A Clustering Algorithm

Consider a set of units $a = (a_1, \ldots, a_{\text{card}(a)})$, where $\text{card}(s)$ is the cardinality of the set $a$. Let $r$ be a pre-defined cluster radius (in our main estimation we choose $r = 150$ km). For a set of units $b$, let $C_b$ be the weighted geographic center of $b$. The weighted geographic center of a set of units is calculated as the weighted average latitude and longitude of all the schools it contains, weighted by the school populations. The distance between two set of units $a$ and $b$, $\text{dist}(a, b)$, is the bird-fly distance between the weighted geographic centers of the respective sets.

We describe the algorithm that builds the set of clusters formed by the units in $a$, $Cl(a)$.

**Algorithm 1 Clustering Algorithm**

Let $\mathcal{A}$ be the collection of sets $\{\{a_1\}, \ldots, \{a_{\text{card}(a)}\}\}$, $d := 0$

while $d < r$ do

Let $d := \min_{b, c \in \mathcal{A}, b \neq c} \text{dist}(b, c)$. Let $(b^*, c^*) := \arg\min_{b, c \in \mathcal{A}, b \neq c} \text{dist}(b, c)$

if $d < r$ then

$\mathcal{A} := \mathcal{A} - (b^* \cup c^*$), $\mathcal{A} := \mathcal{A} \cup \{b^*, c^*\}$

end if

end while

$Cl(a) := \mathcal{A}$

B Computational Experiments on Strategic Bundling

As described in Section 5.2 of the paper the results in the multiproduct monopolist literature do not quite match our setting, because many of them (1) consider two units only; (2) focus on pure bundling; and (3) assume additive opponents’ bids and costs. In Appendix B.1, we show results from numerical experiments

*We thank Sang Won Kim for superb research assistance.
that deal with the first two limitations. The numerical experiments in Appendix B.2 deal with the last limitation.

B.1 Several Units, Flexible Package Bidding

In the first set of experiments we consider the computational experiments developed by Chu et al. (2009) to study factors that affect the magnitude of package discounts. These experiments analyze the problem of a multiproduct monopolist that chooses prices for all possible packages in a set of units and sells to consumers with random valuations. Consumers’ valuations and costs are assumed to be additive. As mentioned before, in this case Cantillon and Pesendorfer (2006) show that the bidder’s problem in a CA is equivalent to the multiproduct monopolist problem. We study the results in Chu et al. (2009) in the context of a CA.

We assume bidders’ statistical model of their opponents’ bids are given by equation (3) of the paper where \( v_i = 1 \), \( \forall i \). Using the notation introduced right after that equation, our main focus is to study how the following factors affect package discounts:

- The pairwise correlations across \( i \) among \( \delta_i \) in a given package.
- The unit cost of the bidders.
- The variance of \( \delta_i \).

Chu et al. (2009) find the optimal bid prices in scenarios with different bid distributions, different cost structures, and auctions ranging from 2 to 5 units. Because we are interested in analyzing the effect of correlations and variance of the bids, we focus on the experiments in which \( (\delta_1, \ldots, \delta_{N(t)}) \) are assumed to be multivariate normally distributed. We focus on experiments with \( N(t) \) equal to 4 or 5 units. These numerical experiments can be grouped into three categories:

1. Experiments for which the \( \delta_i \)'s are all independent random variables, or all pairwise positively correlated, or all pairwise negatively correlated. Their variances are the same across units and experiments.

2. Experiments for which the pairwise correlation of the \( \delta_i \)'s vary across pairs of units; some pairs are independent and others are positively or negatively correlated. Their variances are the same across units and experiments.

3. A third set of experiments that is similar to the first set but the variance of the \( \delta_i \)'s change across units and experiments.

We first identified “nested packages” on every experiment and calculated the discount from adding a unit \( i \) to package \( a \) as:

\[
Disc(i, a) = b_i + b_a - b_{i \cup a}
\]

For the first set of experiments, we compared the magnitude of the discounts among the cases with independent, positive correlation and negative correlation, for packages of different sizes. Note that the

---

1 More details about the model parameters can be found in Chu et al. (2009).
average is taken over different instances that include auctions of different size, different means of the \( \delta_i \)'s, and different bidders’ marginal costs of serving the units; the averages are computed over the same set of instances for each level of correlation. Table 1 shows this comparison.

<table>
<thead>
<tr>
<th>Package size</th>
<th>Negative</th>
<th>Indep.</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.65</td>
<td>0.74</td>
<td>0.49</td>
</tr>
<tr>
<td>3</td>
<td>2.15</td>
<td>1.12</td>
<td>0.75</td>
</tr>
<tr>
<td>4</td>
<td>2.31</td>
<td>1.29</td>
<td>0.94</td>
</tr>
<tr>
<td>5</td>
<td>2.39</td>
<td>1.36</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Table 1: Average discounts for different levels of pairwise correlation.

Table 1 shows two interesting patterns. First, the discounts tend to increase as the size of the package increases, suggesting that the effect of strategic bundling is observed beyond the two-unit case. Second, as the correlation among the units becomes more negative, the discounts become larger. We also used these experiments to study the effect of costs. The following table compares the discounts for experiments for which the bidder’s marginal costs are below or above the mean of the \( \delta_i \)'s distribution. The results show that strategic discounts are mostly relevant when the bidder has relative cost advantages.

<table>
<thead>
<tr>
<th>Costs</th>
<th>Negative</th>
<th>Indep.</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>below mean</td>
<td>3.45</td>
<td>1.56</td>
<td>1</td>
</tr>
<tr>
<td>above mean</td>
<td>0.4</td>
<td>0.41</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 2: Average discounts for different levels of pairwise correlation and costs.

To study the second set of experiments, we estimated a regression with \( \bar{Disc}(i,a) \equiv Disc(i,a)/(|a|+1) \) as the dependent variable (we normalize discounts by the size of the package), and an independent variable measuring the average pairwise correlation of unit \( i \) with units in package \( a \) (\( AvgCorr \)). Each experiment –indexed by \( k \)– is specified by: (i) a total number of units in the auction (from 2 to 5); (ii) a multivariate normal distribution with mean \( \mu \) and variance-covariance \( \Omega \) that determines the distribution of the minimum of opponents’ bids, \( \delta_i \) \( (\Omega \text{ has 0.25 in the diagonal and zeros, negative and positive values off the diagonal ranging from -0.5 to 0.5)} \); and (iii) the marginal costs of each unit (either 0 or 0.2 for all units). The regression is given by:

\[
\bar{Disc}_k(i,a) = \beta AvgCorr_k(i,a) + \rho_{ka} + \varepsilon_{kia}.
\]

We considered the discounts in different types of packages: (i) packages of two units, with marginal costs for both units below the mean of the respective \( \delta_i \) (low cost scenario) or both units above the mean (high cost scenario); (ii) similar cases for packages of 3 or more units. For packages of 3 or more, we also include the fixed effect \( \rho_{ka} \) to control for the experiment conditions and fixed effects for packages \( (a).^2 \)

Table 3 shows the results from these regressions. The results suggest that when the unit costs are low, discounts tend to be larger when the units are more negatively correlated. This effect is not present when

---

^2Including this fixed effects for the regressions on packages of two units does not change the point estimates but reduces their precision.

3
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 units, low cost</td>
<td>2 units, high cost</td>
<td>&gt;=3 units, low cost</td>
<td>&gt;=3, high cost</td>
</tr>
<tr>
<td>Avgcorr</td>
<td>-0.179***</td>
<td>0.049*</td>
<td>-0.135***</td>
<td>0.173***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.023)</td>
<td>(0.007)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>N</td>
<td>10575</td>
<td>1260</td>
<td>55101</td>
<td>11331</td>
</tr>
</tbody>
</table>

Table 3: Effect of average correlation in discounts.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 units</td>
<td>&gt;=3 units</td>
</tr>
<tr>
<td>Var</td>
<td>0.099***</td>
<td>0.103***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>N</td>
<td>11340</td>
<td>51660</td>
</tr>
</tbody>
</table>

Table 4: Effect of variance in discounts.

the firm has higher costs. Moreover, the results are similar when considering larger packages of 3 or more units.

Finally, we used the third set of experiment to analyze the effect of the variance of units in the package discounts. Each experiment –indexed by $k$– is characterized by a total number of units (from 3 to 5), the variance of each unit (from 0.25 to 1.75) and a pairwise correlation coefficient (0, -0.5 or 0.5). We estimated the following regression:

$$\overline{Disc}_k(i,a) = \beta Var_k(i) + \rho_{ka} + \epsilon_{ki,a}. $$

The idea is to estimate how the variance of a unit ($Var_k(i)$) affects the discount of adding a unit to package $a$, controlling for the package $a$ in each experimental condition $k$ through the fixed-effect $\rho_{ka}$. The results of this estimation are reported in Table 4. The results suggest that the discount increases with the variance of the unit added to the package.

In summary, the computational experiments in these section suggest that:

- When firms have a relative cost advantage, they tend to provide larger discounts when combining units which are more negatively correlated.

- The discount of adding a unit to a package increases with the variance of the added unit.\(^4\)

\(^3\)Marginal costs and mean of the competitors’ bid distribution are zero in all the experiments.

\(^4\)In subsequent analysis, because units are heterogeneous in their average bid prices, we measure the variability of the bid price with the coefficient of variation.
This is consistent with the intuition from the multiproduct monopolist literature that a reduction in the variance from combining units makes it easier for a firm with lower costs to extract surplus, which translates into larger package discounts.

B.2 Sub-additive bids and synergies

The second set of computational experiments addresses a limitation of the previous analysis: in the CA for school meals the bids are subadditive and there are cost synergies. In this section we compute the optimal bid prices of a firm that competes in a CA assuming that its opponents make package discounts and the firm exhibits cost synergies. Because of the computational complexity involved in computing optimal bid prices we restrict our analysis to two units. To simplify, both units have the same volume. We assume the firm’s costs are given by \( c = (c_1, c_2, c_{12}) \). We allow for cost synergies, so we can have \( c_{12} < c_1 + c_2 \). Opponents’ bids (per meal) are modeled like:

\[
b_{af} = \sum_{i \in a} \delta_{if} \frac{v_i}{v_a} - g_f 1[|a| > 1],
\]

where \( (\delta_{1f}, \delta_{2f}) \) are i.i.d. samples from a multivariate normal distribution with mean vector \( \mu \) and covariance matrix \( \Omega \). The term \( g_f 1[|a| > 1] \) allows for opponents’ package discounts; \( g_f \) are i.i.d. samples from a uniform distribution \([g, g]\). We assume there are 10 bidders.

To determine the optimal bid prices we use simulation based optimization (Fu et al., 2005). In particular, we construct a grid of bid prices. For each bid price in the grid, \( (b_1^j, b_2^j, b_{12}^j) \), expected profits are given by \( \sum_{j \in \{1, 2, 12\}} (b_j - c_j) P_r[\text{win } j] \), where the uncertainty involved in computing \( P_r[\text{win } j] \) is with respect to the opponents’ bids. For each bid price in the grid, we compute expected profits using simulation by sampling opponents’ bids from equation (1) (we use a relative precision of 0.2%). The optimal bid prices maximize expected profits among all the grid points.

We computed the optimal bid prices for many different scenarios. We consider the following range of unit costs: \( (c_1, c_2) \in [4, 14]^2 \). Cost synergies are contained in the interval \( c_1 + c_2 - c_{12} \in [0, 4] \). We fixed average opponents’ single unit prices \( \mu(\delta_{1f}) = 15 \). The standard deviation of individual bid prices range in \( \sigma(\delta_{1f}) \in \{1, 2, 4\} \), and the correlation among individual bid prices \( \rho(\delta_{1f}, \delta_{2f}) \in \{-0.8, -0.4, 0, 0.4, 0.8\} \). Finally the upper bound for the opponents’ package discounts is in the range \( g \in [0, 6] \) while the lower bound is fixed to \( g = 0 \). We have 8101 scenarios in total. We use the following grid: given the firm’s cost vector \( (c_1, c_2, c_{12}) \), each candidate of bid price vector, or grid point, which is indexed by \( j \), \( (b_1^j, b_2^j, b_{12}^j) \), lies within the following range with incremental size of \( h = 0.1 \):

\[
c_1 \leq b_1^j \leq \mu(\delta_{1f}) + 3\sigma(\delta_{1f})
\]
\[
c_2 \leq b_2^j \leq \mu(\delta_{2f}) + 3\sigma(\delta_{2f})
\]
\[
c_{12} \leq b_{12}^j \leq \max_{i,k} \{b_i^j + b_k^j\}
\]

The solution in many of the experiments sets prices for stand-alone bids which have zero probability

\(^5\)To make the comparison of the profits more reliable, we used the same random number seed for all grid points.
Table 5: Effect of coefficient of variation and variance in discounts.

<table>
<thead>
<tr>
<th></th>
<th>(1) Low Cost</th>
<th>(2) High Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td>3.388**</td>
<td>1.927**</td>
</tr>
<tr>
<td></td>
<td>(0.359)</td>
<td>(0.272)</td>
</tr>
<tr>
<td>Corr</td>
<td>-0.766**</td>
<td>-0.284**</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Observations</td>
<td>944</td>
<td>97</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
* p < 0.01, ** p < 0.001

The results suggest that the patterns suggested by the analysis with additive bids (section B.1) continue to hold under sub-additive bids and cost synergies for \( N = 2 \): (1) discounts tend to increase when the correlation between units decreases and when the coefficient of variation of the units increase; and (2) the incentives for strategic bundling are more pronounced for firms with a relative cost advantage.

References


---

*In experiments where the optimal solution has irrelevant stand-alone bids, the discount can be infinite. We also analyzed all experiments by fixing the discount at the minimum value at which the stand-alone bids are irrelevant and the observed patterns in the results are similar.