Submitted to

manuscript

Robust Workforce Management with Crowdsourced Delivery

Chun Cheng
School of Economics and Management, Dalian University of Technology, Dalian, 116024, China
chun.cheng@polymtl.ca

Melvyn Sim
Department of Analytics & Operations, NUS Business School, National University of Singapore
dscsimm@nus.edu.sg

Yue Zhao
Institute of Operations Research and Analytics, National University of Singapore
yuezhao@u.nus.edu

We investigate how crowdsourced delivery platforms with both contracted and ad-hoc couriers can effectively manage their workforce to meet delivery demands amidst uncertainties. Our objective is to minimize the hiring costs of contracted couriers and the crowdsourcing costs of ad-hoc couriers while considering the uncertain availability and behavior of the latter. Due to the complication of calibrating these uncertainties through data-driven approaches, we instead introduce a basic reduced information model to estimate the upper bound of the crowdsourcing cost and a generalized reduced information model to obtain a tighter bound. Subsequently, we formulate a robust satisficing model associated with the generalized reduced information model and show that a binary search algorithm can tackle the model exactly by solving a modest number of convex optimization problems. Our numerical tests using Solomon's data sets show that reduced information models provide decent approximations for practical delivery scenarios. Simulation tests further demonstrate that the robust satisficing model has better out-of-sample performance than the empirical optimization model that minimizes the total cost under historical scenarios.

Key words: Workforce management, crowdsourced delivery, uncertain ad-hoc couriers, data-driven robust satisficing

1. Introduction

The rapid expansion of e-commerce has stimulated e-retailers and local businesses to enhance their logistics operations and deliver goods promptly and reliably in a cost-effective way. As a result, many have turned to crowdsourced delivery, which involves independent individuals using their own vehicles to deliver goods. For instance, Amazon launched the Amazon Flex program in 2015, which utilizes crowdsourced couriers for last-mile deliveries. Similarly, Walmart piloted crowdsourced delivery in two US cities in 2018 under the name Spark Delivery. Many third-party logistics companies that rely on crowdsourced delivery have also emerged, such as Postmates, Uber

Eats, and Deliv. These platforms vary in terms of payment methods, target markets, and the types of items delivered. Leveraging crowdsourced delivery resources allows platforms to swiftly adjust their delivery capacity to cope with fluctuating demand while also providing cost advantages.

The use of independent couriers for crowdsourced delivery poses challenges due to their unknown availability and job bidding behavior. To mitigate the adverse effects caused by this type of uncertainty, many platforms have opted for a hybrid workforce model that combines ad-hoc crowdsourced couriers with pre-hired couriers, such as employees or crowdsourced couriers who agree to work for a certain period with a guaranteed minimum payment (Yildiz and Savelsbergh 2019, Ulmer and Savelsbergh 2020). In the Amazon Flex program, for instance, deliveries are organized into blocks lasting from 2 to 6 hours, and crowdsourced couriers search for available delivery blocks through the platform's app, make requests for interested offers, and receive confirmation from Amazon. When couriers' block times are approaching, they go to the pick-up sites, load packages, and deliver them to customers. A hybrid model enables platforms to manage their workforce more effectively to provide reliable customer service.

Nevertheless, the employment decision of pre-hired couriers presents a challenge as future customer orders remain uncertain to the platform. Over-hiring couriers incur unnecessary costs for the platform, while under-hiring may require the hiring of ad-hoc couriers at premium prices to ensure timely delivery.

In this study, we present a robust satisficing framework for addressing delivery platforms' workforce management problem, with the aim of maximizing the robustness of achieving cost targets under uncertainty. Our approach distinguishes between two types of couriers: *contracted* couriers, who are hired before the planning horizon, and *ad-hoc* couriers, who are hired during the operational stage. The proposed framework accounts for the uncertainty of ad-hoc couriers' availability and job bidding behavior, as well as the cost associated with hiring ad-hoc couriers when necessary. By formulating the problem as a robust optimization model, our framework provides a tool for decision-making that enables platforms to effectively manage their workforce resources and balance their cost objectives with service quality requirements.

The field of crowdsourced last-mile delivery has garnered increasing attention recently. Alnaggar et al. (2021) analyze the current industry status of crowdsourced delivery platforms based on matching mechanisms, target markets, and compensation schemes. They indicate that for a centralized system with an hourly compensation scheme (*e.g.*, Amazon Flex), a significant challenge is forecasting delivery needs and the number of couriers required to fulfill them. Although platforms can provide on-demand delivery blocks in case contracted couriers are insufficient, there is a higher level of risk involved since the availability of on-demand couriers is not guaranteed. Savelsbergh and Ulmer (2022) identify the challenges and opportunities in crowdsourced delivery planning and operations. The tactical challenge is ensuring that the required crowdsourced delivery capacity is available, while the operational question is how to adjust delivery capacity if the anticipated capacity is not materialized or if demand exceeds expectations. Besides uncertain availability, couriers' behavior is also uncertain, as they may accept or reject a delivery task and deviate from planned routes (Liu et al. 2021). To reduce uncertainties in delivery capacity, the authors suggest that planners can determine a set of delivery shifts and offer them to couriers for commitment before the operational period begins. Additionally, a dynamic compensation mechanism can be designed to adjust the availability of ad-hoc couriers.

Behrendt et al. (2022a) study the crowdsourced same-day delivery problem, deciding the fleet sizing of contracted couriers, pricing of ad-hoc couriers in the planning stage, and order allocation in the operational phase. They assume that order and ad-hoc courier arrivals follow Poisson processes and focus on evaluating the benefits of utilizing a hybrid workforce and various order allocation policies. Goyal et al. (2023) address a multistage problem involving the determination of contracted courier fleet sizes at each warehouse in the first stage, followed by assignment and routing decisions for both contracted and ad-hoc couriers once orders and ad-hoc couriers have arrived. They formulate the problem as a multistage stochastic integer program and develop an approximate dynamic programming method. These two studies assume that contracted couriers are available for the entire planning horizon once hired. In contrast, our work determines fleet sizes for each shift, accounting for varying uncertainty in each period.

The works of Ulmer and Savelsbergh (2020) and Behrendt et al. (2022b) are highly relevant to our paper. Specifically, Ulmer and Savelsbergh (2020) address the workforce scheduling problem for contracted couriers, determining the optimal number of shifts and their start time and duration. The authors consider stochastic arrivals of orders and ad-hoc couriers, as well as the duration an ad-hoc courier is willing to work. Their objective is to minimize working hours while ensuring a minimum percentage of orders are fulfilled. The authors employ continuous approximation and value function approximation methods, utilizing scenarios to represent realizations of random variables. Behrendt et al. (2022b) propose a prescriptive machine learning method for a similar problem. They leverage the sample average approximation (SAA) method offline to generate solutions for various instances, then use a machine learning model to produce online solutions for new demand and ad-hoc courier arrival forecasts. Their objective is to minimize courier payments and penalty costs for late deliveries. We note that besides the uncertain availability, our work also considers ad-hoc couriers' uncertain job bidding behavior. Moreover, we do not assume any probability information of random variables and use a data-driven robust optimization method to mitigate the impact of distributional ambiguity on the risk-based objective function.

Many studies on crowdsourced delivery have focused solely on addressing operational-level allocation and routing problems, without taking tactical decisions into account. For example, Archetti et al. (2016) address the vehicle routing problem (VRP) with occasional drivers, referring to drivers who are willing to make deliveries with a detour on their way to their destination. Dayarian and Savelsbergh (2020) investigate a same-day delivery problem that involves dynamically arriving online orders and in-store customers who, in addition to shopping, also deliver online orders as a supplement to company drivers. The authors propose rolling horizon dispatching approaches, with and without incorporating probabilistic information about future arrivals of orders and in-store customers. Similarly, Torres et al. (2022) tackle the VRP with stochastic crowd vehicles and customer presence, formulating the problem as a two-stage stochastic model and developing a column generation heuristic for solving large-size instances.

To the best of our knowledge, we are the first to use a data-driven decision framework to make tactical and operational planning decisions for the crowdsourced delivery problem with a hybrid workforce, considering ad-hoc couriers' uncertain availability and job bidding behavior. Our work is built upon the recent development of data-driven robust optimization, which aims to address issues in optimization under uncertainty. One prominent issue is the *Optimizer's curse* (Smith and Winkler 2006), a phenomenon that inferior results are always expected in the out-of-sample test if one uses the empirical distribution from training data to solve an optimization problem. A similar issue also appears in SAA (Birge and Louveaux 2011) through which the objective estimated in stochastic optimization is optimistically biased. Although the bias can be reduced with more samples, it would be prohibitive to do so in the data-driven setting where data is collected over time.

Data-driven robust optimization approach has been developed recently to address this issue (Bertsimas et al. 2018, Mohajerin Esfahani and Kuhn 2018, Gao and Kleywegt 2022). This method can effectively mitigate the risk of uncertainty by optimizing the worst-case objective within an ambiguity set that includes probability distributions of certain properties, such as having moment information matching the empirical distribution or being close to it based on specific distance metrics. Notably, Mohajerin Esfahani and Kuhn (2018) and Gao and Kleywegt (2022) propose the ambiguity set based on the Wasserstein distance that enjoys statistical guarantee to capture the true distribution. Nonetheless, the practical issue of determining the radius of the Wasserstein ambiguity set has been acknowledged. Instead of relying on the theoretical value of the statistical guarantee, cross-validation is often required to tune the radius parameter for better out-of-sample performance.

More recently, Long et al. (2022) propose the robust satisficing model, which is based on target-oriented optimization instead of conventional utility maximization. Contrary to robust optimization that specifies an ambiguity set of probability distributions, the robust satisficing model has no restriction on the distributions and minimizes the uncertainty risk of not achieving a specific target. Targets play an important role in the human decision process, especially in complex environments full of uncertainty and risks (Simon 1955, Mao 1970, Chen and Tang 2022). In articulating preferences, the robust satisficing model requires the decision-maker to set performance targets, which are more interpretable and practical to specify. The idea of robust satisficing has been emerging recently in both theoretical developments (Chen et al. 2022, Liu et al. 2022, Sim et al. 2021) and practical applications (Goh and Hall 2013, Zhang et al. 2021, Zhou et al. 2022).

We propose using the robust satisficing framework to address the workforce management problem, which involves the hiring of contracted couriers for each shift before the planning horizon. Whenever the contracted couriers are insufficient during the operational stage, the platform hires ad-hoc couriers based on their bidding for delivery jobs. Because the bidding strategy of ad-hoc couriers is unknown, it would be prohibitive to precisely determine the expected crowdsourcing costs associated with the delivery workforce in the planning horizon. The robust satisficing model aims to meet the specified expected cost target and perform as well as possible under distributional ambiguity. The contributions of our work are summarized as follows.

1. To characterize the empirical distribution of the crowdsourcing costs using historical bidding records made by the ad-hoc couriers, we first propose a basic reduced information model that evaluates the upper bound of the crowdsourcing cost. We prove that the bound is tight under the single payment value bidding scheme. We further introduce a generalized reduced information model to obtain a tighter bound on the crowdsourcing cost that would improve with the number of breakpoints.
2. Based on the basic reduced information model, we propose a data-driven robust satisficing model that incorporates a time additive Wasserstein metric to characterize distributional ambiguity in the workforce management problem. We can reformulate the proposed robust satisficing model as a single deterministic tractable convex optimization problem. We also extend this to the generalized reduced information model and demonstrate that a binary search algorithm can be applied to tackle the robust satisficing model exactly via solving a modest number of convex optimization problems.
3. We provide statistical justifications for the robust satisficing model that are based on target attainment. For small deviations from the target, we provide a target attainment confidence guarantee based on the probability bound of the Wasserstein distance. For large shortfalls from the target, we provide a concise expression of probability guarantee, which does not depend on the number of breakpoints of the generalized reduced information model.

4. Numerical tests on Solomon's data sets show that the basic reduced information model can provide decent approximations of the crowdsourcing costs for practical delivery problems when multiple payment values are allowed. The generalized reduced information model further improves the bounds, controlling all the gaps under 5%. Moreover, through simulated data, we demonstrate that the proposed robust satisficing model provides better out-of-sample performance than the empirical model, especially when dealing with high levels of uncertainty and risk.

Notation. We use boldface lowercase letters for vectors (*e.g.*, $\boldsymbol{\theta}$), and calligraphic letters for sets (*e.g.*, $\mathcal{X}$). We use $|\cdot|$ to denote the cardinality of a finite set. We use $[N]$ to denote the running index $\{1, 2, \ldots, N\}$ for $N \in \mathbb{N}$, and $[0] = \emptyset$. A random variable $\tilde{v}$ is denoted with a tilde sign such as $\tilde{v} \sim \mathbb{P}, \mathbb{P} \in \mathcal{P}_0$, where $\mathcal{P}_0$ represents the set of all possible distributions. For a multivariate random variable, we use $\mathcal{P}_0(\mathcal{Z})$ to represent the set of all distributions for the multivariate random variable that has support $\mathcal{Z} \subseteq \mathbb{R}^n$. Specifically, we use $\tilde{\boldsymbol{z}} \sim \mathbb{P}, \mathbb{P} \in \mathcal{P}_0(\mathcal{Z})$ to define $\tilde{\boldsymbol{z}}$ as a multivariate random variable with support $\mathcal{Z}$ and distribution $\mathbb{P}$. We use $\mathbb{E}_{\mathbb{P}}[\tilde{v}]$ to denote the expectation of a random variable, $\tilde{v} \sim \mathbb{P}$, over its distribution. Finally, $\mathbf{0}$ ($\mathbf{1}$) denotes the vector of all zeros (ones) and $\boldsymbol{e}_i$ denotes the ith basis vector. The dimensions of these vectors should be clear from the context.

2. Workforce management with crowdsourced delivery

We are examining a delivery workforce management platform that operates over a planning horizon consisting of T periods. The platform hires contracted couriers for each shift before the planning horizon. During the operational stage, if the contracted couriers are insufficient to handle the delivery jobs, the platform engages ad-hoc couriers based on their bidding for the available jobs. Before the planning horizon, the platform creates shifts to offer to crowdsourced couriers. If a courier signs up for a shift and the platform also confirms her/his enrollment, s/he becomes a contracted courier and will be managed by the platform during the operational period.

Without loss of generality, we assume that a sufficient number of couriers are registered because the platform can release the work shifts several days or weeks earlier than the start of the planning horizon. The platform has created I different work shifts; for each shift, $i \in [I]$, we define $\mathcal{S}_i \subseteq [T]$ as the set of time periods within the time horizon covered by the shift. For each time period $t \in [T]$, we define the vector $\boldsymbol{a}_t \in \{0,1\}^I$, where $a_{ti} = 1$ if $t \in \mathcal{S}_i$, and $a_{ti} = 0$, otherwise, for all $i \in [I]$. We let the decision variable $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{Z}_+^I$, where x_i represent the number of contracted couriers hired to work at shift $i \in [I]$, determined before the start of the time horizon. The feasible set $\mathcal{X}$ can encapsulate the detailed constraints associated with individual contracted couriers. The number of contracted workers at time $t \in [T]$ is $\boldsymbol{a}_t^\top \boldsymbol{x}$. The compensation to the contracted courier in the ith,

$i \in [I]$ work shift for the planning horizon is w_i. Hence, the total compensation for hiring contracted workers would be $\boldsymbol{w}^\top \boldsymbol{x}$.

Before the beginning of each period, $t \in [T]$, the set of packages has arrived and must be delivered by the end of the tth period. The delivery workforce management platform would lexicographically minimize the number of couriers and then the total traveling distance while adhering to delivery time windows and couriers' capacity constraints. Hence, the solutions to the VRP provide the number of jobs $J_t \in \mathbb{Z}_+$, where each job is a set of packages to be delivered by one courier by the end of the time period. As the employment decision, $\boldsymbol{x}$ is made before customer orders are realized, there may exist situations where the hired contracted couriers cannot complete all the delivery tasks. Specifically, at the beginning of period t, there would be potential $(J_t - \boldsymbol{a}_t^\top \boldsymbol{x})^+$ demands that the contracted couriers could not fulfill.

To provide high-quality service to customers, the platform guarantees that all packages realized at the beginning of a period will be delivered by the end of that period. Hence, the platform considers another matching mechanism by hiring ad-hoc couriers. Specifically, the platform releases all the J_t jobs to ad-hoc couriers who would bid for these jobs. In particular, the ad-hoc couriers specify their desired compensations for each job from a given list of N possible payments, $p_1, p_2, \cdots, p_N > 0$, with $p_n \leq p_{n+1}$. Let K_t denote the number of ad-hoc couriers participating in bidding at period t. We define the corresponding bidding set $\mathcal{B}_t$ as the set of tuples $(k, j, n) \in [K_t] \times [J_t] \times [N]$, with each tuple (k, j, n) representing courier k has bidden for job j for payment p_n. For convenience, we also define the projection of $\mathcal{B}_t$ on the tuples $(k, j) \in [K_t] \times [J_t]$ as follows

$$\bar{\mathcal{B}}_t := \{(k, j) \in [K_t] \times [J_t] \ \mid \ \exists n \in [N] : (k, j, n) \in \mathcal{B}_t\}.$$

Assumption 1. *We assume that p_N is high enough so that we can always find someone to take up any delivery job for that price. Consequently, we can assume that every job $j \in [J_t]$ can be assigned to an unique bidder $k_j \in [K_t]$, so that $(k_j, j) \in \bar{\mathcal{B}}_t$ for all $j \in [J_t]$ and $|\{k_j \mid j \in [J_t]\}| = J_t$.*

To ensure Assumption 1 always holds in practice, we can temporarily assign each job $j \in [J_t]$ to a phantom courier at payment p_N, and whenever a phantom courier is assigned, we can always replace it with someone who is willing to deliver at that price.

After the bids have been gathered, the platform would assign jobs to ad-hoc couriers. For a given employment decision of contracted workers $\boldsymbol{x} \in \mathcal{X}$, and a realized bidding set $\mathcal{B}_t$, we can obtain the information, J_t, K_t, and $\bar{\mathcal{B}}_t$. Subsequently, the planner decides the employment of ad-hoc couriers with the minimum crowdsourcing cost, $f_t(\boldsymbol{x}, \mathcal{B}_t)$, where

$$
\begin{aligned}
f_t(\boldsymbol{x}, \mathcal{B}_t) = \min \ & \sum_{(k,j,n)\in\mathcal{B}_t} p_n s_{kj} \\
\text{s.t.} \ & \sum_{k:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \le 1 && \forall j \in [J_t], \\
& \sum_{j:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \le 1 && \forall k \in [K_t], \\
& \sum_{(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \ge J_t - \boldsymbol{a}_t^\top \boldsymbol{x} \\
& s_{kj} \ge 0 && \forall (k,j) \in \bar{\mathcal{B}}_t,
\end{aligned}
\tag{1}
$$

in which the first set of constraints ensures that job $j \in [J_t]$ is assigned to at most one ad-hoc courier, while the second set of constraints requires that each ad-courier $k \in [K_t]$ is assigned to at most one job. The third set of constraints demands that all jobs are assigned to contracted or ad-hoc couriers. Since this is a network flow optimization problem, if $\boldsymbol{a}_t^\top \boldsymbol{x} \in \mathbb{Z}$, then there exist binary optimal solutions for the decision variables s_{kj}, $(k,j) \in \bar{\mathcal{B}}_t$ such that $s_{kj} = 1$ if courier k's bid for job j is accepted, and $s_{kj} = 0$ otherwise. All the remaining jobs are consequently assigned to contracted couriers.

During the planning horizon, we do not know the future arrivals of packages and how ad-hoc couriers would bid for the jobs. We can denote $(\tilde{\mathcal{B}}_1, \dots, \tilde{\mathcal{B}}_T)$ as the joint random bidding sets for all time periods, and its true distribution $\mathbb{Q}^\star$, $(\tilde{\mathcal{B}}_1, \dots, \tilde{\mathcal{B}}_T) \sim \mathbb{Q}^\star$ is unobservable to the decision maker. Hence, it is impossible to determine the true optimum solution in the following *ideal optimization problem*.

$$
\begin{aligned}
Z^\star = \min \ & \boldsymbol{w}^\top \boldsymbol{x} + \mathbb{E}_{\mathbb{Q}^\star}\left[\sum_{t\in[T]} f_t(\boldsymbol{x}, \tilde{\mathcal{B}}_t)\right] \\
\text{s.t.} \ & \boldsymbol{x} \in \mathcal{X}.
\end{aligned}
\tag{2}
$$

Nevertheless, we have access to Ω historical sample path records of the bidding sets $\mathcal{B}_t^\omega$, for all $\omega \in [\Omega]$, $t \in [T]$. Accordingly, we denote $\hat{\mathbb{Q}}$ as the empirical distribution such that

$$
\mathbb{E}_{\hat{\mathbb{Q}}}\left[\tilde{\mathcal{B}}_t = \mathcal{B}_t^\omega \quad \forall t \in [T]\right] = \frac{1}{\Omega} \qquad \forall \omega \in [\Omega].
$$

To obtain the decisions for the contracted couriers, $\boldsymbol{x}$, we can solve the *empirical optimization problem* that minimizes the average cost for the planning horizon as follows

$$
\begin{aligned}
\min \ & \boldsymbol{w}^\top \boldsymbol{x} + \mathbb{E}_{\hat{\mathbb{Q}}}\left[\sum_{t\in[T]} f_t(\boldsymbol{x}, \tilde{\mathcal{B}}_t)\right] \\
\text{s.t.} \ & \boldsymbol{x} \in \mathcal{X},
\end{aligned}
\tag{3}
$$

which is equivalent to the following linear optimization problem,

$$
\begin{aligned}
\hat{Z} = \min\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \sum_{t \in [T]} \sum_{(k,j,n) \in \mathcal{B}_t^\omega} p_n s_{tkj}^\omega \\
\text{s.t.}\ & \sum_{k:(k,j) \in \bar{\mathcal{B}}_t^\omega} s_{tkj}^\omega \leq 1 && \forall j \in [J_t^\omega], t \in [T], \omega \in [\Omega], \\
& \sum_{j:(k,j) \in \bar{\mathcal{B}}_t^\omega} s_{tkj}^\omega \leq 1 && \forall k \in [K_t^\omega], t \in [T], \omega \in [\Omega], \\
& \sum_{(k,j) \in \bar{\mathcal{B}}_t^\omega} s_{tkj}^\omega \geq J_t^\omega - \boldsymbol{a}_t^\top \boldsymbol{x} && \forall t \in [T], \omega \in [\Omega], \\
& s_{tkj}^\omega \geq 0 && \forall (k,j) \in \bar{\mathcal{B}}_t^\omega, t \in [T], \omega \in [\Omega], \\
& \boldsymbol{x} \in \mathcal{X}.
\end{aligned}
\tag{4}
$$

3. Reduced information models

Apart from being a large-scale linear optimization problem, we are unable to extend Problem (4) to a data-driven robust optimization model due to the difficulties of characterizing the statistics associated with the random bidding sets, $(\tilde{\mathcal{B}}_1, \ldots, \tilde{\mathcal{B}}_T)$. As such, we propose a reduced information model that evaluates an upper bound of the crowdsourcing cost function. Specifically, after the bids have been gathered at the tth period, we solve the following assignment problem,

$$
\begin{aligned}
\min\ & \sum_{(k,j,n) \in \mathcal{B}_t} p_n s_{kj} \\
\text{s.t.}\ & \sum_{k:(k,j) \in \bar{\mathcal{B}}_t} s_{kj} = 1 \ \forall j \in [J_t], \\
& \sum_{j:(k,j) \in \bar{\mathcal{B}}_t} s_{kj} \leq 1 \ \forall k \in [K_t], \\
& s_{kj} \geq 0 \qquad \forall (k,j) \in \bar{\mathcal{B}}_t,
\end{aligned}
\tag{5}
$$

and obtain its optimum binary solution, $s_{kj}^* \in \{0,1\}$, $(k,j) \in \bar{\mathcal{B}}_t$. Observe that under Assumption 1, Problem (5) is a feasible assignment problem. Subsequently, we determine the *basic reduced information vector*, $\boldsymbol{z}_t \in \mathbb{R}^N$ where

$$
z_{tn} = \sum_{(k,j):(k,j,n) \in \mathcal{B}_t} s_{kj}^* \qquad \forall n \in [N]. \tag{6}
$$

Speaking intuitively, z_{tn} is the maximum number of ad-hoc couriers, based on the optimal assignment solution of Problem (5), who could be assigned for the jobs for p_n payment.

Theorem 1. *The crowdsourcing cost function with basic reduced information,*

$$
\begin{aligned}
g_t(\boldsymbol{x}, \boldsymbol{z}_t) = \min\ & \boldsymbol{p}^\top \boldsymbol{y} \\
\text{s.t.}\ & \mathbf{1}^\top \boldsymbol{y} \geq \mathbf{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \\
& \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{z}_t,
\end{aligned}
\tag{7}
$$

is an upper bound of the crowdsourcing cost function, i.e.,

$$f_t(\boldsymbol{x}, \mathcal{B}_t) \le g_t(\boldsymbol{x}, \boldsymbol{z}_t).$$

Moreover, the bound is tight if there exists an optimal binary solution of Problem (1) *such that* $s_{kj} \le s^*_{kj}$ *for all* $(k,j) \in \bar{\mathcal{B}}_t$.

Proof of Theorem 1. Observe that the value of $g_t(\boldsymbol{x}, \boldsymbol{z}_t)$ corresponds to the allocation of $J_t - \boldsymbol{a}_t^\top \boldsymbol{x}$ ad-hoc couriers using the solution $\boldsymbol{s}^*$, which has J_t assigned couriers, and then removing $\boldsymbol{a}_t^\top \boldsymbol{x}$ of the most expensive bidders. Such an assignment is a feasible solution to Problem (1). Hence,

$$f_t(\boldsymbol{x}, \mathcal{B}_t) \le g_t(\boldsymbol{x}, \boldsymbol{z}_t).$$

Now suppose there exists an optimal solution of Problem (1) such that $\boldsymbol{s} \le \boldsymbol{s}^*$. Then we can construct the solution,

$$y_n = \sum_{(k,j):(k,j,n)\in\mathcal{B}_t} s_{kj} \qquad \forall n \in [N],$$

which is feasible in Problem (7) and its objective

$$\boldsymbol{p}^\top \boldsymbol{y} = \sum_{(k,j,n)\in\mathcal{B}_t} p_n s_{kj}$$

coincides with the optimum objective of Problem (1). Hence, in this case, we also have $f_t(\boldsymbol{x}, \mathcal{B}_t) \ge g_t(\boldsymbol{x}, \boldsymbol{z}_t)$. $\square$

The bidding scheme mandated by the platform influences the accuracy of the reduced information model in evaluating the crowdsourcing cost function. In a single payment value bidding, each ad-hoc courier, $k \in [K_t]$ can bid for any number of jobs but for one payment value $r_k \in \{p_1, \ldots, p_N\}$ for any job assigned by the platform.

Theorem 2 **(Single payment value bidding)**. *Under the single payment value bidding scheme, the basic reduced information model evaluates the crowdsourcing cost exactly, i.e.,* $g_t(\boldsymbol{x}, \boldsymbol{z}_t) = f_t(\boldsymbol{x}, \mathcal{B}_t)$ *for any number of assignments* $J_t - \boldsymbol{a}_t^\top \boldsymbol{x} \in \{0, 1, \cdots, J_t\}$.

Proof of Theorem 2. We first define the set,

$$\mathcal{F} := \left\{ \mathcal{K} \subseteq [K_t] \;\middle|\; \begin{array}{l} \forall k \in \mathcal{K}, \exists j_k \in [J_t] : \\ (k, j_k) \in \bar{\mathcal{B}}_t, \\ |\mathcal{K}| = |\{j_k | k \in \mathcal{K}\}| \end{array} \right\},$$

so that each element in $\mathcal{F}$ is a set of couriers in which every courier in the set could be assigned to a unique job.

Given a set of selected couriers $\mathcal{K} \in \mathcal{F}$, we consider the following assignment problem that minimizes the total cost,

$$\begin{aligned} G(\mathcal{K}) = \min \quad & \sum_{(k,j,n)\in\mathcal{B}_t, k\in\mathcal{K}} p_n s_{kj} \\ \text{s.t.} \quad & \sum_{k:(k,j)\in\bar{\mathcal{B}}_t, k\in\mathcal{K}} s_{kj} \leq 1 \ \forall j \in [J_t], \\ & \sum_{j:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} = 1 \qquad \forall k \in \mathcal{K}, \\ & \sum_{j:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} = 0 \qquad \forall k \in [K_t]\backslash\mathcal{K}, \\ & s_{kj} \geq 0 \qquad\qquad \forall (k,j) \in \bar{\mathcal{B}}_t, \end{aligned}$$

where the second collection of constraints specifies that any courier in $\mathcal{K}$ must be assigned with one job, and the third collection of constraints ensures that any courier outside $\mathcal{K}$ cannot be assigned with any job. The above assignment problem is a minimum-cost network flow optimization problem. From discrete convex analysis, the minimum-cost network flow problem is an M-convex problem, *i.e.*, $G:\mathcal{F}\to\mathbb{R}$ is an M-convex function (see example 2.3 in Murota 1998 or section 4.1 in Chen and Li 2021). A feasible solution $\mathcal{K}$ is a *local minimum* in the sense that we cannot replace any courier $k_1 \in \mathcal{K}$ with a courier $k_2 \in [K_t]\backslash\mathcal{K}$ such that $r_{k_2} < r_{k_1}$ and $(\mathcal{K}\cup\{k_2\})\backslash\{k_1\}\in\mathcal{F}$. By Theorem 4.6 in Murota (1998), any local minimum of the M-convex function is also a global minimum. Note that whenever referring to the local/global minimum of G, we assume the domain of G is contained in a hyperplane $\{\mathcal{K}\in\mathcal{F}: |\mathcal{K}| = \omega\}$ for some $\omega \in \{0,1,\cdots,J_t\}$.

Next, we consider a sequence of optimal solutions for the following problems,

$$\begin{aligned} f^w = \min \quad & \sum_{(k,j,n)\in\mathcal{B}_t} p_n s_{kj} \\ \text{s.t.} \quad & \sum_{k:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \leq 1 \ \forall j \in [J_t], \\ & \sum_{j:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \leq 1 \ \forall k \in [K_t], \\ & \sum_{(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \geq w \\ & s_{kj} \geq 0 \qquad\qquad \forall (k,j) \in \bar{\mathcal{B}}_t, \end{aligned} \tag{8}$$

for $w \in \{0,1,\ldots,J_t\}$, and we use $\boldsymbol{s}^w$ to denote the corresponding optimal solution. Obviously, we have $\min_{\mathcal{K}\in\mathcal{F},|\mathcal{F}|=\omega} G(\mathcal{K}) \geq f^\omega$. For a given $w \in [J_t]$, let $\mathcal{K}_w$ be the corresponding set of selected couriers,

$$\mathcal{K}_w = \{k \in [K_t] \ \mid \ \exists j \in [J_t] : s^w_{kj} = 1\},$$

and each courier, $k \in \mathcal{K}_w$ is assigned to the job $j_k \in [J_t]$ so that $s^w_{kj_k} = 1$. Observe that $\mathcal{K}_w \in \mathcal{F}$, $|\mathcal{K}_w| = w$, and the total payment, $\sum_{k \in \mathcal{K}_w} r_k$ does not depend on how the w jobs are being assigned to the couriers in $\mathcal{K}_w$. Hence,

$$f^w = \min_{\mathcal{K} \in \mathcal{F}, |\mathcal{K}| = \omega} G(\mathcal{K}) = G(\mathcal{K}_\omega).$$

Now let $\mathcal{K}_{w-1} = \mathcal{K}_w \backslash \{k_0\}$ for some $k_0 \in \arg\max_{k \in \mathcal{K}_w} \{r_k\}$. We claim the set $\mathcal{K}_{w-1}$ is also a local minimum of G (in domain $\{\mathcal{K} \in \mathcal{F} : |\mathcal{K}| = \omega - 1\}$). Suppose this is not the case, we can replace some courier $k_1 \in \mathcal{K}_{w-1}$ with a courier $k_2 \in [K_t] \backslash \mathcal{K}_{w-1}$ such that $r_{k_2} < r_{k_1}$ and $\bar{\mathcal{K}} \cup \{k_2\} \in \mathcal{F}$, where $\bar{\mathcal{K}} = \mathcal{K}_w \backslash \{k_0, k_1\}$. Observe that $k_2 \neq k_0$ since $r_{k_0} \geq r_k$ for all $k \in \mathcal{K}_{w-1}$. To arrive at the contradiction that $\mathcal{K}_{w-1}$ is not a local minimum, we consider an assignment solution, $\bar{\boldsymbol{s}}$ for couriers in $\bar{\mathcal{K}}$ that based on $\boldsymbol{s}^w$ as follows

$$\bar{s}_{kj} = \begin{cases} s^w_{kj} \text{ if } k \notin \{k_0, k_1\} \\ 0 \quad \text{otherwise} \end{cases} \qquad \forall (k,j) \in \bar{\mathcal{B}}_t.$$

Since the assignment problem is a network flow problem, we can construct the residual network associated with the solution $\bar{\boldsymbol{s}}$ (see, *e.g.*, Ahuja et al. 1988). Moreover, because $\bar{\mathcal{K}} \cup \{k_2\} \in \mathcal{F}$, we can find a feasible assignment solution using a max-flow algorithm by sending a unit of residual flow from node k_2, along an augmenting path on the residual network associated with $\bar{\boldsymbol{s}}$ (refer to augmenting path algorithm for max-flow in Ahuja et al. 1988), that will terminate at one of the unassigned jobs, $\ell \in [J_t] \backslash \bar{\mathcal{J}}$ where $\bar{\mathcal{J}}$ is the set of jobs that assigned to couriers in $\bar{\mathcal{K}}$ based on $\bar{\boldsymbol{s}}$. If $\ell \neq j_{k_0}$, then we would have $\bar{\mathcal{K}} \cup \{k_0, k_2\} \in \mathcal{F}$. However, it contradicts that $\mathcal{K}_w$ is the local minimum since we can replace $k_1 \in \mathcal{K}_w$ with k_2 to achieve a lower total payment. On the other hand, if $\ell = j_{k_0} \neq j_{k_1}$, then $\bar{\mathcal{K}} \cup \{k_1, k_2\} \in \mathcal{F}$. However, this will also contradict that $\mathcal{K}_w$ is local minimum; since $r_{k_0} \geq r_{k_1} > r_{k_2}$, we can replace $k_0 \in \mathcal{K}_w$ with k_2 to achieve a lower total payment. Therefore, by contradiction, we must have $\bar{\mathcal{K}} \cup \{k_2\} \notin \mathcal{F}$, implying that the set $\mathcal{K}_{w-1}$ is local minimum. Consequently, it is also global minimum and we have $f^{w-1} = \min_{\mathcal{K} \in \mathcal{F}, |\mathcal{K}| = \omega - 1} G(\mathcal{K}) = G(\mathcal{K}_{w-1})$.

Notice that the assignment optimization problem in G only uses couriers in $\mathcal{K}_{w-1}$ which rules out the courier with the highest cost from $\mathcal{K}_\omega$. Therefore, if $G(\mathcal{K}_w)$ equals the optimal objective value of the basic reduced information model (7) with $\mathbf{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} = \omega$ assignments, then $G(\mathcal{K}_{w-1})$ must equal the optimal objective value of the basic reduced information model with $\omega - 1$ assignments. Also, from previous analysis, if $\mathcal{K}_\omega \in \arg\min_{\mathcal{K} \in \mathcal{F}, |\mathcal{K}| = \omega} G(\mathcal{K})$ and $f^w = G(\mathcal{K}_w)$, then $\mathcal{K}_{\omega-1} \in \arg\min_{\mathcal{K} \in \mathcal{F}, |\mathcal{K}| = \omega - 1} G(\mathcal{K})$ and $f^{w-1} = G(\mathcal{K}_{w-1})$. Obviously, $G(\mathcal{K}_{J_t})$ equals the optimal objective value of the basic reduced information model with J_t assignments, $\mathcal{K}_{J_t} \in \arg\min_{\mathcal{K} \in \mathcal{F}, |\mathcal{K}| = J_t} G(\mathcal{K})$ and $f^{J_t} = G(\mathcal{K}_{J_t})$. By mathematical induction, the basic reduced information model evaluates the crowdsourcing cost exactly for any number of assignments $w \in \{0, 1, \cdots, J_t\}$. $\square$

Given the historical information $\mathcal{B}_t^\omega$, for $\omega \in [\Omega]$, $t \in [T]$, by solving Problem (5), we can obtain the corresponding basic reduced information vector, $\boldsymbol{z}_t^\omega$, which is used to evaluate the upper bound of the crowdsourcing cost function. The basic reduced information allows us to characterize the underlying random variables associated with the delivery workforce management problem. For each period, $t \in [T]$, the random variable $\tilde{\boldsymbol{z}}_t$ represents the random basic reduced information vector associated with the bidding sets. For convenience, we define $\tilde{\boldsymbol{z}} := (\tilde{\boldsymbol{z}}_t)_{t\in[T]}$ and its support set is

$$\mathcal{Z} := \{(\boldsymbol{z}_t)_{t\in[T]} | \boldsymbol{z}_t \in \mathcal{Z}_t \ \forall t \in [T]\},$$

$$\mathcal{Z}_t = \{\boldsymbol{z} \in \mathbb{R}_+^N \mid \mathbf{1}^\top \boldsymbol{z} \le \bar{z}_t\},$$

where $\bar{z}_t$ is the maximum number of jobs that would ever arrive at the time period t. We denote the Ω historical realizations of the random variables by $\boldsymbol{z}_t^\omega \in \mathcal{Z}_t$, $\omega \in [\Omega]$, $t \in [T]$. Correspondingly, we define the empirical distribution $\hat{\mathbb{P}} \in \mathcal{P}_0(\mathcal{Z})$, $\tilde{\boldsymbol{z}} \sim \hat{\mathbb{P}}$ such that for all $\omega \in [\Omega]$, $\hat{\mathbb{P}}[\tilde{\boldsymbol{z}}_t = \boldsymbol{z}_t^\omega \quad \forall t \in [T]] = \frac{1}{\Omega}$. Specifically, with the basic reduced information, we solve the following empirical optimization problem,

$$\begin{aligned} \bar{Z}_0 = \min \ & \boldsymbol{w}^\top \boldsymbol{x} + \mathbb{E}_{\hat{\mathbb{P}}}\left[\sum_{t\in[T]} g_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}_t)\right] \\ \text{s.t.} \ & \boldsymbol{x} \in \mathcal{X}. \end{aligned} \tag{9}$$

which is an upper bound to the empirical optimization problem (3), and it is equivalent to the following linear optimization problem,

$$\begin{aligned} \bar{Z}_0 = \min \ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega\in[\Omega]} \sum_{t\in[T]} \boldsymbol{p}^\top \boldsymbol{y}_t^\omega \\ \text{s.t.} \ & \mathbf{1}^\top \boldsymbol{y}_t^\omega \ge \mathbf{1}^\top \boldsymbol{z}_t^\omega - \boldsymbol{a}_t^\top \boldsymbol{x} && \forall t \in [T], \omega \in [\Omega], \\ & \mathbf{0} \le \boldsymbol{y}_t^\omega \le \boldsymbol{z}_t^\omega && \forall t \in [T], \omega \in [\Omega], \\ & \boldsymbol{x} \in \mathcal{X}. \end{aligned}$$

For a bidding platform where multiple payment values are allowed, the basic reduced information model may not be exact. Unfortunately, the relative performance gap can be unbounded. Consider an instance with $N = 3$, $J_t = K_t = 2$, $\boldsymbol{p} = (\epsilon, 1, 3)$, for some small $\epsilon > 0$ and $\mathcal{B}_t = \{(1,1,1),(1,2,2),(2,1,2),(2,2,3)\}$. Hence, $\boldsymbol{z}_t = (0,2,0)$. For $\boldsymbol{a}_t^\top \boldsymbol{x} = 1$, observe that $f_t(\boldsymbol{x}, \mathcal{B}_t) = \epsilon$, while $g_t(\boldsymbol{x}, \boldsymbol{z}_t) = 1$, implying that the relative performance gap, $g_t(\boldsymbol{x}, \boldsymbol{z}_t)/f_t(\boldsymbol{x}, \mathcal{B}_t) = 1/\epsilon$ can be arbitrarily large. Nevertheless, we can narrow the gap through a more general reduced information model, which will be discussed in the next section.

Generalized reduced information model

We can further improve the basic reduced information model to obtain a tighter bound on the crowdsourcing cost function. To do so, for each period $t \in [T]$, we consider L_t breakpoints $u_{t\ell} \in [0, \bar{z}_t] \cap \mathbb{Z}_+$, $\ell \in [L_t]$, with $u_{t1} = 0$. Subsequently, we derive an approximation of the crowdsourcing cost function that would be tight if $\boldsymbol{a}_t^\top \boldsymbol{x} = u_{t\ell}$ for some $\ell \in [L_t]$. In particular, given the bidding information $\mathcal{B}_t$, $t \in [T]$, we determine the set of generalized reduced information vectors, $\boldsymbol{z}_t^\ell \in \mathbb{R}^N$, $\ell \in [L_t]$ where

$$z_{tn}^\ell = \sum_{(k,j):(k,j,n)\in\mathcal{B}_t} s_{kj}^{t\ell} \qquad \forall n \in [N], \ell \in [L_t]$$

and

$$\begin{aligned} \boldsymbol{s}^{t\ell} \in \arg\min \ & \sum_{(k,j,n)\in\mathcal{B}_t} p_n s_{kj} \\ \text{s.t.} \ & \sum_{k:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \le 1 && \forall j \in [J_t], \\ & \sum_{j:(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \le 1 && \forall k \in [K_t], \\ & \sum_{(k,j)\in\bar{\mathcal{B}}_t} s_{kj} \ge J_t - u_{t\ell} \\ & s_{kj} \ge 0 && \forall (k,j) \in \bar{\mathcal{B}}_t. \end{aligned}$$

Observe that since $u_{t1} = 0$, $\boldsymbol{z}_t^1$ corresponds to the basic reduced information vector. We also note that whenever $u_{t\ell} \ge J_t$, we have $\boldsymbol{z}_t^\ell = \boldsymbol{0}$.

We now extend to a generalized reduced information model by first characterizing the underlying random variable. As a generalization, we define $\tilde{\boldsymbol{z}} := (\tilde{\boldsymbol{z}}_t^\ell)_{t\in[T],\ell\in[L_t]}$, where $\tilde{\boldsymbol{z}}_t^\ell$ represents the random reduced information vector associated with the breakpoint $\ell \in [L_t]$ at the tth period. We consider breakpoints separable support sets

$$\mathcal{Z} := \left\{ (\boldsymbol{z}_t^\ell)_{t\in[T],\ell\in[L_t]} | \boldsymbol{z}_t^\ell \in \mathcal{Z}_t^\ell \ \forall t \in [T], \ell \in [L_t] \right\},$$

where

$$\mathcal{Z}_t^\ell = \{ \boldsymbol{z} \in \mathbb{R}_+^N \mid \boldsymbol{1}^\top \boldsymbol{z} \le \bar{z}_{t\ell} \},$$

and $\bar{z}_{t\ell} = \bar{z}_t - u_{t\ell} \ge 1$. Hence, noting that $u_1 = 0$, with $L = 1$, the random variable $\tilde{\boldsymbol{z}}$ is a generalization over the random basic reduced information vector. Correspondingly, the empirical distribution is $\hat{\mathbb{P}} \in \mathcal{P}_0(\mathcal{Z})$, $\tilde{\boldsymbol{z}} \sim \hat{\mathbb{P}}$ such that for all $\omega \in [\Omega]$,

$$\hat{\mathbb{P}} \left[\tilde{\boldsymbol{z}}_t^\ell = \boldsymbol{z}_t^{\ell\omega} \ \ \forall t \in [T], \ell \in [L_t] \right] = \frac{1}{\Omega},$$

where $\boldsymbol{z}_t^{\ell\omega} \in \mathcal{Z}_t^\ell$, $\ell \in [L_t]$ is the historical realization of the generalized reduced information associated with the bidding set $\mathcal{B}_t^\omega$ at $t \in [T]$.

THEOREM 3. *For any $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{\eta}, \boldsymbol{\gamma} \in \mathbb{R}_+^{L_t}$ such that $\mathbf{1}^\top \boldsymbol{\gamma} = 1$ and $\mathbf{1}^\top \boldsymbol{\eta} = \boldsymbol{a}_t^\top \boldsymbol{x}$, we have*

$$f_t(\boldsymbol{x}, \mathcal{B}_t) \leq \sum_{\ell \in [L_t]} h_t(\gamma_\ell, \eta_\ell - u_{t\ell}\gamma_\ell, \boldsymbol{z}_t^\ell), \tag{10}$$

where

$$\begin{aligned} h_t(\gamma, \eta, \boldsymbol{z}) = \min\ & \boldsymbol{p}^\top \boldsymbol{y} \\ \text{s.t.}\ & \mathbf{1}^\top \boldsymbol{y} \geq \mathbf{1}^\top \boldsymbol{z}\gamma - \eta \\ & \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{z}\gamma. \end{aligned} \tag{11}$$

Equality holds if $\gamma_{\ell^*} = 1$, $\boldsymbol{a}_t^\top \boldsymbol{x} = u_{t\ell^*}$ for some $\ell^* \in [L_t]$. Moreover,

$$\begin{aligned} \mathbb{E}_{\hat{\mathbb{P}}}\left[g_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}_t^1)\right] \geq \min\ & \mathbb{E}_{\hat{\mathbb{P}}}\left[\sum_{\ell \in [L_t]} h_t(\gamma_\ell, \eta_\ell - u_{t\ell}\gamma_\ell, \tilde{\boldsymbol{z}}_t^\ell)\right] \\ \text{s.t.}\ & \mathbf{1}^\top \boldsymbol{\gamma} = 1 \\ & \mathbf{1}^\top \boldsymbol{\eta} = \boldsymbol{a}_t^\top \boldsymbol{x} \\ & \boldsymbol{\eta}, \boldsymbol{\gamma} \in \mathbb{R}_+^{L_t}. \end{aligned} \tag{12}$$

Proof of Theorem 3. Observe that since $\eta_\ell \geq 0$, we have $h_t(0, \eta_\ell, z_t^\ell) = 0$. Moreover, for all $\gamma_\ell > 0$, we have

$$\begin{aligned} h_t(\gamma_\ell, \eta_\ell - u_{t\ell}\gamma_\ell, \boldsymbol{z}_t^\ell) &= \min\left\{ \boldsymbol{p}^\top \boldsymbol{y} \;\middle|\; \begin{array}{l} \mathbf{1}^\top \boldsymbol{y} \geq \mathbf{1}^\top \boldsymbol{z}_t^\ell \gamma_\ell - (\eta_\ell - u_{t\ell}\gamma_\ell) \\ \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{z}_t^\ell \gamma_\ell \end{array} \right\} \\ &= \gamma_\ell \min\left\{ \boldsymbol{p}^\top \boldsymbol{y} \;\middle|\; \begin{array}{l} \mathbf{1}^\top \boldsymbol{y} \geq \mathbf{1}^\top \boldsymbol{z}_t^\ell - (\eta_\ell/\gamma_\ell - u_{t\ell}) \\ \mathbf{0} \leq \boldsymbol{y} \leq \boldsymbol{z}_t^\ell \end{array} \right\} \\ &= \gamma_\ell h_t(1, \eta_\ell/\gamma_\ell - u_{t\ell}, \boldsymbol{z}_t^\ell). \end{aligned}$$

Moreover, following the same argument in the proof of Theorem 1, observe that for $\eta \geq u_{t\ell}$ the value of the function $h_t(1, \eta - u_{t\ell}, \boldsymbol{z}_t^\ell)$ corresponds to the allocation of $J_t - \eta$ ad-hoc couriers using the solution $\boldsymbol{s}^{t\ell}$, which has $J_t - u_{t\ell}$ assigned couriers, and removing $\eta - u_{t\ell}$ of the most expensive bidders, with at most one partial removal if η is fractional. However, when $\eta < u_{t\ell}$, we have $h_t(1, \eta - u_{t\ell}, \boldsymbol{z}_t^\ell) = \infty$, since the underlying minimization problem would be infeasible. Hence, we have

$$h_t(1, \eta - u_{t\ell}, \boldsymbol{z}_t^\ell) \geq \bar{f}_t(\eta, \mathcal{B}_t),$$

where

$$\begin{aligned} \bar{f}_t(\eta, \mathcal{B}_t) = \min\ & \sum_{(k,j,n) \in \mathcal{B}_t} p_n s_{kj} \\ \text{s.t.}\ & \sum_{k:(k,j) \in \bar{\mathcal{B}}_t} s_{kj} \leq 1 && \forall j \in [J_t], \\ & \sum_{j:(k,j) \in \bar{\mathcal{B}}_t} s_{kj} \leq 1 && \forall k \in [K_t], \\ & \sum_{(k,j) \in \bar{\mathcal{B}}_t} s_{kj} \geq J_t - \eta \\ & s_{kj} \geq 0 && \forall (k,j) \in \bar{\mathcal{B}}_t. \end{aligned} \tag{13}$$

Therefore, any feasible solution to Problem (10)

$$\begin{aligned}
&\sum_{\ell\in[L_t]} h_t(\gamma_\ell, \eta_\ell - u_{t\ell}\gamma_\ell, \boldsymbol{z}_t^\ell) \\
=& \sum_{\ell\in[L_t]:\gamma_\ell>0} \gamma_\ell h_t(1, \eta_\ell/\gamma_\ell - u_{t\ell}, \boldsymbol{z}_t^\ell) \\
\geq& \sum_{\ell\in[L_t]:\gamma_\ell>0} \gamma_\ell \bar{f}_t(\eta_\ell/\gamma_\ell, \mathcal{B}_t) \\
\geq& \ \bar{f}_t(\mathbf{1}^\top\boldsymbol{\eta}, \mathcal{B}_t) \\
=& \ \bar{f}_t(\boldsymbol{a}_t^\top\boldsymbol{x}, \mathcal{B}_t) \\
=& \ f_t(\boldsymbol{x}, \mathcal{B}_t)
\end{aligned}$$

where the last inequality is due to the function $\bar{f}_t(\eta, \mathcal{B}_t)$ being convex in η, $\boldsymbol{\gamma}\geq\mathbf{0}$ and $\mathbf{1}^\top\boldsymbol{\gamma}=1$.

Suppose $\gamma_{\ell^*}=1$, $\boldsymbol{a}_t^\top\boldsymbol{x}=u_{t\ell^*}$ for some $\ell^*\in[L_t]$, then we must have $\gamma_\ell=\eta_\ell=0$ for $\ell\in[L_t], \ell\neq\ell^*$, so that $\eta_{\ell^*}=\boldsymbol{a}_t^\top\boldsymbol{x}$ and

$$\begin{aligned}
f_t(\boldsymbol{x}, \mathcal{B}_t) &\leq \sum_{\ell\in[L_t]} h_t(\gamma_\ell, \eta_\ell - u_{t\ell}\gamma_\ell, \boldsymbol{z}_t^\ell) \\
&= h_t(\gamma_{\ell^*}, \eta_{\ell^*} - u_{t\ell^*}\gamma_{\ell^*}, \boldsymbol{z}_t^{\ell^*}) \\
&= h_t(1, 0, \boldsymbol{z}_t^{\ell^*}) \\
&= f_t(\boldsymbol{x}, \mathcal{B}_t),
\end{aligned}$$

where the final equality follows from the same argument in Theorem 1, since the optimal solution of Problem (1) is the same as $\boldsymbol{s}^{\ell^*}$ in which the reduced information $\boldsymbol{z}_t^{\ell^*}$ is derived. To show the bound (12), it suffices to note that

$$h_t(1, \boldsymbol{a}_t^\top\boldsymbol{x} - u_{t1}, \boldsymbol{z}_t^1) = h_t(1, \boldsymbol{a}_t^\top\boldsymbol{x}, \boldsymbol{z}_t^1) = g_t(\boldsymbol{x}, \boldsymbol{z}_t^1).$$

□

In the same vein, we can also extend the generalized reduced information model, which solves the following empirical optimization problem,

$$\begin{aligned}
Z_0 = \min\ & \boldsymbol{w}^\top\boldsymbol{x} + \mathbb{E}_{\hat{\mathbb{P}}}\left[\sum_{t\in[T]}\sum_{\ell\in[L_t]} h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \tilde{\boldsymbol{z}}_t^\ell)\right] \\
\text{s.t.}\ & \mathbf{1}^\top\boldsymbol{\eta}_t = \boldsymbol{a}_t^\top\boldsymbol{x} && \forall t\in[T], \\
& \mathbf{1}^\top\boldsymbol{\gamma}_t = 1 && \forall t\in[T], \\
& \boldsymbol{\eta}_t, \boldsymbol{\gamma}_t \in \mathbb{R}_+^{L_t} && \forall t\in[T], \\
& \boldsymbol{x}\in\mathcal{X},
\end{aligned}$$

or equivalently

$$
\begin{aligned}
Z_0 = \min\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega}\sum_{\omega\in[\Omega]}\sum_{t\in[T]}\sum_{\ell\in[L_t]} \boldsymbol{p}^\top \boldsymbol{y}_t^{\ell\omega} \\
\text{s.t.}\ & \mathbf{1}^\top \boldsymbol{y}_t^{\ell\omega} \ge \mathbf{1}^\top \boldsymbol{z}_t^{\ell\omega}\gamma_{t\ell} - \eta_{t\ell} + u_{t\ell}\gamma_{t\ell} && \forall t\in[T], \omega\in[\Omega], \ell\in[L_t], \\
& \mathbf{0} \le \boldsymbol{y}_t^{\ell\omega} \le \boldsymbol{z}_t^{\ell\omega}\gamma_{t\ell} && \forall t\in[T], \omega\in[\Omega], \ell\in[L_t], \\
& \mathbf{1}^\top \boldsymbol{\eta}_t = \boldsymbol{a}_t^\top \boldsymbol{x} && \forall t\in[T], \\
& \mathbf{1}^\top \boldsymbol{\gamma}_t = 1 && \forall t\in[T], \\
& \boldsymbol{\eta}_t, \boldsymbol{\gamma}_t \in \mathbb{R}_+^{L_t} && \forall t\in[T], \\
& \boldsymbol{x}\in\mathcal{X}.
\end{aligned}
\tag{14}
$$

Moreover, as the result of Theorem 3, it provides an improvement over the basic information model of Problem (9), *i.e.*, $\hat{Z} \le Z_0 \le \bar{Z}_0$. The approximation of the empirical average crowdsourcing cost is exact if there exists an optimal solution $\boldsymbol{x}$ of Problem (4) such that

$$\boldsymbol{a}_t^\top \boldsymbol{x} \in \{u_{t\ell} | \ell\in[L_t]\} \qquad \forall t\in[T].$$

The reduced information model allows us to characterize the distribution of the underlying random variable, which would enable us to formulate a robust model to mitigate the uncertainty associated with the data-driven empirical optimization problem. The following section will show how we can develop data-driven robust optimization models using the basic and generalized reduced information models. Our numerical tests based on Solomon's data sets for VRPs in Section 5 show that the basic reduced information model performs quite well. Hence, because of its simplicity, we will first focus on the basic reduced information approach for solving the robust delivery workforce management problem under uncertainty. We will also discuss extending the result to the generalized reduced information model.

4. Data-driven robust model

As we do not know how future uncertainty would evolve, the solution to the empirical optimization problem may not necessarily perform as well when the actual uncertainty is realized in the future. This overfitting phenomenon is known as the *optimizer's curse* (Smith and Winkler 2006). To build a data-driven robust model, we first focus on the basic reduced information model where the true distribution of the random basic reduced information vector, $\tilde{\boldsymbol{z}} \sim \mathbb{P}^\star$, $\mathbb{P}^\star \in \mathcal{P}_0(\mathcal{Z})$ is unobservable to the decision maker.

Because the empirical distribution, $\hat{\mathbb{P}}$, is not the true distribution, $\mathbb{P}^\star$, the empirical optimization models would often yield inferior solutions in out-of-sample performance evaluations. To overcome this issue, Mohajerin Esfahani and Kuhn (2018) and Gao and Kleywegt (2022) have proposed a data-driven robust optimization model with an ambiguity set that is characterized by a type-1

Wasserstein metric $\Delta(\mathbb{P},\hat{\mathbb{P}})$, which evaluates the statistical distance between a candidate distribution $\mathbb{P}$ from the empirical distribution, $\hat{\mathbb{P}}$. In particular, suppose $\mathbb{P}^\Omega$ denotes the distribution that governs the distribution of the independent samples $\tilde{\boldsymbol{z}}^1,\ldots,\tilde{\boldsymbol{z}}^\Omega$ drawn from $\mathbb{P}^\star$ for which the empirical distribution $\hat{\mathbb{P}}$ is constructed. Then under some light-tall distribution assumption, $\mathbb{P}^\Omega\left[\Delta(\mathbb{P}^\star,\hat{\mathbb{P}})>r\right]$ would diminish rapidly to zero as the statistical distance r increases (Fournier and Guillin 2015). The robust optimization problem for the basic reduced information model is as follows,

$$
\begin{aligned}
\bar{Z}_r = \min\ & \boldsymbol{w}^\top\boldsymbol{x} + \sup_{\substack{\mathbb{P}\in\mathcal{P}_0(\mathcal{Z}):\\ \Delta(\mathbb{P},\hat{\mathbb{P}})\le r}} \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} g_t(\boldsymbol{x},\tilde{\boldsymbol{z}}_t)\right]\\
\text{s.t.}\ & \boldsymbol{x}\in\mathcal{X},
\end{aligned}
$$

where in practice, the size parameter r is determined via cross-validation to achieve better out-of-sample performance than the non-robust optimization model.

Robust satisficing

Instead of restricting the distribution to within the vicinity of the empirical distribution, in addressing the issues of robustness in data-driven optimization problems, Long et al. (2022) propose a robust satisficing model specified by a target $\tau\ge\bar{Z}_0$. In particular, the robust satisficing model for the problem is as follows

$$
\begin{aligned}
\kappa_\tau = \min\ & k\\
\text{s.t.}\ & \boldsymbol{w}^\top\boldsymbol{x} + \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]} g_t(\boldsymbol{x},\tilde{\boldsymbol{z}}_t)\right] \le \tau + k\Delta(\mathbb{P},\hat{\mathbb{P}})\ \forall\mathbb{P}\in\mathcal{P}_0(\mathcal{Z})\\
& \boldsymbol{x}\in\mathcal{X}, k\ge 0.
\end{aligned}
\tag{15}
$$

To obtain a tractable model for our robust satisficing problem, we propose the following *time additive Wasserstein metric*,

$$
\Delta(\mathbb{P},\hat{\mathbb{P}}) := \inf_{\mathbb{Q}\in\mathcal{P}_0(\mathcal{Z}^2)}\left\{\mathbb{E}_{\mathbb{Q}}\left[\sum_{t\in[T]}\frac{1}{\bar{\tilde{z}}_t}||\tilde{\boldsymbol{z}}_t-\tilde{\boldsymbol{\xi}}_t||_\infty\right]\ \middle|\ (\tilde{\boldsymbol{z}},\tilde{\boldsymbol{\xi}})\sim\mathbb{Q}, \tilde{\boldsymbol{z}}\sim\mathbb{P}, \tilde{\boldsymbol{\xi}}\sim\hat{\mathbb{P}}\right\},
$$

where $\|\cdot\|_\infty$ is the L_∞-norm.

Since the basic reduced information model is a conservative approximation of the actual problem, the objective of the robust satisficing problem κ_τ can be associated with how well the actual expected cost, when evaluated on the true unobservable distribution, would not excessively exceed the threshold τ. More importantly, compared to specifying the size of the ambiguity set of the robust optimization model, it is more interpretable and intuitive to specify the target parameter τ of the robust satisficing model using $\bar{Z}_0$ as the reference that the target should exceed by.

PROPOSITION 1. *Under the time additive Wasserstein metric, the robust satisficing problem* (15) *is equivalent to the following robust optimization problem*

$$
\begin{aligned}
\kappa_\tau = \min\ & k \\
\text{s.t.}\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \sum_{t \in [T]} \sup_{\boldsymbol{z}_t \in \mathcal{Z}_t} \left\{ g_t(\boldsymbol{x}, \boldsymbol{z}_t) - \frac{k}{\bar{z}_t} \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty \right\} \le \tau \\
& \boldsymbol{x} \in \mathcal{X}, k \ge 0.
\end{aligned} \tag{16}
$$

Proof of Proposition 1. Based on the definition of the Wasserstein metric, we can rewrite the first group of constraints in Problem (15) as

$$
\boldsymbol{w}^\top \boldsymbol{x} + \mathbb{E}_{\mathbb{Q}} \left[\sum_{t \in [T]} \left(g_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}_t) - \frac{k}{\bar{z}_t} ||\tilde{\boldsymbol{z}}_t - \tilde{\boldsymbol{\xi}}_t||_\infty \right) \right] \le \tau \qquad \forall \mathbb{Q} \in \mathcal{P}_0(\mathcal{Z}^2) : (\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{\xi}}) \sim \mathbb{Q}, \tilde{\boldsymbol{\xi}} \sim \hat{\mathbb{P}},
$$

or equivalently

$$
\boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \mathbb{E}_{\mathbb{P}^\omega} \left[\sum_{t \in [T]} \left(g_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}_t) - \frac{k}{\bar{z}_t} ||\tilde{\boldsymbol{z}}_t - \boldsymbol{z}_t^\omega||_\infty \right) \right] \le \tau \qquad \forall \mathbb{P}^\omega \in \mathcal{P}_0(\mathcal{Z}).
$$

Since $\mathcal{P}_0(\mathcal{Z})$ contains all Dirac distributions whose unit mass concentrates on any $\boldsymbol{z} \in \mathcal{Z}$, we can also express this as a robust constraint as follows,

$$
\begin{aligned}
& \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \sup_{\boldsymbol{z} \in \mathcal{Z}} \left\{ \sum_{t \in [T]} \left(g_t(\boldsymbol{x}, \boldsymbol{z}_t) - \frac{k}{\bar{z}_t} ||\boldsymbol{z}_t - \boldsymbol{z}_t^\omega||_\infty \right) \right\} \le \tau \\
\Longleftrightarrow\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \sum_{t \in [T]} \sup_{\boldsymbol{z}_t \in \mathcal{Z}_t} \left\{ g_t(\boldsymbol{x}, \boldsymbol{z}_t) - \frac{k}{\bar{z}_t} ||\boldsymbol{z}_t - \boldsymbol{z}_t^\omega||_\infty \right\} \le \tau.
\end{aligned}
$$

□

Note that the basic reduced crowdsourcing cost function, $g_t(\boldsymbol{x}, \boldsymbol{z}_t)$ for $t \in [T]$ is determined by solving the linear optimization problem (7) after the realization of $\boldsymbol{z}_t$. Hence, its optimal recourse $\boldsymbol{y}_t$ is a mapping of $\boldsymbol{z}_t$, and we can model Problem (16) as an adaptive robust optimization problem as follows:

$$
\begin{aligned}
\min\ & k \\
\text{s.t.}\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \sum_{t \in [T]} \sup_{\boldsymbol{z}_t \in \mathcal{Z}_t} \left\{ \boldsymbol{p}^\top \boldsymbol{y}_t(\boldsymbol{z}_t) - \frac{k}{\bar{z}_t} \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty \right\} \le \tau \\
& \mathbf{1}^\top \boldsymbol{y}_t(\boldsymbol{z}_t) \ge \mathbf{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} && \forall \boldsymbol{z}_t \in \mathcal{Z}_t, t \in [T], \\
& \mathbf{0} \le \boldsymbol{y}_t(\boldsymbol{z}_t) \le \boldsymbol{z}_t && \forall \boldsymbol{z}_t \in \mathcal{Z}_t, t \in [T], \\
& \boldsymbol{y}_t : \mathbb{R}^N \to \mathbb{R} && \forall t \in [T], \\
& \boldsymbol{x} \in \mathcal{X}, k \ge 0.
\end{aligned}
$$

Adaptive robust optimization problems are generally computationally challenging problems. Hence, these problems are often solved approximately by replacing the recourse decision $\boldsymbol{y}_t$ with linear decision rules or affine recourse adaptations (Ben-Tal et al. 2004, Chen et al. 2020). However, because the second stage optimization problem (7) does not have complete recourse, such an approximation may not obtain a feasible solution of the robust satisficing problem (Long et al. 2022) for any reasonable chosen target, $\tau > \bar{Z}_0$. Hence, it is surprising that we can model the robust satisficing problem (16) exactly as the following deterministic optimization problem.

THEOREM 4. *The robust satisficing problem* (16) *is equivalent to the following deterministic optimization problem*

$$
\begin{aligned}
\kappa_\tau = \min\ & k \\
\text{s.t.}\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega}\sum_{\omega\in[\Omega]}\sum_{t\in[T]} r_t^\omega \le \tau, \\
& r_t^\omega \ge \bar{z}_t \phi_{tn}^\omega + (\boldsymbol{\theta}_{tn}^\omega)^\top \boldsymbol{z}_t^\omega - \boldsymbol{a}_t^\top \boldsymbol{x} p_n \quad \forall \omega\in[\Omega], t\in[T], n\in[N], \\
& \bar{z}_t \|\boldsymbol{\theta}_{tn}^\omega\|_1 \le k \quad \forall \omega\in[\Omega], t\in[T], n\in[N], \\
& \mathbf{1}\phi_{tn}^\omega + \boldsymbol{\theta}_{tn}^\omega \ge \boldsymbol{q}_n \quad \forall \omega\in[\Omega], t\in[T], n\in[N], \\
& \phi_{tn}^\omega \ge 0, \boldsymbol{\theta}_{tn}^\omega \in \mathbb{R}^N \quad \forall \omega\in[\Omega], t\in[T], n\in[N], \\
& r_t^\omega \ge 0 \quad \forall \omega\in[\Omega], t\in[T], \\
& \boldsymbol{x}\in\mathcal{X}, k\ge 0,
\end{aligned} \tag{17}
$$

where the vector $\boldsymbol{q}_n \in \mathbb{R}^N$, $n\in[N]$ *has elements*

$$
q_{nm} = \begin{cases} p_m & \text{if } m\in[n-1] \\ p_n & \text{otherwise} \end{cases} \qquad \forall m\in[N].
$$

If the empirical optimization problem (9) *is solvable, then the robust satisficing problem is also feasible for all* $\tau \ge \bar{Z}_0$, *and* $\kappa_\tau \in [0, \bar{\kappa}]$, *where*

$$
\bar{\kappa} = (\mathbf{1}^\top \boldsymbol{p}) \max_{t\in[T]} \{\bar{z}_t\}.
$$

Proof of Theorem 4. From Proposition 1, we can express the robust satisficing model as

$$
\begin{aligned}
\kappa_\tau = \min\ & k \\
\text{s.t.}\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega}\sum_{\omega\in[\Omega]}\sum_{t\in[T]} r_t^\omega \le \tau \\
& r_t^\omega \ge \sup_{\boldsymbol{z}_t\in\mathcal{Z}_t} \left\{ g_t(\boldsymbol{x},\boldsymbol{z}_t) - \frac{k}{\bar{z}_t}\|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty \right\} \quad \forall\omega\in[\Omega], t\in[T], \\
& \boldsymbol{x}\in\mathcal{X}, k\ge 0.
\end{aligned}
$$

Recall that p_n is non-decreasing in $n \in [N]$, and let $p_0 = 0$. By strong duality of linear optimization, we have

$$
\begin{aligned}
g_t(\boldsymbol{x}, \boldsymbol{z}_t) &= \min_{\boldsymbol{y}_t} \left\{ \boldsymbol{p}^\top \boldsymbol{y}_t \mid \boldsymbol{1}^\top \boldsymbol{y}_t \geq \boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x}, \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\} \\
&= \max_{\lambda \geq 0} \min_{\boldsymbol{y}_t} \left\{ (\boldsymbol{p} - \lambda \boldsymbol{1})^\top \boldsymbol{y}_t + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \mid \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\} \\
&= \max \left\{ \begin{array}{l} \max\limits_{\lambda \in [0, p_N]} \min\limits_{\boldsymbol{y}_t} \left\{ (\boldsymbol{p} - \lambda \boldsymbol{1})^\top \boldsymbol{y}_t + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \mid \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\}, \\ \qquad \max\limits_{\lambda \geq p_N} \min\limits_{\boldsymbol{y}_t} \left\{ (\boldsymbol{p} - \lambda \boldsymbol{1})^\top \boldsymbol{y}_t + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \mid \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\} \end{array} \right\} \\
&= \max \left\{ \begin{array}{l} \max\limits_{\lambda \in [0, p_N]} \min\limits_{\boldsymbol{y}_t} \left\{ (\boldsymbol{p} - \lambda \boldsymbol{1})^\top \boldsymbol{y}_t + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \mid \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\}, \\ \qquad \max\limits_{\lambda \geq p_N} \left\{ \boldsymbol{p}^\top \boldsymbol{z}_t - \lambda (\boldsymbol{1}^\top \boldsymbol{z}_t + \boldsymbol{a}_t^\top \boldsymbol{x} - \boldsymbol{1}^\top \boldsymbol{z}_t) \right\} \end{array} \right\} \\
&= \max \left\{ \begin{array}{l} \max\limits_{\lambda \in [0, p_N]} \min\limits_{\boldsymbol{y}_t} \left\{ (\boldsymbol{p} - \lambda \boldsymbol{1})^\top \boldsymbol{y}_t + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \mid \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\}, \\ \qquad \max\limits_{\lambda = p_N} \left\{ \boldsymbol{p}^\top \boldsymbol{z}_t - \lambda (\boldsymbol{a}_t^\top \boldsymbol{x}) \right\} \end{array} \right\} \qquad (\text{since } \boldsymbol{a}_t^\top \boldsymbol{x} \geq 0) \\
&= \max_{\lambda \in [0, p_N]} \min_{\boldsymbol{y}_t} \left\{ (\boldsymbol{p} - \lambda \boldsymbol{1})^\top \boldsymbol{y}_t + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \mid \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\} \\
&= \max_{n \in [N]} \max_{\lambda \in [p_{n-1}, p_n]} \min_{\boldsymbol{y}_t} \left\{ \sum_{m \in [N]} (p_m - \lambda) y_{tm} + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \mid \boldsymbol{0} \leq \boldsymbol{y}_t \leq \boldsymbol{z}_t \right\} \\
&= \max_{n \in [N]} \max_{\lambda \in [p_{n-1}, p_n]} \left\{ \sum_{m \in [n-1]} (p_m - \lambda) z_{tm} + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \right\} \\
&= \max_{n \in [N]} \max_{\lambda \in \{p_{n-1}, p_n\}} \left\{ \sum_{m \in [n-1]} (p_m - \lambda) z_{tm} + \lambda \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) \right\} \\
&= \max \left\{ 0, \max_{n \in [N]} \left\{ \sum_{m \in [n-1]} (p_m - p_n) z_{tm} + \left(\boldsymbol{1}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} \right) p_n \right\} \right\} \\
&= \max \left\{ 0, \max_{n \in [N]} \left\{ \boldsymbol{q}_n^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} p_n \right\} \right\}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& r_t^\omega \geq \sup_{\boldsymbol{z}_t \in \mathcal{Z}_t} \left\{ g_t(\boldsymbol{x}, \boldsymbol{z}_t) - k \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t \right\} \\
\iff & \left\{ \begin{array}{ll} r_t^\omega \geq \boldsymbol{q}_n^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} p_n - k \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t & \forall \boldsymbol{z}_t \in \mathcal{Z}_t, n \in [N] \\ r_t^\omega \geq -k \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t & \forall \boldsymbol{z}_t \in \mathcal{Z}_t. \end{array} \right.
\end{aligned}
$$

Observe that since $\boldsymbol{z}_t^\omega \in \mathcal{Z}_t$, we have

$$
\begin{aligned}
& r_t^\omega \geq -k \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t \quad \forall \boldsymbol{z}_t \in \mathcal{Z}_t \\
\iff & r_t^\omega \geq 0.
\end{aligned}
$$

We next show that for any $\boldsymbol{q} \in \mathbb{R}^N$,

$$\begin{aligned}
&\sup_{\boldsymbol{z}_t \in \mathcal{Z}_t} \left\{ \boldsymbol{q}^\top \boldsymbol{z}_t - k \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t \right\} \\
=& \sup_{\substack{\boldsymbol{z}_t \geq \boldsymbol{0} \\ \boldsymbol{1}^\top \boldsymbol{z}_t \leq \bar{z}_t}} \left\{ \boldsymbol{q}^\top \boldsymbol{z}_t - k \left\|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\right\|_\infty / \bar{z}_t \right\} \\
=& \sup_{\substack{\boldsymbol{z}_t \geq \boldsymbol{0} \\ \boldsymbol{1}^\top \boldsymbol{z}_t \leq \bar{z}_t}} \inf_{\|\boldsymbol{\theta}\|_1 \leq k/\bar{z}_t} \left\{ \boldsymbol{q}^\top \boldsymbol{z}_t - \boldsymbol{\theta}^\top \left(\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\right) \right\} \\
=& \inf_{\|\boldsymbol{\theta}\|_1 \leq k/\bar{z}_t} \sup_{\substack{\boldsymbol{z}_t \geq \boldsymbol{0} \\ \boldsymbol{1}^\top \boldsymbol{z}_t \leq \bar{z}_t}} \left\{ \boldsymbol{q}^\top \boldsymbol{z}_t - \boldsymbol{\theta}^\top \left(\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\right) \right\} \\
=& \inf_{\|\boldsymbol{\theta}\|_1 \leq k/\bar{z}_t} \sup_{\substack{\boldsymbol{z}_t \geq \boldsymbol{0} \\ \boldsymbol{1}^\top \boldsymbol{z}_t \leq \bar{z}_t}} \left\{ \left(\boldsymbol{q} - \boldsymbol{\theta}\right)^\top \boldsymbol{z}_t + \boldsymbol{\theta}^\top \boldsymbol{z}_t^\omega \right\} \\
=& \inf_{\substack{\|\boldsymbol{\theta}\|_1 \leq k/\bar{z}_t \\ \boldsymbol{1}\phi + \boldsymbol{\theta} \geq \boldsymbol{q}, \phi \geq 0}} \left\{ \bar{z}_t \phi + \boldsymbol{\theta}^\top \boldsymbol{z}_t^\omega \right\}.
\end{aligned}$$

Therefore, for all $t \in [T], \omega \in [\Omega], n \in [N]$,

$$\begin{aligned}
& r_t^\omega \geq \boldsymbol{q}_{tn}^\top \boldsymbol{z}_t - \boldsymbol{a}_t^\top \boldsymbol{x} p_n - k \|\boldsymbol{z}_t - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t \;\; \forall \boldsymbol{z}_t \in \mathcal{Z}_t \\
\iff & \begin{cases} r_t^\omega \geq \bar{z}_t \phi_{tn}^\omega + \left(\boldsymbol{\theta}_{tn}^\omega\right)^\top \boldsymbol{z}_t^\omega - \boldsymbol{a}_t^\top \boldsymbol{x} p_n \\ \bar{z}_t \left\|\boldsymbol{\theta}_{tn}^\omega\right\|_1 \leq k \\ \boldsymbol{1}\phi_{tn}^\omega + \boldsymbol{\theta}_{tn}^\omega \geq \boldsymbol{q}_{tn} \\ \text{for some } \phi_{tn}^\omega \in \mathbb{R}_+, \boldsymbol{\theta}_{tn}^\omega \in \mathbb{R}^N. \end{cases}
\end{aligned}$$

Finally, to show feasibility for $\tau \geq \bar{Z}_0$, we consider a restricted feasible set of Problem (17) with $\phi_{tn}^\omega = 0$ and $\boldsymbol{q}_n = \boldsymbol{\theta}_{tn}^\omega$ for all $n \in [N], \omega \in [\Omega], t \in [T]$. Hence, we can express the restricted feasible set as

$$\begin{aligned}
\mathcal{Q} &= \left\{ \boldsymbol{x} \in \mathcal{X} \;\middle|\; \begin{array}{ll} \exists r_t^\omega \geq 0 \;\forall \omega \in [\Omega], t \in [T]: & \\ \boldsymbol{w}^\top \boldsymbol{x} + \dfrac{1}{\Omega} \displaystyle\sum_{\omega \in [\Omega]} \sum_{t \in [T]} r_t^\omega \leq \tau & \\ r_t^\omega \geq \boldsymbol{q}_n^\top \boldsymbol{z}_t^\omega - \boldsymbol{a}_t^\top \boldsymbol{x} p_n & \forall \omega \in [\Omega], t \in [T], n \in [N] \end{array} \right\} \\
&= \left\{ \boldsymbol{x} \in \mathcal{X} \;\middle|\; \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \sum_{t \in [T]} \max \left\{ 0, \max_{n \in [N]} \left\{ \boldsymbol{q}_n^\top \boldsymbol{z}_t^\omega - \boldsymbol{a}_t^\top \boldsymbol{x} p_n \right\} \right\} \leq \tau \right\} \\
&= \left\{ \boldsymbol{x} \in \mathcal{X} \;\middle|\; \boldsymbol{w}^\top \boldsymbol{x} + \mathbb{E}_{\hat{\mathbb{P}}} \left[\sum_{t \in [T]} g_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}_t) \right] \leq \tau \right\},
\end{aligned}$$

so that any $\boldsymbol{x} \in \mathcal{Q}$ would be feasible in Problem (17). Hence, if the empirical optimization problem (9) is solvable, then its solution would be feasible in the robust satisficing problem for all $\tau \geq \bar{Z}_0$. Moreover, when $k \geq \kappa_\tau \geq \bar{z}_t \|\boldsymbol{q}_n\|_1$,

$$\begin{aligned}
& \boldsymbol{q}_n^\top \boldsymbol{z} - k \|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t \\
=& \boldsymbol{q}_n^\top \boldsymbol{z}_t^\omega + \boldsymbol{q}_n^\top (\boldsymbol{z} - \boldsymbol{z}_t^\omega) - k \|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t \\
\leq& \boldsymbol{q}_n^\top \boldsymbol{z}_t^\omega + \|\boldsymbol{q}_n\|_1 \|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty - k \|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_t \\
\leq& \boldsymbol{q}_n^\top \boldsymbol{z}_t^\omega.
\end{aligned}$$

On the other hand,

$$\sup_{\boldsymbol{z}_t\in\mathcal{Z}_t}\{\boldsymbol{q}_n^\top\boldsymbol{z}_t - k\|\boldsymbol{z}_t-\boldsymbol{z}_t^\omega\|_\infty/\bar{z}_t\}\geq \boldsymbol{q}_n^\top\boldsymbol{z}_t^\omega$$

since $\boldsymbol{z}_t^\omega\in\mathcal{Z}_t$. Therefore, when $k\geq\kappa_\tau$,

$$\begin{aligned}
& r_t^\omega \geq \sup_{\boldsymbol{z}_t\in\mathcal{Z}_t}\{g_t(\boldsymbol{x},\boldsymbol{z}_t)-k\|\boldsymbol{z}_t-\boldsymbol{z}_t^\omega\|_\infty/\bar{z}_t\}\\
\iff & \begin{cases} r_t^\omega\geq \boldsymbol{q}_n^\top\boldsymbol{z}_t-\boldsymbol{a}_t^\top\boldsymbol{x}p_n-k\|\boldsymbol{z}_t-\boldsymbol{z}_t^\omega\|_\infty/\bar{z}_t\ \forall\boldsymbol{z}_t\in\mathcal{Z}_t, n\in[N]\\ r_t^\omega\geq 0\end{cases}\\
\iff & \begin{cases} r_t^\omega\geq \boldsymbol{q}_n^\top\boldsymbol{z}_t^\omega-\boldsymbol{a}_t^\top\boldsymbol{x}p_n\ \forall n\in[N]\\ r_t^\omega\geq 0\end{cases}\\
\iff & r_t^\omega\geq g_t(\boldsymbol{x},\tilde{\boldsymbol{z}}_t^\omega),
\end{aligned}$$

indicating the robust satisficing problem (16) coincides with the empirical optimization problem (9), thus the κ_τ can never be larger than $\bar{\kappa}$ when for any $\tau\geq\bar{Z}_0$. □

Although the basic robust satisficing model is feasible for any $\tau\geq\bar{Z}_0$, since $\bar{Z}_0\geq\hat{Z}$, it would not be feasible for $\tau\in[\hat{Z},\bar{Z}_0)$, which can be an issue if $\bar{Z}_0$ is significantly larger than $\hat{Z}$. To reduce the conservativeness, we have to consider solving the generalized reduced information model that would achieve $Z_0=\hat{Z}$. A computationally viable approach is to use the optimal solution of Problem (4), $\boldsymbol{x}$ and consider $L_t=2$ breakpoints with $u_{t2}=\boldsymbol{a}_t^\top\boldsymbol{x}\ \forall t\in[T]$. Based on Theorem 3, we have $Z_0=\hat{Z}$. Next, we show how to extend the robust satisficing model to incorporate the generalized reduced information.

Robust model with generalized reduced information

We now extend to the generalized reduced information model with $\tilde{\boldsymbol{z}}=(\tilde{\boldsymbol{z}}_t^\ell)_{t\in[T],\ell\in[L_t]}$. To obtain a computationally tractable model, we consider the following $\boldsymbol{\gamma}$*-weighted time-addictive* Wasserstein metric,

$$\Delta_{\boldsymbol{\gamma}}(\mathbb{P},\hat{\mathbb{P}}):=\inf_{\mathbb{Q}\in\mathcal{P}_0(\mathcal{Z}^2)}\left\{\mathbb{E}_{\mathbb{Q}}\left[\sum_{t\in[T]}\sum_{\ell\in[L_t]}\frac{1}{\bar{z}_{t\ell}}\gamma_{t\ell}||\tilde{\boldsymbol{z}}_t^\ell-\tilde{\boldsymbol{\xi}}_t^\ell||_\infty\right]\ \middle|\ (\tilde{\boldsymbol{z}},\tilde{\boldsymbol{\xi}})\sim\mathbb{Q},\tilde{\boldsymbol{z}}\sim\mathbb{P},\tilde{\boldsymbol{\xi}}\sim\hat{\mathbb{P}}\right\}.$$

Accordingly, we propose the following robust satisficing model associated with the generalized reduced information model,

$$\begin{aligned}
\kappa_\tau = \min\ & k\\
\text{s.t.}\ & \boldsymbol{w}^\top\boldsymbol{x}+\mathbb{E}_{\mathbb{P}}\left[\sum_{t\in[T]}\sum_{\ell\in[L_t]}h_t(\gamma_{t\ell},\eta_{t\ell}-u_{t\ell}\gamma_{t\ell},\tilde{\boldsymbol{z}}_t^\ell)\right]\leq\tau+k\Delta_{\boldsymbol{\gamma}}(\mathbb{P},\hat{\mathbb{P}})\ \forall\mathbb{P}\in\mathcal{P}_0(\mathcal{Z}),\\
& \mathbf{1}^\top\boldsymbol{\eta}_t=\boldsymbol{a}_t^\top\boldsymbol{x} && \forall t\in[T],\\
& \mathbf{1}^\top\boldsymbol{\gamma}_t=1 && \forall t\in[T],\\
& \boldsymbol{\eta}_t,\boldsymbol{\gamma}_t\in\mathbb{R}_+^{L_t} && \forall t\in[T],\\
& \boldsymbol{x}\in\mathcal{X}.
\end{aligned}\tag{18}$$

Observe that when $\eta_{t1} = \boldsymbol{a}_t^\top \boldsymbol{x}$, and $\gamma_{t1} = 1$, Problem (18) will recover the solution of the basic reduced information robust satisficing problem (15). Hence, Problem (18) has a lower objective value, which leads to lower probabilities of violating the target at various levels in the following result.

We note that, unlike the basic reduced information model, we cannot solve the generalized robust satisficing model as a single convex optimization problem even if $\mathcal{X}$ is a convex set. Nevertheless, the following result shows how we can solve the more general robust model via a binary search on a bounded interval of k.

THEOREM 5. *The robust satisficing problem* (18) *is equivalent to the following deterministic optimization problem*

$$\begin{aligned} \kappa_\tau = \min\ & k \\ \text{s.t.}\ & \rho(k) \le \tau \\ & k \ge 0, \end{aligned} \tag{19}$$

where

$$\begin{aligned} \rho(k) = \min\ & \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \sum_{t \in [T]} \sum_{\ell \in [L_t]} r_t^{\ell\omega} \\ \text{s.t.}\ & r_t^{\ell\omega} \ge \alpha_{tn}^{\ell\omega}(k)\gamma_{t\ell} - (\eta_{t\ell} - u_{t\ell}\gamma_{t\ell})p_n && \forall \omega \in [\Omega], t \in [T], n \in [N], \ell \in [L_t], \\ & \mathbf{1}^\top \boldsymbol{\eta}_t = \boldsymbol{a}_t^\top \boldsymbol{x} && \forall t \in [T], \\ & \mathbf{1}^\top \boldsymbol{\gamma}_t = 1 && \forall t \in [T], \\ & \boldsymbol{\eta}_t \in \mathbb{R}^{L_t}, \boldsymbol{\gamma}_t \in \mathbb{R}_+^{L_t} && \forall t \in [T], \\ & r_t^{\ell\omega} \ge 0 && \forall \omega \in [\Omega], t \in [T], \ell \in [L_t], \\ & \boldsymbol{x} \in \mathcal{X}, \end{aligned} \tag{20}$$

and

$$\begin{aligned} \alpha_{tn}^{\ell\omega}(k) = \max\ & \boldsymbol{q}_n^\top \boldsymbol{z} - \tfrac{k}{\bar{z}_{t\ell}} \|\boldsymbol{z} - \boldsymbol{z}_t^{\ell\omega}\|_\infty \\ \text{s.t}\ & \mathbf{1}^\top \boldsymbol{z} \le \bar{z}_{t\ell} \\ & \boldsymbol{z} \ge \mathbf{0}. \end{aligned}$$

Moreover, if the empirical optimization problem (14) *is solvable, then for all* $\tau \ge Z_0$*, the robust satisficing problem is also feasible, and* $\kappa_\tau \in [0, \bar{\kappa}]$ *for the same* $\bar{\kappa}$ *defined in Theorem 4.*

Proof of Theorem 5. Observe that

$$\boldsymbol{w}^\top \boldsymbol{x} + \mathbb{E}_{\mathbb{P}} \left[\sum_{t \in [T]} \sum_{\ell \in [L_t]} h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \tilde{\boldsymbol{z}}_t^\ell) \right] \le \tau + k\Delta_{\boldsymbol{\gamma}}(\mathbb{P}, \hat{\mathbb{P}}) \quad \forall \mathbb{P} \in \mathcal{P}_0(\mathcal{Z})$$

is equivalent to

$$\boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega \in [\Omega]} \mathbb{E}_{\mathbb{P}^\omega} \left[\sum_{t \in [T]} \sum_{\ell \in [L_t]} \left(h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \tilde{\boldsymbol{z}}_t^\ell) - \frac{k\gamma_{t\ell}}{\bar{z}_{t\ell}} ||\tilde{\boldsymbol{z}}_t^\ell - \boldsymbol{z}_t^{\ell\omega}||_\infty \right) \right] \le \tau \quad \forall \mathbb{P}^\omega \in \mathcal{P}_0(\mathcal{Z}).$$

Since $\mathcal{P}_0(\mathcal{Z})$ contains all Dirac distributions whose unit mass concentrates on any $\boldsymbol{z} \in \mathcal{Z}$, and given that the support set is breakpoints separable, we have the following equivalent constraint,

$$\boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega\in[\Omega]} \sum_{t\in[T]} \sum_{\ell\in[L_t]} \sup_{\boldsymbol{z}_t^\ell \in \mathcal{Z}_t^\ell} \left\{ h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \boldsymbol{z}_t^\ell) - \frac{k\gamma_{t\ell}}{\bar{z}_{t\ell}} ||\boldsymbol{z}_t^\ell - \boldsymbol{z}_t^{\ell\omega}||_\infty \right\} \le \tau.$$

Therefore Problem (18) has the following equivalent formulation

$$\begin{aligned}
&\min\ k \\
&\text{s.t.}\ \ \boldsymbol{w}^\top \boldsymbol{x} + \frac{1}{\Omega} \sum_{\omega\in[\Omega]} \sum_{t\in[T]} \sum_{\ell\in[L_t]} r_t^{\ell\omega} \le \tau \\
&\qquad r_t^{\ell\omega} \ge \sup_{\boldsymbol{z}\in\mathcal{Z}_t^\ell} \left\{ h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \boldsymbol{z}_t^\ell) - k\gamma_{t\ell} \|\boldsymbol{z} - \boldsymbol{z}_t^{\ell\omega}\|_\infty / \bar{z}_{t\ell} \right\} && \forall \omega \in [\Omega], t \in [T], n \in [N], \ell \in [L_t], \\
&\qquad \mathbf{1}^\top \boldsymbol{\eta}_t = \boldsymbol{a}_t^\top \boldsymbol{x} && \forall t \in [T], \\
&\qquad \mathbf{1}^\top \boldsymbol{\gamma}_t = 1 && \forall t \in [T], \\
&\qquad \boldsymbol{\eta}_t \in \mathbb{R}^{L_t}, \boldsymbol{\gamma}_t \in \mathbb{R}_+^{L_t} && \forall t \in [T], \\
&\qquad \boldsymbol{x} \in \mathcal{X}.
\end{aligned} \tag{21}$$

Consequently, following the similar analysis in Theorem 4, we have

$$\begin{aligned}
& r_t^{\ell\omega} \ge \sup_{\boldsymbol{z}_t^\ell \in \mathcal{Z}_t^\ell} \left\{ h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \boldsymbol{z}_t^\ell) - k\gamma_{t\ell} \|\boldsymbol{z}_t^\ell - \boldsymbol{z}_t^{\ell\omega}\|_\infty / \bar{z}_{t\ell} \right\} \\
\iff & \begin{cases} r_t^{\ell\omega} \ge \sup\limits_{\boldsymbol{z}\in\mathcal{Z}_t^\ell} \left\{ \boldsymbol{q}_n^\top \boldsymbol{z} - k\|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_{t\ell} \right\} \gamma_{t\ell} - (\eta_{t\ell} - u_{t\ell}\gamma_{t\ell})\, p_n \ \forall n \in [N] \\ r_t^{\ell\omega} \ge \sup\limits_{\boldsymbol{z}\in\mathcal{Z}_t^\ell} \left\{ -k\|\boldsymbol{z} - \boldsymbol{z}_t^{\ell\omega}\|_\infty / \bar{z}_{t\ell} \right\} \gamma_{t\ell} \end{cases} \\
\iff & \begin{cases} r_t^{\ell\omega} \ge \alpha_{tn}^{\ell\omega}(k)\gamma_{t\ell} - (\eta_{t\ell} - u_{t\ell}\gamma_{t\ell})\, p_n & \forall n \in [N] \\ r_t^{\ell\omega} \ge 0. \end{cases}
\end{aligned}$$

Observe that since $\boldsymbol{z}_t^{\ell\omega} \in \mathcal{Z}_t^\ell$, we have

$$\alpha_{tn}^{\ell\omega}(k) \ge \boldsymbol{q}_n^\top \boldsymbol{z}_t^{\ell\omega}.$$

However, we note that if $k \ge \bar{z}_{t\ell}\|\boldsymbol{q}_n\|_1$, then

$$\begin{aligned}
& \boldsymbol{q}_n^\top \boldsymbol{z} - k\|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_{t\ell} \\
= & \boldsymbol{q}_n^\top \boldsymbol{z}_t^{\ell\omega} + \boldsymbol{q}_n^\top (\boldsymbol{z} - \boldsymbol{z}_t^\omega) - k\|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_{t\ell} \\
\le & \boldsymbol{q}_n^\top \boldsymbol{z}_t^{\ell\omega} + \|\boldsymbol{q}_n\|_1 \|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty - k\|\boldsymbol{z} - \boldsymbol{z}_t^\omega\|_\infty / \bar{z}_{t\ell} \\
\le & \boldsymbol{q}_n^\top \boldsymbol{z}_t^{\ell\omega},
\end{aligned}$$

which implies $\alpha_{tn}^{\ell\omega}(k) = \boldsymbol{q}_n^\top \boldsymbol{z}_t^{\ell\omega}$. Hence, if $k \geq \bar{\kappa} \geq \bar{z}_{t\ell}(\mathbf{1}^\top \boldsymbol{p}) \geq \bar{z}_{t\ell}\|\boldsymbol{q}_n\|_1$ for all $t \in [T], \ell \in [L_t]$, then the following holds

$$
\begin{aligned}
&\begin{cases}
r_t^{\ell\omega} \geq \alpha_{tn}^{\ell\omega}(k)\gamma_{t\ell} - (\eta_{t\ell} - u_{t\ell}\gamma_{t\ell})\, p_n & \forall n \in [N] \\
r_t^{\ell\omega} \geq 0.
\end{cases} \\
\iff &\begin{cases}
r_t^{\ell\omega} \geq \boldsymbol{q}_n^\top \boldsymbol{z}_t^{\ell\omega}\gamma_{t\ell} - (\eta_{t\ell} - u_{t\ell}\gamma_{t\ell})\, p_n & \forall n \in [N] \\
r_t^{\ell\omega} \geq 0.
\end{cases} \\
\iff &\ r_t^{\ell\omega} \geq h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \boldsymbol{z}_t^{\ell\omega}).
\end{aligned}
$$

Hence, when $k \geq \bar{\kappa}$, the empirical optimization problem (14) is the same as Problem (20) so that $\rho(k) = Z_0$. Therefore, the robust satisficing problem is feasible for $\tau \geq Z_0$, and κ_τ would not exceed $\bar{\kappa}$. □

Observe that if $\mathcal{X}$ is a polyhedron, then Problem (20) would be a linear optimization problem. Consequently, using binary search, we can solve the generalized robust satisficing model via solving a modest number of convex optimization problems.

Statistical justification

We now provide the statistical justification for the robust satisficing model.

THEOREM 6. *Consider the random bidding sets $(\tilde{\mathcal{B}}_1, \ldots, \tilde{\mathcal{B}}_t) \sim \mathbb{Q}^\star$ which generate the random variable associated with the generalized reduced information $\tilde{\boldsymbol{z}} \sim \mathbb{P}^\star$. Let $\mathbb{P}^\Omega$ be the distribution that governs the distribution of independent samples $\tilde{\boldsymbol{z}}^\omega$, $\omega \in \Omega$ drawn from $\mathbb{P}^\star$. The following holds for the optimal solution to Problem* (18)*:*

1. **Confidence guarantee of Fournier and Guillin (2015).** *For any $\tau \geq Z_0$,*

$$
\mathbb{P}^\Omega\left[\boldsymbol{w}^\top\boldsymbol{x} + \mathbb{E}_{\mathbb{Q}^\star}\left[\sum_{t\in[T]} f_t(\boldsymbol{x}, \tilde{\mathcal{B}}_t)\right] > \tau + Tr\kappa_\tau\right] \leq \begin{cases} c_1\exp(-c_2\Omega r^{NL}) & \forall r \in [0,1] \\ 0 & \forall r > 1, \end{cases}
$$

for some positive c_1 and c_2 that depend on $\mathbb{E}_{\mathbb{P}^\star}[\exp(||\tilde{\boldsymbol{z}}||_\infty)]$ and NL, $L = \sum_{t\in[T]} L_t$.

2. **Confidence guarantee for significant target shortfalls.** *For given $\tau \geq Z_0$,*

$$
\mathbb{P}^\Omega\left[\boldsymbol{w}^\top\boldsymbol{x} + \mathbb{E}_{\mathbb{Q}^\star}\left[\sum_{t\in[T]} f_t(\boldsymbol{x}, \tilde{\mathcal{B}}_t)\right] > \tau + Tr\kappa_\tau\right] \leq \begin{cases} \exp\left(-2\Omega\,(r-\mu)^2\right) & \forall r \in [\mu,1] \\ 0 & \forall r > 1, \end{cases}
$$

where

$$
\mu := \mathbb{E}_{(\tilde{\boldsymbol{z}},\tilde{\boldsymbol{\xi}})\sim\mathbb{P}^\star\times\mathbb{P}^\star}\left[\frac{1}{T}\sum_{t\in[T]}\max_{\ell\in[L_t]}\left\{\frac{1}{\bar{z}_{t\ell}}\|\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{\xi}}_t^\ell\|_\infty\right\}\right],
$$

noting that $\mu \leq 1$.

Proof of Theorem 6. For any optimal solution $\boldsymbol{x}, \boldsymbol{\gamma}, \boldsymbol{\eta}, \kappa_\tau$ to Problem (18) given τ, we have from Theorem 3,

$$\mathbb{E}_{\mathbb{Q}^\star}\left[\sum_{t\in[T]} f_t(\boldsymbol{x}, \tilde{\mathcal{B}}_t)\right] \le \mathbb{E}_{\mathbb{P}^\star}\left[\sum_{t\in[T]}\sum_{\ell\in[L_t]} h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \tilde{\boldsymbol{z}}_t^\ell)\right].$$

Hence, for any $r \ge 0$

$$\begin{aligned}
&\mathbb{P}^\Omega\left[\boldsymbol{w}^\top\boldsymbol{x} + \mathbb{E}_{\mathbb{Q}^\star}\left[\sum_{t\in[T]} f_t(\boldsymbol{x}, \tilde{\mathcal{B}}_t)\right] > \tau + Tr\kappa_\tau\right]\\
\le\ &\mathbb{P}^\Omega\left[\boldsymbol{w}^\top\boldsymbol{x} + \mathbb{E}_{\mathbb{Q}^\star}\left[\sum_{t\in[T]}\sum_{\ell\in[L_t]} h_t(\gamma_{t\ell}, \eta_{t\ell} - u_{t\ell}\gamma_{t\ell}, \tilde{\boldsymbol{z}}_t^\ell)\right] > \tau + Tr\kappa_\tau\right]\\
\le\ &\mathbb{P}^\Omega\left[\Delta_{\boldsymbol{\gamma}}(\mathbb{P}^\star, \hat{\mathbb{P}}) > Tr\right] \le \mathbb{P}^\Omega\left[\bar{\Delta}(\mathbb{P}^\star, \hat{\mathbb{P}}) > Tr\right],
\end{aligned}$$

where

$$\bar{\Delta}(\mathbb{P}, \hat{\mathbb{P}}) := \inf_{\mathbb{Q}\in\mathcal{P}_0(\mathcal{Z}^2)}\left\{\mathbb{E}_{\mathbb{Q}}\left[\sum_{t\in[T]}\max_{\ell\in[L_t]}\left\{\frac{1}{\bar{z}_{t\ell}}\|\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{\xi}}_t^\ell\|_\infty\right\}\right] \;\middle|\; (\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{\xi}}) \sim \mathbb{Q}, \tilde{\boldsymbol{z}} \sim \mathbb{P}, \tilde{\boldsymbol{\xi}} \sim \hat{\mathbb{P}}\right\},$$

noting that $\Delta_{\boldsymbol{\gamma}}(\mathbb{P}^\star, \hat{\mathbb{P}}) \le \bar{\Delta}(\mathbb{P}^\star, \hat{\mathbb{P}})$ is due to $\mathbf{1}^\top\boldsymbol{\gamma}_t = 1$ and $\boldsymbol{\gamma}_t \ge \mathbf{0}$. By the definition of $\bar{\Delta}$, we have

$$\mathbb{P}^\Omega\left[\bar{\Delta}(\mathbb{P}^\star, \hat{\mathbb{P}}) > Tr\right] \le \mathbb{P}^\Omega\left[\frac{1}{\Omega}\sum_{\omega\in[\Omega]}\mathbb{E}_{\tilde{\boldsymbol{z}}\sim\mathbb{P}^\star}\left[\sum_{t\in[T]}\max_{\ell\in[L_t]}\left\{\frac{1}{\bar{z}_{t\ell}}\|\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{z}}_t^{\ell\omega}\|_\infty\right\}\right] > Tr\right].$$

Since $\tilde{\boldsymbol{z}}_t^\ell, \tilde{\boldsymbol{z}}_t^{\ell\omega} \in \mathcal{Z}_t$ almost everywhere, we have

$$\begin{aligned}
\|\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{z}}_t^{\ell\omega}\|_\infty &\le \max_{\substack{\boldsymbol{x},\boldsymbol{y}\ge\mathbf{0}\\ \mathbf{1}^\top\boldsymbol{x}\le\bar{z}_{t\ell}, \mathbf{1}^\top\boldsymbol{y}\le\bar{z}_{t\ell}}}\ \max_{n\in[N]}\{|\boldsymbol{e}_n^\top(\boldsymbol{x}-\boldsymbol{y})|\}\\
&\le \max_{\substack{\boldsymbol{x},\boldsymbol{y}\ge\mathbf{0}\\ \|\boldsymbol{x}\|_1\le\bar{z}_{t\ell}, \|\boldsymbol{y}\|_1\le\bar{z}_{t\ell}}}\ \max_{n\in[N]}\left\{\max\{\boldsymbol{e}_n^\top\boldsymbol{x}, \boldsymbol{e}_n^\top\boldsymbol{y}\}\right\} \qquad \text{(since } \boldsymbol{x},\boldsymbol{y}\ge\mathbf{0}\text{)}\\
&\le \max_{n\in[N]}\left\{\max\left\{\max_{\|\boldsymbol{x}\|_1\le\bar{z}_{t\ell}}\{\boldsymbol{e}_n^\top\boldsymbol{x}\}, \max_{\|\boldsymbol{y}\|_1\le\bar{z}_{t\ell}}\{\boldsymbol{e}_n^\top\boldsymbol{y}\}\right\}\right\}\\
&= \max_{n\in[N]}\{\max\{\bar{z}_{t\ell}\|\boldsymbol{e}_n\|_\infty, \bar{z}_{t\ell}\|\boldsymbol{e}_n\|_\infty\}\} = \bar{z}_{t\ell},
\end{aligned}$$

where $\boldsymbol{e}_n$ is the nth unit basis vector. Therefore,

$$\mathbb{P}^\Omega\left[\frac{1}{\Omega}\sum_{\omega\in[\Omega]}\mathbb{E}_{\tilde{\boldsymbol{z}}\sim\mathbb{P}^\star}\left[\sum_{t\in[T]}\max_{\ell\in[L_t]}\left\{\frac{1}{\bar{z}_{t\ell}}\|\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{z}}_t^{\ell\omega}\|_\infty\right\}\right] \le T\right] = 1.$$

Consequently, when $r > 1$, we must have

$$\mathbb{P}^\Omega\left[\bar{\Delta}(\mathbb{P}^\star, \hat{\mathbb{P}}) > Tr\right] = 0.$$

1. We now complete the proof for the first probability bound. Before proving this result, we need to recap the result on the probability bound of the Wasserstein distance in Fournier and Guillin (2015) as follows.

Suppose the actual data-generating distribution $\mathbb{P}^\star$, $\tilde{\boldsymbol{z}} \sim \mathbb{P}^\star$ is a light-tailed distribution such that

$$\mathbb{E}_{\mathbb{P}^\star}[\exp(\|\tilde{\boldsymbol{z}}\|^\alpha)] < \infty \tag{22}$$

for some $\alpha > 1$ and $\mathbb{P}^\Omega$ is the distribution that governs the distribution of independent samples $\hat{\boldsymbol{z}}_1, \ldots, \hat{\boldsymbol{z}}_\Omega$ drawn from $\mathbb{P}^\star$, which constitutes the empirical distribution $\hat{\mathbb{P}}$. Then for any $R \in (0,1)$,

$$\mathbb{P}^\Omega\left[\Delta_1(\mathbb{P}^\star, \hat{\mathbb{P}}) > R\right] \leq c_1 \exp(-c_2 \Omega R^{\max\{n_z, 2\}}) \tag{23}$$

for some positive constants, c_1 and c_2 that only depend on α, $\mathbb{E}_{\mathbb{P}^\star}[\exp(\|\tilde{\boldsymbol{z}}\|^\alpha)]$, n_z being the dimension of $\tilde{\boldsymbol{z}}$, and Δ_1 is the (type-1) Wasserstein metric defined by

$$\Delta_1(\mathbb{P}, \hat{\mathbb{P}}) := \inf_{\mathbb{Q} \in \mathcal{P}_0(\mathcal{Z}^2)} \left\{ \mathbb{E}_{\mathbb{Q}}\left[||\tilde{\boldsymbol{z}} - \tilde{\boldsymbol{\xi}}||\right] \;\middle|\; (\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{\xi}}) \sim \mathbb{Q}, \tilde{\boldsymbol{z}} \sim \mathbb{P}, \tilde{\boldsymbol{\xi}} \sim \hat{\mathbb{P}} \right\}.$$

This result serves as a critical step to build the target attainment confidence guarantee since our result eventually relies on the probability bound of the Wasserstein distance. It remains to derive an upper bound for $\mathbb{P}^\Omega\left[\bar{\Delta}(\mathbb{P}^\star, \hat{\mathbb{P}}) > Tr\right]$. Since $\tilde{\boldsymbol{z}} \in \mathcal{Z}$ almost everywhere, it is easy to see

$$\|\tilde{\boldsymbol{z}}\|_\infty = \max_{\substack{t \in [T], \ell \in [L_t] \\ n \in [N]}} z_{tn}^\ell \leq \max_{t \in [T], \ell \in [L_t]} \bar{z}_{t\ell} = \bar{z} \quad a.e.,$$

then

$$\mathbb{E}_{\mathbb{P}^\star}[\exp(\|\tilde{\boldsymbol{z}}\|_\infty^\alpha)] \leq \exp(\bar{z}^\alpha) < \infty$$

for any $\alpha \geq 1$. We can choose $\alpha = 1$ to meet the assumption in inequality (22). Since $\min_{t \in [T]} \bar{z}_{t\ell} \geq 1$, we have

$$T\|\tilde{\boldsymbol{z}} - \tilde{\boldsymbol{\xi}}\|_\infty \geq \sum_{t \in [T]} \max_{\ell \in [L_t]} \left[\frac{1}{\bar{z}_{t\ell}} ||\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{\xi}}_t^\ell||_\infty\right] \quad a.e.,$$

therefore

$$T\Delta_1(\mathbb{P}^\star, \hat{\mathbb{P}}) \geq \bar{\Delta}(\mathbb{P}^\star, \hat{\mathbb{P}}) \implies \mathbb{P}^\Omega\left[\bar{\Delta}(\mathbb{P}^\star, \hat{\mathbb{P}}) > Tr\right] \leq \mathbb{P}^\Omega\left[\Delta_1(\mathbb{P}^\star, \hat{\mathbb{P}}) > r\right] \quad \forall r \in [0,1].$$

Further notice the dimension of $\tilde{\boldsymbol{z}}$ is $N(\sum_{t \in [T]} L_t) \geq 2$, the following probability bound hold by inequality (23)

$$\mathbb{P}^\Omega\left[\boldsymbol{w}^\top \boldsymbol{x} + \mathbb{E}_{\mathbb{Q}^\star}\left[\sum_{t \in [T]} f_t(\boldsymbol{x}, \tilde{\mathcal{B}}_t)\right] > \tau + Tr\kappa_\tau\right] \leq c_1 \exp(-c_2 \Omega r^{NL}) \quad \forall r \in [0,1],$$

where $L = \sum_{t \in [T]} L_t$, c_1 and c_2 are constants that depend on $\mathbb{E}_{\mathbb{P}^\star}[\exp(||\tilde{\boldsymbol{z}}||_\infty)]$ and NL.

2. We next prove the result for the second probability bound. For ease of expression, we can define the random variables

$$\tilde{\nu}^\omega := \mathbb{E}_{\tilde{\boldsymbol{z}} \sim \mathbb{P}^\star}\left[\sum_{t \in [T]} \max_{\ell \in [L_t]} \left\{\frac{1}{\bar{z}_{t\ell}} \|\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{z}}_t^{\ell\omega}\|_\infty\right\}\right] \quad \forall \omega \in [\Omega],$$

then $\tilde{\nu}^1, \tilde{\nu}^2, \cdots, \tilde{\nu}^\Omega$ are independent random variables and

$$0 \leq \mathbb{E}_{\mathbb{P}^\Omega}\left[\tilde{\nu}^\omega\right] = \mu \leq \operatorname{ess\,sup} \tilde{\nu}^\omega \leq \frac{1}{T}\sum_{t\in[T]} \frac{\bar{z}_{t\ell}}{\bar{z}_{t\ell}} = 1 \quad \forall \omega \in [\Omega].$$

We have the following for any $r \in [\mu, 1]$

$$\begin{aligned}
&\mathbb{P}^\Omega\left[\frac{1}{\Omega}\sum_{\omega\in[\Omega]} \mathbb{E}_{\tilde{\boldsymbol{z}}\sim\mathbb{P}^\star}\left[\sum_{t\in[T]}\max_{\ell\in[L_t]}\left\{\frac{1}{\bar{z}_{t\ell}}\|\tilde{\boldsymbol{z}}_t^\ell - \tilde{\boldsymbol{z}}_t^{\ell\omega}\|_\infty\right\}\right] \geq Tr\right] \\
&= \mathbb{P}^\Omega\left[\frac{1}{\Omega}\sum_{\omega\in[\Omega]}\tilde{\nu}^\omega \geq r\right] \\
&\leq \inf_{\theta\geq 0}\left\{\exp\left(-\theta r\right)\mathbb{E}_{\mathbb{P}^\Omega}\left[\exp\left(\theta\sum_{\omega\in[\Omega]}\tilde{\nu}^\omega/\Omega\right)\right]\right\} \\
&= \inf_{\theta\geq 0}\left\{\exp\left(-\theta r\right)\mathbb{E}_{\mathbb{P}^\Omega}\left[\prod_{\omega\in[\Omega]}\exp\left(\theta\tilde{\nu}^\omega/\Omega\right)\right]\right\} \\
&= \inf_{\theta\geq 0}\left\{\exp\left(-\theta r\right)\left(\prod_{\omega\in[\Omega]}\mathbb{E}_{\mathbb{P}^\Omega}\left[\exp\left(\theta\tilde{\nu}^\omega/\Omega\right)\right]\right)\right\} \\
&\leq \inf_{\theta\geq 0}\left\{\exp\left(-\theta r\right)\left(\prod_{\omega\in[\Omega]}\exp\left(\frac{\theta\mathbb{E}_{\mathbb{P}^\Omega}\left[\tilde{\nu}^\omega\right]}{\Omega} + \frac{\theta^2}{8\Omega^2}\right)\right)\right\} \\
&= \inf_{\theta\geq 0}\left\{\exp\left(\frac{\theta^2}{8\Omega} - \theta r + \frac{\theta\sum_{\omega\in[\Omega]}\mathbb{E}_{\mathbb{P}^\Omega}\left[\tilde{\nu}^\omega\right]}{\Omega}\right)\right\} \\
&\leq \inf_{\theta\geq 0}\left\{\exp\left(\frac{\theta^2}{8\Omega} - \theta(r-\mu)\right)\right\} \\
&= \exp\Big(-2\Omega\left(r-\mu\right)^2\Big),
\end{aligned}$$

where the first equation holds by the definition of $\tilde{\nu}^\omega$, the second inequality holds by Markov inequality, the third equation holds trivially, the fourth equation holds since $\tilde{\nu}^1, \tilde{\nu}^2, \cdots, \tilde{\nu}^\Omega$ are independent, the fifth inequality holds by Hoeffding's lemma, the sixth equation and the seventh inequality hold trivially, and the last equation holds by calculating the minimum of the quadratic function. $\square$

We present two probability bounds in this theorem. For the confidence bound of Fournier and Guillin (2015) to dominate the second bound for $r \in [0, \mu]$, we would require enough samples such that $c_1 \exp(-c_2 \Omega \mu^{NL}) < 1$. However, since data is collected over time in the workforce management problem, we expect the number of data samples available to be far fewer than necessary for the confidence guarantee to be practically useful; if the planning horizon is seven days, the weekly demand data may only provide 52 samples per year. Hence, in practical situations, we expect the deviation between in-sample and out-of-sample performance to be significant.

Nevertheless, for justifying our robust satisficing model, the second probability bound provides the assurance that, for a fixed number of samples Ω, the probability of target shortfalls exceeding $Tr\kappa_\tau$ decreases exponentially in $(r-\mu)^2$ and diminish to 0 when $r>1$. Although this bound is not useful when $r<\mu$, it is indeed a better bound than the first one, especially when μ is a small number. This also means that greater shortfalls may occur but with exponentially decreasing probability. Minimizing the violation probability is consistent with the objective of the robust satisficing model, which aims to achieve the lowest possible value of κ_τ. Notably, we provide a simple expression specific to the robust satisficing model rather than using the results of Fournier and Guillin (2015). More importantly, unlike the confidence guarantee of Fournier and Guillin (2015), the second bound is independent of the number of breakpoints in the generalized reduced information model. Additionally, introducing more breakpoints can result in a lower value of κ_τ and a reduced violation probability bound, further motivating the use of the robust satisficing model constructed from the $\boldsymbol{\gamma}$-weighted time-addictive Wasserstein metric.

While Theorem 6 provides a useful theoretical framework for the robust satisficing model, it does not show how the target should be set. In particular, a practical implementation may require additional techniques, such as cross-validation on the target parameter to optimize out-of-sample performance.

5. Numerical studies

In this section, our focus is on evaluating the performance of the basic and generalized reduced information models across multiple payment values. We also compare the robust satisficing model and the empirical model using simulated data.

Evaluation of the reduced information models using Solomon's data sets

In Section 3, we demonstrate through a pathological example that the basic reduced information model could have an arbitrarily large relative performance gap with respect to the true model when multiple payment values are allowed. However, in practical delivery problems, the performance gap is usually acceptable. To further investigate the impact of multiple payment values, we conduct experiments using Solomon's data sets (Solomon 1987) for the VRP with time windows, where hierarchical objectives are considered, with the primary aim being to minimize the number of vehicles and then the travel distance. The solutions are available at `https://sun.aei.polsl.pl/ /~zjc/best-solutions-solomon.html`. We only consider the C1, R1, and RC1 data sets, as these contain instances where more than ten vehicles are used. We do not consider the R2, C2, and RC2 data sets because they only involve a small number of vehicles (2, 3, or 4). Thus, there are 29 instances in total.

For each instance, we consider a list of $N = 5$ possible payments, with $p_n = \lfloor 0.8d_{min} + (n-1)\frac{(1.2d_{max}-0.8d_{min})}{(N-1)} \rfloor$, where d_{min} and d_{max} represent the minimal and maximal route lengths, respectively. Note that the route length does not include the travel distance from the last customer to the depot. We set $K_t = 2J_t$ and consider four cases for bidding jobs. In each case, every ad-hoc courier could provide $N' \in \{2, \ldots, 5\}$ payment values. For each courier $j \in [J_t]$, we randomly generate their coordinates within the minimal and maximal coordinates of all nodes, then calculate the travel distances required to complete each job, which is the sum of the distance between a courier's current location and the depot, and the route length associated with a job. Subsequently, we sort the travel distances of each courier in non-decreasing order, denoting the resulting list as D_j. For the first $\lfloor \frac{|J_t|}{N'} \rfloor$ jobs in D_j, we calculate the average travel distance $\bar{d}$ and set the courier's payment for each of these jobs as the one that has the minimal absolute deviation from $\bar{d}$. We then follow a similar procedure to set the payment for the second $\lfloor \frac{|J_t|}{N'} \rfloor$ jobs until N' payments are specified. We acknowledge that this bidding scheme is simplistic and intended for illustrative purposes only. More complex bidding schemes can be developed by incorporating other influential factors, such as the courier's destination after finishing a job and the properties of the jobs (*e.g.*, the number of customers included in a job and package weights).

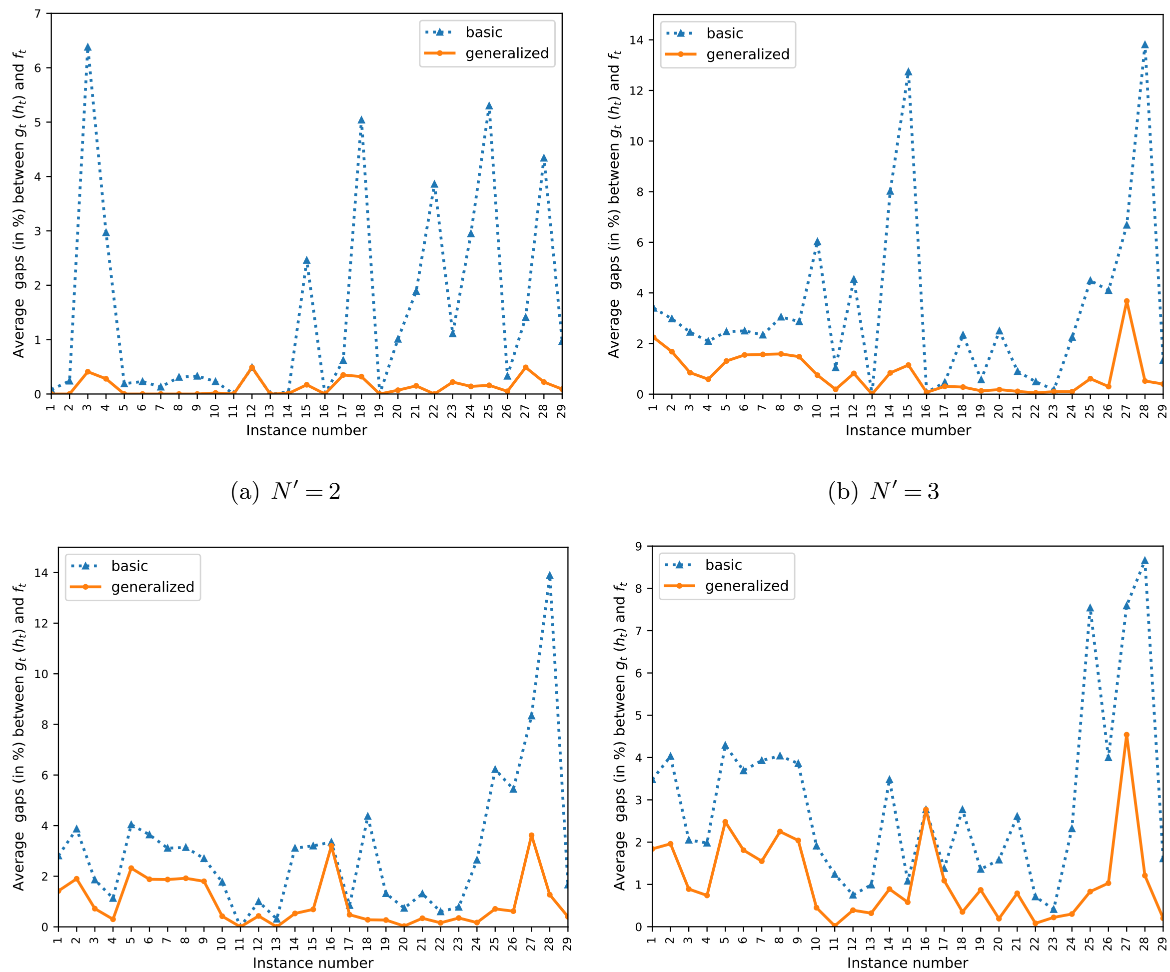
Subsequently, we solve Problems (1) and (7) for $\boldsymbol{a}_t^\top \boldsymbol{x} \in \{0, \ldots, J_t - 1\}$. For a combination of $(N', \boldsymbol{a}_t^\top \boldsymbol{x})$, we randomly generate 20 instances, where each instance differentiates from the other in the payment values provided by a courier and the sets of bidding jobs under each payment. The average percentage gaps are reported in Table 1, where gaps are calculated by $\frac{g_t(\boldsymbol{x},\boldsymbol{z}_t)-f_t(\boldsymbol{x},\mathcal{B}_t)}{f_t(\boldsymbol{x},\mathcal{B}_t)}$. Moreover, we also present the results under $N' = 1$ to give some intuition of Theorem 2. To provide insights into the differences among payment values, we give the price ratio between the maximal value p_5 and the minimal p_1 in the table.

Table 1 Average gaps (in %) between $g_t(\boldsymbol{x}, \boldsymbol{z}_t)$ and $f_t(\boldsymbol{x}, \mathcal{B}_t)$ under multiple payment values

Number	Instance	p_5/p_1	$N'=1$	$N'=2$	$N'=3$	$N'=4$	$N'=5$	Number	Instance	p_5/p_1	$N'=1$	$N'=2$	$N'=3$	$N'=4$	$N'=5$
1	C101	2.81	0.00	0.08	3.40	2.81	3.48	16	R107	3.00	0.00	0.00	0.07	3.35	2.77
2	C102	2.81	0.00	0.24	2.98	3.87	4.03	17	R108	2.02	0.00	0.62	0.47	0.85	1.38
3	C103	4.25	0.00	6.38	2.45	1.86	2.05	18	R109	2.42	0.00	5.04	2.34	4.37	2.77
4	C104	4.25	0.00	2.97	2.09	1.13	1.98	19	R110	2.05	0.00	0.00	0.57	1.32	1.36
5	C105	2.81	0.00	0.19	2.47	4.04	4.29	20	R111	2.67	0.00	1.01	2.50	0.74	1.58
6	C106	2.81	0.00	0.23	2.50	3.64	3.69	21	R112	2.02	0.00	1.89	0.90	1.31	2.61
7	C107	2.81	0.00	0.13	2.34	3.11	3.93	22	RC101	4.07	0.00	3.86	0.49	0.60	0.71
8	C108	2.81	0.00	0.31	3.05	3.14	4.04	23	RC102	2.62	0.00	1.11	0.17	0.78	0.41
9	C109	2.81	0.00	0.33	2.87	2.70	3.86	24	RC103	3.08	0.00	2.95	2.25	2.64	2.32
10	R101	6.42	0.00	0.23	6.03	1.77	1.91	25	RC104	2.53	0.00	5.30	4.49	6.22	7.54
11	R102	9.71	0.00	0.00	1.05	0.00	1.24	26	RC105	3.67	0.00	0.33	4.10	5.44	4.00
12	R103	5.70	0.00	0.49	4.54	1.00	0.75	27	RC106	2.87	0.00	1.41	6.68	8.34	7.59
13	R104	2.09	0.00	0.00	0.00	0.31	0.99	28	RC107	3.46	0.00	4.34	13.81	13.89	8.66
14	R105	4.23	0.00	0.04	8.02	3.11	3.48	29	RC108	2.69	0.00	0.97	1.34	1.66	1.61
15	R106	3.48	0.00	2.46	12.74	3.20	1.08								

Table 1 indicates that the price ratio p_5/p_1 ranges from 2.02 to 9.71, relatively high values compared to the average earnings of Amazon Flex drivers, which is stated to be between 18 to

25 dollars per hour, resulting in a ratio of around 1.39. Despite this, our approximation model performs well, with average performance gaps of less than 5% observed in 101 out of 116 cases where multiple payment values are allowed. In 35 cases, the average gaps are less than 1%. When the single payment value bidding scheme is employed, the reduced information model generates exact solutions, as illustrated in columns $N' = 1$. Based on these results, we conclude that our basic reduced information model can provide good approximations for Problem (1) when applied to practical delivery problems.



(a) $N' = 2$ (b) $N' = 3$

(c) $N' = 4$ (d) $N' = 5$

Figure 1 Average gaps (in %) between the reduced information models and the crowdsourcing cost function

To provide a comparative analysis of the performance of the generalized information model, we conduct further experiments using the same instances using $L = 2$ breakpoints, where $u_0 = 0$ and $u_2 = \lfloor J_t/2 \rfloor$. Figure 1 plots the average percentage gaps between the basic/generalized reduced information model and the crowdsourcing cost function. The results reveal that the generalized model provides tighter bounds than the basic model and sometimes significantly improves the bounds. For example, when $N' = 3$, the average gap between the basic model and the crowdsourcing cost function is approximately 14% in instance RC107 (*i.e.*, instance number 28); however, the generalized model reduces the gap to 0.5%. Moreover, we observe that the generalized model produces solutions with performance gaps of less than 5% for all the instances. When $N' = 2$, the average gap of the generalized model is less than 0.5% for all the instances. Therefore, we can conclude that the generalized reduced information model can provide better approximations for Problem (1) and be useful in tackling practical delivery problems.

Mitigating ambiguity via robust satisficing

In this numerical study, we simulate a delivery platform that makes workforce management decisions for a planning horizon of 8 periods, representing a typical workday of 8 hours. There are a total of 26 work shifts, each lasting between 1 to 4 periods. The compensation for each contracted courier in the ith work shift is given by $w_i = 20\sigma_i \times 0.9^{\sigma_i - 1}$, where the basic compensation for each period is 20 and σ_i is the number of periods that shift i covers. A discount factor of 0.9 is used to distinguish shifts; otherwise, hiring a person for a shift of a certain number of periods is equivalent to hiring that person for multiple continuous shifts, each lasting for one period. It is also reasonable in real-world applications that shorter shift has higher compensation per hour since it is more flexible.

For each period $t \in [T]$, where $T = 8$, the number of jobs J_t has lower and upper bounds $J_{min} = 10$ and $J_{max} = 50$. We randomly generate J_t from a normal distribution with mean $\frac{J_{min}+J_{max}}{2} = 30$ and standard deviation being the same as the mean, and truncated within the interval $[J_{min}, J_{max}]$ to ensure that J_t is an integer in $\{J_{min}, J_{min}+1, \ldots, J_{max}\}$. Each courier $k \in [K_t]$, where $K_t = \lfloor 1.2 J_t \rfloor$, has a travel cost denoted by $\bar{c}_{kt}$ from their current location to the depot, which is a random number between $\bar{c}_l = 0$ and $\bar{c}_u = 20$. The payment c_{jt} for each job $j \in [J_t]$ in period $t \in [T]$ is randomly generated from the interval $[c_l, c_u] = [20, 50]$. The delivery platform provides ad-hoc couriers with $N = 5$ possible bidding prices. Let $p_{min} = c_l + 0.5(\bar{c}_l + \bar{c}_u)$, $p_{max} = c_u + 0.5(\bar{c}_l + \bar{c}_u)$, and $\Delta p = (p_{max} - p_{min})/(N-1)$. Then the bidding price is set to $p_n = p_{min} + (n-1)\Delta p$ for each $n \in [N]$. The bidding price of each courier $k \in [K_t]$ for job $j \in [J_t]$ is set as the smallest value in $\{p_1, \ldots, p_N\}$ that is higher than or equal to the courier's expected payment for taking that job, which is calculated as $p_{kjt} = \bar{c}_{kt} + c_{jt}$. To reflect the fluctuation of price in different time intervals, we let the courier's expected payment p_{kjt} increase and decrease by 20% in periods $t \in \{1, 8\}$ and $t \in \{4, 5\}$ to simulate the peak and off-peak periods, respectively. If all the provided prices are smaller than p_{kjt}, the bidding price is set to p_N. To evaluate the performance of the proposed model, we generate 7 training samples and 200 testing samples for each instance.

To compare the performance of the robust satisficing (RS) model and the empirical model (EMP) associated with the generalized reduced information model, we begin by solving the assignment problem to obtain the generalized reduced information vectors $\boldsymbol{z}_t^{\ell\omega}$, for all $t \in [T]$, $\ell \in [L_t]$, and $\omega \in [\Omega]$. We then solve the empirical model (14) and obtain its objective value Z_0. Next, we consider a target ratio r and set $\tau = rZ_0$. Using a binary search algorithm, we solve problem (19) to produce the solution of the robust satisficing model (18). When solving the generalized reduced information model, we consider three cases for setting the breakpoints in each period:

1. $L_t = 1$ with $u_{t1} = 0$. In this case, the generalized reduced information model coincides with the basic reduced information model (7).
2. $L_t = 2$ with $u_{t1} = 0$ and $u_{t2} = \lfloor 0.5 \sum_{\omega \in [\Omega]} J_t^\omega / \Omega \rfloor$, where J_t^ω is the number of jobs in period $t \in [T]$ under training sample $\omega \in [\Omega]$.
3. $L_t = 2$ with $u_{t1} = 0$ and $u_{t2} = \boldsymbol{a}_t^\top \hat{\boldsymbol{x}}$, where $\hat{\boldsymbol{x}}$ is the solution to the empirical optimization problem (3) with reformulation in (4). Note that if we set the breakpoint to be $\hat{\boldsymbol{x}}$, the generalized reduced information model (14) solves exactly the empirical optimization problem (3).

For ease of notation, we use symbols B, G1, and G2 to denote the above three settings of breakpoints and append them to EMP and RS to represent the empirical and robust satisficing models associated with the breakpoints under these settings. Finally, we evaluate the performance of both RS and EMP models in testing samples by solving problem (1) and obtain the second-stage crowdsourcing costs. To accelerate the resolution of models, we relax the feasible set $\mathcal{X}$ of the decision variable $\boldsymbol{x}$ from $\mathbb{Z}_+^I$ to $\mathbb{R}_+^I$. Our preliminary results show that the differences in the out-of-sample performance of integer solutions (round real-number solutions to the nearest integers) and real-number solutions are negligible. Thus, we conduct the following experiments by setting $\boldsymbol{x} \in \mathbb{R}_+^I$.

Figure 2 demonstrates the out-of-sample performance comparison between the RS and empirical models. To mitigate the impact of randomness, we generate 30 instances and report the average performance. The vertical axis of the figure represents the value (in %) of the difference in test objective between MODEL and EMP_G2, divided by the test objective of EMP_G2, denoted by $\frac{\text{MODEL}-\text{EMP\_G2}}{\text{EMP\_G2}}$ for short, where MODEL $\in$ {RS_B, RS_G1, RS_G2, EMP_B, EMP_G1, EMP_G2}. We observe that the performance of EMP_B, EMP_G1, and EMP_G2 are comparable in the out-of-sample test, and RS models produce better out-of-sample performance than the corresponding EMP models when the target is set slightly higher than the objective of EMP models. Moreover, RS_G2 exhibits superior performance to RS_B and RS_G1 even though EMP_G2 is marginally worse than EMP_B and EMP_G1.

To further evaluate the performance of the RS model, we introduce more uncertainty in the number of jobs in the testing samples. Specifically, we set the low and upper bounds to $J_{min} =$

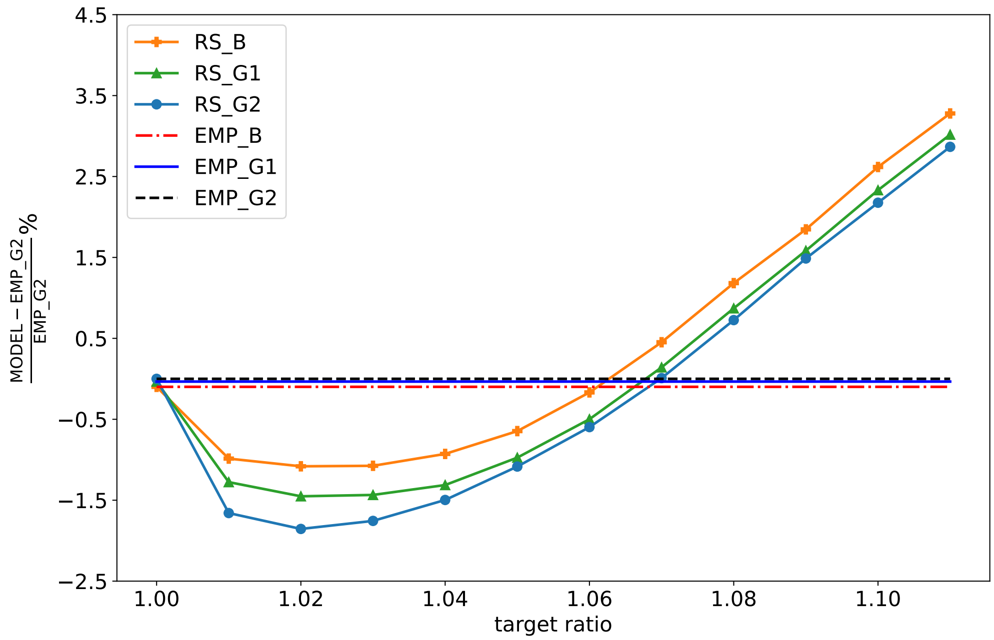


Figure 2 **Comparison of out-of-sample performance between the robust satisficing and the empirical models.**

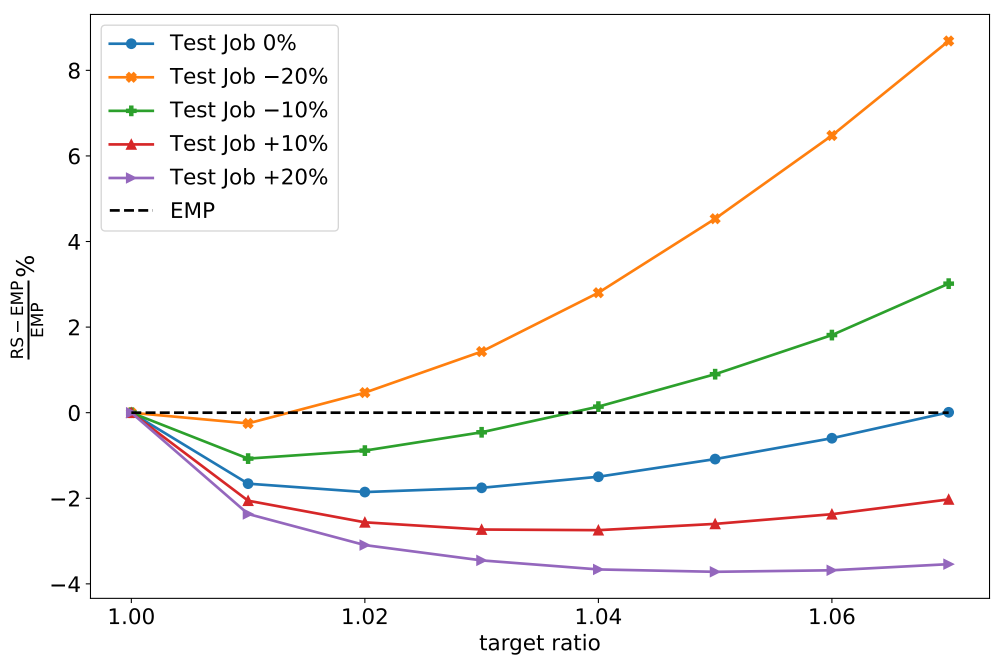


Figure 3 **Comparison of out-of-sample performance between the robust satisficing and the empirical models with different deviations of job numbers in testing samples.**

$\lfloor 10(1-\zeta) \rfloor$ and $J_{max} = \lfloor 50(1+\zeta) \rfloor$, where $\zeta \in \{-20\%, -10\%, 0\%, +10\%, +20\%\}$ represents the percentage variations of job numbers, and generate J_t following the previous approach. We evaluate the out-of-sample performance of the RS_G2 and the EMP_G2 models and report the average comparison over 30 random instances in Figure 3. When the testing samples have a more significant variation in job numbers, the benefit of using the RS_G2 model becomes more apparent. When the variation is minor, the RS_G2 model with a small target performs better than the EMP_G2 model. Thus, we advise that decision-makers adopt a modestly conservative target in the RS model, say, 1.02, which would be advantageous across almost all the scenarios, including low and high levels of uncertainty and risk. Moreover, if a greater variation in the job numbers is expected in the future, then a conservative target in the robust satisficing model could lead to a more significant improvement in the actual performance.

In sum, our experiments using Solomon's data sets suggest that the performance gap between the reduced information models and the true model is usually acceptable for practical delivery problems. More importantly, through simulated data, we show that the robust satisficing model can produce better out-of-sample performance than the empirical model when the target is set slightly higher than the objective of the empirical model.

6. Conclusion

To offer affordable and reliable delivery services, e-commerce platforms and local businesses are increasingly turning to crowdsourced delivery resources. However, the uncertainties surrounding ad-hoc couriers' availability and job bidding behavior have presented a challenge to the management of the workforce. In this study, we propose a robust satisficing framework that accounts for the uncertainty of ad-hoc couriers and the cost associated with hiring them when necessary. Our framework provides a practical tool for decision-making that enables platforms to effectively manage their workforce resources and balance their cost objectives with service quality requirements. Future work in this area could explore the integration of predictive analytics into the robust satisficing framework to provide better insights into future demand and couriers' availability, which could further enhance the platform's decision-making capability.

References

Ahuja RK, Magnanti TL, Orlin JB (1988) Network flows .

Alnaggar A, Gzara F, Bookbinder JH (2021) Crowdsourced delivery: A review of platforms and academic literature. *Omega* 98:102139.

Archetti C, Savelsbergh M, Speranza MG (2016) The vehicle routing problem with occasional drivers. *European Journal of Operational Research* 254(2):472–480.

Behrendt A, Savelsbergh M, Wang H (2022a) Crowdsourced same-day delivery: Joint planning and coordination for centralized and decentralized couriers. Available at Optimization-Online: `https://optimization-online.org/?p=20103`.

Behrendt A, Savelsbergh M, Wang H (2022b) A prescriptive machine learning method for courier scheduling on crowdsourced delivery platforms. *Transportation Science (forthcoming)*.

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* 99(2):351–376.

Bertsimas D, Gupta V, Kallus N (2018) Data-driven robust optimization. *Mathematical Programming* 167:235–292.

Birge JR, Louveaux F (2011) *Introduction to Stochastic Programming* (Springer Science & Business Media).

Chen L, Sim M, Zhang X, Zhou M (2022) Robust explainable prescriptive analytics. Available at SSRN: `https://ssrn.com/abstract=4106222`.

Chen LG, Tang Q (2022) Supply chain performance with target-oriented firms. *Manufacturing & Service Operations Management* 24(3):1714–1732.

Chen X, Li M (2021) Discrete convex analysis and its applications in operations: A survey. *Production and Operations Management* 30(6):1904–1926.

Chen Z, Sim M, Xiong P (2020) Robust stochastic optimization made easy with rsome. *Management Science* 66(8):3329–3339.

Dayarian I, Savelsbergh M (2020) Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders. *Production and Operations Management* 29(9):2153–2174.

Fournier N, Guillin A (2015) On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields* 162(3):707–738.

Gao R, Kleywegt A (2022) Distributionally robust stochastic optimization with wasserstein distance. *Mathematics of Operations Research (forthcoming)*.

Goh J, Hall NG (2013) Total cost control in project management via satisficing. *Management Science* 59(6):1354–1372.

Goyal A, Zhang Y, Benjaafar S (2023) Crowdsourcing last-mile delivery with hybrid fleets under uncertainties of demand and driver supply: Optimizing profitability and service level. Available at SSRN: `https://ssrn.com/abstract=4322670`.

Liu F, Chen Z, Wang S (2022) Globalized distributionally robust counterpart. Available at SSRN: `https://ssrn.com/abstract=4221952`.

Liu S, He L, Max Shen ZJ (2021) On-time last-mile delivery: Order assignment with travel-time predictors. *Management Science* 67(7):4095–4119.

Long Z, Sim M, Zhou M (2022) Robust satisficing. *Operations Research* 71(1):61–82.

Mao J (1970) Survey of capital budgeting: Theory and practice. *The Journal of Finance* 25(2):349–360.

Mohajerin Esfahani P, Kuhn D (2018) Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming* 171(1):115–166.

Murota K (1998) Discrete convex analysis. *Mathematical Programming* 83(1):313–371.

Savelsbergh MW, Ulmer MW (2022) Challenges and opportunities in crowdsourced delivery planning and operations. *4OR* 20(1):1–21.

Sim M, Tang Q, Zhou M, Zhu T (2021) The analytics of robust satisficing. Available at SSRN: `https://ssrn.com/abstract=3829562`.

Simon HA (1955) A behavioral model of rational choice. *The Quarterly Journal of Economics* 99–118.

Smith JE, Winkler RL (2006) The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science* 52(3):311–322.

Solomon MM (1987) Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research* 35(2):254–265.

Torres F, Gendreau M, Rei W (2022) Vehicle routing with stochastic supply of crowd vehicles and time windows. *Transportation Science* 56(3):631–653.

Ulmer MW, Savelsbergh M (2020) Workforce scheduling in the era of crowdsourced delivery. *Transportation Science* 54(4):1113–1133.

Yildiz B, Savelsbergh M (2019) Service and capacity planning in crowd-sourced delivery. *Transportation Research Part C: Emerging Technologies* 100:177–199.

Zhang Y, Zhang Z, Lim A, Sim M (2021) Robust data-driven vehicle routing with time windows. *Operations Research* 69(2):469–485.

Zhou M, Sim M, Lam SW (2022) Advance admission scheduling via resource satisficing. *Production and Operations Management* 31(11):4002–4020.