DISCUSSION PAPERS IN ECONOMICS

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 are useful in measuring profit efficiency. Second, we propose a nonparametric conditional quantile estimator for the profit function of order α based on a recently proposed class of nonparametric kernel estimators introduced by Mynbaev and Martins-Filho (2010). We establish consistency and asymptotic normality of our estimator. Our measures of efficiency are more robust to outliers since our estimated profit functions of order α do not envelope the data. Under some smoothness assumption on the distribution function, the bias of our proposed estimator converges to zero faster than that of the estimator which uses traditional kernels. A Monte-Carlo simulation seems to support our results and show a better performance of our estimator compared to its competitors in most scenarios.

Keywords and phrases. Profit Function, Nonparametric conditional quantile estimation, Efficiency Analysis.

JEL classifications. C13, C14.

*I am deeply indebted to Carlos Martins-Filho for his advise and support. I would also like to thank Xiaodong Liu, Shawna Grosskopf and Rolf Färe for their help and suggestions. All errors are my own.

†Department of Economics, 256 UCB, Boulder, CO 80309. (Email: shan.zhou@colorado.edu; Website: `http://sites.google.com/site/zhoush54/home`)

1 Introduction

Efficient behavior permeates firm decision making in the classical microeconomic theory of the firm. For example, given a technology, firms are assumed to produce maximum output given a set of inputs; minimize cost and and maximize profit. However, there exists voluminous empirical evidence that suggests not all producers engage in successful optimizing behavior. In various settings it is useful to measure and study the magnitude and nature of the inefficiency that pulls firms away from the relevant efficient frontier, be it a production, cost or profit function. Starting with Farrell (1957), a vast literature has emerged focusing on production or technical efficiency. These studies usually postulate a common production frontier for all firms, and measure the technical efficiency by a suitably defined distance between the production plan of each firm and the frontier. However, the ultimate objective of a firm is to maximize profit. Besides technical efficiency, a major component in a firm's profit-seeking behavior involves allocative efficiency, which captures the ability of choosing optimal proportions of inputs and outputs in the production process. In this paper, we provide a way of measuring and estimating a profit function and profit efficiency, which represents the combined effect of both technical and allocative efficiency. The basic idea, inspired by Aragon et al. (2005), is to describe profit functions at different efficiency levels as quantile functions of a suitably defined conditional distribution. We then estimate them by an improved nonparametric kernel method.

Empirical and theoretical models for measuring efficiency and estimating frontiers have fallen into two broad categories: stochastic and nonstochastic models of frontier. The basic principle in stochastic models is to describe the variable of interest (output, cost, profit) as being generated by a sum of the function of interest and a non-observed error term consisting of a noise and an inefficiency term. It was originally proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977) in the context of production functions. Greene (2008) and Kumbhakar and Lovell (2000) provide extensive reviews and applications of these models. Estimation of these models is normally conducted by maximum likelihood methods. Considering the possibility of mis-specification of parametric frontier models, Fan et al. (1996) investigate semiparametric estimation of a stochastic frontier model with nonparametric production function. Recent developments in the estimation 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 and any deviation from the frontier is caused by inefficiency. The most popular nonparametric efficiency estimators are based on the idea of estimating the attainable set by the smallest set within some class that envelops the observed data. Data envelopment analysis (DEA) and free disposal hull (FDH) estimators are among the most popular and have been widely used in efficiency analysis since Charnes et al. (1978). The FDH estimator of the frontier is the free disposal hull of the observations and the DEA estimator is the convex cone of the FDH estimator. They rely on linear programming methods to search for the most efficient units, which are then connected to form a minimum enveloping frontier. DEA and FDH are very appealing to researchers because they rely on very few assumptions and are easy to implement; however, they suffer some critical drawbacks.

Park et al. (2010, 2000) and Simar and Vanhems (2012) obtain general asymptotic properties and convergence rates of FDH and DEA estimators under certain assumptions. Convergence rates are also obtained in Kneip et al. (1998); Korostelev et al. (1995) and Gijbels et al. (1999) in special cases. Similar to other nonparametric estimators, DEA and FDH estimators suffer the "curse of dimensionality." The convergence speed of these estimators becomes much slower as the dimensionality of the problem increases. Another major drawback of these methods is that the estimation of the frontier is highly influenced by the most efficient firms, outliers or extreme values. Hence, these methods are not robust and highly sensitive to a small set of observations. Recently, Simar and Zelenyuk (2011) propose stochastic versions of the FDH and DEA estimators by allowing noise into the model, but the properties of these estimators remain unknown.

Considering these drawbacks, Cazals et al. (2002) introduce the concept of production frontier of order m and provide a robust envelopment estimator. Instead of the full production frontier, they consider the expected maximum output among m firms drawn from the population of firms using less than a given level of inputs. A new probabilistic interpretation of the frontiers and the efficiency scores is provided in the paper. Daouia and Simar (2005, 2007) and Daouia et al. (2010) further extend this idea and link frontier estimation to extreme value theory. Inspired by this idea, Aragon et al. (2005) introduces the quantile approach in the production frontier analysis. They define a production function of continuous order α based on conditional quantiles of a distribution that describes the generation of inputs and outputs of a production process. Estimators for these conditional quantiles are by nature much more robust to outliers as they do not envelope all observations. Martins-Filho and Yao (2008) improves this method by introducing a smooth kernel estimator at the cost of introducing a bias term vanishing with the sample size. In this paper, inspired by Aragon et al. (2005), we define a profit frontier of continuous order α and propose an easy-to-implement nonparametric estimator for these profit frontiers. Our estimator is based on the kernel estimator proposed in Martins-Filho and Yao (2008), but we improve on their estimation method by using a new class of kernels proposed by Mynbaev and Martins-Filho (2010) that promise to reduce the order of the bias of the class of estimators under study.

Throughout the paper we consider competitive firms with technology represented by a production function $y = f(x)$ where $y \in \mathbb{R}_+^{d_1}$ is an output vector and $x \in \mathbb{R}_+^{d_2}$ is an input vector. Given a vector of input prices $w = (w_1, \cdots, w_{d_2}) \in \mathbb{R}_{++}^{d_2}$ and a vector of output prices $p = (p_1, \cdots, p_{d_1}) \in \mathbb{R}_+^{d_1}$, profit is given by

$$\pi = pf(x) - wx.$$

We assume that for all prices $w \in \mathbb{R}_{++}^{d_2}$ and output price $p \in \mathbb{R}_+^{d_1}$, profit is bounded above as a function of x, i.e. there exists $0 < B_\pi < +\infty$ such that $0 \leq \pi \leq B_\pi$. One can maximize the profit with respect to x to get the input demand functions $x^* = x(p, w)$. If the maximum exists, then the maximized profit is given by

$$\pi^* = \pi(p, w) = pf(x(p, w)) - wx(p, w),$$

π^* depend on p and w which are exogenous to the firm's decision making process. $\pi(p, w)$ is 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 to assume a parametric form for the profit function. Second, the production function is monotonic nondecreasing with respect to its arguments (inputs), while the profit function is nondecreasing with respect to some of its arguments (output prices) and nonincreasing with other arguments (input prices). Perhaps, due to these difficulties there is a much smaller literature devoted to the analysis of profit efficiency. Existing empirical studies, such as Ali et al. (1994) and Maudos et al. (2002) are mostly based on parametric stochastic profit frontier models with very high probability of misspecification. We show in this paper that these problems can be solved by our nonparametric quantile approach.

The rest of the paper contains five additional sections and an appendix. Section 2 describes the model and its estimation in detail. Section 3 provides the main assumptions and theorems that establish the asymptotic behavior of our estimators. Section 4 contains a small Monte Carlo study that implements the estimator, investigates its finite sample properties and compares performances of smooth and nonsmooth estimators. Section 5 provides a conclusion and some directions for future work. The proofs for all propositions and theorems are collected in the appendix, where a set of auxiliary lemmas are also given.

2 Model and Estimation

2.1 Profit function of order α

Let $\{(\Pi_i, P_i, W_i)\}_{i=1}^n$ be a sequence of independent and identically distributed random vectors defined in the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and having the same distribution function as (Π, P, W), which is denoted by F with associated density function f. $\Pi_i \in \mathbb{R}$ denotes profit, $P_i \in \mathbb{R}^{d_1}_{++}$ denotes a vector of output prices and $W_i \in \mathbb{R}^{d_2}_{++}$ denotes a vector of input prices associated with a firm or producing unit i. We denote the support of f by Ψ and focus on the set $\Psi^* = \{(\pi, p, w) \in \Psi : \mathcal{P}(P \leq p, W \geq w) > 0\}$. Given $C_{p,w} = \{P \leq p, W \geq w\}$ we let

$$F(\pi|C_{p,w}) = \mathcal{P}(\Pi \leq \pi|P \leq p, W \geq w) = \frac{\mathcal{P}(\Pi \leq \pi, P \leq p, W \geq w)}{\mathcal{P}(P \leq p, W \geq w)}. \tag{1}$$

and give the following probabilistic definition of a profit function

$$\pi(p, w) := \inf\{\pi \in [0, B_\pi] : F(\pi|C_{p,w}) = 1\}. \tag{2}$$

As defined, the profit function $\pi(p, w)$ is the "smallest" function that is larger than or equal to the highest attainable profit given input price larger than or equal to w and output prices less than or equal to p (vector inequalities are all taken element-wise). By its definition, for any (Π_i, P_i, W_i) with $P_i \leq p$ and $W_i \geq w$, we must have $\Pi_i \leq \pi(p, w)$ with probability 1. That is, $\pi(p, w)$ envelopes all data points.

Similar to Aragon et al. (2005), in the context of production functions, our definition of profit function suggests the alternative concept of a profit function of continuous order $\alpha \in [0, 1]$, as the quantile function of order α of the conditional distribution of Π given $C_{p,w}$. Thus, we define

$$\pi_\alpha(p, w) := F^{-1}(\alpha|C_{p,w}) = \inf\{\pi \in [0, B_\pi] : F(\pi|C_{p,w}) \geq \alpha\} \tag{3}$$

where $F^{-1}(\cdot|C_{p,w})$ is the generalized inverse of $F(\cdot|C_{p,w})$. We call $\pi_\alpha(p, w)$ the profit function of order α. It is apparent that the profit function in (2) corresponds to that in (3) when $\alpha = 1$. By definition $F^{-1}(\alpha|C_{p,w})$ is the profit threshold exceeded by $100(1-\alpha)\%$ of firms that face input prices larger than or equal to w and output prices less than or equal to p. If the conditional distribution $F(\cdot|C_{p,w})$ is strictly increasing, then we have the following result.

Proposition 1. *Assume that for every (p, w) such that $\mathcal{P}(P \leq p, W \geq w) > 0$, the conditional distribution function $F(\cdot|C_{p,w})$ is strictly increasing on the support $[0, \pi(p, w)]$. Then, for any $(\pi, p, w) \in \Psi^*$, we have $\pi = \pi_\alpha(p, w)$ with $\alpha = F(\pi|C_{p,w})$.*

Proposition 1 shows that any vector $(\pi, p, w) \in \Psi^*$ belongs to some profit function of order α.

That is, the quantile curves $\{(\pi_\alpha(p,w),p,w)|\mathcal{P}(P \le p, W \ge w) > 0, \alpha \in (0,1]\}$ cover the entire set Ψ^* of attainable profits, input and output prices. Given a firm or production unit associated with $(\pi_\alpha(p,w),p,w)$, its profit is larger than $100\alpha\%$ of all units facing the same or less favorable prices (higher input prices and lower output prices) and and less than $100(1-\alpha)\%$ all other firms or production units. Thus, the order of the conditional quantile curve to which (π,p,w) belongs, gives a measure of "profit efficiency" of the firm or production unit (π,p,w) relative to all other production units facing the same or less favorable prices.

It is clear that, for any fixed (p,w) such that $\mathcal{P}(P \le p, W \ge w) > 0$, $\pi_\alpha(p,w)$ is a monotone nondecreasing function of α. The following proposition shows that as $\alpha \to 1$, $\pi_\alpha(p,w)$ converge to $\pi(p,w)$ pointwise, and under additional regularity condition, the convergence is uniform over a suitably defined set.

Proposition 2. *For any fixed (p,w) such that $\mathcal{P}(P \le p, W \ge w) > 0$, $\lim_{\alpha\to 1}\pi_\alpha(p,w) = \pi(p,w)$. If, in addition, for every $\alpha \in [0,1]$, $\pi_\alpha(p,w)$ is continuous on the interior of the support of the marginal density of (P,W), denoted by S_0, then for any compact subset $\Phi \subset S_0$*

$$\sup_{(p,w)\in\Phi} |\pi_\alpha(p,w) - \pi(p,w)| \to 0 \text{ as } \alpha \to 1.$$

The most natural measure of profit efficiency of a firm or production unit i, compares its realized profit Π_i to the profits attained by all firms facing output prices $p \le P_i$, the output prices faced by unit i, and input prices $w \ge W_i$, the input prices faced by unit i for $\alpha = 1$. However, in an attempt to decrease the sensitivity of our measurement of profit efficiency to outliers or extreme values, we introduce a new measure of efficiency that compares the profit of a production unit to a profit function of order α. Thus, we say that the firm or production unit i is α-profit efficient if its profit $\Pi_i \ge \pi_\alpha(P_i,W_i)$. Otherwise, such firm is labeled α-profit inefficient. Thus, we can define an α-efficiency score as $e_\alpha(\Pi_i,P_i,W_i) = \Pi_i/\pi_\alpha(P_i,W_i)$. Note, that different from efficiency scores that emerge from traditional frontiers that envelope all possible triples (Π_i,P_i,W_i), $e_\alpha(\Pi_i,P_i,W_i)$ may be greater than 1, since the profit function of order α does not provide an upper bound for the profits of all firms facing prices $p \le P_i$ and $w \ge W_i$.

The concept of profit functions of order α 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 it is "efficient". As a result, it may be necessary to consider production capacity when assessing profit efficiency. In these cases, we can adjust our definition of profit frontier by comparing units with the same or smaller production capacities. Thus, if we use y to measure production capacity, then our definition of profit function becomes

$$\pi(y, p, w) := \inf\{\pi \in [0, B_\pi] : F(\pi | C_{y,p,w}) = 1\}$$

and

$$\pi_\alpha(y, p, w) := F^{-1}(\alpha | C_{y,p,w}) = \inf\{\pi \in [0, B_\pi] : F(\pi | C_{y,p,w}) \geq \alpha\}$$

where $C_{y,p,w}$ is a conditional set $\{Y \leq y, P \leq p, W \geq w\}$ and $F(\pi | C_{y,p,w})$ is a conditional distribution

$$F(\pi | C_{y,p,w}) = \mathcal{P}(\Pi \leq \pi | Y \leq y, P \leq p, W \geq w) = \frac{\mathcal{P}(\Pi \leq \pi, Y \leq y, P \leq p, W \geq w)}{\mathcal{P}(Y \leq y, P \leq p, W \geq w)}.$$

Therefore, as before, $\pi(y, p, w)$ represents the smallest function that is larger than or equal to the highest attainable profit given input prices larger than or equal to w, output prices less than or equal to p, and with capacity less than or equal to y. $\pi_\alpha(y, p, w)$ is defined by comparing it with all units with same or smaller production capacities facing the same or less favorable input and output prices.

Second, consider a firm that has monopoly power. Then, market demand affects output prices and output jointly, and only input prices can be viewed as exogenous to the firm's profit maximizing problem. In this case, we can define

$$\pi_\alpha(\xi, w) := \inf\{\pi \in [0, B_\pi] : F(\pi | C_{\xi,w}) \geq \alpha\}$$

where ξ represent the elasticity of market demand and $F(\pi|C_{\xi,w}) = \mathcal{P}(\Pi \leq \pi|\Upsilon \geq \xi, W \geq w)$. The analysis and estimation procedures defined in the following subsection can easily be extended to these alternative settings with minor modifications.

2.2 Estimation

In order to estimate profit functions of order α, i.e., $\pi_\alpha(p, w)$, we first need an estimator for a conditional cumulative distribution function $F(\pi|C_{p,w})$. In a production function setting, Aragon et al. (2005) propose a simple estimator based on the empirical distribution function. Their empirical estimator is not smooth and as a result, it might be difficult to identify differences between firms that are similar in terms of profit efficiency. In the same setting Martins-Filho and Yao (2008) proposed a smooth kernel based estimator. The smoothness it provides might reduce the finite sample variance compared to the empirical estimator, but introduces a bias that does not vanish at the parametric rate. Here, we follow Martins-Filho and Yao (2008), but provide an alternative kernel that may produce biases of lower order. For convenience, we define the functions $P(\pi, p, w) = \mathcal{P}(\Pi \leq \pi, P \leq p, W \geq w)$ and $P_{PW}(p, w) = \mathcal{P}(P \leq p, W \geq w)$. We estimate $F(\pi|C_{p,w})$ by integrating a smooth kernel density estimator constructed using the observations $\{(\Pi_i, P_i, W_i)\}_{i \in \{i: P_i \leq p, W_i \geq w\}}$. Thus, we define

$$\hat{F}(\pi|C_{p,w}) = \begin{cases} 0 & \text{if } \pi \leq 0 \\ \frac{\hat{P}(\pi,p,w)}{\hat{P}_{PW}(p,w)} & \text{if } \pi > 0 \end{cases} \tag{4}$$

with

$$\hat{P}(\pi, p, w) = (nh_n)^{-1} \sum_{i=1}^{n} \left(\int_0^\pi M_k \left(\frac{\Pi_i - \gamma}{h_n} \right) d\gamma \right) I(P_i \leq p, W_i \geq w) \tag{5}$$

and

$$\hat{P}_{PW}(p, w) = n^{-1} \sum_{i=1}^{n} I(P_i \leq p, W_i \geq w). \tag{6}$$

h_n is a nonstochastic sequence of bandwidths such that $0 < h_n \to 0$ as $n \to \infty$, $I(A)$ is the indicator function for the set A and M_k for $k = 1, 2, \cdots$ is a class of kernels defined by Mynbaev and Martins-Filho (2010). The kernels M_k are defined as

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

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 function. It is easy to show that $M_k(x)$ is a kernel function for all k in that $\int M_k(x)dx = 1$. The main advantage of the definition of $M_k(x)$ is that it allows us to express the bias of our estimator in terms of higher order finite differences of the density function (See the proof in the Lemma 1). It is clear that $\hat{F}(\pi|C_{p,w})$ depends on k through the dependence of $\hat{P}(\pi, p, w)$ on k. As a result, we are defining a class of estimators for $F(\pi|C_{p,w})$. The choice of k depends on the smoothness assumption on the distribution function, and will be discussed in the next section. Also, note that our estimator uses a smooth nonparametric estimator of the distribution function in the direction of profit π, but still uses an empirical distribution function in the direction of p and w. In the context of a production function, Martins-Filho and Yao (2008) showed that smooth kernel based estimator implemented in output direction has a parametric ($\sqrt{n}$) rate of convergence. In the next section we will show that our estimator has the same convergence. Note that it is possible to smooth estimators in the directions of prices as well, but as a result the estimator would suffer from the well-known "curse of dimensionality."

Assuming that $\pi_\alpha(p, w)$ is the unique root of $F(\cdot|C_{p,w}) = \alpha$, we denote its estimator by $\pi_{\alpha,n}(p, w)$, the root of

$$\hat{F}(\pi_{\alpha,n}(p, w)|C_{p,w}) = \alpha \text{ for } \alpha \in (0, 1],\ p \in \mathbb{R}_+^{d_1} \text{ and } w \in \mathbb{R}_{++}^{d_2}. \tag{8}$$

By absolute continuity of $F(\cdot|C_{p,w}))$, smoothness of the seed kernel, and the Mean Value Theorem, we write

$$\pi_{\alpha,n}(p, w) - \pi_\alpha(p, w) = \frac{F(\pi_\alpha(p, w)|C_{p,w}) - \hat{F}(\pi_\alpha(p, w)|C_{p,w})}{\hat{f}(\bar{\pi}_{\alpha,n}(p, w)|C_{p,w})}$$

where

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

and $\bar{\pi}_{\alpha,n}(p,w) = \lambda \pi_{\alpha,n}(p,w) + (1-\lambda)\pi_\alpha(p,w)$ for some $\lambda \in (0,1)$. In the following section we provide some asymptotic characterizations for our estimator, including consistency and asymptotic normality.

3 Asymptotic characterization of $\pi_{\alpha,n}$

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. *$\{(\Pi_i, P_i, W_i)\}_{i=1}^n$ is a sequence of independent random vectors taking values in a compact set $\Psi^* = [0, B_\pi] \times S_{PW}$ where S_{PW} is a compact set in $\mathbb{R}_+^{d_1} \times \mathbb{R}_{++}^{d_2}$. For any i, (Π_i, P_i, W_i) have the same joint distribution F and joint density function f as the vector (Π, P, W), f is defined on $\mathbb{R} \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ with support Ψ^*.*

Assumption 2. *(i) The seed kernel $K(\cdot)$ is a bounded symmetric density with compact support $[-B_K, B_K]$ and $\int_{-B_K}^{B_K} \gamma K(\gamma) d\gamma = 0$. (ii) $\int_{-B_K}^{B_K} \gamma^2 K(\gamma) d\gamma = \sigma_K^2$. (iii) For any $\gamma, \gamma' \in [-B_K, B_K]$, we have $|K(\gamma) - K(\gamma')| \le m_K |\gamma - \gamma'|$ for some $0 < m_K < \infty$. (iv) For all $\zeta, \zeta' \in \mathbb{R}$, we have $|\kappa(\zeta) - \kappa(\zeta')| \le m_\kappa |\zeta - \zeta'|$ for some $0 < m_\kappa < \infty$, where $\kappa(\zeta) = \int_{-B_K}^{\zeta} K(\gamma) d\gamma$. (v) For fixed k, $\int |K(t)| t^{2k} dt < \infty$.*

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 $[-B_M, B_M]$. $\int_{-B_M}^{B_M} \gamma M_k(\gamma) d\gamma = 0$; (ii) $\int_{-B_M}^{B_M} \gamma^2 M_k(\gamma) d\gamma := \sigma_M^2 = 2\sigma_K^2 \sum_{s=1}^{k} \lambda_{k,s} s^2$; (iii) For any $\gamma, \gamma' \in [-B_M, B_M]$, we have $|M_k(\gamma) - M_k(\gamma')| \leq m_M |\gamma - \gamma'|$ for some $0 < m_M < \infty$; (iv) For any $\zeta, \zeta' \in \mathbb{R}$, we have $|\kappa_M(\zeta) - \kappa_M(\zeta')| \leq m_\kappa |\zeta - \zeta'|$ for some $0 < m_\kappa < \infty$, where $\kappa_M(\zeta) = \int_{-B_M}^{\zeta} M_k(\gamma) d\gamma$.[1]

Assumption 3. *For all π and $\pi' \in G$, where G is a compact set, we have*

$$\left| \int_{\pi^{-1}([\pi,\pi'])} d(P, W) \right| \leq m_{\pi^{-1}} |\pi' - \pi|$$

for some $0 < m_{\pi^{-1}} < \infty$.[2]

Assumption 3 is similar to Assumption 4 in Martins-Filho and Yao (2008). It imposes a Lipschitz type condition on the inverse image π^{-1} of π.

Assumption 4. *(i) The joint distribution density function f is continuous in Ψ^*, $0 < f(\pi, p, w) < B_f$ for all $(\pi, p, w) \in \Psi^*$. (ii) For all (π, p, w) and $(\pi', p, w) \in \Psi^*$, we have $|f(\pi', p, w) - f(\pi, p, w)| \leq m_f |\pi' - \pi|$ for some $0 < m_f < \infty$. (iii) For all (p, w) such that $\mathcal{P}(P \leq p, W \geq w) > 0$ and for all $\alpha \in (0, 1]$, $f(\pi_\alpha(p, w)|C_{p,w}) > 0$, where $f(\cdot|C_{p,w})$ is the derivative of $F(\cdot|C_{p,w})$.*

Assumption 5. *Given p, w, for all $\pi \in (0, B_\pi)$,*

5A *Fix k, there exist functions $H_{2k}(\pi, p, w) > 0$ and $\varepsilon_{2k}(\pi, p, w) > 0$ such that*

$$|\Delta_h^{2k} F_f(\pi, p, w)| \leq H_{2k}(\pi, p, w) h^{2k}$$

for all $|h| \leq \varepsilon_{2k}(\pi, p, w)$. Here, $F_f(\pi, p, w) = \int_{-\infty}^{\pi} f(\gamma, p, w) d\gamma$ and $\Delta_h^{2k} F_f(\pi, p, w) = \sum_{s=-k}^{k} c_{k,s} F_f(\pi + sh, p, w)$ with $c_{k,s} = (-1)^{s+k} C_{2k}^{s+k}$.

5B *f is continuously differentiable with respect to π. $f^{(1)}(\pi, p, w) < \infty$, where $f^{(1)}(\pi, p, w)$ represent the first order derivative of f with respect to π.*

[1] Verification of these properties results from writing $M_k(\cdot)$ as a linear combination of K, i.e., $M_k(x) = \sum_{s=1}^{k} (\lambda_{ks}/s)(K(x/s) + K(-(x/s)))$ where $\lambda_{ks} = -(c_{k,s}/c_{k,0})$.

[2] Here, for any two sets $A \subseteq D_{p,w} := [0, p] \times [w, \infty)$ and $B \subseteq [0, \pi(p, w)]$, Define $\pi(A) = \{\pi(p, w) : (p, w) \in A\}$ and $\pi^{-1}(B) = \{(p, w) \in D_{p,w} : \pi(p, w) \in B\}$.

Assumption 5A imposes an order $2k$ Lipschitz condition on $F_f(\pi, p, w)$ with respect to π. From the proof of Theorem 1 in Mynbaev and Martins-Filho (2010) we know that boundedness of $F_f^{(2k)}(\pi, p, w)$ 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 specific structure for the asymptotic bias and variance by using a Taylor expansion.

3.2 Asymptotic Properties

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

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

The next theorem establishes consistency of $\pi_{\alpha,n}$.

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 $\min_{\{i:P_i\leq p, W_i\geq w\}} \Pi_i \geq h_n B_M$. Under Assumption 1-4 along with Assumption 5A (or 5B), if $H_{2k}(\pi, p, w)$, $F_f(\pi, p, w)$ and $\varepsilon_{2k}(\pi, p, w)$ are bounded for all $(\pi, p, w) \in \Psi^*$, we have*

$$\pi_{\alpha,n}(p, w) - \pi_\alpha(p, w) = o_p(1). \tag{9}$$

The next theorem shows that under suitable normalization and centering $\pi_{\alpha,n}(p, w)$ 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 $\min_{\{i:P_i\leq p, W_i\geq w\}} \Pi_i \geq h_n B_M$. Then,* *(i) Under Assumption 1-4 and Assumption 5B, we have*

$$v_n(p,w)^{-1}\sqrt{n}(\pi_{\alpha,n}(p,w)-\pi_\alpha(p,w)-B_n(p,w)) \xrightarrow{d} N(0,1)$$

where

$$B_n(p,w) = -\frac{1}{2}h_n^2\sigma_M^2\frac{\int_{\pi^{-1}([\pi_\alpha(p,w),\pi(p,w)])} f^{(1)}(\pi_\alpha(p,w),P,W)d(P,W)}{P_{PW}(p,w)f(\pi_\alpha(p,w)|C_{p,w})} + o(h_n^2),$$

$$\begin{aligned} v_n^2(p,w) &= \frac{1}{(P_{PW}(p,w)f(\pi_\alpha(p,w)|C_{p,w}))^2}(F(\pi_\alpha(p,w),p,w) - \frac{F^2(\pi_\alpha(p,w),p,w)}{P_{PW}(p,w)} \\ &\quad -2h_n\sigma_\kappa\int_{\pi^{-1}([\pi_\alpha(p,w),\pi(p,w)])} f(\pi_\alpha(p,w),P,W)d(P,W)) + o(h_n), \end{aligned}$$

with $\sigma_\kappa = \int \gamma\kappa_M(\gamma)M_k(\gamma)d\gamma$*, and* $f^{(1)}(\pi,P,W)$ *denotes the first derivative of* f *with respect to* π*.*

(ii) Under Assumption 1-4 and Assumption 5A, we have

$$|B_n(p,w)| \le ch_n^{2k}\left[\int_{D_{p,w}} H_{2k}(\pi_\alpha(p,w),P,W)d(P,W) + \int_{D_{p,w}}\sup_{\pi\in\mathbb{R}}|F_f(\pi,P,W)|\varepsilon_{2k}^{-2k}(\pi_\alpha(p,w),P,W)d(P,W)\right]$$

where c represent an arbitrary nonnegative constant.

Part (i) of Theorem 2 shows the explicit structure for bias and variance when $k=1$. Part (ii) shows that the bias decays to zero faster when we impose a stronger Lipschitz smoothness condition on the distribution function and increase the value of parameter k accordingly. From Theorem 2, we first observe that our estimator is $\sqrt{n}$ asymptotically normal although it is based on kernel smoothing. In another word, the convergence speed of our estimator is independent of the dimensionality of the problem. Therefore, our estimator does not suffer the "curse of dimensionality". Second, note that the extra smoothness of our estimator provides a smaller variance compared to the empirical estimator at the cost of introducing a bias which vanishes asymptotically (See Aragon et al. (2005)). Finally, the order of the bias term is controlled by the smoothness assumptions on the density function. Note that under appropriate assumptions, the bias term is smaller than the order h_n^{2k}. Hence we can reduce the bias by increase the parameter k. For our estimator, the "smoother" the 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

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

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 $\beta = 1/3$. 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 $N_{p,w} = \sum_{i=1}^{n} I(P_i \leq p, W_i \geq w)$. For $j = 1, ..., N_{p,w}$, get the order statistic of the observation $\Pi_{(i_j)}$ such that $\Pi_{(i_1)} \leq \Pi_{(i_2)} \leq ... \leq$

$\Pi_{(i_{N_{p,w}})}$. The empirical conditional distribution $\hat{F}_e(\pi|C_{p,w})$ is

$$\begin{aligned}\hat{F}_e(\pi|C_{p,w}) &= \frac{\sum_{j=1}^{N_{p,w}} I(\Pi_{(i_j)} \le \pi)}{N_{p,w}} \\ &= \begin{cases} 0 & \text{if } \pi < \Pi_{(i_1)} \\ m/N_{p,w} & \text{if } \Pi_{(i_m)} \le \pi < \Pi_{(i_{m+1})},\ 1 \le m \le N_{p,w}-1 \\ 1 & \text{if } \pi \ge \Pi_{(i_{N_{p,w}})} \end{cases}\end{aligned}$$

Thus the empirical estimator for the conditional quantile $\pi_\alpha(p,w)$ can be computed as follows

$$\hat{\pi}_{e,\alpha}(p,w) = \begin{cases} \Pi_{(i_{\{\alpha N_{p,w}\}})} & \text{if } \alpha N_{p,w} \in \mathbb{N} \\ \Pi_{(i_{\{[\alpha N_{p,w}]+1\}})} & \text{otherwise} \end{cases}$$

where $[\alpha N_{p,w}]$ denotes the integer part of $\alpha N_{p,w}$.

The implementation of our estimator requires choices of kernel function as well as bandwidth. We use the Epanechnikov function $K(x) = \frac{3}{4}(1-x^2)I(|x| \le 1)$ as the seed kernel. It is easy to show this kernel function satisfies Assumption 2. The bandwidth is chosen by minimizing the asymptotic approximation of our estimator's mean integrated squared error (AMISE). For $k=1$, we get the global optimal bandwidth with respect to α as

$$h_n^* = \left(\frac{2\sigma_\kappa \int_0^1 \frac{I_2(p,w,\alpha)}{f^2(\pi_\alpha(p,w)|p,w)} d\alpha}{(\sigma_K^2)^2 \int_0^1 \frac{I_1^2(p,w,\alpha)}{f^2(\pi_\alpha(p,w)|p,w)} d\alpha} \right)^{1/3} n^{-1/3},$$

where

$$\begin{aligned} I_1(p,w,\alpha) &= \int_{\pi^{-1}([\pi_\alpha(p,w),\pi(p,w)])} f^{(1)}(\pi_\alpha(p,w),P,W)d(P,W), \text{ and} \\ I_2(p,w,\alpha) &= \int_{\pi^{-1}([\pi_\alpha(p,w),\pi(p,w)])} f(\pi_\alpha(p,w),P,W)d(P,W). \end{aligned}$$

In our simulations, since we know the true distribution, we can compute h_n^* directly. In practice, use of h_n^* requires the estimation of the unknown distribution. Applying a similar method described 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

	|Bias|			RMSE		
n=200	Kernel	Kernel	Empirical	Kernel	Kernel	Empirical
α	k=1	k=2		k=1	k=2	
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	Kernel	Empirical	Kernel	Kernel	Empirical
α	k=1	k=2		k=1	k=2	
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 $\alpha = 0.25, 0.5$ 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 $\alpha = 0.5$, 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 α. This is consistent with the result in Theorem 2 which states the bias decays faster as k increases.

We also observe that as α 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 α grows. As a result, when α is close to 1, profit functions of order α become more difficult to estimate. Note that the performance of our smoothed estimator is especially poor when $\alpha = 0.99$. 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 α can outperform the empirical estimator in most cases as long as α 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 α 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 α frontier with $\alpha \in (0, 1)$ 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 $\alpha \in [0, 1]$. 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 $(\pi, p, w) \in \Psi^*$, if $\pi < \pi_\alpha(p, w) = \inf\{\pi \in [0, B_\pi] : F(\pi|C_{p,w}) \geq \alpha\}$, then $\pi \notin \{\pi \in [0, B_\pi] : F(\pi|C_{p,w}) \geq \alpha\}$. That is, $F(\pi|C_{p,w}) < \alpha$. If $\pi > \pi_\alpha(p, w)$, there exist some $\varepsilon > 0$ such that $\pi > \pi_\alpha(p, w) + \varepsilon$. By the definition of $\pi_\alpha(p, w)$, for any $\varepsilon > 0$, there exist some $\pi_0 \in \{\pi \in [0, B_\pi] : F(\pi|C_{p,w}) \geq \alpha\}$ such that $\pi_0 < \pi_\alpha(p, w) + \varepsilon$. By the strict monotonicity of $F(\cdot|C_{p,w})$, $F(\pi|C_{p,w}) > F(\pi_\alpha(p, w) + \varepsilon|C_{p,w}) > F(\pi_0|C_{p,w}) \geq \alpha$. The result then follows. □

Proposition 2 *Proof.* (i) Since $\{\pi_\alpha(p, w)\}_{0\leq\alpha\leq1}$ is monotone nondecreasing in α, and $\sup_{0\leq\alpha\leq1}\{\pi_\alpha(p, w)\} = \pi(p, w)$.The result then follows. (ii) Let Φ be a compact set interior to the support of (P, W). Define $\phi_n(p, w) = \pi_{1-\frac{1}{n}}(p, w)$. Since $\{\pi_\alpha(p, w)\}_{0\leq\alpha\leq1}$ is monotone nondecreasing in α, for any $n \in \mathbb{N}$, $\phi_n(p, w) \leq \phi_{n+1}(p, w)$. From (i), $\lim_{n\to\infty} \phi_n(p, w) = \pi(p, w)$ pointwise. By Dini's Theorem, $\sup_{(p,w)\in\Phi} |\phi_n(p, w) - \pi(p, w)| \to 0$. Thus, for any $\varepsilon > 0$, there exist some N such that when $n > N$, $\sup_{(p,w)\in\Phi} |\phi_n(p, w) - \pi(p, w)| < \varepsilon$. That is, there exist $\delta = 1 - \frac{1}{N}$ such that when $|\alpha - 1| < \delta$, $\sup_{(p,w)\in\Phi} |\pi_\alpha(p, w) - \pi(p, w)| < \varepsilon$. □

Proposition 3 *Proof.* (i) First, note that by definition when $\pi = 0$ we have $\hat{F}(\pi|C_{p,w}) = 0$. If $0 < \pi_1 \leq \pi_2$, we only need to prove $\hat{P}(\pi_2, p, w) - \hat{P}(\pi_1, p, w) \geq 0$, since the denominator does not depend on π. By (5),

$$\begin{aligned} \hat{P}(\pi_2, p, w) - \hat{P}(\pi_1, p, w) &= (nh_n)^{-1} \sum_{i=1}^{n} (\int_0^{\pi_2} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma - \int_0^{\pi_1} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma) I(P_i \leq p, W_i \geq w) \\ &\geq 0 \end{aligned}$$

since M_k is a symmetric density. (ii) For any $\pi_0 \in [0, B_\pi]$, let $\pi - \pi_0 < \delta$ for some $\delta > 0$. Then,

$$\begin{aligned} |\hat{P}(\pi, p, w) - \hat{P}(\pi_0, p, w)| &= (nh_n)^{-1} \sum_{i=1}^{n} (\int_0^{\pi} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma - \int_0^{\pi_0} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma) I(P_i \leq p, W_i \geq w) \\ &\leq (nh_n)^{-1} \sum_{i=1}^{n} (\int_0^{\pi_0+\delta} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma - \int_0^{\pi_0} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma) I(P_i \leq p, W_i \geq w) \\ &= (nh_n)^{-1} \sum_{i=1}^{n} (\int_{\pi_0}^{\pi_0+\delta} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma) I(P_i \leq p, W_i \geq w) \\ &\leq {h_n}^{-1}\delta \cdot \sup_{\varphi\in[-B_M, B_M]} M_k(\varphi) < \varepsilon \end{aligned}$$

where the last inequality follows for any $\epsilon > 0$, since δ 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 $n > N(p, w)$, $h_n{}^{-1} \lim_{\pi\to\infty} \int_0^\pi M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma = 1$. Now, note that

$$h_n{}^{-1} \lim_{\pi\to\infty} \int_0^\pi M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma = \lim_{\pi\to\infty} \int_{-\frac{\Pi_i}{h_n}}^{\frac{\pi-\Pi_i}{h_n}} M_k(\varphi)d\varphi.$$

Since $h_n \to 0$ as $n \to \infty$, there exists $N(p, w)$ for any (p, w), such that for all $n > N(p, w)$, $-\frac{\Pi_i}{h_n} < B_M$ and $\frac{\pi-\Pi_i}{h_n} > B_M$. The result follows from the fact that M_k is a density. □

Let $c > 0$ represent an arbitrary constant and $D_{p,w} = [0, p] \times [w, \infty)$. By the definition of $\pi(p, w)$, for any point $(p', w') \in D_{p,w}$, we have $\pi(p', w') \leq \pi(p, w)$. Therefore, we can write $D_{p,w} = \pi^{-1}([0, \pi(p, w)])$. Denote $P(\pi, p, w) = \mathcal{P}(\Pi \leq \pi, P \leq p, W \geq w)$ and $\hat{P}(\pi, p, w)$ as defined in (5). $\sigma_\kappa = \int_{-B_M}^{B_M} M_k(\gamma)\gamma\kappa_M(\gamma)d\gamma$. In order to prove Theorems 1 and 2, we need the following lemmas.

Lemma 1. *Under Assumption 1-4 and Assumption 5A, we have: (a)*

$$|E(\hat{P}(\pi, p, w)) - P(\pi, p, w)| \leq ch_n^{2k} \left[\int_{D_{p,w}} H_{2k}(\pi, P, W)d(P, W) + \int_{D_{p,w}} \sup_{\pi\in\mathbb{R}} |F_f(\pi, P, W)| \varepsilon_{2k}^{-2k}(\pi, P, W)d(P, W) \right].$$

If we assume furthermore $H_{2k}(\pi, p, w)$, $F_f(\pi, p, w)$ and $\varepsilon_{2k}(\pi, p, w)$ are bounded for all $(\pi, p, w) \in \Psi^*$, we have $|E(\hat{P}(\pi, p, w)) - P(\pi, p, w)| = O(h_n^{2k})$ and (b)

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

where both $|R_{1k}(\pi, p, w, h)|$ and $|R_{2k}(\pi, p, w, h)|$ satisfy

$$|R_{ik}(\pi, p, w, h)| \leq ch_n^{2k}[\int_{D_{p,w}} H_{2k}(\pi, P, W)d(P, W) + \int_{D_{p,w}} \sup_{\pi\in\mathbb{R}} |F_f(\pi, P, W)| \varepsilon_{2k}^{-2k}(\pi, P, W)d(P, W)].$$

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

$$\begin{aligned}
E(\hat{P}(\pi,p,w)) &= E[(nh_n)^{-1}\sum_{i=1}^{n}(\int_0^{\pi} M_k(\frac{\Pi_i-\gamma}{h_n})d\gamma)I(P_i \le p, W_i \ge w)] \\
&= {h_n}^{-1}\int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}\int_0^{\pi} M_k(\frac{\Pi-\gamma}{h_n})d\gamma I(P \le p, W \ge w)f(\Pi,P,W)d\Pi d(P,W) \\
&= \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}\int_{-B_M}^{\frac{\pi-\Pi_i}{h_n}} M_k(\varphi)d\varphi I(P \le p, W \ge w)f(\Pi,P,W)d\Pi d(P,W) \\
&= \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}\kappa_M(\frac{\pi-\Pi}{h_n})I(P \le p, W \ge w)f(\Pi,P,W)d\Pi d(P,W)
\end{aligned}$$

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

$$\begin{aligned}
&\int_{-\infty}^{\infty}\kappa_M(\frac{\pi-\Pi}{h_n})I(P \le p, W \ge w)f(\Pi,P,W)d\Pi \\
=\;& h_n\int_{-\infty}^{\infty}\kappa_M(\varphi)I(P \le p, W \ge w)f(\pi-h_n\varphi,P,W)d\varphi \\
=\;& -\int_{-\infty}^{\infty}\kappa_M(\varphi)I(P \le p, W \ge w)dF_f(\pi-h_n\varphi,P,W) \\
=\;& -\kappa_M(\varphi)I(P \le p, W \ge w)F_f(\pi-h_n\varphi,P,W)|_{\varphi=-\infty}^{\varphi=\infty} + \int_{-\infty}^{\infty}F_f(\pi-h_n\varphi,P,W)I(P \le p, W \ge w)d\kappa_M(\varphi) \\
=\;& 0 + \int_{-\infty}^{\infty}M_k(\varphi)F_f(\pi-h_n\varphi,P,W)I(P \le p, W \ge w)d\varphi \\
=\;& -\frac{1}{c_{k,0}}\int_{-\infty}^{\infty}\sum_{|s|=1}^{k}\frac{c_{k,s}}{|s|}K(\varphi/s)F_f(\pi-h_n\varphi,P,W)I(P \le p, W \ge w)d\varphi \\
=\;& -\frac{1}{c_{k,0}}\int_{-\infty}^{\infty}K(t)\sum_{|s|=1}^{k}c_{k,s}F_f(\pi-sh_nt,P,W)I(P \le p, W \ge w)dt
\end{aligned}$$

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

$$
\begin{aligned}
& E(\hat{P}(\pi,p,w)) - P(\pi,p,w) \\
= & -\frac{1}{c_{k,0}} \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}} \int_{-\infty}^{\infty} K(t) \sum_{|s|=1}^{k} c_{k,s} F_f(\pi - sh_n t, P, W) I(P \le p, W \ge w) dt d(P,W) \\
- & \int_{D_{p,w}} \int_{-\infty}^{\pi} f(\gamma, P, W) d\gamma d(P,W) \\
= & -\frac{1}{c_{k,0}} \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}} \int_{-\infty}^{\infty} K(t) \sum_{|s|=1}^{k} c_{k,s} F_f(\pi - sh_n t, P, W) I(P \le p, W \ge w) dt d(P,W) \\
+ & \sum_{|s|=1}^{k} \frac{c_{k,s}}{c_{k,0}} \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}} \int_{-\infty}^{\infty} K(t) dt F_f(\pi, P, W) I(P \le p, W \ge w) d(P,W) \\
= & -\frac{1}{c_{k,0}} \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}} \int_{-\infty}^{\infty} K(t) \Delta_{h_n t}^{2k} F_f(\pi, P, W) I(P \le p, W \ge w) dt d(P,W)
\end{aligned}
$$

By Assumption 5A, we have

$$
\begin{aligned}
|E(\hat{P}(\pi,p,w)) - P(\pi,p,w)| \le & \; c \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}} \int_{-\infty}^{\infty} |K(t) \Delta_{h_n t}^{2k} F_f(\pi, P, W)| I(P \le p, W \ge w) dt d(P,W) \\
\le & \; c \int_{D_{p,w}} \Big(\int_{|h_n t| \le \varepsilon_{2k}(\pi,P,W)} + \int_{|h_n t| > \varepsilon_{2k}(\pi,P,W)} \Big) |K(t) \Delta_{h_n t}^{2k} F_f(\pi, P, W)| dt d(P,W) \\
\le & \; c \Big[\int_{D_{p,w}} \int_{|h_n t| \le \varepsilon_{2k}(\pi,P,W)} |K(t)| (h_n t)^{2k} dt H_{2k}(\pi, P, W) d(P,W) \\
& + \int_{D_{p,w}} \sup_{\pi\in\mathbb{R}} |F_f(\pi, P, W)| \int_{|h_n t| > \varepsilon_{2k}(\pi,P,W)} |K(t)| dt d(P,W) \Big]
\end{aligned}
$$

Since for any $N > 0$,

$$
\int_{|t|>N} |K(t)| dt \le \int_{|t|>N} |K(t)| |\frac{t}{N}|^{2k} dt \le N^{-2k} \int_{-\infty}^{\infty} |K(t)| t^{2k} dt
$$

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

$$
\begin{aligned}
& |E(\hat{P}(\pi,p,w)) - P(\pi,p,w)| \\
\le & \; c h_n^{2k} \Big[\int_{D_{p,w}} H_{2k}(\pi, P, W) d(P,W) + \int_{D_{p,w}} \sup_{\pi\in\mathbb{R}} |F_f(\pi, P, W)| \varepsilon_{2k}^{-2k}(\pi, P, W) d(P,W) \Big]
\end{aligned}
$$

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

$$\begin{aligned} V_{1n} &= E[{h_n}^{-2}(\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma)^2 I(P_i \le p, W_i \ge w)] \\ V_{2n} &= (E[h_n^{-1}\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma I(P_i \le p, W_i \ge w)])^2 \end{aligned}$$

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

$$\begin{aligned} V_{1n} &= E[{h_n}^{-2}(\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma)^2 I(P_i \le p, W_i \ge w)] \\ &= {h_n}^{-2}\int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}(\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma)^2 f(\Pi,P,W)I(P\le p, W\ge w)d\Pi d(P,W) \\ &= \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}(\int_{-B_M}^{\frac{\pi-\Pi}{h_n}} M_k(\varphi)d\varphi)^2 f(\Pi,P,W)I(P\le p, W\ge w)d\Pi d(P,W) \\ &= \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}(\kappa_M(\frac{\pi-\Pi}{h_n}))^2 f(\Pi,P,W)I(P\le p, W\ge w)d\Pi d(P,W). \end{aligned}$$

Integrating by parts

$$\begin{aligned} &\int_{-\infty}^{\infty}(\kappa_M(\frac{\pi-\Pi}{h_n}))^2 f(\Pi,P,W)I(P\le p, W\ge w)d\Pi \\ =\ & h_n\int_{-\infty}^{\infty}(\kappa_M(\varphi))^2 f(\pi-h_n\varphi,P,W)I(P\le p, W\ge w)d\varphi \\ =\ & -\int_{-\infty}^{\infty}(\kappa_M(\varphi))^2 I(P\le p, W\ge w)dF_f(\pi-h_n\varphi,P,W) \\ =\ & -(\kappa_M(\varphi))^2 I(P\le p, W\ge w)F_f(\pi-h_n\varphi,P,W)|_{\varphi=-\infty}^{\varphi=\infty} \\ +\ & \int_{-\infty}^{\infty}F_f(\pi-h_n\varphi,P,W)I(P\le p, W\ge w)d(\kappa_M(\varphi))^2 \\ =\ & 0 + 2\int_{-\infty}^{\infty}\kappa_M(\varphi)M_k(\varphi)F_f(\pi-h_n\varphi,P,W)I(P\le p, W\ge w)d\varphi \\ =\ & -\frac{2}{c_{k,0}}\int_{-\infty}^{\infty}\sum_{|s|=1}^{k}\frac{c_{k,s}}{|s|}\kappa_M(\varphi)K(\varphi/s)F_f(\pi-h_n\varphi,P,W)I(P\le p, W\ge w)d\varphi \\ =\ & -\frac{2}{c_{k,0}}\int_{-\infty}^{\infty}K(t)\sum_{|s|=1}^{k}\kappa_M(st)c_{k,s}F_f(\pi-sh_nt,P,W)I(P\le p, W\ge w)dt \end{aligned}$$

Note that

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

Thus,

$$V_{1n} = -\frac{2}{c_{k,0}} \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}} \int_{-\infty}^{\infty} K(t)\kappa(t) \sum_{|s|=1}^{k} c_{k,s} F_f(\pi - sh_n t, P, W) I(P \le p, W \ge w) dtd(P,W)$$

Again, integrating by parts,

$$\int_{-\infty}^{\infty} K(t)\kappa(t)dt = 1/2,$$

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

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

where $|R_{1k}(\pi,p,w,h)| \le ch_n^{2k}[\int_{D_{p,w}} H_{2k}(\pi,P,W)d(P,W) + \int_{D_{p,w}} \sup_{\pi\in\mathbb{R}} |F_f(\pi,P,W)|\varepsilon_{2k}^{-2k}(\pi,P,W)d(P,W)]$. From part (a),

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

where $|R_{2k}(\pi,p,w,h)| \le ch_n^{2k}[\int_{D_{p,w}} H_{2k}(\pi,P,W)d(P,W) + \int_{D_{p,w}} \sup_{\pi\in\mathbb{R}} |F_f(\pi,P,W)|\varepsilon_{2k}^{-2k}(\pi,P,W)d(P,W)]$. As a result,

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

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

$$ch_n^{2k}[\int_{D_{p,w}} H_{2k}(\pi,P,W)d(P,W) + \int_{D_{p,w}} \sup_{\pi\in\mathbb{R}} |F_f(\pi,P,W)|\varepsilon_{2k}^{-2k}(\pi,P,W)d(P,W)].$$

$\square$

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 π, 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)*

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

(b)

$$V(\hat{P}(\pi,p,w)) = \begin{cases} n^{-1}P(\pi,p,w)(1-P(\pi,p,w)) - 2n^{-1}h_n\sigma_\kappa \int_{\pi^{-1}((\pi,\pi(p,w)])} f(\pi,P,W)d(P,W) + o(h_n/n) \\ \qquad \text{if } 0 < \pi < \pi(p,w) \\ n^{-1}P(\pi,p,w)(1-P(\pi,p,w)) + o(h_n/n) \qquad \text{if } \pi \geq \pi(p,w) \end{cases}$$

where $P(\pi,p,w) = \mathcal{P}(\Pi \leq \pi, P \leq p, W \geq w)$ and $\hat{P}(\pi,p,w)$ is defined in (5). $\sigma_\kappa = \int_{-B_M}^{B_M} M_k(\gamma)\gamma\kappa_M(\gamma)d\gamma$.

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

$$\begin{aligned} E(\hat{P}(\pi,p,w)) &= E[(nh_n)^{-1}\sum_{i=1}^{n}(\int_0^\pi M_k(\frac{\Pi_i-\gamma}{h_n})d\gamma)I(P_i \leq p, W_i \geq w)] \\ &= h_n^{-1}\int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}}\int_{-\infty}^{\infty}\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma I(P \leq p, W \geq w)f(\Pi,P,W)d\Pi d(P,W) \\ &= h_n^{-1}\int_{D_{p,w}}\int_{[0,\pi(P,W)]}\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma f(\Pi,P,W)d\Pi d(P,W) \\ &= \int_{D_{p,w}}\int_{[0,\pi(P,W)]}\int_{-B_M}^{\frac{\pi-\Pi_i}{h_n}} M_k(\varphi)d\varphi f(\Pi,P,W)d\Pi d(P,W) \\ &= \int_{D_{p,w}}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W). \end{aligned}$$

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

For case (1),

$$\begin{aligned}
E(\hat{P}(\pi,p,w)) &= \int_{D_{p,w}}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\\
&= \int_{\pi^{-1}([0,\pi(p,w)])}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\\
&= \int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\\
&\quad+\int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\\
&= A_{1n}+A_{2n}.
\end{aligned}$$

Note that for the last term, for $\Pi<\pi$, $\kappa_M(\frac{\pi-\Pi}{h_n})\to 1$ as $n\to\infty$. By Assumptions 2 and 4, $|\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)|<\infty$. By Lebesgue's dominated convergence theorem,

$$\begin{aligned}
A_{2n} &= \int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\\
&= \int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi)}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\\
&\to \int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi]}f(\Pi,P,W)d\Pi d(P,W)
\end{aligned}$$

Now,

$$A_{1n}=\int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})\frac{\partial F_f(\Pi,P,W)}{\partial\Pi}d\Pi d(P,W)$$

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

$$\begin{aligned}
&\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})\frac{\partial F_f(\Pi,P,W)}{\partial\Pi}d\Pi\\
=& \int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})dF_f(\Pi,P,W)\\
=& \kappa_M(\frac{\pi-\Pi}{h_n})dF_f(\Pi,P,W)|_{\Pi=0}^{\Pi=\pi(P,W)}-\int_{[0,\pi(P,W)]}F_f(\Pi,P,W)d\kappa_M(\frac{\pi-\Pi}{h_n})\\
=& \kappa_M(\frac{\pi-\pi(P,W)}{h_n})F_f(\pi(P,W),P,W)+\frac{1}{h_n}\int_{[0,\pi(P,W)]}F_f(\Pi,P,W)M_k(\frac{\pi-\Pi}{h_n})d\Pi\\
=& \kappa_M(\frac{\pi-\pi(P,W)}{h_n})F_f(\pi(P,W),P,W)+\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}}F_f(\pi-h_n\gamma,P,W)M_k(\gamma)d\gamma
\end{aligned}$$

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

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

where

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

Now,

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

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

$$E_{11,n} \to \int_{\pi^{-1}([0,\pi])} F_f(\pi(P,W), P, W)d(P,W) = \int_{\pi^{-1}([0,\pi])} \int_{[0,\pi(P,W)]} f(\Pi, P, W)d\Pi d(P,W).$$

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

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

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

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

$$\begin{aligned}
h_n^{-1} E_{3n} &= \int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])} f(\pi, P, W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma d\gamma d(P, W) \\
&= \int_{\pi^{-1}([0,\pi))} f(\pi, P, W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma d\gamma d(P, W) \\
&\quad + \int_{\pi^{-1}((\pi,\pi(p,w)])} f(\pi, P, W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma d\gamma d(P, W) \\
&= E_{31,n} + E_{32,n}
\end{aligned}$$

For $E_{31,n}$, when $(P, W) \in \pi^{-1}([0, \pi))$, $\frac{\pi-\pi(P,W)}{h_n} \to +\infty$ and $\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma d\gamma \to 0$ as $n \to \infty$. For $E_{32,n}$, when $(P, W) \in \pi^{-1}((\pi, \pi(p, w)])$, $\frac{\pi-\pi(P,W)}{h_n} \to -\infty$ and $\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma d\gamma \to 0$ as $n \to \infty$ by the symmetry of $M_k(.)$. As a result, $h_n^{-1} E_{3n} \to 0$.

$$\begin{aligned}
h_n^{-2} E_{4n} &= \frac{1}{2} \int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])} f^{(1)}(\pi, P, W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma^2 d\gamma d(P, W) \\
&= \frac{1}{2} \int_{\pi^{-1}([0,\pi])} f^{(1)}(\pi, P, W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma^2 d\gamma d(P, W) \\
&\quad + \frac{1}{2} \int_{\pi^{-1}([\pi,\pi(p,w)])} f^{(1)}(\pi, P, W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma) \gamma^2 d\gamma d(P, W) \\
&= E_{41,n} + E_{42,n}
\end{aligned}$$

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

$$\begin{aligned}
E(\hat{P}(\pi,p,w)) &= E_{1n}+E_{2n}+E_{3n}+E_{4n}+A_{2n}+o(h_n^2) \\
&= \int_{\pi^{-1}([0,\pi))}\int_{[0,\pi]} f(\Pi,P,W)d\Pi d(P,W) + \int_{\pi^{-1}((\pi,\pi(p,w)])}\int_{[0,\pi]} f(\Pi,P,W)d\Pi d(P,W) \\
&\quad + \int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi]} f(\Pi,P,W)d\Pi d(P,W) \\
&\quad + \frac{1}{2}h_n^2\sigma_M^2 \int_{\pi^{-1}((\pi,\pi(p,w)])} f^{(1)}(\pi,P,W)d(P,W) + o(h_n^2) \\
&= P(\pi,p,w) + \frac{1}{2}h_n^2\sigma_M^2 \int_{\pi^{-1}((\pi,\pi(p,w)])} f^{(1)}(\pi,P,W)d(P,W) + o(h_n^2) \\
&= F(\pi,p,w) - F_{\Pi P}(\pi,p) + \frac{1}{2}h_n^2\sigma_M^2 \int_{\pi^{-1}((\pi,\pi(p,w)])} f^{(1)}(\pi,P,W)d(P,W) + o(h_n^2)
\end{aligned}$$

For case (2), when $\pi > \pi(p,w)$, $\pi > \pi(P,W)$ for all $(P,W) \in D_{p,w}$, $\kappa_M(\frac{\pi-\pi(P,W)}{h_n}) \to 1$, $\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)d\gamma \to 0$, $\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)\gamma d\gamma \to 0$ and $\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)\gamma^2 \to 0$. By LDC,

$$\begin{aligned}
\int_{D_{p,w}} \kappa_M(\frac{\pi-\pi(P,W)}{h_n})F_f(\pi(P,W),P,W)d(P,W) &\to \int_{D_{p,w}} F_f(\pi(P,W),P,W)d(P,W) = P(\pi,p,w) \\
\int_{D_{p,w}} F_f(\pi,P,W)\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)d\gamma d(P,W) &\to 0 \\
h_n\int_{D_{p,w}} f(\pi,P,W)\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)\gamma d\gamma d(P,W) &\to 0 \\
\frac{1}{2}h_n^2\int_{D_{p,w}} f^{(1)}(\pi,P,W)\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)\gamma^2 d\gamma d(P,W) &\to 0
\end{aligned}$$

Therefore, if $\pi > \pi(p,w)$, $E(\hat{P}(\pi,p,w)) = P(\pi,p,w) + o(h_n^2)$.

For case (3), the proof willl be the same as case (1) except that the set $(\pi,\pi(p,w)] = \phi$.

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

$$\begin{aligned} V_{1n} &= E[{h_n}^{-2}(\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma)^2 I(P_i \le p, W_i \ge w)] \\ V_{2n} &= (E[h_n^{-1}\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma I(P_i \le p, W_i \ge w)])^2 \end{aligned}$$

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

$$\begin{aligned} V_{1n} &= E[{h_n}^{-2}(\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma)^2 I(P_i \le p, W_i \ge w)] \\ &= {h_n}^{-2}\int_{D_{p,w}}\int_{[0,\pi(P,W)]}(\int_0^\pi M_k(\frac{\Pi-\gamma}{h_n})d\gamma)^2 f(\Pi,P,W)d\Pi d(P,W) \\ &= \int_{D_{p,w}}\int_{[0,\pi(P,W)]}(\int_{-B_M}^{\frac{\pi-\Pi}{h_n}} M_k(\varphi)d\varphi)^2 f(\Pi,P,W)d\Pi d(P,W) \\ &= \int_{D_{p,w}}\int_{[0,\pi(P,W)]}(\kappa_M(\frac{\pi-\Pi}{h_n}))^2 f(\Pi,P,W)d\Pi d(P,W) \end{aligned}$$

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

For case (1),

$$\begin{aligned} V_{1n} &= \int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])}\int_{[0,\pi(P,W)]}(\kappa_M(\frac{\pi-\Pi}{h_n}))^2 f(\Pi,P,W)d\Pi d(P,W) \\ &\quad + \int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi]}(\kappa_M(\frac{\pi-\Pi}{h_n}))^2 f(\Pi,P,W)d\Pi d(P,W) \\ &= \tilde{A}_{1n} + \tilde{A}_{2n}. \end{aligned}$$

Note that for the last term, for $\Pi < \pi$, $\kappa_M(\frac{\pi-\Pi}{h_n}) \to 1$ as $n \to \infty$. By assumptions 2 and 3, $|\kappa_M(\frac{\pi-\Pi}{h_n})^2 f(\Pi,P,W)| < \infty$. By Lebesgue's dominated convergence theorem,

$$\tilde{A}_{2n} \to \int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi]} f(\Pi,P,W)d\Pi d(P,W)$$

Now,

$$\tilde{A}_{1n} = \int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])}\int_{[0,\pi(P,W)]}(\kappa_M(\frac{\pi-\Pi}{h_n}))^2 \frac{\partial F_f(\Pi,P,W)}{\partial \Pi} d\Pi d(P,W)$$

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

$$\begin{aligned}
& \int_{[0,\pi(P,W)]} (\kappa_M(\frac{\pi-\Pi}{h_n}))^2 \frac{\partial F_f(\Pi,P,W)}{\partial \Pi} d\Pi \\
= & \int_{[0,\pi(P,W)]} (\kappa_M(\frac{\pi-\Pi}{h_n}))^2 dF_f(\Pi,P,W) \\
= & (\kappa_M(\frac{\pi-\Pi}{h_n}))^2 dF_f(\Pi,P,W)|_{\Pi=0}^{\Pi=\pi(P,W)} - \int_{[0,\pi(P,W)]} F_f(\Pi,P,W) d(\kappa_M(\frac{\pi-\Pi}{h_n}))^2 \\
= & (\kappa_M(\frac{\pi-\pi(P,W)}{h_n}))^2 F_f(\pi(P,W),P,W) + \frac{2}{h_n}\int_{[0,\pi(P,W)]} F_f(\Pi,P,W)\kappa_M(\frac{\pi-\Pi}{h_n}) M_k(\frac{\pi-\Pi}{h_n}) d\Pi \\
= & (\kappa_M(\frac{\pi-\pi(P,W)}{h_n}))^2 F_f(\pi(P,W),P,W) + 2\int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} F_f(\pi-h_n\gamma,P,W)\kappa_M(\gamma)M_k(\gamma)d\gamma
\end{aligned}$$

By Taylor's theorem, $F_f(\pi - h_n\gamma, P, W) = F_f(\pi, P, W) - h_n\gamma f(\pi, P, W) + o(h_n)$, Hence

$$\tilde{A}_{1n} = V_{11n} + V_{12n} + V_{13n} + o(h_n)$$

where

$$\begin{aligned}
V_{11n} &= \int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])} (\kappa_M(\frac{\pi-\pi(P,W)}{h_n}))^2 F_f(\pi(P,W),P,W) d(P,W) \\
V_{12n} &= 2\int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])} F_f(\pi,P,W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)\kappa_M(\gamma)d\gamma d(P,W) \\
V_{13n} &= -2h_n\int_{\pi^{-1}([0,\pi)\cup(\pi,\pi(p,w)])} f(\pi,P,W) \int_{\frac{\pi-\pi(P,W)}{h_n}}^{\frac{\pi}{h_n}} M_k(\gamma)\gamma\kappa_M(\gamma)d\gamma d(P,W)
\end{aligned}$$

Using the same argument as in the proof of part (a),

$$\begin{aligned}
V_{11n} &\to \int_{\pi^{-1}([0,\pi))} \int_{[0,\pi]} f(\Pi,P,W)d\Pi d(P,W) \\
V_{12n} &\to 2\int_{\pi^{-1}((\pi,\pi(p,w)])} F_f(\pi,P,W) \int_{-B_M}^{B_M} M_k(\gamma)\kappa_M(\gamma)d\gamma d(P,W)
\end{aligned}$$

Now,

$$\begin{aligned}\int_{-B_M}^{B_M} M_k(\gamma)\kappa_M(\gamma)d\gamma &= \int_{-B_M}^{B_M} \kappa_M(\gamma)d\kappa_M(\gamma) \\ &= \kappa_M(\gamma)^2|_{-B_M}^{B_M} - \int_{-B_M}^{B_M} \kappa_M(\gamma)d\kappa_M(\gamma) \\ &= 1 - \int_{-B_M}^{B_M} \kappa_M(\gamma)d\kappa_M(\gamma)\end{aligned}$$

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

$$V_{12n} \to \int_{\pi^{-1}((\pi,\pi(p,w)])} \int_{[0,\pi]} f(\Pi, P, W)d\Pi d(P, W)$$

Similarly,

$$V_{13n} \to 2h_n\sigma_\kappa \int_{\pi^{-1}((\pi,\pi(p,w)])} f(\Pi, P, W)d\Pi d(P, W)$$

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

Lemma 3. *Let h_n be a sequence of nonstochastic bandwidths such that $0 < h_n \to 0$ as $n \to \infty$. Given $w \in \mathbb{R}_{++}^{d_2}$, $p \in \mathbb{R}_{+}^{d_1}$ and there exist some $N(p, w)$ such that when $n > N(p, w)$ $\min_{\{i:P_i \le p, W_i \ge w\}} \Pi_i \ge h_n B_M$. Under Assumptions 1-4 along with Assumption 5B (or 5A) and if $H_{2k}(\pi, p, w)$, $F_f(\pi, p, w)$ and $\varepsilon_{2k}(\pi, p, w)$ are bounded for all $(\pi, p, w) \in \Psi^*$, we have (a) $\sup_{\pi\in[0,\pi(p,w)]} |\hat{P}(\pi, p, w) - E(\hat{P}(\pi, p, w))| = o_p(1)$ and (b) $\sup_{\pi\in[0,\pi(p,w)]} |E(\hat{P}(\pi, p, w) - P(\pi, p, w))| = o(1)$.*

Proof. (a): For given $w \in \mathbb{R}_{++}^{d_2}$, $p \in \mathbb{R}_{+}^{d_1}$ there exist some $N(p, w)$ such that when $n > N(p, w)$

$$\min_{\{i:P_i \le p, W_i \ge w\}} \Pi_i \ge h_n B_M.$$

Then

$$\begin{aligned}
\hat{P}(\pi,p,w) &= (nh_n)^{-1}\sum_{i=1}^{n}(\int_0^{\pi} M_k(\frac{\Pi_i-\gamma}{h_n})d\gamma)I(P_i\le p, W_i\ge w)\\
&= {h_n}^{-1}\int_0^{\pi} M_k(\frac{\Pi-\gamma}{h_n})d\gamma I(P_i\le p, W_i\ge w)\\
&= {h_n}^{-1}\int_{-\frac{\Pi_i}{h_n}}^{\frac{\pi-\Pi_i}{h_n}} M_k(\varphi)d\varphi I(P_i\le p, W_i\ge w)\\
&= {h_n}^{-1}\int_{-B_M}^{\frac{\pi-\Pi_i}{h_n}} M_k(\varphi)d\varphi I(P_i\le p, W_i\ge w)\\
&= {h_n}^{-1}\kappa_M(\frac{\pi-\Pi_i}{h_n})I(P_i\le p, W_i\ge w)
\end{aligned}$$

Since $[0,\pi(p,w)]$ is compact, there exist $\pi_0\in[0,\pi(p,w)]$ and r_π such that $[0,\pi(p,w)]\subset B(\pi_0,r_\pi)$ where $B(\pi_0,r_\pi)=\{\pi\in\mathbb{R}:|\pi-\pi_0|<r_\pi\}$. Furthermore, for all $\pi\in[0,\pi(p,w)]$,

$$[0,\pi(p,w)]\subset\cup_{\{\pi:\pi\in[0,\pi(p,w)]\}}B(\pi,(\frac{n}{h_n^a})^{-\frac{1}{2}})$$

with $a>0$. By the Heine-Borel Theorem, there exists $\{B(\pi_l,(\frac{n}{h_n^a})^{-\frac{1}{2}})\}_{l=1}^{L_n}$ such that

$$[0,\pi(p,w)]\subset\cup_{l=1}^{L_n}B(\pi_l,(\frac{n}{h_n^a})^{-\frac{1}{2}})$$

with $L_n<r_\pi(\frac{n}{h_n^a})^{\frac{1}{2}}$. Therefore, any $\pi\in[0,\pi(p,w)]$, there exists some $l\in\{1...L_n\}$, such that $\pi\in B(\pi_l,(\frac{n}{h_n^a})^{-\frac{1}{2}})$. Then we have

$$\begin{aligned}
& |\hat{P}(\pi,p,w)-E(\hat{P}(\pi,p,w))|\\
\le\ & |\hat{P}(\pi,p,w)-\hat{P}(\pi_l,p,w)|+|\hat{P}(\pi_l,p,w)-E(\hat{P}(\pi_l,p,w))|\\
& +|E(\hat{P}(\pi_l,p,w))-E(\hat{P}(\pi,p,w))|\\
=\ & P_{1n}+P_{2n}+P_{3n}
\end{aligned}$$

For $\pi \in B(\pi_l, (\frac{n}{h_n^a})^{-\frac{1}{2}})$, we have

$$\begin{aligned} P_{1n} &= |\hat{P}(\pi, p, w) - \hat{P}(\pi_l, p, w)| \\ &\leq {h_n}^{-1} \sum_{i=1}^{n} |\kappa_M(\frac{\pi - \Pi_i}{h_n}) - \kappa_M(\frac{\pi_l - \Pi_i}{h_n})| I(P_i \leq p, W_i \geq w) \\ &\leq {h_n}^{-1} m_\kappa |\pi - \pi_l| \\ &\leq m_\kappa (n h_n^{2-a})^{-\frac{1}{2}}, \end{aligned}$$

by Assumption 2 and the fact that $I(P_i \leq p, W_i \geq w) \leq 1$. Similarly,

$$\begin{aligned} P_{3n} &= |E(\hat{P}(\pi_l, p, w)) - E(\hat{P}(\pi, p, w))| \\ &= |\int_{D_{p,w}} \int_{[0,\pi(P,W)]} \kappa_M(\frac{\pi_l - \Pi}{h_n}) f(\Pi, P, W) d\Pi d(P, W) \\ &\quad - \int_{D_{p,w}} \int_{[0,\pi(P,W)]} \kappa_M(\frac{\pi - \Pi}{h_n}) f(\Pi, P, W) d\Pi d(P, W)| \\ &\leq \int_{D_{p,w}} \int_{[0,\pi(P,W)]} |\kappa_M(\frac{\pi_l - \Pi}{h_n}) - \kappa_M(\frac{\pi - \Pi}{h_n})| f(\Pi, P, W) d\Pi d(P, W) \\ &\leq m_\kappa (n h_n^{2-a})^{-\frac{1}{2}} \mathcal{P}\{P \leq p, W \geq w, \Pi \leq \pi(P, W)\} \\ &\leq m_\kappa (n h_n^{2-a})^{-\frac{1}{2}} \end{aligned}$$

Let $a = 1$, given $nh_n \to \infty$, we have $P_{1n} = o_p(1)$ and $P_{3n} = o(1)$. For any $l \in \{1, ...L_n\}$,

$$\begin{aligned} P_{2n} &= |\hat{P}(\pi_l, p, w) - E(\hat{P}(\pi_l, p, w))| \\ &\leq \max_{1 \leq l \leq L_n} |\hat{P}(\pi_l, p, w) - E(\hat{P}(\pi_l, p, w))| \end{aligned}$$

We need to show that for any $\varepsilon > 0$, there exists some $\Delta_\varepsilon > 0$, such that

$$\mathcal{P}\{(\frac{n}{\ln(n)})^{\frac{1}{2}} \max_{1 \leq l \leq L_n} |\hat{P}(\pi_l, p, w) - E(\hat{P}(\pi_l, p, w))| \geq \Delta_\varepsilon\} < \varepsilon.$$

Note that

$$\mathcal{P}\{(\frac{n}{\ln(n)})^{\frac{1}{2}} \max_{1 \leq l \leq L_n} |\hat{P}(\pi_l, p, w) - E(\hat{P}(\pi_l, p, w))| \geq \Delta_\varepsilon\} \leq \sum_{l=1}^{L_n} \mathcal{P}\{(\frac{n}{\ln(n)})^{\frac{1}{2}} |\hat{P}(\pi_l, p, w) - E(\hat{P}(\pi_l, p, w))| \geq \Delta_\varepsilon\}.$$

Write $|\hat{P}(\pi_l,p,w) - E(\hat{P}(\pi_l,p,w))| = |\frac{1}{n}\sum_{i=1}^n W_{in}|$ where

$$W_{in} = \kappa_M(\frac{\pi_l - \Pi_i}{h_n})I(P_i \le p, W_i \ge w) - E[\kappa_M(\frac{\pi_l - \Pi_i}{h_n})I(P_i \le p, W_i \ge w)]$$

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

$$\mathcal{P}\{(\frac{n}{\ln(n)})^{\frac{1}{2}}|\hat{P}(\pi_l,p,w) - E(\hat{P}(\pi_l,p,w))| \ge \Delta_\varepsilon\} < 2\exp(-\frac{n\Delta_\varepsilon^2 \cdot (\frac{n}{\ln(n)})^{-1}}{2\bar{\sigma}_n^2 + \frac{4}{3}\Delta_\varepsilon \cdot (\frac{n}{\ln(n)})^{-\frac{1}{2}}})$$

with $\bar{\sigma}_n^2 = n^{-1}\sum_{i=1}^n V(W_{in}) \to P(\pi_l,p,w)(1-P(\pi_l,p,w))$ by Lemma 1 or 2. Thus $2\bar{\sigma}_n^2 + \frac{4}{3}\Delta_\varepsilon \cdot (\frac{n}{\ln(n)})^{-\frac{1}{2}} \to 2P(\pi_l,p,w)(1-P(\pi_l,p,w))$, Hence provided that $\Delta_\varepsilon^2 > 2P(\pi_l,p,w)(1-P(\pi_l,p,w))$,

$$\begin{aligned} P_{2n} &\le L_n \mathcal{P}\{(\frac{n}{\ln(n)})^{\frac{1}{2}}|\hat{P}(\pi_l,p,w) - E(\hat{P}(\pi_l,p,w))| \ge \Delta_\varepsilon\} \\ &< r_\pi(\frac{n}{h_n})^{\frac{1}{2}} \cdot 2\exp(-\ln(n)) = r_\pi(nh)^{-\frac{1}{2}} \end{aligned}$$

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

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

$$\begin{aligned} E(\hat{P}(\pi,p,w) &= \int_{D_{p,w}}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W) \\ &= \int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W) \\ &\quad + \int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W) \\ &\quad + \int_{\pi^{-1}((\pi,\pi(p,w)])}\int_{[0,\pi]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W) \\ &\quad + \int_{\pi^{-1}((\pi,\pi(p,w)])}\int_{[\pi,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W) \end{aligned}$$

Therefore, by triangular inequality, we have

$$\sup_{\pi\in[0,\pi(p,w)]}|E(\hat{P}(\pi,p,w) - P(\pi,p,w))| = \sup_{\pi\in[0,\pi(p,w)]}G_{1n} + \sup_{\pi\in[0,\pi(p,w)]}G_{2n} + \sup_{\pi\in[0,\pi(p,w)]}G_{3n} + \sup_{\pi\in[0,\pi(p,w)]}G_{4n}$$

where

$$
\begin{aligned}
G_{1n} &= |\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)-\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}f(\Pi,P,W)d\Pi d(P,W)| \\
G_{2n} &= |\int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)-\int_{\pi^{-1}(\{\pi\})}\int_{[0,\pi(P,W)]}f(\Pi,P,W)d\Pi d(P,W)| \\
G_{3n} &= |\int_{\pi^{-1}((\pi,\pi(p,w)])}\int_{[0,\pi]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)-\int_{\pi^{-1}((\pi,\pi(p,w)])}\int_{[0,\pi]}f(\Pi,P,W)d\Pi d(P,W)| \\
G_{4n} &= |\int_{\pi^{-1}((\pi,\pi(p,w)])}\int_{[\pi,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)|
\end{aligned}
$$

For the first term, when $(P,W)\in\pi^{-1}([0,\pi))$, $\Pi\le\pi(P,W)<\pi$. This implies $\kappa_M(\frac{\pi-\Pi}{h_n})\to 1$ as $n\to\infty$. First, by LDC,

$$
\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\to\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}f(\Pi,P,W)d\Pi d(P,W).
$$

Second, $\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)$ is increasing with n. Furthermore, By the Lipschitz condition imposed on $\kappa_M(.)$,

$\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)$ is a continuous function in π. As a result, by Dini's Theorem,

$$
\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}\kappa_M(\frac{\pi-\Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)\to\int_{\pi^{-1}([0,\pi))}\int_{[0,\pi(P,W)]}f(\Pi,P,W)d\Pi d(P,W)
$$

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

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

$$
\delta(\varepsilon,p,w)=\min\{F(\pi_\alpha(p,w)+\varepsilon|C_{p,w})-F(\pi_\alpha(p,w)|C_{p,w}),F(\pi_\alpha(p,w)|C_{p,w})-F(\pi_\alpha(p,w)-\varepsilon|C_{p,w})\}>0
$$

For all $\omega \in A$,

(1) when $F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w}) > 0$, we have $\pi_{\alpha,n}(p,w) > \pi_\alpha(p,w) + \varepsilon$. By monotonicity,

$$F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w}) > F(\pi_\alpha(p,w) + \varepsilon|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w}) \geq \delta(\varepsilon,p,w).$$

(2) Similarly, when $F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w}) < 0$, we have

$$F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w}) < F(\pi_\alpha(p,w) - \varepsilon|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w}) \leq -\delta(\varepsilon,p,w).$$

As a result, For $\omega \in A$, $|F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w})| > \delta(\varepsilon,p,w)$. i.e. $A \subseteq B = \{\omega : |F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w})| > \delta(\varepsilon,p,w)\}$. Thus, $\mathcal{P}(A) \leq \mathcal{P}(B)$. Therefore, we just need to prove $|F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w})| = o_p(1)$. Note that

$$\begin{aligned}
& |F(\pi_{\alpha,n}(p,w)|C_{p,w}) - F(\pi_\alpha(p,w)|C_{p,w})| \\
= \; & |F(\pi_{\alpha,n}(p,w)|C_{p,w}) - \hat{F}(\pi_{\alpha,n}(p,w)|C_{p,w})| \\
\leq \; & \sup_{\pi \in \mathbb{R}_+} |F(\pi|C_{p,w}) - \hat{F}(\pi|C_{p,w})| \\
\leq \; & \sup_{\pi \in \mathbb{R}_+} |\frac{P(\pi,p,w)}{P_{PW}(p,w)} - \frac{\hat{P}(\pi,p,w)}{\hat{P}_{PW}(p,w)}| \\
\leq \; & \sup_{\pi \in \mathbb{R}_+} |\frac{P(\pi,p,w)}{P_{PW}(p,w)} - \frac{P(\pi,p,w)}{\hat{P}_{PW}(p,w)}| + \sup_{\pi \in \mathbb{R}_+} |\frac{P(\pi,p,w)}{\hat{P}_{PW}(p,w)} - \frac{\hat{P}(\pi,p,w)}{\hat{P}_{PW}(p,w)}| \\
\leq \; & \sup_{\pi \in \mathbb{R}_+} P(\pi,p,w)|\frac{1}{P_{PW}(p,w)} - \frac{1}{\hat{P}_{PW}(p,w)}| + |\frac{1}{\hat{P}_{PW}(p,w)}| \sup_{\pi \in \mathbb{R}_+} |P(\pi,p,w) - \hat{P}(\pi,p,w)| \\
\leq \; & P_{PW}(p,w)|\frac{1}{P_{PW}(p,w)} - \frac{1}{\hat{P}_{PW}(p,w)}| + |\frac{1}{\hat{P}_{PW}(p,w)}| \sup_{\pi \in \mathbb{R}_+} |P(\pi,p,w) - \hat{P}(\pi,p,w)|
\end{aligned}$$

Note that $\hat{P}_{PW}(p,w) - P_{PW}(p,w) = o_p(1)$ by the properties of indicator function. By Slutsky theorem we have $\frac{1}{P_{PW}(p,w)} - \frac{1}{\hat{P}_{PW}(p,w)} = o_p(1)$. Since $\hat{P}_{PW}(p,w) = O_p(1)$, we just need to prove $\sup_{\pi \in \mathbb{R}_+} |P(\pi,p,w) - \hat{P}(\pi,p,w)| = o_p(1)$.

$$\sup_{\pi \in \mathbb{R}_+} |P(\pi,p,w) - \hat{P}(\pi,p,w)| \leq \sup_{\pi \in [0,\pi(p,w)]} |P(\pi,p,w) - \hat{P}(\pi,p,w)| + \sup_{\pi \in (\pi(p,w),\infty)} |P(\pi,p,w) - \hat{P}(\pi,p,w)|$$

From Lemma 3, $\sup_{\pi\in[0,\pi(p,w)]}|P(\pi,p,w)-\hat{P}(\pi,p,w)| = o_p(1)$. For all $\pi\in(\pi(p,w),\infty)$,

$$\begin{aligned} P(\pi,p,w) &= \mathcal{P}(\Pi\leq\pi, P\leq p, W\geq w) \\ &= \mathcal{P}(\Pi\leq\pi(p,w), P\leq p, W\geq w) \\ &= \mathcal{P}(P\leq p, W\geq w) \\ &= P_{PW}(p,w) \end{aligned}$$

Given $\min_{\{i:P_i\leq p, W_i\geq w\}}\Pi_i \geq h_nB_M$, and for any i, $\Pi_i\leq\pi(p,w)<\pi$. There exist $N(p,w)$ such that for all $n>N(p,w)$,

$$\begin{aligned} \hat{P}(\pi,p,w) &= (nh_n)^{-1}\sum_{i=1}^{n}(\int_0^{\pi} M_k(\frac{\Pi_i-\gamma}{h_n})d\gamma)I(P_i\leq p, W_i\geq w) \\ &= n^{-1}\sum_{i=1}^{n}\int_{-\frac{\Pi_i}{h_n}}^{\frac{\pi-\Pi_i}{h_n}} M_k(\varphi)d\varphi I(P_i\leq p, W_i\geq w) \\ &= n^{-1}\sum_{i=1}^{n}\int_{-B_M}^{B_M} M_k(\varphi)d\varphi I(P_i\leq p, W_i\geq w) \\ &= n^{-1}\sum_{i=1}^{n} I(P_i\leq p, W_i\geq w) \\ &= \hat{P}_{PW}(p,w) \end{aligned}$$

As a result, as $\hat{P}_{PW}(p,w)\to P_{PW}(p,w)$ as $n\to\infty$

$$\begin{aligned} & \sup_{\pi\in(\pi(p,w),\infty)}|P(\pi,p,w)-\hat{P}(\pi,p,w)| \\ \leq\ & \sup_{\pi\in(\pi(p,w),\infty)}|P(\pi,p,w)-P_{PW}(p,w)| + \sup_{\pi\in(\pi(p,w),\infty)}|\hat{P}(\pi,p,w)-P_{PW}(p,w)| \\ =\ & o_p(1) \end{aligned}$$

The result then follows. □

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

$$\begin{aligned}\pi_{\alpha,n}(p,w)-\pi_\alpha(p,w) &= \frac{\hat F(\pi_{\alpha,n}(p,w)|C_{p,w})-\hat F(\pi_\alpha(p,w)|C_{p,w})}{\hat f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})}\\ &= \frac{F(\pi_\alpha(p,w)|C_{p,w})-\hat F(\pi_\alpha(p,w)|C_{p,w})}{\hat f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})}\end{aligned}$$

where $\hat f(\pi|C_{p,w})=\frac{\partial\hat F(\pi|C_{p,w})}{\partial\pi}$ and $\bar\pi_{\alpha,n}(p,w)=\lambda\pi_{\alpha,n}(p,w)+(1-\lambda)\pi_\alpha(p,w)$ for some $\lambda\in(0,1)$. Write

$$\pi_{\alpha,n}(p,w)-\pi_\alpha(p,w)=(A_n+C_n)(\frac{1}{f(\pi_\alpha(p,w)|C_{p,w})}+\beta_n)$$

where

$$\begin{aligned}A_n &= F(\pi_\alpha(p,w)|C_{p,w})-\frac{E(\hat P(\pi_\alpha(p,w),p,w))}{E(\hat P_{PW}(p,w))}\\ C_n &= \frac{E(\hat P(\pi_\alpha(p,w),p,w))}{E(\hat P_{PW}(p,w))}-\hat F(\pi_\alpha(p,w)|C_{p,w})\\ \beta_n &= \frac{1}{\hat f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})}-\frac{1}{f(\pi_\alpha(p,w)|C_{p,w})}\end{aligned}$$

The theorem follows if (a) $\beta_n=o_p(1)$; (b) $A_n=-\frac{1}{2}h_n^2\sigma_M^2\frac{\int_{\pi^{-1}([\pi_\alpha(p,w),\pi(p,w)])}f^{(1)}(\pi_\alpha(p,w),P,W)d(P,W)}{P_{PW}(p,w)}+o(h_n^2)$; (c) $(\frac{s_n(p,w)}{\hat P_{PW}(p,w)})^{-1}\sqrt{n}C_n\to N(0,1)$ where

$$\begin{aligned}s_n^2(p,w) &= P(\pi_\alpha(p,w),p,w)-\frac{(P(\pi_\alpha(p,w),p,w))^2}{P_{PW}(p,w)}\\ &\quad -2h_n\sigma_\kappa\int_{\pi^{-1}([\pi_\alpha(p,w),\pi(p,w)])}f(\pi_\alpha(p,w),P,W)d(P,W)+o(h_n)\end{aligned}$$

(a) By Slutsky theorem, it is suffice to prove $\hat f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})-f(\pi_\alpha(p,w)|C_{p,w})=o_p(1)$. Since $\pi_{\alpha,n}(p,w)-\pi_\alpha(p,w)=o_p(1)$ by theorem 1, also, $\bar\pi_{\alpha,n}(p,w)-\pi_\alpha(p,w)=o_p(1)$.

$$\begin{aligned}&|\hat f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})-f(\pi_\alpha(p,w)|C_{p,w})|\\ \le\ &|\hat f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})-f(\bar\pi_\alpha(p,w)|C_{p,w})|+|f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})-f(\pi_\alpha(p,w)|C_{p,w})|\\ \le\ &|\hat f(\bar\pi_{\alpha,n}(p,w)|C_{p,w})-f(\bar\pi_\alpha(p,w)|C_{p,w})|+o_p(1)\end{aligned}$$

by continuity of f. Therefore it is suffice to prove $\sup_{\pi\in G}|\hat f(\pi|C_{p,w})-f(\pi|C_{p,w})|=o_p(1)$. where

G is a compact set and $G \subset (0, \pi(p,w))$.

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

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

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

$$
\sup_{\pi\in G} \int_{\pi^{-1}((\pi,\pi(p,w)])} f(\pi,P,W)d(P,W) \leq B_f \int_{\pi^{-1}((\pi,\pi(p,w)])} d(P,W) = O(1)
$$

by Assumptions 3 and 4.

Denote $Q_n(p,w) = (nh_n)^{-1}\sum_{i=1}^n M_k(\frac{\Pi_i-\pi}{h_n})I(P_i \leq p, W_i \geq w)$, Thus,

$$
\begin{aligned}
& \sup_{\pi\in G} \left|Q_n(p,w) - \int_{\pi^{-1}((\pi,\pi(p,w)])} f(\pi,P,W)d(P,W)\right| \\
\leq \ & \sup_{\pi\in G} |Q_n(p,w) - E(Q_n(p,w))| \\
& + \sup_{\pi\in G} \left|E(Q_n(p,w)) - \int_{D_{p,w}} \kappa_M(\frac{\pi(P,W)-\pi}{h_n}) f(\pi,P,W)d(P,W)\right| \\
& + \sup_{\pi\in G} \left|\int_{\pi^{-1}((\pi,\pi(p,w)])} \kappa_M(\frac{\pi(P,W)-\pi}{h_n}) f(\pi,P,W)d(P,W) - \int_{\pi^{-1}((\pi,\pi(p,w)])} f(\pi,P,W)d(P,W)\right| \\
& + \sup_{\pi\in G} \left|\int_{\pi^{-1}([0,\pi])} \kappa_M(\frac{\pi(P,W)-\pi}{h_n}) f(\pi,P,W)d(P,W)\right| \\
= \ & Q_{1n} + Q_{2n} + Q_{3n} + Q_{4n}
\end{aligned}
$$

Follow the similar proof process as in Lemma 3 (a), we can prove that $Q_{1n} = O_p((\frac{\ln n}{nh_n})^{\frac{1}{2}})$ if

$nh_n^2 \to \infty$. For any (p, w), there exist some $N(p, w)$ such that when $n > N(p, w)$

$$\begin{aligned} E(Q_n(p,w)) &= {h_n}^{-1} \int_{D_{p,w}} \int_{[0,\pi(P,W)]} M_k(\frac{\Pi - \pi}{h_n}) f(\Pi, P, W) d\Pi d(P,W) \\ &= \int_{D_{p,w}} \int_{-\frac{\pi}{h_n}}^{\frac{\pi(P,W)-\pi}{h_n}} M_k(\varphi) f(\pi + h_n\varphi, P, W) d\varphi d(P,W) \\ &= \int_{D_{p,w}} \int_{-B_M}^{\frac{\pi(P,W)-\pi}{h_n}} M_k(\varphi) f(\pi + h_n\varphi, P, W) d\varphi d(P,W) \end{aligned}$$

By Taylor's theorem, for any $\pi \in G$,

$$\begin{aligned} & |\int_{D_{p,w}} \int_{-B_M}^{\frac{\pi(P,W)-\pi}{h_n}} M_k(\varphi)(f(\pi + h_n\varphi, P, W) - f(\pi, P, W)) d\varphi d(P,W)| \\ \leq & \int_{D_{p,w}} \int_{-B_M}^{\frac{\pi(P,W)-\pi}{h_n}} M_k(\varphi)|(f(\pi + h_n\varphi, P, W) - f(\pi, P, W))| d\varphi d(P,W) \\ \leq & m_f h_n \int_{D_{p,w}} \int_{-B_M}^{B_M} M_k(\varphi)|\varphi| d\varphi d(P,W) + o(h_n) \\ = & O(h_n) \end{aligned}$$

Therefore, $Q_{2n} = o(1)$. Since when $(P, W) \in \pi^{-1}([0, \pi])$, $\kappa_M(\frac{\pi(P,W)-\pi}{h_n}) \to 0$ and when $(P, W) \in \pi^{-1}((\pi, \pi(p, w)])$, $\kappa_M(\frac{\pi(P,W)-\pi}{h_n}) \to 1$. By LDC, for any $\pi \in G$,

$$\int_{\pi^{-1}((\pi,\pi(p,w)])} \kappa_M(\frac{\pi(P,W) - \pi}{h_n}) f(\pi, P, W) d(P,W) \to \int_{\pi^{-1}((\pi,\pi(p,w)])} f(\pi, P, W) d(P,W)|$$

and

$$\int_{\pi^{-1}([0,\pi])} \kappa_M(\frac{\pi(P,W) - \pi}{h_n}) f(\pi, P, W) d(P,W) \to 0$$

Therefore, $Q_{3n} = o(1)$ and $Q_{4n} = o(1)$. In sum, Noting that $\frac{1}{\hat{P}_{PW}(p,w)} = O_p(1)$, we have

$$\sup_{\pi \in G} |\hat{f}(\pi | C_{p,w}) - f(\pi | C_{p,w})| = o_p(1)$$

As a result, $\beta_n = o_p(1)$.

(b):

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

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

$$A_{2n} = -\frac{1}{2}h_n^2\sigma_M^2 \int_{\pi^{-1}((\pi_\alpha(p,w),\pi(p,w)])} f^{(1)}(\pi_\alpha(p,w),P,W)d(P,W) + o(h_n^2)$$

The result then follows.

(c):

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

where

$$Z_{in} = \frac{1}{\sqrt{n}}(\frac{E(\hat{P}(\pi_\alpha(p,w),p,w))}{P_{PW}(p,w)}I(P_i \le p, W_i \ge w) - \frac{1}{h_n}\int_0^{\pi_\alpha(p,w)} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma I(P_i \le p, W_i \ge w))$$

Here,

$$E(Z_{in}) = \frac{1}{\sqrt{n}}(E(\hat{P}(\pi_\alpha(p,w),p,w)) - E(\hat{P}(\pi_\alpha(p,w),p,w))) = 0$$

$$\sum_{i=1}^{n} E(Z_{in}^2) = s_n^2(p,w) = s_{1n} + s_{2n} + s_{3n}$$

where

$$\begin{aligned} s_{1n} &= \frac{\{E(\hat{P}(\pi_\alpha(p,w),p,w))\}^2}{P_{PW}(p,w)^2} E(I(P_i \le p, W_i \ge w)) = \frac{\{E(\hat{P}(\pi_\alpha(p,w),p,w))\}^2}{P_{PW}(p,w)} \\ s_{2n} &= E[(\frac{1}{h_n}\int_0^{\pi_\alpha(p,w)} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma I(P_i \le p, W_i \ge w)))^2] = E[(\hat{P}(\pi_\alpha(p,w),p,w)^2] \\ s_{3n} &= -2\frac{E(\hat{P}(\pi_\alpha(p,w),p,w)}{P_{PW}(p,w)} E(\frac{1}{h_n}\int_0^{\pi_\alpha(p,w)} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma I(P_i \le p, W_i \ge w)) \\ &= -2s_{1n} \end{aligned}$$

By Lemma 2,

$$\begin{aligned} E(\hat{P}(\pi_\alpha(p,w),p,w)) &= P(\pi_\alpha(p,w),p,w) + \frac{1}{2}h_n^2\sigma_M^2 \int_{\pi^{-1}((\pi,\pi(p,w)])} f^{(1)}(\pi,P,W)d(P,W) + o(h_n^2); \\ E[(\hat{P}(\pi_\alpha(p,w),p,w)^2] &= P(\pi_\alpha(p,w),p,w) - 2h_n\sigma_\kappa \int_{\pi^{-1}((\pi_\alpha(p,w),\pi(p,w)])} f(\pi_\alpha(p,w),P,W)d(P,W) + o(h_n) \end{aligned}$$

As a result,

$$\begin{aligned} s_{1n} &= \frac{1}{P_{PW}(p,w)}(P(\pi_\alpha(p,w),p,w) + \frac{1}{2}h_n^2\sigma_M^2 \int_{\pi^{-1}((\pi,\pi(p,w)])} f^{(1)}(\pi,P,W)d(P,W) + o(h_n^2))^2 \\ &= \frac{(P(\pi_\alpha(p,w),p,w))^2}{P_{PW}(p,w)} + o(h_n) \\ s_{2n} &= P(\pi_\alpha(p,w),p,w) - 2h_n\sigma_\kappa \int_{\pi^{-1}((\pi_\alpha(p,w),\pi(p,w)])} f(\pi_\alpha(p,w),P,W)d(P,W) + o(h_n) \\ s_{3n} &= -2s_{1n} = -2\frac{(P(\pi_\alpha(p,w),p,w))^2}{P_{PW}(p,w)} + o(h_n) \end{aligned}$$

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

By Liapounov's CLT, $\sum_{i=1}^{n} \frac{Z_{in}}{s_n(p,w)} \xrightarrow{d} N(0,1)$ if $\lim_{n\to\infty} \sum_{i=1}^{n} E(|\frac{Z_{in}}{s_n(p,w)}|^{2+\delta}) = 0$ for some $\delta > 0$.

$$
\sum_{i=1}^{n} E(|\frac{Z_{in}}{s_n(p,w)}|^{2+\delta}) \leq \sum_{i=1}^{n} E(|Z_{in}|^{2+\delta}|\frac{1}{s_n(p,w)}|^{2+\delta})
$$

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

$$
\begin{aligned}
\sum_{i=1}^{n} E(|Z_{in}|^{2+\delta}) &\leq 2^{1+\delta}(n^{-2/\delta}E(|\frac{E(\hat{P}(\pi_\alpha(p,w),p,w))}{P_{PW}(p,w)}I(P_i \leq p, W_i \geq w)|^{2+\delta}) \\
+ &\quad n^{-2/\delta}E(|\frac{1}{h_n}\int_0^{\pi_\alpha(p,w)} M_k(\frac{\Pi_i - \gamma}{h_n})d\gamma I(P_i \leq p, W_i \geq w)|^{2+\delta})) \\
&= 2^{1+\delta}(n^{-2/\delta}E|\frac{E(\hat{P}(\pi_\alpha(p,w),p,w))}{P_{PW}(p,w)}|^{2+\delta}E(I(P_i \leq p, W_i \geq w))) \\
&\quad + n^{-2/\delta}\int_{D_{p,w}}\int_{[0,\pi(P,W)]} \kappa_M(\frac{\pi_\alpha(p,w) - \Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W)
\end{aligned}
$$

Since $E(I(P_i \leq p, W_i \geq w) = O(1)$,

$$
\begin{aligned}
n^{-2/\delta}E|\frac{E(\hat{P}(\pi_\alpha(p,w),p,w))}{P_{PW}(p,w)}|^{2+\delta} &= n^{-2/\delta}\frac{|E(\hat{P}(\pi_\alpha(p,w),p,w))|^{2+\delta}}{P_{PW}(p,w)^{2+\delta}} \\
&= O(n^{-2/\delta})
\end{aligned}
$$

Since $\kappa_M(.) \leq 1$, $f < B_f$ and $\pi \leq B_\pi$,

$$
\begin{aligned}
& n^{-2/\delta}\int_{D_{p,w}}\int_{[0,\pi(P,W)]} \kappa_M(\frac{\pi_\alpha(p,w) - \Pi}{h_n})f(\Pi,P,W)d\Pi d(P,W) \\
\leq\ & n^{-2/\delta}B_f\int_{D_{p,w}}\int_{[0,\pi(P,W)]} d\Pi d(P,W) \\
\leq\ & n^{-2/\delta}B_f\int_{\pi^{-1}[0,B_\pi]}\int_{[0,B_\pi]} d\Pi d(P,W) \\
=\ & O(n^{-2/\delta})
\end{aligned}
$$

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,

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

The result then follows. □

References

Aigner, D., Lovell, C. and Schmidt, P. (1977). Formulation and estimation of stochastic frontiers production function models, *Journal of Econometrics* **6**: 21–37.

Ali, F., Parikh, A. and Shah, M. (1994). Measurement of profit efficiency using behavioural and stochastic frontier approaches, *Applied Economics* **26**: 181–188.

Aragon, Y., Daouia, A. and Thomas-Agnan, C. (2005). Nonparametric frontier estimation: a conditional quantile-based approach, *Econometric Theory* **21**: 358–389.

Cazals, C., Florens, J. P. and Simar, L. (2002). Nonparametric frontier estimation: A robust appoach, *Journal of Econometrics* **106**: 1–25.

Charnes, A., Cooper, W. and Rhodes, E. (1978). Measuring the efficiency of decision making units, *European Journal of Operational Research* **2**: 429–444.

Daouia, A., Florens, J. P. and Simar, L. (2010). Frontier estimation and extreme value theory, *Bernoulli* **16**: 1039–1063.

Daouia, A. and Simar, L. (2005). Robust nonparametric estimators of monotone boundaries, *Journal of Multivariate Analysis* **96**: 311–331.

Daouia, A. and Simar, L. (2007). Nonparametric efficiency analysis: A multivariate conditional quantile approach, *Journal of Econometrics* **140**: 375–400.

Fan, Y., Li, Q. and Weersink, A. (1996). Semiparametric estimation of stochastic production frontier models, *Journal of Business and Economic Statistics* **14**: 460–468.

Farrell, M. (1957). The measurement of productive efficiency, *Journal of the Royal Statistical Society* **A 120**: 253–290.

Gijbels, I., Mammen, E., Park, B. and Simar, L. (1999). On estimation of monotone and concave frontier functions, *Journal of American Statistical Association* **94**: 220–228.

Greene, W., H. (2008). The econometric approach to efficiency analysis, *Oxford University Press* .

Kneip, A., Park, B. and Simar, L. (1998). A note on the convergence of nonparametric dea efficiency measures, *Econometric Theory* **14**: 783–793.

Korostelev, A., Simar, L. and Tsybakov, A. B. (1995). Efficient 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 profile likelihood, *Econometric Reviews* **34**: 413–451.

Maudos, J., Pastor, J. M., Perez, F. and Quesada, J. (2002). Cost and profit efficiency in european banks, *Journal of International Financial Markets, Institutions and Money* **12**: 33–58.

Meeusen, W. and van den Broeck, J. (1977). Efficiency estimation from cobb-douglas production functions with composed error, *International Economic Review* **18**: 435–444.

Mynbaev, K. and Martins-Filho, C. (2010). Bias reduction in kernel density estimation via lipschitz condition, *Journal of Nonparametric Statistics* **22**: 219–235.

Park, U. B., Jeong, S.-O. and Simar, L. (2010). Asymptotic distribution of conical-hull estimators of directional edges, *Annals of Statistics* **38**: 1320–1340.

Park, U. B., Simar, L. and Weiner, C. (2000). The fdh estimator for productivity efficiency scores c asymptotic properties, *Econometric Theory* **16**: 855–877.

Simar, L. and Vanhems, A. (2012). Probabilistic characterization of directional distances and their robust versions, *Journal of Econometrics* **166**: 342–354.

Simar, L. and Zelenyuk, V. (2011). Stochastic fdh/dea estimators for frontier analysis, *Journal of Productivity Analysis* **36**: 1–20.