_k) \\ +q_n(a_k,a_l,a_j) + q_n(a_l,a_k,a_j) \end{bmatrix} = 2 \begin{Bmatrix} E_j[q_n(a_j,a_k,a_l)] \\ +E_j[q_n(a_k,a_j,a_l)] \\ +E_j[q_n(a_k,a_l,a_j)] \end{Bmatrix}$$

$$= \underbrace{(-2)b_n^{-3} \int \begin{Bmatrix} E\left[(Y-cW)\mathbb{K}\left(\frac{x-X}{b_n}\right)\right] E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] \\ -\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n} h_c^{\mathrm{ft}}(t/b_n) E[\Pi_j^*(t/b_n)] K^{\mathrm{ft}}(t)dt\right\} E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] \\ -E\left[(Y-cW)\mathbb{K}\left(\frac{x-X}{b_n}\right)\right]\left\{\frac{-\mathrm{i}}{2\pi}\int t e^{-\mathrm{i}tx/b_n} f^{\mathrm{ft}}(t/b_n) E[\Pi_j^*(t/b_n)] K^{\mathrm{ft}}(t)dt\right\} \end{Bmatrix} dx}_{=:c_r}$$

$$+ \underbrace{(-2)b_n^{-3} \int \begin{Bmatrix} (Y_j-cW_j)\mathbb{K}\left(\frac{x-X_j}{b_n}\right) E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] + E\left[(Y-cW)\mathbb{K}\left(\frac{x-X}{b_n}\right)\right]\mathbb{K}'\left(\frac{x-X_j}{b_n}\right) \\ +\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n} h_c^{\mathrm{ft}}(t/b_n) \Pi_j^*(t/b_n) K^{\mathrm{ft}}(t)dt\right\} E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] \\ +E\left[(Y-cW)\mathbb{K}\left(\frac{x-X}{b_n}\right)\right]\left\{\frac{-\mathrm{i}}{2\pi}\int t e^{-\mathrm{i}tx/b_n} f^{\mathrm{ft}}(t/b_n) \Pi_j^*(t/b_n) K^{\mathrm{ft}}(t)dt\right\} \end{Bmatrix} dx}_{=:r_n^*(a_j)},$$

where the last step follows from $\Pi_j(t/b_n) = \Pi_j^*(t) - E[\Pi_j^*(t)]$ (so $E[\Pi_j(t/b_n)] = 0$) with $\Pi_j^*(t) = -\frac{\mu_{1,j}(t)}{\mu_1(t)} + \mathrm{i}\int_0^t \left\{-\frac{\mu_3(s)\mu_{2,j}(s)}{\mu_2^2(s)} + \frac{\mu_{3,j}(s)}{\mu_2(s)}\right\} ds$.

Since c_r is non-stochastic, to characterize the behavior of U_n^*, it is sufficient to focus on $r_n^*(a_j)$, for which we have

$$r_n^*(a_j) = 2b_n^{-3} \int \begin{Bmatrix} \int \mathbb{K}\left(\frac{x-X_j}{b_n}\right) E\left[(Y-cW)\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] \\ -(Y_j-cW_j)\mathbb{K}\left(\frac{x-X_j}{b_n}\right) E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] \\ -\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n} f^{\mathrm{ft}}(t/b_n)\Pi_j^*(t/b_n)K^{\mathrm{ft}}(t)dt\right\} E\left[(Y-cW)\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] \\ +\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n} h_c^{\mathrm{ft}}(t/b_n)\Pi_j^*(t/b_n)K^{\mathrm{ft}}(t)dt\right\} E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] \end{Bmatrix} dx$$

$$=2b_n^{-1} \int \begin{Bmatrix} \int \mathbb{K}\left(\frac{x-X_j}{b_n}\right) \left\{b_n^{-2} E\left[(Y-cW)\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right]\right\} \\ -(Y_j-cW_j)\mathbb{K}\left(\frac{x-X_j}{b_n}\right) \left\{b_n^{-2} E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right]\right\} \\ -\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n} f^{\mathrm{ft}}(t/b_n)\Pi_j^*(t/b_n)K^{\mathrm{ft}}(t)dt\right\} \left\{b_n^{-2} E\left[(Y-cW)\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right]\right\} \\ +\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n} h_c^{\mathrm{ft}}(t/b_n)\Pi_j^*(t/b_n)K^{\mathrm{ft}}(t)dt\right\} \left\{b_n^{-2} E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right]\right\} \end{Bmatrix} dx$$

$$=\frac{1}{\pi} \int \begin{Bmatrix} \int \left\{\{h_c'\}^{\mathrm{ft}}(-t) - (Y_j-cW_j)\{f'\}^{\mathrm{ft}}(-t)\right\} \frac{e^{\mathrm{i}tX_j}}{f_\epsilon^{\mathrm{ft}}(t)} \\ +\left\{f^{\mathrm{ft}}(t)\{h_c'\}^{\mathrm{ft}}(-t) - \{f'\}^{\mathrm{ft}}(-t)h_c^{\mathrm{ft}}(t)\right\}\Pi_j^*(t) \end{Bmatrix} dt + v_{n,1}(a_j) + v_{n,2}(a_j),$$

where the first step uses the integration by parts and $v_{n,1}(a_j)$ and $v_{n,2}(a_j)$ are defined as

$$v_{n,1}(a_j) = 2b_n^{-1}\int \left\{ \begin{array}{c} \left\{ \begin{array}{c} \frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n}\left[\frac{e^{\mathrm{i}tX_j/b_n}}{f_\epsilon^{\mathrm{ft}}(t/b_n)} + f^{\mathrm{ft}}(t/b_n)\Pi_j^*(t/b_n)\right]K^{\mathrm{ft}}(t)dt \\ \times\left\{b_n^{-2}E\left[(Y-cW)\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] - h_c'(x)\right\} \end{array}\right\} \\ -\left\{ \begin{array}{c} \frac{1}{2\pi}\int e^{-\mathrm{i}tx/b_n}\left[\frac{(Y_j-cW_j)e^{\mathrm{i}tX_j/b_n}}{f_\epsilon^{\mathrm{ft}}(t/b_n)} + h_c^{\mathrm{ft}}(t/b_n)\Pi_j^*(t/b_n)\right]K^{\mathrm{ft}}(t)dt \\ \times\left\{b_n^{-2}E\left[\mathbb{K}'\left(\frac{x-X}{b_n}\right)\right] - f'(x)\right\} \end{array}\right\} \end{array}\right\} dx,$$

$$v_{n,2}(a_j) = \frac{1}{\pi}\int \left\{ \begin{array}{c} \int\left\{\{h_c'\}^{\mathrm{ft}}(-t) - (Y_j-cW_j)\{f'\}^{\mathrm{ft}}(-t)\right\}\frac{e^{\mathrm{i}tX_j}}{f_\epsilon^{\mathrm{ft}}(t)} \\ +\left\{f^{\mathrm{ft}}(t)\{h_c'\}^{\mathrm{ft}}(-t) - \{f'\}^{\mathrm{ft}}(-t)h_c^{\mathrm{ft}}(t)\right\}\Pi_j^*(t) \end{array}\right\}\left\{K^{\mathrm{ft}}(tb_n)-1\right\}dt.$$

Since $Var[\xi_{c,j}] < \infty$ under Assumption (5), $Var[v_{n,2}(a_j)] = o(1)$ as $K^{\mathrm{ft}}(tb_n) \to 1$ as $n \to \infty$, and the conclusion follows if

$$Var[v_{n,1}(a_j)] = o(1). \tag{A.7}$$

To show (A.7), using Lemma 4 and 5, we can write $v_{n,1}(a_j) = v_{n,1,1}(a_j) + v_{n,1,2}(a_j)$, where

$$v_{n,1,1}(a_j) = \frac{b_n^{\alpha-2}}{\pi(\alpha-1)!}\int\left\{ \begin{array}{c} \int e^{-\mathrm{i}tx/b_n}\left\{\mathrm{i}f^{\mathrm{ft}}(t/b_n)\int_0^{t/b_n}\left\{-\frac{\{f^{\mathrm{ft}}\}'(s)}{f^{\mathrm{ft}}(s)} + \mathrm{i}X_j\right\}\frac{e^{\mathrm{i}sW_j}}{f^{\mathrm{ft}}(s)f_\nu^{\mathrm{ft}}(s)}ds\right\}K^{\mathrm{ft}}(t)dt \\ \times\int K(u)\{h_c^{(\alpha)}(x+b_n\bar{u}_h) - h_c^{(\alpha)}(x)\}u^{\alpha-1}du \end{array}\right\}dx,$$

$$v_{n,1,2}(a_j) = \frac{-b_n^{\alpha-2}}{\pi(\alpha-1)!}\int\left\{ \begin{array}{c} \int e^{-\mathrm{i}tx/b_n}\left\{ \begin{array}{c} \left\{(Y_j-cW_j) - \frac{h_c^{\mathrm{ft}}(t/b_n)}{f^{\mathrm{ft}}(t/b_n)}\right\}\frac{e^{\mathrm{i}tX_j/b_n}}{f_\epsilon^{\mathrm{ft}}(t/b_n)} \\ +\mathrm{i}h_c^{\mathrm{ft}}(t/b_n)\int_0^{t/b_n}\left\{-\frac{\{f^{\mathrm{ft}}\}'(s)}{f^{\mathrm{ft}}(s)} + \mathrm{i}X_j\right\}\frac{e^{\mathrm{i}sW_j}}{f^{\mathrm{ft}}(s)f_\nu^{\mathrm{ft}}(s)}ds \end{array}\right\}K^{\mathrm{ft}}(t)dt \\ \times\int K(u)\{f^{(\alpha)}(x+b_n\bar{u}_f) - f^{(\alpha)}(x)\}u^{\alpha-1}du \end{array}\right\}dx$$

For $v_{n,1,1}(a_j)$, we have

$$\begin{aligned}
&Var[v_{n,1,1}(a_j)] \leq E[|v_{n,1,1}(a_j)|^2] \\
=&E\left[\left|\frac{b_n^{\alpha-1}}{\pi(\alpha-1)!}\int\left\{ \begin{array}{c} \int e^{-\mathrm{i}\tilde{t}x}\left\{\mathrm{i}f^{\mathrm{ft}}(\tilde{t})\int_0^{\tilde{t}}\left\{-\frac{\{f^{\mathrm{ft}}\}'(s)}{f^{\mathrm{ft}}(s)} + \mathrm{i}X_j\right\}\frac{e^{\mathrm{i}sW_j}}{f^{\mathrm{ft}}(s)f_\nu^{\mathrm{ft}}(s)}ds\right\}K^{\mathrm{ft}}(\tilde{t}b_n)d\tilde{t} \\ \times\int K(u)\{h_c^{(\alpha)}(x+b_n\bar{u}_h) - h_c^{(\alpha)}(x)\}u^{\alpha-1}du \end{array}\right\}dx\right|^2\right] \\
\leq&E\left[\left\{b_n^{\alpha-1}\int\left\{ \begin{array}{c} \int\left\{\int_0^{\tilde{t}}\left\{\frac{|\{f^{\mathrm{ft}}\}'(s)|}{|f^{\mathrm{ft}}(s)|} + |X_j|\right\}\frac{1}{|f^{\mathrm{ft}}(s)||f_\nu^{\mathrm{ft}}(s)|}ds\right\}|K^{\mathrm{ft}}(\tilde{t}b_n)|d\tilde{t} \\ \times\int K(u)|\underbrace{h_c^{(\alpha)}(x+b_n\bar{u}_h) - h_c^{(\alpha)}(x)}_{\leq m(x)b_n|u|}||u|^{\alpha-1}du \end{array}\right\}dx\right\}^2\right] \\
=&O\left(\frac{b_n^{2(\alpha-1)}}{\{\inf_{|t|\leq b_n^{-1}}|f_\nu^{\mathrm{ft}}(t)|\}^2\{\inf_{|t|\leq b_n^{-1}}|f^{\mathrm{ft}}(t)|\}^4}\right),
\end{aligned}$$

where the second step follows from the change of variables $\tilde{t} = t/b_n$ and the last step uses the fact that K^{ft} is supported on $[-1, 1]$ as in Assumption (3). By similar argument, we can show

$$Var[v_{n,1,2}(a_j)] = O\left(\frac{b_n^{2(\alpha-1)}\{\inf_{|t|\le b_n^{-1}} |f^{\text{ft}}(t)|\}^{-2}}{\min\left\{\{\inf_{|t|\le b_n^{-1}} |f_\epsilon^{\text{ft}}(t)|\}^2, \{\inf_{|t|\le b_n^{-1}} |f_\nu^{\text{ft}}(t)|\}^2\{\inf_{|t|\le b_n^{-1}} |f^{\text{ft}}(t)|\}^2\right\}} \right),$$

and (A.7) follows from Assumption (4).

B Lemmas

Lemma 1. *Under Assumption (1), for $\iota = 1, 2, 3$,*

$$\sup_{|t| \leq b_n^{-1}} |\hat{\delta}_\iota(t)| = O_p\left(n^{-1/2}\log(1/b_n)\right).$$

Proof. See Lemma 2 in Kurisu and Otsu (2022). □

Lemma 2. *Under Assumptions (1) and (4),*

$$\sup_{|t| \leq b_n^{-1}} |\hat{\Pi}(t)| = O_p\left(\frac{n^{-1/2}\log(1/b_n)\{\inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|\}^{-1}}{\min\left\{\inf_{|t| \leq b_n^{-1}} |f_\epsilon^{\mathrm{ft}}(t)|, \{\inf_{|t| \leq b_n^{-1}} |f_\nu^{\mathrm{ft}}(t)|\}^2 \inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)| b_n\right\}}\right),$$

$$\sup_{|t| \leq b_n^{-1}} |\hat{\Pi}^{\mathrm{res}}(t)| = O_p\left(\frac{n^{-1}\log(1/b_n)^2\{\inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|\}^{-1}}{\min\left\{\inf_{|t| \leq b_n^{-1}} |f_\epsilon^{\mathrm{ft}}(t)|, \{\inf_{|t| \leq b_n^{-1}} |f_\nu^{\mathrm{ft}}(t)|\}^4 \{\inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|\}^3 b_n^2\right\}}\right).$$

Proof. The first statement follows by

$$\begin{aligned}
\sup_{|t| \leq b_n^{-1}} |\hat{\Pi}(t)| &= O_p\left(\frac{\sup_{|t| \leq b_n^{-1}} |\hat{\delta}_1(t)|}{\inf_{|t| \leq b_n^{-1}} |\mu_1(t)|} + b_n^{-1}\left\{\frac{\sup_{|t| \leq b_n^{-1}} |\mu_3(t)| \sup_{|t| \leq b_n^{-1}} |\hat{\delta}_2(t)|}{\{\inf_{|t| \leq b_n^{-1}} |\mu_2(t)|\}^2} + \frac{\sup_{|t| \leq b_n^{-1}} |\hat{\delta}_3(t)|}{\inf_{|t| \leq b_n^{-1}} |\mu_2(t)|}\right\}\right) \\
&= O_p\left(\frac{n^{-1/2}\log(1/b_n)\{\inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|\}^{-1}}{\min\left\{\inf_{|t| \leq b_n^{-1}} |f_\epsilon^{\mathrm{ft}}(t)|, \{\inf_{|t| \leq b_n^{-1}} |f_\nu^{\mathrm{ft}}(t)|\}^2 \inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)| b_n\right\}}\right),
\end{aligned}$$

where the last step uses Lemma 1, $\frac{\inf_{|t| \leq b_n^{-1}} |\mu_1(t)|}{\inf_{|t| \leq b_n^{-1}} |f_\epsilon^{\mathrm{ft}}(t)| \inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|} \geq 1$, $\frac{\inf_{|t| \leq b_n^{-1}} |\mu_2(t)|}{\inf_{|t| \leq b_n^{-1}} |f_\nu^{\mathrm{ft}}(t)| \inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|} \geq 1$, and $\sup_{|t| \leq b_n^{-1}} |\mu_3(t)| = O(1)$ under Assumption (1).

For the second statement, observe that,

$$\begin{aligned}
\sup_{|t| \leq b_n^{-1}} |\bar{\phi}(t)| =& O_p\left(\begin{array}{c} b_n^{-1}\left\{\frac{\sup_{|t| \leq b_n^{-1}} |\mu_3(t)| \sup_{|t| \leq b_n^{-1}} |\hat{\delta}_2(t)|}{\{\inf_{|t| \leq b_n^{-1}} |\mu_2(t)|\}^2} + \frac{\sup_{|t| \leq b_n^{-1}} |\hat{\delta}_3(t)|}{\inf_{|t| \leq b_n^{-1}} |\mu_2(t)|}\right\} \\ \times \left\{1 + \frac{\sup_{|t| \leq b_n^{-1}} |\hat{\delta}_2(t)|}{\inf_{|t| \leq b_n^{-1}} |\mu_2(t) + \hat{\delta}_2(t)|}\right\} \end{array}\right), \\
=& O_p\left(\frac{n^{-1/2} b_n^{-1}\log(1/b_n)}{\{\inf_{|t| \leq b_n^{-1}} |f_\nu^{\mathrm{ft}}(t)|\}^2 \{\inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|\}^2}\right) = o_p(1),
\end{aligned}$$

where the second step uses Lemma 1, $\frac{\inf_{|t| \leq b_n^{-1}} |\mu_2(t)|}{\inf_{|t| \leq b_n^{-1}} |f_\nu^{\mathrm{ft}}(t)| \inf_{|t| \leq b_n^{-1}} |f^{\mathrm{ft}}(t)|} \geq 1$, and $\sup_{|t| \leq b_n^{-1}} |\mu_3(t)| = O(1)$ under Assumption (1), and the last step follows from Assumption (4), which implies

$\sup_{|t|\le b_n^{-1}} e^{|\bar{\phi}(t)|} = O_p(1)$. The conclusion then follows by

$$\sup_{|t|\le b_n^{-1}} |\hat{\Pi}^{\mathrm{res}}(t)| = O_p\left(\begin{array}{c} \frac{\{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_1(t)|\}^2}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)+\hat{\delta}_1(t)|} + b_n^{-1}\left\{\frac{\sup_{|t|\le b_n^{-1}}|\mu_3(t)|\sup_{|t|\le b_n^{-1}}|\hat{\delta}_2(t)|}{\{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|\}^2} + \frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_3(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|}\right\} \\ \times\left\{\begin{array}{c} \frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_1(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)|} + \frac{\{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_1(t)|\}^2}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)+\hat{\delta}_1(t)|} + \frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_2(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_2(t)+\hat{\delta}_2(t)|} \\ + \frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_2(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_2(t)+\hat{\delta}_2(t)|}\left\{\frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_1(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)|} + \frac{\{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_1(t)|\}^2}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)+\hat{\delta}_1(t)|}\right\} \end{array}\right\} \\ + b_n^{-2}\left\{\frac{\sup_{|t|\le b_n^{-1}}|\mu_3(t)|\sup_{|t|\le b_n^{-1}}|\hat{\delta}_2(t)|}{\{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|\}^2} + \frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_3(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|}\right\}^2 \\ \times\left\{1+\frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_2(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_2(t)+\hat{\delta}_2(t)|}\right\}^2\left\{1+\frac{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_1(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)|} + \frac{\{\sup_{|t|\le b_n^{-1}}|\hat{\delta}_1(t)|\}^2}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)+\hat{\delta}_1(t)|}\right\} \end{array}\right)$$

$$= O_p\left(\frac{n^{-1}\log(1/b_n)^2\{\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|\}^{-1}}{\min\left\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|, \{\inf_{|t|\le b_n^{-1}}|f_\nu^{\mathrm{ft}}(t)|\}^4\{\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|\}^3 b_n^2\right\}}\right),$$

where the last step uses Lemma 1, $\frac{\inf_{|t|\le b_n^{-1}}|\mu_1(t)|}{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|} \ge 1$, $\frac{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|}{\inf_{|t|\le b_n^{-1}}|f_\nu^{\mathrm{ft}}(t)|\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|} \ge 1$, and $\sup_{|t|\le b_n^{-1}}|\mu_3(t)| = O(1)$ under Assumption (1). □

Lemma 3. *Under Assumption (1),*

$$\sup_{|t|\le b_n^{-1}} E[|\Pi_1(t)|^2] = O\left(\frac{\{\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|\}^{-2}}{\min\left\{\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2, \{\inf_{|t|\le b_n^{-1}}|f_\nu^{\mathrm{ft}}(t)|\}^2 b_n^2\right\}}\right),$$

which implies that for $s = 0, 1$,

$$\sup_{|t|\le b_n^{-1}} E[|(Y_1 - cW_1)^s\Pi_1(t)|] = O\left(\frac{\{\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|\}^{-1}}{\min\left\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|, \inf_{|t|\le b_n^{-1}}|f_\nu^{\mathrm{ft}}(t)|b_n\right\}}\right).$$

Proof. The second statement follows by the first statement, the Cauchy-Schwartz inequality and Assumption (1). The conclusion then follows by

$$\sup_{|t|\le b_n^{-1}} E[|\Pi_1(t)|^2] \le E\left[\sup_{|t|\le b_n^{-1}}\left|-\frac{\delta_{1,1}(t)}{\mu_1(t)} + \mathrm{i}\int_0^t\left\{-\frac{\mu_3(s)\delta_{2,1}(s)}{\mu_2^2(s)} + \frac{\delta_{3,1}(s)}{\mu_2(s)}\right\}ds\right|^2\right]$$

$$\le E\left[\left(\frac{\sup_{|t|\le b_n^{-1}}|\delta_{1,1}(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_1(t)|} + b_n^{-1}\left\{\frac{\sup_{|t|\le b_n^{-1}}|\mu_3(t)|\sup_{|t|\le b_n^{-1}}|\delta_{2,1}(t)|}{\{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|\}^2} + \frac{\sup_{|t|\le b_n^{-1}}|\delta_{3,1}(t)|}{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|}\right\}\right)^2\right]$$

$$= O\left(\frac{E[\sup_{|t|\le b_n^{-1}}|\delta_{1,1}(t)|^2]}{\{\inf_{|t|\le b_n^{-1}}|\mu_1(t)|\}^2} + \frac{\{\sup_{|t|\le b_n^{-1}}|\mu_3(t)|\}^2 E[\sup_{|t|\le b_n^{-1}}|\delta_{2,1}(t)|^2]}{b_n^2\{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|\}^2} + \frac{E[\sup_{|t|\le b_n^{-1}}|\delta_{3,1}(t)|^2]}{b_n^2\{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|\}^2}\right)$$

$$= O\left(\frac{\{\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|\}^{-2}}{\min\left\{\{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\}^2, \{\inf_{|t|\le b_n^{-1}}|f_\nu^{\mathrm{ft}}(t)|\}^2 b_n^2\right\}}\right),$$

where the last step uses $\frac{\inf_{|t|\le b_n^{-1}}|\mu_1(t)|}{\inf_{|t|\le b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(t)|\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|}\ge 1$, $\frac{\inf_{|t|\le b_n^{-1}}|\mu_2(t)|}{\inf_{|t|\le b_n^{-1}}|f_\nu^{\mathrm{ft}}(t)|\inf_{|t|\le b_n^{-1}}|f^{\mathrm{ft}}(t)|}\ge 1$, $\sup_{|t|\le b_n^{-1}}|\mu_3(t)|=O(1)$ and $E[\sup_{|t|\le b_n^{-1}}|\delta_{\iota,1}(t)|^2]<\infty$ for $\iota=1,2,3$ under Assumption (1). □

Lemma 4. *Under Assumptions (1) and (3), for* $s,k=0,1$,

$$E\left[(Y-W)^s\mathbb{K}^{(k)}\left(\frac{x-X}{b_n}\right)\right]=E\left[(Y-W)^sK^{(k)}\left(\frac{x-X^*}{b_n}\right)\right].$$

Proof. For $s,k=0,1$,

$$\begin{aligned}E\left[(Y-W)^s\mathbb{K}^{(k)}\left(\frac{x-X}{b_n}\right)\right]&=E\left[(Y-W)^s\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}t\left(\frac{x-X}{b_n}\right)}\frac{K^{\mathrm{ft}}(t)(-\mathrm{i}t)^k}{f_\epsilon^{\mathrm{ft}}(t/b_n)}dt\right\}\right]\\&=E\left[(Y-W)^s\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}t\left(\frac{x-X^*}{b_n}\right)}K^{\mathrm{ft}}(t)(-\mathrm{i}t)^kdt\right\}\right]\\&=E\left[(Y-W)^sK^{(k)}\left(\frac{x-X^*}{b_n}\right)\right],\end{aligned}$$

where the second step follows from the independence between ϵ and Y, and the last step follows from the fact $\{K^{(k)}\}^{\mathrm{ft}}(t)=K^{\mathrm{ft}}(t)(-\mathrm{i}t)^k$ for $k=0,1$. □

Lemma 5. *Under Assumptions (2) and (3), for* $k=0,1$,

$$\begin{aligned}b_n^{-(k+1)}E\left[K^{(k)}\left(\frac{x-X^*}{b_n}\right)\right]&=f^{(k)}(x)+\frac{b_n^{\alpha-k}}{(\alpha-k)!}\int K(u)\{f^{(\alpha)}(x+b_n\bar{u}_f)-f^{(\alpha)}(x)\}u^{\alpha-k}du,\\b_n^{-(k+1)}E\left[(Y-cW)K^{(k)}\left(\frac{x-X^*}{b_n}\right)\right]&=h_c^{(k)}(x)+\frac{b_n^{\alpha-k}}{(\alpha-k)!}\int K(u)\{h_c^{(\alpha)}(x+b_n\bar{u}_h)-h_c^{(\alpha)}(x)\}u^{\alpha-k}du,\end{aligned}$$

for some $\bar{u}_f$ *and* $\bar{u}_h$ *such that* $\max\{|\bar{u}_f|,|\bar{u}_h|\}\le|u|$.

Proof. Since the arguments are similar, we focus on the second statement when $k=1$, which follows from

$$\begin{aligned}&b_n^{-2}E\left[(Y-cW)K'\left(\frac{x-X^*}{b_n}\right)\right]=b_n^{-2}\int h_c(x^*)K'\left(\frac{x-x^*}{b_n}\right)dx^*\\&=-b_n^{-1}\int h_c(x^*)dK\left(\frac{x-x^*}{b_n}\right)=b_n^{-1}\int K\left(\frac{x^*-x}{b_n}\right)h_c'(x^*)dx^*\\&=\int K(u)\underbrace{h_c'(x+b_nu)}_{\sum_{l=0}^{\alpha-1}\frac{h_c^{(l+1)}(x)}{l!}b_n^lu^l+\frac{b_n^{\alpha-1}}{(\alpha-1)!}\{h_c^{(\alpha)}(x+b_n\bar{u}_f)-h_c^{(\alpha)}(x)\}u^{\alpha-1}}du\\&=h_c'(x)+\frac{b_n^{\alpha-1}}{(\alpha-1)!}\int K(u)\{h_c^{(\alpha)}(x+b_n\bar{u}_f)-h_c^{(\alpha)}(x)\}u^{\alpha-1}du,\end{aligned}$$

where the third step follows from the integration by parts and the symmetry of the kernel function K, the fourth step follows from the change of variables $u = (x^* - x)/b_n$, and the last step follows from the property of the kernel function K as in Assumption (3). $\square$

References

[1] Adusumilli, K., D. Kurisu, T. Otsu and Y.J. Whang (2020) Inference on distribution functions under measurement error, *Journal of Econometrics*, 215, 131-164.

[2] Adusumilli, K. and T. Otsu (2018) Nonparametric instrumental regression with errors in variables, *Econometric Theory*, 34, 1256-1280.

[3] Ahn, H. and J. L. Powell (1993) Semiparametric estimation of censored selection models with a nonparametric selection mechanism, *Journal of Econometrics*, 58, 3-29.

[4] Arellano, M., R. Blundell and S. Bonhomme (2017) Earnings and consumption dynamics: a nonlinear panel data framework, *Econometrica*, 85, 693-734.

[5] Bissantz N., L. Dümbgen, H. Holzmann and A. Munk (2007) Non-parametric confidence bands in deconvolution density estimation, *Journal of the Royal Statistical Society*, B 69, 483-506.

[6] Bissantz, N. and H. Holzmann (2008) Statistical inference for inverse problems, *Inverse Problems*, 24, Article 034009.

[7] Blundell, R., L. Pistaferri and I. Preston (2008) Consumption inequality and partial insurance, *American Economic Review*, 98, 1887-1921.

[8] Bonhomme, S. and J.M. Robin (2010) Generalized non-parametric deconvolution with an application to earnings dynamics, *The Review of Economic Studies*, 77, 491-533.

[9] Chen, X., H. Hong and D. Nekipelov (2011) Nonlinear models of measurement errors, *Journal of Economic Literature*, 49, 901-937.

[10] Carroll, C.D. (2009) Precautionary saving and the marginal propensity to consume out of permanent income, *Journal of Monetary Economics*, 56, 780-790.

[11] Carroll, R.J. and P. Hall (1988) Optimal rates of convergence for deconvolving a density, *Journal of the American Statistical Association*, 83, 1184-1186.

[12] Carroll, R.J., D. Ruppert, L.A. Stefanski and C.M. Crainiceanu (2006) *Measurement error in nonlinear models: a modern perspective*, CRC press.

[13] Comte, F. and J. Kappus (2015) Density deconvolution from repeated measurements without symmetry assumption on the errors, *Journal of Multivariate Analysis*, 140, 31-46.

[14] Comte, F. and C. Lacour (2011) Data-driven density estimation in the presence of additive noise with unknown distribution, *Journal of the Royal Statistical Society*, B 73, 601-627.

[15] Delaigle, A. , J. Fan and R.J. Carroll (2009) A design-adaptive local polynomial estimator for the errors-in-variables problem, *Journal of the American Statistical Association*, 104, 348-359.

[16] Delaigle, A., P. Hall and F. Jamshidi (2015) Confidence bands in non-parametric errors-in-variables regression, *Journal of the Royal Statistical Society*, B 77, 149-169.

[17] Delaigle, A., P. Hall and A. Meister (2008) On deconvolution with repeated measurements, *The Annals of Statistics*, 36, 665-685.

[18] Delaigle, A. and A. Meister (2007) Nonparametric regression estimation in the heteroscedastic errors-in-variables problem, *Journal of the American Statistical Association*, 102, 1416-1426.

[19] Diggle P.J. and P. Hall (1993) A Fourier approach to nonparametric deconvolution of a density estimate, *Journal of the Royal Statistical Society*, B 55, 523-531.

[20] Dong, H., T. Otsu and L.N. Taylor (2021) Average derivative estimation under measurement error, *Econometric Theory*, 37,1004-1033.

[21] Dong, H., T. Otsu and L.N. Taylor (2022) Estimation of varying coefficient models with measurement error, *Journal of Econometrics*, 230, 388-415.

[22] Efromovich, S. (1997) Density estimation for the case of supersmooth measurement error, *Journal of the American Statistical Association*, 92, 526-535.

[23] Fan, J. (1991a) On the optimal rates of convergence for nonparametric deconvolution problems, *The Annals of Statistics*, 19, 1257-1272.

[24] Fan, J. (1991b) Asymptotic normality for deconvolution kernel density estimators, *Sankhyā A*, 53, 97-110.

[25] Fan, J. and E. Masry (1992) Multivariate regression estimation with errors-in-variables: asymptotic normality for mixing processes, *Journal of Multivariate Analysis*, 43, 237-271.

[26] Fan, J. and Y.K. Truong (1993) Nonparametric regression with errors in variables, *The Annals of Statistics*, 21, 1990-1925.

[27] Fan, Y. (1995) Average derivative estimation with errors-in-variables, *Journal of Nonparametric Statistics*, 4, 395-407.

[28] Guvenen, F., S. Ozkan and J. Song (2014) The nature of countercyclical income risk, *Journal of Political Economy*, 122, 621-660.

[29] Guvenen, F., F. Karahan, S. Ozkan and J. Song (2021) What do data on millions of US workers reveal about lifecycle earnings dynamics, *Econometrica*, 89,2303-2339.

[30] Horowitz, J.L. (2009) *Semiparametric and nonparametric methods in econometrics*, Springer.

[31] Horowitz, J.L. and M. Markatou (1996) Semiparametric estimation of regression models for panel data, *The Review of Economic Studies*, 63, 145-168.

[32] Hu, Y., R. Moffitt and Y. Sasaki (2019) Semiparametric estimation of the canonical permanent-transitory model of earnings dynamics, *Quantitative Economics*, 10, 1495-1536.

[33] Hu, Y. and Y. Sasaki (2015) Closed-form estimation of nonparametric models with non-classical measurement errors, *Journal of Econometrics*, 185, 392-408.

[34] Johannes, J. (2009) Deconvolution with unknown error distribution, *The Annals of Statistics*, 37, 2301-2323.

[35] Kato, K. and Y. Sasaki (2018) Uniform confidence bands in deconvolution with unknown error distribution, *Journal of Econometrics*, 207, 129-161.

[36] Kato, K. and Y. Sasaki (2019) Uniform confidence bands for nonparametric errors-in-variables regression, *Journal of Econometrics*, 213, 516-555.

[37] Kato, K., Y. Sasaki and T. Ura (2021) Robust Inference in Deconvolution, *Quantitative Economics*, 21, 109-142.

[38] Kotlarski, I. (1967) On characterizing the gamma and the normal distribution, *Pacific Journal of Mathematics*, 20, 69-76.

[39] Kurisu, D. and T. Otsu (2022) On the uniform convergence of deconvolution estimators from repeated measurements, *Econometric Theory, 38, 172-193.*

[40] Li, T. (2002) Robust and consistent estimation of nonlinear errors-in-variables models, *Journal of Econometrics*, 110, 1-26.

[41] Li, T. and Q. Vuong (1998) Nonparametric estimation of the measurement error model using multiple indicators, *Journal of Multivariate Analysis*, 65, 139-165.

[42] Lounici, K. and R. Nickl (2011) Global uniform risk bounds for wavelet deconvolution estimators, *The Annals of Statistics*, 39, 201-231.

[43] Meister, A. (2009) *Deconvolution problems in nonparametric statistics*, Springer.

[44] Neumann, M.H. and O. Hössjer (1997) On the effect of estimating the error density in nonparametric deconvolution, *Journal of Nonparametric Statistics*, 7, 307-330.

[45] Powell, J.L., J.H. Stock and T.M. Stoker (1989) Semiparametric estimation of index coefficients, *Econometrica*, 57, 1403-1430.

[46] Schennach, S.M. (2004) Nonparametric regression in the presence of measurement error, *Econometric Theory*, 20, 1046-1093.

[47] Schennach, S.M. (2016) Recent advances in the measurement error literature, *Annual Review of Economics*, 8, 341-377.

[48] Schennach, S.M. (2021) Measurement systems, *Journal of Economic Literature,* forthcoming.

[49] Schennach, S.M. and Y. Hu (2013) Nonparametric identification and semiparametric estimation of classical measurement error models without side information, *Journal of the American Statistical Association*, 108, 177-186.

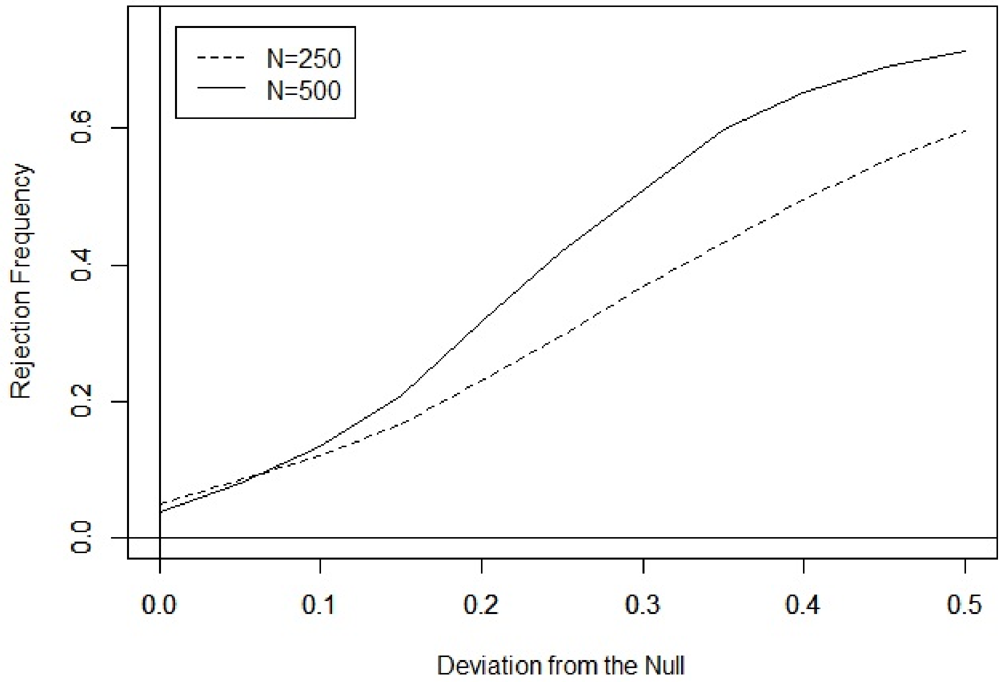
[50] Schennach, S.M., H. White and K. Chalak (2012) Local indirect least squares and average marginal effects in nonseparable structural systems, *Journal of Econometrics*, 166, 282-302.

[51] Schmidt-Hieber J., A. Munk and L. Dümbgen (2013) Multiscale methods for shape constraints in deconvolution: confidence statements for qualitative features, *The Annals of Statistics*, 41, 1299-1328.

[52] Stefanski, L.A. and R.J. Carroll (1990) Deconvolving kernel density estimators, *Statistics*, 21, 169-184.

[53] van Es, B. and S. Gugushvili (2008) Weak convergence of the supremum distance for supersmooth kernel deconvolution, *Statistics & Probability Letters*, 78, 2932-2938.

Figure 1: Monte Carlo Simulation Results



Notes: The vertical axis measures the rejection frequency for $H_0 : \theta_1 \geq 0$ with the nominal size of 0.05. The horizontal axis measures the deviation $\delta \in [0.0, 0.5]$ from the null hypothesis $H_0 : \theta_1 \geq 0$. The dashed (respectively, solid) line indicates the results with $N = 250$ (respectively, 500).

Table 1: Summary Statistics of U.S. PSID for 2013, 2015, 2017, and 2019.

Year	Log Income	Log Consumption	X	W	Y
2013	10.682	10.501	0.167	0.145	0.033
	(1.086)	(0.762)	(0.730)	(0.772)	(0.429)
2015	10.766	10.537			
	(1.061)	(0.739)			
2017	10.857	10.578			
	(1.021)	(0.696)			
2019	10.920	10.640			
	(1.042)	(0.708)	Observations = 5976		

Notes: Displayed values are the sample means. Parentheses enclose sample standard deviations.