Online Supplement: Robust Vehicle Pre-Allocation with Uncertain Covariates

Appendix A: Linear Program Formulation of SAA

$$
\begin{aligned}
\Pi^{\mathrm{SAA}} = \max_{x_{ij}\geq 0, y_{jt}} & \left\{ -\sum_{j\in[M]}\sum_{i\in[N]} w_{ij}x_{ij} + \frac{1}{T}\sum_{j\in[M]}\sum_{t\in[T]} r_j y_{jt} \right\} \\
\text{s.t.} \quad & \sum_{j\in[M]} x_{ij} \leq S_i, \forall i\in[N] \\
& y_{jt} \leq \hat{z}_{jt}, \forall j\in[M], t\in[T] \\
& y_{jt} \leq \sum_{i\in[N]} x_{ij}, \forall j\in[M], t\in[T]
\end{aligned}
$$

Appendix B: Proofs

B.1. Proof of Proposition 1

It is sufficient to show that for all $\mathbb{Q}\in\bar{\mathbb{F}}$, $\mathbb{Q}\in\hat{\mathbb{F}}$ as well. If $\mathbb{Q}\in\bar{\mathbb{F}}$, we have

$$
\mathbb{E}_{\mathbb{Q}}(\tilde{\mathbf{z}}) = \sum_{l\in[L]} \mathbb{E}_{\mathbb{Q}}(\tilde{\mathbf{z}} \mid \tilde{\mathbf{v}}\in\Omega_l)p_l = \sum_{l\in[L]} \boldsymbol{\mu}_l p_l = \boldsymbol{\mu},
$$

and

$$
\begin{aligned}
\mathbb{E}_{\mathbb{Q}}(\tilde{z}_j-\mu_j)^2 = & \sum_{l\in[L]} \mathbb{E}_{\mathbb{Q}}((\tilde{z}_{jl}-\mu_j)^2 \mid \tilde{\mathbf{v}}\in\Omega_l)p_l \\
= & \sum_{l\in[L]} \mathbb{E}_{\mathbb{Q}}((\tilde{z}_{jl})^2 \mid \tilde{\mathbf{v}}\in\Omega_l)p_l - \mu_j^2 \\
\leq & \sum_{l\in[L]} (\sigma_{jl}^2+\mu_{jl}^2)p_l - \mu_j^2 \\
= & \ \sigma_j^2+\mu_j^2-\mu_j^2 \\
= & \ \sigma_j^2, \forall j\in[M].
\end{aligned}
$$

In addition, combining $\cup_{l\in[L]}\mathcal{Z}_l \subseteq \mathcal{Z}$ and $\mathbb{Q}(\tilde{\mathbf{z}}\in\mathcal{Z}_l \mid \tilde{\mathbf{v}}\in\Omega_l)=1$ for all $l\in[L]$, we have $\mathbb{Q}(\tilde{\mathbf{z}}\in\mathcal{Z})=1$. Hence, $\mathbb{Q}\in\hat{\mathbb{F}}$. $\square$

B.2. Proof of Lemma 1

First, we have $\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij} = \sum_{i\in[N]} x_{ij} - (\sum_{i\in[N]} x_{ij} - \tilde{z}_j)^+$. Therefore, with the ambiguity set $\bar{\mathbb{F}}$, the objective function in problem (4) can be rewritten as

$$
\sum_{j\in[M]}\sum_{i\in[N]} (r_j-w_{ij})x_{ij} - \sup_{\mathbb{Q}\in\bar{\mathbb{F}}} \mathbb{E}_{\mathbb{Q}}[\sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij}-\tilde{z}_j)^+] \tag{B.1}
$$

From Proposition 1 in Bertsimas et al. (2018), we know $\bar{\mathbb{F}} = \Pi_{(\tilde{\mathbf{z}},\tilde{\mathbf{v}})}\mathbb{G}$. Therefore, we have $\sup_{\mathbb{Q}\in\bar{\mathbb{F}}} \mathbb{E}_{\mathbb{Q}}[\sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij}-\tilde{z}_j)^+] = \sup_{\mathbb{Q}\in\mathbb{G}} \mathbb{E}_{\mathbb{Q}}[\sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij}-\tilde{z}_j)^+]$. For any $\mathbb{Q}\in\mathbb{G}$, let $\mathbb{Q}_l$ be the conditional probability distribution of $(\tilde{\mathbf{z}},\tilde{\mathbf{u}})$ given scenario $l\in[L]$. For given x_{ij}, by the law of total expectation, we have

$$
\mathbb{E}_{\mathbb{Q}}[\sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij}-\tilde{z}_j)^+] = \sum_{l\in[L]} p_l \mathbb{E}_{\mathbb{Q}_l}[\sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij}-\tilde{z}_{jl})^+]. \tag{B.2}
$$

Therefore, with the constraints specified in ambiguity set $\mathbb{G}$, $\sup\limits_{\mathbb{Q}\in\mathbb{G}} \mathbb{E}_{\mathbb{Q}}[\sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij} - \tilde{z}_j)^+]$ is equivalent to the following problem:

$$
\begin{aligned}
\sup_{\mathbb{Q}_1,\ldots,\mathbb{Q}_L} \quad & \sum_{l\in[L]} p_l \mathbb{E}_{\mathbb{Q}_l}[\sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij} - \tilde{z}_{jl})^+] && \text{(B.3)}\\
\text{s.t.} \quad & \mathbb{E}_{\mathbb{Q}_l}(\tilde{\mathbf{z}}) = \boldsymbol{\mu}_l, \forall l\in[L] && (\cdots \text{ dual variable } \boldsymbol{\delta}_l \in \mathbb{R}^M)\\
& \mathbb{E}_{\mathbb{Q}_l}(\tilde{u}_{jl}) \le \sigma_{jl}^2, \quad \forall j\in[M], l\in[L] && (\cdots \text{ dual variable } \gamma_{jl}\in\mathbb{R})\\
& \mathbb{Q}_l[(\tilde{\mathbf{z}},\tilde{\mathbf{u}})\in\mathcal{W}_l] = 1, \forall l\in[L] && (\cdots \text{ dual variable } \alpha_l\in\mathbb{R})
\end{aligned}
$$

With the dual variables defined above, by strong duality, problem (B.3) is equivalent to the following dual formulation:

$$
\begin{aligned}
\min_{\boldsymbol{\alpha},\boldsymbol{\delta}_l,\boldsymbol{\gamma}_l} \quad & \sum_{l\in[L]} \alpha_l + \sum_{l\in[L]} \left(\boldsymbol{\delta}_l^{'} \boldsymbol{\mu}_l + \sum_{j\in[M]} \gamma_{jl}\sigma_{jl}^2 \right)\\
\text{s.t.} \quad & \alpha_l + \boldsymbol{\delta}_l^{'}\mathbf{z} + \boldsymbol{\gamma}_l^{'}\mathbf{u} \ge p_l \sum_{j\in[M]} r_j(\sum_{i\in[N]} x_{ij} - z_{jl})^+, \forall (\mathbf{z},\mathbf{u})\in\mathcal{W}_l, l\in[L]\\
& \boldsymbol{\alpha}\in\mathbb{R}^L, \boldsymbol{\delta}_l\in\mathbb{R}^M, \boldsymbol{\gamma}_l\in\mathbb{R}_+^M, \forall l\in[L]
\end{aligned}
$$

Therefore, by replacing the second term in (B.1), we can reformulate problem (4) into problem (6). □

B.3. Proof of Proposition 2

First, note that any feasible solution, i.e., x_{ij}, $\boldsymbol{\alpha}, \boldsymbol{\delta}_l, \boldsymbol{\gamma}_l$, $\forall l\in[L]$, and $y_j^l(\mathbf{z},\mathbf{u})$, $\forall j\in[M], l\in[L]$ to problem (7), x_{ij}, $\boldsymbol{\alpha}, \boldsymbol{\delta}_l, \boldsymbol{\gamma}_l$ are also feasible to problem (6). Therefore, we have $\hat{\Pi} \le \Pi^{\text{SDR}}$.

Next, we prove $\hat{\Pi} \ge \Pi^{\text{SDR}}$. Let $\hat{x}_{ij}$, $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\delta}}_l, \hat{\boldsymbol{\gamma}}_l, \forall l\in[L]$ be a feasible solution to problem (6). By the definition of $\mathcal{W}_l$, the constraint (6b) is equivalent to

$$
\hat{\alpha}_l \ge \max_{(\mathbf{z},\mathbf{u})\in\mathcal{W}_l} \left(-\hat{\boldsymbol{\delta}}_l^{'}\mathbf{z} - \hat{\boldsymbol{\gamma}}_l^{'}\mathbf{u} + p_l \sum_{j\in[M]} r_j(\sum_{i\in[N]} \hat{x}_{ij} - z_{jl})^+ \right) = \sum_{j\in[M]} M_{jl}, \forall l\in[L],
$$

where $M_{jl} = \max\limits_{(z_{jl},u_{jl})\in\mathcal{W}_{jl}} \{-\hat{\delta}_{jl}z_{jl} - \hat{\gamma}_{jl}u_{jl} + p_l r_j(\sum_{i\in[N]} \hat{x}_{ij} - z_{jl})^+\}$, and $\mathcal{W}_{jl} = \{(z_{jl},u_{jl})\in\mathbb{R}\times\mathbb{R} \mid z_{jl}\in\mathcal{Z}_{jl}, (z_{jl}-\mu_{jl})^2 \le u_{jl}\}$.

We can construct a feasible solution to (7) by letting $x_{ij} = \hat{x}_{ij}$, $\boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}, \boldsymbol{\delta}_l = \hat{\boldsymbol{\delta}}_l, \boldsymbol{\gamma}_l = \hat{\boldsymbol{\gamma}}_l, \forall l\in[L]$, and

$$
\begin{cases}
y_j^{0l} = \frac{M_{jl}}{r_j p_l}, & \forall j\in[M], l\in[L]\\
y_{jj}^{1l} = \frac{\hat{\delta}_{jl}}{r_j p_l}, & \forall j\in[M], l\in[L]\\
y_{jj}^{2l} = \frac{\hat{\gamma}_{jl}}{r_j p_l}, & \forall j\in[M], l\in[L]\\
y_{jk}^{1l} = 0, & \forall j,k\in[M], k\ne j, l\in[L]\\
y_{jk}^{2l} = 0, & \forall j,k\in[M], k\ne j, l\in[L]
\end{cases}
$$

Its feasibility is shown as follows. For any $l\in[L]$, $j\in[M]$

$$
\begin{aligned}
& y_j^l(\mathbf{z},\mathbf{u}) - (\textstyle\sum_{i\in[N]} x_{ij} - z_{jl})^+\\
= & \tfrac{1}{r_j p_l} \left(M_{jl} + \hat{\delta}_{jl} z_{jl} + \hat{\gamma}_{jl} u_{jl} - r_j p_l (\textstyle\sum_{i\in[N]} \hat{x}_{ij} - z_{jl})^+ \right)\\
= & \tfrac{1}{r_j p_l} \left(M_{jl} - (-\hat{\delta}_{jl} z_{jl} - \hat{\gamma}_{jl} u_{jl} + r_j p_l (\textstyle\sum_{i\in[N]} \hat{x}_{ij} - z_{jl})^+) \right)\\
\ge & \ 0, \forall (\mathbf{z},\mathbf{u})\in\mathcal{W}_l,
\end{aligned}
$$

where the last inequality is by the fact that M_{jl} is defined as the maximum of $-\hat{\delta}_{jl}z_{jl} - \hat{\gamma}_{jl}u_{jl} + r_j p_l(\sum_{i\in[N]}\hat{x}_{ij} - z_{jl})^+$ over $\mathcal{W}_{jl}$.

In addition,

$$\begin{aligned}
&\alpha_l + \left(\boldsymbol{\delta}_l'\mathbf{z} + \boldsymbol{\gamma}_l'\mathbf{u}\right) - p_l\textstyle\sum_{j\in[M]} r_j y_j^l(\mathbf{z},\mathbf{u})\\
=\ &\hat{\alpha}_l + \left(\hat{\boldsymbol{\delta}}_l'\mathbf{z} + \hat{\boldsymbol{\gamma}}_l'\mathbf{u}\right) - p_l\textstyle\sum_{j\in[M]} r_j y_j^l(\mathbf{z},\mathbf{u})\\
=\ &\hat{\alpha}_l + \left(\hat{\boldsymbol{\delta}}_l'\mathbf{z} + \hat{\boldsymbol{\gamma}}_l'\mathbf{u}\right) - (\textstyle\sum_{j\in[M]}(M_{jl} + \hat{\delta}_{jl}z_{jl} + \hat{\gamma}_{jl}u_{jl}))\\
=\ &\hat{\alpha}_l - \textstyle\sum_{j\in[M]} M_{jl} \geq 0, \forall(\mathbf{z},\mathbf{u})\in\mathcal{W}_l.
\end{aligned}$$

Therefore, we obtain $\hat{\Pi} \geq \Pi^{\mathrm{SDR}}$. Together with $\hat{\Pi} \leq \Pi^{\mathrm{SDR}}$, we complete our proof. $\square$

Appendix C: Formulation of Robust Vehicle Allocation Problem with Covariance information

To include the covariance matrix, denoted by $\mathbf{\Sigma}$, we construct the following ambiguity set:

$$\mathbb{F}_{\mathbf{\Sigma}} = \left\{ \mathbb{Q}\in\mathcal{P}_0(\mathbb{R}^M\times\mathbb{R}^I) \,\middle|\, \begin{array}{l} (\tilde{\mathbf{z}},\tilde{\mathbf{v}})\sim\mathbb{Q}\\ \mathbb{E}_{\mathbb{Q}}(\tilde{\mathbf{z}}) = \boldsymbol{\mu}\\ \mathbb{E}_{\mathbb{Q}}(\tilde{\mathbf{z}}-\boldsymbol{\mu})(\tilde{\mathbf{z}}-\boldsymbol{\mu})' \preceq \mathbf{\Sigma}\\ \mathbb{Q}(\tilde{\mathbf{z}}\in\mathcal{Z}) = 1 \end{array} \right\},$$

where constraint $\mathbb{E}_{\mathbb{Q}}(\tilde{\mathbf{z}}-\boldsymbol{\mu})(\tilde{\mathbf{z}}-\boldsymbol{\mu})' \preceq \mathbf{\Sigma}$ forces the second moment matrix of $\tilde{\mathbf{z}}$ to lie in a positive semi-definite cone defined with a matrix inequality (Delage and Ye 2010). Following Wiesemann et al. (2014), we further introduce the following lifted ambiguity set:

$$\mathbb{G}_{\mathbf{\Sigma}} = \left\{ \mathbb{Q}\in\mathcal{P}_0(\mathbb{R}^M\times\mathbb{R}^I) \,\middle|\, \begin{array}{l} (\tilde{\mathbf{z}},\tilde{\mathbf{v}})\sim\mathbb{Q}\\ \mathbb{E}_{\mathbb{Q}}(\tilde{\mathbf{z}}) = \boldsymbol{\mu}\\ \mathbb{E}_{\mathbb{Q}}(\boldsymbol{U}) \preceq \mathbf{\Sigma}\\ \mathbb{Q}(\tilde{\mathbf{z}}\in\mathcal{Z}, (\tilde{\mathbf{z}}-\boldsymbol{\mu})(\tilde{\mathbf{z}}-\boldsymbol{\mu})' \preceq \boldsymbol{U}) = 1 \end{array} \right\},$$

where $\mathcal{Z} = [\mathbf{0}, \bar{\mathbf{z}}]$. We have the following result.

Proposition 1. *The distributionally robust vehicle allocation model with $\mathbb{F}_{\mathbf{\Sigma}}$ is equivalent to the following semi-definite program:*

$$\begin{aligned}
\Pi_{\Sigma} = \max\ & -\sum_{j\in[M]}\sum_{i\in[N]} w_{ij}x_{ij} - \alpha - \boldsymbol{\phi}'\boldsymbol{\mu} - \langle \boldsymbol{Q}, \mathbf{\Sigma}\rangle \qquad\qquad \text{(C.1)}\\
\text{s.t.}\ & \alpha + \boldsymbol{\beta}(\mathcal{M})'\boldsymbol{x} - \boldsymbol{y}(\mathcal{M})'\bar{\mathbf{z}} - t(\mathcal{M}) + 2\boldsymbol{g}(\mathcal{M})'\boldsymbol{\mu} \geq 0, \forall\mathcal{M}\in\mathcal{P}(M)\\
& \boldsymbol{\phi}' + \boldsymbol{\gamma}(\mathcal{M})' + \boldsymbol{y}(\mathcal{M})' - 2\boldsymbol{g}(\mathcal{M})' \geq \mathbf{0}', \forall\mathcal{M}\in\mathcal{P}(M)\\
& \boldsymbol{Q} = \mathbf{\Gamma}(\mathcal{M}), \forall\mathcal{M}\in\mathcal{P}(M)\\
& \boldsymbol{y}(\mathcal{M}) \geq \mathbf{0}, \forall\mathcal{M}\in\mathcal{P}(M)\\
& \begin{pmatrix} \mathbf{\Gamma}(\mathcal{M}) & \boldsymbol{g}(\mathcal{M})\\ \boldsymbol{g}(\mathcal{M})' & t(\mathcal{M}) \end{pmatrix} \succeq \mathbf{0}, \forall\mathcal{M}\in\mathcal{P}(M)\\
& \boldsymbol{Q}\in\mathbb{S}_+^M, \boldsymbol{y}(\mathcal{M})\in\mathbb{R}_+^M, \boldsymbol{g}(\mathcal{M})\in\mathbb{R}^M, \boldsymbol{\phi}\in\mathbb{R}^M\\
& t(\mathcal{M})\in\mathbb{R}, \mathbf{\Gamma}(\mathcal{M})\in\mathbb{S}_+^M, x_{ij}\in\mathbb{R}_+, \alpha\in\mathbb{R}\\
& \sum_{j\in[M]} x_{ij} \leq S_i, \forall i\in[N].
\end{aligned}$$

where $\boldsymbol{x} = (\sum_{i\in[N]}x_{i1},\ldots,\sum_{i\in[N]}x_{iM})'$. $\beta(\mathcal{M}) = (r_1',\ldots,r_M')'$, where $r_j' = r_j$ if $j\in\mathcal{M}$; otherwise $r_j' = 0$. $\gamma(\mathcal{M}) = (\hat{r}_1,\ldots,\hat{r}_M)'$, where $\hat{r}_j = 0$ if $j\in\mathcal{M}$; otherwise $\hat{r}_j = r_j$. $\mathbb{S}_+^M$ is the set of positive semidefinite matrices.

Proof. With the ambiguity set $\mathbb{F}_{\boldsymbol{\Sigma}}$, we aim to solve the following optimization problem:

$$\begin{aligned} \max_{x_{ij}\geq 0} & - \sum_{j\in[M]}\sum_{i\in[N]} w_{ij}x_{ij} + \inf_{\mathbb{Q}\in\mathbb{F}_{\boldsymbol{\Sigma}}} \mathbb{E}_{\mathbb{Q}}\left[\sum_{j\in[M]} r_j(\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij})\right] \\ \text{s.t.} & \quad \sum_{j\in[M]} x_{ij} \leq S_i, \forall i\in[N]. \end{aligned} \tag{C.2}$$

To solve (C.2), the key is to solve $\inf_{\mathbb{Q}\in\mathbb{F}_{\boldsymbol{\Sigma}}} \mathbb{E}_{\mathbb{Q}}\left[\sum_{j\in[M]} r_j(\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij})\right]$. From Wiesemann et al. (2014), we have $\inf_{\mathbb{Q}\in\mathbb{F}_{\boldsymbol{\Sigma}}} \mathbb{E}_{\mathbb{Q}}\left[\sum_{j\in[M]} r_j(\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij})\right] = \inf_{\mathbb{Q}\in\mathbb{G}_{\boldsymbol{\Sigma}}} \mathbb{E}_{\mathbb{Q}}\left[\sum_{j\in[M]} r_j(\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij})\right] =$
$-\sup_{\mathbb{Q}\in\mathbb{G}_{\boldsymbol{\Sigma}}} \mathbb{E}_{\mathbb{Q}}\left[-\sum_{j\in[M]} r_j(\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij})\right]$. With the constraints specified in ambiguity set $\mathbb{G}_{\boldsymbol{\Sigma}}$, $\sup_{\mathbb{Q}\in\mathbb{G}_{\boldsymbol{\Sigma}}} \mathbb{E}_{\mathbb{Q}}\left[-\sum_{j\in[M]} r_j(\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij})\right]$ is equivalent to the following problem:

$$\begin{aligned} \sup \quad & \mathbb{E}_{\mathbb{Q}}\left[-\sum_{j\in[M]} r_j(\tilde{z}_j \wedge \sum_{i\in[N]} x_{ij})\right] && \text{(C.3)} \\ \text{s.t.} \quad & \mathbb{E}_{\mathbb{Q}}(\tilde{\mathbf{z}}) = \boldsymbol{\mu}, && (\cdots \text{ dual variable } \boldsymbol{\phi}\in\mathbb{R}^M) \\ & \mathbb{E}_{\mathbb{Q}}(\boldsymbol{U}) \preceq \boldsymbol{\Sigma}, && (\cdots \text{ dual variable } \boldsymbol{Q}\in\mathbb{S}_+^M) \\ & \mathbb{Q}(\tilde{\mathbf{z}}\in\mathcal{Z}, (\tilde{\mathbf{z}}-\boldsymbol{\mu})(\tilde{\mathbf{z}}-\boldsymbol{\mu})' \preceq \boldsymbol{U}) = 1, && (\cdots \text{ dual variable } \alpha\in\mathbb{R}) \end{aligned}$$

With the dual variables defined above, by strong duality, problem (C.3) is equivalent to the following dual formulation

$$\begin{aligned} \min_{\alpha,\boldsymbol{\phi},\boldsymbol{Q}} \quad & \alpha + \boldsymbol{\phi}'\boldsymbol{\mu} + \langle \boldsymbol{Q}, \boldsymbol{\Sigma}\rangle \\ \text{s.t.} \quad & \alpha + \boldsymbol{\phi}'\mathbf{z} + \langle \boldsymbol{Q}, \boldsymbol{U}\rangle \geq -\sum_{j\in[M]} r_j(z_j \wedge \sum_{i\in[N]} x_{ij}), \forall \mathbf{z}\in\mathcal{Z}, (\mathbf{z}-\boldsymbol{\mu})(\mathbf{z}-\boldsymbol{\mu})' \preceq \boldsymbol{U} \\ & \alpha\in\mathbb{R}, \boldsymbol{\phi}\in\mathbb{R}^M, \boldsymbol{Q}\in\mathbb{S}_+^M, \end{aligned} \tag{C.4}$$

where $\langle \boldsymbol{Q}, \boldsymbol{U}\rangle = trace(\boldsymbol{Q}^T\boldsymbol{U})$. Let $\boldsymbol{x} = (\sum_{i\in[N]} x_{i1}, ..., \sum_{i\in[N]} x_{iM})'$. $\beta(\mathcal{M}) = (r'_1, ..., r'_M)'$, where $r'_j = r_j$ if $j\in\mathcal{M}$; otherwise $r'_j = 0$. $\gamma(\mathcal{M}) = (\hat{r}_1, ..., \hat{r}_M)'$, where $\hat{r}_j = 0$ if $j\in\mathcal{M}$; otherwise $\hat{r}_j = r_j$. For given $\mathcal{M}$, the constraint $\alpha + \boldsymbol{\phi}'\mathbf{z} + \langle \boldsymbol{Q}, \boldsymbol{U}\rangle \geq -\sum_{j\in[M]} r_j(z_j \wedge \sum_{i\in[N]} x_{ij}), \forall \mathbf{z}\in\mathcal{Z}, (\mathbf{z}-\boldsymbol{\mu})(\mathbf{z}-\boldsymbol{\mu})' \preceq \boldsymbol{U}$ is equivalent to the following problem:

$$\begin{aligned} \min \quad & \alpha + (\boldsymbol{\phi}' + \boldsymbol{\gamma}(\mathcal{M})')\mathbf{z} + \langle \boldsymbol{Q}, \boldsymbol{U}\rangle + \boldsymbol{\beta}(\mathcal{M})'\boldsymbol{x} \geq 0 \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{U} & \mathbf{z}-\mu \\ (\mathbf{z}-\mu)' & 1 \end{pmatrix} \succeq \mathbf{0} \\ & \mathbf{z} \geq \mathbf{0} \\ & \mathbf{z} \leq \bar{\mathbf{z}} \end{aligned}$$

The dual formulation of the above problem is

$$\begin{aligned} \max \quad & \alpha + \boldsymbol{\beta}(\mathcal{M})'\boldsymbol{x} - \boldsymbol{y}(\mathcal{M})'\bar{\mathbf{z}} - t(\mathcal{M}) + 2\boldsymbol{g}(\mathcal{M})'\boldsymbol{\mu} \geq 0 \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{\Gamma}(\mathcal{M}) & \boldsymbol{g}(\mathcal{M}) \\ \boldsymbol{g}(\mathcal{M})' & t(\mathcal{M}) \end{pmatrix} \succeq \mathbf{0} \\ & \boldsymbol{Q} = \boldsymbol{\Gamma}(\mathcal{M}) \\ & \boldsymbol{\phi}' + \boldsymbol{\gamma}(\mathcal{M})' + \boldsymbol{y}(\mathcal{M})' - 2\boldsymbol{g}(\mathcal{M})' \geq \mathbf{0}' \\ & \boldsymbol{y}(\mathcal{M}) \geq \mathbf{0} \end{aligned}$$

Substituting this dual formulation and then performing the outer and inner maximization jointly, we obtain the formulation (C.1). □

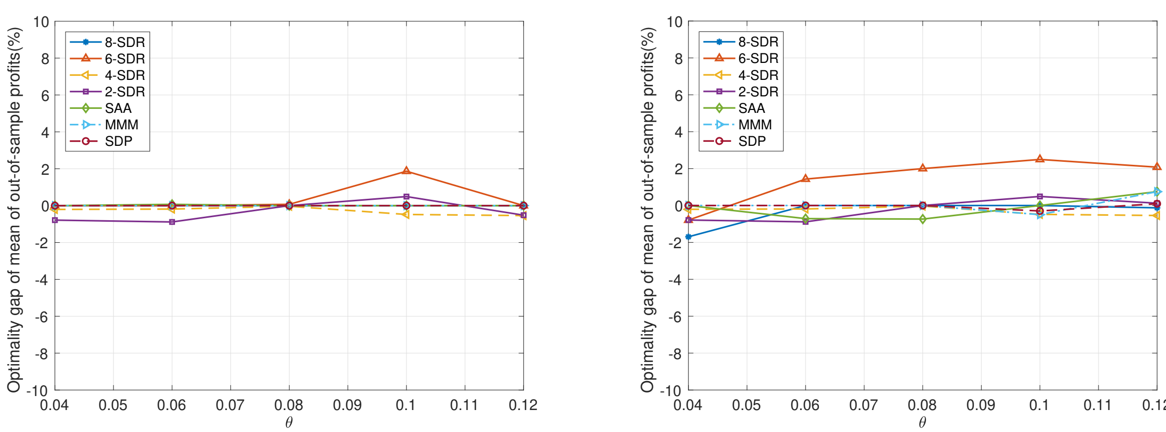
Appendix D: Performance Comparison between the Integer Solution and the Heuristic Rounding Solution of Case Study

We compute the optimality gap between the mean of out-of-sample profits with different solutions for all the models to examine the performance of the heuristic rounding solution compared to the optimal integer solution. Specifically, let π_{int}, π_{con} and π_{rou} denote the mean of out-of-sample profits with optimal integer solution, continuous solution and heuristic rounding solution, respectively. Then the optimality gap of the mean of out-of-sample profits in Figure 1(a) is defined as $\frac{\pi_{\text{con}}-\pi_{\text{int}}}{\pi_{\text{int}}}$, and the one in Figure 1(b) is defined as $\frac{\pi_{\text{rou}}-\pi_{\text{int}}}{\pi_{\text{int}}}$. The smaller the gap is, the closer the two solutions perform.

From Figure 1(a), we can observe that for all the models with all θ, the max optimality gap between the mean of out-of-sample profits of the optimal integer solution and the heuristic rounding solution is 2% and the average gap is about 0.03%.

From Figure 1(b), we can observe that for all the models with all θ, the max optimality gap between the mean of out-of-sample profits with the optimal integer solution and the one with the heuristic rounding solution is 2.5% and the average gap is about 0.06%.

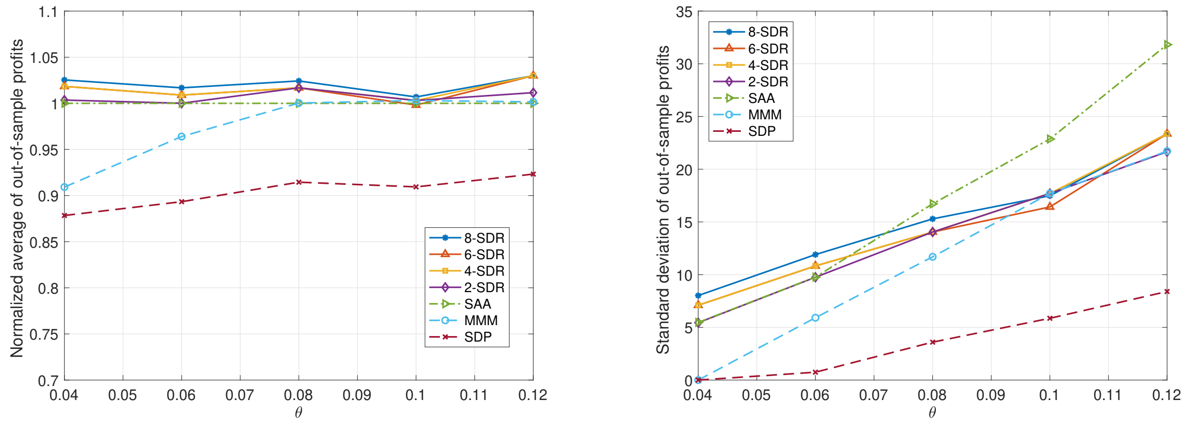
In summary, the results show that the optimality gap between the performance of the optimal integer solution and the heuristic rounding solution is negligible. Therefore, in practice, to obtain an integer solution, one can directly round the continuous solutions instead of solving the model with integer constraints that result in much more computational time.



(a) Integer solution v.s. Continuous solution

(b) Integer solution v.s. Rounded continuous solution

Figure 1 Out-of-sample performance difference of the SDR, SAA, SDP and MMM models for case study



(a) Mean of out-of-sample profits (b) Standard deviation of out-of-sample profits

Figure 2 Out-of-sample performance of the four models with the rounded continuous solution for case study

Appendix E: Data and Computational Results of the Numerical Study

For exposition simplicity, let Z1=Alexandra Hill, Z2=Bendemeer, Z3=Cairnhill, Z4=Chatsworth, Z5=lavender, Z6=Leonie Hill, Z7=Marina Centre, Z8=Nassim, Z9=Robertson Quay, Z10=Tiong Bahru Station.

E.1. Correlation Coefficient Matrix

The covariance matrix of the in-sample simulated data in the numerical simulation is

$$\mathbf{\Sigma} = \begin{pmatrix} 4238 & 730 & 669 & 596 & 708 \\ 730 & 2590 & -405 & -368 & -191 \\ 669 & -405 & 1682 & 1110 & 936 \\ 596 & -368 & 1110 & 4153 & 336 \\ 708 & -191 & 936 & 336 & 2404 \end{pmatrix}$$

The corresponding correlation coefficient matrix is

$$\mathbf{\Sigma}_{\text{corr}} = \begin{pmatrix} 1 & -0.2203 & 0.2506 & 0.1422 & 0.2219 \\ -0.2203 & 1 & -0.194 & -0.1123 & -0.0765 \\ 0.2506 & -0.194 & 1 & 0.42 & 0.4656 \\ 0.1422 & -0.1123 & 0.42 & 1 & 0.1065 \\ 0.2219 & -0.0765 & 0.4656 & 0.1065 & 1 \end{pmatrix}$$

E.2. Performance Results of the Numerical Simulation

Table 1 lists the mean and the standard deviation (abbreviated as std) of out-of-sample profits of 24 instances of the numerical experiments.

q	Δ	θ	S	SAA		SDR		MMM		SDP	
				mean	std	mean	std	mean	std	mean	std
0.1	-0.2			28.97	130.00	39.29	120.68	-13.98	162.25	-13.98	162.25
0.1	0	0.05	400	161.90	34.56	165.29	28.60	138.75	61.97	138.75	61.97
0.1	0.2			194.60	6.76	192.04	5.29	202.21	12.17	202.21	12.17
0.1	-0.2			115.90	224.65	116.26	224.86	116.07	224.79	116.07	224.79
0.1	0	0.1	400	311.18	112.40	311.12	111.98	311.06	112.21	311.06	112.21
0.1	0.2			421.63	27.89	421.10	26.89	422.13	26.70	422.13	26.70
0.2	-0.2			49.77	98.26	53.13	96.13	23.21	115.96	23.21	115.96
0.2	0	0.05	400	138.08	44.70	138.73	43.40	131.91	56.94	131.91	56.94
0.2	0.2			167.90	22.79	168.41	22.56	171.93	27.20	171.93	27.20
0.2	-0.2			132.61	178.36	131.42	178.24	88.30	210.56	88.30	210.56
0.2	0	0.1	400	301.01	101.37	300.16	101.85	275.65	130.72	275.65	130.72
0.2	0.2			372.07	50.60	372.06	50.99	375.60	67.77	375.60	67.77
0.1	-0.2			54.02	9.60	59.08	0	56.95	5.18	58.70	1.69
0.1	0	0.05	100	59.31	0	59.08	0	59.23	0	59.15	0
0.1	0.2			59.31	0	59.08	0	59.23	0	59.15	0
0.1	-0.2			112.45	11.20	116.50	4.72	113.57	9.66	113.62	9.60
0.1	0	0.1	100	118.62	0	118.41	0	118.57	0	118.57	0
0.1	0.2			118.62	0	118.41	0	118.57	0	118.57	0
0.2	-0.2			56.42	7.12	55.61	3.74	54.73	8.49	54.74	8.47
0.2	0	0.05	100	59.02	0	57.47	0	58.24	0	58.24	0
0.2	0.2			59.02	0	57.47	0	58.24	0	58.24	0
0.2	-0.2			115.02	8.28	115.41	0	109.75	15.45	109.75	15.42
0.2	0	0.1	100	118.04	0	115.41	0	116.83	0	116.83	0
0.2	0.2			118.04	0	115.41	0	116.83	0	116.83	0

Table 1 **Performance results of the SDR, SDP, MMM and SAA models in simulation experiments**

E.3. Performance Results of Case Study

Table 2 lists the mean and std of out-of-sample profits of the SDR models with different scenarios, SAA, SDP and MMM of the case study.

θ	8-SDR		6-SDR		4-SDR		2-SDR		SAA		MMM		SDP	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
0.04	42.07	8.17	41.80	7.29	41.82	7.36	41.51	6.40	41.24	6.21	37.72	0.00	36.07	0.00
0.06	65.69	12.19	65.28	11.21	65.37	11.38	64.81	10.32	63.72	11.75	62.46	6.27	57.73	0.87
0.08	89.80	15.57	89.23	14.47	89.35	14.69	88.76	13.58	86.33	19.80	87.90	12.15	80.08	3.62
0.10	114.37	17.87	113.53	17.50	113.56	17.72	113.12	16.63	111.73	25.70	113.57	17.77	103.4	6.22
0.12	141.58	25.46	140.65	25.15	140.85	25.27	139.10	19.54	135.33	35.61	137.77	23.21	127.51	8.95

Table 2 **Performance results of SDR, SDP, MMM and SAA models in case study**

E.4. Parameters Estimation Results of Case Study

Table 3 lists the estimation results of the average per order revenue $\hat{\mathbf{r}}$, and the booking fee $\hat{\mathbf{b}}$.

Economic parameter	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10
$\hat{\mathbf{r}}$	15.0	14.1	16.1	14.9	13.8	15.8	15.3	16.4	15.8	13.2
$\hat{\mathbf{b}}$	3	3	3	3	3	3	3	3	3	3

Table 3 Estimated compensation fee and average per order revenue in the case study

Tables 4-8 list the estimation results of the scenario probability, mean and standard deviation of demands at 10 demand regions for the training data.

L	p	Z1		Z2		Z3		Z4		Z5		Z6		Z7		Z8		Z9		Z10	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	1	5.7	7.0	7.7	9.4	6.2	9.1	9.3	12.2	3.0	4.4	86.0	30.5	8.4	7.9	11.0	14.1	9.8	10.5	7.0	5.6

Table 4 Estimation results of the MMM case

L	p	Z1		Z2		Z3		Z4		Z5		Z6		Z7		Z8		Z9		Z10	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	$\frac{64}{77}$	4.4	5.1	5.0	4.9	3.6	3.2	6.3	5.0	2.3	3.2	78.0	11.4	7.1	7.0	7.4	6.8	7.8	6.8	6.1	4.3
2	$\frac{13}{77}$	12.2	10.7	21.1	14.1	18.8	16.3	24.0	22.8	6.2	7.6	125.7	56.3	14.5	9.4	28.9	24.3	19.8	17.9	11.7	8.7

Table 5 Estimation results of the 2-SDR case

L	p	Z1		Z2		Z3		Z4		Z5		Z6		Z7		Z8		Z9		Z10	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	$\frac{67}{77}$	4.3	5.1	5.6	5.6	4.0	3.8	6.5	4.9	2.4	3.2	79.1	11.3	7.4	6.9	7.7	7.0	8.1	6.6	5.9	4.3
2	$\frac{3}{77}$	4.7	4.1	4.1	7.1	4.8	6.3	4.4	7.7	0.2	0.3	60.2	15.5	4.9	2.6	6.7	4.1	2.6	4.5	6.4	3.3
3	$\frac{4}{77}$	13.9	8.4	24.6	7.3	16.2	4.1	21.9	8.6	5.2	3.6	128.9	14.6	18.8	12.3	29.7	6.3	16.2	2.5	13.0	2.6
4	$\frac{3}{77}$	26.4	3.5	36.1	15.8	42.6	16.4	59.9	14.0	16.0	10.5	209.5	39.3	19.3	9.9	64.7	20.5	47.4	17.5	24.0	4.9

Table 6 Estimation results of the 4-SDR case

L	p	Z1		Z2		Z3		Z4		Z5		Z6		Z7		Z8		Z9		Z10	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	$\frac{4}{77}$	4.0	5.5	14.0	9.8	9.9	6.9	8.0	3.1	2.8	2.1	90.5	8.7	10.6	5.1	11.7	8.8	10.7	2.1	4.0	3.9
2	$\frac{63}{77}$	4.3	5.1	5.1	4.9	3.6	3.2	6.4	5.0	2.4	3.2	78.4	11.1	7.2	7.0	7.4	6.8	7.9	6.8	6.0	4.3
3	$\frac{3}{77}$	4.7	4.1	4.1	7.1	4.8	6.3	4.4	7.7	0.2	0.3	60.2	15.5	4.9	2.6	6.7	4.1	2.6	4.5	6.4	3.3
4	$\frac{2}{77}$	9.6	0.9	22.6	8.1	16.2	3.6	20.7	6.8	3.9	5.5	116.3	0.2	13.2	8.2	27.2	0.4	16.2	2.3	12.0	3.6
5	$\frac{2}{77}$	18.2	11.7	26.5	9.0	16.2	6.1	23.1	13.0	6.4	1.7	141.5	0.1	24.5	16.0	32.2	9.7	16.2	3.6	14.0	1.6
6	$\frac{3}{77}$	26.4	3.5	36.1	15.8	42.6	16.4	59.9	14.0	16.0	10.5	209.5	39.3	19.3	9.9	64.7	20.5	47.4	17.5	24.0	4.9

Table 7 Estimation results of the 6-SDR case

L	p	Z1		Z2		Z3		Z4		Z5		Z6		Z7		Z8		Z9		Z10	
		mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	$\frac{2}{77}$	0.9	1.3	3.2	4.5	0.0	0.0	3.1	4.4	0.0	0.0	61.2	21.4	2.5	1.6	0.4	0.6	7.4	10.5	4.1	2.1
2	$\frac{2}{77}$	2.2	2.7	19.7	10.9	14.3	3.5	5.6	1.6	2.2	1.6	97.2	6.4	9.4	5.3	14.3	14.1	10.0	3.4	2.6	1.6
3	$\frac{61}{77}$	4.5	5.2	5.2	4.9	3.8	3.2	6.5	5.0	2.5	3.2	78.9	10.4	7.3	7.1	7.6	6.8	7.9	6.7	6.1	4.4
4	$\frac{3}{77}$	4.7	4.1	4.1	7.1	4.8	6.3	4.4	7.7	0.2	0.3	60.2	15.5	4.9	2.6	6.7	4.1	2.6	4.5	6.4	3.3
5	$\frac{2}{77}$	5.9	8.3	8.3	6.0	5.4	7.1	10.3	2.0	3.3	3.0	83.7	0.6	11.9	6.6	9.2	2.4	11.5	0.5	5.5	5.9
6	$\frac{2}{77}$	9.6	0.9	22.6	8.1	16.2	3.6	20.7	6.8	3.9	5.5	116.3	0.2	13.2	8.2	27.2	0.4	16.2	2.3	12.0	3.6
7	$\frac{2}{77}$	18.2	11.7	26.5	9.0	16.2	6.1	23.1	13.0	6.4	1.7	141.5	0.1	24.5	16.0	32.2	9.7	16.2	3.6	14.0	1.6
8	$\frac{3}{77}$	26.4	3.5	36.1	15.8	42.6	16.4	59.9	14.0	16.0	10.5	209.5	39.3	19.3	9.9	64.7	20.5	47.4	17.5	24.0	4.9

Table 8 Estimation results of the 8-SDR case

References

Bertsimas, D., Sim, M., and Zhang, M. (2018). Adaptive distributionally robust optimization. *Management Science*, 65(2):604–618.

Delage, E. and Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612.

Wiesemann, W., Kuhn, D., and Sim, M. (2014). Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376.