PHBS WORKING PAPER SERIES

Faster Uniform Convergence Rates for Deconvolution Estimators From Repeated Measurements

Liang Chen
Peking University

Minyuan Zhang
Shanghai University of Finance and Economics

September 2022

Working Paper 20220905

Abstract

Recently, Kurisu and Otsu (2022, Econometric Theory 38(1), 172-193) derived the uniform convergence rates for the nonparametric deconvolution estimators proposed by Li and Vuong (1998, Journal of Multivariate Analysis 65, 139-165). This paper shows that faster uniform convergence rates can be established for their estimators under the same assumptions. In addition, a new class of deconvolution estimators based on a variant of Kotlarski's identity is also proposed. It is shown that in some cases, these new estimators can have faster uniform convergence rates than the existing estimators.

Keywords: Nonparametric deconvolution, Kotlarski's identity, rate of convergence
JEL Classification: C14, C31, C33

Peking University HSBC Business School
University Town, Nanshan District
Shenzhen 518055, China

Faster Uniform Convergence Rates for Deconvolution Estimators From Repeated Measurements*

Liang Chen[1] and Minyuan Zhang[2]

[1] *Peking Unviersity HSBC Business School*
[2] *School of Economics, Shanghai University of Finance and Economics*

August 31, 2022

Abstract

Recently, Kurisu and Otsu (2022, *Econometric Theory* 38(1), 172-193) derived the uniform convergence rates for the nonparametric deconvolution estimators proposed by Li and Vuong (1998, *Journal of Multivariate Analysis* 65, 139-165). This paper shows that faster uniform convergence rates can be established for their estimators under the same assumptions. In addition, a new class of deconvolution estimators based on a variant of Kotlarski's identity is also proposed. It is shown that in some cases, these new estimators can have faster uniform convergence rates than the existing estimators.

Keywords: nonparametric deconvolution; Kotlarski's identity; rate of convergence.

JEL codes: C14, C31, C33.

*Address correspondence to Liang Chen, Peking University HSBC Business School, No. 2199 Lishui Road, Shenzhen, Guangdong 518055, China; e-mail: chenliang@phbs.pku.edu.cn.

1 Introduction

In empirical research, many data sets are contaminated by measurement errors. Thus, there is a large body of literature that studies the identification and estimation of measurement error models. Chen, Hong, and Nekipelov (2011) and Hu (2017) provide excellent reviews of recent developments and applications of measurement errors models in economics.

One of the most fundamental problems in this literature is how to recover the distribution of an unobserved variable (denoted by X) from one or more mismeasured variables. The classical *nonparametric deconvolution* approach (see Fan 1991, for example) assumes that there is an observed measurement (denoted by Y) contaminated by an independent measurement error (denoted by ϵ), i.e., $Y = X + \epsilon$, and that the distribution of ϵ is known. However, the assumption of knowing the error's distribution is too strong for most applications. Thus, many studies on structural identification in economics rely on Kotlarski's identity (see Kotlarski 1967 and Evdokimov and White 2012), which provides an explicit identification result for the unknown distributions of X and ϵ based on two different contaminated measurements — see Li, Perrigne, and Vuong (2000), Bonhomme and Robin (2010), Arcidiacono et al. (2011), Kennan and Walker (2011), Krasnokutskaya (2011) and Arellano and Bonhomme (2012) for examples.

Based on Kotlarski's identity, a class of nonparametric deconvolution estimators for the characteristic functions and density functions of the error-free variable and the measurement errors (denoted by $\varphi_X, \varphi_\epsilon, f_X, f_\epsilon$) was first proposed by Li and Vuong (1998). The uniform convergence rates of the LV estimators were established, and it was shown that these rates depend on the smoothness of f_X and f_ϵ, which is characterized by the decay rates of their characteristic functions. Recently, based on a new maximal inequality for the multivariate empirical characteristic function process, Kurisu and Otsu (2022) (KO, hereafter) showed that the LV estimators can achieve faster uniform convergence rates under weaker assumptions (e.g., unbounded supports of X and ϵ).

The main goal of this study is to establish faster uniform convergence rates for nonparametric deconvolution estimators based on Kotlarski's identity. This paper makes two main contributions. First, it is shown that the uniform convergence rates of the LV estimators established in KO can be further improved under the same assumptions, and the faster convergence rates obtained in this paper can provide some new insights. For example, we find that the uniform convergence rate of the LV estimator for φ_ϵ only depends on the smoothness of the distribution of X, while in LV and KO this rate also depends on the smoothness of the distribution of ϵ. Second, we propose a new class of nonparametric deconvolution estimators based on a variant of Kotlarski's identity, where the roles of the measurement errors and the error-free variable are switched. Under very similar assumptions, the uniform convergence rates of these new estimators are established, and they are compared with the rates of the LV estimators. It is found that these new estimators can have much faster convergence rates than the LV estimators in some cases. Moreover, since the deconvolution estimators are often used as the first-step nonparametric estimators in the estimation of some semiparametric models, the faster rates of convergence obtained in this paper will facilitate the asymptotic analysis of these two-step semiparametric estimators (see Newey 1994 for example).

The rest of the paper is organized as follows. Section 2 defines the models and the estimators. The uniform convergence rates of the estimators are presented and compared in Section 3. Section 4 gives a heuristic explanation of how we can obtain the faster convergence rates for the LV estimators. Finally, Section 5 concludes.

2 The Models and The Estimators

Let X be an error-free random variable of interest, and let Y_1, Y_2 be two noisy measurements of X with measurement errors ϵ_1 and ϵ_2 respectively. In particular, the model can be written as

$$\begin{aligned} Y_1 &= X + \epsilon_1, \\ Y_2 &= X + \epsilon_2. \end{aligned} \tag{1}$$

The following assumption is standard in the literature to derive Kotlarski's identity.

Assumption 1. *Let $\varphi_X(\cdot)$ and $\varphi_\epsilon(\cdot)$ denote the characteristic functions of X and ϵ respectively.*
(i) X, ϵ_1 and ϵ_2 are mutually independent.
(ii) The distributions of ϵ_1 and ϵ_2 are identical to the distribution of ϵ, and $\mathbb{E}[\epsilon] = 0$.
(iii) $\varphi_X(\cdot)$ and $\varphi_\epsilon(\cdot)$ are non-vanishing everywhere.
(iv) $\mathbb{E}|X| < \infty$ and $\mathbb{E}|\epsilon| < \infty$.

Let $\psi(u_1, u_2) = \mathbb{E}[e^{\mathsf{i}u_1Y_1 + \mathsf{i}u_2Y_2}]$ be the characteristic function of (Y_1, Y_2), where $\mathsf{i} = \sqrt{-1}$, then it is well known that Kotlarski's identify provides an explicit identification result for φ_X and φ_ϵ.

Lemma 1. *Under Assumption 1,*

$$\varphi_X(t) = \exp\left(\int_0^t \frac{\partial\psi(0, u_2)/\partial u_1}{\psi(0, u_2)} du_2\right) = \exp\left(\int_0^t \frac{\mathsf{i}\mathbb{E}[Y_1 e^{\mathsf{i}sY_2}]}{\mathbb{E}[e^{\mathsf{i}sY_2}]} ds\right), \tag{2}$$

and $\varphi_\epsilon(t) = \psi(t, 0)/\varphi_X(t) = \psi(0, t)/\varphi_X(t)$.

Let $\{Y_{j1}, Y_{j2}\}_{j=1}^n$ be an i.i.d sample from the joint distribution of (Y_1, Y_2), and let $f_X(\cdot)$ and $f_\epsilon(\cdot)$ denote the density functions of X and ϵ respectively. In a seminal work, Li and Vuong

(1998) proposed the following estimators for φ_X and f_X based on Lemma 1:

$$\hat{\varphi}_X(t) = \exp\left(\int_0^t \frac{\sum_j \mathrm{i}Y_{j1}e^{\mathrm{i}sY_{j2}}}{\sum_j e^{\mathrm{i}sY_{j2}}}ds\right), \quad \text{and} \quad \hat{f}_X(x) = \frac{1}{2\pi}\int e^{-\mathrm{i}tx}\hat{\varphi}_X(t)\varphi_K(h_n t)dt,$$

where φ_K is the Fourier transform of a kernel function K and h_n is a bandwidth parameter. Moreover, φ_ϵ and f_ϵ can be estimated by

$$\hat{\varphi}_\epsilon(t) = \frac{\hat{\psi}(t,0)}{\hat{\varphi}_X(t)}, \quad \text{and} \quad \hat{f}_\epsilon(u) = \frac{1}{2\pi}\int e^{-\mathrm{i}tu}\hat{\varphi}_\epsilon(t)\varphi_K(h_n t)dt,$$

where $\hat{\psi}(t,0) = n^{-1}\sum_{j=1}^n e^{\mathrm{i}tY_{j1}}$.

Alternatively, model (1) can be written as

$$\begin{aligned} Y_1 =& \epsilon_1 + X, \\ Y_1 - Y_2 =& \epsilon_1 - \epsilon_2, \end{aligned} \tag{3}$$

where Y_1 and $Y_1 - Y_2$ can be viewed as the measurements, and ϵ_1 is treated as the error-free variable. Thus, one can derive an alternative set of estimators for $(\varphi_X, f_X, \varphi_\epsilon, f_\epsilon)$ based on the application of Kotlarski's identity to (3). In particular, the following identification result gives an alternative explicit formula for φ_ϵ.

Lemma 2. *Under Assumption 1,*

$$\varphi_\epsilon(t) = \exp\left(\int_0^t \frac{\mathrm{i}\mathbb{E}[Y_1 e^{\mathrm{i}s(Y_1-Y_2)}]}{\mathbb{E}[e^{\mathrm{i}s(Y_1-Y_2)}]}ds - \mathrm{i}t\mathbb{E}[Y_1]\right).$$

Based on the above result, the following estimators for $(\varphi_X, f_X, \varphi_\epsilon, f_\epsilon)$ are proposed:

$$\tilde{\varphi}_\epsilon(t) = \exp\left(\int_0^t \frac{\sum_j \mathrm{i}Y_{j1}e^{\mathrm{i}s(Y_{j1}-Y_{j2})}}{\sum_j e^{\mathrm{i}s(Y_{j1}-Y_{j2})}}ds - \frac{1}{n}\sum_j \mathrm{i}tY_{j1}\right),$$

$$\tilde{f}_\epsilon(u) = \frac{1}{2\pi}\int e^{-\mathrm{i}tu}\tilde{\varphi}_\epsilon(t)\varphi_K(h_n t)dt,$$

$$\tilde{\varphi}_X(t) = \frac{\hat{\psi}(t,0)}{\tilde{\varphi}_\epsilon(t)}, \quad \tilde{f}_X(x) = \frac{1}{2\pi}\int e^{-\mathrm{i}tx}\tilde{\varphi}_X(t)\varphi_K(h_n t)dt.$$

3 Main Results

The main purpose of this section is to establish the uniform convergence rates of $(\hat{\varphi}_X, \hat{\varphi}_\epsilon)$ and $(\tilde{\varphi}_X, \tilde{\varphi}_\epsilon)$. Following the literature, we require one of the following assumptions to hold for the characteristic functions of X and ϵ.

Assumption 2. ***(Ordinary Smooth)*** *There exist some positive constants $\beta_x > 1$, $C_x \geq c_x$, ω_x, $\beta_\epsilon > 1$, $C_\epsilon \geq c_\epsilon$ and ω_ϵ such that*

(i) $c_x|t|^{-\beta_x} \leq |\varphi_X(t)| \leq C_x|t|^{-\beta_x}$ for all $|t| \geq \omega_x$.

(ii) $c_\epsilon|t|^{-\beta_\epsilon} \leq |\varphi_\epsilon(t)| \leq C_\epsilon|t|^{-\beta_\epsilon}$ for all $|t| \geq \omega_\epsilon$.

Assumption 3. ***(Supersmooth)*** *There exist some positive constants $\rho_x, C_x \geq c_x, \omega_x, \mu_x, \rho_\epsilon, C_\epsilon \geq c_\epsilon, \omega_\epsilon, \mu_\epsilon$ and some constants β_x, β_ϵ such that*

(i) $c_x|t|^{\beta_x}\exp(-|t|^{\rho_x}/\mu_x) \leq |\varphi_X(t)| \leq C_x|t|^{\beta_x}\exp(-|t|^{\rho_x}/\mu_x)$ for all $|t| \geq \omega_x$.

(ii) $c_\epsilon|t|^{\beta_\epsilon}\exp(-|t|^{\rho_\epsilon}/\mu_\epsilon) \leq |\varphi_\epsilon(t)| \leq C_\epsilon|t|^{\beta_\epsilon}\exp(-|t|^{\rho_\epsilon}/\mu_\epsilon)$ for all $|t| \geq \omega_\epsilon$.

The concepts of *ordinary smooth* (OS, hereafter) and *supersmooth* (SS, hereafter) distributions were first introduced by Fan (1991) in the study of the optimal convergence rates for nonparametric deconvolution estimators. In his setup, there is only one measurement for X and the distribution of the measurement error is known. Moreover, the smoothness condition is imposed on the error's characteristic function, while the density function of X is assumed to satisfy a Lipschitz condition. In this paper, we follow Li and Vuong (1998) and KO to distinguish the following four cases[1]:

Case 1: φ_X and φ_ϵ are both OS.
Case 2: φ_X is OS and φ_ϵ is SS.
Case 3: φ_X is SS and φ_ϵ is OS.
Case 4: φ_X and φ_ϵ are both SS.

In addition, the following assumption is imposed.

Assumption 4. *(i) The kernel function K satisfies Assumption of KO.*[2] *(ii) $E[|X|^{3+\delta}] < \infty$, $E[|\epsilon|^{3+\delta}] < \infty$ for some $\delta > 0$.*

Note that Assumption 4(ii), which strengthens Assumption 1(iv), is required to apply the maximal inequality for multivariate empirical characteristic function processes (Lemma 1 of KO).[3] The next four subsections provide the uniform convergence rates of $(\hat{\varphi}_X, \hat{\varphi}_\epsilon)$ and $(\tilde{\varphi}_X, \tilde{\varphi}_\epsilon)$ for Case 1 to Case 4, which are the main theoretical results of this paper.

[1] For simplicity, in this paper a characteristic function is called OS (SS) if it satisfies Assumption 2 (Assumption 3). But it should be kept in mind that the smoothness condition is actually imposed on the density functions through the decay rate of the characteristic functions.

[2] This condition is only required for establishing the convergence rates of the estimated density functions.

[3] To apply Lemma 1 of KO with $\boldsymbol{k} = (1, 0)$, one needs $\mathbb{E}[|Y_1|^{2+\eta}|Y_2|^{1+\eta/2}] < \infty$ for some $\eta > 0$, which is ensured by Assumption 4(ii). However, KO only assumes that $\mathbb{E}|Y_1|^{2+\eta} < \infty$, which we believe is not sufficient.

3.1 Case 1: OS φ_X and OS φ_ϵ

Theorem 1. *Suppose that Assumptions 1, 2 and 4 hold, then:*

$$\begin{aligned}\sup_{|t|\le T_n} |\hat{\varphi}_X(t) - \varphi_X(t)| &= O_p\left(n^{-1/2} T_n^{\beta_x+\beta_\epsilon+1} \log T_n\right), \\ \sup_{|t|\le T_n} |\hat{\varphi}_\epsilon(t) - \varphi_\epsilon(t)| &= O_p\left(n^{-1/2} T_n^{2\beta_x+1} \log T_n\right),\end{aligned}$$

if $n^{-1/2} T_n^{2\beta_x+\beta_\epsilon+1} \log T_n \to 0$*, and*

$$\begin{aligned}\sup_{|t|\le T_n} |\tilde{\varphi}_X(t) - \varphi_X(t)| &= O_p\left(n^{-1/2} T_n^{(3\beta_\epsilon-\beta_x+1)\vee\beta_\epsilon} \log T_n\right), \\ \sup_{|t|\le T_n} |\tilde{\varphi}_\epsilon(t) - \varphi_\epsilon(t)| &= O_p\left(n^{-1/2} T_n^{2\beta_\epsilon+1} \log T_n\right),\end{aligned}$$

if $n^{-1/2} T_n^{3\beta_\epsilon+1} \log T_n \to 0$.

3.2 Case 2: OS φ_X and SS φ_ϵ

Theorem 2. *Suppose that Assumptions 1, 2(i), 3(ii) and 4 hold, then:*

$$\begin{aligned}\sup_{|t|\le T_n} |\hat{\varphi}_X(t) - \varphi_X(t)| &= O_p\left(n^{-1/2} T_n^{\beta_x-\beta_\epsilon+1} \exp(T_n^{\rho_\epsilon}/\mu_\epsilon) \log T_n\right), \\ \sup_{|t|\le T_n} |\hat{\varphi}_\epsilon(t) - \varphi_\epsilon(t)| &= O_p\left(n^{-1/2} T_n^{2\beta_x+1} \log T_n\right),\end{aligned}$$

if $n^{-1/2} T_n^{2\beta_x-\beta_\epsilon+1} \exp(T_n^{\rho_\epsilon}/\mu_\epsilon) \log T_n \to 0$*, and*

$$\begin{aligned}\sup_{|t|\le T_n} |\tilde{\varphi}_X(t) - \varphi_X(t)| &= O_p\left(n^{-1/2} T_n^{-\beta_x-3\beta_\epsilon+1} \exp(3T_n^{\rho_\epsilon}/\mu_\epsilon) \log T_n\right), \\ \sup_{|t|\le T_n} |\tilde{\varphi}_\epsilon(t) - \varphi_\epsilon(t)| &= O_p\left(n^{-1/2} T_n^{-2\beta_\epsilon+1} \exp(2T_n^{\rho_\epsilon}/\mu_\epsilon) \log T_n\right),\end{aligned}$$

if $n^{-1/2} T_n^{-3\beta_\epsilon+1} \exp(3T_n^{\rho_\epsilon}/\mu_\epsilon) \log T_n \to 0$.

3.3 Case 3: SS φ_X and OS φ_ϵ

Theorem 3. *Suppose that Assumptions 1, 2(ii), 3(i) and 4 hold, then:*

$$\begin{aligned}\sup_{|t|\le T_n} |\hat{\varphi}_X(t) - \varphi_X(t)| &= O_p\left(n^{-1/2} T_n^{-\beta_x+\beta_\epsilon+1} \exp(T_n^{\rho_x}/\mu_x) \log T_n\right), \\ \sup_{|t|\le T_n} |\hat{\varphi}_\epsilon(t) - \varphi_\epsilon(t)| &= O_p\left(n^{-1/2} T_n^{-2\beta_x+1} \exp(2T_n^{\rho_x}/\mu_x) \log T_n\right),\end{aligned}$$

if $n^{-1/2}T_n^{-2\beta_x+\beta_\epsilon+1}\exp(2T_n^{\rho_x}/\mu_x)\log T_n \to 0,$ *and*

$$\begin{aligned}\sup_{|t|\le T_n}|\tilde\varphi_X(t)-\varphi_X(t)| &= O_p\left(n^{-1/2}T_n^{\beta_\epsilon}\log T_n\right),\\ \sup_{|t|\le T_n}|\tilde\varphi_\epsilon(t)-\varphi_\epsilon(t)| &= O_p\left(n^{-1/2}T_n^{2\beta_\epsilon+1}\log T_n\right),\end{aligned}$$

if $n^{-1/2}T_n^{3\beta_\epsilon+1}\log T_n \to 0.$

3.4 Case 4: SS φ_X and SS φ_ϵ

Theorem 4. *Suppose that Assumptions 1, 3 and 4 hold, then:*

$$\begin{aligned}\sup_{|t|\le T_n}|\hat\varphi_X(t)-\varphi_X(t)| &= O_p\left(n^{-1/2}T_n^{-\beta_x-\beta_\epsilon+1}\exp(T_n^{\rho_x}/\mu_x+T_n^{\rho_\epsilon}/\mu_\epsilon)\log T_n\right),\\ \sup_{|t|\le T_n}|\hat\varphi_\epsilon(t)-\varphi_\epsilon(t)| &= O_p\left(n^{-1/2}T_n^{-2\beta_x+1}\exp(2T_n^{\rho_x}/\mu_x)\log T_n\right),\end{aligned}$$

if $n^{-1/2}T_n^{-2\beta_x-\beta_\epsilon+1}\exp(2T_n^{\rho_x}/\mu_x+T_n^{\rho_\epsilon}/\mu_\epsilon)\log T_n \to 0,$ *and*

$$\begin{aligned}\sup_{|t|\le T_n}|\tilde\varphi_X(t)-\varphi_X(t)| &= O_p\left(n^{-1/2}(a_n\vee b_n)\log T_n\right),\\ \sup_{|t|\le T_n}|\tilde\varphi_\epsilon(t)-\varphi_\epsilon(t)| &= O_p\left(n^{-1/2}T_n^{-2\beta_\epsilon+1}\exp(2T_n^{\rho_\epsilon}/\mu_\epsilon)\log T_n\right),\end{aligned}$$

if $n^{-1/2}T_n^{-3\beta_\epsilon+1}\exp(3T_n^{\rho_\epsilon}/\mu_\epsilon)\log T_n \to 0,$ *where*

$$a_n = T_n^{\beta_x-3\beta_\epsilon+1}\exp(3T_n^{\rho_\epsilon}/\mu_\epsilon - T_n^{\rho_x}/\mu_x) \quad \textit{and} \quad b_n = T_n^{-\beta_\epsilon}\exp(T_n^{\rho_\epsilon}/\mu_\epsilon).$$

3.5 Summary of Results and Some Remarks

The uniform convergence rates of $(\hat\varphi_X,\hat\varphi_\epsilon)$ and $(\tilde\varphi_X,\tilde\varphi_\epsilon)$ for Case 1 to Case 4 are summarized in Table 1 below. For comparison, we also include the convergence rates established in KO. There are three main takeaways from the results in Table 1.

First, for the LV estimators, the rates obtained in this paper are obviously faster than those obtained in KO, which have been shown to be faster than the rates proved by Li and Vuong (1998) and Bonhomme and Robin (2010) under more restrictive conditions (see Remark 1 and Remark 2 of KO). For example, if we set $T_n = O\left((n/\log\log n)^{\alpha/2(1+\beta_x+\beta_\epsilon)}\right)$ with $0<\alpha<1/2$ as in Li and Vuong (1998), then our Theorem 1 implies that

$$\sup_{|t|\le T_n}|\hat\varphi_X(t)-\varphi_X(t)| = O_P\left(\left(\frac{n}{\log\log n}\right)^{-\frac{1}{2}+\frac{\alpha}{2}}\right),$$

while the results of KO and Li and Vuong (1998) are

$$O_P\left(\left(\frac{n}{\log\log n}\right)^{-\frac{1}{2}+\alpha-\frac{\alpha}{2(1+\beta_x+\beta_\epsilon)}}\right) \qquad \text{and} \qquad O_P\left(\left(\frac{n}{\log\log n}\right)^{-\frac{1}{2}+\alpha}\right)$$

respectively.

Second, depending on the values of (μ_x, ρ_x, β_x) and $(\mu_\epsilon, \rho_\epsilon, \beta_\epsilon)$, the convergence rates of $(\tilde{\varphi}_X, \tilde{\varphi}_\epsilon)$ can be faster than those of $(\hat{\varphi}_X, \hat{\varphi}_\epsilon)$, especially in Case 3 where the rates of the $(\hat{\varphi}_X, \hat{\varphi}_\epsilon)$ are exponential functions of T_n while the rates of the $(\tilde{\varphi}_X, \tilde{\varphi}_\epsilon)$ are polynomial functions of T_n.

Third, with the faster convergence rates obtained in this paper, some interesting new insights can be gained. It can be seen that for $\hat{\varphi}_\epsilon$, the convergence rates only depends on the smoothness of the distribution of X, while for $\tilde{\varphi}_\epsilon$, the convergence rates only depends on the smoothness of the distribution of ϵ. Thus, if one is interested in estimating the distribution of the measurement errors, $\hat{\varphi}_\epsilon$ should be used if φ_X is OS, while $\tilde{\varphi}_\epsilon$ is preferred if φ_ϵ is OS. In the case where both φ_X and φ_ϵ are OS or SS, these two estimators are very similar in terms of convergence rates.

Remark 1 (The density functions)**.** *To save space, we only report the results for the estimated characteristic functions. Given Theorem 1 to Theorem 4, faster uniform convergence rates of the estimated densities can be easily obtained. For example, with $T_n = h_n^{-1}$, in Case 1, we can show that the uniform convergence rates of $\hat{f}_X$ and $\hat{f}_\epsilon$ are*

$$O_p\left(n^{-1/2}T_n^{\beta_x+\beta_\epsilon+2}\log T_n + T_n^{1-\beta_x}\right) \qquad \textit{and} \qquad O_p\left(n^{-1/2}T_n^{2\beta_x+2}\log T_n + T_n^{1-\beta_\epsilon}\right)$$

respectively. The proofs are similar to those of KO, and therefore they are omitted.

Remark 2 (Regularized estimators)**.** *Comte and Kappus (2015) proposed a regularized version of the LV estimators. KO showed that these regularized estimators have the same uniform convergence rates as the LV estimators. With very small modifications of our proofs, it can be shown that the uniform convergence rates established in this paper also apply to the regularized LV estimators.*

Table 1: Comparison of Uniform Convergence Rates

	Case 1	Case 2	Case 3	Case 4
$\hat{\varphi}_{X,KO}$	$T_n^{2\beta_x+2\beta_\epsilon}$	$T_n^{2\beta_x-2\beta_\epsilon}\exp(2T_n^{\rho_\epsilon}/\mu_\epsilon)$	$T_n^{-2\beta_x+2\beta_\epsilon}\exp(2T_n^{\rho_x}/\mu_x)$	$T_n^{-2\beta_x-2\beta_\epsilon}\exp(2T_n^{\rho_x}/\mu_x+2T_n^{\rho_\epsilon}/\mu_\epsilon)$
$\hat{\varphi}_X$	$T_n^{\beta_x+\beta_\epsilon}$	$T_n^{\beta_x-\beta_\epsilon}\exp(T_n^{\rho_\epsilon}/\mu_\epsilon)$	$T_n^{-\beta_x+\beta_\epsilon}\exp(T_n^{\rho_x}/\mu_x)$	$T_n^{-\beta_x-\beta_\epsilon}\exp(T_n^{\rho_x}/\mu_x+T_n^{\rho_\epsilon}/\mu_\epsilon)$
$\tilde{\varphi}_X$	$T_n^{(3\beta_\epsilon-\beta_x)\vee(\beta_\epsilon-1)}$	$T_n^{-\beta_x-3\beta_\epsilon}\exp(3T_n^{\rho_\epsilon}/\mu_\epsilon)$	$T_n^{\beta_\epsilon-1}$	$\tilde{a}_n\vee\tilde{b}_n$
$\hat{\varphi}_{\epsilon,KO}$	$T_n^{3\beta_x+2\beta_\epsilon}$	$T_n^{3\beta_x-2\beta_\epsilon}\exp(2T_n^{\rho_\epsilon}/\mu_\epsilon)$	$T_n^{-3\beta_x+2\beta_\epsilon}\exp(3T_n^{\rho_x}/\mu_x)$	$T_n^{-3\beta_x-2\beta_\epsilon}\exp(3T_n^{\rho_x}/\mu_x+2T_n^{\rho_\epsilon}/\mu_\epsilon)$
$\hat{\varphi}_\epsilon$	$T_n^{2\beta_x}$	$T_n^{2\beta_x}$	$T_n^{-2\beta_x}\exp(2T_n^{\rho_x}/\mu_x)$	$T_n^{-2\beta_x}\exp(2T_n^{\rho_x}/\mu_x)$
$\tilde{\varphi}_\epsilon$	$T_n^{2\beta_\epsilon}$	$T_n^{-2\beta_\epsilon}\exp(2T_n^{\rho_\epsilon}/\mu_\epsilon)$	$T_n^{2\beta_\epsilon}$	$T_n^{-2\beta_\epsilon}\exp(2T_n^{\rho_\epsilon}/\mu_\epsilon)$

Note: This table presents the uniform convergence rates of $(\hat{\varphi}_X,\hat{\varphi}_\epsilon)$ and $(\tilde{\varphi}_X,\tilde{\varphi}_\epsilon)$ for Case 1 to Case 4, up to a common factor: $n^{-1/2}T_n\log T_n$. For comparison, we also include the rates established in KO, which are denoted by $\hat{\varphi}_{\cdot,KO}$. $\tilde{a}_n = T_n^{\beta_x-3\beta_\epsilon}\exp(3T_n^{\rho_\epsilon}/\mu_\epsilon - T_n^{\rho_x}/\mu_x)$, and $\tilde{b}_n = T_n^{-\beta_\epsilon-1}\exp(T_n^{\rho_\epsilon}/\mu_\epsilon)$.

4 A Heuristic Explanation for the Faster Convergence Rate

In this section, we provide a heuristic explanation of why the faster convergence rates can be obtained for the LV estimators in this paper. The discussion below focuses on Case 1 and the estimation of φ_X, but the idea behind applies to the other three cases and the other estimators.

Define $\hat{a}(s) = n^{-1}\sum_{j=1}^n \mathrm{i}Y_{j1}e^{\mathrm{i}sY_{j2}}$, $a(s)=\mathbb{E}[\mathrm{i}Y_1e^{\mathrm{i}sY_2}]$, $\hat{b}(s)=n^{-1}\sum_{j=1}^n e^{\mathrm{i}sY_{j2}}$, $b(s)=\mathbb{E}[e^{\mathrm{i}sY_2}]$, and

$$\Delta(t)=\int_0^t\left(\frac{\hat{a}(s)}{\hat{b}(s)}-\frac{a(s)}{b(s)}\right)ds,\quad \Delta_1(t)=\int_0^t\frac{\hat{a}(s)-a(s)}{b(s)}ds,$$

$$\Delta_2(t)=\int_0^t a(s)\left(\frac{1}{\hat{b}(s)}-\frac{1}{b(s)}\right)ds,\quad \Delta_3(t)=\int_0^t(\hat{a}(s)-a(s))\left(\frac{1}{\hat{b}(s)}-\frac{1}{b(s)}\right)ds.$$

Note that $\Delta(t)=\Delta_1(t)+\Delta_2(t)+\Delta_3(t)$. Then it can be shown that (see equation A.2 of KO)[4]

$$\begin{aligned}|\hat{\varphi}_X(t)-\varphi_X(t)| &\le 2|\varphi_X(t)|\cdot|\Delta(t)|\cdot\mathbb{I}\{|\Delta(t)|\le 1\}+2\mathbb{I}\{|\Delta(t)|>1\} &(4)\\ &\le 4|\Delta(t)|\lesssim|\Delta_1(t)|+|\Delta_2(t)|+|\Delta_3(t)|. &(5)\end{aligned}$$

First, if one follows the proof of KO that starts with inequality (5), then we have:

$$\sup_{|t|\le T_n}|\hat{\varphi}_X(t)-\varphi_X(t)|\lesssim\sup_{|t|\le T_n}|\Delta_1(t)|+\sup_{|t|\le T_n}|\Delta_2(t)|+\sup_{|t|\le T_n}|\Delta_3(t)|.$$

Compared to the proof of KO, this paper establishes a faster convergence rate for the second term on the right-hand side of the above inequality, which turns out to be the dominating term.

[4]To facilitate the discussion we assume that $|\hat{\varphi}_X(t)|\le 1$ here.

Note that

$$\begin{aligned}
\sup_{|t|\le T_n} |\Delta_2(t)| &\le \sup_{|t|\le T_n} \int_0^t \left| \frac{a(s)[\hat{b}(s)-b(s)]}{\hat{b}(s)b(s)} \right| ds & (6)\\
&\le \sup_{|t|\le T_n} \left|\hat{b}(t)-b(t)\right| \cdot \frac{1}{\inf_{|t|\le T_n} |\hat{b}(t)|} \cdot \underbrace{\int_0^{T_n} \left|\frac{a(s)}{b(s)}\right| ds}_{O_P(T_n^{\beta_x+1})} & (7)\\
&\le \underbrace{\sup_{|t|\le T_n} \left|\hat{b}(t)-b(t)\right|}_{O_P(n^{-1/2}\log T_n)} \cdot \underbrace{\frac{1}{\inf_{|t|\le T_n} |\hat{b}(t)|}}_{O_P(T_n^{\beta_x+\beta_\epsilon})} \cdot \underbrace{\frac{1}{\inf_{|t|\le T_n} |b(t)|}}_{O_P(T_n^{\beta_x+\beta_\epsilon})} \cdot \underbrace{\int_0^{T_n} |a(s)|\, ds}_{O_P(T_n)}. & (8)
\end{aligned}$$

The main difference is that we use the sharper inequality (7) to derive the rate while KO use (8). In particular, the first two terms of (7) and (8) can be shown to be $O_P(n^{-1/2}\log T_n)$ and $O_P(T_n^{\beta_x+\beta_\epsilon})$ respectively. Moreover, the third term of (8) is $O_P(T_n^{\beta_x+\beta_\epsilon})$ and the last term of (8) is $O_P(T_n)$. This gives the derived uniform convergence rate for $\hat{\varphi}_X(t)$ in KO. In comparison, we show that the third term of (7) is $O_P(T_n^{\beta+1})$, leading to a faster convergence rate.

Intuitively, the reason why (7) provides a better bound is that the large values of $1/|b(s)|$ is partially offset by small values of $|a(s)|$ when s approaches infinity. Such an effect is ignored when one uses (8) and considers the uniform bounds for $1/|b(s)|$ and $|a(s)|$ separately. To see this, note that

$$\begin{aligned}
\frac{a(s)}{b(s)} &= \frac{\mathrm{i}\mathbb{E}[Y_1 e^{\mathrm{i}sY_2}]}{\mathbb{E}[e^{\mathrm{i}sY_2}]} = \frac{\mathrm{i}\mathbb{E}[X e^{\mathrm{i}sY_2}] + \mathrm{i}\mathbb{E}[\epsilon_1 e^{\mathrm{i}sY_2}]}{\mathbb{E}[e^{\mathrm{i}sY_2}]}\\
&= \frac{\mathrm{i}\mathbb{E}[X e^{\mathrm{i}sX} e^{\mathrm{i}s\epsilon_2}] + \mathrm{i}\mathbb{E}[\epsilon_1 e^{\mathrm{i}sX} e^{\mathrm{i}s\epsilon_2}]}{\mathbb{E}[e^{\mathrm{i}sX} e^{\mathrm{i}s\epsilon_2}]}\\
&= \frac{\mathrm{i}\mathbb{E}[X e^{\mathrm{i}sX}]}{\mathbb{E}[e^{\mathrm{i}sX}]}.
\end{aligned}$$

Thus, it follows that

$$\int_0^{T_n} \left|\frac{a(s)}{b(s)}\right| ds \le T_n \cdot \mathbb{E}[|X|] \cdot \frac{1}{\inf_{|t|\le T_n} |\varphi_X(t)|} = O_P(T_n^{\beta_x+1}).$$

which gives the rate in (7).

Second, our proof starts with inequality (4) instead of (5) to show that the uniform convergence rate of $\hat{\varphi}_X(t)$ is actually decided by $\sup_{|t|\le T_n} |\varphi_X(t)| \cdot |\Delta_2(t)|$ rather than $\sup_{|t|\le T_n} |\Delta_2(t)|$. Since $|\varphi_X(t)|$ is very small as $|t|$ becomes large, the convergence rate can be further improved.

Remark 3 (Tighter bound for $a(s)/b(s)$)**.** *The discussion above implies that the uniform convergence rate for $\hat{\varphi}_X$ can be further improved if a sharper bound on $a(s)/b(s)$ can be obtained.*

For example:

(i) When $X \sim$*Normal* (μ, σ^2) *(SS distribution), it can be shown that* $a(s)/b(s) = -\sigma^2 s + \mathrm{i}\mu$ *and* $|a(s)/b(s)| \leq C(1+|s|)$ *for some* $C > 0$*, thus the last term of* (7) *becomes* $O_P(T_n^2)$ *instead of* $O_P(T_n^{1-\beta_x} \exp(T_n^{\rho_x}/\mu_x))$.

(ii) When $X \sim$*Cauchy* (μ, θ) *(SS distribution), it can be shown that* $a(s)/b(s) = \mathrm{i}\mu - sgn(s)\theta$ *and* $|a(s)/b(s)| \leq C$ *for some* $C > 0$*, thus the last term of* (7) *becomes* $O_P(T_n)$ *instead of* $O_P(T_n^{1-\beta_x} \exp(T_n^{\rho_x}/\mu_x))$.

(iii) When X *follows Student's* t *distribution with* ν *degrees of freedom (SS distribution), it can be shown that* $a(s)/b(s) = -\sqrt{\nu} \cdot sgn(s) \cdot K_{\nu/2-1}(\sqrt{\nu}|s|)/K_{\nu/2}(\sqrt{\nu}|s|)$*, where* K_ν *is the modified Bessel function of the second kind, and* $|a(s)/b(s)| \leq C$ *for some* $C > 0$*, thus the last term of* (7) *becomes* $O_P(T_n)$ *instead of* $O_P(T_n^{1-\beta_x} \exp(T_n^{\rho_x}/\mu_x))$.

(iv) When $X \sim$*Laplace* (μ, b) *(OS distribution), it can be shown that* $a(s)/b(s) = -2b^2 s/(1+b^2 s^2) + \mathrm{i}\mu$ *and* $|a(s)/b(s)| \leq C$ *for some* $C > 0$*, thus the last term of* (7) *becomes* $O_P(T_n)$ *instead of* $O_P(T_n^{\beta_x+1})$.

(v) When $X \sim$*Gamma* (k, θ) *(OS distribution), it can be shown that* $a(s)/b(s) = \mathrm{i}k\theta/(1-\mathrm{i}s\theta)$ *and* $|a(s)/b(s)| \leq C$ *for some* $C > 0$*, thus the last term of* (7) *becomes* $O_P(T_n)$ *instead of* $O_P(T_n^{\beta_x+1})$.

Based on the above examples, a natural conjecture is that $|a(s)/b(s)| < C$ *when* φ_X *is OS and* $|a(s)/b(s)| < C(1+|s|^\gamma)$ *when* φ_X *is SS, where* $C > 0$ *and* $\gamma \geq 0$ *are some constants. If this conjecture is true,*[5] *it is obvious from* (7) *that a faster uniform convergence rate can be obtained for* $\hat{\varphi}_X$*, especially when* φ_X *is SS. However, a formal justification for this conjecture is beyond the scope of this paper.*

5 Conclusions

As pointed out in the concluding remarks of KO, deconvolution estimators based on Kotlarski's identity are usually used as nonparametric plug-in components in many other studies (see Li 2002, Li and Hsiao 2004, Adusumilli et al. 2020 and Otsu and Taylor 2021 for examples). Thus a faster uniform convergence rate for these estimators is essential for understanding the asymptotic properties of the related test statistics and semiparametric estimators. Based on a maximal inequality for the multivariate normalized empirical characteristic function process developed in KO, this paper shows that the LV estimators can achieve faster uniform convergence rates than those obtained in existing studies. Moreover, a new class of nonparametric deconvolution estimators is proposed based on a variant of Kotlarski's identity, and these estimators are shown to have faster convergence rates than the LV estimators in some cases. Finally, we point out that

[5] A similar condition is imposed in Assumption ASYM of Evdokimov (2010).

a even faster uniform convergence rate can be achieved by imposing a inequality for ordinary smooth and supersmooth distributions. However, a formal justification of this inequality is beyond the scope of this paper and is left for future research.

A Proofs of the Main Results

A.1 Proof of Lemma 2

Proof of Lemma 2. First note that

$$\varphi_{\epsilon_1}(t) = \exp\left(\ln \varphi_{\epsilon_1}(t) - \ln 1\right) = \exp\left(\int_0^t \frac{\partial \ln \varphi_{\epsilon_1}(s)}{\partial s} ds\right) = \exp\left(\int_0^t \frac{\mathrm{i}\mathbb{E}[\epsilon_1 e^{\mathrm{i}s\epsilon_1}]}{\mathbb{E}[e^{\mathrm{i}s\epsilon_1}]} ds\right),$$

where the integrand is well-defined due to Assumption 1.

Next, we have

$$\begin{aligned}\frac{\mathbb{E}[\epsilon_1 e^{\mathrm{i}s\epsilon_1}]}{\mathbb{E}[e^{\mathrm{i}s\epsilon_1}]} =& \frac{\mathbb{E}[\epsilon_1 e^{\mathrm{i}s\epsilon_1}] \cdot \mathbb{E}[e^{-\mathrm{i}s\epsilon_2}]}{\mathbb{E}[e^{\mathrm{i}s\epsilon_1}] \cdot \mathbb{E}[e^{-\mathrm{i}s\epsilon_2}]} = \frac{\mathbb{E}[\epsilon_1 e^{\mathrm{i}s(Y_1-Y_2)}] + \mathbb{E}[X e^{\mathrm{i}s(Y_1-Y_2)}] - \mathbb{E}[X e^{\mathrm{i}s(Y_1-Y_2)}]}{\mathbb{E}[e^{\mathrm{i}s(Y_1-Y_2)}]} \\ =& \frac{\mathbb{E}[Y_1 e^{\mathrm{i}s(Y_1-Y_2)}]}{\mathbb{E}[e^{\mathrm{i}s(Y_1-Y_2)}]} - \mathbb{E}[X] = \frac{\mathbb{E}[Y_1 e^{\mathrm{i}s(Y_1-Y_2)}]}{\mathbb{E}[e^{\mathrm{i}s(Y_1-Y_2)}]} - \mathbb{E}[Y_1].\end{aligned}$$

This completes the proof.[6] □

A.2 Proof of Theorem 1

Recall that $a(s), b(s), \hat{a}(s), \hat{b}(s), \Delta_1(t), \Delta_2(t), \Delta_3(t)$ are defined in Section 4.

Lemma 3. *Suppose that Assumptions 1, 2 and 4 hold, then:*
(i) $\sup_{|t|\le T_n} |\Delta_1(t)| = O_P(n^{-1/2} T_n^{\beta_x+\beta_\epsilon+1} \log T_n)$.
(ii) $\sup_{|t|\le T_n} |\Delta_2(t)| = O_P(n^{-1/2} T_n^{2\beta_x+\beta_\epsilon+1} \log T_n)$.
(iii) $\sup_{|t|\le T_n} |\Delta_3(t)| = O_P(n^{-1} T_n^{2\beta_x+2\beta_\epsilon+1} (\log T_n)^2)$.

Proof. First, for $\Delta_1(t)$, note that

$$\sup_{|t|\le T_n} |\Delta_1(t)| \le \sup_{|t|\le T_n} |\hat{a}(t) - a(t)| \cdot \int_0^{T_n} \frac{1}{|b(s)|} ds.$$

Lemma 1 of KO and Assumption 2 imply that

$$\sup_{|t|\le T_n} |\hat{a}(t) - a(t)| = O_P(n^{-1/2} \log T_n), \quad \text{and} \quad \int_0^{T_n} \frac{1}{|b(s)|} ds = O(T_n^{\beta_x+\beta_\epsilon+1}).$$

Thus, the first result follows.

[6] Note that in this proof we only need the characteristic function of ϵ to be non-vanishing everywhere.

Second, as discussed in Section 4, we have

$$\sup_{|t|\le T_n} |\Delta_2(t)| \le \sup_{|t|\le T_n} |\hat{b}(t) - b(t)| \cdot \frac{1}{\inf_{|t|\le T_n} |\hat{b}(t)|} \cdot \int_0^{T_n} \left| \frac{a(s)}{b(s)} \right| ds.$$

From the discussion in Section 4, we have

$$\int_0^{T_n} \left| \frac{a(s)}{b(s)} \right| ds \le T_n \cdot \mathbb{E}[|X|] \cdot \frac{1}{\inf_{|t|\le T_n} |\varphi_X(t)|} = O_P(T_n^{\beta_x+1}).$$

Moreover, Lemma 1 of KO and Assumption 2 imply that

$$\sup_{|t|\le T_n} |\hat{b}(t) - b(t)| = O_P(n^{-1/2}\log T_n), \quad \text{and} \quad \frac{1}{\inf_{|t|\le T_n} |\hat{b}(t)|} = O_P(T_n^{\beta_x+\beta_\epsilon}).$$

Thus, the second result follows.

Finally, the last result follows from

$$\sup_{|t|\le T_n} |\Delta_3(t)| \le T_n \cdot \sup_{|t|\le T_n} |\hat{a}(t) - a(t)| \cdot \sup_{|t|\le T_n} |\hat{b}(t) - b(t)| \cdot \frac{1}{\inf_{|t|\le T_n} |\hat{b}(t)|} \cdot \frac{1}{\inf_{|t|\le T_n} |b(t)|}.$$

$\square$

Proof of Theorem 1. We first establish the convergence rates for $\hat{\varphi}_X$ and $\hat{\varphi}_\epsilon$.

Step 1: Convergence rate for $\hat{\varphi}_X$

Similar to the proof of Theorem 1 in KO, one can show that

$$|\hat{\varphi}_X(t) - \varphi_X(t)| \le 2|\varphi_X(t)||\Delta(t)| + |\hat{\varphi}_X(t) - \varphi_X(t)|\,\mathbb{I}\{|\Delta(t)| > 1\}$$

Thus, for any sequence $\{c_n\}$ of positive constants that converges to 0 as n diverges, we have

$$P\left[\sup_{|t|\le T_n} |\hat{\varphi}_X(t) - \varphi_X(t)| > c_n\right] \le P\left[\sup_{|t|\le T_n} |\varphi_X(t)||\Delta(t)| > c_n/4\right] + P\left[\sup_{|t|\le T_n} \mathbb{I}\{|\Delta(t)| > 1\} > 0\right].$$

Note that

$$P\left[\sup_{|t|\le T_n} \mathbb{I}\{|\Delta(t)| > 1\} > 0\right] \le P\left[\sup_{|t|\le T_n} |\Delta(t)| > 1\right],$$

and the right-hand side of the above inequality goes to 0 by Lemma 3. Then, to establish the uniform convergence rate of $\hat{\varphi}_X(t)$, it suffices to show that

$$\sup_{|t|\le T_n} |\varphi_X(t)||\Delta(t)| = O_P(n^{-1/2}T_n^{\beta_x+\beta_\epsilon+1}\log T_n). \tag{A.1}$$

Let $\lambda_x \in [\omega_x, T_n]$ such that $C_x\lambda_x^{-\beta_x} \leq \inf_{|s|\leq\omega_x} |\varphi_X(s)|$, and let $\lambda_\epsilon \in [\omega_\epsilon, T_n]$ such that $C_\epsilon\lambda_\epsilon^{-\beta_\epsilon} \leq \inf_{|s|\leq\omega_\epsilon} |\varphi_\epsilon(s)|$. Then for $\lambda = \max\{\lambda_x, \lambda_\epsilon\}$ and any $t \in [\lambda, T_n]$, we have $\inf_{|s|\leq t} |\varphi_X(s)| = \inf_{\omega_x\leq|s|\leq t} |\varphi_X(s)|$ because

$$\inf_{\omega_x\leq|s|\leq t} |\varphi_X(s)| \leq \inf_{\lambda\leq|s|\leq t} |\varphi_X(s)| \leq C_x\lambda^{-\beta_x} \leq C_x\lambda_x^{-\beta_x} \leq \inf_{|s|\leq\omega_x} |\varphi_X(s)|.$$

Thus, it follows from Assumption 2 that

$$\inf_{|s|\leq t} |\varphi_X(s)| \geq c_x t^{-\beta_x} \text{ for any } t \in [\lambda, T_n]. \tag{A.2}$$

Similarly, it can be shown that

$$\inf_{|s|\leq t} |\varphi_\epsilon(s)| \geq c_\epsilon t^{-\beta_\epsilon} \text{ for any } t \in [\lambda, T_n]. \tag{A.3}$$

To prove (A.1), note that

$$\sup_{|t|\leq T_n} |\varphi_X(t)||\Delta(t)| \leq \sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)||\Delta(t)| + \sup_{|t|\leq\lambda} |\varphi_X(t)| \cdot \sup_{|t|\leq\lambda} |\Delta(t)|, \tag{A.4}$$

where the second term on the right-hand side of the above inequality is $O_P(n^{-1/2})$ by Lemma 3 and the property of characteristic functions, and the first term is bounded by

$$\begin{aligned}\sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)||\Delta(t)| \leq \sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)||\Delta_1(t)| \\ + \sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)||\Delta_2(t)| + \sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)||\Delta_3(t)|.\end{aligned} \tag{A.5}$$

By the definition of $\Delta_2(t)$, we have

$$\begin{aligned}
& \sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)|\,|\Delta_2(t)| \\
\leq\ & \sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)| \int_0^{|t|} \left| \frac{a(s)[\hat{b}(s) - b(s)]}{\hat{b}(s)b(s)} \right| ds \\
\leq\ & \sup_{\lambda\leq|t|\leq T_n} |\varphi_X(t)| \cdot \sup_{|s|\leq|t|} \left| \frac{a(s)}{b(s)^2} \right| \cdot |t| \sup_{|s|\leq|t|} \left| \frac{b(s)[\hat{b}(s) - b(s)]}{\hat{b}(s)} \right| \\
\leq\ & \sup_{|s|\leq T_n} \left| \frac{b(s)[\hat{b}(s) - b(s)]}{\hat{b}(s)} \right| \cdot \sup_{\lambda\leq|t|\leq T_n} |t| \cdot |\varphi_X(t)| \cdot g_t \\
\leq\ & C_x \cdot \sup_{|s|\leq T_n} \left| \frac{b(s)[\hat{b}(s) - b(s)]}{\hat{b}(s)} \right| \cdot \sup_{\lambda\leq|t|\leq T_n} |t|^{1-\beta_x} \cdot g_t
\end{aligned}$$

where

$$g_t := \sup_{|s|\le|t|}\left|\frac{a(s)}{b(s)^2}\right| \le \sup_{|s|\le|t|}\left|\frac{a(s)}{b(s)}\right| \cdot \sup_{|s|\le|t|}\left|\frac{1}{b(s)}\right| \le \frac{\sup_{|s|\le|t|}\left|\mathrm{i}\mathbb{E}[Xe^{\mathrm{i}sX}]\right|}{\inf_{|s|\le|t|}|\varphi_X(s)|^2\cdot\inf_{|s|\le|t|}|\varphi_\epsilon(s)|}.$$

By (A.2) and (A.3), for $|t|\ge\lambda$ we have $g_t \lesssim |t|^{2\beta_x+\beta_\epsilon}$. It follows that

$$\sup_{\lambda\le|t|\le T_n}|t|^{1-\beta_x}\cdot g_t \lesssim T_n^{\beta_x+\beta_\epsilon+1}.$$

Moreover, similar to the proof of Lemma 3, it can be shown that

$$\sup_{|s|\le T_n}\left|\frac{b(s)[\hat{b}(s)-b(s)]}{\hat{b}(s)}\right| = O_P(n^{-1/2}\log T_n).$$

Thus we have

$$\sup_{\lambda\le|t|\le T_n}|\varphi_X(t)|\,|\Delta_2(t)| = O_P(n^{-1/2}T_n^{\beta_x+\beta_\epsilon+1}\log T_n). \tag{A.6}$$

Finally, it follows from Lemma 3 that

$$\sup_{\lambda\le|t|\le T_n}|\varphi_X(t)|\,|\Delta_1(t)| \le \sup_{|t|\le T_n}|\Delta_1(t)| = O_P(n^{-1/2}T_n^{\beta_x+\beta_\epsilon+1}\log T_n), \tag{A.7}$$

and

$$\sup_{\lambda\le|t|\le T_n}|\varphi_X(t)|\,|\Delta_3(t)| \le \sup_{|t|\le T_n}|\Delta_3(t)| = o_P(n^{-1/2}T_n^{\beta_x+\beta_\epsilon+1}\log T_n). \tag{A.8}$$

Therefore, (A.1) follows from (A.4) to (A.8), and this completes the proof for $\hat{\varphi}_X$.

Step 2: Convergence rate for $\hat{\varphi}_\epsilon$

Note that

$$\begin{aligned}\hat{\varphi}_\epsilon(t)-\varphi_\epsilon(t) =& \frac{\hat{\varphi}_{Y_1}(t)}{\hat{\varphi}_X(t)} - \frac{\varphi_{Y_1}(t)}{\varphi_X(t)}\\ =& \frac{\hat{\varphi}_{Y_1}(t)-\varphi_{Y_1}(t)}{\varphi_X(t)} + \varphi_{Y_1}(t)\cdot\left(\frac{1}{\hat{\varphi}_X(t)}-\frac{1}{\varphi_X(t)}\right) + (\hat{\varphi}_{Y_1}(t)-\varphi_{Y_1}(t))\cdot\left(\frac{1}{\hat{\varphi}_X(t)}-\frac{1}{\varphi_X(t)}\right)\\ :=& \Delta_4(t)+\Delta_5(t)+\Delta_6(t).\end{aligned}$$

First, by Lemma 1 of KO and Assumption 2,

$$\sup_{|t|\le T_n}|\Delta_4(t)| \le \frac{\sup_{|t|\le T_n}|\hat{\varphi}_{Y_1}(t)-\varphi_{Y_1}(t)|}{\inf_{|t|\le T_n}|\varphi_X(t)|} = O_P(n^{-1/2}T_n^{\beta_x}\log T_n). \tag{A.9}$$

Second, by the result of the previous step,

$$\sup_{|t|\le T_n} |\Delta_6(t)| \le \frac{\sup_{|t|\le T_n} |\hat{\varphi}_{Y_1}(t) - \varphi_{Y_1}(t)| \cdot \sup_{|t|\le T_n} |\hat{\varphi}_X(t) - \varphi_X(t)|}{\inf_{|t|\le T_n} |\varphi_X(t)| \cdot \inf_{|t|\le T_n} |\hat{\varphi}_X(t)|}$$
$$= O_P(n^{-1/2} T_n^{2\beta_x+1} \log T_n) \cdot O_P(n^{-1/2} T_n^{\beta_x+\beta_\epsilon} \log T_n). \quad \text{(A.10)}$$

Finally, we will show that

$$\sup_{|t|\le T_n} |\Delta_5(t)| = O_P(n^{-1/2} T_n^{2\beta_x+1} \log T_n), \quad \text{(A.11)}$$

which together with (A.9) and (A.10) lead to the desired result.

To prove (A.11), note that

$$\begin{aligned}
\sup_{|t|\le T_n} |\Delta_5(t)| &= \sup_{|t|\le T_n} \left| \frac{\varphi_{Y_1}(t)\,(\hat{\varphi}_X(t) - \varphi_X(t))}{\hat{\varphi}_X(t)\varphi_X(t)} \right| \\
&= \sup_{|t|\le T_n} \left\{ \left| \frac{\varphi_\epsilon(t)}{\varphi_X(t)} \right| \cdot |\hat{\varphi}_X(t) - \varphi_X(t)| \cdot \left| \frac{\varphi_X(t)}{\hat{\varphi}_X(t)} \right| \right\} \\
&\le \sup_{|t|\le T_n} \left\{ \left| \frac{\varphi_\epsilon(t)}{\varphi_X(t)} \right| \cdot |\hat{\varphi}_X(t) - \varphi_X(t)| \right\} \cdot \sup_{|t|\le T_n} \left| \frac{\varphi_X(t)}{\hat{\varphi}_X(t)} \right|.
\end{aligned}$$

Since, $|\hat{\varphi}_X(t) - \varphi_X(t)| \le 2|\varphi_X(t)| \cdot |\Delta(t)| + |\hat{\varphi}_X(t) - \varphi_X(t)|\, \mathbb{I}(|\Delta(t)| > 1)$, it follows that

$$\sup_{|t|\le T_n} |\Delta_5(t)| \le 2 \sup_{|t|\le T_n} \{|\varphi_\epsilon(t)| \cdot |\Delta(t)|\} \cdot \sup_{|t|\le T_n} \left| \frac{\varphi_X(t)}{\hat{\varphi}_X(t)} \right|$$
$$+ \sup_{|t|\le T_n} \left\{ \left| \frac{\varphi_\epsilon(t)}{\varphi_X(t)} \right| \right\} \cdot \sup_{|t|\le T_n} \mathbb{I}(|\Delta(t)| > 1) \cdot O_P(1). \quad \text{(A.12)}$$

Since Lemma 3 implies that $\sup_{|t|\le T_n} |\Delta(t)| = o_P(1)$, the second term on the right-hand side of (A.12) is equal to 0 with probability approaching 1. It is easy to show that $\sup_{|t|\le T_n} |\varphi_X(t)/\hat{\varphi}_X(t)| = O_P(1)$. Moreover, similar to the proof of the previous step, it can be shown that

$$\begin{aligned}
& \sup_{|t|\le T_n} \{|\varphi_\epsilon(t)| \cdot |\Delta(t)|\} \\
\le\; & \sup_{|t|\le T_n} \{|\varphi_\epsilon(t)| \cdot |\Delta_1(t)|\} + \sup_{|t|\le T_n} \{|\varphi_\epsilon(t)| \cdot |\Delta_2(t)|\} + \sup_{|t|\le T_n} \{|\varphi_\epsilon(t)| \cdot |\Delta_3(t)|\} \\
=\; & O_P(n^{-1/2} T_n^{\beta_x+1} \log T_n) + O_P(n^{-1/2} T_n^{2\beta_x+1} \log T_n) + o_P(n^{-1/2} T_n^{2\beta_x+1} \log T_n).
\end{aligned}$$

This completes the proof.

Next, we establish the convergence rates for $\tilde{\varphi}_X$ and $\tilde{\varphi}_\epsilon$.

Step 3: Convergence rate for $\tilde{\varphi}_\epsilon$

Define

$$c(s) = \mathbb{E}\left[\mathrm{i}Y_1 e^{\mathrm{i}s(Y_1-Y_2)}\right], \quad \tilde{c}(s) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{i}Y_{j1} e^{\mathrm{i}s(Y_{j1}-Y_{j2})},$$

$$d(s) = \mathbb{E}\left[e^{\mathrm{i}s(Y_1-Y_2)}\right], \quad \tilde{d}(s) = \frac{1}{n}\sum_{j=1}^{n} e^{\mathrm{i}s(Y_{j1}-Y_{j2})},$$

$$\tilde{\Delta}(t) = \int_0^t \left(\frac{\tilde{c}(s)}{\tilde{d}(s)} - \frac{c(s)}{d(s)}\right) ds - \frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{i}tY_{j1} - \mathbb{E}[\mathrm{i}tY_{j1}]\right),$$

$$\tilde{\Delta}_1(t) = \int_0^t \frac{\tilde{c}(s) - c(s)}{d(s)} ds, \quad \tilde{\Delta}_2(t) = \int_0^t c(s)\left(\frac{1}{\tilde{d}(s)} - \frac{1}{d(s)}\right) ds,$$

$$\tilde{\Delta}_3(t) = \int_0^t (\tilde{c}(s) - c(s))\left(\frac{1}{\tilde{d}(s)} - \frac{1}{d(s)}\right) ds, \quad \tilde{\Delta}_4(t) = \frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{i}tY_{j1} - \mathbb{E}[\mathrm{i}tY_{j1}]\right).$$

Then $\tilde{\Delta}(t) = \tilde{\Delta}_1(t) + \tilde{\Delta}_2(t) + \tilde{\Delta}_3(t) - \tilde{\Delta}_4(t)$, and $\tilde{\varphi}_\epsilon(t) - \varphi_\epsilon(t) = \varphi_\epsilon(t) \cdot (e^{\tilde{\Delta}(t)} - 1)$.

Similar to the proof of Lemma 3, it can be shown that

$$\sup_{|t|\le T_n} |\tilde{\Delta}_1(t)| = O_P(n^{-1/2}T_n^{2\beta_\epsilon+1}\log T_n), \quad \sup_{|t|\le T_n} |\tilde{\Delta}_2(t)| = O_P(n^{-1/2}T_n^{3\beta_\epsilon+1}\log T_n), \tag{A.13}$$

$$\sup_{|t|\le T_n} |\tilde{\Delta}_3(t)| = O_P(n^{-1}T_n^{4\beta_\epsilon+1}(\log T_n)^2), \quad \sup_{|t|\le T_n} |\tilde{\Delta}_4(t)| = O_P(n^{-1/2}T_n). \tag{A.14}$$

Since

$$|\tilde{\varphi}_\epsilon(t) - \varphi_\epsilon(t)| \le 2|\varphi_\epsilon(t)||\tilde{\Delta}(t)|\mathbb{I}\{|\tilde{\Delta}(t)| \le 1\} + |\tilde{\varphi}_\epsilon(t) - \varphi_\epsilon(t)|\mathbb{I}\{|\tilde{\Delta}(t)| > 1\},$$

and (A.13) and (A.14) imply that the second term on the right-hand side of the above inequality is equal to 0 with probability approaching 1, it suffices to show that

$$\sup_{|t|\le T_n} |\varphi_\epsilon(t)||\tilde{\Delta}(t)| = O_P(n^{-1/2}T_n^{2\beta_\epsilon+1}\log T_n). \tag{A.15}$$

First, by (A.3), we have

$$
\begin{aligned}
& \sup_{|t|\le T_n} |\varphi_\epsilon(t)||\tilde{\Delta}_1(t)| \\
\le\ & \sup_{|t|\le T_n} |\varphi_\epsilon(t)| \int_0^{|t|} \frac{|\tilde{c}(s)-c(s)|}{|d(s)|} ds \\
\le\ & \sup_{|t|\le \lambda} |\varphi_\epsilon(t)| \int_0^{|t|} \frac{|\tilde{c}(s)-c(s)|}{|d(s)|} ds + \sup_{\lambda<|t|\le T_n} |\varphi_\epsilon(t)| \int_0^{|t|} \frac{|\tilde{c}(s)-c(s)|}{|d(s)|} ds \\
\le\ & O_P(n^{-1/2}) + \sup_{\lambda<|t|\le T_n} |\varphi_\epsilon(t)| \cdot \frac{|t|}{\inf_{|s|\le|t|}|d(s)|} \cdot \sup_{|s|\le|t|} |\tilde{c}(s)-c(s)| \\
\le\ & O_P(n^{-1/2}) + C_\epsilon c_\epsilon^{-2} \sup_{|s|\le T_n} |\tilde{c}(s)-c(s)| \cdot \sup_{\lambda<|t|\le T_n} |t|^{1+\beta_\epsilon} \\
=\ & O_P(n^{-1/2}) + O_P(n^{-1/2} T_n^{\beta_\epsilon+1} \log T_n).
\end{aligned}
$$

Second, using a similar argument,

$$
\begin{aligned}
& \sup_{|t|\le T_n} |\varphi_\epsilon(t)||\tilde{\Delta}_2(t)| \\
\le\ & \sup_{|t|\le T_n} |\varphi_\epsilon(t)| \int_0^{|t|} \left| c(s) \left(\frac{1}{\tilde{d}(s)} - \frac{1}{d(s)} \right) \right| ds \\
\le\ & \sup_{|t|\le \lambda} |\varphi_\epsilon(t)| \int_0^{|t|} |c(s)| \left| \frac{1}{\tilde{d}(s)} - \frac{1}{d(s)} \right| ds + \sup_{\lambda<|t|\le T_n} |\varphi_\epsilon(t)| \int_0^{|t|} \left| c(s) \left(\frac{1}{\tilde{d}(s)} - \frac{1}{d(s)} \right) \right| ds \\
\le\ & O_P(n^{-1/2}) + \sup_{|s|\le T_n} \frac{|d(s)|}{|\tilde{d}(s)|} \cdot \sup_{|s|\le T_n} |\tilde{d}(s)-d(s)| \cdot C_\epsilon \sup_{\lambda<|t|\le T_n} |t|^{1-\beta_\epsilon} \sup_{|s|<|t|} \left| \frac{c(s)}{d^2(s)} \right| \\
=\ & O_P(n^{-1/2}) + O_P(n^{-1/2} T_n^{2\beta_\epsilon+1} \log T_n).
\end{aligned}
$$

Third, it follows from (A.14) that

$$
\sup_{|t|\le T_n} |\varphi_\epsilon(t)||\tilde{\Delta}_3(t)| = o_P(n^{-1/2} T_n^{2\beta_\epsilon+1} \log T_n),
$$

$$
\sup_{|t|\le T_n} |\varphi_\epsilon(t)||\tilde{\Delta}_4(t)| = O_P(n^{-1/2} T_n).
$$

Thus, (A.15) follows from the above results. This completes the proof.

Step 4: Convergence rate for $\tilde{\varphi}_X$

Note that

$$
\begin{aligned}
\tilde{\varphi}_X(t) - \varphi_X(t) =& \frac{\hat{\varphi}_{Y_1}(t)}{\tilde{\varphi}_\epsilon(t)} - \frac{\varphi_{Y_1}(t)}{\varphi_\epsilon(t)} \\
=& \frac{\hat{\varphi}_{Y_1}(t) - \varphi_{Y_1}(t)}{\varphi_\epsilon(t)} + \varphi_{Y_1}(t) \cdot \left(\frac{1}{\tilde{\varphi}_\epsilon(t)} - \frac{1}{\varphi_\epsilon(t)} \right) + (\hat{\varphi}_{Y_1}(t) - \varphi_{Y_1}(t)) \cdot \left(\frac{1}{\tilde{\varphi}_\epsilon(t)} - \frac{1}{\varphi_\epsilon(t)} \right) \\
:=& \tilde{\Delta}_5(t) + \tilde{\Delta}_6(t) + \tilde{\Delta}_7(t).
\end{aligned}
$$

First, by Lemma 1 of KO and Assumption 2,

$$
\sup_{|t|\le T_n} |\tilde{\Delta}_5(t)| \le \frac{\sup_{|t|\le T_n} |\hat{\varphi}_{Y_1}(t) - \varphi_{Y_1}(t)|}{\inf_{|t|\le T_n} |\varphi_\epsilon(t)|} = O_P(n^{-1/2} T_n^{\beta_\epsilon} \log T_n). \tag{A.16}
$$

Second, by the result of Step 3,

$$
\begin{aligned}
\sup_{|t|\le T_n} |\tilde{\Delta}_7(t)| \le& \frac{\sup_{|t|\le T_n} |\hat{\varphi}_{Y_1}(t) - \varphi_{Y_1}(t)| \cdot \sup_{|t|\le T_n} |\tilde{\varphi}_\epsilon(t) - \varphi_\epsilon(t)|}{\inf_{|t|\le T_n} |\varphi_\epsilon(t)| \cdot \inf_{|t|\le T_n} |\tilde{\varphi}_\epsilon(t)|} \\
&= O_P(n^{-1/2} T_n^{2\beta_\epsilon+1} \log T_n) \cdot O_P(n^{-1/2} T_n^{2\beta_\epsilon} \log T_n).
\end{aligned}
$$

Third, similar to the proof of Step 2, we have

$$
\begin{aligned}
\sup_{|t|\le T_n} |\tilde{\Delta}_6(t)| &\le \sup_{|t|\le T_n} \left| \frac{\varphi_X(t)}{\varphi_\epsilon(t)} \right| \cdot \left| \frac{\varphi_\epsilon(t)}{\tilde{\varphi}_\epsilon(t)} \right| \cdot |\tilde{\varphi}_\epsilon(t) - \varphi_\epsilon(t)| \\
&\le O_P(1) \cdot \sup_{|t|\le T_n} |\varphi_X(t)||\tilde{\Delta}(t)| + O_P(1) \cdot \sup_{|t|\le T_n} \left| \frac{\varphi_\epsilon(t)}{\tilde{\varphi}_\epsilon(t)} \right| \cdot \sup_{|t|\le T_n} \left| \frac{\varphi_X(t)}{\varphi_\epsilon(t)} \right| \mathbb{I}\{|\tilde{\Delta}(t)| > 1\}.
\end{aligned}
$$

Since the second term on the right-hand side of the last inequality is equal to 0 with probability approaching 1, to establish the order of $\sup_{|t|\le T_n} |\tilde{\Delta}_6(t)|$, it suffices to show that

$$
\sup_{|t|\le T_n} |\varphi_X(t)||\tilde{\Delta}(t)| = O_P(n^{-1/2} T_n^{(3\beta_\epsilon - \beta_x + 1)\vee 0} \log T_n). \tag{A.17}
$$

Similar to the proof of Step 3, it can be shown that

$$
\sup_{|t|\le T_n} |\varphi_X(t)||\tilde{\Delta}_1(t)| = O_P(n^{-1/2} T_n^{(2\beta_\epsilon - \beta_x + 1)\vee 0} \log T_n), \tag{A.18}
$$

$$
\sup_{|t|\le T_n} |\varphi_X(t)||\tilde{\Delta}_2(t)| = O_P(n^{-1/2} T_n^{(3\beta_\epsilon - \beta_x + 1)\vee 0} \log T_n), \tag{A.19}
$$

$$
\sup_{|t|\le T_n} |\varphi_X(t)||\tilde{\Delta}_3(t)| = o_P(n^{-1/2} T_n^{(3\beta_\epsilon - \beta_x + 1)\vee 0} \log T_n), \tag{A.20}
$$

$$
\sup_{|t|\le T_n} |\varphi_X(t)||\tilde{\Delta}_4(t)| = O_P(n^{-1/2} T_n). \tag{A.21}
$$

Thus, (A.17) follows from (A.18) to (A.21), and we have

$$\sup_{|t|\le T_n} |\tilde{\Delta}_6(t)| = O_P(n^{-1/2} T_n^{(3\beta_\epsilon - \beta_x + 1)\vee 0} \log T_n). \tag{A.22}$$

The desired result then follows from (A.16) and (A.22). □

A.3 Proofs of Other Theorems

The proofs of Theorem 2 to Theorem 4 are similar to the proof of Theorem 1. Therefore, we relegate the proofs of these theorems to the Online Appendix.

References

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

Arcidiacono, P., E. M. Aucejo, H. Fang, and K. I. Spenner (2011). Does affirmative action lead to mismatch? a new test and evidence. *Quantitative Economics 2*(3), 303–333.

Arellano, M. and S. Bonhomme (2012). Identifying distributional characteristics in random coefficients panel data models. *The Review of Economic Studies 79*(3), 987–1020.

Bonhomme, S. and J.-M. Robin (2010). Generalized non-parametric deconvolution with an application to earnings dynamics. *The Review of Economic Studies 77*(2), 491–533.

Chen, X., H. Hong, and D. Nekipelov (2011). Nonlinear models of measurement errors. *Journal of Economic Literature 49*(4), 901–37.

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

Evdokimov, K. (2010). Identification and estimation of a nonparametric panel data model with unobserved heterogeneity. *Working paper*.

Evdokimov, K. and H. White (2012). Some extensions of a lemma of kotlarski. *Econometric Theory 28*(4), 925–932.

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *The Annals of Statistics*, 1257–1272.

Hu, Y. (2017). The econometrics of unobservables: Applications of measurement error models in empirical industrial organization and labor economics. *Journal of Econometrics 200*(2), 154–168.

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

Kotlarski, I. (1967). On characterizing the gamma and the normal distribution. *Pacific Journal of Mathematics 20*(1), 69–76.

Krasnokutskaya, E. (2011). Identification and estimation of auction models with unobserved heterogeneity. *The Review of Economic Studies 78*(1), 293–327.

Kurisu, D. and T. Otsu (2022). On the uniform convergence of deconvolution estimators from repeated measurements. *Econometric Theory 38*(1), 172–193.

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

Li, T. and C. Hsiao (2004). Robust estimation of generalized linear models with measurement errors. *Journal of Econometrics 118*(1-2), 51–65.

Li, T., I. Perrigne, and Q. Vuong (2000). Conditionally independent private information in ocs wildcat auctions. *Journal of Econometrics 98*(1), 129–161.

Li, T. and Q. Vuong (1998). Nonparametric estimation of the measurement error model using multiple indicators. *Journal of Multivariate Analysis 65*(2), 139–165.

Newey, W. K. (1994). The asymptotic variance of semiparametric estimators. *Econometrica: Journal of the Econometric Society*, 1349–1382.

Otsu, T. and L. Taylor (2021). Specification testing for errors-in-variables models. *Econometric Theory 37*(4), 747–768.