Working Paper No. 15-05

Nonparametric Conditional Quantile Estimation for Profit Frontier Analysis

Shan Zhou
University of Colorado Boulder

October 2015
Revised November 2015

Department of Economics

University of Colorado Boulder
Boulder, Colorado 80309

Nonparametric Estimation of Profit Frontier of order α

Shan Zhou[†]

University of Colorado, Boulder

November 8, 2015

Abstract. In this paper, we introduce the concept of a profit frontier of continuous order $\alpha \in [0,1]$ and provide an easy to implement nonparametric estimator for such profit frontiers. From a statistical perspective the estimator we propose is, in essence, the estimator for a conditional quantile with a suitably defined conditioning set. Inspired by Aragon et al. (2005) in a production function setting, instead of studying a traditional profit frontier, whose estimation might be very sensitive to outliers and extreme values, we define a class of profit functions of order α based on conditional quantiles of an appropriate distribution of profit, input and output prices. We show these quantiles

of nonparametric stochastic frontier models include, among others, Kumbhakar et al. (2007) and Martins-Filho and Yao (2013). A critical drawback of stochastic frontier models is that they generally require strong distributional assumptions regarding the inefficiency and noise terms. In addition, by assumption the stochastic frontier model error terms have a non-zero conditional expectation, and the average production relation is maintained for all firms. However, it is highly possible that the relationship might vary at different efficiency levels.

On the other hand, nonstochastic frontier models assume that all observations lie inside the frontier of order

what we call the profit function throughout the paper. Given the existence of inefficiency, our objective is to estimate profit functions and assess firms' efficiency levels. There are a number of differences in estimating a profit function compared to estimating a production function. First, the derivation of the profit function relies on many assumptions on market structure and it is difficult or producing unit i. We denote the support of f by and focus on the set $= f(;p;w)\ 2 \ :$ $P(P \ p;W \ w) > 0g$. Given $C_{p;w} = fP \ p;W \ wg$ we let

$$F(\ jC_{p;w}) = P(\quad jP \ p;W \ w) = \frac{P(\quad ;P \ p;W \ w)}{P(P \ p;W \ w)}: \tag{1}$$

and give the following probabilistic de nition of a pro t function

$$(p;w) := \inf f \ 2\ [0;B \ ] : F(\ jC_{p;w}) = 1g: \tag{2}$$

As de ned, the pro t function $(p;w$

That is, the quantile curves $f($

The concept of profit functions of order can be easily extended to settings where additional constraints on profit and technology are appropriate. we give two examples. First, firms can face different production capacities and by consequence different profit functions. If a firm has small production capacity, the value of profit would be small compared to a representative firm, even if where represent the elasticity of market demand and $F(\; jC_{;w}) = P(\quad j \quad ; W \quad w):$ The and Martins-Filho (2010). The kernels M_k are defined as

$$M_k(x) = \frac{1}{c_{k;0}} \sum_{|s|=1}^{k} \frac{c_{k;s}}{|s|} K\left(\frac{x}{s}\right) \tag{7}$$

where $c_{k;s} = (-1)^{s+k} C_{2k}^{s+k}$, C_{2k}^{s+k} are the binomial coefficients and $K(\cdot)$ is a traditional (seed) kernel where

$$\hat{f}(\ |C_{p,w}) = \frac{\partial \hat{F}(\ |C_{p,w})}{\partial} = \begin{cases} 0; & \text{if } \ = 0 \\ \frac{(nh_n)^{-1}\sum_{i=1}^{n} M_k(\frac{\ \ - \ _i}{h_n}) I(P_i \le p, W_i \le w)}{n^{-1}\sum_{i=1}^{n} I(P_i \le p, W_i \le w)}; & \text{if } \ > 0 \end{cases}$$

and $_{;n}(p;w) = \ _{;n}(p;w) + (1 \ \ ) \ (p;w)$ for some $\ 2\ (0;1)$. In the following section we provide some asymptotic characterizations for our estimator, including consistency and asymptotic normality.

3 Asymptotic characterization of $_{;n}$

In this section we provide theorems establishing asymptotic properties of our estimators. All proofs of the theorems and required lemmas can be found in Appendix. We begin by listing and discussing assumptions that are sufficient to establish our main theorems.

3.1 Assumptions

Assumption 1. *$f(\ _i;P_i;W_i)g_{i=1}^{n}$ is a sequence of independent random vectors taking values in a compact set $\ = [0;B\ ]\ \ S_{PW}$ where S_{PW} is a compact set in $R_{+}^{d_1}\ \ R_{++}^{d_2}$. For any i, $(\ _i;P_i;W_i)$ have the same joint distribution F and joint density function f as the vector $(\ ;P;W)$, f is defined on $R\ \ R^{d_1}\ \ R^{d_2}$ with support $\ $.*

Assumption 2. *(i) The seed kernel $K(\ )$ is a bounded symmetric density with compact support $[\ B_K;B_K]$ and $\int_{B_K}^{B_K}\ K(\ )d\ = 0$. (ii) $\int_{B_K}^{B_K}\ ^2K(\ )d\ = \ _K^2$. (iii) For any $\ ;\ ^0 2\ [\ B_K;B_K]$, we have $jK(\ )\ \ K(\ ^0)j\ \ m_K j\ \ \ ^0 j$ for some $0 < m_K < 1$. (iv) For all $\ ;\ ^0 2$ R, we have $j\ (\ )\ \ (\ ^0)j\ \ m\ j\ \ \ ^0 j$ for some $0 < m\ < 1$, where $\ (\ ) = \int_{B_K}\ K(\ )d\ $. (v) For fixed k, $\int jK(t)jt^{2k}dt < 1$.*

The first assumption is standard in the deterministic frontier literature. Assumption 2 is the same as Martins-Filho and Yao (2008) except (v). We need Assumption 2 (v) for restricting the order of bias(See the similar assumption in Mynbaev and Martins-Filho (2010)). Note that (7) implies that for any $k \in \mathbb{N}$, the above assumptions also hold for kernel M_k. That is, (i) $M_k(\cdot)$ is a symmetric bounded kernel function with compact support $[-B_M; B_M]$. $\int_{-B_M}^{B_M} u M_k(u)du = 0$; (ii) $\int_{-B_M}^{B_M} u^2 M_k(u)du := \sigma_M^2 = 2\sigma_K^2 \sum_{s=1}^{k} \lambda_{k;s} s^2$; (iii) For any $u; u' \in [-B_M; B_M]$, we have $|M_k(u) - M_k(u')| \leq m_M |u - u'|$ for some $0 < m_M < \infty$; (iv) For any $u; u' \in \mathbb{R}$, we have $|\mu_M(u) - \mu_M(u')| \leq m_\mu |u - u'|$ for some $0 < R$

Assumption 5A imposes an order $2k$ Lipschitz condition on $F_f(\ ; p; w)$ with respect to . From the proof of Theorem 1 in Mynbaev and Martins-Filho (2010) we know that boundedness of $F_f^{(2k)}(\ ; p; w)$ implies a Lipshitz condition of order $2k$. As a result, Assumption 5B is a more strict condition than 5A in the special case $k = 1$. Given Assumption 5A, we can restrict the order of the bias for our estimator to h^{2k}. Given Assumption 5B, we can obtain a speci c structure for the asymptotic bias and variance by using a Taylor expansion.

3.2 Asymptotic Properties

We start by showing that $\hat{F}(\ | C_{p;w})$ is asymptotically a proper distribution function for kernels that satisfy Assumption 2.

Proposition 3. *Under Assumption 2, we have: (i) $\hat{F}(\ | C_{p;w})$ is nondecreasing in ; (ii) $\hat{F}(\ | C_{p;w})$ is right continuous; (iii) $\lim_{\ \to 0} \hat{F}(\ | C_{p;w}) = 0$; (iv) For any $(p; w)$, there exists some $N(p; w)$ such that for all $n > N(p; w)$, we have $\lim_{\ \to \infty} \hat{F}(\ | C_{p;w}) = 1$.*

The next theorem establishes consistency of $\ _{;n}$.

Theorem 1. *Let h_n be a nonstochastic sequence of bandwidths such that $0 < h_n \to 0$ as $n \to \infty$. Given $w \in \mathbb{R}^{d_2}_{++}$, $p \in \mathbb{R}^{d_1}_{+}$, suppose there exist $N(p; w)$ such that when $n > N(p; w)$ we have $\min_{\{i: P_i \ \ p; W_i \ \ w\}} \ _i \ \ h_n B_M$. Under Assumption 1-4 along with Assumption 5A (or 5B), if $H_{2k}(\ ; p; w)$, $F_f(\ ; p; w)$ and $\varepsilon_{2k}(\ ; p; w)$ are bounded for all $(\ ; p; w) \in$, we have*

$$\ _{;n}(p; w) \quad \ (p; w) = o_p(1). \tag{9}$$

The next theorem shows that under suitable normalization and centering $\ _{;n}(p; w)$ is asymptotically distributed as as standard normal.

Theorem 2. *Let h_n be a nonstochastic sequence of bandwidths such that $nh_n^2 \to \infty$ and $nh_n^4 = O(1)$ as $n \to \infty$. Given $w \in \mathbb{R}^{d_2}_{++}$, $p \in \mathbb{R}^{d_1}_{+}$, suppose there exist $N(p; w)$ such that when $n > N(p; w)$ we have $\min_{\{i: P_i \ \ p; W_i \ \ w\}} \ _i \ \ h_n B_M$. Then,*

density function is, the faster the bias term would vanish.

4 Monte Carlo Study

4.1 Setup and Implementation

In this section, we design and conduct a small Monte-Carlo simulation to implement our estimator and investigate some of its finite sample properties. We also compare the performance of our smooth estimator and an similar estimator based on the empirical estimation. The data generating process is given by

$$
\begin{aligned}
\pi_i &= \pi(P_i, W_i)R_i \quad i = 1, \ldots, n \\
R_i &= \exp(-Z_i), \quad Z_i \sim Exp(\theta)
\end{aligned}
$$

where π_i represents profit, P_i and W_i represent output and input prices. In this simulation, we assume both output and input price are scalars. Prices are uniformly drawn from a meshgrid $[p_l, p_u] \times [w_l, w_u] = [1, 3] \times [1, 3]$. $R_i = \exp(-Z_i)$ represents efficiency score for each unit i. Z_i are independently generated from an exponential distribution with parameter $\theta = 1/3$. As a result the density function of R_i is $f(r) = 3r^2$ with support $(0, 1]$ and a mean 0.75. $\pi(p, w)$ is the profit function. In this simulation we consider the functional form $\pi(p, w) = p^{6/5} w^{-6/5}$. One can easily verify this function satisfies all properties of a profit function: a) nondecreasing in p and nonincreasing in w; b) convex in both p and w; c) homogenous of degree one, and d) continuous. Several experimental designs are considered: We estimate profit frontiers of order $\alpha = 0.25, 0.5, 0.75$ and 0.99 using M_k kernel functions with $k = 1, 2$ as well as an empirical distribution. In each experiment, We consider two sample sizes $n = 200$ and $n = 400$ and perform 2000 iterations to obtain the averaged absolute value of bias and root mean squared error of each estimator.

The empirical profit frontier of order α is estimated as follows: Let $N_{p,w} = \sum_{i=1}^{n} I(P_i \le p, W_i \ge w)$. For $j = 1, \ldots, N_{p,w}$, get the order statistic of the observation $\pi_{(i_j)}$ such that $\pi_{(i_1)} \le \pi_{(i_2)} \le \ldots$

$(i_{N_{p;w}})$. The empirical conditional distribution $\hat{F}_e(\ jC_{p;w})$ is

$$\hat{F}_e(\ jC_{p;w}) = \frac{\sum_{j=1}^{N_{p;w}} I(\ _{(i_j)} \qquad)}{}$$

in Mynbaev and Martins-Filho (2010), we can estimate I_1, I_2 and f a suitably defined Rosenblatt density estimator. The optimal bandwidths for the estimators with higher k are yet to be obtained. We use the same bandwidth as $k = 1$.

4.2 Results and Analysis

Table 1 gives the bias and root mean square error of our smoothed estimator with order of kernel $k = 1$ and $k = 2$ compared with the empirical estimator evaluated at prices $p = 2$ and $w = 2$.

Table 1: Bias and RMSE under Each Experiment Design

	$\lvert$Bias$\rvert$			RMSE		
n=200	Kernel k=1	Kernel k=2	Empirical	Kernel k=1	Kernel k=2	Empirical
0.25	.018	.019	.021	.024	.024	.027
0.50	.020	.021	.024	.033	.033	.037
0.75	.027	.027	.030	.031	.032	.037
0.99	.132	.261	.084	.175	.358	.095
n=400	Kernel k=1	Kernel k=2	Empirical	Kernel k=1	Kernel k=2	Empirical
0.25	.014	.013	.015	.017	.016	.019
0.50	.015	.012	.017	.018	.016	.019
0.75	.019	.016	.021	.023	.021	.028
0.99	.083	.098	.057	.102	.121	.068

The simulations seem to confirm our asymptotic results. In particular, the root mean squared error of all estimators decreases with the sample size, confirming our asymptotic results. Our smoothed kernel estimator outperforms the empirical estimator in the cases with $= 0.25; 0.5$ and 0.75. Although we do not use the optimal bandwidth, the performance of the estimator with kernel order $k = 2$ is quite good. When the sample size is 200, the performance of estimators with $k = 1$ and $k = 2$ are very close. When the sample size grows from 200 to 400 we observe a larger improvement for the estimator with $k = 2$. For example, with $= 0.5$, the bias of the estimator with $k = 2$ decreases from .021 to .012, while the bias of the estimator with $k = 1$ just decreases from .020 to .015. We find the similar results for all . This is consistent with the result in Theorem 2 which states the bias decays faster as k increases.

We also observe that as increases, all estimators show larger bias and mean square error. This can be interpreted as resulting from the fact that there are less effective data available as grows. As a result, when is close to 1, profit functions of order become more difficult to estimate. Note that the performance of our smoothed estimator is especially poor when $= 0.99$. This is most likely due to the fact that our distribution function has compact support, and it is not smooth near the boundary. Therefore, the smoothed estimator can generate large biases.

In summary, our simulation results indicate the proposed smooth estimator for the profit function of order can outperform the empirical estimator in most cases as long as is not very close to 1. Additionally, increases in the order k of the M_k kernel may increase the convergence speed of the bias. However, we do not suggest to use our method in approximating the full frontier where is approaching to 1. Note that the full frontier is not required in estimating the efficiency in our method. According to the analysis in section 2, any frontier with $\in (0, 1)$ can be served as a standard in the efficiency analysis.

5 Conclusion and Discussion

In this paper we consider the construction and estimation of a profit function of continuous order $\in [0, 1]$. We define a class of such profit functions based on conditional quantiles of an appropriate distribution of profit, input and output prices. We show that they are useful in measuring and assessing profit efficiency. We show that our estimator is consistent and asymptotically normal with a parametric convergence speed of $\sqrt{n}$. Furthermore, the bias of our estimator decays to zero faster than the traditional kernel estimators. A Monte-Carlo simulation is performed to implement our estimator; investigate its finite sample performance and compare it to the empirical estimator. Simulation results seem to confirm the asymptotic results we have obtained and also seems to indicate that our proposed estimator can outperform its competitors in most cases. However, our estimator seems to possess large boundary bias. Decreasing the boundary bias would be a desirable direction for future work. The choice of optimal bandwidth when $k > 1$ is another issue to address. It is also desirable to study the decomposition of technique efficiency and allocative efficiency.

Appendix - Proofs and auxiliary lemmas

Proposition 1 *Proof.* For any $(\ ; p; w) \in$, if $<$ $(p; w) = \inf\{ \ \in [0; B\ ] : F(\ |C_{p;w}) \ \ \}$, then $\notin \{ \ \in [0; B\ ] : F(\ |C_{p;w}) \ \ \}$. That is, $F(\ |C_{p;w}) <$. If $>$ $(p; w)$, there exist some $\varepsilon > 0$ such that $> \ (p; w) + \varepsilon$. By the definition of $(p; w)$, for any $\varepsilon > 0$, there exist some $_0 \in \{ \ \in [0; B\ ] : F(\ |C_{p;w}) \ \ \}$ such that $_0 < \ (p; w) + \varepsilon$. By the strict monotonicity of $F(|C_{p;w})$, $F(\ |C_{p;w}) > F(\ (p; w) + \varepsilon|C_{p;w}) > F(\ _0|C_{p;w})$. The result then follows. □

Proposition 2 *Proof.* (i) Since $\{ \ (p; w)\}_{0 \ \ 1}$ is monotone nondecreasing in , and $\sup_{0 \ \ 1}\{ \ (p; w)\} =$ $(p; w)$.The result then follows. (ii) Let be a compact set interior to the support of $(P; W)$. Define $_n(p; w) = \ _{1 \ \frac{1}{n}}(p; w)$. Since $\{ \ (p; w)\}_{0 \ \ 1}$ is monotone nondecreasing in , for any $n \in \mathbb{N}$, $_n(p; w) \ \ _{n+1}(p; w)$. From (i), $\lim_{n \to \infty} \ _n(p; w) = \ (p; w)$ pointwise. By Dini's Theorem, $\sup_{(p;w) \in} | \ _n(p; w) \ \ (p; w)| \to 0$. Thus, for any $\varepsilon > 0$, there exist some N such that when $n > N$, $\sup_{(p;w) \in} | \ _n(p; w) \ \ (p; w)| < \varepsilon$. That is, there exist $= 1 \ \frac{1}{N}$ such that when $| \ \ 1| <$, $\sup_{(p;w) \in} | \ (p; w) \ \ (p; w)| < \varepsilon$. □

Proposition 3 *Proof.* (i) First, note that by definition when $= 0$ we have $\hat{F}(\ |C_{p;w}) = 0$. If $0 < \ _1 \ \ _2$, we only need to prove $\hat{P}(\ _2; p; w) \ \ \hat{P}(\ _1; p; w) \ \ 0$, since the denominator does not depend on . By (5),

$$\hat{P}(\ _2; p; w) \ \ \hat{P}(\ _1; p; w) = (nh_n)^{\ 1} \sum_{i=1} \Big(\int_0^{\ _2} M_k(\frac{ \ _i}{h_n}) d \ \ \int_0^{\ _1} M_k(\frac{ \ _i}{h_n}) d \ \Big) I(P_i \ \ p; W_i \ \ w)$$
$$0$$

since M_k is a symmetric density. (ii) For any $_0 \in [0; B\ ]$, let $| \ \ _0| <$ for some > 0. Then,

$$\begin{aligned} |\hat{P}(\ ; p; w) \ \ \hat{P}(\ _0; p; w)| &= (nh_n)^{\ 1} \sum_{i=1} \Big(\int_0 M_k(\frac{ \ _i}{h_n}) d \ \ \int_0^{\ _0} M_k(\frac{ \ _i}{h_n}) d \ \Big) I(P_i \ \ p; W_i \ \ w) \\ &\ (nh_n)^{\ 1} \sum_{i=1} \Big(\int_0^{\ _0 +} M_k(\frac{ \ _i}{h_n}) d \ \ \int_0^{\ _0} M_k(\frac{ \ _i}{h_n}) d \ \Big) I(P_i \ \ p; W_i \ \ w) \\ &= (nh_n)^{\ 1} \sum_{i=1} \Big(\int_{\ _0}^{\ _0 +} M_k(\frac{ \ _i}{h_n}) d \ \Big) I(P_i \ \ p; W_i \ \ w) \\ &\ h_n^{\ 1} \sup_{\varphi \in [\ B_M; B_M]} M_k(\varphi) < \varepsilon \end{aligned}$$

where the last inequality follows for any > 0, since can be made as small as desired. (iii) follows directly from (i) and (ii). For (iv) we need only prove that for any $(p; w)$, there exists some $N(p; w)$ such that for all $n > N(p; w)$, $h_n^{-1} \lim_{\to \infty} \int_0 M_k(\frac{i}{h_n}) d$

Proof. (a) Since $h_n \to 0$ as $n \to \infty$, there exist $N(p;w) \in \mathbb{R}_+$ such that for all $n > N(p;w)$,

$$
\begin{aligned}
E(\hat{P}(\theta;p;w)) &= E\Big[(nh_n)^{-1}\sum_{i=1}^{n}\Big(\int_0^{\theta} M_k\Big(\frac{\phi-\theta_i}{h_n}\Big)d\phi\Big) I(P_i \le p; W_i \le w)\Big] \\
&= h_n^{-1}\int_{\mathbb{R}^{d_1}}\int_{\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}\int_0^{\theta} M_k\Big(\frac{\phi-\psi}{h_n}\Big)d\phi\; I(P \le p; W \le w) f(\psi;P;W)d\psi\; d(P;W) \\
&= \int_{\mathbb{R}^{d_1}}\int_{\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}\int_{-B_M}^{\frac{\theta-\psi}{h_n}} M_k(\varphi)d\varphi\; I(P \le p; W \le w) f(\psi;P;W)d\psi\; d(P;W) \\
&= \int_{\mathbb{R}^{d_1}}\int_{\mathbb{R}^{d_2}}\int_{-\infty}^{\infty} \Phi_M\Big(\frac{\theta-\psi}{h_n}\Big) I(P \le p; W \le w) f(\psi;P;W)d\psi\; d(P;W)
\end{aligned}
$$

Let $F_f(\theta;p;w) = \int_{-\infty}^{\theta} f(\psi;p;w)d\psi$. Using integration by parts,

$$
\begin{aligned}
&\int_{-\infty}^{\infty} \Phi_M\Big(\frac{\theta-\psi}{h_n}\Big) I(P \le p; W \le w) f(\psi;P;W)d\psi \\
&= h_n \int_{-\infty}^{\infty} \Phi_M(\varphi) I(P \le p; W \le w) f(\theta - h_n\varphi;P;W)d\varphi \\
&= -\int_{-\infty}^{\infty} \Phi_M(\varphi) I(P \le p; W \le w) dF_f(\theta - h_n\varphi;P;W) \\
&= -\Phi_M(\varphi) I(P \le p; W \le w) F_f(\theta - h_n\varphi;P;W)\Big|_{\varphi=-\infty}^{\varphi=\infty} + \int_{-\infty}^{\infty} F_f(\theta - h_n\varphi;P;W) I(P \le p; W \le w) d\Phi_M(\varphi) \\
&= 0 + \int_{-\infty}^{\infty} M_k(\varphi) F_f(\theta - h_n\varphi;P;W) I(P \le p; W \le w) d\varphi \\
&= \frac{1}{c_{k;0}}\int_{-\infty}^{\infty}\sum_{|s|=1}^{k}\frac{c_{k;s}}{|s|} K(\varphi/s) F_f(\theta - h_n\varphi;P;W) I(P \le p; W \le w) d\varphi \\
&= \frac{1}{c_{k;0}}\int_{-\infty}^{\infty} K(t)\sum_{|s|=1}^{k} c_{k;s} F_f(\theta - sh_n t;P;W) I(P \le p; W \le w) dt
\end{aligned}
$$

Since $\sum_{|s|=1}^{k} \frac{c_{k;s}}{c_{k;0}} = 1$, $\int_{1}^{1} K(t)dt = 1$, we have

$$
\begin{aligned}
& E(\hat{P}(\; ;p;w)) \quad P(\; ;p;w) \\
= & \frac{1}{c_{k;0}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{1}^{1} K(t) \sum_{|s|=1}^{k} c_{k;s} F_f(\quad sh_n t; P; W) I(P \quad p; W \quad w) dt d(P;W) \\
& \int_{D_{p;w}} \int_{1} f(\; ;P;W) d \;\; d(P;W) \\
= & \frac{1}{c_{k;0}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{1}^{1} K(t) \sum_{|s|=1}^{k} c_{k;s} F_f(\quad sh_n t; P; W) I(P \quad p; W \quad w) dt d(P;W) \\
+ & \sum_{|s|=1}^{k} \frac{c_{k;s}}{c_{k;0}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{1}^{1} K(t) dt F_f(\; ;P;W) I(P \quad p;W \quad w) d(P;W) \\
= & \frac{1}{c_{k;0}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{1}^{1} K(t) \;\; {}^{2k}_{h_n t} F_f(\; ;P;W) I(P \quad p;W \quad w) dt d(P;W)
\end{aligned}
$$

By Assumption 5A, we have

$$
\begin{aligned}
|E(\hat{P}(\; ;p;w)) \quad P(\; ;p;w)| \quad & c \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{1}^{1} |K(t) \;\; {}^{2k}_{h_n t} F_f(\; ;P;W)| I(P \quad p;W \quad w) dt d(P;W) \\
& c \int_{D_{p;w}} (\int_{|h_n t| \quad "_{2k}(\; ;P;W)} + \int_{|h_n t| > "_{2k}(\; ;P;W)}) |K(t) \;\; {}^{2k}_{h_n t} F_f(\; ;P;W)| dt d(P;W) \\
& c[\int_{D_{p;w}} \int_{|h_n t| \quad "_{2k}(\; ;P;W)} |K(t)| (h_n t)^{2k} dt H_{2k}(\; ;P;W) d(P;W) \\
& + \int_{D_{p;w}} \sup_{2\mathbb{R}} |F_f(\; ;P;W)| \int_{|h_n t| > "_{2k}(\; ;P;W)} |K(t)| dt d(P;W)]
\end{aligned}
$$

Since for any $N > 0$,

$$
\int_{|t|>N} |K(t)| dt \quad \int_{|t|>N} |K(t)| |\frac{t}{N}|^{2k} dt \quad N^{\;\;2k} \int_{1}^{1} |K(t)| t^{2k} dt
$$

Assume that $\int_{1}^{1} |K(t)| t^{2k} dt < 1$, we have

$$
\begin{aligned}
& |E(\hat{P}(\; ;p;w)) \quad P(\; ;p;w)| \\
& c h_n^{2k} [\int_{D_{p;w}} H_{2k}(\; ;P;W) d(P;W) + \int_{D_{p;w}} \sup_{2\mathbb{R}} |F_f(\; ;P;W)| "^{\;2k}_{2k}] [\qquad) I(P;W) \\
& = \quad H_{E} \;\; 2
\end{aligned}
$$

(b) Note that $V(\hat{P}(\ ;p;w)) = \frac{1}{n}(V_{1n} \quad V_{2n})$, where

$$V_{1n} = E[h_n^{\ 2}(\int_0 M_k(\frac{}{h_n})d\ )^2 I(P_i \quad p; W_i \quad w)]$$

$$V_{2n} = (E[h_n^{1}\int_0 M_k(\frac{}{h_n})d\ I(P_i \quad p; W_i \quad w)])^2$$

From part (a), we know the limiting behavior of V_{2n}. Now, for V_{1n} since $h_n\ !\ 0$ as $n\ !\ 1$, there exist $N(p;w)\ 2\ \mathbb{R}_+$ such that for all $n > N(p;w)$,

$$\begin{aligned}
V_{1n} &= E[h_n^{\ 2}(\int_0 M_k(\frac{}{h_n})d\ )^2 I(P_i \quad p; W_i \quad w)] \\
&= h_n^{\ 2}\int_{\mathbb{R}^{d_1}}\int_{\mathbb{R}^{d_2}}\int_1^1 (\int_0 M_k(\frac{}{h_n})d\ )^2 f(\ ;P;W) I(P \quad p; W \quad w)d\ \ d(P;W) \\
&= \int_{\mathbb{R}^{d_1}}\int_{\mathbb{R}^{d_2}}\int_1^1 (\int_{B_M}^{\frac{}{h_n}} M_k(')d')^2 f(\ ;P;W) I(P \quad p; W \quad w)d\ \ d(P;W) \\
&= \int_{\mathbb{R}^{d_1}}\int_{\mathbb{R}^{d_2}}\int_1^1 (\ _M(\frac{}{h_n}))^2 f(\ ;P;W) I(P \quad p; W \quad w)d\ \ d(P;W):
\end{aligned}$$

Integrating by parts

$$\int_1^1 (\ _M(\frac{}{h_n}))^2 f(\ ;P;W) I(P \quad p; W \quad w)d$$

$$= h_n \int_1^1 (\ _M('))^2 f(\quad h_n \quad P \quad p; W \quad w)d)) \quad \int_1^1 (\ _M(\ '))^2 f\ (P \quad p; W \quad w)dk(\quad h_n$$

Note that

$$_M(st) \;=\; \int_{B_M}^{st} M_k(v)dv = \sum_{|s|=1}^{k} \frac{c_{k;s}}{c_{k;0}} \int_{B_M}^{st} \frac{1}{|s|} K(\frac{v}{s})dv \;=\; \sum_{|s|=1}^{k} \frac{c_{k;s}}{c_{k;0}} \int_{B_K}^{t} K(u)du = \;(t)$$

Thus,

$$V_{1n} = \frac{2}{c_{k;0}} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{\infty}^{\infty} K(t)\;(t) \sum_{|s|=1}^{k} c_{k;s} F_f(\quad sh_n t; P; W) I(P \quad p; W \quad w)dtd(P; W)$$

Again, integrating by parts,

$$\int_{\infty}^{\infty} K(t)\;(t)dt = 1{=}2;$$

since $0 \quad (t) \quad 1$, $\int_{\infty}^{\infty} |K(t)\;(t)|t^{2k}dt < \infty$. Similarto the proof in part (a) with $K(t)\;(t)$ instead of $K(t)$,

$$V_{1n} = P(\;;p;w) + R_{1k}(\;;p;w;h)$$

where $|R_{1k}(\;;p;w;h)| \quad ch_n^{2k}[\int_{D_{p;w}} H_{2k}(\;;P;W)d(P;W) + \int_{D_{p;w}} \sup_{\;2\mathbb{R}} |F_f(\;;P;W)|\text{"}_{2k}^{2k}(\;;P;W)d(P;W)]$.
From part (a),

$$\begin{aligned} V_{2n} &= [E(\hat{P}(\;;p;w))]^2 \\ &= [P(\;;p;w) + R_{2k}(\;;p;w;h)]^2 \end{aligned}$$

where $|R_{2k}(\;;p;w;h)| \quad ch_n^{2k}[\int_{D_{p;w}} H_{2k}(\;;P;W)d(P;W) + \int_{D_{p;w}} \sup_{\;2\mathbb{R}} |F_f(\;;P;W)|\text{"}_{2k}^{2k}(\;;P;W)d(P;W)]$.
As a result,

$$\begin{aligned} V(\hat{P}(\;;p;w)) &= \frac{1}{n}(V_{1n} \quad V_{2n}) \\ &= \frac{1}{n} P(\;;p;w)(1 \quad P(\;;p;w)) \quad \frac{2}{n} P(\;;p;w)R_{2k}(\;;p;w;h) + \frac{1}{n} R_{1k}(\;;p;w;h) \\ &\quad \frac{1}{n} R_{2k}^2(\;;p;w;h) \end{aligned}$$

where $|R_{1k}(\;;p;w;h)|$ and $|R_{2k}(\;;p;w;h)|$ are lead than or equal to

$$ch_n^{2k}[\int_{D_{p;w}} H_{2k}(\;;P;W)d(P;W) + \int_{D_{p;w}} \sup_{2\mathbb{R}} |F_f(\;;P;W)|\text{"}_{2k}^{2k}(\;;P;W)d(P;W)]:$$

□

Lemma 1 gives the order of the bias and variance as functions of k. Thus as we increase k, the speed of decay of bias and variance increases. If we assume f has bounded first order derivative with respect to , by applying Taylor's Theorem, the next lemma provides a more explicit structure for bias and variance when $k = 1$.

Lemma 2. *For $k = 1$, under Assumption 1-4 and Assumption 5B, we have: (a)*

$$E(\hat{P}(;p;w)) = \begin{cases} P(;p;w) + \frac{1}{2}h_n^2 \ {}_M^2 \int_{1((; (p;w)])} f^{(1)}(;P;W)d(P;W) + o(h_n^2) & \text{if } 0 < \ < \ (p;w) \\ P(;p;w) + o(h_n^2) & \text{if } \quad (p;w) \end{cases}$$

(b)

$$V(\hat{P}(;p;w)) = \begin{cases} n^{-1}P(;p;w)(1 - P(;p;w)) - 2n^{-1}h_n \int_{1((; (p;w)])} f(;P;W)d(P;W) + o(h_n/n) & \text{if } 0 < \ < \ (p;w) \\ n^{-1}P(;p;w)(1 - P(;p;w)) + o(h_n/n) & \text{if } \quad (p;w) \end{cases}$$

where $P(;p;w) = P(\quad ;P \quad p;W \quad w)$ and $\hat{P}(;p;w)$ is defined in (5). $\quad = \int_{B_M}^{B_M} M_k( ) \ _M( )d$.

Proof. (a) Since $h_n \to 0$ as $n \to \infty$, there exist $N(p;w) \in \mathbb{R}_+$ such that for all $n > N(p;w)$,

$$\begin{aligned} E(\hat{P}(;p;w)) &= E[(nh_n)^{-1} \sum_{i=1}^{n} (\int_{-\infty}^{0} M_k(\frac{\ \ -\ _i}{h_n})d \ )I(P_i \quad p;W_i \quad w)] \\ &= h_n^{-1} \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} \int_{-\infty}^{0} \int M_k(\frac{\ \ -\ }{h_n})d \ \ I(P \quad p;W \quad w)f(;P;W)d \ \ d(P;W) \\ &= h_n^{-1} \int_{D_{p;w}} \int_{[0; (P;W)]} \int_{0} M_k(\frac{\ \ -\ }{h_n})d \ \ f(;P;W)d \ \ d(P;W) \\ &= \int_{D_{p;w}} \int_{[0; (P;W)]} \int_{B_M}^{\frac{\ -\ i}{h_n}} M_k(')d' f(;P;W)d \ \ d(P;W) \\ &= \int_{D_{p;w}} \int_{[0; (P;W)]} \ _M(\frac{\ \ -\ }{h_n}) f(;P;W)d \ \ d(P;W). \end{aligned}$$

We consider 3 cases: (1) $0 < \ < \ (p;w)$; (2) $\ > \ (p;w)$; (3) $\ = \ (p;w)$.

For case (1),

$$
\begin{aligned}
E(\hat{P}(\ ;p;w)) &= \int_{D_{p;w}} \int_{[0;\ (P;W)]} {}_M(\frac{}{h_n}) \text{[redacted]} \\
&= \int_{{}^{1}([0;\ (p;w)])} \int_{[0;\ (P;W)]} {}_M \text{[redacted]} \\
&= \int_{{}^{1}([0;\ )[(\ ;\ (p;w)])} \int_{[0;\ (P;W} \text{[redacted]} \\
&\quad + \int_{{}^{1}(f\ \ g)} \int_{[0;\ ]} {}_M(\frac{}{h_n} \text{[redacted]} \\
&= A_{1n} + A_{2n}.
\end{aligned}
$$

Note that for the last term, for $\int < \ ,\ {}_M(\frac{}{h_n})\ !$ [redacted]

$j\ {}_M(\frac{}{h_n}) f(\ ;P;W) j < 1$ $\int_{g)\ [0}$ [redacted]

By Taylor's theorem, $F_f(\quad h_n \; ;P;W) = F_f(\; ;P;W) \quad h_n \; f(\; ;P;W) + \frac{1}{2}h_n^2 \; {}^2 f^{(1)}(\; ;P;W) + o(h_n^2)$, Hence

$$A_{1n} = E_{1n} + E_{2n} + E_{3n} + E_{4n} + o(h_n^2)$$

where

$$
\begin{aligned}
E_{1n} &= \int_{{}^{1}([0; \;) \cup (\; ; \; (p;w)])} {}_M(\frac{(P;W)}{h_n}) F_f(\; (P;W);P;W) d(P;W) \\
E_{2n} &= \int_{{}^{1}([0; \;) \cup (\; ; \; (p;w)])} F_f(\; ;P;W) \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\;) d \; d(P;W) \\
E_{3n} &= h_n \int_{{}^{1}([0; \;) \cup (\; ; \; (p;w)])} f(\; ;P;W) \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\;) \; d \; d(P;W) \\
E_{4n} &= \frac{1}{2}h_n^2 \int_{{}^{1}([0; \;) \cup (\; ; \; (p;w)])} f^{(1)}(\; ;P;W) \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\;) \; {}^2 d \; d(P;W)
\end{aligned}
$$

Now,

$$
\begin{aligned}
E_{1n} &= \int_{{}^{1}([0; \;) \cup (\; ; \; (p;w)])} {}_M(\frac{(P;W)}{h_n}) F_f(\; (P;W);P;W) d(P;W) \\
&= \int_{{}^{1}([0; \;))} {}_M(\frac{(P;W)}{h_n}) F_f(\; (P;W);P;W) d(P;W) \\
&\quad + \int_{{}^{1}((\; ; \; (p;w)])} {}_M(\frac{(P;W)}{h_n}) F_f(\; (P;W);P;W) d(P;W) \\
&= E_{11;n} + E_{12;n}
\end{aligned}
$$

For $E_{11;n}$, note that when $(P;W) \in \; {}^{1}([0; \;))$, $\frac{(P;W)}{h_n} \to +\infty$ and ${}_M(\frac{(P;W)}{h_n}) \to 1$ as $n \to \infty$. By assumption 2 and assumption 3, $| {}_M(\frac{(P;W)}{h_n}) F_f(\; (P;W);P;W)| < \infty$. Thus by Lebesgue's dominated convergence theorem we have

$$E_{11;n} \to \int_{{}^{1}([0; \;])} F_f(\; (P;W);P;W) d(P;W) = \int_{{}^{1}([0; \;])} \int_{[0; \; (P;W)]} f(\; ;P;W) d \; d(P;W):$$

For $E_{12;n}$, note that when $(P;W) \in \; {}^{1}((\; ; \; (p;w)])$, $\frac{(P;W)}{h_n} \to \; \infty$ and ${}_M(\frac{(P;W)}{h_n}) \to 0$ as

$n \to \infty$. As a result, $E_{1n} \to \int_{^{1}([0;\ )\cup(\ ;\ (p;w)])} \int_{[0;\ (P;W)]} f(\ ;P;W)d\ d(P;W)$.

$$
\begin{aligned}
E_{2n} &= \int_{^{1}([0;\ )\cup(\ ;\ (p;w)])} F_f(\ ;P;W) \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\ )d\ d(P;W) \\
&= \int_{^{1}([0;\ ))} F_f(\ ;P;W) \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\ )d\ d(P;W) \\
&\quad + \int_{^{1}((\ ;\ (p;w)])} F_f(\ ;P;W) \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\ )d\ d(P;W) \\
&= E_{21;n} + E_{22;n}
\end{aligned}
$$

For $E_{21;n}$, when $(P;W) \in\ ^{1}([0;\ ))$, $\frac{(P;W)}{h_n} \to +\infty$ and $\int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\ )d\ \to 0$ as $n \to \infty$. For $E_{22;n}$, when $(P;W) \in\ ^{1}((\ ;\ (p;w)])$, $\frac{(P;W)}{h_n} \to\ \infty$ and $\int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\ )d\ \to 1$ as $n \to \infty$. As a result, $E_{2n} \to \int_{^{1}([0;\ )\cup(\ ;\ (p;w)])} \int_{[0;\ ]} f(\ ;P;W)d\ d(P;W)$.

$$
h_n{}^{1}E_{3n} = \int_{^{1}([0;\ )\cup(\ ;\ (p;w)])} f(\ ;P;W) \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} M_k(\ )\ d\ d
$$

Similarly, when $(P;W) \in {}^{-1}([0; \;))$, $\frac{(P;W)}{h_n} \to +\infty$ and $E_{41;n} \to 0$ as $n \to \infty$. When $(P;W) \in {}^{-1}((\;;\;(p;w)])$, $\frac{(P;W)}{h_n} \to \infty$ and $\int^{\frac{}{h_n}}_{\frac{(P;W)}{h_n}} M_k(\;)^2 d \to {}^2_M$ as $n \to \infty$ by the symmetry of $M_k(:)$. As a result, $h_n{}^2 E_{4n} \to \frac{1}{2}\,{}^2_M \int_{{}^{-1}((\;;\;(p;w)])} f^{(1)}(\;;P;W)d(P;W)$. Therefore, if $0 < \;<\; (p;w)$,

$$E(\hat{P}(\;;p;w)) = E_{1n} + E_{2n} + E_{3n} + E_{4n} + A_{2n} + o(h_n^2)$$

$$= \int_{{}^{-1}([0;\;))} \int_{[0;\;]} {}^{(1)}$$

9 2]TJ/F32 7.9701 Tf 3.2910.9091 Tf 10.9 [(2)]TJ/F32 7.9701 Tf 0 -7.201 Td [(n)]TJ/F15 10.9091 Tf 5.

n

(b) Note that $V(\hat{P}(\ ;p;w)) = \frac{1}{n}(V_{1n} \quad V_{2n})$, where

$$V_{1n} = E[h_n^{\ \ 2}(\int_0 M_k(\frac{\quad}{h_n})d\ )^2 I(P_i \quad p; W_i \quad w)]$$

$$V_{2n} = (E[h_n^{\ 1} \int_0 M_k(\frac{\quad}{h_n})d\ \ I(P_i \quad p; W_i \quad w)])^2$$

From part (a), we know the limiting behavior of V_{2n}, now for V_{1n}, Since $h_n\ !\ 0$ as $n\ !\ 1$, there exist $N(p;w)\ 2\ \mathbb{R}_+$ such that for all $n > N(p;w)$,

$$\begin{aligned}
V_{1n} &= E[h_n^{\ \ 2}(\int_0 M_k(\frac{\quad}{h_n})d\ )^2 I(P_i \quad p; W_i \quad w)] \\
&= h_n^{\ \ 2} \int_{D_{p;w}} \int_{[0;\ \ (P;W)]} (\int_0 M_k(\frac{\quad}{h_n})d\ )^2 f(\ ;P;W)d\ \ d(P;W) \\
&= \int_{D_{p;w}} \int_{[0;\ \ (P;W)]} (\int_{B_M}^{\frac{\quad}{h_n}} M_k(')d')^2 f(\ ;P;W)d\ \ d(P;W) \\
&= \int_{D_{p;w}} \int_{[0;\ \ (P;W)]} (\ _M(\frac{\quad}{h_n}))^2 f(\ ;P;W)d\ \ d(P;W)
\end{aligned}$$

Like part (a), we also consider 3 cases when (1) $0 < \ < \ (p;w)$; (2) $\ > \ (p;w)$; (3) $\ = \ (p;w)$.

For case (1),

$$V_{1n} = \int_{^1([0;\ )[(\ ;\ (p;w)])} \int_{[0;\ (P;W)]} (\ _M \mathrm{M}$$

5.822 0 Td [( ))]TJ/F53 10.9091 Tf 4.24R8 14

p;nTJ/F55 10.9091 Tf 10.905 Td56 9d [( ))]TJ/F21 10.9091 Tf 7.42 Td (d [( )]TJ/F32 7.97/F15 10.9091 Tf 10.905 Td56 9d [( ))]TJ/F21 10.9091 Tff 6.515 0 Td [(( )]TJ/F53 1057 1]TJ/F53 109F53 10.9091 Tf 5.438 7.38 Td [( )]TJ/F21 10.9091 Tf 9.034 0 Td [( )

h

where $F_f(\ ;P;W) = \int_0 f(\ ;p;w)d$. Using integration by parts,

$$
\begin{aligned}
&\int_{[0;\ (P;W)]} (\ _M(\frac{}{h_n}))^2 \frac{@F_f(\ ;P;W)}{@} d \\
= &\int_{[0;\ (P;W)]} (\ _M(\frac{}{h_n}))^2 dF_f(\ ;P;W) \\
= &(\ _M(\frac{}{h_n}))^2 dF_f(\ ;P;W) j_{=0}^{=\ (P;W)} \int_{[0;\ (P;W)]} F_f(\ ;P;W) d(\ _M(\frac{}{h_n}))^2 \\
= &(\ _M(\frac{(P;W)}{h_n}))^2 F_f(\ (P;W);P;W) + \frac{2}{h_n} \int_{[0;\ (P;W)]} F_f(\ ;P;W)\ _M(\frac{}{h_n}) M_k(\frac{}{h_n}) d \\
= &(\ _M(\frac{(P;W)}{h_n}))^2 F_f(\ (P;W);P;W) + 2 \int_{\frac{(P;W)}{h_n}}^{\overline{h_n}} F_f(\ \ h_n\ ;P;W)\ _M(\ ) M_k(\ ) d \\
&))
\end{aligned}
$$

By Taylor's s71 Tf -431.923 -40.923 Td [(By)-333(T)83(a)28(ylor's)-34.e0.90or34.e0.90m, 6.286 -1.636 Td [(n)]3 -71

Now,

$$
\begin{aligned}
\int_{B_M}^{B_M} M_k(\ )\ _M(\ )d\ &= \int_{B_M}^{B_M}\ _M(\ )d\ _M(\ ) \\
&= \ _M(\ )^2 j^{B_M}_{B_M} \int_{B_M}^{B_M}\ _M(\ )d\ _M(\ ) \\
&= 1 \int_{B_M}^{B_M}\ _M(\ )d\ _M(\ )
\end{aligned}
$$

As a result, $\int_{B_M}^{B_M} M_k(\ )\ _M(\ )d\ = 1{=}2$. Therefore,

$$V_{12n}\ !\ \int_{^1((\ ;\ (p;w)])\ \ [0;\ ]} \int f(\ ;P;W)d\ d(P;W)$$

Similarly,

$$V_{13n}\ !\ 2h_n \int_{^1((\ ;\ (p;w)])} f(\ ;P;W)d\ d(P;W)$$

The result then follows. Case (2) and (3) follow similarly. □

Lemma 3. *Let h_n be a sequence of nonstochastic bandwidths such that $0 < h_n\ !\ 0$ as $n\ !$ 1. Given $w($*

h

6[(kB)]215]TJ/F5i10.909 4 12.65 0 Td [16[(k) 215]TJ/F53 10.9091 Tf 4.243 0 Td [0 (
M(1n)][(f)]TJ/F1o.9701.91 Tf16.97.0 Tdd 8(032 7.93; 10.9091 T7.091 17.933 Td

Then

$$
\begin{aligned}
\hat{P}(\;;p;w) &= (nh_n)^{1}\sum_{i=1}^{n}\Big(\int_{0} M_k(\frac{i}{h_n})d\;\Big)I(P_i\quad p;W_i\quad w) \\
&= h_n^{1}\int_{0} M_k(\frac{}{h_n})d\; I(P_i\quad p;W_i\quad w) \\
&= h_n^{1}\int_{\frac{i}{h_n}}^{\frac{i}{h_n}} M_k(')d'\,I(P_i\quad p;W_i\quad w) \\
&= h_n^{1}\int_{B_M}^{\frac{i}{h_n}} M_k(')d'\,I(P_i\quad p;W_i\quad w) \\
&= h_n^{1}\;{}_M(\frac{i}{h_n})I(P_i\quad p;W_i\quad w)
\end{aligned}
$$

Since [0

For $\quad 2\ B(\ _{l};(\frac{n}{h_n^a})\ \ ^{\frac{1}{2}})$, we have

$$
\begin{aligned}
P_{1n} &= j\hat{P}(\ ;p;w) \quad \hat{P}(\ _{l};p;w)j \\
& \quad h_n^{\ \ 1} \sum_{i=1}^{n} j\ _M(\frac{\ \ \ _i}{h_n}) \quad _M(\frac{\ _l \ \ _i}{h_n})jI(P_i \quad p;W_i \quad w) \\
& \quad h_n^{\ \ 1}
\end{aligned}
$$

Write $|\hat{P}(\ _l;p;w) \quad E(\hat{P}(\ _l;p;w))| = |\frac{1}{n}\sum_{i=1}^{n} W_{in}|$ where

$$W_{in} = \ _M(\frac{\ _l \quad \ _i}{h_n})I(P_i \quad p;W_i \quad w) \quad E[\ _M(\frac{\ _l \quad \ _i}{h_n})I(P_i \quad p;W_i \quad w)]$$

Obviously, $E(W_{in}) = 0$, $|W_{in}| \quad 2$ since both $I(.)$ and $\ _M(.)$ are less or equal to one. By Bernstein's inequality we have

$$P\{(\frac{n}{\ln(n)})^{\frac{1}{2}}|\hat{P}(\ _l;p;w) \quad E(\hat{P}(\ _l;p;w))| \qquad {}_{"}g\} < 2\exp(\frac{n\ ^2_{"}\ (\frac{n}{\ln(n)})\ ^1}{2\ ^2_n + \frac{4}{3}\ \ _{"}\ (\frac{n}{\ln(n)})\ ^{\frac{1}{2}}})$$

with $\ ^2_n = n\ ^1\sum_{i=1}^{n} V(W_{in}) \to P(\ _l;p;w)(1 \quad P(\ _l;p;w))$ by Lemma 1 or 2. Thus $2\ ^2_n + \frac{4}{3}\ _{"}\ (\frac{n}{\ln(n)})\ ^{\frac{1}{2}} \to 2P(\ _l;p;w)(1 \quad P(\ _l;p;w))$, Hence provided that $\ ^2_{"} > 2P(\ _l;p;w)(1 \quad P(\ _l;p;w))$,

$$\begin{aligned} P_{2n} & \qquad L_n P\{(\frac{n}{\ln(n)})^{\frac{1}{2}}|\hat{P}(\ _l;p;w) \quad E(\hat{P}(\ _l;p;w))| \qquad {}_{"}g \\ & < \ r\ (\frac{n}{h_n})^{\frac{1}{2}}\ 2\exp(\ \ln(n)) = r\ (nh)\ ^{\frac{1}{2}} \end{aligned}$$

Therefore, $P_{2n} = o_p(1)$ and as a result, $\sup_{\ \in[0;\ (p;w)]}|\hat{P}(\ ;p;w) \quad E(\hat{P}(\ ;p;w))| = o_p(1)$.

(b) Note that for $\ \in [0;\ (p;w)]$,

$$\begin{aligned} E(\hat{P}(\ ;p;w)) &= \int_{D_{p;w}}\int_{[0;\ (P;W)]} \ _M(\frac{\quad}{h_n})f(\ ;P;W)d\ d(P;W) \\ &= \int_{\ ^1([0;\ ))}\int_{[0;\ (P;W)]} \ _M(\frac{\quad}{h_n})f(\ ;P;W)d\mathrm{d}(P;W \end{aligned}$$

where

$$
\begin{aligned}
G_{1n} &= \Big| \int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W) - \int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} f(\ ;P;W)d\ \ d(P;W) \Big| \\
G_{2n} &= \Big| \int_{{}^{1}(f\ g)} \int_{[0;\ (P;W)]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W) - \int_{{}^{1}(f\ g)} \int_{[0;\ (P;W)]} f(\ ;P;W)d\ \ d(P;W) \Big| \\
G_{3n} &= \Big| \int_{{}^{1}((\ ;\ (p;w)])} \int_{[0;\ ]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W) - \int_{{}^{1}((\ ;\ (p;w)])} \int_{[0;\ ]} f(\ ;P;W)d\ \ d(P;W) \Big| \\
G_{4n} &= \Big| \int_{{}^{1}((\ ;\ (p;w)])} \int_{[\ ;\ (P;W)]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W) \Big|
\end{aligned}
$$

For the rst term, when $(P;W) \in {}^{1}([0;\ ))$, $(P;W) <$. This implies ${}_{M}(\frac{\quad}{h_n}) \to 1$ as $n \to \infty$. First, by LDC,

$$\int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W) \to \int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} f(\ ;P;W)d\ \ d(P;W).$$

Second, $\int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W)$ is increasing with n. Furthermore, By the Lipschitz condition imposed on ${}_{M}(\cdot)$,

$\int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W)$ is a continuous function in . As a result, by Dini's Theorem,

$$\int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} {}_{M}(\frac{\quad}{h_n}) f(\ ;P;W)d\ \ d(P;W) \to \int_{{}^{1}([0;\ ))} \int_{[0;\ (P;W)]} f(\ ;P;W)d\ \ d(P;W)$$

uniformly. Thus, $\sup_{\ \in[0;\ (p;w)]} G_{1n} = o(1)$. Similarly, we can prove that $\sup_{\ \in[0;\ (p;w)]} G_{2n} = o(1)$ and $\sup_{\ \in[0;\ (p;w)]} G_{3n} = o(1)$. For the last term, note when $\ \in [\ ;\ (P;W)]$, ${}_{M}(\frac{\quad}{h_n}) \to 0$ Similarly, by LDC and Dini's theorem, $\sup_{\ \in[0;\ (p;w)]} G_{4n} = o(1)$. □

Theorem 1 *Proof.* First we consider the event set $A = \{\omega : |\ _{;n}(p;w) - \ (p;w)| > \varepsilon\}$. Given $(p;w)$, provided that $(p;w)$ is unique, for any $\varepsilon > 0$, we have $F(\ (p;w) + \varepsilon|C_{p;w}) > F(\ (p;w)|C_{p;w}) > F(\ (p;w) - \varepsilon|C_{p;w})$. For $\omega \in A = \{\omega : |\ _{;n}(p;w) - \ (p;w)| > \varepsilon\}$, $\ _{;n}(p;w) > \ (p;w) + \varepsilon$ or $\ _{;n}(p;w) < \ (p;w) - \varepsilon$. By the monotonicity of $F(\cdot|C_{p;w})$, $F(\ _{;n}(p;w)|C_{p;w}) \geq F(\ (p;w) + \varepsilon|C_{p;w})$ or $F(\ _{;n}(p;w)|C_{p;w}) \leq F(\ (p;w) - \varepsilon|C_{p;w})$. Let

$$(\varepsilon;p;w) = \min\{F(\ (p;w)+\varepsilon|C_{p;w}) - F(\ (p;w)|C_{p;w}); F(\ (p;w)|C_{p;w}) - F(\ (p;w) - \varepsilon|C_{p;w})\} > 0$$

Theorem 2 *Proof.* (i) By Mean Value Theorem,

$$
\begin{aligned}
{}_{;n}(p;w) \qquad (p;w) &= \frac{\hat{F}(\quad {}_{;n}(p;w)jC_{p;w}) \quad \hat{F}(\quad (p;w)jC_{p;w})}{\hat{f}(\quad {}_{;n}(p;w)jC_{p;w})} \\
&= \frac{F(\quad (p;w)jC_{p;w}) \quad \hat{F}(\quad (p;w)jC_{p;w})}{\hat{f}(\quad {}_{;n}(p;w) \quad)}
\end{aligned}
$$

G is a compact set and $G \quad (0; \quad (p;w))$.

Note that $f(\ |C_{p;w}) = \frac{\int_{\ ^{1}((\ ;\ (p;w)])} f(\ ;P;W)d(P;W)}{P_{PW}(p;w)}$ since when $(P;W) \in \ ^{1}([0;\ ])$, $\quad (P;W)$. $F(\ |C_{p;w}) = 1$ and $\frac{\partial F(\ |C_{p;w})}{\partial} = 0$

$$
\begin{aligned}
& \sup_{\in G} |\hat{f}(\ |C_{p;w}) \quad f(\ |C_{p;w})| \\
= \quad & \sup_{\in G} \left| \frac{(nh_n)^{\ 1} \sum_{i=1}^{n} M_k(\frac{i}{h_n}) I(P_i \quad p; W_i \quad w)}{\hat{P}_{PW}(p;w)} \quad \frac{\int_{\ ^{1}((\ ;\ (p;w)])} f(\ ;P;W)d(P;W)}{P_{PW}(p;w)} \right| \\
& \frac{1}{\hat{P}_{PW}(p;w)} \sup_{\in G} |(nh_n)^{\ 1} \sum_{i=1} M_k(\frac{i}{h_n}) I(P_i \quad p; W_i \quad w) \quad \int_{\ ^{1}((\ ;\ (p;w)])} f(\ ;P;W)d(P;W)| \\
& + \left| \frac{1}{P_{PW}(p;w)} \quad \frac{1}{\hat{P}_{PW}(p;w)} \right| \sup_{\in G} \int_{\ ^{1}((\ ;\ (p;w)])} f(\ ;P;W)d(P;W)
\end{aligned}
$$

Since $\frac{1}{P_{PW}(p;w)} \quad \frac{1}{\hat{P}_{PW}(p;w)} = o_p(1)$ by Slutsky theorem,

$$
\sup_{\in G} \int_{\ ^{1}((\ ;\ (p;w)])} f(\ ;P;W)d(P;W) \quad B_f \int_{\ ^{1}((\ ;\ (p;w)])} d(P;W) = O(1)
$$

by Assumptions 3 and 4.

Denote $Q_n(p;w) = (nh_n)^{\ 1} \sum_{i=1}^{n} M_k(\frac{i}{h_n}) I(P_i \quad p; W_i \quad w)$, Thus,

$$
\begin{aligned}
& \sup_{\in G} |Q_n(p;w) \quad \int_{\ ^{1}((\ ;\ (p;w)])} f(\ ;P;W)d(P;W)| \\
& \sup_{\in G} |Q_n(p;w) \quad E(Q_n(p;w))| \\
& + \sup_{\in G} |E(Q_n(p;w)) \quad \int_{D_{p;w}} \ _M(\frac{(P;W)}{h_n}) f(\ ;P;W)d(P;W)| \\
& + \sup_{\in G} \left| \int_{\ ^{1}((\ ;\ (p;w)])} \ _M(\frac{(P;W)}{h_n}) f(\ ;P;W)d(P;W) \quad \int_{\ ^{1}((\ ;\ (p;w)])} f(\ ;P;W)d(P;W) \right| \\
& + \sup_{\in G} \left| \int_{\ ^{1}([0;\ ])} \ _M(\frac{(P;W)}{h_n}) f(\ ;P;W)d(P;W) \right| \\
= \quad & Q
\end{aligned}
$$

nh_n^2 ! 1. For any $(p; w$

(b):

$$
\begin{aligned}
A_n &= F(\quad (p;w)|C_{p;w}) \quad \frac{E(\hat{P}(\quad (p;w);p;w))}{E(\hat{P}_{PW}(p;w))} \\
&= \frac{E(\hat{P}_{PW}(p;w))F(\quad (p;w)|C_{p;w})}{E(\hat{P}_{PW}(p;w))} \quad \frac{P(\quad (p;w);p;w)}{E(\hat{P}_{PW}(p;w))} \\
&\quad + \frac{P(\quad (p;w);p;w)}{E(\hat{P}_{PW}(p;w))} \quad \frac{E(\hat{P}(\quad (p;w);p;w))}{E(\hat{P}_{PW}(p;w))} \\
&= \frac{1}{E(\hat{P}_{PW}(p;w))}[(E(\hat{P}_{PW}(p;w))F(\quad (p;w)|C_{p;w}) \quad P(\quad (p;w);p;w)) \\
&\quad +(P(\quad (p;w);p;w) \quad E(\hat{P}(\quad (p;w);p;w)))] \\
&= \frac{1}{E(\hat{P}_{PW}(p;w))}(A_{1n} + A_{2n})
\end{aligned}
$$

we know $E(\hat{P}_{PW}(p;w)) = P_{PW}(p;w)$. $A_{1n} = 0$. Since given $\quad 2\ (0;1)$, $\quad (p;w)\ 2\ (0; \quad (p;w))$, by Lemma 2,

$$
A_{2n} = \frac{1}{2}h_n^2 \ {}^2_M \int_{{}^1((\quad (p;w); \quad (p;w)])} f^{(1)}(\quad (p;w);P;W)d(P;W) + o(h_n^2)
$$

The result then follows.

(c):

$$
\begin{aligned}
\sqrt{n}C_n &= \sqrt{n}(\frac{E(\hat{P}(\quad (p;w);p;w))}{E(\hat{P}_{PW}(p;w))} \quad \hat{F}(\quad (p;w)|C_{p;w})) \\
&= \sqrt{n}(\frac{E(\hat{P}(\quad (p;w);p;w))\hat{P}_{PW}(p;w)}{E(\hat{P}_{PW}(p;w))\hat{P}_{PW}(p;w)} \quad \frac{\hat{P}(\quad (p;w);p;w)}{\hat{P}_{PW}(p;w)}) \\
&= \frac{1}{\hat{P}_{PW}(p;w)} \sum_{i=1}^{n} Z_{in}
\end{aligned}
$$

where

$$
Z_{in} = \frac{1}{\sqrt{n}}(\frac{E(\hat{P}(\quad (p;w);p;w))}{P_{PW}(p;w)}I(P_i \quad p;W_i \quad w) \quad \frac{1}{h_n}\int_0^{ \quad (p;w)} M_k(\frac{ \quad i}{h_n})d \quad I(P_i \quad p;W_i \quad w))
$$

Here,

$$
\begin{aligned}
E(Z_{in}) &= \sqrt{\frac{1}{n}}(E(\hat{P}(\quad(p;w);p;w)) \quad E(\hat{P}(\quad(p;w);p;w))) \\
&= 0
\end{aligned}
$$

$$\sum_{i=1}^{n} E($$

$$
\begin{aligned}
\sum_{i=1} E(Z_{in}^2) &= s_{1n} + s_{2n} + s_{3n} \\
&= P(\quad(p;w);p;w) \quad \frac{(P(\quad(p;w);p;w))^2}{P_{PW}(p;w)} \\
&\quad 2h_n \int_{^1((\quad(p;w);\ (p;w)])} f(\quad(p;w);P;W)d(P;W) + o(h_n)
\end{aligned}
$$

By Liapounov's CLT, $\sum_{i=1}^{n} \frac{Z_{in}}{s_n(p;w)} \overset{d}{!} N(0;1)$ if $\lim_{n\,!\,1} \sum_{i=1}^{n} E(j\frac{Z_{in}}{s_n(p;w)}j^{2+}\ ) = 0$ for some $\quad > 0$.

$$
\sum_{i=1} E(j\frac{Z_{in}}{s_n(p;w)}j^{2+}\ ) \quad \sum_{i=1} E(jZ_{in}j^{2+}\ j\frac{1}{s_n(p;w)}j^{2+}\ )
$$

Since $s_n(p;w) = O(1)$, we just need to prove $\lim_{n\,!\,1} \sum_{i=1}^{n} E(jZ_{in}j^{2+}\ ) = 0$. By C_r Inequality,

$$
\begin{aligned}
\sum_{i=1} E(jZ_{in}j^{2+}\ ) &\quad 2^{1+}\ (n^{\ 2=}\ E(j\frac{E(\hat{P}(\quad(p;w);p;w))}{P_{PW}(p;w)}I(P_i\quad p;W_i\quad w)j^{2+}\ ) \\
&+ \quad n^{\ 2=}\ E(j\frac{1}{h_n}\int_0^{(p;w)} M_k(\frac{i}{h_n})d\ \ I(P_i\quad p;W_i\quad w)j^{2+}\ )) \\
&= 2^{1+}\ (n^{\ 2=}\ Ej\frac{E(\hat{P}(\quad(p;w);p;w))}{P_{PW}(p;w)}j^{2+}\ E(I(P_i\quad p;W_i\quad w))) \\
&\quad +n^{\ 2=} \int_{D_{p;w}} \int_{[0;\ (P;W)]} {}_M(\frac{(p;w)}{h_n})f(\ ;P;W)d\ \ d(P;W)
\end{aligned}
$$

Since $E(I(P_i\quad p;W_i\quad w) = O(1)$,

$$
\begin{aligned}
n^{\ 2=}\ Ej\frac{E(\hat{P}(\quad(p;w);p;w))}{P_{PW}(p;w)}j^{2+} &= n^{\ 2=}\ \frac{jE(\hat{P}(\quad(p;w);p;w))j^{2+}}{P_{PW}(p;w)^{2+}} \\
&= O(n^{\ 2=}\ )
\end{aligned}
$$

Since ${}_M(:)\quad 1$, $f < B_f$ and $\quad B\ $,

$$
\begin{aligned}
& n^{\ 2=} \int_{D_{p;w}} \int_{[0;\ (P;W)]}^{f)]} {}_M(\frac{(p;w)}{h_n})f(\ ;P;W)d\ \ d(P;W) \\
& n^{\ 2=}\ B_f \int_{D_{p;w}} \int_{[0;\ (P;W)]} d\ \ d(P;W) \\
& n^{\ 2=}\ B_f \int^{D_{p;w}}_{^1[0;B\ \ ]}
\end{aligned}
$$

The result then follows.

(ii) Note that in the proof of part (i), $A_n = \frac{1}{E(\hat{P}_{PW}(p;w))}(A_{1n} + A_{2n})$ is the bias term and $A_{1n} = 0$.
by Lemma 1,

$$\begin{aligned} |A_{2n}| &= |P(\ (p;w);p;w) \quad E(\hat{P}(\ (p;w);p;w))| \\ & \quad ch_n^{2k}[\int_{D_{p;w}} H_{2k}(\ (p;w);P;W)d(P;W) + \int_{D_{p;w}} \sup_{\ \in \mathbb{R}} |F_f(\ ;P;W)|\varepsilon_{2k}^{2k}(\ (p;w);P;W)d(P;W)] \end{aligned}$$

The result then follows. □

References

Korostelev, A., Simar, L. and Tsybakov, A. B. (1995). E cient estimation of monotone boundaries, *Annals of Statistics* **23**: 476{89.

Kumbhakar, S. C., Park, B. U., Simar, L. and Tsionas, E. (2007). Nonparametric stochastic frontiers: A local maximum likelihood approach, *Journal of Econometrics* **137**: 1{27.

Kumbhakar, S. and Lovell, C. A. K. (2000). Stochastic frontier analysis, *Cambridge University Press* .

Martins-Filho, C. and Yao, F. (2008). A smooth nonparametric conditional quantile frontier estimator, *Journal of Econometrics* **143**: 317{333.

Martins-Filho, C. and Yao, F. (2013). Semiparametric stochastic frontier estimation via pro le likelihood, *Econometric Reviews* **34**: 413{451.

Maudos, J., Pastor, J. M., Perez, F. and Quesada, J. (2002). Cost and pro t e ciency in european banks, *Journal of International Financial Markets, Institutions and Money* **12**: 33{58.

Meeusen, W. and van den Broeck, J. (1977). E ciency estimation from cobb-douglas production