E-companion for "Offline-Channel Planning in Smart Omnichannel Retailing"

EC.1. Proofs of Statements

The proofs are ordered based on their first appearances in the main text.

EC.1.1. Proof of Proposition 1.

Proof: We prove by showing a reduction of the *Uncapacitated Facility Location* (UFL) problem to problem (**LA**). Because UFL is known to be NP-hard, we have that problem (**LA**) is NP-hard.

To see the reduction, first note that given $\boldsymbol{X}$ and $\boldsymbol{Z}$, the optimal expected revenue $\mathcal{Q}(\boldsymbol{X},\boldsymbol{Z})$ is decomposable by each $i\in\mathcal{I}$. Indeed, define

$$\mathscr{Y}_i(X_i)=\left\{\boldsymbol{Y}\in\{0,1\}^K \,\middle|\, \sum_{k\in\mathcal{K}} a_k Y_k \le c_i X_i \right\}.$$

Immediately, we have $\mathscr{Y}_i(0)=\{\mathbf{0}\}$ and $\mathscr{Y}_i(1)=\{\boldsymbol{Y}\in\{0,1\}^K \mid \sum_{k\in\mathcal{K}} a_k Y_k\le c_i\}$. For given $\boldsymbol{X}$ and $\boldsymbol{Z}$, the objective of $\mathcal{Q}(\boldsymbol{X},\boldsymbol{Z})$ can be rewritten as

$$\begin{aligned}
\mathcal{Q}(\boldsymbol{X},\boldsymbol{Z}) =& \max_{\boldsymbol{Y}_i\in\mathscr{Y}_i(X_i),\forall i\in\mathcal{I}} \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}} r_k\mu_j\theta_{l|j}\left(\lambda_{0|jl}(\boldsymbol{Z})p_{k|0l}+\sum_{i\in\mathcal{I}_j}\lambda_{i|jl}(\boldsymbol{Z})p_{k|il}(\boldsymbol{Y}_i)\right)\\
=&\mathcal{Q}_0(\boldsymbol{Z})+\max_{\boldsymbol{Y}_i\in\mathscr{Y}_i(1),\forall i\in\mathcal{I}(\boldsymbol{X})} \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}} r_k\mu_j\theta_{l|j}\sum_{i\in\mathcal{I}_j\cap\mathcal{I}(\boldsymbol{X})}\lambda_{i|jl}(\boldsymbol{Z})p_{k|il}(\boldsymbol{Y}_i)\\
=&\mathcal{Q}_0(\boldsymbol{Z})+\max_{\boldsymbol{Y}_i\in\mathscr{Y}_i(1),\forall i\in\mathcal{I}(\boldsymbol{X})} \sum_{i\in\mathcal{I}(\boldsymbol{X})}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}} r_k\left(\sum_{j\in\mathcal{J}:i\in\mathcal{I}_j}\mu_j\theta_{l|j}\lambda_{i|jl}(\boldsymbol{Z})\right)p_{k|il}(\boldsymbol{Y}_i)\\
=&\mathcal{Q}_0(\boldsymbol{Z})+\sum_{i\in\mathcal{I}(\boldsymbol{X})}\max_{\boldsymbol{Y}_i\in\mathscr{Y}_i(1)}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}} r_k\widehat{\mu}_{il}(\boldsymbol{Z})p_{k|il}(\boldsymbol{Y}_i)\\
:=&\sum_{i\in\mathcal{I}(\boldsymbol{X})\cup\{0\}}\mathcal{Q}_i(\boldsymbol{Z}).
\end{aligned}$$

Next, note that each $\mathcal{Q}_i(\boldsymbol{Z})$ further reduces to

$$\sum_{j\in\mathcal{J}:i\in\mathcal{I}_j} Z_{ij}\,\mathcal{Q}_i\mu_j\frac{e^{\beta_i-\alpha d_{ij}}}{1+e^{\beta_i-\alpha d_{ij}}}$$

when customers are assumed to be homogeneous. Note that $\mathcal{Q}_i$'s can be obtained by solving separate optimization problems. Consequently, we have $\mathcal{Q}(\boldsymbol{X},\boldsymbol{Z})=\sum_{i\in\mathcal{I},j\in\mathcal{J}:i\in\mathcal{I}_j} Z_{ij}\mathcal{D}_{ij}$, where $\mathcal{D}_{ij}=\mathcal{Q}_i\mu_j e^{\beta_i-\alpha d_{ij}}/(1+e^{\beta_i-\alpha d_{ij}})$. Plugging into problem (**LA**), we obtain a linear objective function consisting of fixed costs of locating facilities and costs associated with virtual travel distances $\mathcal{D}_{ij}$. Then, it can be verified that constraint (4) with properly chosen coefficients ensures customer assignments based on proximity. Therefore, any instance of problem (**LA**) converts into an instance of the UFL problem. Conversely, every UFL instance reduces to an instance of problem (**LA**) with homogeneous customers with a reverse construction. □

EC.1.2. Proof of Proposition 2.

First by our assumption, a customer from region j will never visit a store in $\mathcal{I}\setminus\{i(j)\}$. Therefore, $\widehat{n}^t_{ijkl}=0$ if $i\in\mathcal{I}\setminus\{i(j)\}$, for all $t\in\mathcal{T}$, $j\in\mathcal{J}$, $l\in\mathcal{L}$ and $k\in\overline{\mathcal{K}}$. Next, given $\widehat{\Omega}$, we consider the following cases for $i\in\{0,i(j)\}$.

Case A ($i=0$ and $k\in\mathcal{K}$). Fix $t\in\mathcal{T}$, $j\in\mathcal{J}$ and $l\in\mathcal{L}$. For product $k\in\mathcal{K}$, we have

$$\begin{aligned}p_{0jkl} &:= \mathbb{P}\{\text{the customer is of class } l|\text{a region-}j \text{ online customer purchases } k\}\\ &= \frac{\mathbb{P}\{\text{a region-}j\text{ and class-}l\text{ online customer purchases } k\}}{\mathbb{P}\{\text{a region-}j\text{ online customer purchases } k\}} = \frac{\widehat{\mu}_j\widehat{\theta}_{l|j}\widehat{\lambda}_{0|jl}P_{k|l}(\emptyset,\widehat{\boldsymbol{v}})}{\sum_{l'\in\mathcal{L}}\widehat{\mu}_j\widehat{\theta}_{l'|j}\widehat{\lambda}_{0|jl'}P_{k|l'}(\emptyset,\widehat{\boldsymbol{v}})},\end{aligned}$$

which yields

$$\widehat{n}^t_{0jkl}=\mathbb{E}[n^t_{0jkl}|\boldsymbol{w}^t]=p_{0jkl}w^t_{0jk}=\frac{\widehat{\theta}_{l|j}\widehat{\lambda}_{0|jl}P_{k|l}(\emptyset,\widehat{\boldsymbol{v}})}{\sum_{l'\in\mathcal{L}}\widehat{\theta}_{l'|j}\widehat{\lambda}_{0|jl'}P_{k|l'}(\emptyset,\widehat{\boldsymbol{v}})}w^t_{0jk}.$$

Case B ($i=i(j)$ and $k\in\mathcal{K}$). Fix $t\in\mathcal{T}$, $j\in\mathcal{J}$ and $l\in\mathcal{L}$, and use i to denote $i(j)$ for notation brevity. Note that

$$\widehat{n}^t_{ijkl}=\mathbb{E}[n^t_{ijkl}|\boldsymbol{w}^t]=\mathbb{E}[a^t|\boldsymbol{w}^t]\widehat{\mu}_j\widehat{\theta}_{l|j}\widehat{\lambda}_{i|jl}P_{k|l}(\mathcal{K},\widehat{\boldsymbol{v}}),$$

and

$$w^t_{ijk}=\mathbb{E}[a^t|\boldsymbol{w}^t]\widehat{\mu}_j\sum_{l'\in\mathcal{L}}\widehat{\theta}_{l'|j}\widehat{\lambda}_{i|jl'}P_{k|l'}(\mathcal{S}^t_i,\widehat{\boldsymbol{v}}).$$

By eliminating $\mathbb{E}[a^t|\boldsymbol{w}^t]$, we have

$$\widehat{n}^t_{ijkl}=\frac{\widehat{\theta}_{l|j}\widehat{\lambda}_{i|jl}P_{k|l}(\mathcal{K},\widehat{\boldsymbol{v}})}{\sum_{l'\in\mathcal{L}}\widehat{\theta}_{l'|j}\widehat{\lambda}_{i|jl'}P_{k|l'}(\mathcal{S}^t_i,\widehat{\boldsymbol{v}})}w^t_{ijk}.$$

Case C ($k=0$). Now we consider the no-purchase option $k=0$. Note that $\mathbb{E}[a^t|\boldsymbol{w}^t]=\sum_{i\in\overline{\mathcal{I}}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}}\widehat{n}^t_{ijkl}+\sum_{i\in\overline{\mathcal{I}}}\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\widehat{n}^t_{ij0l}$. Note that we have

$$\widehat{n}^t_{0j0l}=\widehat{\mu}_j\widehat{\theta}_{l|j}\widehat{\lambda}_{0|jl}P_{0|l}(\emptyset,\widehat{\boldsymbol{v}})\mathbb{E}[a^t|\boldsymbol{w}^t], \tag{EC.1}$$

$$\widehat{n}^t_{i(j),j0l}=\widehat{\mu}_j\widehat{\theta}_{l|j}\widehat{\lambda}_{i(j)|jl}P_{0|l}(\mathcal{K},\widehat{\boldsymbol{v}})\mathbb{E}[a^t|\boldsymbol{w}^t]. \tag{EC.2}$$

Thus we have

$$\mathbb{E}[a^t|\boldsymbol{w}^t]=\sum_{i\in\overline{\mathcal{I}}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}}\widehat{n}^t_{ijkl}+\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\widehat{\mu}_j\widehat{\theta}_{l|j}\left(\widehat{\lambda}_{0|jl}P_{0|l}(\emptyset,\widehat{\boldsymbol{v}})+\widehat{\lambda}_{i(j)|jl}P_{0|l}(\mathcal{K},\widehat{\boldsymbol{v}})\right)\mathbb{E}[a^t|\boldsymbol{w}^t],$$

$$\begin{aligned}\mathbb{E}[a^t|\boldsymbol{w}^t]&=\frac{\sum_{i\in\overline{\mathcal{I}}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}}\widehat{n}^t_{ijkl}}{1-\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\widehat{\mu}_j\widehat{\theta}_{l|j}\left(\widehat{\lambda}_{0|jl}P_{0|l}(\emptyset,\widehat{\boldsymbol{v}})+\widehat{\lambda}_{i(j)|jl}P_{0|l}(\mathcal{K},\widehat{\boldsymbol{v}})\right)}\\ &=\frac{\sum_{i\in\overline{\mathcal{I}}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}}\widehat{n}^t_{ijkl}}{\sum_{i'\in\overline{\mathcal{I}}}\sum_{j'\in\mathcal{J}}\sum_{l'\in\mathcal{L}}\sum_{k'\in\mathcal{K}}\widehat{\mu}_{j'}\widehat{\theta}_{l'|j'}\widehat{\lambda}_{i'|j'l'}P_{k'|l'}}\end{aligned}$$

Plugging the above equation into (EC.1) and (EC.2) completes the proof. □

EC.1.3. Proof of Proposition 3

Based on the definition of Λ_t, we have

$$\Lambda_t^* = \mathbb{E}[a^t|\boldsymbol{w}^t] = \sum_{i\in\overline{\mathcal{I}}}\sum_{j\in\mathcal{J}}\sum_{k\in\overline{\mathcal{K}}}\sum_{l\in\mathcal{L}}\widehat{n}_{ijkl}^t.$$

As (8) is decomposable, $\boldsymbol{\mu}^*$, $\boldsymbol{\theta}^*$ and $\boldsymbol{\lambda}^*$ can be obtained by solving the following problems:

$$\boldsymbol{\mu}^* = \arg\max_{\boldsymbol{\mu}\ge 0}\left\{\sum_{j\in\mathcal{J}}\left(\sum_{t\in\mathcal{T}}\sum_{i\in\overline{\mathcal{I}}}\sum_{l\in\mathcal{L}}\sum_{k\in\overline{\mathcal{K}}}\widehat{n}_{ijkl}^t\right)\ln\mu_j; \sum_{j\in\mathcal{J}}\mu_j = 1\right\}$$

$$\boldsymbol{\theta}_j^* = \arg\max_{\boldsymbol{\theta}_j\ge 0}\left\{\sum_{l\in\mathcal{L}}\left(\sum_{t\in\mathcal{T}}\sum_{k\in\overline{\mathcal{K}}}\sum_{i\in\overline{\mathcal{I}}}\widehat{n}_{ijkl}^t\right)\ln\theta_{l|j} : \sum_{l\in\mathcal{L}}\theta_{l|j} = 1\right\}, j\in\mathcal{J}$$

$$\boldsymbol{\lambda}_{\cdot|jl}^* = \arg\max_{\boldsymbol{\lambda}_{\cdot|jl}\ge 0}\left\{\sum_{t\in\mathcal{T}}\sum_{i\in\{i(j),0\}}\sum_{k\in\overline{\mathcal{K}}}\widehat{n}_{ijkl}^t\ln\lambda_{i|jl} : \sum_{i\in\{0,i(j)\}}\lambda_{i|jl} = 1, \lambda_{i|jl} = 0\ i\in\mathcal{I}\setminus\{i(j)\}\right\}, j\in\mathcal{J}, l\in\mathcal{L},$$

$$\boldsymbol{v}_{l,\text{on}}^* = \arg\max_{\boldsymbol{v}_{l,\text{on}}\ge 0}\sum_{k\in\mathcal{K}}\left(\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\widehat{n}_{0jkl}^t\right)\ln v_{kl,\text{on}} - \sum_{k\in\overline{\mathcal{K}}}\left(\sum_{t=1}^{T}\sum_{j\in\mathcal{J}}\widehat{n}_{0jkl}^t\right)\ln\left(1+\sum_{k'\in\mathcal{K}}v_{k'l,\text{on}}\right),\ l\in\mathcal{L}$$

$$\boldsymbol{v}_{l,\text{off}}^* = \arg\max_{\boldsymbol{v}_{l,\text{off}}\ge 0}\sum_{k\in\mathcal{K}}\left(\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\widehat{n}_{ijkl}^t\right)\ln v_{kl,\text{off}} - \sum_{k\in\overline{\mathcal{K}}}\left(\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\widehat{n}_{ijkl}^t\right)\ln\left(1+\sum_{k'\in\mathcal{K}}v_{k'l,\text{off}}\right),\ l\in\mathcal{L}$$

which have closed-form solutions given by

$$\mu_j^* = \frac{\sum_{t\in\mathcal{T}}\sum_{i\in\overline{\mathcal{I}}}\sum_{l\in\mathcal{L}}\sum_{k\in\overline{\mathcal{K}}}\widehat{n}_{ijkl}^t}{\sum_{j'\in\mathcal{J}}\sum_{t\in\mathcal{T}}\sum_{i\in\overline{\mathcal{I}}}\sum_{l\in\mathcal{L}}\sum_{k\in\overline{\mathcal{K}}}\widehat{n}_{ij'kl}^t},\quad j\in\mathcal{J}$$

$$\theta_{l|j}^* = \frac{\sum_{t\in\mathcal{T}}\sum_{k\in\overline{\mathcal{K}}}\sum_{i\in\overline{\mathcal{I}}}\widehat{n}_{ijkl}^t}{\sum_{l'\in\mathcal{L}}\sum_{t\in\mathcal{T}}\sum_{k\in\overline{\mathcal{K}}}\sum_{i\in\overline{\mathcal{I}}}\widehat{n}_{ijkl'}^t},\quad l\in\mathcal{L}, j\in\mathcal{J}$$

$$\lambda_{i|jl}^* = \begin{cases}\frac{\sum_{t\in\mathcal{T}}\sum_{k\in\overline{\mathcal{K}}}\widehat{n}_{ijkl}^t}{\sum_{t\in\mathcal{T}}\sum_{i'\in\overline{\mathcal{I}}}\sum_{k\in\overline{\mathcal{K}}}\widehat{n}_{i'jkl}^t} & i\in\{0,i(j)\},\\ 0 & i\in\mathcal{I}\setminus\{i(j)\},\end{cases}\quad j\in\mathcal{J}, l\in\mathcal{L}$$

$$v_{kl,\text{on}}^* = \frac{\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\widehat{n}_{0jkl}^t}{\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}\widehat{n}_{0j0l}^t},\quad k\in\mathcal{K}, l\in\mathcal{L}$$

$$v_{kl,\text{off}}^* = \frac{\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\widehat{n}_{ijkl}^t}{\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}\widehat{n}_{ij0l}^t},\quad k\in\mathcal{K}, l\in\mathcal{L}.$$

Thus we finish the proof. □

EC.1.4. Proof of Lemma 1.

Proof: Note that (i) for any fixed j and l, because there is up to one $i\in\mathcal{I}_j$ such that $Z_{ij}=1$ (due to constraints (4)), then we have $\lambda_{i|jl}(\boldsymbol{Z}) = \widehat{\lambda}_{i|jl}Z_{ij}$, for all $i\in\mathcal{I}_j$, and (ii) $e^{u_{kl}+\delta_{kl}Y_{ik}} = e^{u_{kl}}Y_{ik,\text{on}} + e^{u_{kl}+\delta_{kl}}Y_{ik,\text{off}}$ by constraints (11) and (12). Therefore, the objective function of problem (**LA**) is equal to

$$\sum_{i\in\mathcal{I}}f_iX_i - \sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}}r_k\mu_j\theta_{l|j}\left(\lambda_{0|jl}(\boldsymbol{Z})p_{k|0l} + \sum_{i\in\mathcal{I}_j}\lambda_{i|jl}(\boldsymbol{Z})p_{k|il}(\mathbf{Y}_i)\right)$$

$$
\begin{aligned}
&= \sum_{i\in\mathcal{I}} f_i X_i - \sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}} \mu_j \theta_{l|j} \left(\lambda_{0|jl}(\boldsymbol{Z}) \frac{\sum_{k\in\mathcal{K}} r_k e^{u_{kl}}}{1+\sum_{k\in\mathcal{K}} e^{u_{kl}}} + \sum_{i\in\mathcal{I}_j} \lambda_{i|jl}(\boldsymbol{Z}) \frac{\sum_{k\in\mathcal{K}} r_k e^{u_{kl}+\delta_{kl}Y_{ik}}}{1+\sum_{k\in\mathcal{K}} e^{u_{kl}+\delta_{kl}Y_{ik}}} \right) \\
&= \sum_{i\in\mathcal{I}} f_i X_i - \overline{r}\sum_{j\in\mathcal{J}} \mu_j \\
&\quad + \sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}} \mu_j \theta_{l|j} \left(\lambda_{0|jl}(\boldsymbol{Z}) \frac{\overline{r}+\sum_{k\in\mathcal{K}} (\overline{r}-r_k) e^{u_{kl}}}{1+\sum_{k\in\mathcal{K}} e^{u_{kl}}} + \sum_{i\in\mathcal{I}_j} \lambda_{i|jl}(\boldsymbol{Z}) \frac{\overline{r}+\sum_{k\in\mathcal{K}} (\overline{r}-r_k) e^{u_{kl}+\delta_{kl}Y_{ik}}}{1+\sum_{k\in\mathcal{K}} e^{u_{kl}+\delta_{kl}Y_{ik}}} \right) \\
&= \sum_{i\in\mathcal{I}} f_i X_i - \overline{r}\sum_{j\in\mathcal{J}} \mu_j \\
&\quad + \sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}} \mu_j \theta_{l|j} \left(\left(1-\sum_{i\in\mathcal{I}_j} \widehat{\lambda}_{i|jl} Z_{ij}\right) R_{0l} + \sum_{i\in\mathcal{I}_j} \widehat{\lambda}_{i|jl} Z_{ij} \frac{\overline{r}+\sum_{k\in\mathcal{K}} (\overline{r}-r_k)\sum_{c\in\mathcal{C}} v_{kl,c} Y_{ik,c}}{1+\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} v_{kl,c} Y_{ik,c}} \right).
\end{aligned}
$$

Dropping the constant term $-\overline{r}\sum_{j\in\mathcal{J}}\mu_j + \sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}}\mu_j\theta_{l|j}R_{0l}$ in the above expression leads to the equivalent objective function (10). □

EC.1.5. Proof of Proposition 4.

Proof: Omitted. □

EC.1.6. Proof of Proposition 5.

Proof: Note that

$$
R_{ij}(\boldsymbol{Y}_i) = \overline{r}\sum_{l\in\mathcal{L}} \theta_{l|j}\widehat{\lambda}_{i|jl} - \sum_{l\in\mathcal{L}} \theta_{l|j}\widehat{\lambda}_{i|jl} \frac{\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} r_k v_{kl,c} Y_{ik,c}}{1+\sum_{k\in\mathcal{K}}\sum_{c\in\mathcal{C}} v_{kl,c} Y_{ik,c}}.
$$

Along the same line of derivation of the reformulation (**LA-MISOCP**) as in Section 5.1, we have (19) and (20). This completes the proof. □

EC.1.7. Proof of Proposition 6

Proof: Fix a permutation $\boldsymbol{\pi}$, it is evident that for any $n \in \{1,2,\ldots,2K\}$, we have $\hat{\sigma}_{\pi_n} \geq \sigma_{\pi_n}$. Then, by noting that $\sqrt{\cdot}$ defines a polymatroid function, we obtain that $\hat{\rho}_{\pi_n} \leq \rho_{\pi_n}$ for all $n \in \{1,2,\ldots,2K\}$. Consequently, any $(\boldsymbol{y}, z)$ satisfying inequality (26) must also satisfy inequality (27). Then, the result follows from the validity of inequality (26) (proved in Atamtürk and Gómez (2020)). □

EC.1.8. Proof of Proposition 7.

Proof: Subsuming the notation $\widehat{b}_1,\ldots,\widehat{b}_{2K}$ and $\widehat{c}$ as defined in the proposition's statement, we can re-express $\hat{\sigma}_{\pi_n}$ as the following LP:

$$
\hat{\sigma}_{\pi_n} = b_0 + \max_{\boldsymbol{y}\geq 0} \sum_{s=1}^{2K} \widehat{b}_s y_s
$$

$$\text{s.t. } y_{2k-1} + y_{2k} = 1 \quad \forall k \in \mathcal{K} \setminus \{n^\circ\}$$
$$\sum_{k \in \mathcal{K}\setminus\{n^\circ\}} a_k y_{2k} \le \widehat{c}.$$

This is because notation $\widehat{b}_1, \ldots, \widehat{b}_{2K}$ is defined in a way that allows the use of the same lower and upper limits of summation for all n. By the strong duality, we have

$$\hat{\sigma}_{\pi_n} = b_0 + \min_{\boldsymbol{u}\in\mathbb{R}^K, v\ge 0} \left\{ \widehat{c}v + \sum_{k\in\mathcal{K}\setminus\{n^\circ\}} u_k \right\}$$
$$\text{s.t. } u_k \ge \widehat{b}_{2k-1} \quad \forall k \in \mathcal{K}\setminus\{n^\circ\} \tag{EC.3}$$
$$u_k + a_k v \ge \widehat{b}_{2k} \quad \forall k \in \mathcal{K}\setminus\{n^\circ\}. \tag{EC.4}$$

Note that constraints (EC.3) and (EC.4) imply that $u_k = \max\{\widehat{b}_{2k-1}, \widehat{b}_{2k} - a_k v\}$ for all $k \in \mathcal{K}\setminus\{n^\circ\}$. Then, it can be verified that $u_k = \widehat{b}_{2k-1}$ for all $k \in \widehat{\mathcal{K}}^-$. Consequently, we have

$$\hat{\sigma}_{\pi_n} = b_0 + \sum_{k\in\widehat{\mathcal{K}}^-} \widehat{b}_{2k-1} + \min_{v\ge 0} f(v), \tag{EC.5}$$

where $f(v) := \widehat{c}v + \sum_{k\in\widehat{\mathcal{K}}^+} \max\{\widehat{b}_{2k-1}, \widehat{b}_{2k} - a_k v\}$.

We now study the monotonicity of $f(v)$. Notice that $f(v)$ is a univariate piece-wise linear function, of which the derivative $f'(v)$ satisfies

$$f'(v) = \widehat{c} - \sum_{t'=t}^{|\widehat{\mathcal{K}}^+|} a_{\widehat{\pi}_{t'}}, \ \forall v \in (v_{t-1}, v_t), \ t = 1, 2, \ldots, |\widehat{\mathcal{K}}^+|, \tag{EC.6}$$

where $v_0 = 0$ and $v_t = (\widehat{b}_{2\widehat{\pi}_t} - \widehat{b}_{2\widehat{\pi}_t - 1})/a_{\widehat{\pi}_t}$ for $t = 1, 2, \ldots, |\widehat{\mathcal{K}}^+|$. By the definition of t^* and (EC.6), we obtain that $f(v)$ takes its minimum at v_{t^*}, that is, $\min_{v\ge 0} f(v) = f(v_{t^*})$. Then, equation (EC.5) yields (29), and this completes the proof. □

EC.2. Algorithms

Here, we present two additional algorithms to improve the performance of the proposed parameter estimation approach. In particular, Algorithm 1 performs model selection by determining the optimal number of latent classes, whereas Algorithm 2 provides an initialization approach that can increase the computational efficiency of the EM algorithm.

EC.3. Related Concepts

Conic programming problems are optimization problems with linear objective functions and conic constraints. A second-order conic constraint on decision variables $\boldsymbol{x} \in \mathbb{R}^n$ is of the form

$$\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\| \le \boldsymbol{c}^T \boldsymbol{x} - d,$$

Algorithm 1 Procedure for model selection

Input $\mathcal{J}$, $\mathcal{K}$, Data $\mathcal{N}$, convergence level ϵ.

Ouput Latent class index set $\mathcal{L}$, parameter set Ω.

Set $\mathcal{L} = \emptyset$, $\text{AIC}^0 = \infty$.

while $\text{AIC}^{|\mathcal{L}|} < \text{AIC}^{|\mathcal{L}|-1}$ **do**

  Set $\mathcal{L} = \mathcal{L} \cup \{|\mathcal{L}| + 1\}$

  EM Procedure: Apply Algorithm 1 and obtain $\widehat{\Omega}$ and $\ln \mathscr{L}_I(\widehat{\Omega})$.

  Compute AIC: $\text{AIC}^{|\mathcal{L}|} = -2\ln \mathscr{L}_{\text{I}}(\widehat{\Omega}) + 2|\widehat{\Omega}|$.

end while

Algorithm 2 Procedure for an initial solution

Input: $\mathcal{J}$, $\mathcal{L}$,$\mathcal{K}$, Data $\mathcal{N}$, constant γ

Output: $\Omega^0 = (\boldsymbol{\Lambda}^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0, \boldsymbol{\theta}^0, \boldsymbol{v}^0)$

for $i \in \mathcal{I} \cup \{0\}$ **do**

  for $1 \le t \le T$ **do**

    Randomly select $\widehat{n}^t_{ijkl}$ from $[0, w^t_{ijk}]$ with $\sum_{l\in\mathcal{L}} \widehat{n}^t_{ijkl} = w^t_{ijk}$.

    Initialize $\widehat{n}^t_{ij0l} = 0$ for $j \in \mathcal{J}$, $l \in \mathcal{L}$.

    Randomly select NonPurchase>0

    for $1 \le r \le \text{NonPurchase}$ **do**

      Randomly select n from $1 \ldots |\mathcal{L}|$

      Randomly select m from $1 \ldots, |\mathcal{J}|$

      Draw $u \sim unif(0,1)$

      if $u < \gamma$ **then**

        Set $\widehat{n}^t_{im0n} += 1$

      end if

    end for

    Ensure $\sum_{j\in\mathcal{J}} \widehat{n}_{ij0l} > 0$ for each $l \in \mathcal{L}$.

  end for

end for

Set $\Omega^0 = (\boldsymbol{\Lambda}^0, \boldsymbol{\lambda}^0, \boldsymbol{\mu}^0, \boldsymbol{\theta}^0, \boldsymbol{v}^0)$ by Proposition 3.

where $\|\cdot\|$ is the Euclidean norm, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{c} \in \mathbb{R}^n$ and $d \in \mathbb{R}$ are parameters. Conic programming problems have been shown to be polynomial time solvable (see for example, primal-dual interior point methods). Furthermore, recent advances in solution methods and software packages have greatly enhanced the solvability of mixed-integer conic problems.

Our reformulation approach is closely related to a special case of second-order conic inequalities, which is the following *rotated cone/hyperbolic inequality*:

$$x_1 x_2 \geq x_3^2, \tag{EC.7}$$

where $x_1, x_2, x_3 \geq 0$. It can be shown that the rotated-cone constraint (EC.7) is equivalent to the following second-order conic inequality:

$$\|(x_1 - x_2, 2x_3)\| \leq x_1 + x_2.$$

To facilitate the discussion on strengthening cuts, we review the definition of polymatroid functions. Let $\vee$ and $\wedge$ denote the componentwise maximum and minimum of two vectors, respectively, that is, for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ $\boldsymbol{x} \vee \boldsymbol{y} = \inf\{\boldsymbol{z} \in \mathbb{R}^n | z_i \geq x_i, z_i \geq y_i, \forall i = \{1, 2, \ldots, n\}\}$, and $\boldsymbol{x} \wedge \boldsymbol{y} = \sup\{\boldsymbol{z} \in \mathbb{R}^n | z_i \leq x_i, z_i \leq y_i, \forall i = \{1, 2, \ldots, n\}\}$. A real-valued function f defined on $\{0,1\}^n$ is a polymatroid function if it is normalized ($f(\mathbf{0}) = 0$), non-decreasing ($f(\boldsymbol{y}_1) \leq f(\boldsymbol{y}_2)$ for $\boldsymbol{y}_1 \leq \boldsymbol{y}_2$ and $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \{0,1\}^n$), and submodular, that is

$$f(\boldsymbol{y}_1 \vee \boldsymbol{y}_2) + f(\boldsymbol{y}_1 \wedge \boldsymbol{y}_2) \leq f(\boldsymbol{y}_1) + f(\boldsymbol{y}_2), \ \forall \boldsymbol{y}_1, \boldsymbol{y}_2 \in \{0,1\}^n \ .$$

References

Atamtürk A, Gómez A (2020) Submodularity in conic quadratic mixed 0–1 optimization. *Oper. Res.* 68(2):609–630.