E-Companion to *Model-Free Assortment Pricing with Transaction Data*

Ningyuan Chen, Andre A. Cire, Ming Hu, Saman Lagzi

EC.1. Additional Tables and Figures

Table EC.1 shows the percentage performance ratio of the conservative, LP relaxation, and cut-off pricing as explained in Section 7.1. All the instances were solved to optimality using Gurobi 9.0.0 in Python with a desktop computer (Intel Core i7-8700, 3.2 GHz). Similarly, Table EC.2 portrays the average performance guarantee of conservative and cut-off pricing approximation algorithms.

Table EC.1 The relative performance of the approximation strategies. Standard errors are reported in parentheses.

	Performance relative to the optimal solution		
(m,n)	Conservative	LP Relaxation	Cut-off
$(50,10)$	12.5% (0.6%)	79.8% (0.6%)	97.6% (0.1%)
$(50,15)$	10.8% (0.4%)	73.3% (0.6%)	97.0% (0.1%)
$(50,20)$	9.1% (0.4%)	70.4% (0.5%)	96.5% (0.1%)
$(50,25)$	9.1% (0.5%)	72.1% (0.5%)	96.0% (0.1%)
$(100,10)$	7.4% (0.3%)	82.4% (0.4%)	99.0% (0.1%)
$(150,10)$	5.3% (0.2%)	85.0% (0.3%)	99.3% (0.1%)
$(200,10)$	4.4% (0.2%)	85.4% (0.3%)	99.6% (0.1%)

Table EC.2 The performance guarantee of approximation strategies. Standard errors are reported in parentheses.

	Theoretical performance guarantee	
(m,n)	Conservative	Cut-off
$(50,10)$	2.2% (0.2%)	50.1% (3.8%)
$(50,15)$	1.9% (0.1%)	49.5% (3.8%)
$(50,20)$	1.7% (0.1%)	49.7% (4.0%)
$(50,25)$	1.9% (0.2%)	50.4% (3.7%)
$(100,10)$	1.0% (0.1%)	49.7% (2.9%)
$(150,10)$	0.6% (0.0%)	50.1% (2.5%)
$(200,10)$	0.5% (0.0%)	49.9% (2.1%)

Table EC.3 contains the calculated performance bound of cut-off pricing using the 31 product categories in the IRI academic data set. We consider the ratios of median purchase price to mean purchase price and $\underline{P}/\bar{P}$ after proper data cleaning. We note that $\underline{P}/\bar{P}$ could be very small in many categories, but this is primarily due to data entry errors. For example, some products are

Figure EC.1 The sensitivity of the performance of the data-driven optimal pricing (OP-MIP) and cut-off pricing relative to the optimal MNL prices estimated from the data with 100 historical customers. The difference in the revenues is converted to percentage by dividing it by the optimal revenue of the model and then averaged.

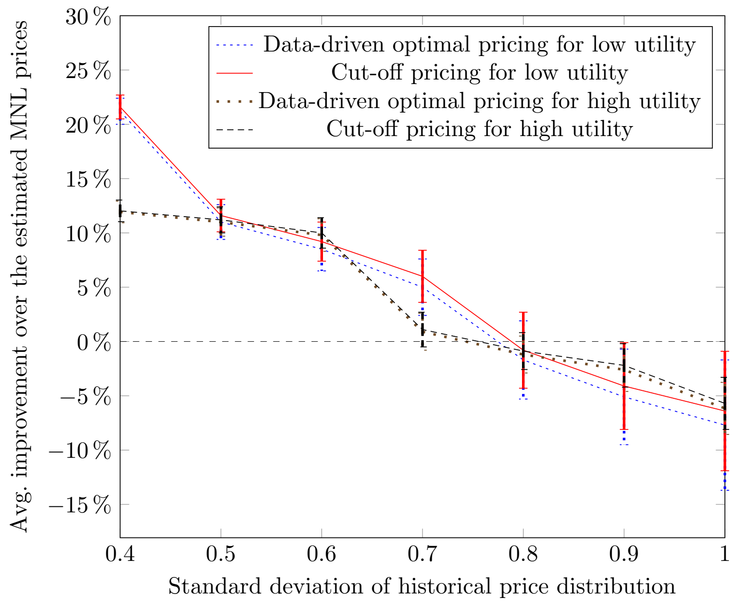

Figure EC.2 The sensitivity of the performance of the data-driven optimal pricing (OP-MIP) and cut-off pricing relative to the optimal MNL prices estimated from the data with 50 historical customers. The data is generated from a mixed logit model.

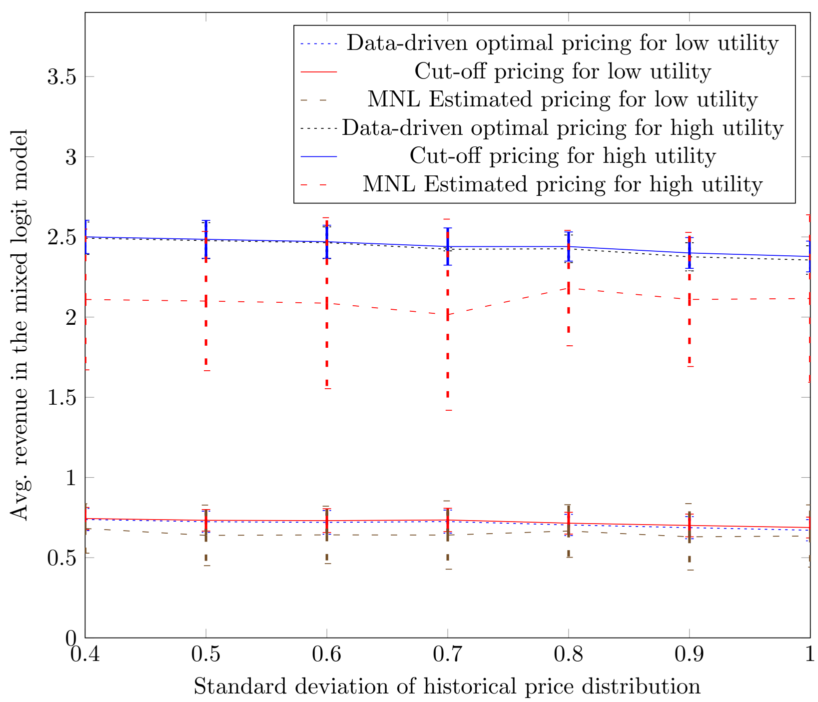

Table EC.3 The implied performance guarantee for cut-off pricing using the IRI data.

	Ratios		
Category	Median/Mean	$\underline{P}/\bar{P}$	Implied performance guarantee
Beer	0.91	0.025	45.7%
Blades	0.772	0.02	38.6%
Carbonated beverages	0.811	0.026	40.6%
Cigarettes	0.466	0.028	23.3%
Coffee	0.98	0.033	49.0%
Cold cereals	0.986	0.069	49.3%
Deodorants	0.97	0.072	48.5%
Dippers	0.769	0.07	38.5%
Facial tissue	0.812	0.04	40.6%
Frozen dinners/entrees	0.842	0.023	42.1%
Frozen pizzas	0.919	0.053	46.0%
Household cleaner	0.964	0.033	48.2%
Hotdogs	0.989	0.037	49.5%
Laundry detergent	0.914	0.031	45.7%
Margarine/spreads/butter blends	0.924	0.07	46.2%
Mayonnaise	0.951	0.09	47.5%
Milk	1.01	0.05	50.5%
Mustard & ketchup	0.953	0.063	47.7%
Paper towels	0.831	0.03	41.5%
Peanut butter	0.861	0.046	43.1%
Razors	1.02	0.145	51.1%
Photography supplies	0.92	0.057	46.0%
Salty snacks	0.982	0.023	49.0%
Shampoo	0.891	0.023	44.5%
Soup	0.942	0.021	47.1%
Spaghetti/Italian sauce	0.887	0.072	44.4%
Sugar substitutes	0.91	0.058	45.5%
Toilet tissue	0.935	0.032	46.8%
Toothbrush	0.733	0.01	36.7%
Toothpaste	0.831	0.013	41.6%
Yogurt	0.711	0.037	35.5%

purchased at \$0.01. We remove the top and bottom 0.001 price quantiles and present the ratio in Table EC.3. The implied performance guarantee column in Table EC.3 contains the implied theoretical performance guarantee by these ratios for each category, obtained through Proposition 10.

We also study the sensitivity of the performance of our data-driven model-free optimal and cut-off pricing algorithms with respect to the dispersion of prices in the data. To that avail, we first revisit the low-utility and high-utility experiments with a limited number of customers, explored in Sections 7.2. For the high-utility experiments, in each of the 200 instances, $\{\alpha_j\}_{j=1}^{10}$ are independent and drawn uniformly at random from the interval $[1,3]$, while historical prices P_{ij} are drawn uniformly at random from the interval $7 \pm \sqrt{3}\sigma$, where σ is the standard deviation of historical prices. In the low-utility experiment, in each of the 200 instances, $\{\alpha_j\}_{j=1}^{10}$ are independent and drawn uniformly at random from the interval $[-2,0]$ and historical prices are drawn uniformly at random from the interval $3.5 \pm \sqrt{3}\sigma$.

Figure EC.3 The sensitivity of the performance of the data-driven cut-off pricing relative to the optimal MNL prices estimated from censored data with 360 historical customers. The difference in the revenues is converted to percentage by dividing it by the optimal revenue of the model and then averaged.

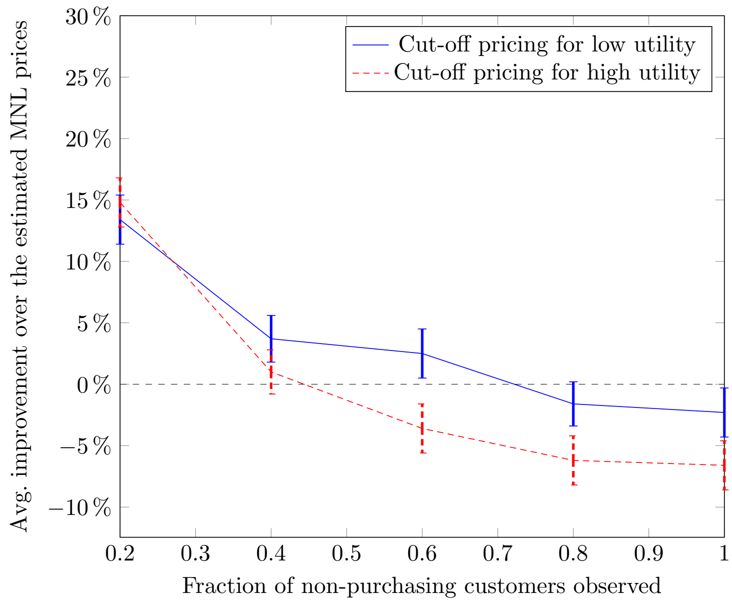

Figure EC.1 portrays our results for both experiments when the number of historical customers is fixed at 100. We observe that in both experiments, the performance of the model-free methodologies compared to the estimated MNL optimal prices deteriorate with the increase of σ. This is intuitive because given that the average historical prices are kept fixed, with the increase of σ (while customers make decisions based on the MNL model), more purchases happen at lower prices, making the model-free approaches more conservative. In the meantime, due to a higher price dispersion in the historical data, the MNL estimation which is correctly specified improves, leading to better performing prices.

We further study the situation where the MNL model is misspecified and revisit the low-utility and high-utility experiments in Section 7.3. Here, like before, we set $\beta_1 = 0.5$ and $\beta_2 = 2$, while in each of the 200 instances, we randomly draw $(\alpha_{1j}, \alpha_{2j})$ independently from $[1,3]$ (for high-utility) and $[-2,0]$ (for low-utility). In the high-utility experiments, historical prices P_{ij} are drawn uniformly at random from the interval $7 \pm \sqrt{3}\sigma$, while in the low-utility experiment, they are drawn uniformly at random from the interval $3.5 \pm \sqrt{3}\sigma$.

Figure EC.2 displays our results for both experiments when the number of historical customers is fixed at 50 and the customers in the historical data made decisions based on the mixed logit model described in Section 7.3. We observe that like before, with the increase of σ the performances of the model-free approach slightly deteriorate, as there will be a larger portion of low price purchases in the historical data which tend to be reflected in the model-free pricing. However, what is interesting is that the performance of the estimated MNL optimal prices does not improve with the increase of σ, which can be attributed to the fact that the MNL model is misspecified here.

Finally, Figure EC.3 shows our results for the case when the data does not record all the non-purchasing customers. In other words, the data is censored and we revisit the low-utility and high-utility experiments with a limited number of customers, explored in Sections 7.2. As it was mentioned in Remark 1, censoring does not affect the prices prescribed by our data-driven methodologies, however, it can significantly affect the estimated MNL optimal prices, even when the the MNL choice model is correctly specified. The result in Figure EC.3 suggests that for a relatively large number of historical customers, 360, in the high-utility experiment from Section 7.3, as long as only 40% or less of the non-purchasing customers are recorded in the data, the data-driven prices prescribed by cut-off pricing outperform estimated MNL optimal prices in terms of the expected revenue. We note that in the low-utility experiment from Section 7.3, this threshold on the fraction of recorded non-purchase customers becomes 60%.

EC.2. Proofs

Proof of Proposition 1. By Proposition 7, the optimal price output by our framework (OP-MIP) satisfies $p^* \in \arg\max_{p\geq 0} p\sum_{i=1}^m \mathbb{I}(P_{i1}\geq p)$. Equivalently, we may write $p^* \in \arg\max_{p\geq 0} p\frac{\sum_{i=1}^m \mathbb{I}(P_{i1}\geq p)}{m}$. By the Law of Large Numbers, we have $\lim_{m\to\infty}\frac{\sum_{i=1}^m \mathbb{I}(P_{i1}\geq p)}{m} = \mathsf{P}(P_{i1}\geq p) = \int_p^{+\infty}(1-F(x))\mathrm{d}G(x)$. Therefore,

$$p^* \in \arg\max_{p\geq 0}\lim_{m\to\infty} p\frac{\sum_{i=1}^m \mathbb{I}(P_{i1}\geq p)}{m} = \arg\max_{p\geq 0} p\int_p^{+\infty} g(x)(1-F(x))\mathrm{d}x.$$

To prove the last part of Proposition 1, we replace $g(x)$ in (3) with an appropriate scaling of $f(x)/(1-F(x))$ using $\Lambda>0$. Hence, when $g(p)\propto f(p)/(1-F(p))$,

$$\arg\max_{p\geq 0} p\int_p^{+\infty} g(x)(1-F(x))\mathrm{d}x = \arg\max_{p\geq 0}\Lambda p\int_p^{+\infty} f(p)\mathrm{d}x = \arg\max_{p\geq 0} R(p).$$

□

Proof of Theorem 1. The proof will proceed in several steps.

1. We will show that the solution to (4) on $[0,a]$ with $p\int_p^a(1-F(x))\mathrm{d}x>0$ is the unique maximizer of (3), p^*.
2. We will show that we can assume without any loss of generality that $a=1$.
3. We will show that if $p^*\leq\hat{p}$ then $\frac{R(p^*)}{R(\hat{p})}\geq\frac{1}{2}$ and for any $\epsilon>0$ construct an example such that $\frac{R(p^*)}{R(\hat{p})}\leq\frac{1}{2}+\epsilon$.
4. We will show that if $p^*>\hat{p}$ then $\frac{R(p^*)}{R(\hat{p})}\geq\frac{\hat{p}}{p^*}$ and for any $\epsilon>0$ construct an example such that $\frac{R(p^*)}{R(\hat{p})}\leq\frac{\hat{p}^{1-\epsilon}}{p^{*1-\epsilon}}$.

Step 1. When $g(p)=1/b$, $p^* = \arg\max_{p\geq 0} p\int_p^b g(x)(1-F(x))\mathrm{d}x = \arg\max_{p\geq 0} p\int_p^a \frac{1}{b}(1-F(x))\mathrm{d}x$. Thus

$$p^* = \arg\max_{p\geq 0} p\int_p^a (1-F(x))\mathrm{d}x = \arg\max_{p\geq 0} pS(p). \tag{EC.1}$$

Note that the function $H(p)\triangleq pS(p)$ on the right hand side of (EC.1) is not zero everywhere and is a continuous, bounded and differentiable function, defined over the compact region of $[0,a]$. On the two ends, it is easy to see $0S(0)=aS(a)=0$. Therefore, the maximizer of $H(p)$ must be in the interior of $[0,a]$ and satisfy the first-order condition for (EC.1):

$$\int_{p^*}^a (1-F(x))\mathrm{d}x - p^*(1-F(p^*)) = S(p^*)-R(p^*)=0.$$

This proves that p^* must satisfy (4).

Step 2. Assume $a\neq 1$. Then, we could define $y=\frac{p}{a}$, $y\in[0,1]$ and define a new distribution $\bar{F}(y)=F(ay)$. We have

$$\bar{R}(y)=y(1-\bar{F}(y))=\frac{a}{a}y(1-F(ay))=\frac{1}{a}p(1-F(p))=\frac{1}{a}R(p).$$

Moreover,

$$\bar{S}(y)=\int_y^1(1-\bar{F}(x))\mathrm{d}x=\frac{a}{a}\int_y^1(1-F(ax))\mathrm{d}x=\frac{a}{a}\int_{ay}^a\frac{1}{a}(1-F(z))\mathrm{d}z=\frac{1}{a}\int_p^a(1-F(z))\mathrm{d}z=\frac{1}{a}S(p).$$

Where the third equation comes from a change of variable in the integral as $z=ax$. Hence, if $a\neq 1$, we can simply rescale the domain of customer valuations by a and define a new distribution over the new domain. Thus, we can assume $a=1$.

Step 3. Assume $p^*\leq\hat{p}$. Notice that

$$p(1-F(p))+\int_p^1(1-F(x))\mathrm{d}x \tag{EC.2}$$

is decreasing in p since

$$\left(p(1-F(p))+\int_p^1(1-F(x))\mathrm{d}x\right)' = 1-F(p)-pf(p)-1+F(p) = -pf(p)\leq 0.$$

Thus, by assumption of $p^*\leq\hat{p}$, we have

$$R(\hat{p})\leq \hat{p}(1-F(\hat{p}))+\int_{\hat{p}}^1(1-F(x))\mathrm{d}x \leq p^*(1-F(p^*))+\int_{p^*}^1(1-F(x))\mathrm{d}x = 2p^*(1-F(p^*))=2R(p^*).$$

The first inequality comes from the fact that $\int_{\hat{p}}^1(1-F(x))\mathrm{d}x\geq 0$. The second inequality follows since (EC.2) is decreasing in p and $p^*\leq\hat{p}$. The prior to the last equality comes from the fact that p^* must satisfy (4).

To prove that this bound is tight, assume $0 < \epsilon_1 < \epsilon$ that is sufficiently small and consider the following distribution F such that:

$$F(p) = \begin{cases} 0, & \text{if } 0 \le p \le 1-\epsilon_1, \\ \frac{1}{\epsilon_1}(p-(1-\epsilon_1)), & \text{if } 1-\epsilon_1 < p \le 1. \end{cases}$$

This leads to:

$$R(p) = \begin{cases} p, & \text{if } 0 \le p \le 1-\epsilon_1, \\ \frac{p(1-p)}{\epsilon_1}, & \text{if } 1-\epsilon_1 < p \le 1. \end{cases} \qquad S(p) = \begin{cases} 1-p-\frac{\epsilon_1}{2}, & \text{if } 0 \le p \le 1-\epsilon_1, \\ \frac{(1-p)^2}{2\epsilon_1}, & \text{if } 1-\epsilon_1 < p \le 1. \end{cases}$$

To prove tightness, we need to show that F is a proper CDF that satisfies Assumptions 1, 2 and 3. It is easy to check that $F(p)$ is a proper CDF as it is nonnegative and increasing in p while $F(0)=0$ and $F(1)=1$, and moreover, for any $b \ge 1$, Assumptions 1 and 2 will be satisfied.

Hence, it remains to show that the bound is tight for F and that F satisfies Assumption 3. Notice that for all $0 \le p \le 1-\epsilon_1$, $R'(p) = (p)' = 1 > 0$ and $S'(p) = -1 < 0$. Meanwhile, $S(0) > R(0) = 0$ and

$$\frac{\epsilon_1}{2} = S(1-\epsilon_1) < R(1-\epsilon_1) = 1-\epsilon_1$$

when $\epsilon_1 < 2/3$. Hence, there is a single point on $[0, 1-\epsilon_1]$ that can satisfy (4). In fact, this point can be calculated as $S(\frac{1}{2}-\frac{\epsilon_1}{4}) = R(\frac{1}{2}-\frac{\epsilon_1}{4})$. Furthermore, $S(1) = R(1) = 0$ and for all $1-\epsilon_1 \le p \le 1$, if $\epsilon_1 < 1/3$,

$$R'(p) = (p - \frac{1}{\epsilon_1}(p^2 - (p - p\epsilon_1)))' = 1 - \frac{2p-1+\epsilon_1}{\epsilon_1} < S'(p) = \frac{p-1+\epsilon_1}{\epsilon_1} - 1 \le 0.$$

Hence $S(p)$ and $R(p)$ cannot cross at any interior point on $[1-\epsilon_1, 1]$ when $\epsilon_1 < 1/3$. Thus the only candidate for p^* if ϵ_1 is sufficiently small is $\frac{1}{2}-\frac{\epsilon_1}{4}$ (and clearly $(\frac{1}{2}-\frac{\epsilon_1}{4})\int_{(\frac{1}{2}-\frac{\epsilon_1}{4})}^1 \big(1-F(z)\big)\mathrm{d}z > 0$) and it is the unique maximizer of $p\int_p^1(1-F(z))\mathrm{d}z$ (we showed this in Step 1.).

Moreover, $R(\frac{1}{2}-\frac{\epsilon_1}{4}) = \frac{1}{2}-\frac{\epsilon_1}{4}$ while if $\epsilon_1 \le \frac{1}{3}$, $\hat{p} = 1-\epsilon_1$ and $R(\hat{p}) = 1-\epsilon_1$, resulting in $R(p^*)/R(\hat{p}) = \frac{\frac{1}{2}-\frac{\epsilon_1}{4}}{1-\epsilon_1} \le \frac{1}{2}+\epsilon$ for ϵ_1 sufficiently small, thus proving the bound is tight (while $R(p)$ is unimodal and has a unique maximizer $\hat{p}$). Hence, we have shown the bound is tight for F and that F satisfies Assumption 3.

Step 4. Assume $p^* > \hat{p}$. Notice that by assumption, p^* is the first and only point on $[0,1]$ that satisfies (4) with $p^*\int_{p^*}^1(1-F(x))\mathrm{d}x > 0$. Moreover since $p^* > \hat{p}$,

$$\int_{\hat{p}}^1 (1-F(z))\mathrm{d}z \ge \int_{p^*}^1 (1-F(z))\mathrm{d}z,$$

while $\hat{p} > 0$. Hence $\hat{p}\int_{\hat{p}}^1(1-F(x))\mathrm{d}x > 0$ and by assumption we have $R(\hat{p}) \ne S(\hat{p})$. Therefore, since

$$\int_0^1 (1-F(z))\mathrm{d}z > 0(1-F(0)),$$

then we have that

$$R(\hat{p}) < \int_{\hat{p}}^{1}(1-F(z))\mathrm{d}z = \frac{\hat{p}}{\hat{p}}\int_{\hat{p}}^{1}(1-F(z))\mathrm{d}z < \frac{p^*}{\hat{p}}\int_{p^*}^{1}(1-F(z))\mathrm{d}z = \frac{p^*}{\hat{p}}R(p^*),$$

where the last inequality follows from Assumption 3 and that p^* is the unique maximizer of (3) and the last equality comes from the fact that p^* must satisfy (4). Thus,

$$\frac{R(p^*)}{R(\hat{p})} \geq \frac{\hat{p}}{p^*}.$$

It remains to be shown that the bound is tight when $p^* > \hat{p}$. Consider any $0 < x_1 < x_2 < 1/e$. Notice that by Assumption 3, $R(p)$ is unimodal on $[0,1]$. Moreover, by assumption we have $p^* > \hat{p}$, hence, for all $x > p^*$, $x(1-F(x)) \leq p^*(1-F(p^*))$. Therefore, for all $x > p^*$, $1-F(x) \leq \frac{R(p^*)}{x}$. Thus, we claim $p^* < \frac{1}{e}$. This is because,

$$R(p^*) = \int_{p^*}^{1}(1-F(z))\mathrm{d}z < \int_{p^*}^{1}\frac{R(p^*)}{z}\mathrm{d}z = R(p^*)\log(\frac{1}{p^*}),$$

where the first equality comes from the fact that p^* must satisfy (4) while the strict inequality follows form the fact that F is assumed to be continuous and $F(1)=0$, proving that $p^* < \frac{1}{e}$. Let $\bar{x} = ex_2$, we know $\bar{x} < 1$. Assume $0 < \epsilon_2 < \epsilon$ is sufficiently small and consider the following distribution F such that:

$$1-F(p) = \begin{cases} 1 - \frac{1-\frac{\delta}{x_1^{2-\epsilon_2}}}{\epsilon_2}p, & \text{if } 0 \leq p \leq \epsilon_2, \\ \frac{\delta}{x_1^{2-\epsilon_2}}, & \text{if } \epsilon_2 \leq p \leq x_1, \\ \frac{\delta}{p^{2-\epsilon_2}}, & \text{if } x_1 \leq p \leq x_2, \\ \frac{\delta}{p x_2^{1-\epsilon_2}}, & \text{if } x_2 \leq p \leq \bar{x}, \\ \frac{\delta}{\bar{x}x_2^{1-\epsilon_2}} - \frac{\frac{\delta}{\bar{x}x_2^{1-\epsilon_2}}}{\epsilon_2}(p-\bar{x}), & \text{if } \bar{x} \leq p \leq \bar{x}+\epsilon_2, \\ 0, & \text{if } \bar{x}+\epsilon_2 \leq p \leq 1. \end{cases}$$

The intuition behind the construction of F is as follows. We want to ensure that $\hat{p} = x_1$ while p^* can be arbitrarily close to x_2. If F is such that at these points $R(\cdot)$ is proportional to p, we would be very close to our goal. It would remain to ensure p^* happens at the unique point at which $R(\cdot)$ and $S(\cdot)$ cross each other. Thus, F must be such that from $\hat{p}$ to x_2, the difference between $R(\cdot)$ and $S(\cdot)$ shrinks and reaches zero at a point slightly larger than x_2.

To formally prove tightness, we need to show that F is a proper CDF that satisfies Assumptions 1, 2 and 3. It is easy to check that $F(p)$ is nonnegative and increasing in p while $F(0)=0$ and $F(1)=1$. Moreover, for any $b \geq 1$, Assumptions 1 and 2 are satisfied.

Hence, it remains to show that the bound is tight for F and it satisfies Assumption 3. Notice that for all $0 \leq p \leq \epsilon_2$, $R(p) = p - \frac{1-\frac{\delta}{x_1^{2-\epsilon_2}}}{\epsilon_2}p^2$ and $R'(p) = 1 - 2\frac{1-\frac{\delta}{x_1^{2-\epsilon_2}}}{\epsilon_2}p > 0$ if $\delta > \frac{1}{2}x_1^{2-\epsilon_2}$. Moreover, for all $\epsilon_2 \le p \le x_1$ we have $R'(p) = \frac{\delta}{x_1^{2-\epsilon_2}} > 0$ while $S(p)$ is decreasing for all $0 \le p \le x_1$. Furthermore, notice that

$$S(x_2) > \int_{x_2}^{\bar{x}} \frac{\delta}{p x_2^{1-\epsilon_2}} \mathrm{d}p = \frac{\delta}{x_2^{1-\epsilon_2}} \log\left(\frac{\bar{x}}{x_2}\right) = \frac{\delta}{x_2^{1-\epsilon_2}} = R(x_2),$$

while

$$\begin{aligned} S(x_1) &= \int_{x_1}^{x_2} \frac{\delta}{p^{2-\epsilon_2}} \mathrm{d}p + S(x_2) > \frac{\delta}{(1-\epsilon_2)} \left(\frac{1}{x_1^{1-\epsilon_2}} - \frac{1}{x_2^{1-\epsilon_2}}\right) + \frac{\delta}{x_2^{1-\epsilon_2}} \\ &= \frac{\delta}{(1-\epsilon_2)} \frac{1}{x_1^{1-\epsilon_2}} - \frac{\epsilon_2 \delta}{(1-\epsilon_2)} \frac{1}{x_2^{1-\epsilon_2}} > \frac{\delta}{x_1^{1-\epsilon_2}} = R(x_1), \end{aligned}$$

thus $S(p)$ and $R(p)$ do not cross on $0 \le p \le x_1$. This is because $R(p)$ is strictly increasing and $S(p)$ is decreasing in this region while $S(0) > R(0)$ and $S(x_1) > R(x_1)$.

Moreover, if $x_1 \le p \le x_2$, $R(p) = \frac{\delta}{p^{1-\epsilon_2}}$ and

$$R'(p) = -(1-\epsilon_2) \frac{\delta}{p^{2-\epsilon_2}} > -\frac{\delta}{p^{2-\epsilon_2}} = S'(p),$$

while $S(x_2) > R(x_2) = \frac{\delta}{x_2^{1-\epsilon_2}}$ thus $S(p)$ and $R(p)$ do not cross on $x_1 \le p \le x_2$ either. This is because if they cross at any point in this regions, since $R'(p) > S'(p)$ in this region, it must be that $S(x_2) \le R(x_2)$ which is a contradiction since we showed earlier that $S(x_2) > R(x_2)$.

When $x_2 \le p \le \bar{x}$, $R'(p) = 0$ while $S'(p) < 0$, thus $S(p)$ and $R(p)$ can cross at most once in this region, as $S(x_2) > \frac{\delta}{x_2^{1-\epsilon_2}} = R(x_2)$ while $S(\bar{x}) < R(\bar{x}) = \frac{\delta}{x_2^{1-\epsilon_2}}$ for ϵ_2 sufficiently small.

Finally, if $\bar{x} \le p \le \bar{x} + \epsilon_2$ and $\epsilon_2 < \frac{\bar{x}}{2}$,

$$R'(p) = \left(\frac{p\delta}{\bar{x} x_2^{1-\epsilon_2}} - \frac{\frac{\delta}{\bar{x} x_2^{1-\epsilon_2}}}{\epsilon_2}(p^2 - p\bar{x})\right)' = \frac{\delta}{\bar{x} x_2^{1-\epsilon_2}} - \frac{\frac{\delta}{\bar{x} x_2^{1-\epsilon_2}}}{\epsilon_2}(2p - \bar{x}) < S'(p) = -\frac{\delta}{\bar{x} x_2^{1-\epsilon_2}} + \frac{\frac{\delta}{\bar{x} x_2^{1-\epsilon_2}}}{\epsilon_2}(p - \bar{x}),$$

while $S(\bar{x} + \epsilon_2) = R(\bar{x} + \epsilon_2) = 0$. Therefore, $S(p)$ and $R(p)$ do not cross on $[\bar{x}, \bar{x} + \epsilon_2)$ since otherwise, because $S'(p) > R'(p)$ in this region when $\epsilon_2 < \frac{\bar{x}}{2}$, it would imply $0 = S(\bar{x} + \epsilon_2) > R(\bar{x} + \epsilon_2)$ which is a contradiction.

Thus, $S(p)$ and $R(p)$ will cross once and that will happen on $[x_2, \bar{x}]$. This crossing point p^* will satisfy $0 < p^* - x_2 < \epsilon$ for some ϵ_2 sufficiently small. This is because $R(x_2) = \frac{\delta}{x_2^{1-\epsilon_2}}$ while $S(x_2)$ will get arbitrarily close to $\frac{\delta}{x_2^{1-\epsilon_2}}$ for a sufficiently small ϵ_2, and $S'(p) \le -\frac{\delta}{\bar{x} x_2^{1-\epsilon_2}}$ in this region while $R(p)$ is constant (while clearly we have $p^* \int_{p^*}^1 (1 - F(x)) \mathrm{d}x > 0$).

Therefore we can conclude that the unique maximizer of $\max_{p \ge 0} p \int_p^1 (1 - F(z)) \mathrm{d}z$, p^* satisfies $0 < p^* - x_2 < \epsilon$ for some ϵ_2 sufficiently small. Moreover, it is easy to see that $R(p)$ is unimodal and has a unique maximizer at $\hat{p} = x_1$. We have

$$\frac{R(p^*)}{R(\hat{p})} = \frac{\frac{\delta}{p^{*1-\epsilon}}}{\frac{\delta}{\hat{p}^{1-\epsilon}}} = \frac{\hat{p}^{1-\epsilon_2}}{p^{*1-\epsilon_2}} \le \frac{\hat{p}^{1-\epsilon}}{p^{*1-\epsilon}}.$$

Hence, we have shown the bound is tight for F and it satisfies Assumption 3. Finally, since we already showed $p^* < \frac{1}{e}$, we can also see that $\frac{R(p^*)}{R(\hat{p})} > e\hat{p}$.

□

Proof of Lemma 1. First we will prove that (3) has a maximizer that is the unique solution to (4) on $[0,a)$. Since $p\int_p^a(1-F(x))\mathrm{d}x$ is a continuous, bounded and differentiable function, not zero everywhere and defined on a compact set, while

$$0\int_0^a(1-F(x))\mathrm{d}x = a\int_a^a(1-F(x))\mathrm{d}x = 0,$$

it must have a maximizer in the interior of $[0,a]$ and this maximizer must satisfy the first-order condition for (3). Hence, we must have

$$\left(p\int_p^a(1-F(x))\mathrm{d}x\right)' = \int_p^a(1-F(x))\mathrm{d}x - p(1-F(p)) = 0. \tag{EC.3}$$

Notice that

$$\int_0^a(1-F(x))\mathrm{d}x - 0(1-F(0)) > 0,$$

while if $p\geq\frac{a}{2}$

$$\int_p^a(1-F(x))\mathrm{d}x - p(1-F(p)) \leq (1-F(p))(a-p) - p(1-F(p)) \leq (1-F(p))(a-2p) \leq 0.$$

Thus (EC.3) must have at least one root in $(0,\frac{a}{2}]$. We will show that it cannot have more than one root in $[0,a)$. Notice that

$$\int_a^a(1-F(x))\mathrm{d}x - a(1-F(a)) = 0$$

while

$$\left(\int_p^a(1-F(x))\mathrm{d}x - p(1-F(p))\right)' = 2F(p) - 2 + pf(p).$$

By assumption, $\frac{pf(p)}{1-F(p)}$ is strictly increasing, thus, $\frac{pf(p)}{1-F(p)}=2$ can only have at most one root. Let $\hat{x}$ be the single root of $\frac{pf(p)}{1-F(p)}=2$. Then, (EC.3) is strictly decreasing for all $p<\hat{x}$ and strictly increasing for all $p>\hat{x}$ and $0\leq\hat{x}<a$ (otherwise we will have $\int_a^a(1-F(x))\mathrm{d}x - a(1-F(a))<0$ which is a contradiction). Then by noticing that $\int_0^a(1-F(x))\mathrm{d}x - 0(1-F(0))>0$ while $\int_a^a(1-F(x))\mathrm{d}x - a(1-F(a))=0$, it turns out that (EC.3) must have exactly one root in $[0,a)$. Hence, $p\int_p^a(1-F(x))\mathrm{d}x$ has a unique point on $[0,a)$ that satisfies (4) and is its maximizer. If we denote this point with p^*, it is clear that $p^*\int_{p^*}^a(1-F(x))\mathrm{d}x>0$.

Now we prove $R(p)$ is unimodal in $[0,a]$ and has a unique maximizer. $R(p)$ is a continuous, bounded and differentiable function defined on a compact set and not zero everywhere while $R(0)=R(a)=0$. Thus, $R(p)$ must have a maximizer in the interior of $[0,a]$ and it must satisfy the first-order condition for $R(p)$. Thus at optimality we must have

$$\Big(p(1-F(p))\Big)' = 1-F(p)-pf(p) = 0,$$

and since $\frac{pf(p)}{1-F(p)}$ is assumed to be strictly increasing, $\frac{pf(p)}{1-F(p)}=1$ (equivalent to the first-order condition) can have at most one root. Thus there is a unique point on $[0,a]$ that satisfies the first-order condition for $R(p)$ and is its unique maximizer, proving that $R(p)$ is unimodal as well.

$\square$

Proof of Lemma 2. First, notice that $R(p)$ is a continuous, bounded and differentiable function defined on a compact set and not zero everywhere while $R(0)=R(a)=0$. Thus, $R(p)$ must have a maximizer in the interior of $[0,a]$ and it must satisfy the first-order condition for $R(p)$. Thus by the definition of $\hat{p}$, it must be that $R'(\hat{p})=1-F(\hat{p})-\hat{p}f(\hat{p})=0$, suggesting that $\frac{1}{\hat{p}}=\frac{f(\hat{p})}{1-F(\hat{p})}$.

Now, for the sake of contradiction, let us assume that $\hat{p}<p^*$. Notice that by assumption, p^* is the first and only point on $[0,a]$ that satisfies (4) with $p^*\int_{p^*}^a(1-F(x))\mathrm{d}x>0$. Moreover since $p^*>\hat{p}$,

$$\int_{\hat{p}}^a(1-F(z))\mathrm{d}z\geq\int_{p^*}^a(1-F(z))\mathrm{d}z,$$

while $\hat{p}>0$. Hence $\hat{p}\int_{\hat{p}}^a(1-F(x))\mathrm{d}x>0$ and by assumption we have $R(\hat{p})\neq S(\hat{p})$, and since $\int_0^a(1-F(z))\mathrm{d}z>0(1-F(0))$, then we have that

$$R(\hat{p})=\hat{p}(1-F(\hat{p}))<\int_{\hat{p}}^a(1-F(z))\mathrm{d}z,$$

suggesting that

$$\frac{1-F(\hat{p})}{\int_{\hat{p}}^a(1-F(z))\mathrm{d}z}<\frac{1}{\hat{p}}=\frac{f(\hat{p})}{1-F(\hat{p})}.$$

However, this is a contradiction since by assumption we have $\frac{h(p)}{1-H(p)}\geq\frac{f(p)}{1-F(p)}$ for all $p\in[0,a]$, which suggests:

$$\frac{1-F(\hat{p})}{\int_{\hat{p}}^a(1-F(z))\mathrm{d}z}=\frac{\frac{1-F(\hat{p})}{\int_0^a(1-F(z))\mathrm{d}z}}{\frac{\int_{\hat{p}}^a(1-F(z))\mathrm{d}z}{\int_0^a(1-F(z))\mathrm{d}z}}=\frac{h(\hat{p})}{1-H(\hat{p})}\geq\frac{f(\hat{p})}{1-F(\hat{p})}.$$

Finally, given that we have $\hat{p}\geq p^*$, it follows directly from Theorem 1 that $\frac{R(p^*)}{R(\hat{p})}\geq\frac{1}{2}$. $\square$

Proof of Proposition 2. Let $\hat{R}(p)\equiv p\int_p^b g(x)(1-F(x))\mathrm{d}x$ and let $\hat{R}_m(p)=p\frac{\sum_{i=1}^m\mathbb{I}(P_{i1}\geq p)}{m}=p(1-\hat{\mathcal{F}}_m(p))$ where $\hat{\mathcal{F}}_m(p)=\frac{\sum_{i=1}^m\mathbb{I}(P_{i1}\geq p)}{m}$ is the empirical CDF corresponding to $\mathcal{F}(p)=\int_0^p g(x)(1-F(x))\mathrm{d}x$. We note that this is rigorous as P_{i1} for all $i\in\mathscr{C}$ is generated independently from the distribution $\mathsf{P}(P_{i1}\geq p)=\int_p^{+\infty}(1-F(x))\mathrm{d}G(x)$.

By applying Hoeffding's inequality and noticing that for all $i\in\mathscr{C}$ and p^* we have $0\leq\frac{p^*\mathbb{I}(P_{i1}\geq p^*)}{m}\leq\frac{b}{m}$, we can write:

$$\mathsf{P}(|\hat{R}_m(p^*)-\hat{R}(p^*)|\geq t)\leq 2e^{\frac{-2t^2m}{b^2}}\ \forall\ t>0. \tag{EC.4}$$

Define the event $E\triangleq\{|\hat{R}_m(p^*)-\hat{R}(p^*)|<t\}$. By assumption, for all $p\in[0,b]$, we have:

$$p\int_p^b g(x)(1-F(x))\mathrm{d}x\leq p^*\int_{p^*}^b g(x)(1-F(x))\mathrm{d}x-\alpha(p-p^*)^2.$$

Therefore, for any p such that $|p-p^*|\geq\epsilon$, we have:

$$\hat{R}(p)\leq\hat{R}(p^*)-\alpha\epsilon^2.$$

Consider any $t\in[0,\alpha\epsilon^2]$. On event E, if $|p_m^*-p^*|\geq\epsilon$, then $\hat{R}(p_m^*)\leq\hat{R}(p^*)-\alpha\epsilon^2$ and we have

$$\hat{R}_m(p_m^*)\geq\hat{R}_m(p^*)\geq\hat{R}(p^*)-t\geq\hat{R}(p_m^*)+\alpha\epsilon^2-t.$$

Where the first inequality follows from the definition of p_m^* while the second inequality happens due to event E. Therefore, $|\hat{R}_m(p_m^*)-\hat{R}(p_m^*)|\geq\hat{R}_m(p_m^*)-\hat{R}(p_m^*)\geq\alpha\epsilon^2-t$.

In other words, we know that $E\cap\{|p_m^*-p^*|\geq\epsilon\}\subseteq E\cap\{|\hat{R}_m(p_m^*)-\hat{R}(p_m^*)|\geq\alpha\epsilon^2-t\}$.

Moreover, since $|\hat{R}_m(p_m^*)-\hat{R}(p_m^*)|\geq\alpha\epsilon^2-t$ implies that $\max_{p\geq0}(|\hat{R}_m(p)-\hat{R}(p)|)\geq\alpha\epsilon^2-t$, we have

$$\mathsf{P}\big(E\cap\{|p_m^*-p^*|\geq\epsilon\}\big)\leq\mathsf{P}\big(E\cap\{\max_{p\geq0}(|\hat{R}_m(p)-\hat{R}(p)|)\geq\alpha\epsilon^2-t\}\big)\leq\mathsf{P}(\max_{p\geq0}(|\hat{R}_m(p)-\hat{R}(p)|)\geq\alpha\epsilon^2-t). \tag{EC.5}$$

Next, we provide an upper bound for $\mathsf{P}(\max_{p\geq0}(|\hat{R}_m(p)-\hat{R}(p)|)\geq\alpha\epsilon^2-t)$. This can be achieved by applying DvoretzkyKieferWolfowitz inequality (Van der Vaart 2000, page 268): since $\hat{F}_m(p)$ is the empirical CDF of $F(p)$ for $p\in[0,b]$, we have:

$$\begin{aligned}\mathsf{P}(\max_{p\geq0}(|\hat{R}_m(p)-\hat{R}(p)|)\geq\alpha\epsilon^2-t)&\leq\mathsf{P}\left(b\left\|\frac{\sum_{i=1}^m\mathbb{I}(P_{i1}\geq p)}{m}-\int_p^b g(x)(1-F(x))\mathrm{d}x\right\|_\infty\geq\alpha\epsilon^2-t\right)\\&\leq 2e^{\frac{-2(\alpha\epsilon^2-t)^2m}{b^2}}.\end{aligned} \tag{EC.6}$$

Thus we have:

$$\begin{aligned}\mathsf{P}(|p_m^*-p^*|\geq\epsilon)&=\mathsf{P}(E\cap\{|p_m^*-p^*|\geq\epsilon\})+\mathsf{P}(E^c\cap\{|p_m^*-p^*|\geq\epsilon\})\\&\leq\mathsf{P}(E\cap\{|p_m^*-p^*|\geq\epsilon\})+\mathsf{P}(E^c)\\&\leq 2e^{\frac{-2(\alpha\epsilon^2-t)^2m}{b^2}}+2e^{\frac{-2t^2m}{b^2}},\end{aligned}$$

where the last inequality follows from (EC.4), (EC.5), and (EC.6).

To finalize the proof, we need to choose $0\leq t\leq\alpha\epsilon^2$ such that the quasi-convex function (when m is large enough)

$$2e^{\frac{-2t^2m}{b^2}}+2e^{\frac{-2(\alpha\epsilon^2-t)^2m}{b^2}}$$

is minimized, which is achieved, when $t=\frac{\alpha\epsilon^2}{2}$. Therefore, we have:

$$\mathsf{P}(|p_m^*-p^*|\geq\epsilon)\leq 4e^{\frac{-\alpha^2\epsilon^4m}{2b^2}}.$$

Hence, if we want $|p_m^*-p^*|\geq\epsilon$ with probability at most δ, we need up to $m=\lceil(\frac{2b^2}{\alpha^2\epsilon^4})\log(\frac{4}{\delta})\rceil$ historical customers, suggesting a sample complexity of $O\big((\frac{b^2}{\alpha^2\epsilon^4})\log(\frac{1}{\delta})\big)$.

$\square$

Proof of Proposition 3. The linear program that results from removing the disjunctive constraint (2) from (DP), i.e.,

$$\min_{r,\mathbf{v}_i}\{r\colon r\geq 0, \mathbf{v}_i\in\mathscr{V}_i\},$$

has the trivial finite optimum $r=0$. Thus, the model (DP-LP) is obtained by applying directly Corollary 2.1.2 by Balas (1998). □

Proof of Lemma 3. We begin with statement (a). If $p_{c_i}\geq P_{ic_i}$, then the valuation $\mathbf{v}_i^{\varnothing}$ defined by $v_{ic_i}^{\varnothing}=P_{ic_i}$ and $v_{ij}^{\varnothing}=0$ for all $j\in\mathscr{P}\setminus\{c_i\}$ belongs to $\mathscr{W}_i^{\varnothing}(\mathbf{p})$. Conversely, if $\mathbf{v}_i^{\varnothing}\in\mathscr{W}_i^{\varnothing}(\mathbf{p})$, the inequalities $v_{ij}^{\varnothing}\leq p_j$ and $v_{ij}^{\varnothing}\geq P_{ic_i}$ from (1) together imply $p_{c_i}\geq P_{ic_i}$.

For statement (b), the valuation $\mathbf{v}_i^{c_i}$ defined by $v_{ic_i}^{c_i}=\max\{P_{ic_i},p_{c_i}\}$ and $v_{ij}^{c_i}=0$ for all $j\in\mathscr{P}\setminus\{c_i\}$ belongs to $\mathscr{W}_i^{c_i}(\mathbf{p})$.

Finally, for statement (c), consider the set of constraints that are satisfied by points in $\mathscr{W}_i^{j}(\mathbf{p})$ after re-arranging the constant terms to the right-hand side of the inequalities:

$$v_{ij}^{j}\geq p_j, \tag{EC.7}$$

$$v_{ij}^{j}-v_{ij'}^{j}\geq p_j-p_{j'}, \qquad \forall j'\in\mathscr{P}, \tag{EC.8}$$

$$v_{ic_i}^{j}\geq P_{ic_i}, \tag{EC.9}$$

$$v_{ic_i}^{j}-v_{ij'}^{j}\geq P_{ic_i}-P_{ij'}, \qquad \forall j'\in\mathscr{P}\setminus\{c_i\}. \tag{EC.10}$$

If $p_j-p_{c_i}\leq P_{ij}-P_{ic_i}$, we construct a valuation $\mathbf{v}_i^{j}\in\mathscr{W}_i^{j}(\mathbf{p})$ where $v_{ij}^{j}=\max\{P_{ij},p_j\}$, $v_{ic_i}^{j}=\max\{P_{ic_i},P_{ic_i}-P_{ij}+p_j\}$, and $v_{ij'}^{j}=0$ for all $j'\in\mathscr{P}\setminus\{c_i,j\}$. In particular, (EC.7) and (EC.9) are satisfied by construction. Assume now $p_j>P_{ij}$. For (EC.8) and (EC.10) with $j'=c_i$, we have

$$v_{ij}^{j}-v_{ic_i}^{j}=p_j-P_{ic_i}+P_{ij}-p_j=P_{ij}-P_{ic_i}\geq p_j-p_{c_i},$$

where the last inequality follows from the statement hypothesis. For $j'\neq c_i$, note that $v_{ij}^{j}-p_j$ is zero while $v_{ij'}^{j}-p_{j'}$ is non-positive in (EC.8), and analogously $v_{ic_i}^{j}-P_{ic_i}$ is positive while $v_{ij'}^{j}-P_{ij'}=-P_{ij'}$ is non-positive. If $p_j\leq P_{ij}$, note that $v_{ij}^{j}=P_{ij}$ and $v_{ic_i}^{j}=P_{ic_i}$, and the same derivations above apply.

Finally, the sufficient conditions of (c) follow directly from (EC.8) and (EC.10) with $j'=c_i$ in (EC.8) and $j'=j$ in (EC.10). □

Proof of Proposition 4. Let $j\in\mathscr{P}$ and denote by G the set of feasible solutions to (DP-LP). The projection of G onto variable x_j is

$$\begin{aligned}\mathrm{Proj}_{x_j}G &= \{x_j\in[0,1]\colon \exists\ ((\mathbf{v}_i^1,\ldots,\mathbf{v}_i^n,\mathbf{v}_i^{\varnothing}),(x_1,\ldots,x_j,\ldots,x_n,x_{\varnothing}))\in G\}\\ &= \{x_j\in[0,1]\colon\ \nexists \mathbf{v}_i^{j}\in\mathscr{W}_i^{j}(\mathbf{p})\Rightarrow x_j=0\}\\ &= \{x_j\in[0,1]\colon (p_j-p_{c_i}>P_{ij}-P_{ic_i})\Rightarrow x_j=0\}\\ &= \{x_j\in[0,1]\colon x_j\leq\mathbb{I}(p_j-p_{c_i}\leq P_{ij}-P_{ic_i})\},\end{aligned}$$

where the second-to-last equality follows from Lemma 3. The same arguments follow for $x_\varnothing$. Since the objective is defined only in terms of $\mathbf{x}$, we can replace the inequalities of G by the projections depicted above, which results in the equivalent formulation

$$f_i(\mathbf{p}) = \min_{\mathbf{x}\geq 0} \quad \sum_{j\in\mathscr{P}} p_j x_j \tag{EC.11}$$

$$\text{s.t.} \quad \sum_{j\in\mathscr{P}} x_j + x_\varnothing = 1, \tag{EC.12}$$

$$x_\varnothing \leq \mathbb{I}(p_{c_i} \geq P_{ic_i}), \tag{EC.13}$$

$$x_j \leq \mathbb{I}(p_j - p_{c_i} \leq P_{ij} - P_{ic_i}), \qquad \forall j \in \mathscr{P}. \tag{EC.14}$$

Finally, note from (EC.12) and (EC.13) that

$$\sum_{j\in\mathscr{P}} x_j = 1 - x_\varnothing \geq 1 - \mathbb{I}(p_{c_i} \geq P_{ic_i}) = \mathbb{I}(p_{c_i} < P_{ic_i})$$

must hold at any feasible solution, and particularly tight at optimality since it is the only constraint that bounds $\mathbf{x}$ from below besides the non-negativity conditions. □

Proof of Proposition 5. Statement (a) follows from the fact that product c_i is always feasible (Lemma 3-(b)) and that the inequality (13) for $j = c_i$ holds with $\mu^*_{ic_i} = 0$ at optimality, while from inequality (11), we have that if $p_{c_i} \geq P_{ic_i}$, $f_i(\mathbf{p}) = 0$. For statement (b), if $f_i(\mathbf{p}) = 0$, the condition trivially holds. Assume $f_i(\mathbf{p}) > 0$. Let $P' = \max_{\{j\in\mathscr{P}: p_j < P^{\max}\}} p_j$. It can be easily shown that P' always exists. Moreover, let $j' \in \mathscr{P}$ be a product such that $p_{j'} \geq P^{\max}$. If $j' = c_i$, then $f_i(\mathbf{p}) = 0$ which is a contradiction. Thus, $p_{c_i} < P^{\max}$. If $j' \neq c_i$, $f_i(\mathbf{p}) < p_{j'}$ because of (a) and $p_{c_i} < p_{j'}$. Reducing the price of j' to P' therefore does not change the optimal value of (DP-C-Dual), and the same argument can be repeated for other products. □

Proof of Proposition 6. It suffices to show that both optimal solution values match when conditioned to a fixed $\mathbf{p} \geq 0$. First, by Proposition 5-(b), we can restrict our analysis to $p_j < P^{\max}$ for all $j \in \mathscr{P}$ without loss of generality.

Consider any customer $i \in \mathscr{C}$. If $p_{c_i} \geq P_{ic_i}$, then we must have $y_{ic_i} = 0$ in (OP-B); otherwise, we can assume $y_{ic_i} = 1$ since that can only be benefitial to the objective. Thus, the objective functions of both models match.

Suppose now that $p_j - p_{c_i} \leq P_{ij} - P_{ic_i}$ for $j \neq c_i$, i.e., product j is feasible to purchase by customer i. We necessarily must have $y_{ij} = 1$ because of (18) and, thus, (15) and (16) match. If otherwise $p_j - p_{c_i} > P_{ij} - P_{ic_i}$, then we may have either $y_{ij} = 0$ or $y_{ij} = 1$. Since the objective of (OP-B) maximizes τ_i, we can assume $y_{ij} = 0$, which can only relax the bound on τ_i in (16). Thus, (15) and (16) also match in this case, i.e., at optimality, the values of τ_i (and hence the optimal values of both models) are the same. □

Proof of Theorem 2. We first show (a). First, (OP-ϵ) with $\epsilon = 0$ is always feasible. That is, for any $\mathbf{p} \geq 0$ and customer $i \in \mathscr{C}$, we set $y_{ic_i} = 1$ if and only if $p_{c_i} \leq P_{ic_i}$, for all $j \in \mathscr{P} \setminus \{c_i\}$, $y_{ij} = 0$ if and only if $p_j - p_{c_i} \geq P_{ij} - P_{ij}$, and τ_i appropriately to satisfy (20). Thus, it remains to show that $g(0)$ is bounded from above. This follows from noting that, for all $i \in \mathscr{C}$, p_{c_i} is bounded by $P_{ic_i} + P^{\max}$ in inequality (21) and that τ_i is bounded by p_{c_i} in inequality (20) with $j = c_i$.

For (b), let $(\mathbf{p}^0, \boldsymbol{\tau}^0, \mathbf{y}^0)$ be an optimal solution tuple and consider any $j' \in \mathscr{P}$ such that $p^0_{j'} = 0$. We show an alternative feasible solution with the same (optimal) value after increasing $p^0_{j'}$ as in the statement. If $y^0_{ic_i} = 0$ for some customer i, then increasing $p^0_{j'}$ does not affect τ_i nor the final solution value, given that y^0_{ij} is adjusted appropriately for $j \neq c_i$ to ensure feasibility. If otherwise $y^0_{ic_i} = 1$ for some customer i, we have two cases:

1. Case 1, $\tau^0_i = 0$. In such a scenario, we can equivalently set $y^0_{ic_i} = 0$ and apply the same adjustments to y^0_{ij} for all $j \neq c_i$ as above, preserving the solution value.
2. Case 2, $\tau^0_i > 0$. Due to inequality (20), we must have $j' \neq c_i$ and $y^0_{ij'} = 0$. Thus, $p^0_{j'} - p^0_{c_i} \geq P_{ij'} - P_{ic_i}$ from inequality (22). Increasing $p^0_{j'}$ thefore just increases the left-hand side of such inequality, and hence does not impact feasibility nor the solution value.

We now show (c). Since the feasible set of (OP-ϵ) relaxes and restricts that of (OP-B) for $\epsilon = 0$ and $\epsilon > 0$, respectively, the inequality

$$0 \leq g(0) - m\tau^* \leq g(0) - \sum_{i \in \mathscr{C}} f_i(\mathbf{p}')$$

follows directly. We will now show that

$$g(0) - \sum_{i \in \mathscr{C}} f_i(\mathbf{p}') \leq \delta' \leq \delta.$$

Consider the ordered vector of prices $\mathbf{p}^0 > 0$ in the statement and the associated tuple $(\mathbf{p}^0, \boldsymbol{\tau}^0, \mathbf{y}^0)$. Let $i \in \mathscr{C}$ be a customer such that $y^0_{ic_i}\tau^0_i > 0$. The inequality (20) is tight for some $j' \leq c_i$, i.e., we can have $y^0_{ij} = 0$ for all $j < j'$. This implies that $p^0_j - p^0_{c_i} \geq P_{ij} - P_{ic_i}$ for those indices due to (22).

Next, for ease of notation, let $\sigma = \delta'/(mn)$ so that $\mathbf{p}' = (p^0_1 - \sigma, p^0_2 - 2\sigma, \ldots, p^0_n - n\sigma)$. For $j < j'$,

$$p'_j - p'_{c_i} = p^0_j - p^0_{c_i} + (c_i - j)\sigma > P_{ij} - P_{ic_i},$$

since $j < c_i$. Thus, when evaluating $f_i(\mathbf{p}')$, the constraints (14) in (DP-D) for $j < j'$ remains non-binding, i.e.,

$$\tau_i \leq p'_j + \mathbb{I}(p'_j - p'_{c_i} > P_{ij} - P_{ic_i})P^{\max} = p'_j + P^{\max}$$

for $j < j'$. Thus, since the price of any product $j \geq j'$ is decreased by $j\sigma \leq n\sigma$, evaluating the new price vector $\mathbf{p}'$ in (DP-D) yields: for all $i \in \mathscr{C}$,

$$f_i(\mathbf{p}') \geq \mathbb{I}(p'_{c_i} < P_{ic_i})(\tau_i^0 - n\sigma) = \mathbb{I}(p^0_{c_i} - c_i\sigma < P_{ic_i})(\tau_i^0 - n\sigma) \geq y^0_{ic_i}\tau_i^0 - n\sigma,$$

where the last inequality follows from the fact that inequality (21) implies $p^0_{c_i} \leq P_{ic_i}$, which implies $p^0_{c_i} - c_i\sigma < P_{ic_i}$. Finally, summing the above inequality over all i, we have:

$$\sum_{i\in\mathscr{C}} f_i(\mathbf{p}') \geq \sum_{i\in\mathscr{C}} (y^0_{ic_i}\tau_i^0 - n\sigma) = g(0) - \delta',$$

concluding the proof. □

Proof of Proposition 7. We first show by contradiction that all prices are the same at optimality. To this end, note from inequality (22) with $\epsilon = 0$ that, for any $i \in \mathscr{C}$, if a product $j \neq c_i$ is *not* eligible for purchase (i.e., $y_{ij} = 0$) then $p_j \geq p_{c_i}$, since all historical prices for i are the same. Consider now an optimal solution $(\mathbf{p}^*, \boldsymbol{\tau}^*, \mathbf{y}^*)$ and let $p^{\min} \equiv \min_{j\in\mathscr{P}} p^*_j$ be the minimum optimal price. Suppose there exists some customer $i \in \mathscr{C}$ who selects a product j such that $\tau^*_i = p^*_j > p^{\min}$. But this implies that $p^{\min} < p^*_j \leq p^*_{c_i}$ (since c_i is always feasible), i.e., we must have $y^*_{ij} = 1$ which by inequality (20) implies that $\tau^*_i \leq p^{\min}$, a contradiction.

Finally, let p^* be the optimal (scalar) price of all products, and suppose $i' \in \mathscr{C}$ is the smallest customer index such that $p^* \leq P_{i'}$. It follows that any customer $i < i'$ does not purchase any product (since $p^* = p^*_{c_i} > P_i$), while all customers $i \geq i'$ yield a revenue of p^* (since $p^* = p^*_{c_i} \leq P_i$). Thus, we must have $p^* = P_{i'}$ at optimality, and the total revenue is $(m - i' + 1)P_{i'}$. The element i^* in the proposition statement is the index that maximizes this revenue. □

Proof of Proposition 8. The solution value of the proposed solution (i.e., a lower bound to the problem) is $\sum_{i\in\mathscr{C}} P_{c_i}$, since all customers would purchase their choice c_i. Due to inequality (20), this is also an upper bound, and hence is optimal. □

Proof of Proposition 9. By definition (28), the total revenue is at least $m\underline{P}$. Conversely, by inequalities (20) and (21), the total revenue is at most $\sum_{i\in\mathscr{C}} P_{ic_i} \leq m\bar{P}$. The ratio hence follows from dividing the lower bound by the upper bound.

We now construct an instance where this ratio is asymptotically tight. Consider $n = 1$ product, $m > 1$ customers, and fix any P_1, P_2 such that $P_1 < P_2$. Customer 1 purchases the product at price P_1, while the remaining customers $2, \ldots, m$ purchase it at P_2. The conservative pricing (28) will set the product's price at P_1, yielding a total revenue of mP_1. For any sufficiently large m, the optimal pricing strategy sets P_2 as the optimal price, yielding a total revenue of $(m-1)P_2$ (since customer 1 will not purchase any product). The performance ratio is therefore $mP_1/(m-1)P_2$. Taking the limit $m \to +\infty$ with respect to this ratio completes the proof. □

Proof of Proposition 10. Without loss of generality, suppose the customer index set $\mathscr{C}$ is ordered according to historical purchase prices, i.e., $0 < \underline{P} = P_{1c_1} \leq P_{2c_2} \leq \cdots \leq P_{mc_m} = \bar{P}$. Let τ^{OPT} and τ^{CP} be the optimal solution value of (OP-MIP) and the total revenue obtained by the cut-off price (30), respectively. It follows from (20) for $j = c_i$ that $\tau^{\text{OPT}} \leq \sum_{i\in\mathscr{C}} P_{ic_i}$. Furthermore, given the price ordering, notice that (29) evaluates to $\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}$. By the cut-off pricing definition (30), we therefore have $\tau^{\text{CP}} \geq \max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}$. This is because due to the incentive-compatibility constraint (22) some customers might not purchase the products priced at the cut-off price p^* and opt for a product that is priced higher than p^*. Thus,

$$\frac{\tau^{\text{CP}}}{\tau^{\text{OPT}}} \geq \frac{\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}}{\sum_{i\in\mathscr{C}} P_{ic_i}}. \tag{EC.15}$$

Next, if we divide both the numerator and denominator of the right-hand side ratio above by the number of customers $m \geq 1$, we obtain

$$\frac{\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}}{\sum_{i\in\mathscr{C}} P_{ic_i}} = \frac{\frac{\max_{i\in\mathscr{C}}(m-i+1)P_{ic_i}}{m}}{\frac{\sum_{i\in\mathscr{C}} P_{ic_i}}{m}} = \frac{\max_{i\in\mathscr{C}}[1-F_{\mathcal{X}}(P_{ic_i})]P_{ic_i}}{\mathbb{E}[\mathcal{X}]}, \tag{EC.16}$$

where $\mathcal{X}$ is a non-negative discrete random variable uniformly distributed on the set $\{P_{1c_1}, \ldots, P_{mc_m}\}$, and $F_{\mathcal{X}}(\cdot)$ and $\mathbb{E}[\mathcal{X}]$ denote the left continuous c.d.f. (i.e.,$F_{\mathcal{X}}(x) \equiv P(\mathcal{X} < x)$) and the expectation of $\mathcal{X}$, respectively. This problem now bears resemblance to the personalized pricing problem studied in Elmachtoub et al. (2020).

Let $R \equiv \max_{i\in\mathscr{C}}[1-F_{\mathcal{X}}(P_{ic_i})]P_{ic_i}$ be the numerator of the ratio above. We have $\underline{P} \leq R \leq \bar{P}$; in particular, the left-hand side inequality holds since $[1-F_{\mathcal{X}}(\underline{P})]\underline{P} = \underline{P}$. We can hence rewrite the expectation term as

$$\begin{aligned}
\mathbb{E}[\mathcal{X}] = \sum_{i\in\mathscr{C}} \frac{1}{m} P_{ic_i} &= \sum_{i\in\mathscr{C}} \frac{1}{m}\left[\int_0^{P_{ic_i}} 1\,\mathrm{d}x\right] \\
&= \sum_{i\in\mathscr{C}} \frac{1}{m}\left[\int_0^{P_{ic_i}} 1\,\mathrm{d}x + \int_{P_{ic_i}}^{+\infty} 0\,\mathrm{d}x\right] \\
&= \sum_{i\in\mathscr{C}} \frac{1}{m}\left[\int_0^{+\infty} \mathbb{I}(x \leq P_{ic_i})\,\mathrm{d}x\right] \\
&= \int_0^{+\infty} \sum_{i\in\mathscr{C}} \left[\frac{1}{m}\mathbb{I}(x \leq P_{ic_i})\right] \mathrm{d}x \\
&= \int_0^{+\infty} [1-F_{\mathcal{X}}(x)]\,\mathrm{d}x \\
&= \int_0^{\bar{P}} [1-F_{\mathcal{X}}(x)]\,\mathrm{d}x \\
&= \int_0^{R} [1-F_{\mathcal{X}}(x)]\,\mathrm{d}x + \int_R^{\bar{P}} [1-F_{\mathcal{X}}(x)]\,\mathrm{d}x
\end{aligned}$$

$$\le R + \int_R^{\bar{P}} [1 - F_{\mathcal{X}}(x)]\,\mathrm{d}x$$
$$\le R + \int_R^{\bar{P}} \frac{R}{x}\,\mathrm{d}x$$
$$= R + R\log\left(\frac{\bar{P}}{R}\right)$$
$$\le R + R\log\left(\frac{\bar{P}}{\underline{P}}\right),$$

where the previous-to-the-last inequality follows because $1 - F_{\mathcal{X}}(x) \le R/x$ for any $x \in [\underline{P}, \bar{P}]$ by definition. From the inequality above, we obtain

$$\frac{\mathbb{E}[\mathcal{X}]}{R} \le 1 + \log\left(\frac{\bar{P}}{\underline{P}}\right) \Leftrightarrow \frac{\tau^{\mathrm{CP}}}{\tau^{\mathrm{OPT}}} \ge \frac{R}{\mathbb{E}[\mathcal{X}]} \ge \frac{1}{1 + \log\left(\frac{\bar{P}}{\underline{P}}\right)}.$$

Finally, from (EC.15) and (EC.16) we have that

$$\frac{\tau^{\mathrm{CP}}}{\tau^{\mathrm{OPT}}} \ge \frac{\max_{i\in\mathscr{C}}[1 - F_{\mathcal{X}}(P_{ic_i})]P_{ic_i}}{\mathbb{E}[\mathcal{X}]} \ge \frac{(1 - F_{\mathcal{X}}(P_{kc_k}))P_{kc_k}}{\mathbb{E}[\mathcal{X}]} \ge \frac{P_{kc_k}}{2\mathbb{E}[\mathcal{X}]} \ge \frac{med(P)}{2\mathbb{E}[\mathcal{X}]}$$

where $med(P)$ denotes the median of $\mathcal{X}$, while $P_{kc_k} = \min_{i\in\mathscr{C}}\{P_{ic_i} : P_{ic_i} \ge med(P)\}$. The prior to the last inequality follows from the fact that $1 - F_{\mathcal{X}}(P_{kc_k}) \ge \frac{1}{2}$ (as $F_{\mathcal{X}}(\cdot)$ is defined to be left continuous), while the last inequality follows from $P_{kc_k} \ge med(P)$, proving the performance bound.

It remains to show that the ratio is asymptotically tight. Consider an instance with any number $m > 0$ of customers and $m - k + 1$ products, where $k \in \mathbb{N}$ is any positive integer such that $k \le m$. The historical observed prices $\mathbf{P}_i$ by customer i, in turn, are defined by the following vectors, where δ is any scalar such that $0 < \delta < 1/(m-1)$:

$$\mathbf{P}_1 = \left(\frac{m}{m}, \frac{m}{m-1}, \frac{m}{m-2}, \dots, \frac{m}{k+1}, \frac{m}{k}\right),$$
$$\mathbf{P}_2 = \left(\frac{m}{m} + \delta, \frac{m}{m-1}, \frac{m}{m-2}, \dots, \frac{m}{k+1}, \frac{m}{k}\right),$$
$$\mathbf{P}_3 = \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2}, \dots, \frac{m}{k+1}, \frac{m}{k}\right),$$
$$\dots$$
$$\mathbf{P}_{m-k} = \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \dots, \frac{m}{k+1}, \frac{m}{k}\right),$$
$$\mathbf{P}_{m-k+1} = \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \dots, \frac{m}{k+1} + \delta, \frac{m}{k}\right),$$
$$\mathbf{P}_{m-k+2} = \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \dots, \frac{m}{k+1} + \delta, \frac{m}{k}\right),$$
$$\dots$$
$$\mathbf{P}_{m} = \left(\frac{m}{m} + \delta, \frac{m}{m-1} + \delta, \frac{m}{m-2} + \delta, \dots, \frac{m}{k+1} + \delta, \frac{m}{k}\right).$$

For the historical purchase choices, each customer $i \in \{1, \ldots, m-k\}$ purchased product $c_i = i$, while all remaining customers $i' \in \{m-k+1, \ldots, m\}$ pick the same product $c_{i'} = m-k+1$.

An upper bound on the optimal revenue τ^{OPT} for this instance is:

$$\tau^{\text{OPT}} \leq \sum_{i \in \mathscr{C}} P_{ic_i} = \sum_{i=1}^{m-k+1} P_{ii} + \sum_{i=m-k+2}^{m} P_{i(m-k+1)} = \sum_{i=0}^{m-k} \frac{m}{m-i} + \sum_{i=m-k+1}^{m-1} \frac{m}{k} = \sum_{i=0}^{m-k} \frac{m}{m-i} + m\frac{k-1}{k}.$$

We now define prices whose resulting total revenue is arbitrarily close to τ^{OPT}. Specifically, consider the price vector

$$\mathbf{p} = \left(\frac{m}{m}, \frac{m}{m-1} - \delta, \frac{m}{m-2} - 2\delta, \ldots, \frac{m}{k} - (m-k)\delta\right).$$

From the definition above, since $p_{c_i} \leq P_{ic_i}$ for all $i \in \mathscr{C}$, by inequality (21) every customer purchases a product. We next show that customer i purchases product c_i. For any $j < c_i$,

$$p_j - p_{c_i} = \frac{m}{m-j+1} - \frac{m}{m-c_i+1} - j\delta + c_i\delta \geq P_{ij} - P_{ic_i},$$

i.e., such products j violates incentive-compatibility constraints (inequality (22) with $\epsilon = 0$) and will not be purchasable by customer i. Moreover, from our choice of $\delta < 1/(m-1)$, it can be easily verified that $p_1 < p_2 < \cdots < p_{m-k+1}$ and therefore $p_j > p_{c_i}$ for all i and $j > c_i$. Product p_{c_i} must be necessarily chosen by customer i under her worst-case valuation ($v_{ic_i} = P_{ic_i}$ and $v_{ik} = 0$ for all $k \in \mathscr{P} \setminus \{c_i\}$) and the revenue from this pricing is hence

$$\sum_{i \in \mathscr{C}} p_{c_i} = \sum_{i=1}^{m-k+1} p_i + (k-1)p_{m-k+1} = \sum_{i=0}^{m-k} \left(\frac{m}{m-i} - i\delta\right) + m\frac{k-1}{k} - (k-1)(m-k)\delta,$$

thus, as $\delta \to 0$, the total revenue obtained from $\mathbf{p}$ approximates that of τ^{OPT}.

We now show that the revenue obtained from the cut-off pricing (30) is m. Suppose that the customer index that solves (29) is i', and hence $p^* = P_{i'c_{i'}}$. If $c_{i'} < m-k+1$, then the cut-off prices are $p_j^{CP} = \frac{m}{k}$ for all $j < c_{i'}$, $p_j^{CP} = \frac{m}{m-j+1}$ for all $c_{i'} \leq j < m-k+1$ and $p_{m-k+1}^{CP} = \frac{m}{k}$. However, if $c_{i'} = m-k+1$, the cut-off prices are $p_j^{CP} = \frac{m}{k}$ for all $j \in \mathscr{P}$. Suppose that $c_{i'} < m-k+1$. By the construction of historical prices,

$$P_{i'c_{i'}} = \frac{m}{m-c_{i'}+1},$$

and therefore $m - c_{i'} + 1$ customers would purchase a product since $P_{ic_i} \geq P_{i'c_{i'}}$ for all $i \geq i'$. Analogously, since by construction for all customers $i < i'$, $P_{ic_i} < P_{i'c_{i'}}$, none of the worst-case historical customers $i < i'$ would purchase any product. Moreover, for any of the worst-case customer $i \geq i'$, the cut-off price (30) of its chosen product c_i is

$$p_{c_i}^{\text{CP}} = \frac{m}{m-c_i+1}.$$

This implies that, for any $i > i'$,

$$p^{\mathrm{CP}}_{c_{i'}} - p^{\mathrm{CP}}_{c_i} = \frac{m}{m-c_{i'}+1} - \frac{m}{m-c_i+1} < \left(\frac{m}{m-c_{i'}+1}+\delta\right) - \frac{m}{m-c_i+1} = P_{i'c_{i'}} - P_{ic_i},$$

i.e., product $c_{i'}$ is incentive-compatible with all customers $i > i'$. Because $p^{\mathrm{CP}}_{c_{i'}}$ is the lowest price across all the products, the total revenue of the cut-off solution

$$\tau^{\mathrm{CP}} = (m-c_{i'}+1)P_{i'c_{i'}} = (m-c_{i'}+1)\frac{m}{m-c_{i'}+1} = m$$

If $c_{i'} = m-k+1$, as mentioned before the cut-off prices are $p^{CP}_j = \frac{m}{k}$ for all $j \in \mathscr{P}$ and it can also be easily confirmed that $\tau^{\mathrm{CP}} = m$ as only the k historical customers who historically purchased product $m-k+1$ will make a purchase, under price $\frac{m}{k}$, in the worst-case.

Finally, as $\delta \to 0$,

$$\begin{aligned}\frac{\tau^{\mathrm{CP}}}{\tau^{\mathrm{OPT}}} \to \frac{m}{\sum_{i=0}^{m-k}\frac{m}{m-i}+m\frac{k-1}{k}} &= \frac{m}{m\left(\sum_{i=0}^{m-k}\frac{1}{m-i}+\frac{k-1}{k}\right)}\\ &= \frac{1}{\frac{1}{m}+\cdots+\frac{1}{k}-\frac{1}{k}+\frac{k}{k}}\\ &= \frac{1}{\sum_{i=k+1}^{m}\frac{1}{i}+\frac{k}{k}}.\end{aligned}$$

For a sufficiently large m and k (e.g., by multiplying both by the same constant), the ratio above can be made sufficiently close to notespar,

$$\frac{1}{\log m - \log k + \frac{1}{2m} - \frac{1}{2k} + 1} = \frac{1}{\log\left(\frac{m}{k}\right) + \frac{1}{2m} - \frac{1}{2k} + 1},$$

where the last equality follows from the fact that for large enough n, $\sum_{i=1}^{n}\frac{1}{i} = \log(n) + \gamma + \frac{1}{2n}$ where γ is the EulerMascheroni constant. The ratio above can approximate $1/(1+\log(\bar{P}/\underline{P}))$ at any desired precision since m/k can be made sufficiently close to $\bar{P}/\underline{P}$, while both numbers are also sufficiently large. Thus, the ratio is asymptotically tight for the constructed instance. □

EC.3. Extending External Validity to the MNL Model

When there are multiple products, the external validity analysis of our approach becomes significantly more complicated, primarily because both the model-free and model-based optimal prices are not tractable. Given such hurdles, we focus on a simple yet relevant special case, where for some $\beta > 0$, the probability of customer i choosing product j from the assortment is given by

$$\frac{\exp(\alpha - \beta P_{ij})}{1+\sum_{k=1}^{n}\exp(\alpha-\beta P_{ik})}. \tag{EC.17}$$

That is, we limit our scope to the situation where customers make decisions based on the MNL choice model and the average attractiveness of the products is symmetric (the products are horizontally differentiated). Moreover, we assume the historical prices satisfy $P_{ij} \equiv P_i$ for all $i \in \mathscr{C}$. Such a uniform price is common in some settings, e.g., as noted by Draganska and Jain (2006), it is a well established policy to offer all flavors in a product line of yogurt at the same price. In this case, implied by Proposition 7, model-free optimal prices are the same for all the products and can be expressed in a closed form for uniformly distributed historical prices. We denote the model-free optimal price of all the products as p^*. Moreover, given that customers make decisions based on the MNL choice model, the model-based optimal prices are equal for all the products, denoted by $\hat{p}$. We also define the expected revenue of the firm when all the products are priced at p under the MNL model (EC.17) as $\Upsilon(p)$. It turns out the insights developed for a single product in Theorem 1 can be extended to this special case with multiple products.

Proposition EC.1. *Suppose the firm sets prices P_i from a distribution with PDF $g(p) = 1/b$ for $p \in [0, b]$, where $b > \hat{p}$. We have $\frac{\Upsilon(p^*)}{\Upsilon(\hat{p})} \geq \frac{1}{2}$. Moreover, the bound is asymptotically tight: for any $\epsilon > 0$ that is sufficiently small, a case can be constructed in which there exists $\bar{\alpha}$ such that when $\alpha > \bar{\alpha}$, we have $\Upsilon(p^*)/\Upsilon(\hat{p}) \leq \frac{1}{2} + \epsilon$.*

We note that the bound is asymptotically tight since the worst-case holds when $\alpha \to \infty$.

Proof of Proposition EC.1. First, we will show that p^* must be unique. By Proposition 7, the optimal price output by our framework OP-MIP satisfies $p^* \in \arg\max_{p\geq 0} p \sum_{i=1}^m \mathbb{I}(P_{ic_i} \geq p)$. Dividing the latter part of this quantity by the number of historical customers m does not affect the value of p^*. In other words, $p^* \in \arg\max_{p\geq 0} p \frac{\sum_{i=1}^m \mathbb{I}(P_{ic_i} \geq p)}{m}$.

By the Law of Large Numbers we have

$$\lim_{m\to\infty} \frac{\sum_{i=1}^m \mathbb{I}(P_{ic_i} \geq p)}{m} = \mathsf{P}(P_{ic_i} \geq p) = \int_p^b \frac{1}{b}(1 - \frac{1}{ne^{\alpha-\beta x}+1})\mathrm{d}x.$$

Therefore

$$p^* \in \arg\max_{p\geq 0} p \lim_{m\to\infty} \frac{\sum_{i=1}^m \mathbb{I}(P_{ic_i} \geq p)}{m} = \arg\max_{0\leq p\leq b} p \int_p^b (1 - \frac{1}{ne^{\alpha-\beta x}+1})\mathrm{d}x$$

Now, notice that since

$$0\int_0^b (1 - \frac{1}{ne^{\alpha-\beta x}+1})\mathrm{d}x = b\int_b^b (1 - \frac{1}{ne^{\alpha-\beta x}+1})\mathrm{d}x = 0,$$

and

$$p\int_p^b (1 - \frac{1}{ne^{\alpha-\beta x}+1})\mathrm{d}x \tag{EC.18}$$

is a continuous and differentiable function, not zero everywhere in $[0,b]$ and defined on the compact region of $[0,b]$, it has a maximizer and its maximizer must be in the interior of $[0,b]$, and must satisfy the first-order condition for (EC.18).

Now we will show there is a unique $p^* \in [0,b]$ that satisfies the first-order condition for (EC.18), and then we will show that $p^* \le \hat{p}$.

To that avail, notice that the first-order condition for (EC.18) with respect to p is equivalent to:

$$\left(p\int_p^b (1-\frac{1}{ne^{\alpha-\beta x}+1})\mathrm{d}x\right)' = \int_p^b (1-\frac{1}{ne^{\alpha-\beta x}+1})\mathrm{d}x - p(1-\frac{1}{ne^{\alpha-\beta p}+1}) = \frac{\log(\frac{ne^{\alpha-\beta p}+1}{ne^{\alpha-\beta b}+1})}{\beta} - p(1-\frac{1}{ne^{\alpha-\beta p}+1}) = 0.$$

To show that p^* must be unique, we need to show

$$q(p) = \frac{\log(\frac{ne^{\alpha-\beta p}+1}{ne^{\alpha-\beta b}+1})}{\beta} - p(1-\frac{1}{ne^{\alpha-\beta p}+1}) = 0$$

has only one root in $[0,b]$. Notice that

$$q(b) = \frac{\log(\frac{ne^{\alpha-\beta b}+1}{ne^{\alpha-\beta b}+1})}{\beta} - b(1-\frac{1}{ne^{\alpha-\beta b}+1}) < 0.$$

Moreover,

$$q(0) = \frac{\log(\frac{ne^{\alpha-\beta 0}+1}{ne^{\alpha-\beta b}+1})}{\beta} - 0(1-\frac{1}{ne^{\alpha-\beta 0}+1}) > 0,$$

thus there must be at least one root.

Now notice that

$$q'(p) = \left(\frac{\log(\frac{ne^{\alpha-\beta p}+1}{ne^{\alpha-\beta b}+1})}{\beta} - p(1-\frac{1}{ne^{\alpha-\beta p}+1})\right)' = \frac{n\beta e^{\alpha}((\beta p-2)e^{\beta p} - 2ne^{\alpha})}{(e^{\beta p}+ne^{\alpha})^2},$$

which is <0 if $p < \frac{w_0(2ne^{\alpha-2})+2}{\beta}$ and is >0 if $p > \frac{w_0(2ne^{\alpha-2})+2}{\beta}$. If $\frac{w_0(2ne^{\alpha-2})+2}{\beta} > b$, then $q(p)$ is strictly decreasing in $[0,b]$ and has exactly one root in $[0,b]$. If $\frac{w_0(2ne^{\alpha-2})+2}{\beta} \le b$, then since $q(b)<0$ and $q(p)$ is strictly increasing on $(\frac{w_0(2ne^{\alpha-2})+2}{\beta}, b]$, $q(p)$ can have only one root, p^*, on $[0,b]$.

Now, we show that $p^* \le \hat{p}$. Assume otherwise, that for some α, b, n and $\beta > 0$, $p^* > \hat{p}$. It is trivial to see that if

$$\frac{\log(ne^{\alpha-\beta p}+1)}{\beta} - p(1-\frac{1}{ne^{\alpha-\beta p}+1}) = 0 \tag{EC.19}$$

has a root, its root must be larger than or equal to p^*. To show that (EC.19) also has a unique root it suffices to notice that

$$\frac{\log(ne^{\alpha-\beta 0}+1)}{\beta} - 0(1-\frac{1}{ne^{\alpha-\beta 0}+1}) > 0,$$

and

$$\lim_{p\to\infty}(\frac{\log(ne^{\alpha-\beta p}+1)}{\beta} - p(1-\frac{1}{ne^{\alpha-\beta p}+1})) = 0,$$

while

$$(\frac{\log(ne^{\alpha-\beta p}+1)}{\beta}-p(1-\frac{1}{ne^{\alpha-\beta p}+1}))'=\frac{n\beta e^{\alpha}((\beta p-2)e^{\beta p}-2ne^{\alpha})}{(e^{\beta p}+ne^{\alpha})^2},$$

which is <0 if $p<\frac{w_0(2ne^{\alpha-2})+2}{\beta}$ and is >0 if $p>\frac{w_0(2ne^{\alpha-2})+2}{\beta}$.

Now let $\bar{p}$ be the single root of (EC.19). We know $\bar{p}\geq p^*>\hat{p}$. Notice that the derivative of (EC.19) with respect to α is

$$\frac{ne^{\alpha}(ne^{\alpha}+(1-\beta p)e^{\beta p})}{\beta(e^{\beta p}+ne^{\alpha})^2},$$

which is <0 if $p>\frac{w_0(ne^{\alpha-1})+1}{\beta}=\hat{p}$. Thus, (EC.19) is strictly decreasing in α at $\bar{p}$. Moreover, $\frac{w_0(ne^{\alpha-1})+1}{\beta}$ is increasing in α, hence

$$\lim_{\alpha\to-\infty}(\frac{\log(ne^{\alpha-\beta\bar{p}}+1)}{\beta}-\bar{p}(1-\frac{1}{ne^{\alpha-\beta\bar{p}}+1}))>0,$$

however this is a contradiction since

$$\frac{\log(ne^{\alpha-\beta p}+1)}{\beta}-p(1-\frac{1}{ne^{\alpha-\beta p}+1})\leq 0$$

for all $p\geq 0$ when $\alpha\to-\infty$. The latter claim follows from the fact that $\frac{\log(ne^{\alpha-\beta p}+1)}{\beta}$ is decreasing in p and $\frac{\log(ne^{\alpha}+1)}{\beta}=0$ as $\alpha\to-\infty$ while $p(1-\frac{1}{ne^{\alpha-\beta p}+1})\geq 0$ if $p\geq 0$. As a result, by contradiction we can conclude that $\bar{p}\leq\hat{p}$ which implies $p^*\leq\hat{p}$ since we know $p^*\leq\bar{p}$.

Now we show the revenue from p^* is at least $\frac{1}{2}$ times the revenue from $\hat{p}$. We have $p^*\leq\hat{p}\leq b$. Notice that

$$\left(\frac{\log(\frac{ne^{\alpha-\beta p}+1}{ne^{\alpha-\beta b}+1})}{\beta}+p(1-\frac{1}{ne^{\alpha-\beta p}+1})\right)'=-\frac{n\beta pe^{\beta p+\alpha}}{(e^{\beta p}+ne^{\alpha})^2}\leq 0 \tag{EC.20}$$

for all $p\geq 0$. Moreover, for all $0\leq p\leq b$,

$$\frac{\log(\frac{ne^{\alpha-\beta p}+1}{ne^{\alpha-\beta b}+1})}{\beta}\geq 0,$$

thus

$$\Upsilon(\hat{p})=\hat{p}(1-\frac{1}{ne^{\alpha-\beta\hat{p}}+1})\leq\frac{\log(\frac{ne^{\alpha-\beta\hat{p}}+1}{ne^{\alpha-\beta b}+1})}{\beta}+\hat{p}(1-\frac{1}{ne^{\alpha-\beta\hat{p}}+1})\leq\frac{\log(\frac{ne^{\alpha-\beta p^*}+1}{ne^{\alpha-\beta b}+1})}{\beta}+p^*(1-\frac{1}{ne^{\alpha-\beta p^*}+1})=2\Upsilon(p^*).$$

Where he second inequality follows from (EC.20) and the fact that $0\leq p^*\leq\hat{p}$. The last equality follows from the fact that p^* must satisfy the first-order condition for (EC.18). Thus, we have shown that $\Upsilon(p^*)\geq\frac{1}{2}\Upsilon(\hat{p})$.

It remains to show the bound is tight. Consider a case where $n=2$, $\beta=1$, and $b\gg\alpha$. Let $0<\epsilon_1\ll\epsilon$. For any $\sigma>0$, there exists $\tilde{\alpha}$ such that if $\alpha>\tilde{\alpha}$, for any $0\leq p\leq(1-\epsilon_1)\alpha$, $|\log(\frac{2e^{\alpha-p}+1}{2e^{\alpha-b}+1})-\log(2e^{\alpha-p})|<\sigma$, and $|p(1-\frac{1}{2e^{\alpha-p}+1})-p|<\sigma$. Thus, for any $\sigma>0$, there exists $\tilde{\alpha}$ such that if $\alpha>\tilde{\alpha}$, then $|p^* - \frac{\alpha+\log(2)}{2}| < \sigma$ (i.e., p^* can be arbitrarily close to the root of $\alpha - p + log(2) - p = 0$). Hence, for any $\sigma > 0$, there exists $\tilde{\alpha}$ such that if $\alpha > \tilde{\alpha}$, $|\Upsilon(p^*) - \frac{\alpha+\log(2)}{2}| < \sigma$. Now, we notice that for any $\sigma > 0$, there exists $\tilde{\alpha}$ such that if $\alpha > \tilde{\alpha}$, then $|\Upsilon((1-\epsilon_1)\alpha) - (1-\epsilon_1)\alpha| < \sigma$. Finally,

$$\lim_{\alpha\to\infty} \frac{\Upsilon(p^*)}{\Upsilon(\hat{p})} \leq \lim_{\alpha\to\infty} \frac{\frac{\alpha+\log(2)}{2}}{(1-\epsilon_1)\alpha} = \frac{1}{2(1-\epsilon_1)} \leq \frac{1}{2} + \epsilon,$$

proving the bound is asymptotically tight. □

EC.4. The Effect of Data Censoring on Pricing

In this section we study the effect of data censoring on pricing. First, we characterize how censoring distorts the estimated customer demand model from the data, in the case of single product pricing. Next, we study a simple, yet illuminating case of demand and we provide a sufficient condition under which the asymptotic model-free optimal price outperforms the asymptotic optimal price estimated from the censored data. Finally, we extend the insights derived for the problem of single product pricing to that of multiple products pricing, when customers are making decisions based on the symmetric MNL choice model.

EC.4.1. Single Product

LEMMA EC.1. *Suppose Assumption 1 holds. Then, if at each historical price only a fraction z of non-purchasing customers are observed, the asymptotic estimated purchase probability of a customer at price p would be $\hat{F}(\cdot) = \frac{1-F(p)}{1-F(p)+zF(p)}$.*

Proof of Lemma EC.1. Let $B(p)$ be the number of customers that purchase the product at price p in the data, and let $N(p)$ indicate the number of non-purchasing customers at price p. By assumption, only $zN(p)$ of these non-purchasing customers are recorded in the data. By definition, at any price p, we have $1 - \hat{F}(p) = \frac{B(p)}{B(p)+zN(p)}$, while $1 - F(p) = \frac{B(p)}{B(p)+N(p)}$. Therefore,

$$B(p) = (1 - \hat{F}(p))(B(p) + zN(p)) = (1 - F(p))(B(p) + N(p)).$$

Thus we have

$$1 - \hat{F}(p) = (1 - F(p))\frac{B(p) + N(p)}{B(p) + zN(p)} = (1 - F(p))\frac{\frac{B(p)+N(p)}{B(p)+N(p)}}{\frac{B(p)+zN(p)}{B(p)+N(p)}} = \frac{1 - F(p)}{1 - F(p) + zF(p)}.$$

Where the last equality follows from the fact that

$$\frac{B(p) + zN(p)}{B(p) + N(p)} = \frac{B(p)}{B(p) + N(p)} + \frac{zN(p)}{B(p) + N(p)} = 1 - F(p) + zF(p).$$

□

Proposition EC.2. *Suppose Assumptions 1 and 2 hold. Moreover, assume* $F(p)=\frac{p}{a}$*. We have:*

$$\frac{R(p^*)}{R(\hat{p})}\geq\frac{8}{9}.$$

Proof of Proposition EC.2. From Proposition 1, we have

$$p^*\in\arg\max_{p\geq 0} p\int_p^{+\infty} g(x)(1-F(x))\mathrm{d}x=\arg\max_{p\geq 0}\frac{p}{ab}\int_p^a(a-x)\mathrm{d}x=\arg\max_{p\geq 0}p(ax-\frac{x^2}{2})\Big|_p^a=\arg\max_{p\geq 0}p(\frac{a^2}{2}-ap+\frac{p^2}{2}).$$

By taking the first-order condition with respect to p we obtain

$$\frac{a^2}{2}-2ap+\frac{3p^2}{2}=0,$$

Thus, $p^*=\frac{a}{3}$, which results in a revenue of $\frac{2a}{9}$, while the optimal revenue is $\frac{a}{4}$, proving the claimed ratio. We note that $p=a$ is also a root of $\frac{a^2}{2}-2ap+\frac{3p^2}{2}=0$. However, $a\int_a^b g(x)(1-F(x))\mathrm{d}x=0$, and hence

$$a\neq\arg\max_{p\geq 0} p\int_p^b g(x)(1-F(x))\mathrm{d}x.$$

□

Corollary EC.1. *Suppose Assumptions 1 and 2 hold. Moreover, let* $F(p)=\frac{p}{a}$*. Then as long as at each historical price, only a fraction* $z<0.25$ *of non-purchasing customers are recorded in the data, the asymptotic optimal model-free price* p^* *generates a higher revenue than that of the model-based optimal price, estimated from the censored data.*

Proof of Corollary EC.1. By Lemma EC.1, the estimated asymptotic purchase probability at price p would be $\frac{a-p}{a-p+zp}$ and the model-based optimal price will be the optimizer of $p(\frac{a-p}{a-p+zp})$. By taking the first order condition, we note that the optimal model-based price would be $\frac{a(1-\sqrt{z})}{1-z}$ leading to a revenue of $\frac{a(1-\sqrt{z})(\sqrt{z}-z)}{(1-z)^2}$. To conclude the proof, we note that from Proposition EC.2, the revenue from p^* would be $\frac{2a}{9}$ and it is easy to verify that $\frac{(1-\sqrt{z})(\sqrt{z}-z)}{(1-z)^2}<\frac{2}{9}$ as long as $z<0.25$. □

EC.4.2. Multiple Products

We note that with a slight abuse of notation, we can extend the insight from Corollary EC.1 to the setting with multiple products explored in Section EC.3. We assume the historical prices satisfy $P_{ij}\equiv P_i$ for all $i\in\mathscr{C}$. Moreover, we assume that customers make decisions based on the MNL choice model specified in Equation (EC.17). Then, the following corollary specifies the estimated MNL optimal price from censored data in terms of $w_0(\cdot)$, where $w_0(\cdot)$ indicates the positive part of the lambert function and is defined as $w_0(x)e^{w_0(x)}=x,\ x\geq 0$.

Corollary EC.2. *Suppose the firm samples the price vector seen by each past customer i, P_i, from a distribution with PDF $g(p) = 1/b$ for $p \in [0, b]$, where $b > \hat{p} = \frac{w_0(ne^{\alpha-1})+1}{\beta}$. If at each historical price vector, only a fraction z of non-purchasing customers are recorded in the data, the asymptotic estimated MNL optimal price can be expressed as $\bar{p} = \frac{w_0(\frac{ne^{\alpha-1}}{z})+1}{\beta}$. Moreover, there exists $\hat{z} > 0$ such that as long as $0 \le z < \hat{z}$ the asymptotic optimal model-free price p^* generates a higher revenue than $\bar{p}$, the model-based optimal price, estimated from the censored data.* □

Proof of Corollary EC.2. With a slight abuse of notation we can notice that Lemma EC.1 is applicable here. Thus, at each given price p, the asymptotic estimated probability of purchase can be expressed as:

$$\hat{F}(p) = \frac{ne^{\alpha-\beta p}}{ne^{\alpha-\beta p} + z}$$

and hence $\bar{p}$ will be the optimizer of:

$$\frac{pne^{\alpha-\beta p}}{ne^{\alpha-\beta p} + z}.$$

By taking the first order condition, it can be easily verified that $\bar{p} = \frac{w_0(\frac{ne^{\alpha-1}}{z})+1}{\beta}$. We note that when there is no censoring in the data, the MNL optimal price is $\hat{p} = \frac{w_0(ne^{\alpha-1})+1}{\beta}$. It follows from Proposition EC.1 that the model-free optimal price p^* will always guarantee at least 50% of the revenue from $\hat{p}$ and is not affected by censoring as mentioned in Remark 1. However, we note that since $0 \le z \le 1$, we have $\bar{p} \ge \hat{p} \ge p^*$ for all z, while $\Upsilon(p)$ is strictly decreasing in p for all $p > \hat{p}$ (which follows from the fact that $\hat{p}$ is the unique point satisfying the first order condition for the pseudo-concave function $\Upsilon(p)$ while $\Upsilon(0) = \Upsilon(\infty) = 0$). Moreover, as the positive part of $w_0(\cdot)$ is known to be a strictly increasing function, $\bar{p}$ is strictly decreasing in z. Therefore, $\Upsilon(\bar{p})$ is strictly increasing in z, becoming zero when $z \to 0$, as $\lim_{z\to 0} \bar{p} = \infty$, leading to $\lim_{z\to 0} \Upsilon(\bar{p}) = 0$. Therefore, there must exist $\hat{z} \in (0, 1]$ such that for all $0 \le z < \hat{z}$, $\Upsilon(\bar{p}) \le \Upsilon(p^*)$. □

EC.5. Supplementary Results and Discussions

Proposition EC.3. *If prices are set to their average historical values, the worst-case fraction between the revenue obtained with this approach and the optimal value w.r.t. (OP-MIP) is asymptotically zero as the number of historical customers grow.*

Proof of Proposition EC.3. Assume we have $n = 1$ product and $m > 1$ historical customers where the purchase price of customers $1, \cdots, (m-1)$, was 1, and the purchase price of customer m was 2. Then, the average historic price of the product is $\frac{m+1}{m}$. At this price, customers $1, \cdots, (m-1)$ will not make a purchase under their worst-case valuations while customer m will purchase the product at price $\frac{m+1}{m}$. However, setting the price of the product at 1 results in a revenue of m, since all customers will purchase the product. Thus, the ratio of the revenue from the average price to the optimal value of (OP-MIP) could be less than or equal to $\frac{m+1}{m^2}$. Taking the limit $m \to +\infty$ with respect to this ratio completes the proof. □

Proposition EC.4. *If prices are set uniformly at random based on the empirical distribution defined by historical prices, the worst-case fraction between the revenue obtained with this approach and the optimal value w.r.t.* (OP-MIP) *is asymptotically zero as the number of historical customers grow.*

Proof of Proposition EC.4. Assume we have n products and n historical customers. The historical prices $\mathbf{P}_i$ observed by customer i, are defined by the following vectors:

$$\mathbf{P}_1 = (1, 2, 2, \ldots, 2),$$
$$\mathbf{P}_2 = (2, 1, 2, \ldots, 2),$$
$$\mathbf{P}_3 = (2, 2, 1, 2, \ldots, 2),$$
$$\ldots$$
$$\mathbf{P}_n = (2, 2, \ldots, 2, 1).$$

For the historical purchase choices, each customer $i \in \{1, \ldots, n\}$ purchased product i.

Assume the vector of prices p is equal to one of $\mathbf{P}_1, \ldots, \mathbf{P}_n$ at random (with equal probability). If $p = \mathbf{P}_1$, then for $i = 1$, we have $P_{1c_1} \leq p_{c_1}$ and hence the revenue from customer 1 is 1. For any customer $i > 1$, we have $P_{ic_i} < p_{c_i}$, hence the revenue form any customer $i > 1$ is zero. Similarly we can show that if $p = \mathbf{P}_i$ for any $i > 1$, the total revenue will be 1, under the worst-case customer valuations.

However, for $p = (1, 1, \ldots, 1)$, the revenue will be n as every customer will make a purchase at price 1. Therefore, the ratio of the expected revenue from the price observed by a random historical customer to the optimal value of (OP-MIP) could be less than or equal to $\frac{1}{n}$. Taking the limit $n \to +\infty$ with respect to this ratio completes the proof. □

In the following example we show that the prices obtained through the LP relaxation of the program (OP-MIP), have a worse worst-case performance than the conservative pricing approach.

Example EC.1. Consider an instance with $m = 3$ customers and $n = 2$ products. The historical prices P_{ij} and the customer choices (in bold) are listed in the table below:

Price	Product 1	Product 2
Customer 1	**1**	2
Customer 2	2	**3**
Customer 3	**1**	3

It can be shown that the LP relaxation yields the price vector $\mathbf{p}^{\text{LP}} = (1.2, 2.3)$ and an upper bound of 4.8 to the optimal value of (OP-MIP). However, when evaluating (OP-MIP) with variables $\mathbf{p}$ fixed to $\mathbf{p}^{\text{LP}}$, both customers 1 and 3 do not purchase any products, while customer 2 purchases product 1 in the worst-case. The following set of valuations, v_{ij}, that are drawn from the IC polyhedra of customers 1, 2 and 3 conform to these customer choices that lead to a worst-case revenue for the firm.

Valuation	Product 1	Product 2
Customer 1	1	0
Customer 2	2	3
Customer 3	1	0

Thus, the total revenue generated from $\mathbf{p}^{\text{LP}}$ is \$1.2. The optimal solution of this instance is \$4.0 certified by prices $\mathbf{p}^* = (1.0, 2.0)$, which leads to an LP solution ratio of 30%, worse than the conservative price ratio of $\underline{P}/\bar{P} = 1/3 \approx 33\%$.

The optimal LP solution associated with variables $\mathbf{y}$ provides insights into the poor performance of the heuristic. In particular, consider inequality (20) for customer 2 ($i = 2$) and product 1 ($j = 1$):

$$\tau_2 \leq p_1 + P_{22}(1 - y_{21}).$$

At the LP optimality, $y_{21} = 0.4$ and the above right-hand side is equal to 3. The inequality is also tight, leading to a (relaxed) revenue of $\tau_2 = 3.0$. However, with integrality constraints, $y_{21} = 1.0$ and the constraint is again tight with $\tau_2 \leq p_1 = 1.2$, which is a significant decrease in revenue. The same issue is identified for the other customers.

More generally, the big-M structure of inequalities (20) tends to result in the optimal LP solution pricing p_j slightly higher than the historical prices for most of the customer i with $c_i = j$, in spite of the fact that the indicator $\mathbb{I}(p_j > P_{ic_j})$ has been encoded in (OP-MIP) to represent the no-purchase option. More precisely, y_{ij} can be set to $1 - \epsilon$ for any sufficiently small ϵ to overcome that condition, and the objective value of the LP is not impacted significantly since the customer is still assumed to purchase the product. However, when the integrality constraint is imposed, either the inequality becomes binding with respect to some price or the no-purchase option is chosen, leading to a smaller expected revenue. □