Online Appendices to "Offline Feature-Based Pricing under Censored Demand: A Causal Inference Approach" by Jingwen Tang, Zhengling Qi, Ethan X. Fang, Cong Shi

Appendix A: Summary of Major Notation

Table 2: Summary of Major Notation for Problem Formulation

P	the price of a product
$\mathcal{P}$	$[p_1, p_2]$ the compact space of price values
Y	the amount of inventory available for sales
$\mathcal{Y}$	the set for inventory vales $\mathcal{Y} \subseteq [0, \infty)$
$D(p)$	the potential demand of a product if the price P is set to be a (deterministic) value p
q	number of dimensions of features associated with the product
X	observed q-dimensional features associated with the product
$\mathcal{X}$	some feature space $\mathcal{X} \subset \mathbb{R}^q$
π	a pricing strategy being a measurable function: $(\mathcal{X}, \mathcal{Y}) \to \mathcal{P}$
Π	the class of all pricing strategies
$D(\pi)$	the potential outcome under a pricing strategy $\pi \in \Pi$
$V(\pi)$	the expected profit of a pricing strategy $\pi \in \Pi$
c	the stockout cost per unit
$f(P \mid X, Y)$	conditional probability density of the price, commonly referred to as the generalized propensity score
$f_{\min}$	an almost surely positive lower bound for the conditional probability density of the price $f(P \mid X, Y)$
$Q(X, Y, P)$	the expected profit of a product given the product features X, inventory amount Y and price P
π^*	the global optimal pricing strategy which maximizes the expected profit $V(\pi)$
$K(u)$	a kernel function $\mathbb{R} \to [0, \infty)$
h	the bandwidth used in kernel approximation
$V_h(\pi)$	approximated expected profit of pricing strategy π using kernel approximation with bandwidth h
C_1	a constant in Assumption 2(a)
C_2	a constant in Assumption 2(b)
C_3	a constant in Lemma 2
S	the observed sales quantity $S = \min\{D, Y\}$
Δ	censor indicator $\Delta = \mathbb{1}(D < Y)$
$R(X, P, S, \Delta)$	the surrogate profit given X, P, S, Δ of a product

Table 3: Summary of Major Notation for Offline Feature-Based Pricing Strategy

$D_{\max}$	no-negative constant, upper bound of D in the assumption b
n	sample size
$\mathcal{D}_n$	n independent and identically distributed samples
$H(t\|X, P)$	the conditional survival function of the demand D, $H(t\|X, P) = \mathbb{P}(D > t \mid X, P)$
$\widehat{H}(t\|X, P)$	estimated conditional survival function using random forests method
ℓ	number of unique sales values in the dataset $\mathcal{D}_n$
$\widehat{\mathbb{E}}[D\|X, P, S, \Delta = 0]$	estimated conditional expectation of demand using estimated $H(t\|X, P)$
$\widehat{R}(X, P, S, \Delta)$	estimated potential profit of a product
$\widehat{f}(P \mid X, Y)$	estimated conditional density function of the price
Π_0	some pre-specified class of pricing strategies
λ_n	a positive tuning parameter possibly depending on the sample size n
$\\|\cdot\\|_{\Pi_0}$	norm of the Hilbert space Π_0
$J(\pi)$	some regularization function on the policy π, set to be $\\|\pi\\|^2_{\Pi_0}$
β_0	constant term in the example of linear pricing strategies
β	parameters of the features in the example of linear pricing strategies, $\beta \in \mathbb{R}^{q+1}$
$\widehat{Q}(X, Y, P)$	estimated expected conditional potential profit $\widehat{Q}(X, Y, P) = \mathbb{E}[R \mid X, Y, P]$
$\widehat{V}_n^{DR}(\pi)$	a doubly robust estimator for estimating $V_h(\pi)$
$\widehat{\pi}$	the estimated global optimal pricing policy by solving (16)
$\widehat{\pi}_n$	the estimated global optimal pricing policy by solving (17)

B	number of survival trees in the random survival forests algorithm
M	number of folds in the cross-fitting technique
m	index of the fold in the cross-fitting technique, $m = 1, \ldots, M$
$m(i)$	the fold containing the i-th observation
$\widehat{Q}^{(-m)}(X, Y, P)$	estimated expected conditional potential profit using data excluding fold m
$\widehat{f}^{(-m)}(P \mid X, Y)$	estimated conditional probability density of the price using data excluding fold m
$\mathcal{D}_n^{(-m)}$	the other $(M-1)$ folds data except k
ϕ_1	parameters of the neural network for the mean of the distribution of $(P \mid X, Y)$
ϕ_2	parameters of the neural network for the covariance of the distribution of $(P \mid X, Y)$
$\widehat{\phi}_1$	estimated ϕ_1 using MLE
$\widehat{\phi}_2$	estimated ϕ_2 using MLE
$\widehat{\mu}^{(-m(i))}(X_i, Y_i)$	estimated mean of the the multi-variate Gaussian distribution of $(P \mid X, Y)$ using $\widehat{\phi}_1$
$\widehat{\sigma}^{(-m(i))}(X_i, Y_i)$	estimated covariance matrix of the multivariate Gaussian distribution of the price $(P \mid X, Y)$ using $\widehat{\phi}_2$
ϕ_3	parameters of the neural network for the pricing policy
$\widehat{\phi}_3$	estimated ϕ_3 maximizing the right hand side of (17)

Table 4: Summary of Major Notation for Regret Analysis and Double Robustness

$\mathbf{Regret}(\widehat{\pi}_n)$	the regret of the pricing strategy $\widehat{\pi}_n$
C_4	a constant in Assumption 4
A	a constant in Assumption 5
v	a constant in Assumption 5
$\widetilde{Q}$	some probability measure on (X, Y)
$\|\cdot\|_{Q,2}$	the L_2-norm under $\widetilde{Q}$ on (X, Y)
F	the envelope function of Π_0
C_5	a constant in Assumption 6(a)
C_6	a constant in Assumption 6(b)
α, β	constants in Assumption 6(b)
ω	the smoothness coefficient of the true Q
π_h^*	the estimated optimal pricing policy by maximizing $V_h(\pi)$

Appendix B: Technical Proofs

Proof of Lemma 1. By definition, we have

$$\begin{aligned}
V(\pi) &= \mathbb{E}\left\{\pi(X,Y) \times \min\{D(\pi), Y\} - c \times (D(\pi) - Y)^+\right\}, \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\pi(X,Y) \times \min\{D(\pi), Y\} - c \times (D(\pi) - Y)^+ \mid X, Y\right\}\right\}, \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\pi(X,Y) \times \min\{D(\pi), Y\} - c \times (D(\pi) - Y)^+ \mid X, Y, P = \pi(X,Y)\right\}\right\}, \\
&= \mathbb{E}\left\{\int_{p_1}^{p_2} \mathbb{E}\left\{p \times \min\{D, Y\} - c \times (D - Y)^+ \mid X, Y, P = p\right\} \mathbb{1}(\pi(X,Y) = p) dp\right\}, \\
&= \mathbb{E}\left\{Q(X, Y, \pi(X,Y))\right\},
\end{aligned}$$

where the third equality is by Assumption 1(c) and the fourth equality is by Assumption 1(a). **Q.E.D.**

Proof of Lemma 2. We first consider the unbounded support of $\mathcal{P}$. As seen from Lemma 1, $V(\pi) = \mathbb{E}\left[Q(X, Y, \pi(X,Y))\right]$. By a similar derivation, we can show that

$$\begin{aligned}
V_h(\pi) &= \mathbb{E}\left\{\frac{Q(X,Y,P) K(\frac{P - \pi(X,Y)}{h})}{h f(P|X,Y)}\right\} \\
&= \mathbb{E}\left\{\int \frac{Q(X,Y,p) K(\frac{p - \pi(X,Y)}{h})}{h} dp\right\} \\
&= \mathbb{E}\left\{\int Q(X, Y, th + \pi(X,Y)) K(t) dt\right\},
\end{aligned}$$

where the last equality is based on the change of variables. Then it can be seen that

$$\begin{aligned}|V_h(\pi)-V(\pi)| &= \left|\mathbb{E}\left\{\int Q(X,Y,th+\pi(X,Y))K(t)dt\right\}-\mathbb{E}\left\{\int Q(X,Y,\pi(X,Y))\right\}\right| \\ &\le \mathbb{E}\left\{\int |Q(X,Y,th+\pi(X,Y))-Q(X,Y,\pi(X,Y))|K(t)dt\right\} \\ &\le \mathbb{E}\left\{\sup_{p_1\le p<p'\le p_2}\left|\frac{Q(X,Y,p)-Q(X,Y,p')}{p'-p}\right|\int |th|K(t)dt\right\} \\ &\le hC_2\int |t|K(t)dt \\ &\le C_3 h,\end{aligned}$$

where the second inequality is given by Assumption 2(b). When $\mathcal{P}$ has a bounded support, we need to normalize the kernel by

$$\widetilde{K}\left(\frac{p-\pi(X,Y)}{h}\right)=K\left(\frac{p-\pi(X,Y)}{h}\right)\Big/\int_{p_1}^{p_2}K\left(\frac{p-\pi(X,Y)}{h}\right)dp.$$

Then by a similar proof as in the unbounded case, we can show that $|V_h(\pi)-V(\pi)|\le C_3h$, which completes our proof. **Q.E.D.**

Proof of Lemma 3. Consider the following quantity:

$$\begin{aligned}&\mathbb{E}\left[(P\times S-c\times(D-Y)^+)\,|X,P,S,Y\right] \\ &=\mathbb{E}\left[(P\times S-c\times(D-Y)^+)\,|X,P,S,Y,\Delta=1\right]\mathbb{1}(\Delta=1) \\ &\quad+\mathbb{E}\left[(P\times S-c\times(D-Y)^+)\,|X,P,S,Y,\Delta=0\right]\mathbb{1}(\Delta=0) \\ &=\mathbb{1}(\Delta=1)PS+\mathbb{1}(\Delta=0)\mathbb{E}\left[(P\times S-c\times(D-Y)^+)\,|X,P,S,Y,\Delta=0\right] \\ &=\mathbb{1}(\Delta=1)PS+\mathbb{1}(\Delta=0)\mathbb{E}\left[(P\times S-c\times(D-Y))\,|X,P,S,D>S,Y=S\right] \\ &=\mathbb{1}(\Delta=1)PS+(P+c)S\mathbb{1}(\Delta=0)-c\mathbb{1}(\Delta=0)\mathbb{E}\left[D|X,P,S,D>S\right] \\ &=PS+cS\mathbb{1}(\Delta=0)-c\mathbb{1}(\Delta=0)\mathbb{E}\left[D|X,P,S,D>S\right]=R,\end{aligned}$$

where the last but two equality is based on Assumption 3(a). In addition, we can rewrite $V(\pi)$ as

$$\begin{aligned}V(\pi)&=\mathbb{E}\left\{\int_{p_1}^{p_2}\mathbb{E}\left[(P\times S-c\times(D-Y)^+)\,|X,Y,P=p\right]\mathbb{1}(\pi(X,Y)=p)dp\right\} \\ &=\mathbb{E}\left\{\int_{p_1}^{p_2}\mathbb{E}\left[R|X,Y,P=p\right]\mathbb{1}(\pi(X,Y)=p)dp\right\} \\ &=\mathbb{E}\left\{R|X,Y,P=\pi(X,Y)\right\},\end{aligned}$$

which concludes our proof. **Q.E.D.**

Proof of Lemma 4. We show this lemma by interchanging the order of integration. Let $h(w\,|\,X,P)$ as the conditional probability density of survival function H. Note that

$$\begin{aligned}&\int_S^{D_{\max}}\frac{H(t|X,P)}{H(S|X,P)}dt=\int_S^{D_{\max}}\int_t^{D_{\max}}\frac{h(w|X,P)}{H(S|X,P)}dwdt \\ &=\int_S^{D_{\max}}\left\{\int_S^{w}\frac{h(w|X,P)}{H(S|X,P)}dt\right\}dw=\int_S^{D_{\max}}(w-S)\frac{h(w|X,P)}{H(S|X,P)}dw \\ &=\int_S^{D_{\max}}w\frac{h(w|X,P)}{H(S|X,P)}dw-S=\mathbb{E}\left[D\,|\,X,P,S,\Delta=0\right]-S,\end{aligned}$$

which concludes the proof. **Q.E.D.**

Proof of Theorem 1. Since by the assumption our estimator $\widehat{R}$ and either $\widetilde{Q}$ or $\widetilde{f}$ are consistent in terms of sup-norm, without loss of generality, we assume $\widehat{R}=R$ and consider either $Q=\widetilde{Q}$ or $f=\widetilde{f}$. If $Q=\widetilde{Q}$, we have

$$\widehat{V}_n^{DR}(\pi) = \frac{1}{nh}\sum_{i=1}^n \int_{p_1}^{p_2} Q(X_i,Y_i,p)K\left(\frac{p-\pi(X_i,Y_i)}{h}\right)dp + \frac{1}{n}\sum_{i=1}^n \frac{1}{h\widetilde{f}(P_i|X_i,Y_i)}K\left(\frac{P_i-\pi(X_i,Y_i)}{h}\right)(R_i - Q(X_i,Y_i,P_i)).$$

By the law of large numbers, we can show that $\widehat{V}_n^{DR}(\pi)$ converges in probability to

$$\frac{1}{h}\mathbb{E}\left[\int_{p_1}^{p_2} Q(X,Y,p)K\left(\frac{p-\pi(X,Y)}{h}\right)dp\right] + \mathbb{E}\left[\frac{1}{h\widetilde{f}(P\,|\,X,Y)}K\left(\frac{P-\pi(X,Y)}{h}\right)(R-Q(X,Y,P))\right] = V_h(\pi),$$

where the equation is given by $\mathbb{E}\left[R-Q(X,Y,P)\,|\,X,Y,P\right]=0$. If $f=\widetilde{f}$, then by the law of large numbers again, we can show that $\widehat{V}_n^{DR}(\pi)$ converges in probability to

$$\begin{aligned}
&\frac{1}{h}\mathbb{E}\left[\int_{p_1}^{p_2} \widetilde{Q}(X,Y,p)K\left(\frac{p-\pi(X,Y)}{h}\right)dp\right] + \mathbb{E}\left[\frac{1}{hf(P\,|\,X,Y)}K\left(\frac{P-\pi(X,Y)}{h}\right)(R-\widetilde{Q}(X,Y,P))\right]\\
&= V_h(\pi) + \frac{1}{h}\mathbb{E}\left[\int_{p_1}^{p_2} \widetilde{Q}(X,Y,p)K\left(\frac{p-\pi(X,Y)}{h}\right)dp\right] - \mathbb{E}\left[\frac{1}{hf(P\,|\,X,Y)}K\left(\frac{P-\pi(X,Y)}{h}\right)\widetilde{Q}(X,Y,P)\right]\\
&= V_h(\pi).
\end{aligned}$$

The proof is complete by noticing that $|V_h(\pi)-V(\pi)| \le C_3 h$ given by Lemma 2. **Q.E.D.**

Proof of Theorem 2. For notational simplicity, let $Z=(X,Y)$ and $\mathcal{Z}=\mathcal{X}\times\mathcal{Y}$. We further let $U(\pi)=-V(\pi)$, $U_h(\pi)=-V_h(\pi)$. By Lemma 2, we can show that

$$\begin{aligned}
V(\pi^*) - V(\widehat{\pi}_n) &= U(\widehat{\pi}_n) - U(\pi^*)\\
&\le U_h(\widehat{\pi}_n) + C_3 h - U_h(\pi^*) + C_3 h + U_h(\pi_h^{\lambda_n}) + \lambda_n J(\pi_h^{\lambda_n}) + \lambda_n J(\widehat{\pi}_n) - \{U_h(\pi_h^{\lambda_n}) + \lambda_n J(\pi_h^{\lambda_n})\}\\
&\le U_h(\pi_h^{\lambda_n}) + \lambda_n J(\pi_h^{\lambda_n}) - U_h(\pi_h^*) + U_h(\widehat{\pi}_n) + \lambda_n J(\widehat{\pi}_n) - \{U_h(\pi_h^{\lambda_n}) + \lambda_n J(\pi_h^{\lambda_n})\} + 2C_3 h\\
&= \Lambda(\lambda_n) + \underbrace{U_h(\widehat{\pi}_n) + \lambda_n J(\widehat{\pi}_n) - \{U_h(\pi_h^{\lambda_n}) + \lambda_n J(\pi_h^{\lambda_n})\}}_{(I)} + 2C_3 h,
\end{aligned}$$

where $\pi_h^{\lambda_n} \in \arg\min_{\pi\in\Pi_0}\{U_h(\pi)+\lambda_n J(\pi)\}$. In the following, we apply the empirical process theory to bound Term (I) on the right hand side of the inequality above. Let

$$\begin{aligned}
\mathcal{G}_\pi \triangleq \Bigg\{ &\int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\pi_h^{\lambda_n}(Z))/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{P-\pi_h^{\lambda_n}(Z)}{h})(R-Q(Z,P)) + \lambda_n J(\pi)\\
&- \int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\pi(Z))/h)}{h}dp - \frac{1}{hf(P|Z)}K(\frac{P-\pi(Z)}{h})(R-Q(Z,P)) - \lambda_n J(\pi_h^{\lambda_n}) \,|\, J(\pi) \lesssim \lambda_n^{-1}, \pi\in\Pi_0 \Bigg\}.
\end{aligned}$$

Based on the definition of $\mathcal{G}_\pi$, we use g_π to denote any generic element in $\mathcal{G}_\pi$. Recall that $J(\pi)=\|\pi\|_{\Pi_0}^2$. We consider a constraint class on π by the following argument. By Assumptions 4 and 6(b), all nuisance functions in (17) are bounded. Then according to the optimization property, we can show that

$$\begin{aligned}
&\frac{1}{nh}\sum_{i=1}^n \int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z_i,p)K(\frac{p-\widehat{\pi}_n(Z_i)}{h})dp + \frac{1}{nh}\sum_{i=1}^n \frac{1}{\widehat{f}^{(-m(i))}(P_i|Z_i)}K(\frac{P_i-\widehat{\pi}_n(Z_i)}{h})(\widehat{R}_i - \widehat{Q}(Z_i,P_i)) + \lambda_n J(\widehat{\pi}_n)\\
&\le \frac{1}{nh}\sum_{i=1}^n \int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z_i,p)K(\frac{p}{h})dp + \frac{1}{nh}\sum_{i=1}^n \frac{1}{\widehat{f}^{(-m(i))}(P_i|Z_i)}K(\frac{P_i}{h})(\widehat{R}_i - \widehat{Q}(Z_i,P_i)),
\end{aligned}$$

which implies that $\lambda_n J(\widehat{\pi}_n) \lesssim 1$. Based on this, we can further show that for any $g_\pi \in \mathcal{G}_\pi$,

$$\|g_\pi\|_\infty \lesssim 1/h + \|\pi\|_{\Pi_0}/h \lesssim \lambda_n^{-\frac{1}{2}}/h,$$

since K is Lipschitz with respect to $\|\bullet\|_{\Pi_0}$ and $\lambda_n \to 0$ with $\lambda_n \leq 1$. The remaining proof consists of two steps. In the first step, we show $\mathbb{E}_n(g_{\widehat{\pi}_n}) \leq \varepsilon_1$, for some $\varepsilon_1 > 0$ with a high probability. In the second step, we aim to show that $\sup_{g_\pi \in \mathcal{G}_\pi} |\mathbb{E}_n(g_\pi) - \mathbb{E}(g_\pi)| \leq \varepsilon_2$, with a high probability for some ε_2. Then combining two, we are able to show $(I) \leq \varepsilon_1 + \varepsilon_2$ with some high probability.

Step 1: We first notice that

$$\begin{aligned}
&\mathbb{E}_n(g_{\widehat{\pi}_n})\\
&= \mathbb{E}_n\left\{\int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\pi_h^{\lambda_n}(Z))/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{P-\pi_h^{\lambda_n}(Z)}{h})(R-Q(Z,P))\right\} + \lambda_n J(\widehat{\pi}_n)\\
&\quad - \mathbb{E}_n\left\{\int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\widehat{\pi}_n(Z))/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{P-\widehat{\pi}_n(Z)}{h})(R-Q(Z,P))\right\} - \lambda_n J(\pi_h^{\lambda_n})\\
&= \mathbb{E}_n\left\{\int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\pi_h^{\lambda_n}(Z))/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{P-\pi_h^{\lambda_n}(Z)}{h})(R-Q(Z,P))\right\}\\
&\quad - \mathbb{E}_n\left\{\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z,p)\frac{K((p-\pi_h^{\lambda_n}(Z))/h)}{h}dp + \frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{P-\pi_h^{\lambda_n}(Z)}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\}\\
&\quad + \mathbb{E}_n\left\{\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z,p)\frac{K((p-\pi_h^{\lambda_n}(Z))/h)}{h}dp + \frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{P-\pi_h^{\lambda_n}(Z)}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\} - \lambda_n J(\pi_h^{\lambda_n})\\
&\quad - \mathbb{E}_n\left\{\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z,p)\frac{K((p-\widehat{\pi}_n(Z))/h)}{h}dp + \frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{P-\widehat{\pi}_n(Z)}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\} + \lambda_n J(\widehat{\pi}_n)\\
&\quad + \mathbb{E}_n\left\{\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z,p)\frac{K((p-\widehat{\pi}_n(Z))/h)}{h}dp + \frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{P-\widehat{\pi}_n(Z)}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\}\\
&\quad - \mathbb{E}_n\left\{\int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\widehat{\pi}_n(Z))/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{P-\widehat{\pi}_n(Z)}{h})(R-Q(Z,P))\right\}\\
&\leq \mathbb{E}_n\left\{\int_{p_1}^{p_2} Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(R-Q(Z,P))\right\}\\
&\quad - \mathbb{E}_n\left\{\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp + \frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\}\\
&\quad + \mathbb{E}_n\left\{\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z,p)\frac{K((\widehat{\pi}_n(Z)-p)/h)}{h}dp + \frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{\widehat{\pi}_n(Z)-P}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\}\\
&\quad - \mathbb{E}_n\left\{\int_{p_1}^{p_2} Q(Z,p)\frac{K((\widehat{\pi}_n(Z)-p)/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{\widehat{\pi}_n(Z)-p}{h})(R-Q(Z,P))\right\},
\end{aligned}$$

where the last inequality is given by the optimization property in (17). In the following, we bound right hand side of the above inequality. It suffices to focus on the first two terms on the right hand side while the other two terms can be bounded similarly.

Specifically, we consider bounding the following term, defined as

$$\begin{aligned}
E_1 \triangleq \mathbb{E}_n&\left\{\int_{p_1}^{p_2} Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(R-Q(Z,P))\right\}\\
&- \mathbb{E}_n\left\{\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp + \frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\}
\end{aligned}$$

We remark that we can write

$$\int_{p_1}^{p_2} Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(R-Q(Z,P))$$
$$= \int_{p_1}^{p_2} \underbrace{Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h} + \frac{\mathbb{1}(P=p)}{hf(p|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-p}{h})(R-Q(Z,p))}_{E_1(p)}dp,$$

where $\mathbb{1}(P=p)$ is indeed a Dirac measure. For a fix p, it can be seen that

$$\begin{aligned}
E_1(p) &= \frac{1}{nh}\sum_{i=1}^{n}(1-\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)})(\widehat{Q}^{-m(i)}(Z_i,P_i)-Q(Z_i,p))K(\frac{\widehat{\pi}_n(Z_i)-p}{h}) \\
&+ \frac{1}{nh}\sum_{i=1}^{n}(\frac{\mathbb{1}(P_i=p)}{\widehat{f}^{-m(i)}(p|Z_i)}-\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)})(R_i-Q(Z_i,P_i))K(\frac{\widehat{\pi}_n(Z_i)-p}{h}) \\
&+ \frac{1}{nh}\sum_{i=1}^{n}(\frac{\mathbb{1}(P_i=p)}{\widehat{f}^{-m(i)}(p|Z_i)}-\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)})(\widehat{R}_i-\widehat{Q}^{-m(i)}(Z_i,P_i)-(R_i-Q(Z_i,P_i)))K(\frac{\widehat{\pi}_n(Z_i)-p}{h}) \\
&+ \frac{1}{nh}\sum_{i=1}^{n}\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)}(\widehat{R}_i-R_i)K(\frac{\widehat{\pi}_n(Z_i)-p}{h}) \\
&\triangleq E_2(p)+E_3(p)+E_4(p)+E_5(p).
\end{aligned}$$

In the following, we bound each of the above four terms. For $E_3(p)$, consider

$$\mathcal{G}_{1,\pi} \triangleq \left\{\int_{p_1}^{p_2}(\frac{\mathbb{1}(P=p)}{\widehat{f}^{-(k)}(p|Z)}-\frac{\mathbb{1}(P=p)}{f(p|Z)})(R-Q(Z,P))K(\frac{\pi(Z)-P}{h})dp \,|\, J(\pi)\le\lambda_n^{-1}, \pi\in\Pi_0\right\}.$$

By the sample splitting, we can show that $\mathbb{E}\left[R-Q(Z,P)|Z,P,f^{-(m(i))}(p|Z)\right]=0$. Therefore we can observe that $\mathbb{E}[g_\pi]=0$ for any $g_\pi\in\mathcal{G}_{1,\pi}$. In addition, the envelop function of $\mathcal{G}_1$, defined as G_1, is proportional to $\int_{p_1}^{p_2}|\frac{\mathbb{1}(P=p)}{\widehat{f}^{-(k)}(p|Z)}-\frac{\mathbb{1}(P=p)}{f(p|Z)}||R-Q(Z,P)|\lambda_n^{-\frac{1}{2}}/hdp$ by the Lipschitz boundness on K in Assumption 2(a). Therefore $\|G_1\|_{2,P}\lesssim n^{-\beta}\lambda_n^{-\frac{1}{2}}/h$ by the error bound condition on $\widehat{f}^{-(m)}(p|Z)$ given in Assumption 6(b). By the entropy condition in Assumption 5 and Lipschitz property of K in Assumption 2(a), we can further show that

$$\sup_{\widetilde{Q}} N(\mathcal{G}_{1,\pi},\widetilde{Q},\varepsilon\|G_1\|_{2,\widetilde{Q}}) \lesssim \left(\frac{1}{\varepsilon}\right)^{v},$$

which implies that

$$J(1,\mathcal{G}_{1,\pi},G_1)\triangleq\int_0^1\sup_{\widetilde{Q}}\sqrt{\log N(\mathcal{G}_{1,\pi},\widetilde{Q},\varepsilon\|G_1\|_{2,\widetilde{Q}})}d\varepsilon\lesssim\sqrt{v}.$$

By leveraging the maximal inequality in the empirical process theory, we can show that

$$\mathbb{E}\sup_{g\in\mathcal{G}_{1,\pi}}|\mathbb{E}_n g|\lesssim\sqrt{v}n^{-\frac{1}{2}}n^{-\beta}\lambda_n^{-\frac{1}{2}}/h.$$

Then by Talagrand's inequality, we can show with probability $1-e^{-x}$,

$$\begin{aligned}
\int_{p_1}^{p_2}E_3(p)dp &\lesssim \frac{1}{h}\left\{\mathbb{E}\sup_{g\in\mathcal{G}_{1,\pi}}|\mathbb{E}_n g| + 2\sqrt{x}\sqrt{\frac{4\sqrt{v}n^{-\frac{1}{2}-\beta}\lambda_n^{-1}/h^2+C_0 n^{-2\beta}\lambda_n^{-1}/h^2}{n}}+\frac{3x\lambda_n^{-\frac{1}{2}}}{nh}\right\} \\
&\lesssim \max\{1,x\}\sqrt{v}n^{-\frac{1}{2}}n^{-\beta}\lambda_n^{-\frac{1}{2}}/h^2.
\end{aligned}$$

Similarly, we can show

$$\int_{p_1}^{p_2}E_2(p)dp\lesssim\max\{1,x\}\sqrt{v}n^{-\frac{1}{2}}n^{-\alpha}\lambda_n^{-\frac{1}{2}}/h^2,$$

with probability at least $1-e^{-x}$. In addition, we can bound $\int_{p_1}^{p_2} E_4(p)dp$ term by Cauchy-Schwarz inequality, i.e., with probability at least $1-2e^{-x}$,

$$\begin{aligned}\int_{p_1}^{p_2} E_4(p)dp &\leq 1/h^2\left(\mathbb{E}_n\left[\frac{1}{\widehat{f}^{-(m)}(P|Z)}-\frac{1}{f(P|Z)}\right]^2\right)^{\frac{1}{2}}\times\left(\mathbb{E}_n\left[\widehat{R}-\widehat{Q}^{-m(i)}(Z,P)-(R-Q(Z,P))\right]^2\right)^{\frac{1}{2}}\lambda_n^{-\frac{1}{2}}\\ &\leq 1/h^2\left(\mathbb{E}_n\left[\frac{1}{\widehat{f}^{-(m)}(P|Z)}-\frac{1}{f(P|Z)}\right]^2\right)^{\frac{1}{2}}\times\left\{\left(\mathbb{E}_n\left[\widehat{Q}^{-m(i)}(Z,P)-Q(Z,P)\right]^2\right)^{\frac{1}{2}}+n^{-\delta}\right\}\lambda_n^{-\frac{1}{2}}\\ &\lesssim \max\{1,x\}\left(n^{-(\alpha+\beta)}+n^{-\beta}n^{-\delta}\right)\lambda_n^{-\frac{1}{2}}/h^2.\end{aligned}$$

The last inequality is due to Bernstein's inequality, i.e.,

$$\mathbb{E}_n\left[\frac{1}{\widehat{f}^{-(k)}(p|Z)}-\frac{1}{f(p|Z)}\right]^2 \lesssim n^{-2\beta}+n^{-\frac{1}{2}-\beta}\sqrt{2x}+\frac{x}{3n},$$

and

$$\mathbb{E}_n\left[\widehat{Q}^{-m(i)}(Z,P)-Q(Z,P)\right]^2 \lesssim n^{-2\beta}+n^{-\frac{1}{2}-\beta}\sqrt{2x}+\frac{x}{3n},$$

by the uniformly bounded assumption in Assumptions 4 and 6(b) and the error bound condition on nuisance function estimation in Assumption 6(b).

For the last term $\int_{p_1}^{p_2} E_5(p)dp$, we can show that with probability at least $1-e^{x}-\varepsilon$,

$$\begin{aligned}\int_{p_1}^{p_2} E_5(p)dp &= \frac{1}{nh}\sum_{i=1}^{n}\frac{1}{f(P_i|Z_i)}(\widehat{R}_i-R_i)K(\frac{\widehat{\pi}_n(Z_i)-p}{h})\\ &\lesssim C_5(\varepsilon)\frac{1}{nh^2}\sum_{i=1}^{n}\mathbb{1}(\Delta_i=0)n^{-\delta}\lambda_n^{-1/2}\\ &\lesssim C_5(\varepsilon)\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\left(\mathbb{P}(\Delta=0)+\sqrt{\frac{x}{n}}\right).\end{aligned}$$

Combining the results above together, we can show that with probability at least $1-5e^{-x}-\varepsilon$,

$$\begin{aligned}\int_{p_1}^{p_2} E_1(p)dp &\lesssim \max\{1,x\}\sqrt{v}n^{-\frac{1}{2}}n^{-\min(\beta,\alpha)}\lambda_n^{-\frac{1}{2}}/h^2\\ &\quad+\max\{1,x\}n^{-(\alpha+\beta)}\lambda_n^{-\frac{1}{2}}/h^2+C_5(\varepsilon)\max\{1,x\}\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\mathbb{P}(\Delta=0).\end{aligned}$$

Similar results can be obtained if we replace $\widehat{\pi}_n$ by $\pi_h^{\lambda_n}$ in E_1. Then we have

$$\begin{aligned}\mathbb{E}_n(g_{\widehat{\pi}_n}) &\lesssim \max\{1,x\}\sqrt{v}n^{-\frac{1}{2}}n^{-\min(\beta,\alpha)}\lambda_n^{-\frac{1}{2}}/h^2\\ &\quad+\max\{1,x\}n^{-(\alpha+\beta)}\lambda_n^{-\frac{1}{2}}/h^2+C_5(\varepsilon)\max\{1,x\}\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\mathbb{P}(\Delta=0),\end{aligned}$$

with probability $1-10\exp(-x)-2\varepsilon$.

Step 2: Again by applying Talagrand's inequality and maximal inequality, we can similarly show that with probability at least $1-e^{-x}$,

$$\sup_{g_\pi\in\mathcal{G}_\pi}|\mathbb{E}_n(g_\pi)-\mathbb{E}(g_\pi)| \lesssim \max\{1,x\}\sqrt{v}\lambda_n^{-\frac{1}{2}}n^{-\frac{1}{2}}/h^2.$$

Summarizing Steps 1 and 2, we can show that with probability $1-e^{-x}-\varepsilon$,

$$\begin{aligned}\mathbf{Regret}(\widehat{\pi}_n) =& V(\pi^*) - V(\widehat{\pi}_n)\\ &\lesssim \Lambda(\lambda_n) + 2C_3 h + \max\{1,x\}\sqrt{v}\lambda_n^{-\frac{1}{2}} n^{-\frac{1}{2}}/h^2\\ &\quad + \max\{1,x\}\sqrt{v} n^{-\frac{1}{2}} n^{-\min(\beta,\alpha)}\lambda_n^{-\frac{1}{2}}/h^2\\ &\quad + \max\{1,x\} n^{-(\alpha+\beta)}\lambda_n^{-\frac{1}{2}}/h^2 + C_5(\varepsilon)\max\{1,x\}\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\mathbb{P}(\Delta=0).\end{aligned}$$

which concludes our proof. **Q.E.D.**

Proof of Theorem 3 For any $0<\delta<1$, with probability at least $1-\delta$, we have

$$\begin{aligned}&\mathbf{Regret}(\widetilde{\pi}_n^\delta) = V(\pi^*) - V(\widetilde{\pi}_n) = \mathbb{E}_{X,Y}\left[Q\left(X,Y,\pi^*(X,Y)\right) - Q\left(X,Y,\widetilde{\pi}_n^\delta(X,Y)\right)\right]\\ \leq& \mathbb{E}_{X,Y}\left[\widehat{Q}\left(X,Y,\pi^*(X,Y)\right) + U\left(X,Y,\pi^*(X,Y),\delta\right) - \widehat{Q}\left(X,Y,\widetilde{\pi}_n^\delta(X,Y)\right) + U\left(X,Y,\widetilde{\pi}_n^\delta(X,Y),\delta\right)\right] \quad (29)\\ \leq& \mathbb{E}_{X,Y}\left[\widehat{Q}\left(X,Y,\pi^*(X,Y)\right) + U\left(X,Y,\pi^*(X,Y),\delta\right) - \widehat{Q}\left(X,Y,\pi^*(X,Y)\right) + U\left(X,Y,\pi^*(X,Y),\delta\right)\right] \quad (30)\\ =& 2\mathbb{E}_{X,Y}\left[U\left(X,Y,\pi^*(X,Y),\delta\right)\right],\end{aligned}$$

where (29) is due to Assumption 7 and (30) is from the definition of $\widetilde{\pi}_n$. Now we define for any $\delta\in(0,1)$,

$$W\left(\pi^*\right) = \mathbb{E}_{X,Y}\left[U\left(X,Y,\pi^*(X,Y),\delta\right)\right],\; W_h\left(\pi^*\right) = \mathbb{E}\left[\frac{U(X,Y,P,\delta)K\left(\frac{\pi^*(X,Y)-P}{h}\right)}{hf(P\mid X,Y)}\right],$$

where $\mathbb{E}$ means the expectation over (X,Y,P). Then we have

$$\begin{aligned}W_h\left(\pi^*\right) =& \mathbb{E}\left[\frac{U(X,Y,p,\delta)K\left(\frac{p-\pi^*(X,Y)}{h}\right)}{hf(P\mid X,Y)}\right] = \mathbb{E}_{X,Y}\left[\int_{\mathcal{P}}\frac{U(X,Y,p,\delta)K\left(\frac{p-\pi^*(X,Y)}{h}\right)}{hf(p\mid X,Y)}f(p\mid X,Y)dp\right]\\ =& \mathbb{E}_{X,Y}\left[\int_{\mathcal{P}}\frac{U(X,Y,p,\delta)K\left(\frac{p-\pi^*(X,Y)}{h}\right)}{h}dp\right].\end{aligned}$$

By letting $\frac{p-\pi^*(X,Y)}{h}=t$, we obtain $W_h\left(\pi^*\right) = \mathbb{E}_{X,Y}\left[\int U\left(X,Y,\pi^*(X,Y)+th,\delta\right)K(t)dt\right]$. Therefore,

$$\begin{aligned}&\left|W_h\left(\pi^*\right) - W\left(\pi^*\right)\right|\\ =& \mathbb{E}_{X,Y}\left[\int U\left(X,Y,\pi^*(X,Y)+th,\delta\right)K(t)dt\right] - \mathbb{E}_{X,Y}\left[U\left(X,Y,\pi^*(X,Y),\delta\right)\right]\\ =& \mathbb{E}_{X,Y}\left[\int U\left(X,Y,\pi^*(X,Y)+th,\delta\right)K(t)dt - U\left(X,Y,\pi^*(X,Y),\delta\right)\right]\\ =& \mathbb{E}_{X,Y}\left[\int U\left(X,Y,\pi^*(X,Y)+th,\delta\right) - U\left(X,Y,\pi^*(X,Y),\delta\right)K(t)dt\right]\\ \leq& \mathbb{E}_{X,Y}\left[\sup_{p_l\leq p<p'\leq p_2}\left|\frac{U(X,Y,p,\delta)-U\left(X,Y,p',\delta\right)}{p-p'}\right|\int |th|K(t)dt\right]\\ \leq& C_8(\delta)h\int |t|K(t)dt = C_9(\delta)h, \quad (31)\end{aligned}$$

where $C_9(\delta) := C_1C_8(\delta)$. The first inequality is by Assumption 9 and the second inequality is by Assumption 2a. Therefore, we have with probability at least $1-\delta$,

$$\mathbf{Regret}(\widetilde{\pi}_n) \leq 2W\left(\pi^*\right) \leq 2C_9(\delta)h + 2W_h\left(\pi^*\right) \leq 2C_9(\delta)h + 2\left\|\frac{U(X,Y,P,\delta)K\left(\frac{\pi^*(X,Y)-P}{h}\right)}{hf(P\mid X,Y)}\right\|_2$$

$$\leq 2C_9(\delta)h + 2\left\|U(X,Y,P,\delta)\right\|_2 \left\| \frac{K\left(\frac{\pi^*(X,Y)-P}{h}\right)}{hf(P\mid X,Y)} \right\|_2,$$

where the first inequality is due to (31), the second inequality is by definition and the last inequality is by Holder's inequality. By Assumption 8, for any h,

$$\mathbf{Regret}(\widetilde{\pi}_n) \leq 2C_9(\delta)h + \frac{2C_7\|U(X,Y,P,\delta)\|_2}{h}.$$

So we have

$$\mathbf{Regret}(\widetilde{\pi}_n) \leq \inf_h \left\{ 2C_9(\delta)h + \frac{2C_7\|U(X,Y,P,\delta)\|_2}{h} \right\} = C_{10}(\delta)\sqrt{\|U(X,Y,P,\delta)\|_2},$$

where the constant $C_{10}(\delta) := 4\sqrt{C_9(\delta)C_7}$. **Q.E.D.**

Appendix C: Dependent Data Scenario and Its Analysis

In this section, we consider the scenario where $\{(X_i, Y_i, P_i, S_i, \Delta_i)\}_{1\leq i\leq n}$ are not i.i.d. copies. Instead, we assume our data come from M centers, where data collected at each center are dependent and cross the center are independent. Specifically, for center $1 \leq m \leq M$, our offline data consist of $\left\{(X_t^{(m)}, Y_t^{(m)}, P_t^{(m)}, S_t^{(m)}, \Delta_t^{(m)})\right\}_{1\leq t\leq n}$. Since data across different centers are independent, we can apply the previously proposed method (17) to learn the optimal pricing strategy that

$$\begin{aligned}\widehat{\pi}_n \in \arg\max_{\pi\in\Pi_0} \Bigg\{ & \frac{1}{nMh}\sum_{i=1}^{nM}\int_{p_1}^{p_2} \widehat{Q}^{(-m(i))}(X_i,Y_i,p) K\left(\frac{p-\pi(X_i,Y_i)}{h}\right) dp \\ & + \frac{1}{nM}\sum_{i=1}^{nM} \frac{1}{h\widehat{f}^{(-m(i))}(P_i|X_i,Y_i)} K\left(\frac{P_i-\pi(X_i,Y_i)}{h}\right)(\widehat{R}_i - \widehat{Q}^{(-m(i))}(X_i,Y_i,P_i)) - \lambda_n J(\pi) \Bigg\},\end{aligned} \tag{32}$$

where $m(i)$ denotes the center containing the i-th observation.

In the following, we provide a theoretical guarantee for our approach under the non-i.i.d. case. For any two σ-fields $\mathcal{A}$ and $\mathcal{B} \subset \mathcal{F}$, we define $\beta(\mathcal{A},\mathcal{B}) := \sup \frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)|$ where the supremum is taken over all pairs of (finite) partitions $\{A_1,\ldots,A_I\}$ and $\{B_1,\ldots,B_J\}$ of Ω such that $A_i \in \mathcal{A}$ for each i and $B_j \in \mathcal{B}$ for each j. We also define the σ-field for the sequence of random variables $X := (X_k, k \in \mathbb{Z})$. For each $n \geq 1$, the dependence coefficient (β-mixing coefficient) is defined as $\beta(n) := \sup_{j\in\mathbb{Z}} \beta\left(\mathcal{F}_{-\infty}^{j}, \mathcal{F}_{j+n}^{\infty}\right)$. Then the random sequence X is said to be β-mixing if $\beta(n) \to 0$ as $n \to \infty$. We now make the following assumption to characterize the dependency among observations in each center.

ASSUMPTION 10. *For each $1 \leq m \leq M$, the stochastic process $\left\{(X_t^{(m)}, Y_t^{(m)}, P_t^{(m)}, S_t^{(m)}, \Delta_t^{(m)})\right\}_{t\geq 1}$ is a stationary and exponential β-mixing process with β-mixing coefficient $\beta(t) \leq \beta_0 \exp(-\beta_1 t)$ for some $\beta_0 \geq 0$ and $\beta_1 > 0$.*

Assumption 10 characterizes the dependence among observations in each center. An exponential β-mixing process is a type of stochastic process that satisfies certain conditions regarding the dependence structure of its random variables. Specifically, a process is β-mixing if the correlation between two variables decreases exponentially fast as the time separation between them increases. The β-mixing coefficient is a measure of the rate at which the correlation between two variables decays as the time separation between them increases. A smaller β-mixing coefficient indicates a faster decay of correlation and hence a weaker dependence structure. In particular, the upper bound on the β-mixing coefficient at the time lag t basically means that the dependence decays to 0 at the slowest possible exponential rate with respect to t. See Bradley (2005) for more details. Without loss of generality, we assume $M = 2$. Then we have the following regret guarantee for our estimated optimal pricing strategy given in (32).

THEOREM 4. *Suppose that Assumptions 1–10 hold. If $\lambda_n \le 1$ and $\alpha+\beta > 1/2$, then for any $\varepsilon \in (0,1)$ with probability at least $1-1/n-\varepsilon$, Algorithm 1 admits the following regret upper bound*

$$\mathbf{Regret}(\widehat{\pi}_n) \lesssim \Lambda(\lambda_n) + 2C_3 h + \log(n)\sqrt{v}\lambda_n^{-\frac{1}{2}} n^{-\frac{1}{2}}/h^2 + C_5(\varepsilon)\log(n)\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\mathbb{P}(\Delta = 0), \tag{33}$$

where the regret is defined in (19). Furthermore, if (21) holds, then with probability at least $1-1/n-\varepsilon$

$$\mathbf{Regret}(\widehat{\pi}_n) \lesssim \log(n) n^{-\frac{\zeta \min(\frac{1}{2},\delta)}{6\zeta+1}}. \tag{34}$$

Under Assumption 10, one can obtain a similar regret result as that in Theorem 2.

LEMMA 5. *Suppose that a stochastic process $\{Z_t\}_{t\ge 1}$ is a stationary and exponential β-mixing process with β-mixing coefficient $\beta(q) \le \beta_0 \exp(-\beta_1 q)$ for some $\beta_0 \ge 0$ and $\beta_1 > 0$. Let $\mathcal{G}$ be a class of measurable functions that take Z_t as input. For any $g \in \mathcal{G}$, assume $\mathbb{E}[g(Z_t)] = 0$ for any $t \ge 0$. Suppose that the envelop function of $\mathcal{G}$ is uniformly bounded by some constant $C > 0$. In addition, if $\mathcal{G}$ belongs to the class of VC-typed functions such that $\sup_{\widetilde{Q}} \mathcal{N}(\mathcal{G}, \widetilde{Q}, \epsilon \| \bullet \|_{\widetilde{Q},2}) \lesssim (1/\epsilon)^{\alpha}$ for a constant $\alpha > 0$. Then with a probability at least $1-1/T$,*

$$\sup_{g\in\mathcal{G}} \left|\frac{1}{T}\sum_{t=1}^{T} g(Z_{i,t})\right| \lesssim \log(T)\sqrt{\frac{\alpha}{T}}.$$

If $\sigma^2 = \sup_{g\in\mathcal{G}} \mathbb{E}[g^2(Z_t)]$ for $1 \le t \le T$, then with probability at least $1-1/T$,

$$\sup_{g\in\mathcal{G}} \left|\frac{1}{T}\sum_{t=1}^{T} g(Z_{i,t})\right| \lesssim \frac{\log(T)(\sqrt{\log(T)\alpha}\|G\|_2 + \sqrt{T\sigma^2} + 1)}{T}.$$

Proof of Lemma 5. To prove the case (i) of the lemma, we focus on bounding $\sup_{g\in\mathcal{G}} |\sum_{t=1}^{T} g(Z_t)|$. Specifically, we apply Berbee's coupling lemma (Berbee 1979) and follow the remark below Lemma 4.1 of Dedecker and Louhichi (2002). Let q be some positive integer. One can always construct a sequence $\{\widetilde{Z}_t\}_{t\ge 1}$ such that with probability at least $1-(T\beta(q))/q$,

$$\sup_{g\in\mathcal{G}} \left|\sum_{t=1}^{T} g(Z_t)\right| = \sup_{g\in\mathcal{G}} \left|\sum_{t=1}^{T} g(\widetilde{Z}_t)\right|.$$

In the same time, the block sequence $\widetilde{X}_k(g) = \{g(\widetilde{Z}_{(k-1)q+j})\}_{1\le j\le q}$ are identically distributed for $k \ge 1$. In addition, the sequence $\{\widetilde{X}_k(g) \mid k = 2\omega, \omega \ge 1\}$ are independent and so are the sequence $\{\widetilde{X}_k(g) \mid k = 2\omega+1, \omega \ge 0\}$. Let $I_r = \{\lfloor T/q \rfloor q + 1, \cdots, T\}$ with Card$(I_r) < q$. Then we can show that with probability at least $1-(T\beta(q))/q$,

$$\sup_{g\in\mathcal{G}} \left|\sum_{t=1}^{T} g(Z_t)\right| \le \sup_{g\in\mathcal{G}} \left|\sum_{t=1}^{q\lfloor T/q \rfloor} g(\widetilde{Z}_t)\right| + \sup_{g\in\mathcal{G}} \left|\sum_{t\in I_r} g(Z_t)\right|.$$

In the following, we always assume that the above inequality holds. Then it is sufficient to bound each of the above two terms separately. First without loss of generality, we assume $\lfloor T/q \rfloor$ is an even number. For the first term, we have

$$\sup_{g\in\mathcal{G}} \left|\sum_{t=1}^{q\lfloor T/q \rfloor} g(\widetilde{Z}_t)\right| \le \sum_{j=1}^{2q} \sup_{g\in\mathcal{G}} \left|\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)\right|.$$

By the previous construction, $\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)|$ is a suprema empirical process of i.i.d. sequences. Then by conditions in Lemma 5 and Mcdiarmid's inequality, we have with probability at least $1-\varepsilon$,

$$\sup_{g\in\mathcal{G}} \left|\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)\right| \lesssim \mathbb{E}\left[\sup_{g\in\mathcal{G}} \left|\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)\right|\right] + \sqrt{\frac{T\log(1/\varepsilon)}{q}}.$$

Given the condition that $\sup_{\widetilde{Q}} \mathcal{N}(\mathcal{G}, \widetilde{Q}, \epsilon \| \bullet \|_{\widetilde{Q},2}) \lesssim (1/\epsilon)^{\alpha}$. By a standard maximal inequality using uniform entropy integral (e.g., Van Der Vaart and Wellner 2011), we can show that with probability at least $1-\varepsilon$,

$$\mathbb{E}\left[\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)|\right] \lesssim \sqrt{\frac{\alpha T}{q}}.$$

By letting $\varepsilon = 1/T$, we can show that with probability at least $1-1/T$,

$$\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)| \lesssim \sqrt{\frac{\alpha T}{q}} + \sqrt{\frac{T\log(T)}{q}}.$$

Next, we can bound $\sup_{g\in\mathcal{G}} |\sum_{t\in I_r} g(Z_t)|$ by Cq. By letting $q \asymp \log(T)$, with probability at least $1-1/T$,

$$\sup_{g\in\mathcal{G}} |\sum_{t=1}^{T} g(Z_t)| \lesssim \log(T)\sqrt{\frac{\alpha T}{\log(T)}} + \log(T)\sqrt{\frac{T\log(T)}{\log(T)}} + \log(T) \lesssim \log(T)\sqrt{T\alpha}.$$

This concludes our proof of case (i) by dividing both sides by T.

In the second part of our proof, we have $\sigma^2 = \sup_{g\in\mathcal{G}} \mathbb{E}[g^2(Z_t)]$ for $1 \le t \le T$. Then by conditions in Lemma 5 and Talagrand's inequality, we have with probability at least $1-\varepsilon$,

$$\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)| \lesssim \mathbb{E}\left[\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)|\right] + \sqrt{2\eta_n \log(1/\varepsilon)} + C\log(1/\varepsilon),$$

where $\eta_n = 2T/q\mathbb{E}\left[\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)|\right] + T/q\sigma^2$. We can deploy another maximal inequality to show that

$$\mathbb{E}\left[\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)|\right] \lesssim J(1, \mathcal{G}, G)\|G\|_2 \lesssim \sqrt{\alpha}\|G\|_2.$$

by letting $\varepsilon = 1/T$, we can show that with probability at least $1-1/T$,

$$\sup_{g\in\mathcal{G}} |\sum_{k=1}^{\lfloor T/q \rfloor/2} g(\widetilde{Z}_i)| \lesssim \sqrt{\alpha}\|G\|_2 + \sqrt{\left(\sqrt{\alpha}\|G\|_2 + \frac{T}{q}\sigma^2\right)\log(T)} + C\log(T).$$

By letting $q \asymp \log(T)$, we can show that with probability at least $1-1/T$,

$$\begin{aligned}\sup_{g\in\mathcal{G}} |\sum_{t=1}^{T} g(Z_t)| &\lesssim \log(T)\sqrt{\alpha}\|G\|_2 + \log(T)\sqrt{\left(\sqrt{\alpha}\|G\|_2 + \frac{T}{\log(T)}\sigma^2\right)\log(T)} + \log(T) \\ &\lesssim \log(T)(\sqrt{\log(T)\alpha}\|G\|_2 + \sqrt{T\sigma^2} + 1)\end{aligned}$$

The result follows by dividing both sides by T. **Q.E.D.**

Proof of Theorem 4. For notational simplicity, let $Z_t = (X_t, Y_t)$ for $1 \le t \le n$ and $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$. We further let $U(\pi) = -V(\pi)$, $U_h(\pi) = -V_h(\pi)$. By similar derivation, we can show that

$$V(\pi^*) - V(\widehat{\pi}_n) \le \Lambda(\lambda_n) + \underbrace{U_h(\widehat{\pi}) + \lambda_n J(\widehat{\pi}) - \{U_h(\pi_h^{\lambda_n}) + \lambda_n J(\pi_h^{\lambda_n})\}}_{(I)} + 2C_3 h,$$

where $\pi_h^{\lambda_n} \in \arg\min_{\pi\in\Pi_0}\{U_h(\pi) + \lambda_n J(\pi)\}$. In the following, we apply Lemma 5 to bound Term (I) on the right hand side of the inequality above. Let

$$\begin{aligned}\mathcal{G}_\pi \triangleq \Bigg\{ &\int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\pi_h^{\lambda_n}(Z))/h)}{h}dp + \frac{1}{hf(P|Z)}K(\frac{P-\pi_h^{\lambda_n}(Z)}{h})(R-Q(Z,P)) + \lambda_n J(\pi) \\ &- \int_{p_1}^{p_2} Q(Z,p)\frac{K((p-\pi(Z))/h)}{h}dp - \frac{1}{hf(P|Z)}K(\frac{P-\pi(Z)}{h})(R-Q(Z,P)) - \lambda_n J(\pi_h^{\lambda_n}) \,|\, J(\pi) \lesssim \lambda_n^{-1}, \pi \in \Pi_0 \Bigg\}.\end{aligned}$$

Recall that $J(\pi)=\|\pi\|_{\Pi_0}^2$. We consider a constraint class on π based on the same argument in Theorem 2. In the first step, we show $\mathbb{E}_n(g_{\widehat{\pi}_n})\le\varepsilon_1$, for some $\varepsilon_1>0$ with a high probability. In the second step, we aim to show that $\sup_{g_\pi\in\mathcal{G}_\pi}|\mathbb{E}_n(g_\pi)-\mathbb{E}(g_\pi)|\le\varepsilon_2$, with a high probability for some ε_2. Then combining two, we are able to show $(I)\le\varepsilon_1+\varepsilon_2$ with some high probability.

Step 1: We can similarly derive that

$$\begin{aligned}
&\mathbb{E}_n(g_{\widehat{\pi}_n})\\
&\le\mathbb{E}_n\left\{\int_{p_1}^{p_2}Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp+\frac{1}{hf(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(R-Q(Z,P))\right\}\\
&\quad-\mathbb{E}_n\left\{\int_{p_1}^{p_2}\widehat{Q}^{(-m(i))}(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp+\frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P))\right\}\\
&\quad+\mathbb{E}_n\left\{\int_{p_1}^{p_2}\widehat{Q}^{(-m(i))}(Z,p)\frac{K((\widehat{\pi}_n(Z)-p)/h)}{h}dp+\frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{\widehat{\pi}_n(Z)-P}{h})(\widehat{R}-\widehat{Q}^{(-m(i))}(Z,P))\right\}\\
&\quad-\mathbb{E}_n\left\{\int_{p_1}^{p_2}Q(Z,p)\frac{K((\widehat{\pi}_n(Z)-p)/h)}{h}dp+\frac{1}{hf(P|Z)}K(\frac{\widehat{\pi}_n(Z)-p}{h})(R-Q(Z,P))\right\},
\end{aligned}$$

In the following, we bound the right-hand side of the above inequality. It suffices to focus on the first two terms on the right-hand side while the other two terms can be bounded similarly.

Specifically, we consider bounding the following term, defined as

$$\begin{aligned}
E_1\triangleq&\,\mathbb{E}_n\left\{\int_{p_1}^{p_2}Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp+\frac{1}{hf(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(R-Q(Z,P))\right\}\\
&-\mathbb{E}_n\left\{\int_{p_1}^{p_2}\widehat{Q}^{(-m(i))}(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp+\frac{1}{h\widehat{f}^{(-m(i))}(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(\widehat{R}-\widehat{Q}^{(-m(i))})(Z,P)\right\}
\end{aligned}$$

We again notice that

$$\begin{aligned}
&\int_{p_1}^{p_2}Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}dp+\frac{1}{hf(P|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-P}{h})(R-Q(Z,P))\\
&=\int_{p_1}^{p_2}\underbrace{Q(Z,p)\frac{K((\pi_h^{\lambda_n}(Z)-p)/h)}{h}+\frac{\mathbb{1}(P=p)}{hf(p|Z)}K(\frac{\pi_h^{\lambda_n}(Z)-p}{h})(R-Q(Z,p))}_{E_1(p)}dp,
\end{aligned}$$

where $\mathbb{1}(P=p)$ is indeed a Dirac measure. For a fix p, it can be seen that

$$\begin{aligned}
E_1(p)=&\frac{1}{nh}\sum_{i=1}^n(1-\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)})(\widehat{Q}^{-m(i)}(Z_i,P_i)-Q(Z_i,p))K(\frac{\widehat{\pi}_n(Z_i)-p}{h})\\
&+\frac{1}{nh}\sum_{i=1}^n(\frac{\mathbb{1}(P_i=p)}{\widehat{f}^{-m(i)}(p|Z_i)}-\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)})(R_i-Q(Z_i,P_i))K(\frac{\widehat{\pi}_n(Z_i)-p}{h})\\
&+\frac{1}{nh}\sum_{i=1}^n(\frac{\mathbb{1}(P_i=p)}{\widehat{f}^{-m(i)}(p|Z_i)}-\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)})(\widehat{R}_i-\widehat{Q}^{-m(i)}(Z_i,P_i)-(R_i-Q(Z_i,P_i)))K(\frac{\widehat{\pi}_n(Z_i)-p}{h})\\
&+\frac{1}{nh}\sum_{i=1}^n\frac{\mathbb{1}(P_i=p)}{f(p|Z_i)}(\widehat{R}_i-R_i)K(\frac{\widehat{\pi}_n(Z_i)-p}{h})\\
\triangleq&\,E_2(p)+E_3(p)+E_4(p)+E_5(p).
\end{aligned}$$

In the following, we bound each of the above four terms. For $E_3(p)$,consider

$$\mathcal{G}_{1,\pi}\triangleq\left\{\int_{p_1}^{p_2}(\frac{\mathbb{1}(P=p)}{\widehat{f}^{-(k)}(p|Z)}-\frac{\mathbb{1}(P=p)}{f(p|Z)})(R-Q(Z,P))K(\frac{\pi(Z)-P}{h})dp\,|\,J(\pi)\le\lambda_n^{-1},\pi\in\Pi_0\right\}.$$

By the problem setting and the independence across centers, we can show that $\mathbb{E}\left[R - Q(Z,P)|Z,P,f^{-(m(i))}(p|Z)\right] = 0$. Therefore we can observe that $\mathbb{E}[g_\pi] = 0$ for any $g_\pi \in \mathcal{G}_{1,\pi}$. In addition, the envelop function of $\mathcal{G}_1$, defined as G_1, is proportional to $\int_{p_1}^{p_2} |\frac{\mathbb{1}(P=p)}{\widehat{f}^{-(k)}(p|Z)} - \frac{\mathbb{1}(P=p)}{f(p|Z)}||R - Q(Z,P)|\lambda_n^{-\frac{1}{2}}/h dp$ by the Lipschitz boundness on K in Assumption2(a). Therefore $\|G_1\|_{2,P} \lesssim n^{-\beta}\lambda_n^{-\frac{1}{2}}/h$ by the error bound condition on $\widehat{f}^{-(m)}(p|Z)$ given in Assumption 6(b). By the entropy condition in Assumption 5 and Lipschitz property of K in Assumption 2(a), we can further show that

$$\sup_{\widetilde{Q}} N(\mathcal{G}_{1,\pi}, \widetilde{Q}, \varepsilon \|G_1\|_{2,\widetilde{Q}}) \lesssim \left(\frac{1}{\varepsilon}\right)^v,$$

which implies that

$$J(1, \mathcal{G}_{1,\pi}, G_1) \triangleq \int_0^1 \sup_{\widetilde{Q}} \sqrt{\log N(\mathcal{G}_{1,\pi}, \widetilde{Q}, \varepsilon\|G_1\|_{2,\widetilde{Q}})} d\varepsilon \lesssim \sqrt{v}.$$

By leveraging the result in Lemma 5, we can show that with probability at least $1 - 1/n$,

$$\int_{p_1}^{p_2} E_3(p)dp \lesssim \log(n)\sqrt{v}n^{-\frac{1}{2}}n^{-\beta}\lambda_n^{-\frac{1}{2}}/h^2.$$

Similarly, we can show

$$\int_{p_1}^{p_2} E_2(p)dp \lesssim \log(n)\sqrt{v}n^{-\frac{1}{2}}n^{-\alpha}\lambda_n^{-\frac{1}{2}}/h^2,$$

with probability at least $1 - 1/n$. In addition, we can bound $\int_{p_1}^{p_2} E_4(p)dp$ term by Cauchy-Schwarz inequality, i.e., with probability at least $1 - 1/n$,

$$\begin{aligned}
\int_{p_1}^{p_2} E_4(p)dp &\le 1/h^2 \left(\mathbb{E}_n\left[\frac{1}{\widehat{f}^{-(m)}(P|Z)} - \frac{1}{f(P|Z)}\right]^2\right)^{\frac{1}{2}} \times \left(\mathbb{E}_n\left[\widehat{R} - \widehat{Q}^{-m(i)}(Z,P) - (R - Q(Z,P))\right]^2\right)^{\frac{1}{2}} \lambda_n^{-\frac{1}{2}} \\
&\le 1/h^2 \left(\mathbb{E}_n\left[\frac{1}{\widehat{f}^{-(m)}(P|Z)} - \frac{1}{f(P|Z)}\right]^2\right)^{\frac{1}{2}} \times \left\{\left(\mathbb{E}_n\left[\widehat{Q}^{-m(i)}(Z,P) - Q(Z,P)\right]^2\right)^{\frac{1}{2}} + n^{-\delta}\right\} \lambda_n^{-\frac{1}{2}} \\
&\lesssim \log(n)\left(n^{-(\alpha+\beta)} + n^{-\beta}n^{-\delta}\right)\lambda_n^{-\frac{1}{2}}/h^2.
\end{aligned}$$

The last inequality is due to Bernstein's inequality in the dependent case using the uniformly bounded assumption in Assumptions 4 and 6(b) and the error bound condition on nuisance function estimation in Assumption 6(b). See theorem 8 of Fu et al. (2022) for more details.

For the last term $\int_{p_1}^{p_2} E_5(p)dp$, we can show that with probability at least $1 - \varepsilon$,

$$\begin{aligned}
\int_{p_1}^{p_2} E_5(p)dp &= \frac{1}{nh}\sum_{i=1}^n \frac{1}{f(P_i|Z_i)}(\widehat{R}_i - R_i)K(\frac{\widehat{\pi}_n(Z_i) - p}{h}) \\
&\lesssim C_5(\varepsilon)\frac{1}{nh^2}\sum_{i=1}^n \mathbb{1}(\Delta_i = 0)n^{-\delta}\lambda_n^{-1/2} \\
&\lesssim C_5(\varepsilon)\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\left(\mathbb{P}(\Delta = 0) + \sqrt{\frac{x}{n}}\right).
\end{aligned}$$

Combining the results above together, we can show that with probability at least $1 - 5/n - \varepsilon$,

$$\begin{aligned}
\int_{p_1}^{p_2} E_1(p)dp \lesssim{}& \log(n)\sqrt{v}n^{-\frac{1}{2}}n^{-\min(\beta,\alpha)}\lambda_n^{-\frac{1}{2}}/h^2 \\
&+ \log(n)n^{-(\alpha+\beta)}\lambda_n^{-\frac{1}{2}}/h^2 + C_5(\varepsilon)\log(n)\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\mathbb{P}(\Delta = 0).
\end{aligned}$$

Similar results can be obtained if we replace $\widehat{\pi}_n$ by $\pi_h^{\lambda_n}$ in E_1. Then we have

$$\mathbb{E}_n(g_{\widehat{\pi}_n}) \lesssim \log(n)\sqrt{v}n^{-\frac{1}{2}}n^{-\min(\beta,\alpha)}\lambda_n^{-\frac{1}{2}}/h^2 + \log(n)n^{-(\alpha+\beta)}\lambda_n^{-\frac{1}{2}}/h^2 + C_5(\varepsilon)\log(n)\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\mathbb{P}(\Delta=0),$$

with probability $1-10/n-2\varepsilon$.

Step 2: Again by applying Lemma 5, we can similarly show that with probability at least $1-1/n$,

$$\sup_{g_\pi \in \mathcal{G}_\pi} |\mathbb{E}_n(g_\pi) - \mathbb{E}(g_\pi)| \lesssim \log(n)\sqrt{v}\lambda_n^{-\frac{1}{2}}n^{-\frac{1}{2}}/h^2.$$

Summarizing Steps 1 and 2, we can show that with probability $1-1/n-\varepsilon$,

$$\begin{aligned}\mathbf{Regret}(\widehat{\pi}_n) =& V(\pi^*) - V(\widehat{\pi}_n) \\ \lesssim& \Lambda(\lambda_n) + 2C_3 h + \log(n)\sqrt{v}\lambda_n^{-\frac{1}{2}}n^{-\frac{1}{2}}/h^2 \\ &+ \log(n)\sqrt{v}n^{-\frac{1}{2}}n^{-\min(\beta,\alpha)}\lambda_n^{-\frac{1}{2}}/h^2 \\ &+ \log(n)n^{-(\alpha+\beta)}\lambda_n^{-\frac{1}{2}}/h^2 + C_5(\varepsilon)\log(n)\frac{n^{-\delta}\lambda_n^{-1/2}}{h^2}\mathbb{P}(\Delta=0).\end{aligned}$$

which proves the first statement. The second statement holds the same argument as in Corollary 1. **Q.E.D.**

Appendix D: Numerical Experiment Supplementary

D.1. Robustness Experiments

Figures 4 – 7 are the results of the experiments in Table 1.

D.2. Sensitivity Analysis

The neural network parameters' sensitivity will be instance-dependent. For the instance we use, after experiments with various neural network parameters for the neural networks generating $\widehat{Q}(X,Y,P)$, $\hat{f}(P\,|\,X,Y)$, and $\widehat{\pi}_n$, we find the outputs are not too sensitive to the parameters. For example, Figures 8 – 11 are the results of the same instance obtained using four sets of parameters, where the numbers in the parentheses denote the number of neurons in each hidden layer and the number of scalars denotes the number of hidden layers used for the neural network for estimating each variable:

Set	$\widehat{Q}(X,Y,P)$	$\hat{f}(P\|X,Y)$	$\widehat{\pi}_n$
1	(100,100)	(48)	(12)
2	(100,100)	(48)	(24)
3	(100,100)	(24)	(12)
4	(200,100)	(48)	(12)

Table 5 Neural Network Parameters

D.3. Running Time

The main time-consuming part of the algorithm lies in 1) in each fold, the training of the neural networks used to estimate $\widehat{Q}^{(-m(i))}(X_i,Y_i,p)$, $\widehat{Q}^{(-m(i))}(X_i,Y_i,P_i)$ and $\widehat{f}^{(-m(i))}(P_i|X_i,Y_i)$; 2) the training of the neural network used to estimate $\widehat{\pi}_n$. Note that the number of folds used in cross-validation is usually not too large. In our case, we choose $K=3$. Below we report the running time to train the neural network for each statistic, averaged over 100 instances of sample size 2000 and also 3 folds if in the inner loop, with the value in the parenthesis being the standard deviation. In general, it takes around 457.28 seconds to run our algorithm for an instance with a sample size of 2000.

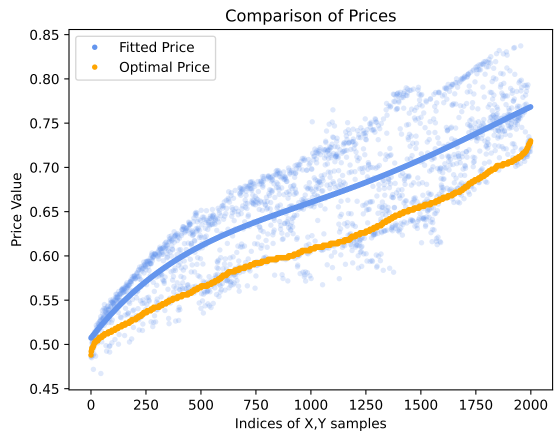

Figure 4 **Instance 1**

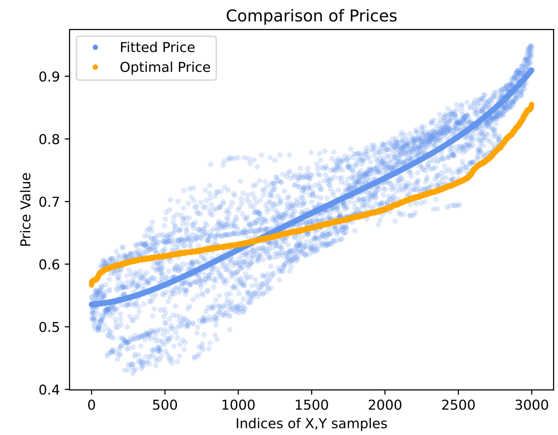

Figure 5 **Instance 2**

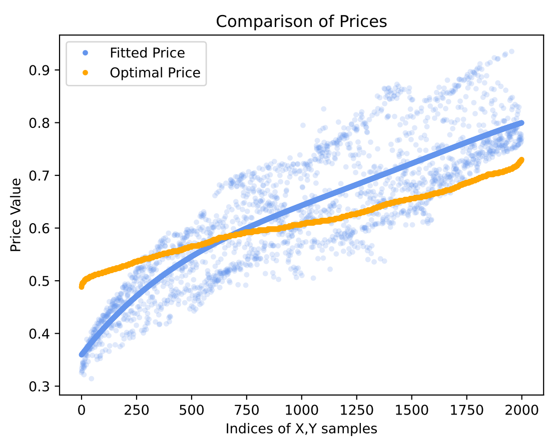

Figure 6 **Instance 3**

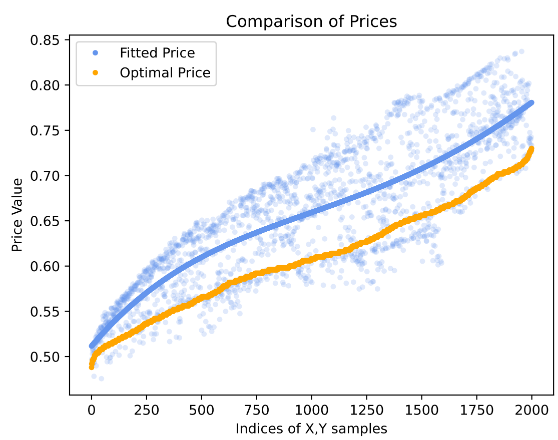

Figure 7 **Instance 4**

Neural Network	$\widehat{Q}^{(-m(i))}(X_i, Y_i, p)$	$\widehat{Q}^{(-m(i))}(X_i, Y_i, P_i)$	$\widehat{f}^{(-m(i))}(P_i \mid X_i, Y_i)$	$\widehat{\pi}_n$
Average Running Time (s)	33.95(13.00)	30.26(12.35)	22.34(1.96)	197.73(16.29)

Table 6 **Running Time**

D.4. Random Survival Forests Description

We briefly describe the random survival forests method (Ishwaran et al. 2008). Define the censoring indicator to be 0 if the data is right-censored and otherwise 1. Given a data set with each record comprising the individual's survival time and the $0-1$ censoring indicator, the random survival forests algorithm consists of the following steps:

1. First, we draw B bootstrap samples from the original data where B is a given parameter. Note that each bootstrap sample excludes on average 37% of the data, called out-of-bag data (OOB data).

2. For each bootstrap sample, grow a survival tree, where p candidate variables randomly selected are used at each node. Then the node is split using the candidate variable that maximizes survival difference between daughter nodes to separate the dissimilar cases, by searching over all possible x variables and split values c, and choosing

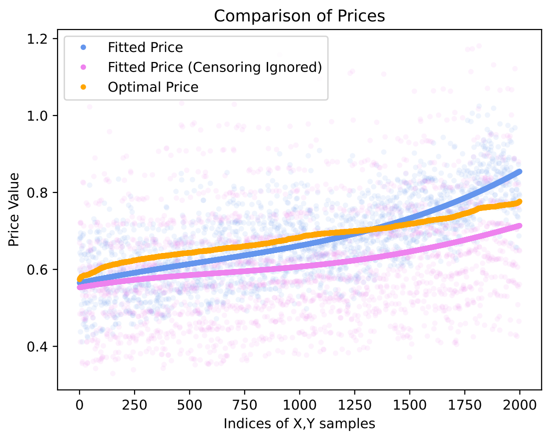

Figure 8 **Set 1**

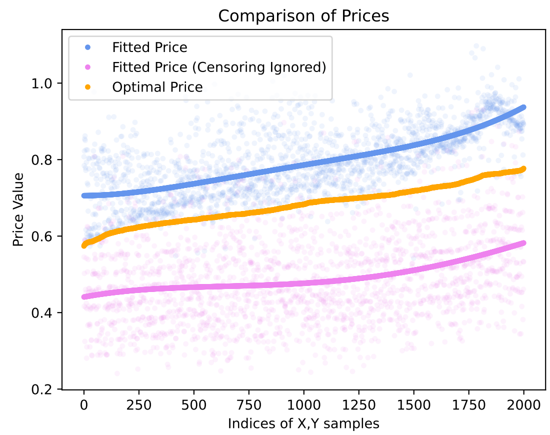

Figure 9 **Set 2**

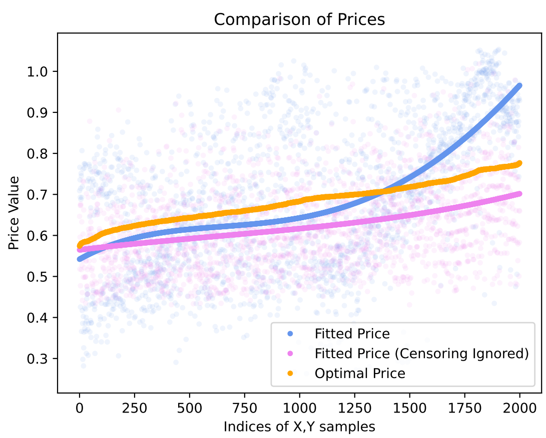

Figure 10 **Set 3**

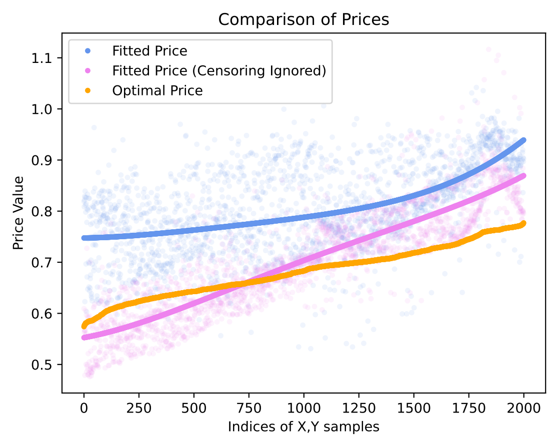

Figure 11 **Set 4**

that x^* and c^* that maximizes survival difference. The tree is grown to a full size such that a terminal node should have at least $d_0 > 0$ unique deaths, where d_0 is also a specified parameter.

3. Calculate a cumulative hazard function (CHF) for each tree. Average to obtain the ensemble CHF.
4. Calculate prediction error for the ensemble CHF using OOB data.

Using the non-parametric random survival forests method above, we can obtain the estimated CHF $h(t \mid X, P)$ for each record (X, P) and thus the survival function $H(t \mid X, P)$ in (10) via $H(t \mid X, P) = e^{-h(t \mid X, P)}$.