# Online Appendix to "Distribution-Free Pricing"

## A. Alternative Information Structure

### A.1. Distribution Symmetry

In this subsection, we consider symmetric distributions. In particular, we consider a class of valuation distributions that are symmetric and share the same mean $\mu$ and standard deviation $\sigma$, denoted by $\mathscr{F}_s \subset \mathscr{F}$. We will comment on the purpose of this consideration after presenting our results.

Proposition OA.1. *Assume $F \in \mathscr{F}_s$. The price heuristic*

$$p_s^* = \begin{cases} \mu - k_s^* \sigma & \text{if } \tau \geq \frac{2(\sqrt{\frac{5+\sqrt{17}}{4}})^3 + \sqrt{\frac{5+\sqrt{17}}{4}}}{2} \approx 4.2, \\ \mu & \text{otherwise,} \end{cases}$$

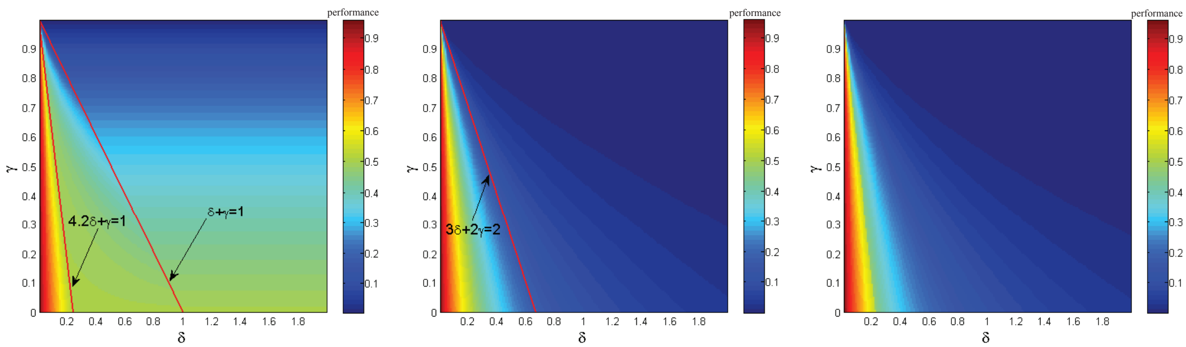
*where* $k_s^* = \sqrt[3]{\frac{1}{2}\tau + \sqrt{\frac{1}{4}\tau^2 + (\frac{1}{6})^3}} + \sqrt[3]{\frac{1}{2}\tau - \sqrt{\frac{1}{4}\tau^2 + (\frac{1}{6})^3}}$, *achieves a sharper performance guarantee. Moreover,* $p^* \leq p_s^*$.

Proposition OA.1 shows that the knowledge of the valuation distribution is symmetric results in less aggressive pricing (i.e., $p_s^* \geq p^*$) and indeed provides more value. The benefits come from two sources. First, technically speaking, instead of applying the one-sided Chebyshev's inequality, we can resort to the two-sided Chebyshev's inequality and obtain sharper lower and upper bounds with the price heuristic $p = \mu - k_s^* \sigma$. Second, by setting the price at the mean valuation, i.e., $p = \mu$, the seller earns a guaranteed profit of $\frac{1}{2}(\mu - c)$, regardless of the valuation distribution as long as it is symmetric. Next, we show the first benefit is limited in the parameter space where the price heuristic is not equal to the mean valuation (i.e., $\tau \geq 4.2$).

Proposition OA.2. *For all $F \in \mathscr{F}_s$, if* $\tau \geq \frac{2(\sqrt{\frac{5+\sqrt{17}}{4}})^3 + \sqrt{\frac{5+\sqrt{17}}{4}}}{2} \approx 4.2$, *then* $\frac{\Pi(p^*;F)}{\Pi(p_s^*;F)} \geq 95\%$.

Proposition OA.2 shows that the first benefit that comes from using the two-sided Chebyshev's inequality is not very significant. That is because symmetry cannot eliminate the extreme case of a symmetric two-point distribution. Moreover, the second benefit is not more valuable than knowing the median of the valuation distribution. In other words, except that the area below the mean is equal to the area above the mean in a distribution, detailed symmetry at other points (other than the mean) is not helpful. One may argue that the symmetric valuation distribution is impractical, even though some theoretical work specifically requires such a distributional assumption in order to derive insights (see, e.g., Fang and Norman 2006). The gain from this somewhat hypothetical exercise is that even though we impose a very restrictive property such as symmetry on the valuation distribution, the benefit can be little beyond knowing the median. This insight is further confirmed

**Figure OA.1 Symmetric or Unimodal Distribution**




(a) (b) (c)

*Note.* (a) Symmetric distribution with both benefits (b) Symmetric distribution with one benefit (not using the median information) (c) Unimodal distribution

when we compare the performance guarantees with and without the knowledge of the unimodality of the valuation distribution. Nevertheless, the percentile information such as the median can be very valuable. By pricing at the $z$-th percentile, $v_z$, of the valuation, the seller can guarantee a profit of $z(v_z - c)$.

Figure OA.1(a) shows the performance of symmetric distributions. It is obviously better than the base model in comparison with Figure 1. By Proposition OA.1, if $\tau = \frac{1-\gamma}{\delta} \leq 4.2$, i.e., to the right of the line $4.2\delta + \gamma = 1$, the price heuristic is $\mu$. Thus we can see that the improvement beyond the base model is due to setting the price to the median $\mu$, namely, the second benefit. Moreover, by Proposition OA.1, if $\tau = \frac{1-\gamma}{\delta} \leq 1$, i.e., to the right of the line $\delta + \gamma = 1$, with the price heuristic being $\mu$, the performance guarantee becomes $\frac{1-\gamma}{2-\gamma}$, independent of $\delta$. To isolate those two benefits, Figure OA.1(b) shows the performance guarantee without letting the price heuristic be equal to the mean. In other words, we display only the first benefit. Comparing Figures 1 and OA.1(b), we see that applying the two-sided Chebyshev's inequality affords little or *even no* benefit beyond applying the one-sided Chebyshev's inequality. In particular, when $\tau = \frac{1-\gamma}{\delta} \leq 1.5$, i.e., to the right of the line $3\delta + 2\gamma = 2$, $k_s^*$ given by equation (S.18) is less than 1, for which case we have no more benefit from applying the two-sided Chebyshev's inequality than the one-sided one.

### A.2. Distribution Unimodality

In this subsection, we consider the situation where the seller has the additional information that the valuation distribution is unimodal, which can be understood as there being one primary market segment. In particular, we consider a class of valuation distributions that are unimodal and share the same mean $\mu$ and standard deviation $\sigma$, denoted by $\mathscr{F}_u \subset \mathscr{F}$.

Proposition OA.3. *Assume $F \in \mathscr{F}_u$. There exists the price heuristic $p_u^*$ that achieves a sharper performance guarantee. Moreover, $p_u^* \geq p^*$ if $\tau \geq \frac{17\sqrt{2}}{3\sqrt{3}} \approx 4.63$, otherwise, $p_u^* \leq p^*$.*

From Proposition OA.3, we can see that if $\tau \geq 4.63$, $p_u^* \geq p^*$, otherwise, $p_u^* \leq p^*$. That is because the class of unimodal distributions is more concentrated around the mean for smaller standard deviation $\sigma$ than the class of general distributions that share the same mean and standard deviation. Hence, when the standard deviation decreases, the price heuristic increases faster towards the mean for the unimodal distributions than for the general class of distributions.

The unimodality may reduce the (two-sided) Chebyshev's bound by a factor of at least 4/9. However, by Proposition OA.3, this occurs when $\tau = \frac{1-\gamma}{\delta}$ is large enough (i.e. $\tau \geq 5.72$), in other words, when $\gamma$ or $\delta$ is small enough. By Corollary 1(a), the performance guarantee given by the base model is decreasing in $\gamma$ and $\delta$, and thus there is little room to improve the performance guarantee with small values of $\gamma$ and $\delta$ by knowing more information of unimodality. More importantly, knowing the unimodality does not technically eliminate the possibility of the extreme case of two-point distributions. In the following proposition, we confirm that the benefit of knowing unimodality is not very significant.

Proposition OA.4. *For all $F \in \mathscr{F}_u$, if $\tau \geq \frac{17\sqrt{2}}{3\sqrt{3}} \approx 4.63$, $\frac{\Pi(p^*;F)}{\Pi(p_u^*;F)} \geq 95\%$. If $2.65 \leq \tau \leq \frac{17\sqrt{2}}{3\sqrt{3}} \approx 4.63$, $\frac{L}{L_u} \geq 87\%$, where $L$ and $L_u$ are the achievable lower bounds for the model without and with unimodality, respectively.*

Whereas a region with a large middle-class base has a unimodal wealth distribution, another region with a high Gini index may have a bipolar wealth distribution. Analogously, the targeted consumer market may also have a valuation distribution that is not unimodal. For example, there could be many vertically separated segments of customers whose valuations are clustered in their segment. Although marketing research, which may be expensive, can be used to investigate whether the customer valuation distribution is unimodal, Proposition OA.4 implies that the additional benefit of having such information is no more than 15%($\approx \frac{1}{87\%} - 1$).

Figure OA.1(c) shows the performance guarantee for unimodal distributions, with $\delta$ ranging from 0 to 2, and $\gamma$ ranging from 0 to 1. In comparison with Figure 1, the gap between knowing the unimodality information and not knowing is not much, as is analytically demonstrated by Proposition OA.4.

## B. Mean-Ranked Clustered Bundling

**Algorithm OA.1** Mean-Ranked Clustered Bundling

**Input**: $c_i$, $\mu_i$, $\sigma_i$, where $i = 1, 2, \ldots, n$

Initialize a binary-matrix $M$ with $2^{n-1}$ rows and $n-1$ columns

**for** $j = 1, 2, \ldots, 2^{n-1}$ **do**

  Find the index of 0 in $j$-th row vector of $M$ and store them in an array $K^j$, update $K^j \leftarrow [0\ K^j\ n]$

  **for** $m = 1, 2, \ldots, length(K^j) - 1$ **do**

    $L_m^j \leftarrow$ the total profit lower bound of a bundle with products $i \in \{K_m^j + 1, K_m^j + 2, \ldots, K_{m+1}^j\}$

  **end for**

  $\pi_j \leftarrow \sum_{m=1}^{length(K^j)-1} L_m^j$

**end for**

Let $j^* = \arg\max_j \pi_j$

**Output**: The total profit lower bound $\pi_{j^*}$, and the corresponding clustered bundling for $j^*$

PROPOSITION OA.5. *Algorithm OA.1 can find the optimal mean-ranked clustered bundling, and its computational complexity is at least $O(2^n)$.*

*Proof of Proposition OA.5.* First, we prove that there are $2^{n-1}$ cases for mean-ranked clustered bundling by induction. Consider $n = 2$, there are two cases: products 1 and 2 are bundled as a cluster, or products 1 and 2 are divided as two clusters. Suppose $n = m$, there are $2^{m-1}$ cases, then we prove it is true for $n = m + 1$. Consider products $1, 2, \ldots, m$, there are $2^{m-1}$ cases. And for any case, the product $m + 1$ can be put into the last cluster or to be a single cluster. Hence, there are $2^{m-1} \times 2 = 2^m$ cases. Second, Algorithm OA.1 considers $2^{n-1}$ cases for mean-ranked clustered bundling. Thus it contains all cases of for mean-ranked clustered bundling. In other words, it outputs the optimal mean-ranked clustered bundling. Lastly, it is easy to see the computational complexity of Algorithm OA.1 is at least $O(2^n)$ since there are $2^{n-1}$ circles in the outer loop. □

### B.1. Fixed Radius Clustering

In this subsection, we consider the heuristic which we called Fixed Radius Clustering (FRC): for a given radius $r$, the difference of means of any two products is no more than $r$ in a cluster. We consider 10000 different radii, and let $r_j = (\mu_n - \mu_1)j/9999$, where $j = 0, 1, \ldots, 9999$. It is easy to see the heuristic reduces to separate sales when $j = 0$ and the pure bundle when $j = 9999$. Thus, this heuristic is more general with the optimized $j^*$, and would have a better performance guarantee than separates sales and the pure bundle. The pseudo-code for the algorithm is given in Algorithm OA.2. There are two interior loops: steps 5-9 and steps 11-13. The first loop is to produce an array $K^j$ to store the bundles, where the $m$-th bundle contains products $i \in \{K_m^j + 1, K_m^j + 2, \ldots, K_{m+1}^j\}$ and $K_m^j$ is the $m$-th element of $K^j$. The second loop applies Theorem 1(b) to calculate $L_m^j$, which is the lower bound on the profit of the $m$-th bundle.

**Algorithm OA.2** Fixed Radius Clustering (FRC)

1: **Input**: $c_i$, $\mu_i$, $\sigma_i$, where $i = 1, 2, \ldots, n$
2: Initialize $r_j = (\mu_n - \mu_1)j/9999$
3: **for** $j = 1, 2, \ldots, 10000$ **do**
4:     Let $m = 1$, $x = \mu_1$
5:     **for** $i = 1, 2, \ldots, n$ **do**
6:         **if** $\mu_i > x + r_j$ **then**
7:             Let $x = \mu_i$, $K_m^j = i - 1$, $m = m + 1$
8:         **end if**
9:     **end for**
10:     Update $K^j \leftarrow [0 \ K^j \ n]$
11:     **for** $m = 1, 2, \ldots, length(K^j) - 1$ **do**
12:         $L_m^j \leftarrow$ the total profit lower bound of a bundle with products $i \in \{K_m^j + 1, K_m^j + 2, \ldots, K_{m+1}^j\}$
13:     **end for**
14:     $\pi_j \leftarrow \sum_{m=1}^{length(K^j)-1} L_m^j$
15: **end for**
16: Let $j^* = \arg\max_j \pi_j$
17: **Output**: The total profit lower bound $\pi_{j^*}$, and the corresponding clustered bundling for $j^*$

Proposition OA.6. *The computational complexity of Algorithm OA.2 is $O(n)$.*

*Proof of Proposition OA.6.* Note that steps 5-9 will run no more than $n$ times. And there can be no more $n$ bundles, i.e., $m \leq n$. Thus steps 11-13 also run no more than $n$ times. Hence, the computational complexity of Algorithm OA.2 is $O(n)$ since it runs the two interior loops "for" no more than $n$ times. □

## B.2. Top-down/Bottom-up Clustering

In this subsection, we consider the Top-down/Bottom-up Clustering algorithm which contains two processes: "Top-down" and "Bottom-up". This is why we call it Top-down/Bottom-up Clustering. If the pure bundle is better than separate sales, we execute the algorithm from the pure bundle (Top-down). We want to split each bundle into two sub-bundles with better total profit lower bounds. If a bundle cannot be split into two sub-bundles with a better performance guarantee, i.e., the bundle has a larger lower bound than the total lower bounds of any two sub-bundles, we keep the bundle unchanged. When all bundles cannot be split into two sub-bundles, the "Top-down" process terminates. If the pure bundle is worse than separate sales, we execute the algorithm from $n$ separate products (Bottom-up). We want to cluster two adjacent bundles into one if it is better to do so. When any two adjacent bundles cannot be bundled with a better performance guarantee, the "Bottom-up" process terminates. The pseudo-code is given in Algorithm OA.3.

For the "Top-down" process, there are two arrays $K$ and $I$, where the former one stores the bundles with the $m$-th bundle containing products $i \in \{K_m+1, K_m+2, \ldots, K_{m+1}\}$ and $K_m$ is the $m$-th element of $K$, and the latter one stores the index of each bundle. Since it starts from the pure bundle, we initialize $K = [0\ n]$ and $I = [1]$ in step 3. The steps 5-9 investigate whether the current bundles can be split if it is better to do so. If a bundle cannot be split into two sub-bundles, let its index be 0, otherwise, let the two new bundles' indices be 1. When all indices are 0, the loop "while" terminates, i.e., we cannot get a better performance guarantee by "Top-down". For the "Bottom-up" process, $K$ stores the bundles. The steps 17-22 investigate whether any two adjacent bundles can be bundled with a better total profit lower bound for the current bundles. If the first two adjacent bundles can be bundled, i.e., $f=1$, we let the two bundles as the first bundle and the third one as the second. If the first two adjacent bundles cannot be bundled, i.e., $f=0$, we let the second bundle as the first bundle and the third one as the second (step 19). After doing this for all bundles, we check whether there exists a case in which two bundles can be bundled (step 23). If any two adjacent bundles cannot be bundled, let $I=0$ and the loop "while" terminates, i.e., we cannot get a better total profit lower bound by "Bottom-up". Figure OA.2 illustrates Algorithm OA.3.

**Algorithm OA.3** Top-down/Bottom-up Clustering (TBC)

1: **Input**: $c_i$, $\mu_i$, $\sigma_i$, where $i = 1, 2, \ldots, n$

2: **if** the pure bundle is better than separate sales (**Top-down**) **then**

3: Initialize two arrays $K = [0\ n]$ and $I = [1]$

4: **while** $I \neq \mathbf{0}$ **do**

5: **for** $m = 1, 2, \ldots, length(K) - 1$ **do**

6: Let $\mathbf{c}^m = [c_{K_m+1}\ c_{K_m+2} \cdots c_{K_{m+1}}]$, $\boldsymbol{\mu}^m = [\mu_{K_m+1}\ \mu_{K_m+2} \cdots \mu_{K_{m+1}}]$, $\boldsymbol{\sigma}^m = [\sigma_{K_m+1}\ \sigma_{K_m+2} \cdots \sigma_{K_{m+1}}]$

7: If $I_m = 1$, compute $f_m = \mathbf{split}(\mathbf{c}^m, \boldsymbol{\mu}^m, \boldsymbol{\sigma}^m)$, otherwise, $f_m = length(\boldsymbol{\mu}^m)$ (Function **split** outputs the optimal size of the first bundle that splits a bundle into two sub-bundles, where the bundles' parameters are $\mathbf{c}^m, \boldsymbol{\mu}^m, \boldsymbol{\sigma}^m$, respectively. The pseudo-code for function **split** is given in Algorithm OA.4.)

8: If $f_m = length(\boldsymbol{\mu}^m)$, let $I_m = 0$, otherwise, let $I_m = 1$; $K'_{2m-1} = K_m$ and $K'_{2m} = K_m + f_m$

9: **end for**

10: Update $K' \leftarrow [K'\ n]$ and $K \leftarrow$ remove the repeated elements in $K'$, $I \leftarrow$ add one element 1 after each element 1 in array $I$

11: **end while**

12: **end if**

13: **if** the pure bundle is worse than separate sales (**Bottom-up**) **then**

14: Initialize an array $K = [1\ 2 \cdots n]$ and $I = 1$

15: **while** $I = 1$ **do**

16: Initialize $m = 1$, $t = 1$, $\mathbf{c}^1 = [c_1\ \ c_2 \cdots c_{K_1}]$, $\boldsymbol{\mu}^1 = [\mu_1\ \ \mu_2 \cdots \mu_{K_1}]$, $\boldsymbol{\sigma}^1 = [\sigma_1\ \ \sigma_2 \cdots \sigma_{K_1}]$, $\mathbf{c}^2 = [c_{K_1+1}\ c_{K_1+2} \cdots c_{K_2}]$, $\boldsymbol{\mu}^2 = [\mu_{K_1+1}\ \mu_{K_1+2} \cdots \mu_{K_2}]$, $\boldsymbol{\sigma}^2 = [\sigma_{K_1+1}\ \sigma_{K_1+2} \cdots \sigma_{K_2}]$

17: **for** $m = 1, 2 \ldots, \mathit{length}(K) - 1$ **do**

18: Compute $f = \mathbf{cluster}(\mathbf{c}^1, \boldsymbol{\mu}^1, \boldsymbol{\sigma}^1, \mathbf{c}^2, \boldsymbol{\mu}^2, \boldsymbol{\sigma}^2)$ (Function **cluster** outputs 1 if bundling two sub-bundles is better, otherwise 0, where the two sub-bundles' parameters are $\mathbf{c}^1, \boldsymbol{\mu}^1, \boldsymbol{\sigma}^1$, and $\mathbf{c}^2, \boldsymbol{\mu}^2, \boldsymbol{\sigma}^2$, respectively. The pseudo-code for function **cluster** is given in Algorithm OA.5.)

19: If $f = 1$, update $\mathbf{c}^1 \leftarrow [\mathbf{c}^1\ \mathbf{c}^2], \boldsymbol{\mu}^1 \leftarrow [\boldsymbol{\mu}^1\ \boldsymbol{\mu}^2], \boldsymbol{\sigma}^1 \leftarrow [\boldsymbol{\sigma}^1\ \boldsymbol{\sigma}^2]$; otherwise, update $\mathbf{c}^1 \leftarrow \mathbf{c}^2, \boldsymbol{\mu}^1 \leftarrow \boldsymbol{\mu}^2, \boldsymbol{\sigma}^1 \leftarrow \boldsymbol{\sigma}^2$, let $K'_t = K_m$ and update $t \leftarrow t + 1$

20: Update $m \leftarrow m + 1$, if $m = \mathit{length}(K)$, **break**

21: Update $\mathbf{c}^2 \leftarrow [c_{K_m+1}\ c_{K_m+2} \cdots c_{K_{m+1}}]$, $\boldsymbol{\mu}^2 \leftarrow [\mu_{K_m+1}\ \mu_{K_m+2} \cdots \mu_{K_{m+1}}]$, $\boldsymbol{\sigma}^2 \leftarrow [\sigma_{K_m+1}\ \sigma_{K_m+2} \cdots \sigma_{K_{m+1}}]$

22: **end for**

23: Update $K' \leftarrow [K'\ n]$, if length($K$)=length($K'$), $I = 0$; otherwise, $I = 1$ and $K \leftarrow K'$

24: **end while**

25: Update $K \leftarrow [0\ K]$

26: **end if**

27: Compute the total profit lower bound of each bundle and store it in $L^*$, let $K^* \leftarrow K$

28: **Output**: The total profit lower bound $L^*$, and $m$-th bundle contains products $i \in \{K^*_m + 1, K^*_m + 2, \ldots, K^*_{m+1}\}$

---

**Algorithm OA.4** function **split**

1: **function** $f = \mathbf{split}(\mathbf{c}, \boldsymbol{\mu}, \boldsymbol{\sigma})$

2: **for** $i = 1, 2, \ldots, \mathit{length}(\boldsymbol{\mu})$ **do**

3: Let $t_1 = \sum_{j=1}^{i} (\boldsymbol{\mu}_j - \mathbf{c}_j) \Big/ \sqrt{\sum_{j=1}^{i} \boldsymbol{\sigma}_j^2}$ and $t_2 = \sum_{j=i+1}^{\mathit{length}(\boldsymbol{\mu})} (\boldsymbol{\mu}_j - \mathbf{c}_j) \Big/ \sqrt{\sum_{j=i+1}^{\mathit{length}(\boldsymbol{\mu})} \boldsymbol{\sigma}_j^2}$ ($t_2 = 0$ for $i = \mathit{length}(\boldsymbol{\mu})$)

4: Compute $k'_1 = \sqrt[3]{t_1 + \sqrt{t_1^2 + 1}} + \sqrt[3]{t_1 - \sqrt{t_1^2 + 1}}$ and $k'_2 = \sqrt[3]{t_2 + \sqrt{t_2^2 + 1}} + \sqrt[3]{t_2 - \sqrt{t_2^2 + 1}}$

5: Compute $L_1 = \sum_{j=1}^{i} (\boldsymbol{\mu}_j - \mathbf{c}_j)(1 - \frac{3}{2t_1} k_1'^3)$ and $L_2 = \sum_{j=i+1}^{\mathit{length}(\boldsymbol{\mu})} (\boldsymbol{\mu}_j - \mathbf{c}_j)(1 - \frac{3}{2t_2} k_2'^3)$ ($L_2 = 0$ for $i = \mathit{length}(\boldsymbol{\mu})$)

6: $\pi_i = L_1 + L_2$

7: **end for**

8: $f$=index of the maximum in array $\pi = [\pi_1\ \pi_2 \cdots \pi_{\mathit{length}(\boldsymbol{\mu})}]$

9: **end**

---

**Algorithm OA.5** function **cluster**

1: **function** $f = \mathbf{cluster}(\mathbf{c}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{c}', \boldsymbol{\mu}', \boldsymbol{\sigma}')$

2: Let $t_1 = \sum_{j=1}^{\mathit{length}(\boldsymbol{\mu})} (\boldsymbol{\mu}_j - \mathbf{c}_j) \Big/ \sqrt{\sum_{j=1}^{\mathit{length}(\boldsymbol{\mu})} \boldsymbol{\sigma}_j^2}$ and $t_2 = \sum_{j=1}^{\mathit{length}(\boldsymbol{\mu}')} (\boldsymbol{\mu}'_j - \mathbf{c}'_j) \Big/ \sqrt{\sum_{j=1}^{\mathit{length}(\boldsymbol{\mu}')} \boldsymbol{\sigma}_j'^2}$

3: Compute $k'_1 = \sqrt[3]{t_1 + \sqrt{t_1^2 + 1}} + \sqrt[3]{t_1 - \sqrt{t_1^2 + 1}}$ and $k'_2 = \sqrt[3]{t_2 + \sqrt{t_2^2 + 1}} + \sqrt[3]{t_2 - \sqrt{t_2^2 + 1}}$

4: Compute $L_1 = \sum_{j=1}^{length(\boldsymbol{\mu})} (\boldsymbol{\mu}_j - \mathbf{c}_j)(1 - \frac{3}{2t_1} k_1'^3)$ and $L_2 = \sum_{j=1}^{length(\boldsymbol{\mu}')} (\boldsymbol{\mu}'_j - \mathbf{c}'_j)(1 - \frac{3}{2t_2} k_2'^3)$

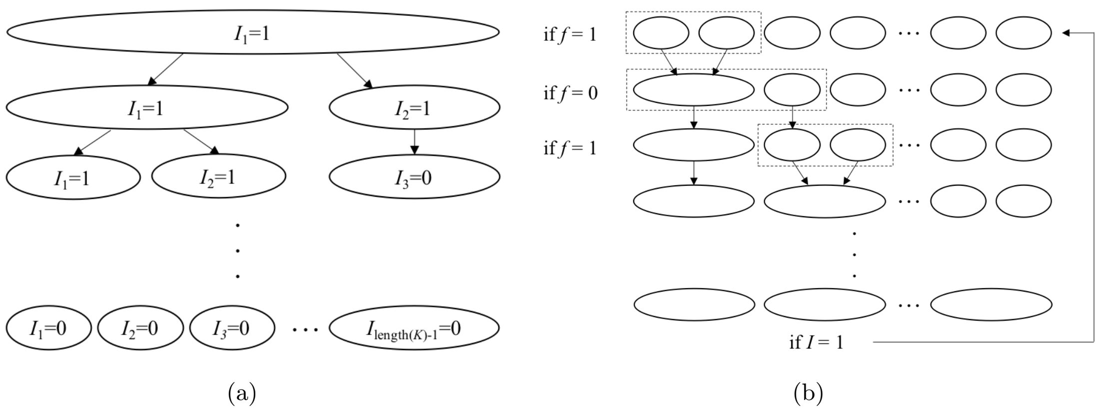
5: Let $\mathbf{c}'' = [\mathbf{c}\ \mathbf{c}'], \boldsymbol{\mu}'' = [\boldsymbol{\mu}\ \boldsymbol{\mu}'], \boldsymbol{\sigma}'' = [\boldsymbol{\sigma}\ \boldsymbol{\sigma}']$ and $t = \sum_{j=1}^{length(\boldsymbol{\mu}'')} (\boldsymbol{\mu}''_j - \mathbf{c}''_j) \Big/ \sqrt{\sum_{j=1}^{length(\boldsymbol{\mu}'')} \boldsymbol{\sigma}''^2_j}$

6: Compute $k = \sqrt[3]{t + \sqrt{t^2+1}} + \sqrt[3]{t - \sqrt{t^2+1}}$ and $L = \sum_{j=1}^{length(\boldsymbol{\mu}'')} (\boldsymbol{\mu}''_j - \mathbf{c}''_j)(1 - \frac{3}{2t} k^3)$

7: If $L_1 + L_2 < L$, $f = 1$; otherwise, $f = 0$

8: **end**

**Figure OA.2 Illustration of Algorithm OA.3**




*Note.* (a) "Top-down" process. (b) "Bottom-up" process.

PROPOSITION OA.7. *The heuristic out of Algorithm OA.3 has a better performance guarantee than separates sales and the pure bundle, and its computational complexity is $O(n^2)$.*

*Proof of Proposition OA.7.* First, consider the pure bundle is better than separate sales. By Algorithm OA.3, it executes the "Top-down" process. Note that the "Top-down" process will be executed if and only if there exists a bundle that can be split into two sub-bundles with a better performance guarantee. Then the output of the "Top-down" process is better than the pure bundle. On the one hand, for each interior loop "for", the algorithm runs $length(K) - 1$ times and the function **split** contains a loop "for" with $length(\mathbf{c}^m) = K_{m+1} - K_m$ circles due to Algorithm OA.4. Thus, step 7 is executed $\sum_{m=1}^{length(K)-1} K_{m+1} - K_m = n$ times for each loop "for". On the other hand, for the outer loop "while", at least one bundle is split. There are $n$ products, and then it runs at most $n - 1$ times. Consequently, its computational complexity is $O(n^2)$.

Second, consider the pure bundle is worse than separate sales. By Algorithm OA.3, it executes the "Bottom-up" process. Note that the "Bottom-up" process will be executed if and only if there exists two adjacent bundles that can be bundled with a better performance guarantee. Then the output of the "Bottom-up" process is better than separate sales. On the one hand, for each interior loop "for", the algorithm runs $length(K)-1\le n$ times and the computational complexity of function **cluster** is $O(1)$ due to Algorithm OA.5. On the other hand, for the outer loop "while", at least one product is bundled. There are $n$ products, and then it runs at most $n-1$ times. Consequently, its computational complexity is $O(n^2)$. □

## C. Proofs

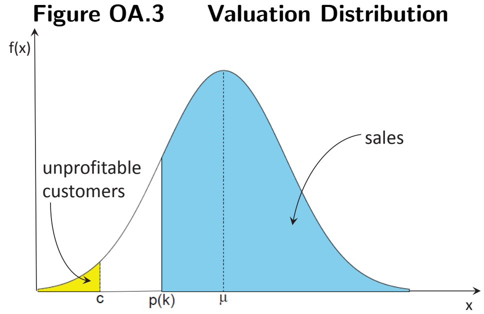
To show Theorem 1, we first prove a version of a one-sided Chebyshev's inequality (on the left tail), also known as Cantelli's inequality, which is often stated in a different form (on the right tail, for example, see Gallego et al. 2007).

Lemma OA.1 (Cantelli's Inequality). *For a random variable $X$ that has mean $\mu$ and standard deviation $\sigma$ and all $a\ge 0$,*

$$\mathsf{P}(X<\mu-a)\le \mathsf{P}(X\le\mu-a)\le\frac{\sigma^2}{\sigma^2+a^2}.$$

*Equivalently,* $\mathsf{P}(V<\mu-k\sigma)\le\mathsf{P}(V\le\mu-k\sigma)\le\frac{1}{1+k^2}$ *for any* $k\ge 0$.

**Figure OA.3 Valuation Distribution**




*Note.* $p(k)=\mu-k\sigma$, $k\ge 0$

*Proof of Lemma OA.1.* If $a=0$, the inequalities trivially hold. Now consider $a>0$. Let $Y=X-\mu$, then $\mathsf{E}[Y]=\mathsf{E}[-Y]=0$ and $\mathrm{var}[Y]=\mathrm{var}[-Y]=\sigma^2$. Let $t=\sigma^2/a\ge 0$. We have

$$\begin{aligned}\mathsf{P}(Y<-a)\le\mathsf{P}(Y\le -a)=\mathsf{P}(-Y\ge a)=\mathsf{P}(-Y+t\ge a+t)=\mathsf{P}\left(\frac{-Y+t}{a+t}\ge 1\right)\\ \le\mathsf{P}\left(\left(\frac{-Y+t}{a+t}\right)^2\ge 1\right)\le\mathsf{E}\left[\left(\frac{-Y+t}{a+t}\right)^2\right]=\frac{\sigma^2+t^2}{(a+t)^2}=\frac{\sigma^2}{\sigma^2+a^2},\end{aligned}$$

where the last inequality follows from Markov's inequality. □

*A sketch of the proof of Theorem 1.* We first give a sketch of the proof here. The upper bound $U$ in Theorem 1(a) is no less than $\mu-c$, i.e., $U\ge\mu-c$, where $\mu-c$ is the expected profit earned from *all* customers under perfect price discrimination (i.e., everyone pays at their own valuation).

Theorem 1(a) implies that the seller may be able to earn *more* than $\mu - c$ by serving only those customers whose valuations are no less than the marginal cost $c$. Hence, the upper bound of the optimal profit should be no more than $\mu - c$ *plus* the cost of serving unprofitable customers. The unprofitable sales volume is $\mathsf{P}(V < c)$ (to the left of $x = c$ in Figure OA.3). By Lemma OA.1, we can obtain an upper bound of $\mathsf{P}(V < c)$ as $\frac{1}{1+\tau^2}$; thus an upper bound of the expected profit is $\mu - c + c\mathsf{P}(V < c) \le \mu - c + \frac{c}{1+\tau^2} = \mu - \frac{\tau^2}{1+\tau^2}c$. In the extreme case $\sigma = 0$, when the valuation distribution is deterministic, the seller can earn the optimal profit $\mu - c \ge 0$ by charging price $p = \mu$. By Theorem 1(a), when $\sigma = 0$, $\tau = \infty$, then $U = \mu - c$; thus we can see the upper bound $U$ is achievable.

Next we apply Lemma OA.1 again to find a lower bound on the seller's profits. Theorem 1(b) suggests that the profit earned by the seller has a lower bound $L$, which comes from charging a price $p^* = \mu - k^*\sigma$, below the mean valuation $\mu$. When the seller sets a price in the form of $p(k) = \mu - k\sigma$ that is less than $\mu$, the sales volumes correspond to the right of $x = p(k)$ in Figure OA.3. By Lemma OA.1, the complement of that area has an upper bound $\frac{1}{1+k^2}$; hence, that area has a lower bound $1 - \frac{1}{1+k^2}$. Then we obtain a lower bound of the profit when charging $p(k) = \mu - k\sigma$:

$$(p(k) - c)\mathsf{P}(V \ge p(k)) \ge (\mu - k\sigma - c)(1 - \frac{1}{1+k^2}) = (\mu - c)(1 - \frac{k}{\tau})(1 - \frac{1}{1+k^2}).$$

The lower bound is maximized by $k^* = \sqrt[3]{\tau + \sqrt{\tau^2+1}} + \sqrt[3]{\tau - \sqrt{\tau^2+1}}$. $\square$

*Proof of Theorem 1(a).* Define a random variable $X$ such that its density function $g(x) = f(x - c)$. Hence the distribution of $X$ is a horizontal shift of that of $V$ towards the left by $c$ units. Thus, the mean and standard deviation of $X$ are $\mu - c$ and $\sigma$, respectively.

Since any $p \ge c$ is better off than $p < c$ which yields negative profit, we only need to consider $p \ge c$ for the profit maximization problem $\max_p \pi(p; F)$. For any $p \ge c$,

$$\begin{aligned}\pi(p;F) = (p-c)\mathsf{P}(V \ge p) &= (p-c)\mathsf{P}(X \ge p-c) \\ &\le \mathsf{E}[X|X \ge p-c] \le \mathsf{E}[X|X \ge 0] = \mu - c - \mathsf{E}[X|X<0], \quad \text{(OA.1)}\end{aligned}$$

where the last equality is due to that $\mu - c = \mathsf{E}[X|X<0] + \mathsf{E}[X|X \ge 0]$.

Since $g(x+c) = f(x)$, then

$$-\mathsf{E}[X|X<0] = \mathsf{E}[c - V|V<c] \le \mathsf{E}[c|V<c] \le c\,\mathsf{P}(V<c) \le \frac{c}{1+((\mu-c)/\sigma)^2} = \frac{c}{1+\tau^2}, \quad \text{(OA.2)}$$

where the first inequality is due to valuation $V \ge 0$ and the last inequality is due to Lemma OA.1 with $k = (\mu - c)/\sigma$ (where $k \ge 0$ due to Assumption (P)). Combining (OA.1) and (OA.2), we have the desired upper bound.

When $\sigma = 0$, we have $\tau = \infty$ and $U = \mu - c$ which is the exact optimal profit the seller earns when facing the deterministic valuation $\mu$. $\square$

*Proof of Theorem 1(b).* Write a price heuristic in the form of $p = \mu - k\sigma \geq c$ for some $k \geq 0$. In other words, $0 \leq k \leq \frac{\mu - c}{\sigma}$. Then, for all $F \in \mathscr{F}$,

$$\max_{p>c} \pi(p;F) \geq \pi(p = \mu - k\sigma; F) = (\mu - k\sigma - c)\mathsf{P}(V \geq \mu - k\sigma) \geq (\mu - c)(1 - \frac{k\sigma}{\mu - c})(1 - \frac{1}{1+k^2}), \tag{OA.3}$$

where the inequality is due to $\mathsf{P}(V < \mu - k\sigma) \leq \frac{1}{1+k^2}$ for $k \geq 0$ from Lemma OA.1. For notation convenience, denote $\varphi(k) \equiv (1 - \frac{k\sigma}{\mu - c})(1 - \frac{1}{1+k^2})$.

Now we want to maximize the lower bound, $(\mu - c)\varphi(k)$, in (OA.3) to obtain the optimal robust price. It can be verified that $\varphi'(k) = \frac{-k^4 - 3k^2 + 2k\tau}{\tau(1+k^2)^2} = -\frac{k}{\tau(1+k^2)^2}(k^3 + 3k - 2\tau)$. The non-zero stationary points of $\varphi(k)$ are characterized by $\varphi'(k) = 0$ which is equivalent to

$$k^3 + 3k = 2\tau. \tag{OA.4}$$

(The function $\varphi(k)$ is minimized at $k = 0$ or $k = \frac{\mu - c}{\sigma}$.) By Cardano's solution for a cubic function, the unique real root of (OA.4) is $k^* = \sqrt[3]{\tau + \sqrt{\tau^2 + 1}} + \sqrt[3]{\tau - \sqrt{\tau^2 + 1}}$. It is easy to see that indeed, $k^* \geq 0$ because $k^* = \frac{2\tau}{(k^*)^2 + 3} \geq 0$. Moreover, since $-k^3 - 3k$ is decreasing in $k$, thus for $k \leq k^*$, $\varphi'(k) \geq 0$ and when $k \geq k^*$, $\varphi'(k) \leq 0$. This demonstrates the function $\varphi(k)$ is maximized at $k = k^*$ with the optimal value

$$\varphi(k^*) = \left(1 - \frac{k^*}{\tau}\right)\left(1 - \frac{1}{1 + (k^*)^2}\right) = \frac{(k^*)^2}{(k^*)^2 + 3} = \frac{(k^*)^3}{2\tau} = 1 - \frac{3}{2\tau}k^*, \tag{OA.5}$$

where the first equality is by definition of $\varphi(k)$ and the rest of identities are all due to (OA.4). By (OA.4) and $k^* \geq 0$, we have $k^{*3} + 3k^* = 2\tau \geq 3k^*$, i.e., $k^* \leq 2\tau/3 \leq \tau = (\mu - c)/\sigma$. Thus, $p^* = \mu - k^*\sigma \geq c$. In view of (OA.3), we immediately have the desired lower bound.

It is easily verifiable that any distribution within the stipulated class of two-point distributions for any $\epsilon$ indeed has mean $\mu$ and standard deviation $\sigma$. With the optimal robust price $p^* = \mu - k^*\sigma$, for a given two point distribution with $k_\epsilon = k^* + \epsilon$, only those customers who have the valuation at the high end of the two points would choose to purchase the product. Thus the resulting profit is $(1 - \frac{1}{1+k_\epsilon^2})(\mu - k^*\sigma - c) = (\mu - c)(1 - \frac{k^*}{\tau})(1 - \frac{1}{1+k_\epsilon^2})$, which is asymptotically approximated to the lower bound $(\mu - c)\varphi(k^*) = (\mu - c)(1 - \frac{k^*}{\tau})(1 - \frac{1}{1+(k^*)^2})$, in view of (OA.3), as $\epsilon \searrow 0$. $\square$

*Proof of Theorem 2.* For all $F \in \mathscr{F}$,

$$\frac{\pi(p^* = \mu - k^*\sigma; F)}{\max_{p>c} \pi(p;F)} \geq \frac{\mu(1-\gamma)(1 - \frac{3}{2\tau}k^*)}{\max_{p>c} \pi(p;F)} \geq \frac{\mu(1-\gamma)(1 - \frac{3}{2\tau}k^*)}{\mu - \frac{\tau^2}{1+\tau^2}c} = \frac{1 - \frac{3}{2\tau}k^*}{\frac{1}{1-\gamma} - \frac{\tau^2\gamma}{(1+\tau^2)(1-\gamma)}} = \frac{1 - \frac{3}{2\tau}k^*}{1 + \frac{\gamma}{(1-\gamma)(1+\tau^2)}},$$

where the first inequality is due to Theorem 1(b) and the second inequality is due to Theorem 1(a). $\square$

*Proof of Corollary 1.* (a) It is trivial to see that $\rho$ depends on the system primitives $c,\mu,\sigma$ only through the ratios $\gamma$ and $\delta$. By (OA.4), $(k^*)^3+3k^*=2\tau=2\frac{1-\gamma}{\delta}$. Since the right-hand-side of this equation is strictly decreasing in $\gamma$ and $\delta$, its unique real root $k^*$ must also be strictly decreasing in $\gamma$ and $\delta$. Moreover, $1-\frac{3}{2\tau}k^*=\frac{(k^*)^2}{(k^*)^2+3}$ is strictly increasing in $k^*$ and $1+\frac{\gamma}{(1-\gamma)(1+\tau^2)}$ is strictly increasing in $\gamma$ and $\delta$. As a result, $\rho=\frac{1-\frac{3}{2\tau}k^*}{1+\frac{\gamma}{(1-\gamma)(1+\tau^2)}}$ is strictly decreasing in $\gamma$ and $\delta$.

(b) By (OA.4), $(k^*)^3+3k^*=2\tau=\frac{2(\mu-c)}{\sigma}$. Since the right-hand-side of this equation is strictly decreasing in $c$, $k^*$ is strictly decreasing in $c$ and $p^*(c,\mu,\sigma)=\mu-k^*\sigma$ is strictly increasing in $c$.

By (OA.4), $(k^*)^3+3k^*=2\tau=\frac{2(\mu-c)}{\sigma}$. Since the right-hand-side of this equation is strictly decreasing in $\sigma$, $k^*$ is strictly decreasing in $\sigma$. Again by (OA.4), $k^*\sigma=\frac{2(\mu-c)}{(k^*)^2+3}$, thus $k^*\sigma=\frac{2(\mu-c)}{(k^*)^2+3}$ is strictly increasing in $\sigma$ and $p^*=\mu-k^*\sigma$ is strictly decreasing in $\sigma$.

By (OA.4), $(k^*)^3+3k^*=2\tau=\frac{2(\mu-c)}{\sigma}$. Taking derivatives with respect to $\mu$ on both sides of this equation, we obtain $(3(k^*)^2+3)\frac{\partial k^*}{\partial\mu}=\frac{2}{\sigma}$, then $\frac{\partial k^*}{\partial\mu}\sigma=\frac{2}{3(k^*)^2+3}$. Since $p^*=\mu-k^*\sigma$, $\frac{\partial p^*}{\partial\mu}=1-\frac{\partial k^*}{\partial\mu}\sigma=1-\frac{2}{3(k^*)^2+3}=\frac{3(k^*)^2+1}{3(k^*)^2+3}>0$, which shows $p^*=\mu-k^*\sigma$ is strictly increasing in $\mu$. □

*Proof of Proposition 1.* (a) Because both the mean and standard deviation of the exponential distribution are $\frac{1}{\lambda}$, $\delta=1$. Thus the optimal robust price is $p^*=\frac{1}{\lambda}(1-k_e(\gamma))$, under which the profit is $\Pi(p^*;F\sim\exp(\lambda))=(\frac{1}{\lambda}(1-k_e(\gamma))-c)e^{-1+k_e(\gamma)}$. Solve the first order condition $\Pi'=(1-\lambda(p-c))e^{-\lambda p}=0$, we have $p^*=\frac{1}{\lambda}+c$ and $\Pi''|_{p=p^*}=-\lambda e^{-\lambda p}-\lambda(1-\lambda(p-c))e^{-\lambda p}=-\lambda e^{-\lambda p^*}<0$. Thus the optimal profit is $\frac{1}{\lambda}e^{-1-c\lambda}$, which is achieved by $p^*=\frac{1}{\lambda}+c$. With $\gamma=\frac{c}{\mu}=c\lambda$, the performance should be $\frac{(\frac{1}{\lambda}(1-k_e(\gamma))-c)e^{-1+k_e(\gamma)}}{\frac{1}{\lambda}e^{-1-c\lambda}}=(1-(k_e(\gamma)+\gamma))e^{(k_e(\gamma)+\gamma)}$. When $\gamma=0$, we have $k_e(\gamma)\approx 0.5961$, and then the performance is $(1-0.5961)e^{0.5961}\approx 0.7331$.

(b) The mean of the uniform distribution is $\frac{a}{2}$, and the standard deviation is $\frac{a}{2\sqrt{3}}$, so the CV $\delta$ is $\frac{1}{\sqrt{3}}$. Thus the optimal robust price is $\frac{a}{2}-k_u(\gamma)\frac{a}{2\sqrt{3}}$, under which the profit is $(\frac{a}{2}-k_u(\gamma)\frac{a}{2\sqrt{3}}-c)\frac{\frac{a}{2}+k_u(\gamma)\frac{a}{2\sqrt{3}}}{a}$. With the uniform valuation, the profit function is $(p-c)\frac{a-p}{a}$, then the optimal price is $\frac{a+c}{2}$, and the optimal profit is $\frac{(\frac{a-c}{2})^2}{a}$. With $\gamma=\frac{c}{\mu}=\frac{2c}{a}$, the performance should be $\frac{(1-\frac{k_u(\gamma)}{\sqrt{3}}-\gamma)(1+\frac{k_u(\gamma)}{\sqrt{3}})}{(1-\frac{\gamma}{2})^2}$. When $\gamma=0$, we have $k_u(\gamma)\approx 0.9064$, and then the performance is $(1-\frac{0.9064}{\sqrt{3}})(1+\frac{0.9064}{\sqrt{3}})\approx 0.7261$. □

*Proof of Proposition 4.* We will apply the optimal robust price in Theorem 1(b). It is sufficient to compare the lower bounds on profitability of selling separately and selling a bundle, because the lower bounds obtained in Theorem 1(b) is asymptotically tight.

Denote $\tau_i=\frac{\mu_i-c_i}{\sigma_i}=\frac{(1-\gamma)\mu_i}{\sigma_i}=\frac{1-\gamma}{\delta_i}$ for all $i$, and $\tau_{\bar{n}}=\frac{(1-\gamma)\sum_i\mu_i}{\sqrt{\sum_i\sigma_i^2}}$. Moreover, let $k_i$ and $k_{\bar{n}}$ be the unique real root to

$$k_i^3+3k_i=2\tau_i=\frac{2}{\sigma_i}\mu_i(1-\gamma),\quad k_{\bar{n}}^3+3k_{\bar{n}}=2\tau_{\bar{n}}=\frac{2\sum_i\mu_i}{\sqrt{\sum_i\sigma_i^2}}(1-\gamma),$$

respectively. By (OA.3) and (OA.5), the profit lower bound for selling $n$ products separately, $L_{S_n}$, is:

$$L_{S_n} = (1-\gamma)\sum_i \mu_i \frac{k_i^2}{k_i^2+3} = (1-\gamma)\sum_i \mu_i \frac{1}{2\tau_i} k_i^3 = \sum_i \frac{1}{2}\sigma_i k_i^3.$$

The profit lower bound for selling a bundle of $n$ products, $L_{B_n}$, is:

$$L_{B_n} = (1-\gamma)\sum_i \mu_i \frac{k_{\bar{n}}^2}{k_{\bar{n}}^2+3} = (1-\gamma)\sum_i \mu_i \frac{1}{2\tau_{\bar{n}}} k_{\bar{n}}^3 = \frac{\sqrt{\sum_i \sigma_i^2}}{2} k_{\bar{n}}^3.$$

Since $\frac{\sqrt{\sum_i \sigma_i^2}}{\sum_i \mu_i} \le \min_i \delta_i$, i.e., $\tau_{\bar{n}} \ge \tau_i$, and then $k_{\bar{n}} \ge k_i$ for all $i$. We have: for all $i$,

$$\frac{k_i^3}{k_{\bar{n}}^3} \le \frac{k_i^3 + 3k_i}{k_{\bar{n}}^3 + 3k_{\bar{n}}} = \frac{\tau_i}{\tau_{\bar{n}}} = \frac{\sqrt{\sum_i \sigma_i^2}}{n\sigma_i}. \tag{OA.6}$$

Then

$$\frac{L_{S_n}}{L_{B_n}} = \sum_i \frac{\sigma_i}{\sqrt{\sum_j \sigma_j^2}} \frac{k_i^3}{k_{\bar{n}}^3} \le \sum_i \frac{1}{n} = 1,$$

where the inequality is due to (OA.6). Now we can conclude that under the condition $\frac{\sqrt{\sum_i \sigma_i^2}}{\sum_i \mu_i} \le \min_i \delta_i$, selling a bundle is guaranteed to generate higher profits than selling separately. □

*Proof of Corollaries 2 and 3.* It is easy to see that the conditions in Corollary 2 are special cases of those in Corollary 3. And for Corollaries 2(b), if $\frac{\max_i \sigma_i}{\min_i \sigma_i} \le \sqrt{n}$, then $\sqrt{n}\min_i \sigma_i \ge \max_i \sigma_i \ge \frac{1}{\sqrt{n}}\sqrt{\sum_i \sigma_i^2}$, i.e., $\sqrt{\sum_i \sigma_i^2} \le n \min_i \sigma_i$. Similarly, for Corollaries 2(c), $\frac{\max_i \mu_i}{\min_i \mu_i} \le \sqrt{n}$ is one sufficient condition for $\sqrt{n}\max_i \mu_i \le \sum_i \mu_i$. To prove Corollary 3, it is sufficient to verify that the condition in Proposition 4 is satisfied.

(a) We have

$$\frac{\sqrt{\sum_i \sigma_i^2}}{\sum_i \mu_i} = \frac{\sqrt{\sum_i \delta_i^2 \mu_i^2}}{\sum_i \mu_i} \le \max_i \delta_i \frac{\sqrt{\sum_i \mu_i^2}}{\sum_i \mu_i} \le \min_i \delta_i,$$

where the last inequality is due to $\frac{\max_i \delta_i}{\min_i \delta_i} \le \frac{\sum_i \mu_i}{\sqrt{\sum_i \mu_i^2}}$.

(b) We have

$$\frac{\sqrt{\sum_i \sigma_i^2}}{\sum_i \mu_i} \le \frac{\sqrt{\sum_i \sigma_i^2}}{n \min_i \mu_i} \le \frac{\min_i \sigma_i}{\max_i \mu_i} \le \frac{\sigma_l}{\mu_l} = \min_i \delta_i,$$

where the second inequality is due to $\frac{\max_i \mu_i}{\min_i \mu_i} \le \frac{n \min_i \sigma_i}{\sqrt{\sum_i \sigma_i^2}}$ and $l = \arg\min_i \delta_i$.

(c) We have

$$\frac{\sqrt{\sum_i \sigma_i^2}}{\sum_i \mu_i} \le \frac{\sqrt{n}\max_i \sigma_i}{\sum_i \mu_i} \le \frac{\min_i \sigma_i}{\max_i \mu_i} \le \min_i \delta_i,$$

where the second inequality is due to $\frac{\max_i \sigma_i}{\min_i \sigma_i} \le \frac{\sum_i \mu_i}{\sqrt{n}\max_i \mu_i}$. □

*Proof of Proposition 5.* The CV of the bundle is $\frac{\delta}{\sqrt{n}}$. Define $\tau_b = \frac{1-\gamma}{\frac{\delta}{\sqrt{n}}} = \sqrt{n}\tau$. Let $n^*$ be the unique root of

$$\frac{(\mu-c)(1-\frac{3}{2\sqrt{n}\tau}k_b^*)}{\mu-\frac{\tau^2}{1+\tau^2}c} = 1-\epsilon, \tag{OA.7}$$

where $k_b^*$ is the unique root of $(k_b^*)^3+3k_b^* = 2\sqrt{n}\tau$. It is easy to see that $n^*$ is decreasing in $\epsilon$, because the left hand side of (OA.7) is increasing in $n$.

By Theorem 1 (b), the lower bound of the robust bundle price $p_b^* = p^*(nc, n\mu, \sqrt{n}\sigma) = n\mu - k_b^*\sqrt{n}\sigma$ is $n(\mu-c)(1-\frac{3}{2\sqrt{n}\tau}k_b^*)$, and the upper bound is $n\big(\mu-\frac{\tau^2}{1+\tau^2}c\big)$. The performance guarantee of the robust bundle price $p_b^*$ is $\frac{(\mu-c)(1-\frac{3}{2\sqrt{n}\tau}k_b^*)}{\mu-\frac{\tau^2}{1+\tau^2}c}$, which is increasing in $n$. Thus if $n>n^*$, the performance guarantee should be more than $1-\epsilon$.

If $\gamma=0$, (OA.7) becomes $1-\frac{3}{2\sqrt{n}\tau}k_b^* = \frac{(k_b^*)^2}{(k_b^*)^2+3} = 1-\epsilon$ where the first equality is due to (OA.5). Then $k_b^* = \sqrt{\frac{3}{\epsilon}-3}$. In view of (OA.4), we have $(k_b^*)^3+3k_b^* = \frac{3}{\epsilon}\sqrt{\frac{3}{\epsilon}-3} = 2\tau_b = 2\sqrt{n}\tau = \frac{2\sqrt{n}}{\delta}$. Hence, $n^* = (\frac{3}{\epsilon})^2(\frac{3}{\epsilon}-3)\frac{\delta^2}{4}$. □

*Proof of Proposition 6.* Denote the correlation coefficient between valuations for products $i$ and $j$ by $\rho_{ij}$, which is non-positive. Then the standard deviation of the bundle of $n$ products under non-positive correlations is

$$\sqrt{\sum_{i=1}^n \sigma_i^2 + 2\sum_{i=1}^n\sum_{j=i+1}^n \sigma_i\sigma_j\rho_{ij}} \le \sqrt{\sum_{i=1}^n \sigma_i^2}.$$

The latter is the standard deviation of the bundle of $n$ independent products. Then $\tau'' = \frac{\sum_{i=1}^n(\mu_i-c_i)}{\sqrt{\sum_{i=1}^n \sigma_i^2+2\sum_{i=1}^n\sum_{j=i+1}^n\sigma_i\sigma_j\rho_{ij}}} \ge \frac{\sum_{i=1}^n(\mu_i-c_i)}{\sqrt{\sum_{i=1}^n\sigma_i^2}} = \tau'$. Let $k'$ and $k''$ be the unique root of $(k')^3+3k' = 2\tau'$ and $(k'')^3+3k'' = 2\tau''$, respectively. Because $k^3+3k$ is increasing in $k$, then $k''\ge k'$. Hence the guaranteed performance lower bound of the bundle under non-positive correlations is even higher than that under independent valuations, i.e.,

$$\sum_{i=1}^n(\mu_i-c_i)\frac{(k'')^2}{(k'')^2+3} \ge \sum_{i=1}^n(\mu_i-c_i)\frac{(k')^2}{(k')^2+3}.$$

Hence under conditions of Proposition 4 or its Corollaries 2 and 3 or Proposition 5, the bundle is guaranteed to generate even higher profits than selling separately whose performance is unaffected by valuation correlations. □

*Proof of Proposition 7.* Since the absolute value of correlation coefficient is less than 1, i.e., $-1\le\rho_{ij}\le 1$,

$$\tau'' = \frac{\sum_{i=1}^n(\mu_i-c_i)}{\sqrt{\sum_{i=1}^n\sigma_i^2+2\sum_{i=1}^n\sum_{j=i+1}^n\sigma_i\sigma_j\rho_{ij}}}$$

$$\geq \frac{\sum_{i=1}^n(\mu_i - c_i)}{\sqrt{\sum_{i=1}^n \sigma_i^2 + 2\sum_{i=1}^n \sum_{j=i+1}^n \sigma_i\sigma_j}} = \frac{\sum_{i=1}^n(\mu_i - c_i)}{\sum_{i=1}^n \sigma_i} = \tau = \frac{1-\gamma}{\delta} = \frac{\mu_i - c_i}{\sigma_i},$$

where the inequality is due to $\rho_{ij} \leq 1$ and the last three identities are due to Assumption (S) and $\frac{\sigma_1}{\mu_1} = \frac{\sigma_2}{\mu_2} = \cdots = \frac{\sigma_n}{\mu_n} = \delta$. Following the same logic as the proof of Proposition 6, we have the desired conclusion. □

*Proof of Proposition 8.* Let $p_i$ and $p_b$ be the prices for product $i$ and the bundle, respectively. Moreover, $p_b < \sum_{i=1}^n p_i$, otherwise, no customer will buy the bundle. Then if the customer buy product $i$ or the pure bundle, we have

$$V_i \geq p_i,$$

or

$$\sum_{i=1}^n V_i \geq p_b.$$

In other words, the fraction of customers who will buy nothing is

$$\mathsf{P}\left(\bigcap_{i=1}^n (V_i < p_i) \bigcap \sum_{i=1}^n V_i < p_b\right).$$

If the customer buy the pure bundle, we have

$$\sum_{i=1}^n V_i - p_b \geq V_i - p_i \quad \forall i, \quad \sum_{i=1}^n V_i \geq p_b,$$

which is equivalent to

$$\sum_{j\neq i}^n V_j \geq p_b - p_i \quad \forall i, \quad \sum_{i=1}^n V_i \geq p_b.$$

Therefore, we can write the profit function as follows:

$$\begin{aligned}
&\pi(p_i, p_b; F_i \in \mathscr{F}_i)\\
\geq &\min_i\{p_i - c_i\}\left[1 - \mathsf{P}\left(\bigcap_{i=1}^n (V_i < p_i) \bigcap \sum_{i=1}^n V_i < p_b\right)\right]\\
&+ \left(p_b - \sum_{i=1}^n c_i - \min_i\{p_i - c_i\}\right)\mathsf{P}\left(\bigcap_{i=1}^n \left(\sum_{j\neq i} V_j \geq p_b - p_i\right) \bigcap \sum_{i=1}^n V_i \geq p_b\right)\\
\geq &\min_i\{p_i - c_i\}\left[1 - \mathsf{P}\left(\bigcap_{i=1}^n (V_i < p_i)\right)\right]\\
&+ \left(p_b - \sum_{i=1}^n c_i - \min_i\{p_i - c_i\}\right)\mathsf{P}\left(\bigcap_{i=1}^n \left(\sum_{j\neq i} V_j \geq p_b - p_i\right) \bigcap \sum_{i=1}^n V_i \geq p_b\right),
\end{aligned} \tag{OA.8}$$

where the first inequality is due to $p_j - c_j \geq \min\{p_i - c_i\}$, and the second inequality is due to

$$1 - \mathsf{P}\left(\bigcap_{i=1}^n (V_i < p_i) \bigcap \sum_{i=1}^n V_i \geq p_b\right) \geq 1 - \mathsf{P}\left(\bigcap_{i=1}^n (V_i < p_i)\right).$$

Next, we will give a lower bound of the last term of (OA.8). We denote $p_i = \mu_i - k_i\sigma_i$ and $p_b = \sum_{i=1}^{n}\mu_i - k_b\sqrt{\sum_{i=1}^{n}\sigma_i^2}$, where $k_i > 0$ and $k_b > 0$. Let $p_b - p_i = \sum_{j\neq i}\mu_j - \bar{k}_i\sqrt{\sum_{j\neq i}\sigma_j^2}$, where $\bar{k}_i > 0$. Thus, $k_i\sigma_i + \bar{k}_i\sqrt{\sum_{j\neq i}\sigma_j^2} = k_b\sqrt{\sum_{i=1}^{n}\sigma_i^2}$. By Cantelli's inequality, we have

$$1 - \mathsf{P}\left(\bigcap_{i=1}^{n}(V_i < p_i)\right) = 1 - \prod_{i=1}^{n}\mathsf{P}(V_i < p_i = \mu_i - k_i\sigma_i) \geq 1 - \prod_{i=1}^{n}\frac{1}{1+k_i^2}. \tag{OA.9}$$

Denote $X_i = \sum_{j\neq i}V_j \geq p_b - p_i$ for $i \leq n$ and $X_{n+1} = \sum_{i=1}^{n}V_i$. We can easily obtain that the correlation between $X_{n+1}$ and $X_i$ is $\rho_{n+1,i} = \sqrt{\sum_{j\neq i}\sigma_j^2}\Big/\sqrt{\sum_{i=1}^{n}\sigma_i^2}$, and the correlation between $X_i$ and $X_j$ is $\rho_{ij} = \sum_{k\neq i,j}\sigma_k^2\Big/\left(\sqrt{\sum_{k\neq i}\sigma_k^2}\sqrt{\sum_{k\neq j}\sigma_k^2}\right)$. By Olkin and Pratt's inequality:

$$\mathsf{P}\left(\bigcap_{i=1}^{n}\frac{|X_i - \mu_i|}{\sigma_i} < k_i\right) \geq 1 - \frac{1}{n^2}\left(\sqrt{u} + \sqrt{n-1}\sqrt{n\sum_{i=1}^{n}\frac{1}{k_i^2} - u}\right)^2,$$

where $u = \sum_{i=1}^{n}\frac{1}{k_i^2} + 2\sum_{i=1}^{n}\sum_{j<i}\frac{\rho_{ij}}{k_ik_j}$ and $\rho_{ij}$ is the correlation between $X_i$ and $X_j$. We have

$$\begin{aligned}
&\mathsf{P}\left(\bigcap_{i=1}^{n}\left(\sum_{j\neq i}V_j \geq p_b - p_i\right)\bigcap\sum_{i=1}^{n}V_i \geq p_b\right)\\
=&\mathsf{P}\left(\bigcap_{i=1}^{n}\left(\sum_{j\neq i}(V_j-\mu_j)\Big/\sqrt{\sum_{j\neq i}\sigma_j^2} \geq -\bar{k}_i\right)\bigcap\sum_{i=1}^{n}(V_i-\mu_i)\Big/\sqrt{\sum_{i=1}^{n}\sigma_i^2} \geq -k_b\right)\\
\geq&\mathsf{P}\left(\bigcap_{i=1}^{n}\left(\left|\sum_{j\neq i}(V_j-\mu_j)\right|\Big/\sqrt{\sum_{j\neq i}\sigma_j^2} \leq \bar{k}_i\right)\bigcap\left|\sum_{i=1}^{n}(V_i-\mu_i)\right|\Big/\sqrt{\sum_{i=1}^{n}\sigma_i^2} \leq k_b\right)\\
\geq&1 - \frac{1}{(n+1)^2}\left(\sqrt{u} + \sqrt{n}\sqrt{(n+1)\left(\sum_{i=1}^{n}\frac{1}{\bar{k}_i^2} + \frac{1}{k_b^2}\right) - u}\right)^2 \equiv T,
\end{aligned} \tag{OA.10}$$

where $u = \sum_{i=1}^{n}\frac{1}{\bar{k}_i^2} + \frac{1}{k_b^2} + 2\left(\sum_{i=1}^{n}\sum_{j<i}\frac{\rho_{ij}}{\bar{k}_i\bar{k}_j} + \sum_{i=1}^{n}\frac{\rho_{n+1,i}}{k_b\bar{k}_i}\right)$.

If $p_i - c_i = \min_j\{p_j - c_j\}$, i.e., $\mu_i - c_i - k_i\sigma_i \leq \mu_j - c_j - k_j\sigma_j$ for $j \neq i$. Combining Equations (OA.8), (OA.9) and (OA.10), we have

$$\pi(p_i, p_b; F_i \in \mathscr{F}_i) \geq (\mu_i - c_i - k_i\sigma_i)\left(1 - \prod_{i=1}^{n}\frac{1}{1+k_i^2}\right) + \left(\sum_{j\neq i}(\mu_j - c_j) - \bar{k}_i\sqrt{\sum_{j\neq i}\sigma_j^2}\right)T.$$

Hence, we can solve the optimization problems (2), and select the largest value of the $n$ problems (one for each $i$), as the lower bound of $\pi(p_i, p_b; F_i \in \mathscr{F}_i)$. Moreover, the distribution-free pricing heuristic can be given by the corresponding optimal solutions. □