AVERAGE DERIVATIVE ESTIMATION UNDER MEASUREMENT ERROR

HAO DONG, TAISUKE OTSU, AND LUKE TAYLOR

Abstract. In this paper, we derive the asymptotic properties of average derivative estimators when the regressors are contaminated with classical measurement error and the density of this error is unknown. Average derivatives of conditional mean functions are used extensively in economics and statistics, most notably in semiparametric index models. As well as ordinary smooth measurement error, we provide results for supersmooth error distributions. This is a particularly important class of error distribution as it includes the popular Gaussian density. We show that under this ill-posed inverse problem, despite using nonparametric deconvolution techniques and an estimated error characteristic function, we are able to achieve a $\sqrt{n}$ rate of convergence for the average derivative estimator. Interestingly, if the measurement error density is symmetric, the asymptotic variance of the average derivative estimator is the same irrespective of whether the error density is estimated or not.

1. Introduction

Since the seminal paper of Powell, Stock and Stoker (1989), average derivatives have enjoyed much popularity. They have found primary use in estimating coefficients in single index models, where Powell, Stock and Stoker (1989) showed that these estimators identify the parameters of interest up-to-scale. They have also been employed to great effect in the estimation of consumer demand functions (see, for example, Blundell, Duncan and Pendakur, 1998, and Yatchew, 2003) and sample selection models (for example, Das, Newey and Vella, 2003). Finally, several testing procedures have also made use of these estimators (see, for example, Härdle, Hildenbrand and Jerison, 1991, and Racine, 1997).

A key benefit of average derivate estimators is their ability to achieve a $\sqrt{n}$ rate of convergence despite being constructed using nonparametric techniques. Powell, Stock and Stoker (1989), among many others, demonstrated this parametric rate in the standard case of correctly measured regressors. Fan (1995) extended this result to allow for regressors contaminated with classical measurement error from the class of ordinary smooth distributions, for example, gamma or Laplace. In that paper, it was shown that average derivative estimators, constructed using deconvolution techniques, were able to retain the $\sqrt{n}$ rate of convergence enjoyed by their correctly

Financial support from the ERC Consolidator Grant (SNP 615882) is gratefully acknowledged (Otsu).

measured counterparts. However, this result relied on knowledge of the true error distribution and did not cover the case of supersmooth error densities, which includes Gaussian error.

Extending these results to supersmooth measurement error is not a trivial extension, and it is not clear a priori whether this parametric rate can be achieved in this case. Indeed, in many estimation and testing problems, convergence rates and asymptotic distributions are fundamentally different between ordinary smooth and supersmooth error densities (see, for example, Fan, 1991, van Es and Uh, 2005, Dong and Otsu, 2018, and Otsu and Taylor, 2019).

Furthermore, no result has been provided regarding the asymptotic properties of average derivative estimators in the more realistic situation where the measurement error density is unknown. Much recent work in the errors-in-variables literature has been aimed at relaxing the assumption of a known measurement error distribution, and deriving the asymptotic properties of estimators and test statistics in this setting (see, for example, Delaigle, Hall and Meister, 2008, Dattner, Reiß and Trabs, 2016, and Kato and Sasaki, 2018).

Measurement error is rife in datasets from all fields. It is a problem that affects economic, medical, social, and physical data sets, to name just a few. In response to the slow convergence rates achieved by nonparametric deconvolution techniques, practitioners may shy away from the use of these estimators in the face of classical measurement error. By showing that we can still obtain a parametric rate of convergence even in the worst case scenario of supersmooth error and an estimated error characteristic function, we hope to encourage greater use of nonparametric estimation in applied work when covariates are contaminated with error.

Moreover, since the curse of dimensionality (which plagues all nonparametric estimators) is exacerbated in the presence of measurement error, the potential gain from using average derivatives is increased when regressors are mismeasured. In particular, in the case of ordinary smooth error densities, the convergence rate of deconvolution estimators, although slower than standard nonparametric estimators, remains polynomial. However, for supersmooth densities, this convergence typically deteriorates to a $\log(n)$ rate.

In the next section, we describe the setup of our model, discuss the assumptions imposed, and provide our main result. All mathematical proofs are relegated to the Appendix.

2. Main result

2.1. **Setup and estimator.** Consider the nonparametric errors-in-variables model

$$\begin{aligned} Y &= g(X^*) + u, \qquad E[u|X^*] = 0, \\ X &= X^* + \epsilon, \end{aligned} \tag{1}$$

where Y is a scalar dependent variable, X^* is an unobservable error-free scalar covariate, X is an observable covariate, u is a regression error term, and ϵ is a measurement error on the covariate. Suppose the density function f of X^* and the regression function g are continuously differentiable, we are interested in estimating the density weighted average derivative

$$\theta = E[g'(X^*)f(X^*)] = -2E[Yf'(X^*)], \tag{2}$$

where g' and f' are the first-order derivatives of g and f, respectively. The second equality follows from using integration by parts (see Lemma 2.1 of Powell, Stock and Stoker, 1989).

The key use of such density weighted average derivatives is in single-index models and partially linear single-index models. Taking $g(X) = g(X_1'\beta, X_2)$ for some unknown link function g with $X = (X_1, X_2)$, we obtain the partially linear case; when X_2 is removed, this becomes the single-index model. Such specifications are very general and cover a wide variety of regression models. For example, binary choice models, truncated and censored dependent variable models, and duration models (see Ichimura, 1993, for a more detailed discussion). They can also be used as a simple dimension reduction solution to the curse of dimensionality.

For identification purposes, it is necessary to make some normalization restriction on β. This is because any scaling factor can be subsumed into g. Hence, this parameter of interest is only identified up to scale. Due to the linear index structure, the density weighted average derivative identifies this scaled β.

If we directly observe X^*, θ can be estimated by the sample analog $-\frac{2}{n}\sum_{j=1}^n Y_j \tilde{f}'(X_j^*)$, where $\tilde{f}'$ is a nonparametric estimator of the derivative f'. However, if X^* is unobservable, this estimator is infeasible. On the other hand, when the density function f_ϵ of the measurement error ϵ is *known* (and ordinary smooth), Fan (1995) suggested estimating θ by evaluating the joint density $h(x, y)$ of (X^*, Y) and the derivative $f'(x)$ in the expression

$$\theta = -2\int\int yf'(x)h(x, y)dxdy, \tag{3}$$

by applying the deconvolution method. Let $\mathrm{i} = \sqrt{-1}$ and f^{ft} be the Fourier transform of a function f. If f_ϵ is known, based on the i.i.d. sample $\{Y_j, X_j\}_{j=1}^n$ of (Y, X), the densities f and h can be estimated by

$$\tilde{f}(x) = \frac{1}{nb_n} \sum_{j=1}^n \mathbb{K}\left(\frac{x - X_j}{b_n}\right), \qquad \tilde{h}(x, y) = \frac{1}{nb_n^2} \sum_{j=1}^n \mathbb{K}\left(\frac{x - X_j}{b_n}\right) K_y\left(\frac{y - Y_j}{b_n}\right),$$

where b_n is a bandwidth, K_y is a (ordinary) kernel function and $\mathbb{K}$ is a deconvolution kernel function defined as

$$\mathbb{K}(x) = \frac{1}{2\pi} \int e^{-\mathrm{i}tx} \frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)} dt.$$

By plugging these estimators into (3), Fan (1995) proposed an estimator of θ and studied its asymptotic properties (again, when f_ϵ^{ft} is known and ordinary smooth).

In this paper, we extend Fan's (1995) result to the cases where (i) f_ϵ is unknown and symmetric around zero but repeated measurements on X^* are available, and (ii) f_ϵ is known and supersmooth. Since the second result is obtained as a by-product of the first one, we hereafter focus on the first case. Suppose we have two independent noisy measurements of the error-free variable X^*, i.e.,

$$X_j = X_j^* + \epsilon_j \quad \text{and} \quad X_j^r = X_j^* + \epsilon_j^r,$$

for $j = 1, \ldots, n$. Under the assumption that f_ϵ is symmetric, its Fourier transform f_ϵ^{ft} can be estimated by (Delaigle, Hall and Meister, 2008)

$$\hat{f}_\epsilon^{\mathrm{ft}}(t) = \left| \frac{1}{n} \sum_{j=1}^n \cos\{t(X_j - X_j^r)\} \right|^{1/2}. \tag{4}$$

By plugging in this estimator, the densities f and h can be estimated by

$$\hat{f}(x) = \frac{1}{nb_n} \sum_{j=1}^n \hat{\mathbb{K}}\left(\frac{x - X_j}{b_n}\right), \qquad \hat{h}(x, y) = \frac{1}{nb_n^2} \sum_{j=1}^n \hat{\mathbb{K}}\left(\frac{x - X_j}{b_n}\right) K_y\left(\frac{y - Y_j}{b_n}\right),$$

where

$$\hat{\mathbb{K}}(x) = \frac{1}{2\pi} \int e^{-\mathrm{i}tx} \frac{K^{\mathrm{ft}}(t)}{\hat{f}_\epsilon^{\mathrm{ft}}(t/b_n)} dt.$$

Then the parameter θ can estimated by

$$\begin{aligned} \hat{\theta} &= -2 \int y \hat{f}'(x) \hat{h}(x, y) dx dy \\ &= -\frac{2}{n^2 b_n^3} \sum_{j=1}^n \sum_{k=1}^n Y_k \int \hat{\mathbb{K}}'\left(\frac{x - X_j}{b_n}\right) \hat{\mathbb{K}}\left(\frac{x - X_k}{b_n}\right) dx, \end{aligned} \tag{5}$$

where $\hat{f}'$ and $\hat{\mathbb{K}}'$ are the first-order derivatives of $\hat{f}$ and $\hat{\mathbb{K}}'$, respectively, and the second equality follows from $\int yK_y((y-Y_k)/b_n)dy = b_nY_k$. Here we have derived the estimator for the case of a continuous Y. However, our estimator $\hat{\theta}$ in (5) can be applied to the case of a discrete Y as well.

Throughout this paper, we will focus on the case of a single covariate to keep the notation simple. The proposed method, however, can easily adapt to the multivariate case. In particular, when there are multiple covariates and one of them is mismeasured, i.e.,

$$Y = g(X^*, Z) + u,$$

where $Z = (Z_1, \dots, Z_D)$ is a vector of D correctly measured covariates, the parameters of interest are

$$\begin{aligned}
\theta_x &= E\left[\left.\frac{\partial g(x,z)}{\partial x}\right|_{(X^*,Z)} f_{X^*,Z}(X^*,Z)\right] = -2E\left[Y\left.\frac{\partial f_{X^*,Z}(x,z)}{\partial x}\right|_{(X^*,Z)}\right],\\
\theta_d &= E\left[\left.\frac{\partial g(x,z)}{\partial z_d}\right|_{(X^*,Z)} f_{X^*,Z}(X^*,Z)\right] = -2E\left[Y\left.\frac{\partial f_{X^*,Z}(x,z)}{\partial z_d}\right|_{(X^*,Z)}\right],
\end{aligned}$$

for $d = 1, \dots, D$, and can be written as

$$\begin{aligned}
\theta_x &= -2\int\int y\frac{\partial f_{X^*,Z}(x,z)}{\partial x}h(x,y,z)dxdydz,\\
\theta_d &= -2\int\int y\frac{\partial f_{X^*,Z}(x,z)}{\partial z_d}h(x,y,z)dxdydz,
\end{aligned}$$

for the joint densities $f_{X^*,Z}$ and h of (X^*, Z) and (X^*, Y, Z), respectively.

Let $K_z : \mathbb{R}^D \to \mathbb{R}$ be a (ordinary) kernel function. If f_ϵ is known, $f_{X^*,Z}$ and h can be estimated by

$$\begin{aligned}
\tilde{f}_{X^*,Z}(x,z) &= \frac{1}{nb_n^{D+1}}\sum_{j=1}^{n}\mathbb{K}\left(\frac{x-X_j}{b_n}\right)K_z\left(\frac{z-Z_j}{b_n}\right),\\
\tilde{h}(x,y,z) &= \frac{1}{nb_n^{D+2}}\sum_{j=1}^{n}\mathbb{K}\left(\frac{x-X_j}{b_n}\right)K_z\left(\frac{z-Z_j}{b_n}\right)K_y\left(\frac{y-Y_j}{b_n}\right),
\end{aligned}$$

and θ_x and θ_d can be estimated by

$$\begin{aligned}
\hat{\theta}_x &= -\frac{2}{n^2b_n^{2D+3}}\sum_{j=1}^{n}\sum_{k=1}^{n}Y_k\int\int\mathbb{K}'\left(\frac{x-X_j}{b_n}\right)\mathbb{K}\left(\frac{x-X_k}{b_n}\right)K_z\left(\frac{z-Z_j}{b_n}\right)K_z\left(\frac{z-Z_k}{b_n}\right)dxdz,\\
\hat{\theta}_d &= -\frac{2}{n^2b_n^{2D+3}}\sum_{j=1}^{n}\sum_{k=1}^{n}Y_k\int\int\mathbb{K}\left(\frac{x-X_j}{b_n}\right)\mathbb{K}\left(\frac{x-X_k}{b_n}\right)\frac{\partial K_z\left(\frac{z-Z_j}{b_n}\right)}{\partial z_d}K_z\left(\frac{z-Z_k}{b_n}\right)dxdz.
\end{aligned}$$

We expect that analogous results to our main theorem can be established for this estimator as well.

2.2. **Asymptotic properties.** We now investigate the asymptotic properties of the average derivative estimator $\hat{\theta}$ in (5). Let $G = gf$. For ordinary smooth measurement error densities, we impose the following assumptions.

Assumption OS.

(1): *$\{Y_j, X_j, X_j^r\}_{j=1}^n$ is an i.i.d. sample of (Y, X, X^r) satisfying (1). $g(\cdot) = E[Y|X^* = \cdot]$ has p continuous, bounded, and integrable derivatives. The density function $f(\cdot)$ of X^* has $(p+1)$ continuous, bounded, and integrable derivatives, where p is a positive integer satisfying $p > \alpha + 1$.*

(2): *(ϵ, ϵ^r) are mutually independent and independent of (Y, X^*), the distributions of ϵ and ϵ^r are identical, absolutely continuous with respect to the Lebesgue measure, and the characteristic function f_ϵ^{ft} is of the form*

$$f_\epsilon^{\text{ft}}(s) \sim \frac{1}{\sum_{v=0}^{\alpha} C_v |s|^v} \quad \textit{for all } t \in \mathbb{R},$$

for some finite constants $C_0, \ldots, C_\alpha$ with $C_0 \neq 0$ and a positive integer α.

(3): *K is differentiable to order $(\alpha + 1)$ and satisfies*

$$\int K(x)dx = 1, \qquad \int x^p K(x)dx \neq 0, \qquad \int x^l K(x)dx = 0, \textit{ for all } l = 1, \ldots, p-1.$$

Also K^{ft} is compactly supported on $[-1, 1]$, symmetric around zero, and bounded.

(4): *$n^{-1/2} b_n^{-2(1+3\alpha)} \log(b_n^{-1})^{-1/2} \to 0$, and $n^{1/2} b_n^p \to 0$ as $n \to \infty$.*

(5): *$Var(r(X, Y)) < \infty$, where*

$$r(X, Y) = \sum_{v=0}^{\alpha} (-\mathrm{i})^v C_v [G^{(v+1)}(X) - Y f^{(v+1)}(X)].$$

The i.i.d. restriction on the data from Assumption (1) is standard in the literature and is imposed merely for ease of derivation rather than necessity. The second part of this assumption requires sufficient smoothness from the regression function and density function of X relative to the smoothness of the measurement error. Assumption (2) is the conventional ordinary smooth assumption for the measurement error. Assumption (3) requires a kernel function of order p to remove the bias term from the nonparametric estimator. The first part of Assumption (4) requires that the bandwidth does not decay to zero too quickly as $n \to \infty$. This is necessary to ensure the asymptotic linearity of the estimator and apply a Hoeffding projection. The particular rate depends on the parameters of the measurement error characteristic function. The second part of Assumption (4) ensures the bandwidth approaches zero sufficiently fast to remove the asymptotic bias from the nonparametric estimator. Finally, Assumption (5) is a high-level assumption on the boundedness of the asymptotic variance of the average derivative estimator.

For the supersmooth case, we impose the following assumptions.

Assumption SS.

(1): $\{Y_j, X_j, X_j^r\}_{j=1}^n$ *is an i.i.d. sample of* (Y, X, X^r) *satisfying (1).* $g(\cdot) = E[Y|X^* = \cdot]$ *and the Lebesgue density* $f(\cdot)$ *of* X^* *are infinitely differentiable.*

(2): (ϵ, ϵ^r) *are mutually independent and independent of* (Y, X^*)*, the distributions of* ϵ *and* ϵ^r *are identical, absolutely continuous with respect to the Lebesgue measure, and the characteristic function* f_ϵ^{ft} *is of the form*

$$f_\epsilon^{\text{ft}}(t) = Ce^{-\mu|t|^\gamma} \quad \textit{for all } t \in \mathbb{R},$$

for some positive constants C *and* μ*, and positive even integer* γ*.*

(3): K *is infinitely differentiable and satisfies*

$$\int K(x)dx = 1, \qquad \int x^l K(x)dx = 0, \ \textit{for all } l \in \mathbb{N}.$$

Also K^{ft} *is compactly supported on* $[-1, 1]$*, symmetric around zero, and bounded.*

(4): $b_n \to 0$ *and* $n^{-1/2} b_n^{-2} e^{6\mu b_n^{-\gamma}} \log(b_n^{-1})^{-1/2} \to 0$ *as* $n \to \infty$*.*

(5): $Var(r(X, Y)) < \infty$*, where*

$$r(X, Y) = \sum_{h=0}^{\infty} \frac{\mu^h}{\mathrm{i}^{h\gamma} C h!} \{G^{(h\gamma+1)}(X) - Y f^{(h\gamma+1)}(X)\}.$$

Many of the same comments as for the ordinary smooth case apply to this setting. However, the second part of Assumption (2) is more restrictive and appears to be necessary. As discussed in Meister (2009), one can show that the class of infinitely differentiable functions still contains a comprehensive nonparametric class of densities' (pp. 44), including, of course, Gaussian and mixtures of Gaussians. For the regression function, all polynomials satisfy this restriction, as well as circular functions, exponentials, and products or sums of such smooth functions. Assumption (2) is the conventional supersmooth assumption for the measurement error, with the non-standard additional constraint on γ being even. Although this rules out the Cauchy distribution (where $\gamma = 1$), importantly, this still contains the canonical Gaussian distribution as well as Gaussian mixtures. van Es and Gugushvili (2008) imposed a similar constraint, although they restrict themselves further to $\gamma = 2$. Assumption (3) requires an infinite-order kernel function; these are often required in supersmooth deconvolution problems. Meister (2009) discussed their construction and noted that the commonly used sinc kernel, $K(x) = \frac{\sin(x)}{\pi x}$, satisfies the requirements. Assumption (4) requires the bandwidth to decay to zero at a logarithmic rate. In particular, because we are using an infinite-order kernel, we can ignore concerns of the bias from the nonparametric estimator and choose a bandwidth of at least $b_n = O\left((13\mu)^{1/\gamma}\log(n)^{-1/\gamma}\right)$ to satisfy this assumption.

Based on these assumptions, our main result is as follows.

Theorem. *Suppose Assumption OS or SS holds true. Then*

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, 4Var(r(X,Y))).$$

The most important aspect of this result is the $\sqrt{n}$ convergence of the estimator. Before this result, Powell, Stock and Stoker (1989) showed the same rate of convergence in the case of correctly measured regressors, and Fan (1995) confirmed this result for ordinary smooth error in the regressors when the error distribution is known. The above theorem shows that the convergence rate of these average derivative estimators does not change when measurement error is introduced. In particular, it does not change in the severely ill-posed case of supersmooth error, nor does it change when the measurement error distribution is estimated. Interestingly, as outlined in the Appendix, the asymptotic variance depends on the symmetry of the measurement error density. When the measurement error is symmetric around zero, remainder terms associated with the estimation error of the measurement error characteristic function vanish, and the asymptotic variance is the same as if the measurement error distribution is known; however, this is not the case for asymmetric distributions.

In pointwise estimation and testing problems, $\sqrt{n}$ convergence is typically not attained. For example, Holzman and Boysen (2006) showed that the integrated squared error of deconvolution estimators has a fundamentally different asymptotic distribution in the face of supersmooth measurement error in comparison to the case of ordinary smooth error. While Fan (1991) showed that deconvolution estimators under supersmooth contamination attain a $\log(n)$ rate of convergence whereas ordinary smooth measurement error results in a polynomial rate of convergence in n. In this paper, we show that this discontinuity in the properties of deconvolution estimators facing supersmooth or ordinary smooth error does not continue to hold for averaged estimators.

As a by-product of the proof, we also establish the asymptotic distribution of Fan's (1995) estimator for θ when the distribution of ϵ is known and supersmooth.

Corollary. *Suppose Assumption SS holds true without the repeated measurement X^r. Then the estimator $\tilde{\theta}$ defined by replacing $\hat{\mathbb{K}}$ in (5) with $\mathbb{K}$ satisfies*

$$\sqrt{n}(\tilde{\theta} - \theta) \xrightarrow{d} N(0, 4Var(r(X, Y))).$$

Appendix A. Proof of theorem (supersmooth case)

Since the arguments are similar, we first present a proof for the supersmooth case. In Section C, we provide a proof for the ordinary smooth case by explaining in detail the parts of the proof that differ to the supersmooth setting.

Let $\hat{\xi}(t) = \frac{1}{n}\sum_{l=1}^{n} \xi_l(t)$ for $\xi_l(t) = \cos(t(X_l - X_l^r))$, and $\xi(t) = |f_\epsilon^{\rm ft}(t)|^2$. Note that $\hat{f}_\epsilon^{\rm ft}(t) = |\hat{\xi}(t)|^{1/2}$ and $f_\epsilon(t) = |\xi(t)|^{1/2}$. By expansions around $\hat{\xi}(t/b_n) = \xi(t/b_n)$, we obtain

$$\begin{aligned}\hat{\mathbb{K}}(x) &= \mathbb{K}(x) + A_1(x) + R_1(x),\\ \hat{\mathbb{K}}'(x) &= \mathbb{K}'(x) + A_2(x) + R_2(x),\end{aligned}$$

where

$$\begin{aligned}
A_1(x) &= -\frac{1}{4\pi}\int e^{-\mathrm{i}tx} K^{\rm ft}(t)\left\{\frac{\hat{\xi}(t/b_n) - \xi(t/b_n)}{|\xi(t/b_n)|^{3/2}}\right\} dt,\\
A_2(x) &= \frac{\mathrm{i}}{4\pi}\int e^{-\mathrm{i}tx} tK^{\rm ft}(t)\left\{\frac{\hat{\xi}(t/b_n) - \xi(t/b_n)}{|\xi(t/b_n)|^{3/2}}\right\} dt,\\
R_1(x) &= -\frac{1}{4\pi}\int e^{-\mathrm{i}tx} K^{\rm ft}(t)\left\{\frac{1}{|\tilde{\xi}(t/b_n)|^{1/2}} - \frac{1}{|\xi(t/b_n)|^{1/2}}\right\}\left\{\frac{\hat{\xi}(t/b_n) - \xi(t/b_n)}{|\xi(t/b_n)|}\right\} dt\\
&\quad -\frac{1}{2\pi}\int e^{-\mathrm{i}tx} K^{\rm ft}(t)\left\{\frac{1}{|\hat{\xi}(t/b_n)|^{1/2}} - \frac{1}{|\xi(t/b_n)|^{1/2}}\right\}\left\{\frac{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}}{|\xi(t/b_n)|^{1/2}}\right\} dt,\\
R_2(x) &= \frac{\mathrm{i}}{4\pi}\int e^{-\mathrm{i}tx} tK^{\rm ft}(t)\left\{\frac{1}{|\tilde{\xi}(t/b_n)|^{1/2}} - \frac{1}{|\xi(t/b_n)|^{1/2}}\right\}\left\{\frac{\hat{\xi}(t/b_n) - \xi(t/b_n)}{|\xi(t/b_n)|}\right\} dt\\
&\quad +\frac{\mathrm{i}}{2\pi}\int e^{-\mathrm{i}tx} tK^{\rm ft}(t)\left\{\frac{1}{|\hat{\xi}(t/b_n)|^{1/2}} - \frac{1}{|\xi(t/b_n)|^{1/2}}\right\}\left\{\frac{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}}{|\xi(t/b_n)|^{1/2}}\right\} dt,
\end{aligned}$$

for some $\tilde{\xi}(t/b_n) \in (\hat{\xi}(t/b_n), \xi(t/b_n))$. Thus, we can decompose

$$\hat{\theta} = -\frac{2}{n^2 b_n^3}\sum_{j=1}^{n}\sum_{k=1}^{n} Y_k \int \hat{\mathbb{K}}'\left(\frac{x - X_j}{b_n}\right)\hat{\mathbb{K}}\left(\frac{x - X_k}{b_n}\right) dx = S + T_1 + \cdots + T_6, \tag{6}$$

where

$$\begin{aligned}
S &= -\frac{2}{n^2 b_n^3}\sum_{j=1}^{n}\sum_{k=1}^{n} Y_k \int \mathbb{K}'\left(\frac{x - X_j}{b_n}\right)\mathbb{K}\left(\frac{x - X_k}{b_n}\right) dx\\
&\quad -\frac{2}{n^2 b_n^3}\sum_{j=1}^{n}\sum_{k=1}^{n} Y_k \int \mathbb{K}'\left(\frac{x - X_j}{b_n}\right) A_1\left(\frac{x - X_k}{b_n}\right) dx\\
&\quad -\frac{2}{n^2 b_n^3}\sum_{j=1}^{n}\sum_{k=1}^{n} Y_k \int A_2\left(\frac{x - X_j}{b_n}\right)\mathbb{K}\left(\frac{x - X_k}{b_n}\right) dx,
\end{aligned}$$

$$T_1 = -\frac{2}{n^2b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int \mathbb{K}'\left(\frac{x-X_j}{b_n}\right) R_1\left(\frac{x-X_k}{b_n}\right)dx,$$

$$T_2 = -\frac{2}{n^2b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int R_2\left(\frac{x-X_j}{b_n}\right) \mathbb{K}\left(\frac{x-X_k}{b_n}\right)dx,$$

$$T_3 = -\frac{2}{n^2b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int A_2\left(\frac{x-X_j}{b_n}\right) A_1\left(\frac{x-X_k}{b_n}\right)dx,$$

$$T_4 = -\frac{2}{n^2b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int R_2\left(\frac{x-X_j}{b_n}\right) A_1\left(\frac{x-X_k}{b_n}\right)dx,$$

$$T_5 = -\frac{2}{n^2b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int A_2\left(\frac{x-X_j}{b_n}\right) R_1\left(\frac{x-X_k}{b_n}\right)dx,$$

$$T_6 = -\frac{2}{n^2b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int R_2\left(\frac{x-X_j}{b_n}\right) R_1\left(\frac{x-X_k}{b_n}\right)dx.$$

First, we show that $T_1, \dots, T_6$ are asymptotically negligible, i.e.,

$$T_1, \dots, T_6 = o_p(n^{-1/2}). \tag{7}$$

For T_2, we decompose $T_2 = T_{2,1} + T_{2,2}$, where

$$T_{2,1} = -\frac{\mathrm{i}}{2\pi n^2 b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int\int \left[\begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\tilde{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\}\end{array}\right] dt\mathbb{K}\left(\frac{x-X_k}{b_n}\right)dx,$$

$$T_{2,2} = -\frac{\mathrm{i}}{\pi n^2 b_n^3}\sum_{j=1}^n\sum_{k=1}^n Y_k \int\int \left[\begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\hat{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1/2}\{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}\}\end{array}\right] dt\mathbb{K}\left(\frac{x-X_k}{b_n}\right)dx.$$

For $T_{2,1}$, we have

$$\begin{aligned}
|n^{1/2}T_{2,1}| &= \left|\frac{1}{2\pi n^{3/2}b_n^2}\sum_{j=1}^n\sum_{k=1}^n Y_k \int \left\{\begin{array}{c} e^{\mathrm{i}t\left(\frac{X_j-X_k}{b_n}\right)} tK^{\mathrm{ft}}(t)K^{\mathrm{ft}}(-t)\{|\tilde{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-3/2}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\}\end{array}\right\} dt\right| \\
&= O_p\left(n^{1/2}b_n^{-2}\sup_{|t|\le b_n^{-1}}\left|\{|\tilde{\xi}(t)|^{-1/2} - |\xi(t)|^{-1/2}\}|\xi(t)|^{-3/2}\{\hat{\xi}(t) - \xi(t)\}\right|\right) \\
&= O_p\left(n^{1/2}b_n^{-2}e^{4\mu b_n^{-\gamma}}\varrho_n^2\right) = o_p(1),
\end{aligned}$$

where the first equality follows from a change of variables, the second equality follows from $\left|e^{\mathrm{i}t\left(\frac{X_j-X_k}{b_n}\right)}\right| = 1$, $\frac{1}{n}\sum_{k=1}^n |Y_k| = O_p(1)$, and $\int |tK^{\mathrm{ft}}(t)K^{\mathrm{ft}}(-t)| < \infty$ (by Assumption SS (3)), the third equality follows from the definition of $\tilde{\xi}(t)$, Assumption SS (2), and Lemma 1, and the last equality follows from Assumption SS (4). A similar argument yields $T_{2,2} = o_p(n^{-1/2})$, and thus $T_2 = o_p(n^{-1/2})$. Also, using similar arguments as for T_2, gives $T_1 = o_p(n^{-1/2})$ and $T_3 = o_p(n^{-1/2})$.

For T_4, we decompose $T_4 = T_{4,1} + T_{4,2}$, where

$$T_{4,1} = \frac{\mathrm{i}}{8\pi^2 n^2 b_n^3} \sum_{j=1}^n \sum_{k=1}^n Y_k \int \left[\begin{array}{c} \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\tilde{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1}\left\{\hat{\xi}(t/b_n) - \xi(t/b_n)\right\} \end{array} \right\} dt \\ \times \int e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} K^{\mathrm{ft}}(t)|\xi(t/b_n)|^{-3/2}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\}dt \end{array} \right] dx,$$

$$T_{4,2} = \frac{\mathrm{i}}{4\pi^2 n^2 b_n^3} \sum_{j=1}^n \sum_{k=1}^n Y_k \int \left[\begin{array}{c} \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\hat{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1/2}\{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}\} \end{array} \right\} dt \\ \times \int e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} K^{\mathrm{ft}}(t)|\xi(t/b_n)|^{-3/2}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\}dt \end{array} \right] dx.$$

For $T_{4,1}$, we have

$$\begin{aligned} |n^{1/2}T_{4,1}| &= O_p\left(n^{1/2} b_n^{-2} \sup_{|t|\le b_n^{-1}} \left| \{|\tilde{\xi}(t)|^{-1/2} - |\xi(t)|^{-1/2}\}|\xi(t)|^{-5/2}\{\hat{\xi}(t) - \xi(t)\}^2 \right| \right) \\ &= O_p\left(n^{1/2} b_n^{-2} e^{5\mu b_n^{-\gamma}} \varrho_n^3 \right) = o_p(1), \end{aligned}$$

where the first equality follows from a change of variables, $\left| e^{\mathrm{i}t\left(\frac{X_j - X_k}{b_n}\right)} \right| = 1$, $\frac{1}{n}\sum_{k=1}^n |Y_k| = O_p(1)$, and $\int |tK^{\mathrm{ft}}(t)K^{\mathrm{ft}}(-t)|dt < \infty$ (by Assumption SS (3)), the second equality follows from the definition of $\tilde{\xi}(t)$, Assumption SS (2), and Lemma 1, and the last equality follows from Assumption SS (4). A similar argument yields $T_{4,2} = o_p(n^{-1/2})$, and thus $T_4 = o_p(n^{-1/2})$. Also, similar arguments as used for T_4 imply $T_5 = o_p(n^{-1/2})$.

For T_6, we decompose $T_6 = T_{6,1} + T_{6,2} + T_{6,3} + T_{6,4}$, where

$$T_{6,1} = \frac{\mathrm{i}}{8\pi^2 n^2 b_n^3} \sum_{j=1}^{n} \sum_{k=1}^{n} Y_k \int \left[\begin{array}{c} \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\tilde{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\} \end{array} \right\} dt \\ \times \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} K^{\mathrm{ft}}(t)\{|\tilde{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\} \end{array} \right\} dt \end{array} \right] dx$$

$$T_{6,2} = \frac{\mathrm{i}}{4\pi^2 n^2 b_n^3} \sum_{j=1}^{n} \sum_{k=1}^{n} Y_k \int \left[\begin{array}{c} \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\hat{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1/2}\{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}\} \end{array} \right\} dt \\ \times \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} K^{\mathrm{ft}}(t)\{|\tilde{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ |\xi(t/b_n)|^{-1}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\} \end{array} \right\} dt \end{array} \right] dx$$

$$T_{6,3} = \frac{\mathrm{i}}{4\pi^2 n^2 b_n^3} \sum_{j=1}^{n} \sum_{k=1}^{n} Y_k \int \left[\begin{array}{c} \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\tilde{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1}\{\hat{\xi}(t/b_n) - \xi(t/b_n)\} \end{array} \right\} dt \\ \times \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} K^{\mathrm{ft}}(t)\{|\hat{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1/2}\{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}\} \end{array} \right\} dt \end{array} \right] dx$$

$$T_{6,4} = \frac{\mathrm{i}}{2\pi^2 n^2 b_n^3} \sum_{j=1}^{n} \sum_{k=1}^{n} Y_k \int \left[\begin{array}{c} \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} tK^{\mathrm{ft}}(t)\{|\hat{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1/2}\{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}\} \end{array} \right\} dt \\ \times \int \left\{ \begin{array}{c} e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} K^{\mathrm{ft}}(t)\{|\hat{\xi}(t/b_n)|^{-1/2} - |\xi(t/b_n)|^{-1/2}\} \\ \times|\xi(t/b_n)|^{-1/2}\{|\hat{\xi}(t/b_n)|^{1/2} - |\xi(t/b_n)|^{1/2}\} \end{array} \right\} dt \end{array} \right] dx$$

Since $T_{6,2}$ and $T_{6,3}$ are cross-product terms, it is enough to focus on $T_{6,1}$ and $T_{6,4}$. For $T_{6,1}$, we have

$$\begin{aligned} |n^{1/2} T_{6,1}| &= O_p\left(n^{1/2} b_n^{-2} \sup_{|t| \le b_n^{-1}} \left| \{|\tilde{\xi}(t)|^{-1/2} - |\xi(t)|^{-1/2}\}^2 |\xi(t)|^{-2} \{\hat{\xi}(t) - \xi(t)\}^2 \right| \right) \\ &= O_p\left(n^{1/2} b_n^{-2} e^{6\mu b_n^{-\gamma}} \varrho_n^4 \right) = o_p(1), \end{aligned}$$

where the first equality follows from a change of variables, $\left| e^{\mathrm{i}t\left(\frac{X_j - X_k}{b_n}\right)} \right| = 1$, $\frac{1}{n}\sum_{k=1}^{n} |Y_k| = O_p(1)$, and $\int |tK^{\mathrm{ft}}(t)K^{\mathrm{ft}}(-t)|dt < \infty$ (by Assumption SS (3)), the second equality follows from the definition of $\tilde{\xi}(t)$, Assumption SS (2), and Lemma 1, and the last equality follows from Assumption SS (4). A similar argument yields $T_{6,4} = o_p(n^{-1/2})$, and thus $T_6 = o_p(n^{-1/2})$.

Combining these results, we obtain (7).

We now consider the term S in (6). Let $d_j = (Y_j, X_j, \xi_j)$ and

$$p_n(d_j, d_k, d_l) = q_n(d_j, d_k, d_l) + q_n(d_j, d_l, d_k) + q_n(d_k, d_j, d_l) + q_n(d_k, d_l, d_j) + q_n(d_l, d_j, d_k) + q_n(d_l, d_k, d_j),$$

where

$$q_n(d_j,d_k,d_l) = -\frac{1}{3b_n^3}\left(\begin{array}{c} \int \mathbb{K}'\left(\frac{x-X_j}{b_n}\right) Y_k \mathbb{K}\left(\frac{x-X_k}{b_n}\right) dx \\ +\frac{\mathrm{i}}{4\pi}\int\left\{\int e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} Y_k \mathbb{K}\left(\frac{x-X_k}{b_n}\right) dx\right\}\left\{\frac{\xi_l(t/b_n)-E[\xi_l(t/b_n)]}{|\xi(t/b_n)|^{3/2}}\right\} tK^{\mathrm{ft}}(t)dt \\ -\frac{1}{4\pi}\int\left\{\int \mathbb{K}'\left(\frac{x-X_j}{b_n}\right) Y_k e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} dx\right\}\left\{\frac{\xi_l(t/b_n)-E[\xi_l(t/b_n)]}{|\xi(t/b_n)|^{3/2}}\right\} K^{\mathrm{ft}}(t)dt \end{array}\right).$$

We then decompose $S = n^{-2}(n-1)(n-2)U + S_1 + S_2 + S_3 + S_4$, where

$$U = \binom{n}{3}^{-1}\sum_{j=1}^{n}\sum_{k=j+1}^{n}\sum_{l=k+1}^{n} p_n(d_j,d_k,d_l),$$

$$S_1 = \frac{6}{n^3}\sum_{j=1}^{n}\sum_{k=j+1}^{n}[q_n(d_j,d_j,d_k)+q_n(d_k,d_k,d_j)], \qquad S_2 = \frac{6}{n^3}\sum_{j=1}^{n}\sum_{k=j+1}^{n}[q_n(d_j,d_k,d_j)+q_n(d_k,d_j,d_k)],$$

$$S_3 = \frac{6}{n^3}\sum_{j=1}^{n}\sum_{k=j+1}^{n}[q_n(d_j,d_k,d_k)+q_n(d_k,d_j,d_j)], \qquad S_4 = \frac{6}{n^3}\sum_{j=1}^{n} q_n(d_j,d_j,d_j).$$

We show that

$$S_1,\dots,S_4 = o_p(n^{-1/2}), \tag{8}$$

in the following way. For S_1, decompose

$$\begin{aligned} & |n^{1/2}S_1| \\ = {} & O(n^{-5/2}b_n^{-3})\left[\begin{array}{c} \left|\sum_{j=1}^n\sum_{k=j+1}^n \int \mathbb{K}'\left(\frac{x-X_j}{b_n}\right) Y_k \mathbb{K}\left(\frac{x-X_k}{b_n}\right) dx\right| \\ +\left|\sum_{j=1}^n\sum_{k=j+1}^n \frac{\mathrm{i}}{4\pi}\int\left\{\int e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} Y_k \mathbb{K}\left(\frac{x-X_k}{b_n}\right) dx\right\}\left\{\frac{\xi_l(t/b_n)-E[\xi_l(t/b_n)]}{|\xi(t/b_n)|^{3/2}}\right\} tK^{\mathrm{ft}}(t)dt\right| \\ +\left|\sum_{j=1}^n\sum_{k=j+1}^n \frac{1}{4\pi}\int\left\{\int \mathbb{K}'\left(\frac{x-X_j}{b_n}\right) Y_k e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} dx\right\}\left\{\frac{\xi_l(t/b_n)-E[\xi_l(t/b_n)]}{|\xi(t/b_n)|^{3/2}}\right\} K^{\mathrm{ft}}(t)dt\right| \end{array}\right] \\ \equiv {} & S_{1,1}+S_{1,2}+S_{1,3}. \end{aligned}$$

To bound $S_{1,1}$, we write

$$\begin{aligned} S_{1,1} &= O(n^{-5/2}b_n^{-2})\left|\sum_{j=1}^{n}\sum_{k=j+1}^{n} Y_k\left\{\int \frac{1}{b_n}e^{-\mathrm{i}(s+t)x/b_n}dx\right\}\int\int \mathrm{i}se^{\mathrm{i}\left(\frac{tX_k+sX_j}{b_n}\right)}\frac{K^{\mathrm{ft}}(s)}{f_\epsilon^{\mathrm{ft}}(s/b_n)}\frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)}dsdt\right| \\ &= O_p\left(n^{-1/2}b_n^{-2}e^{2\mu b_n^{-\gamma}}\right) = o_p(1), \end{aligned}$$

where the second equality follows from a change of variables, $\left|e^{\mathrm{i}\left(\frac{tX_k+sX_j}{b_n}\right)}\right| = 1$, $\frac{1}{n}\sum_{k=1}^n |Y_k| = O_p(1)$, and Assumption SS (3), and the last equality follows from Assumption SS (4). For $S_{1,2}$, a similar argument as used for T_3 can be used to show

$$S_{1,2} = O_p\left(n^{-1/2}b_n^{-2}e^{4\mu b_n^{-\gamma}}\varrho_n\right) = o_p(1).$$

Furthermore, the same arguments can be used to show $S_2, S_3, S_4 = o_p(n^{-1/2})$.

We now analyze the main term U. Let $r_n(d_j) = E[p_n(d_j, d_k, d_l)|d_j]$ and $\hat{U} = \theta + \frac{3}{n}\sum_{j=1}^{n}\{r_n(d_j) - E[r_n(d_j)]\}$. By Ahn and Powell (1993, Lemma A.3), if

$$E[p_n(d_j, d_k, d_l)^2] = o(n), \tag{9}$$

then it holds

$$U = \theta + \frac{3}{n}\sum_{j=1}^{n}\{r_n(d_j) - E[r_n(d_j)]\} + o_p(n^{-1/2}). \tag{10}$$

For (9), note that

$$\begin{aligned}
& E[p_n(d_j, d_k, d_l)^2] \\
\leq\ & \frac{1}{3b_n^6} E\left[\left\{\int \mathbb{K}'\left(\frac{x - X_j}{b_n}\right) Y_k \mathbb{K}\left(\frac{x - X_k}{b_n}\right) dx\right\}^2\right] \\
& + \frac{1}{3b_n^6} E\left[\left\{\frac{\mathrm{i}}{4\pi}\int \left\{\int e^{-\mathrm{i}t\left(\frac{x-X_j}{b_n}\right)} Y_k \mathbb{K}\left(\frac{x - X_k}{b_n}\right) dx\right\}\left\{\frac{\xi_l(t/b_n) - E[\xi_l(t/b_n)]}{|\xi(t/b_n)|^{3/2}}\right\} t K^{\mathrm{ft}}(t) dt\right\}^2\right] \\
& + \frac{1}{3b_n^6} E\left[\left\{\frac{1}{4\pi}\int \left\{\int \mathbb{K}'\left(\frac{x - X_j}{b_n}\right) Y_k e^{-\mathrm{i}t\left(\frac{x-X_k}{b_n}\right)} dx\right\}\left\{\frac{\xi_l(t/b_n) - E[\xi_l(t/b_n)]}{|\xi(t/b_n)|^{3/2}}\right\} K^{\mathrm{ft}}(t) dt\right\}^2\right] \\
\equiv\ & P_1 + P_2 + P_3.
\end{aligned}$$

For P_1,

$$\begin{aligned}
P_1 &= \frac{1}{3b_n^4}\int\int\int\int \left\{\int \mathbb{K}'(z)\mathbb{K}\left(z + \frac{s_j + t_j - s_k - t_k}{b_n}\right) dz\right\}^2 E[Y^2|X^* = s_k] \\
&\quad \times f(s_k) f(s_j) f_v(t_k) f_v(t_j) ds_k ds_j dt_k dt_j \\
&= \frac{1}{12\pi^2 b_n^4}\int\int\left\{\int\int e^{-\mathrm{i}(w_1+w_2)\left(\frac{s_j - s_k}{b_n}\right)} E[Y^2|X^* = s_k] f(s_k) f(s_j) ds_k ds_j\right\} \\
&\quad \times \frac{w_1 w_2 |K^{\mathrm{ft}}(w_1)|^2 |K^{\mathrm{ft}}(w_2)|^2}{|f_\epsilon^{\mathrm{ft}}(w_1/b_n)|^2 |f_\epsilon^{\mathrm{ft}}(w_2/b_n)|^2} dw_1 dw_2 \\
&= O\left(b_n^{-4} e^{4\mu b_n^{-\gamma}}\right),
\end{aligned}$$

where the first equality follows by the change of variables $z = \frac{x - s_j - t_j}{b_n}$, the second equality follows by Lemma 2, and the penultimate equality follows from Assumption SS (2). Thus Assumption SS (4) guarantees $P_1 = o(n)$.

For P_2, Lemma 2 implies

$$\begin{aligned}
& \int\left\{\frac{\mathrm{i}}{4\pi}\int t e^{-\mathrm{i}tz} \frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)^3}\{\xi_l(t/b_n) - E[\xi_l(t/b_n)]\} dt\right\} \mathbb{K}(z - c) dz \\
=\ & \frac{\mathrm{i}}{4\pi}\int \frac{w e^{-\mathrm{i}wc} |K^{\mathrm{ft}}(w)|^2}{|f_\epsilon^{\mathrm{ft}}(w/b_n)|^4}\{\xi_l(w/b_n) - E[\xi_l(w/b_n)]\} dw.
\end{aligned}$$

Then we can write

$$
\begin{aligned}
P_2 &= \frac{1}{3b_n^6} E\left[Y_k^2 \left\{ \int \left\{ \frac{\mathrm{i}}{4\pi} \int t e^{-\mathrm{i}tz} \frac{K^{\mathrm{ft}}(t)}{f_\epsilon^{\mathrm{ft}}(t/b_n)^3} \{\xi_l(t/b_n) - E[\xi_l(t/b_n)]\} dt \right\} \mathbb{K}\left(\frac{x - X_k}{b_n}\right) dx \right\}^2\right] \\
&= \frac{1}{3b_n^6} \int \cdots \int E\left[\left\{ \frac{\mathrm{i}}{4\pi} \int\int \left\{ \begin{array}{c} t e^{-\mathrm{i}t\left(\frac{x - s_j - u_j}{b_n}\right)} K^{\mathrm{ft}}(t) \\ \times f_\epsilon^{\mathrm{ft}}(t/b_n)^{-3} \{\xi_l(t/b_n) - E[\xi_l(t/b_n)]\} \end{array} \right\} dt \mathbb{K}\left(\frac{x - s_k - u_k}{b_n}\right) dx \right\}^2\right] \\
&\quad \times E[Y^2 | X^* = s_k] f(s_k) f(s_j) f_v(u_k) f_v(u_j) ds_k ds_j du_k du_j \\
&= \frac{1}{12\pi^2 b_n^4} \int\int \left\{ \int\int e^{-\mathrm{i}(w_1 + w_2)\left(\frac{s_j - s_k}{b_n}\right)} E[Y^2 | X^* = s_k] f(s_k) f(s_j) ds_k ds_j \right\} \\
&\quad \times \frac{w_1 w_2 |K^{\mathrm{ft}}(w_1)|^2 |K^{\mathrm{ft}}(w_2)|^2}{|f_\epsilon^{\mathrm{ft}}(w_1/b_n)|^6 |f_\epsilon^{\mathrm{ft}}(w_2/b_n)|^6} E[\{\xi_l(w_1/b_n) - E[\xi_l(w_1/b_n)]\}\{\xi_l(w_2/b_n) - E[\xi_l(w_2/b_n)]\}] dw_1 dw_2 \\
&= O\left(b_n^{-4} e^{12\mu b_n^{-\gamma}} \log(b_n^{-1})^{-1}\right) = o(n),
\end{aligned}
$$

where the third equality follows from a similar argument as for P_1 combined with Kato and Sasaki (2018, Lemma 4) to bound $\{\xi_l(w_1/b_n) - E[\xi_l(w_1/b_n)\}$, and the last equality follows from Assumption SS (4). The order of P_3 can be shown in an almost identical manner, and we obtain (9).

Combining (6), (7), (8), (10), and a direct calculation to characterize $r_n(d_j) = E[p_n(d_j, d_k, d_l)|d_j]$, it follows

$$
\begin{aligned}
&\sqrt{n}(\hat{\theta} - \theta) = \frac{3}{\sqrt{n}} \sum_{j=1}^n \{r_n(d_j) - E[r_n(d_j)]\} + o_p(1), \\
&= \frac{2}{\sqrt{n} b_n^3} \sum_{j=1}^n \{\eta_j - E[\eta_j]\} - \frac{1}{2\pi\sqrt{n} b_n^3} \sum_{j=1}^n \int \Delta(t) \left\{ \frac{\xi_j(t/b_n) - E[\xi_j(t/b_n)]}{|\xi(t/b_n)|^{3/2}} \right\} K^{\mathrm{ft}}(t) dt + o_p(1), \qquad (11)
\end{aligned}
$$

where

$$
\begin{aligned}
\eta_j &= \int \mathbb{K}\left(\frac{x - X_j}{b_n}\right) \left\{ E\left[Y K'\left(\frac{x - X^*}{b_n}\right)\right] - Y_j E\left[K'\left(\frac{x - X^*}{b_n}\right)\right] \right\} dx, \\
\Delta(t) &= \int\int \mathrm{i}t E\left[e^{-\mathrm{i}t\left(\frac{x - X}{b_n}\right)}\right] E\left[Y K\left(\frac{x - X^*}{b_n}\right)\right] - E\left[K'\left(\frac{x - X^*}{b_n}\right)\right] E\left[Y e^{-\mathrm{i}t\left(\frac{x - X}{b_n}\right)}\right] dx.
\end{aligned}
$$

For the first term in (11), note that

$$
\begin{aligned}
\frac{\eta_j}{b_n^3} &= \frac{1}{b_n^2}\int \mathbb{K}(z)\{q_1(X_j+b_nz)-Y_jq_2(X_j+b_nz)\}dz \\
&= \sum_{l=0}^{+\infty}\frac{(-1)^l b_n^l}{l!}\int\int \mathbb{K}(z)K(w)(w-z)^l dwdz\{G^{(l+1)}(X_j)-Y_jf^{(l+1)}(X_j)\} \\
&= \sum_{l=0}^{+\infty}\frac{(-1)^l b_n^l}{l!}\int \mathbb{K}(z)z^l dz\{G^{(l+1)}(X_j)-Y_jf^{(l+1)}(X_j)\} \\
&= \sum_{h=0}^{\infty}\frac{\mu^h}{\mathrm{i}^{h\gamma}Ch!}\{G^{(h\gamma+1)}(X_j)-Y_jf^{(h\gamma+1)}(X_j)\},
\end{aligned}
$$

where the first equality follows by $q_1(x)=E\left[YK'\left(\frac{x-X^*}{b_n}\right)\right]$ and $q_2(x)=E\left[K'\left(\frac{x-X^*}{b_n}\right)\right]$ and the change of variable $z=\frac{x-X_j}{b_n}$, the second equality follows by Lemma 4, the third equality follows from Assumption SS (3), and the last equality follows by Lemma 5.

Let $\Xi_j(t)=\frac{\xi_j(t/b_n)-E[\xi_j(t/b_n)]}{|\xi(t/b_n)|}$. For the second term in (11), we have

$$
\begin{aligned}
&\int \Delta(t)\left\{\frac{\xi_j(t/b_n)-E[\xi_j(t/b_n)]}{|\xi(t/b_n)|^{3/2}}\right\}K^{\mathrm{ft}}(t)dt \\
=\ & \mathrm{i}\int tf^{\mathrm{ft}}(t/b_n)\left\{\int\int e^{-\mathrm{i}tx/b_n}K\left(\frac{x-x^*}{b_n}\right)G(x^*)dxdx^*\right\}\Xi_j(t)K^{\mathrm{ft}}(t)dt \\
& -\int G^{\mathrm{ft}}(t/b_n)\left\{\int\int e^{-\mathrm{i}tx/b_n}K'\left(\frac{x-x^*}{b_n}\right)f(x^*)dxdx^*\right\}\Xi_j(t)K^{\mathrm{ft}}(t)dt \\
=\ & \mathrm{i}\int tf^{\mathrm{ft}}(t/b_n)\left\{\int\int e^{-\mathrm{i}tx/b_n}K\left(\frac{x-x^*}{b_n}\right)G(x^*)dxdx^*\right\}\Xi_j(t)K^{\mathrm{ft}}(t)dt \\
& -\mathrm{i}\int tG^{\mathrm{ft}}(t/b_n)\left\{\int\int e^{-\mathrm{i}tx/b_n}K\left(\frac{x-x^*}{b_n}\right)f(x^*)dxdx^*\right\}\Xi_j(t)K^{\mathrm{ft}}(t)dt \\
=\ & \mathrm{i}b_n\int tK^{\mathrm{ft}}(t)f^{\mathrm{ft}}(t/b_n)K^{\mathrm{ft}}(-t)G^{\mathrm{ft}}(-t/b_n)\Xi_j(t)dt \\
& -\mathrm{i}b_n\int tK^{\mathrm{ft}}(-t)f^{\mathrm{ft}}(-t/b_n)K^{\mathrm{ft}}(t)G^{\mathrm{ft}}(t/b_n)\Xi_j(t)dt \\
=\ & 0,
\end{aligned}
$$

where the first equality follows from the definition of $\Delta(t)$, the second equality follows by integration by parts, that is $\int e^{-\mathrm{i}tx/b_n}K'\left(\frac{x-x^*}{b_n}\right)dx=\mathrm{i}t\int e^{-\mathrm{i}tx/b_n}K\left(\frac{x-x^*}{b_n}\right)dx$, the third equality follows from a change of variables, and the last equality follows from symmetry of $\xi_j(t)$ and $\xi(t)$ (which implies symmetry of $\Xi_j(t)$).

Therefore, the conclusion follows by the central limit theorem.

Appendix B. Lemmas

Lemma 1. *[Kato and Sasaki, 2018, Lemma 4] Under Assumption SS,*

$$\sup_{|t|\leq b_n^{-1}} |\hat{f}_\epsilon^{\mathrm{ft}}(t) - f_\epsilon^{\mathrm{ft}}(t)| = O_p(\varrho_n),$$

where $\varrho_n = n^{-1/2}\log(b_n^{-1})^{1/2}$.

Lemma 2. *Under Assumption SS (1) and (3), it holds*

$$\int \mathbb{K}'(z)\mathbb{K}(z+c)dz = \frac{\mathrm{i}}{2\pi}\int \frac{we^{-\mathrm{i}wc}|K^{\mathrm{ft}}(w)|^2}{|f_\epsilon^{\mathrm{ft}}(w/b_n)|^2}dw,$$

for any constant c.

Proof. Observe that

$$\begin{aligned}\int \mathbb{K}'(z)\mathbb{K}(z-c)dz &= \int\left(\frac{-\mathrm{i}}{2\pi}\int w_1 e^{-\mathrm{i}w_1 z}\frac{K^{\mathrm{ft}}(w_1)}{f_\epsilon^{\mathrm{ft}}(w_1/b_n)}dw_1\right)\left(\frac{1}{2\pi}\int e^{-\mathrm{i}w_2 z}\frac{e^{-\mathrm{i}w_2 c}K^{\mathrm{ft}}(w_2)}{f_\epsilon^{\mathrm{ft}}(w_2/b_n)}dw_2\right)dz \\ &= \frac{-\mathrm{i}}{2\pi}\int\int\left(\frac{1}{2\pi}\int e^{-\mathrm{i}(w_1+w_2)z}dz\right)\frac{w_1 e^{-\mathrm{i}w_2 c}K^{\mathrm{ft}}(w_1)K^{\mathrm{ft}}(w_2)}{f_\epsilon^{\mathrm{ft}}(w_1/b_n)f_\epsilon^{\mathrm{ft}}(w_2/b_n)}dw_1dw_2 \\ &= \frac{\mathrm{i}}{2\pi}\int\frac{we^{-\mathrm{i}wc}|K^{\mathrm{ft}}(w)|^2}{|f_\epsilon^{\mathrm{ft}}(w/b_n)|^2}dw,\end{aligned}$$

where the last equality follows by $\int \delta(w-b)f(w)dw = f(b)$ with Dirac delta function $\delta(w) = \frac{1}{2\pi}\int e^{-\mathrm{i}wx}dx$. ∎

Lemma 3. *Under Assumption SS (1)-(3), it holds*

$$|\int \mathbb{K}'(z)\mathbb{K}(z)dz| = O(e^{2\mu b_n^{-\gamma}}).$$

Proof. By Lemma 2, we have

$$|\int \mathbb{K}'(z)\mathbb{K}(z)dz| = \frac{1}{2\pi}\left|\int\frac{w|K^{\mathrm{ft}}(w)|^2}{|f_\epsilon^{\mathrm{ft}}(w/b_n)|^2}dw\right| = O\left(\left(\inf_{|w|\leq b_n^{-1}}|f_\epsilon^{\mathrm{ft}}(w)|\right)^{-1}\right),$$

where the second equality follows from compactness of support of K^{ft} (Assumption SS (3)). The conclusion follows by Assumption SS (2). ∎

Lemma 4. *Under Assumption SS (1) and (3), it holds*

$$E\left[Y_k K'\left(z+\frac{X_j-X_k^*}{b_n}\right)\Big| X_j\right] = \sum_{l=0}^{+\infty}\frac{(-1)^l b_n^{l+2} G^{(l+1)}(X_j)}{l!}\int K(w)(w-z)^l dw,$$

$$E\left[K'\left(z+\frac{X_j-X_k^*}{b_n}\right)\Big| X_j\right] = \sum_{l=0}^{+\infty}\frac{(-1)^l b_n^{l+2} f^{(l+1)}(X_j)}{l!}\int K(w)(w-z)^l dw.$$

Proof. Since $G = gf$ is infinitely differentiable (Assumption SS (1)), we have

$$\begin{aligned} E\left[Y_k K'\left(z+\frac{X_j-X_k^*}{b_n}\right)\Big| X_j\right] &= \int G(s)K'\left(z+\frac{X_j-s}{b_n}\right)ds \\ = \quad -b_n \int G(X_j-b_n(w-z))K'(w)dw &= b_n^2\int K(w)G'(X_j-b_n(w-z))dw \\ = \quad \sum_{l=0}^{+\infty}\frac{(-1)^l b_n^{l+2}G^{(l+1)}(X_j)}{l!}&\int K(w)(w-z)^l dw, \end{aligned}$$

where the second equality follows by the change of variable $w = z + \frac{X_j - s}{b_n}$, the third equality follows by integration by parts, the fourth equality follows by an expansion of $G'(X_j - b_n(w-z))$ around X_j. The second statement can be proved by similar arguments. $\square$

Lemma 5. *Under Assumptions SS (1)-(3), it holds*

$$\int \mathbb{K}(z)z^p dz = \begin{cases} \frac{\mu^{p/\gamma}p!}{b_n^p \mathrm{i}^p C(p/\gamma)!} & \text{for } p = h\gamma \text{ with } h = 0,1,\ldots, \\ 0 & \text{for other positive integers.} \end{cases}$$

Proof. First, note that

$$\begin{aligned} \mathbb{K}(z) &= \frac{1}{2\pi C}\int e^{-\mathrm{i}tz}e^{\mu|t/b_n|^\gamma}K^{\mathrm{ft}}(|t|)dt = \sum_{h=0}^{+\infty}\frac{\mu^h}{Ch!b_n^{h\gamma}}\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tz}|t|^{h\gamma}K^{\mathrm{ft}}(|t|)dt\right\} \\ &= \sum_{h=0}^{+\infty}\frac{\mu^h}{Ch!(-\mathrm{i}b_n)^{h\gamma}}\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tz}(K^{(h\gamma)})^{\mathrm{ft}}(|t|)dt\right\} \\ &= \sum_{h=0}^{+\infty}\frac{\mu^h}{Ch!(-\mathrm{i}a_n)^{h\gamma}}\left\{\frac{1}{2\pi}\int e^{-\mathrm{i}tz}(K^{(h\gamma)})^{\mathrm{ft}}(t)dt\right\} = \sum_{h=0}^{+\infty}\frac{\mu^h}{Ch!(-\mathrm{i}a_n)^{h\gamma}}K^{(h\gamma)}(z), \end{aligned}$$

where the first equality follows by Assumption SS (2) and $K^{\mathrm{ft}}(t) = K^{\mathrm{ft}}(-t)$, the second equality follows by $e^u = \sum_{h=0}^{+\infty}\frac{u^h}{h!}$, the third equality follows by $(K^{(l)})^{\mathrm{ft}}(t) = (-\mathrm{i}t)^l K^{\mathrm{ft}}(t)$ (see, e.g., Lemma A.6 of Meister, 2009), the fourth equality follows by $(K^{(h\gamma)})^{\mathrm{ft}}(-t) = (K^{(h\gamma)})^{\mathrm{ft}}(t)$, which is from $K^{\mathrm{ft}}(t) = K^{\mathrm{ft}}(-t)$, $(K^{(l)})^{\mathrm{ft}}(t) = (-\mathrm{i}t)^l K^{\mathrm{ft}}(t)$, and the assumption that γ is even. Thus, we have

$$\int \mathbb{K}(z)z^p dz = \sum_{h=0}^{+\infty}\frac{\mu^h}{Ch!(-\mathrm{i}b_n)^{h\gamma}}\int z^p K^{(h\gamma)}(z)dz,$$

and the conclusion follows by Assumption SS (3) and using the integration by parts. □

Appendix C. Proof of theorem (ordinary smooth case)

The steps in this proof are the same as that for the supersmooth case, as such, we only explain parts of the proof that differ. Furthermore, in the proof of the supersmooth case we endevour to obtain expressions in terms of f_ϵ^{ft} wherever possible. This allows us to skip to this final step in each asymptotic argument, and requires us only to input the relevant form for f_ϵ^{ft}. This proof also leverages much of the work from Fan (1995) but extends this by allowing for an estimated measurement error density.

As in the proof of the supersmooth case, we have

$$\hat{\theta} \;=\; -\frac{2}{n^2 b_n^3}\sum_{j=1}^{n}\sum_{k=1}^{n} Y_k \int \hat{\mathbb{K}}'\left(\frac{x-X_j}{b_n}\right)\hat{\mathbb{K}}\left(\frac{x-X_k}{b_n}\right) dx = S + T_1 + \cdots + T_6,$$

where $S, T_1, \ldots, T_6$ are defined in Section A. We were able to show that

$$|n^{1/2}T_2| \;=\; O\left(n^{1/2} b_n^{-2}\left(\sup_{|t|\le b_n^{-1}} |\hat{f}_\epsilon^{\text{ft}}(t) - f_\epsilon^{\text{ft}}(t)|\right)^2 \left(\inf_{|t|\le b_n^{-1}} |f_\epsilon^{\text{ft}}(t)|^4\right)^{-1}\right).$$

Then we have $T_2 = o_p(n^{-1/2})$ by Lemma 1 and Assumption OS (2). The rest of $T_1, T_3, \ldots T_6$ are shown to be of order $o_p(n^{-1/2})$ in a similar way.

Again, decompose $S = n^{-2}(n-1)(n-2)U + S_1 + \cdots + S_4$, where all objects are defined in the proof of the supersmooth case. We can show the asymptotic negligibility of $S_1, \ldots, S_4$ as follows. We again decompose $|n^{1/2}S_1| = S_{1,1} + S_{1,2} + S_{1,3}$. To bound $S_{1,1}$, we write

$$S_{1,1} \;=\; O_p\left(n^{-1/2} b_n^{-2}\left(\inf_{|t|\le b_n^{-1}} |f_\epsilon^{\text{ft}}(t)|\right)^{-2}\right) = o_p(1).$$

where the second equality follows from Assumption OS (2) and (4). Recall from the proof of the supersmooth case

$$S_{1,2} \;=\; O_p\left(n^{-1/2} b_n^{-2}\left(\sup_{|t|\le b_n^{-1}} |\hat{f}_\epsilon^{\text{ft}}(t) - f_\epsilon^{\text{ft}}(t)|\right)\left(\inf_{|t|\le b_n^{-1}} |f_\epsilon^{\text{ft}}(t)|^4\right)^{-1}\right) = o_p(1).$$

The asymptotic negligibility of $S_{1,3}$ can be shown in an almost identical way. The same arguments can also be used to show $S_2, S_3, S_4 = o_p(n^{-1/2})$.

As in the supersmooth case, we also need to show $E[p_n(d_j, d_k, d_l)^2] = o(n)$ in order to write $U = \theta + \frac{3}{n}\sum_{j=1}^{n}\{r_n(d_j) - E[r_n(d_j)]\} + o_p(1)$. We begin by decomposing $E[p_n(d_j, d_k, d_l)^2] =$

$P_1 + P_2 + P_3$, where these objects are defined in the supersmooth proof. For P_1,

$$P_1 = O\left(b_n^{-4}\left(\inf_{|w|\leq b_n^{-1}} |f_\epsilon^{\text{ft}}(w)|^2\right)^{-2}\right) = o(n),$$

by Assumption OS (2) and (4). For P_2, we can write

$$P_2 = O\left(b_n^{-4}\left(\inf_{|w|\leq b_n^{-1}} |f_\epsilon^{\text{ft}}(w)|^6\right)^{-2} \log(b_n)^{-2}\right) = o(n),$$

by Assumption OS (2) and (4). The order of P_3 can be shown in an almost identical manner.

Then, it follows

$$\sqrt{n}(\hat{\theta} - \theta) = \frac{3}{\sqrt{n}} \sum_{j=1}^{n} \{r_n(d_j) - E[r_n(d_j)]\} + o_p(1),$$

and the remainder of the proof for the supersmooth case applies as it does not depend on the form of f_ϵ^{ft}.

References

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

[2] Blundell, R., Duncan, A. and K. Pendakur (1998) Semiparametric estimation and consumer demand, *Journal of Applied Econometrics*, 13, 435-461.

[3] Das, M., Newey, W. K. and F. Vella (2003) Nonparametric estimation of sample selection models, *Review of Economic Studies*, 70, 33-58.

[4] Dattner, I., Reiß, M. and M. Trabs (2016) Adaptive quantile estimation in deconvolution with unknown error distribution, *Bernoulli*, 22, 143-192.

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

[6] Dong, H. and T. Otsu (2018) Nonparametric estimation of additive model with errors-in-variables, Working paper.

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

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

[9] Härdle, W., Hildenbrand, W. and M. Jerison (1991) Empirical evidence on the law of demand, *Econometrica*, 1525-1549.

[10] Holzmann, H. and L. Boysen (2006) Integrated square error asymptotics for supersmooth deconvolution, *Scandinavian Journal of Statistics*, 33, 849-860.

[11] Ichimura, H. (1993) Semiparametric least squares (SLS) and weighted SLS estimation of single-index models, *Journal of Econometrics*, 58, 71-120.

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

[13] Meister, A (2009) *Deconvolution Problems in Nonparametric Statistics*, Springer.

[14] Otsu, T. and L. Taylor (2019) Specification testing for errors-in-variables models, Working paper.

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

[16] Racine, J. (1997) Consistent significance testing for nonparametric regression, *Journal of Business & Economic Statistics*, 15, 369-378.

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

[18] van Es, B. and H.-W. Uh (2005) Asymptotic normality for kernel type deconvolution estimators, *Scandinavian Journal of Statistics*, 32, 467-483.

[19] Yatchew, A. (2003) *Semiparametric Regression for the Applied Econometrician*, Cambridge University Press.

Department of Economics, Southern Methodist University, 3300 Dyer Street, Dallas, TX 75275, US.

Email address: haod@smu.edu

Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, UK.

Email address: t.otsu@lse.ac.uk

Department of Economics and Business Economics, Fuglesangs Allé 4 building 2631, 12 8210 Aarhus V, Denmark

Email address: lntaylor@econ.au.dk