Reliable Facility Location Under Event-Correlated Uncertainty

Chun Cheng
Institute of Supply Chain Analytics, Dongbei University of Finance and Economics, Dalian 116025, China
chun.cheng@polymtl.ca

Qinxiao Yu
Economics and Management College, Civil Aviation University of China, Tianjin 300300, China
Corresponding Author. yuqinxiao@tju.edu.cn

Yossiri Adulyasak
HEC Montréal and GERAD, Montréal H3T 2A7, Canada
yossiri.adulyasak@hec.ca

Louis-Martin Rousseau
Polytechnique Montréal and CIRRELT, Montréal H3C 3A7, Canada
louis-martin.rousseau@cirrelt.net

Supply chains are often prone to disruptions caused by rare but high-impact events such as the recent global pandemic. To enhance supply chain robustness and reliability, we study a capacitated facility location problem where facility capacity and customer demand are subject to uncertainties simultaneously. In this problem, the decision maker seeks to select a subset of facilities from a given set of candidate sites to open at the system design phase, to serve customer demand during the operational phase. This problem can be formulated as a two-stage stochastic programming model when the joint distribution of the two types of uncertain parameters is perfectly known. Yet, in practice, the exact probability distribution is often unknown and can only be estimated from historical observations, resulting in severe inaccuracy. To address this issue, we assume that the probability distribution lies in a scenario-wise ambiguity set, where only partial distributional information is known. The utilized ambiguity set can capture uncertainties caused by different random events or different magnitudes of the same event type, and it can also be used to explicitly represent the correlation between the provider-side and the receiver-side uncertainties. We apply an adaptation policy to the resulting distributionally robust optimization (DRO) model and reformulate it to a mixed-integer linear programming model. We further extend the framework to the location and inventory pre-positioning problem with uncertainties in disaster operations management. Extensive simulation tests and numerical experiments based on a case study are conducted to validate the scenario-wise DRO framework. Results show that it can provide a better trade-off between cost and service level than other modeling schemes, demonstrating that the proposed DRO model offers a practical decision-making tool for enhancing supply chain robustness via facility location.

Key words: facility location, capacity failure, demand uncertainty, scenario-wise ambiguity, distributionally robust optimization

1. Introduction

Disruptions caused by unexpected events or random factors, such as natural disasters or industrial accidents can simultaneously affect facility operations and customer demand at a large scale. Specifically, facilities' capacity can be partially or completely diminished by random events. Meanwhile, customer demand patterns may also deviate from those in the nominal disruption-free scenario, e.g., after a disaster, the demand for daily necessities and medical supplies may increase and the demand for luxuries may decrease or even vanish (Ergun et al. 2011, An et al. 2014). A recent example is the supply chain disruptions caused by the coronavirus pandemic, where supplies of personal protective equipment and medical devices have been restricted due to shutdowns of manufacturing plants, whereas demand has increased (Besson 2020). To enhance supply chain robustness, decision makers should take these uncertainties into account at the system design phase, to avoid costly recourse actions during the operational phase.

Moreover, during the lifetime of a supply chain system, its operations can be affected by multiple types of random events, e.g., natural disasters and man-made interference, and each having a different impact and leading to different levels of uncertainty, thus calling for different recourse actions. Even under the same type of disruption events, the impact can vary significantly. Thus, we consider it important for decision makers to incorporate *event-correlated* uncertainties when robustifying the supply chain system, to hedge against different levels of damage resulting from different types of random events as well as from different magnitudes of the same event type.

Facility location, one of the most significant strategic decisions in a supply chain system, determines the locations of new facilities from a set of candidate sites at the system design phase, in order to meet customer demand during the operational phase. However, the aforementioned random events may affect facilities' operations such that they do not have enough capacity to satisfy all the demand, leading to poor service level. Meanwhile, customer demand patterns may also be influenced by random events, which further aggravates the impact of random events on facilities, especially when demand surges after a disruption event. Such an unexpected interruption and disruption can potentially put the supply chain of a firm at risk, and an inability to incorporate disruption risk in the decision-making process can conceivably result in far more expensive long-run operating costs (Chopra and Sodhi 2014). Therefore, to improve supply chain robustness and reliability, it is crucial to consider uncertainties in facility location problems (FLPs) so that the firm can reliably provide satisfactory customer service (i.e., meet demand as much as possible) at a reasonable cost in both the nominal and disruption situations.

To tackle problems under uncertainties, there are three frameworks we can use: stochastic programming (SP), robust optimization (RO), and distributionally robust optimization (DRO) (Chen et al. 2020). In stochastic models, the uncertain parameters are denoted as random variables with

a *known* probability distribution. The objective function is to minimize/maximize the expected cost/profit over this distribution. Due to limited samples, it is often difficult to estimate the occurrence probability of some random events, especially natural disasters. Thus, solutions produced by stochastic models may lead to disappointing results in out-of-sample tests (Smith and Winkler 2006). In contrast, the RO approach does not rely on the probability distribution. Instead, it assumes that the uncertain parameters belong to an uncertainty set. The RO models then optimize the performance of the *worst-case scenario* within the uncertainty set, which often leads to over-conservative solutions. As an alternative paradigm, DRO unifies SP and RO to overcome their deficiencies. The DRO approach assumes that the probability distribution of uncertain parameters belongs to a family of distributions (i.e., an ambiguity set) that shares common distributional information, which is often available and reliable (Popescu 2007). Optimization is subsequently performed to hedge against the *worst-case distribution* within the ambiguity set. As the DRO method does not require perfect distribution information for the uncertain parameters, it is more compatible with the uncertainties considered in this study than the SP method. Moreover, compared to the RO method, the DRO approach can take full advantage of available data to produce less-conservative solutions, which result in significantly lower operating costs in the long run.

1.1. Our Contributions

In this paper, we complement the existing literature on reliable facility location by proposing a two-stage distributionally robust model for the capacitated fixed-charge location problem (CFLP). The model captures both provider-side and receiver-side uncertainties, which are associated with the capacity at facilities and the demand at customers. Instead of assuming that the joint probability distribution of uncertain parameters is perfectly known, we assume that it lies in a scenario-wise ambiguity set, where some of the statistical information in each scenario can be estimated from historical observations which can potentially be used in conjunction with human inputs when historical data is insufficient. Our study makes the following contributions to the literature:

- We use a DRO modeling paradigm to tackle the FLPs with simultaneous provider-side and receiver-side uncertainties. Unlike stochastic models, it does not require perfect distribution information of uncertain parameters. Different from RO models, it can make use of historical data to mitigate the conservatism of robust solutions.
- We construct a scenario-wise ambiguity set to characterize the randomness of parameters, which can be used to capture the impacts of different types of disruption events or different magnitudes of the same event type. More importantly, the applied ambiguity set can be used to represent the correlation between the two types of uncertainty considered (via the event causing these simultaneous uncertainties) instead of assuming they are independent as is often

seen in the literature. We also shed light on how to construct the scenario-wise ambiguity set from historical observations for practitioners.

- We apply a scenario-wise adaptation policy for the second-stage recourse problem of the DRO model, which can further alleviate the conservatism of robust solutions. We reformulate the resulting adaptive DRO model together with the scenario-wise ambiguity set to a mixed-integer linear programming (MILP) model, which is solvable by off-the-shelf solvers.
- We extend the proposed modeling framework to a location and inventory pre-positioning problem with uncertainty in disaster operations management, for which numerical results are also provided using instances based on a case study.
- We conduct both simulation tests and a case study to validate the scenario-wise DRO framework. Results show that the DRO model takes much less computing time in comparison with the sample average approximation (SAA) model; meanwhile, it achieves a better trade-off between cost and service level in out-of-sample tests than other modeling schemes, demonstrating that the scenario-wise DRO model provides a practical decision-making tool for enhancing supply chain robustness.

The rest of this paper is organized as follows. Section 2 reviews related literature. Section 3 describes the problem and presents different formulations to model the problem. Section 4 develops the solution method for the scenario-wise DRO model. Numerical tests and analyses are provided in Section 5, which is followed by conclusions in Section 6.

2. Literature Review

This section reviews related research on FLPs under uncertainty and the DRO approach.

2.1. Facility Location Problems Under Uncertainty

In FLPs, there are generally three types of uncertainty: provider-side uncertainty, in-between uncertainty, and receiver-side uncertainty (Shen et al. 2011). Provider-side uncertainty involves uncertain supply capacity and uncertain lead time. In-between uncertainty refers to uncertain travel time or cost and uncertain transportation capacity on arcs (due to failures of the transportation network). Receiver-side uncertainty is the random demand at customers. These three types of uncertainty have been widely considered in FLPs. For example, FLP under facility disruptions (Snyder and Daskin 2005, An et al. 2014, Lu et al. 2015, Azad and Hassini 2019, Du et al. 2020), FLP under edge failures (Xie and Ouyang 2019, Matthews et al. 2019), FLP with uncertain travel time/cost (Nikoofal and Sadjadi 2010, Gao and Qin 2016, Mišković et al. 2017), and FLP with uncertain demand (Atamtürk and Zhang 2007, Shehadeh and Sanci 2021, Basciftci et al. 2021, Saif and Delage 2021, Zhang et al. 2021b).

Although there are many works on FLPs under uncertainty, most of them consider one type of uncertainty at a time and only a limited number of prior works study simultaneous uncertainties. However, the presence of multiple simultaneous uncertainties is common in realistic applications of facility location. For example, as occurred during the coronavirus pandemic, the demand for some products such as domestic toilet rolls surges, while productivity decreases due to employee absences and the shutdown of factories (Montgomery 2020). Thus, it is important to consider multiple uncertainties at the system design phase to enhance supply chain resilience. Noyan et al. (2016) explore a network design problem that determines the locations and capacities of relief centers. They consider customers' uncertain demand for relief items and the uncertain capacity of transportation links (e.g., roads and bridges). A two-stage SP model is constructed for the problem, which is solved by a Benders decomposition-based branch-and-cut algorithm. Elçi and Noyan (2018) explore a stochastic pre-disaster relief network design problem, which decides the locations and capacities of the response facilities and their inventory levels. Multiple types of parameters are subject to uncertainty—demand, travel time, supply capacity, unit transportation cost, and shortage penalty cost—which are represented by scenarios. The authors use a chance-constrained two-stage mean-risk SP model for the problem. Zetina et al. (2017) study the uncapacitated hub location problem with uncertain demand and transportation costs using the static RO method, where budgeted uncertainty sets are utilized to capture randomness. Wang et al. (2020) apply the adaptive DRO method to both the uncapacitated and capacitated hub location problems under demand and cost uncertainty. Taherkhani et al. (2021) investigate a profit-maximizing capacitated hub location problem, where demands and revenues are subject to uncertainty. The authors use both stochastic and robust approaches to model the problem. Mazahir and Ardestani-Jaafari (2020) study a global sourcing problem under compliance legislation, where a supplier's compliance capability to a market is uncertain (resulting in arc disruptions) and customer demand is also uncertain. They use a two-stage RO method to formulate the problem. Cheng et al. (2021) study a CFLP with facility disruptions and uncertain demand using a two-stage RO method. Specifically, the authors employ a budgeted uncertainty set to characterize uncertainties, where a binary random variable is used to denote whether a facility is (completely) disrupted or not.

Based on the reviewed papers, we realize that the dependency between facility disruption and uncertain customer demand is often neglected in the literature on facility location under uncertainty, whereas this relationship often holds in practical applications. Moreover, the DRO modeling framework and its corresponding solution method for FLPs with multiple types of uncertainty have not been explored. We further emphasize that our work differs from that of Cheng et al. (2021) in three aspects: First, our ambiguity set can capture the dependency between facility disruption and uncertain demand, whereas these two types of uncertainties are assumed to be independent

in their uncertainty set. Second, facilities are assumed to lose their capacities completely in Cheng et al. (2021) once disruptions happen, whereas our work assumes that facilities can be completely or partially disrupted. Our setting provides more flexibility because facilities may still be able to meet part of the demand in some disruption scenarios. Third, different modeling paradigms are used to solve the CFLP under uncertainties; thus, different solution methods are required. In addition, although the DRO modeling paradigm has been applied by Saif and Delage (2021) to solve the FLP, the authors only consider uncertain demand. In particular, they construct a Wasserstein ambiguity set, which can be considered equivalent to the scenario-wise ambiguity set with one scenario. Thus, our model generalizes the approach considered in Saif and Delage (2021) by considering simultaneous supply and demand uncertainties through the construction of a scenario-wise ambiguity set. The quality of the robust solutions obtained by our proposed model with the scenario-wise ambiguity set is presented in Section 5.

2.2. Distributionally Robust Optimization

The concept of DRO can be dated back to the work of Scarf (1957), which addresses an ambiguity-averse newsvendor problem. Due to the development of RO and statistics, tractable reformulations for important classes of DRO models have been developed only recently (Saif and Delage 2021). A key step in employing the DRO method is the construction of the ambiguity set. In general, there are two types of uncertainty sets: *moment-based* and *statistical distance-based*. The moment-based ambiguity set includes all the probability distributions that meet specified moment constraints, e.g., the first and the second moments. This type of ambiguity set has been widely used in solving various application problems, for example portfolio optimization (Delage and Ye 2010), single machine scheduling (Chang et al. 2017), and facility location (Shehadeh and Sanci 2021), among others. In recent years, the statistical distance-based ambiguity set has gained in popularity, which includes all the probability distributions that are within a certain *distance* from the given distribution. The proposed distance metrics encompass the Wasserstein metric (Mohajerin Esfahani and Kuhn 2018) and the $\phi-$divergence (Ben-Tal et al. 2013). This class of ambiguity set is also used in different applications, e.g., facility location (Saif and Delage 2021), the unit commitment problem (Hou et al. 2018), and the vehicle routing problem (Zhang et al. 2021c), to name a few.

Recently, Chen et al. (2020) introduce a new DRO framework, *robust stochastic optimization (RSO)*, which unifies scenario tree-based stochastic optimization and DRO in a single framework. The equipped event-wise ambiguity set is rich enough to cover several types of ambiguity sets, including those generated by statistical-based or machine learning-based methods. The works of Hao et al. (2020), Cheng et al. (2020), Shehadeh and Sanci (2021), Perakis et al. (2022), and Li et al. (2022) are novel applications of the RSO framework. In particular, Hao et al. (2020) study

a vehicle allocation problem with uncertain demand, which is affected by weather conditions, i.e., the demand presents different patterns with respect to the weather (sunny or rainy). Shehadeh and Sanci (2021) consider a distributionally robust CFLP with bimodal random demand. Specifically, the authors assume that customer demand belongs to exactly two spatially distinct distributions—one before the occurrence of an event and one after it happens. In contrast to the work of Shehadeh and Sanci (2021), our study considers both provider-side and receiver-side uncertainty, and more importantly, our ambiguity set can encompass more than two scenarios to differentiate the impacts of different events, or different magnitudes of the same event type. In addition, we utilize a scenario-wise adaptation policy for the second-stage recourse problem, while Shehadeh and Sanci (2021) adopt a static policy—the recourse decisions are the same for all realizations of the uncertain scenario. Perakis et al. (2022) explore a multi-item joint pricing and production problem, where a cluster-wise ambiguity set is constructed by applying the K-means clustering algorithm to demand residuals. Li et al. (2022) study a multi-period inventory routing problem with uncertain demand, where the first-moment information of demand under each scenario is included in the ambiguity set.

Notation. We denote by $[S] \triangleq \{1, \ldots, S\}$ the set of positive running indices up to S. Boldface lowercase and uppercase characters represent vectors and matrices with appropriate dimensions, respectively. $\boldsymbol{a}^T$ is the transpose of $\boldsymbol{a}$. $\boldsymbol{a} \cdot \boldsymbol{b}$ denotes the dot product of two vectors with the same dimension. We use $\mathcal{P}(\mathbb{R}^I)$ to denote the set of all distributions supported on $\mathbb{R}^I$. A random variable, $\tilde{d}_i$ is denoted with a tilde sign and we use $\tilde{\boldsymbol{d}} \in \mathbb{P}, \mathbb{P} \in \mathcal{P}(\Omega), \Omega \subseteq \mathbb{R}^I$ to define $\tilde{\boldsymbol{d}}$ as an I-dimensional random vector with support Ω and distribution $\mathbb{P}$.

3. Problem Definition and Formulation

In this section, we describe the problem and formulate it as a two-stage SP model and a two-stage DRO model. We also provide an example to show the value of the scenario-wise ambiguity set.

3.1. Problem Definition

We consider a CFLP, where $[J]$ is the set of candidate facilities and $[I]$ is the set of customers. Parameter f_j is the fixed cost of locating a facility at site $j \in [J]$, and c_j is the corresponding capacity of facility $j \in [J]$ if it is open. Parameter d_i is the demand quantity of customer $i \in [I]$. Parameter t_{ij} is the unit transportation cost for serving customer $i \in [I]$ by facility $j \in [J]$. Parameter p_i is the unit penalty cost at customer $i \in [I]$ for unmet demand. Binary variable $y_j = 1$ if facility $j \in [J]$ is open, $y_j = 0$ otherwise. Continuous variable x_{ij} denotes the product quantity transported from facility $j \in [J]$ to customer $i \in [I]$. Continuous variable u_i is the unmet demand at customer $i \in [I]$.

In a deterministic environment, facility capacity and customer demand are perfectly known at the time of making location decisions. In practice, these two parameter types are often subject

to uncertainty during the operational stage of a supply chain system. To reflect this reality, we denote facilities' uncertain capacities as a random vector $\tilde{\boldsymbol{c}} = (\tilde{c}_1, \ldots, \tilde{c}_J)^T$ and customers' uncertain demand as a random vector $\tilde{\boldsymbol{d}} = (\tilde{d}_1, \ldots, \tilde{d}_I)^T$. The event that causes these uncertainties is represented by a random variable $\tilde{e} \in \mathbb{R}$. If only one type of event is considered, $\tilde{e}$ can represent the magnitude or the level of influence of the event. We denote the joint capacity, demand, and related event as $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{e}) \in \mathbb{R}^J \times \mathbb{R}^I \times \mathbb{R}$.

3.2. Two-stage Stochastic Programming Model

If the decision maker has perfect knowledge of the joint distribution of $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{e})$, say, $\mathbb{P} \in \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I \times \mathbb{R})$, then we can formulate the problem as the following two-stage SP model:

$$\min_{\boldsymbol{y}} \sum_{j\in[J]} f_j y_j + \mathbb{E}_{\mathbb{P}}\left[h(\boldsymbol{y}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}})\right] \tag{1a}$$

$$\text{s.t. } y_j \in \{0,1\} \qquad \forall j \in [J], \tag{1b}$$

where $h(\boldsymbol{y}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}})$ is the second-stage recourse cost. Under a given location decision $\boldsymbol{y}$ and a realization $(\boldsymbol{c}, \boldsymbol{d})$ of uncertain parameters $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}})$, we have

$$h(\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}) = \min_{\boldsymbol{x},\boldsymbol{u}} \sum_{i\in[I]}\sum_{j\in[J]} t_{ij}x_{ij} + \sum_{i\in[I]} p_i u_i \tag{2a}$$

$$\text{s.t. } \sum_{j\in[J]} x_{ij} + u_i \geq d_i \qquad \forall i \in [I], \tag{2b}$$

$$\sum_{i\in[I]} x_{ij} \leq c_j y_j \qquad \forall j \in [J], \tag{2c}$$

$$x_{ij} \geq 0 \qquad \forall i \in [I], j \in [J], \tag{2d}$$

$$u_i \geq 0 \qquad \forall i \in [I]. \tag{2e}$$

The objective function (1a) minimizes the sum of the first-stage location cost and the expected second-stage recourse cost. Equation (2a) indicates that the recourse cost is composed of two parts: transportation costs for serving demand and penalty costs for unmet demand. Constraints (2b) impose that the sum of quantity received and unmet demand at each customer must be equal to or larger than that customer's demand. Constraints (2c) specify that only open facilities can serve customers and that each facility's capacity constraint must be respected. Constraints (2d)–(2e) define the non-negativity of variables.

The SP model (1) relies on the assumption that $\mathbb{P}$ is perfectly known. However, in practice, the probability distribution needs to be estimated from historical data. One commonly used method is SAA. Specifically, we assume that there are L samples of historical observations denoted as $\mathcal{L} = \{(\hat{\boldsymbol{c}}_1, \hat{\boldsymbol{d}}_1, \hat{e}_1), (\hat{\boldsymbol{c}}_2, \hat{\boldsymbol{d}}_2, \hat{e}_2), \ldots, (\hat{\boldsymbol{c}}_L, \hat{\boldsymbol{d}}_L, \hat{e}_L)\}$. The SAA method uses the empirical distribution to

approximate the true distribution $\mathbb{P}$, where each sample has an equal occurrence probability $1/L$. Thus, the SAA model of (1) is

$$\min \sum_{j\in[J]} f_j y_j + \frac{1}{L}\sum_{l=1}^{L}\left(\sum_{i\in[I]}\sum_{j\in[J]} t_{ij}x_{ijl} + \sum_{i\in[I]} p_i u_{il}\right) \tag{3a}$$

$$\text{s.t.} \sum_{j\in[J]} x_{ijl} + u_{il} \geq \hat{d}_{il} \qquad \forall i\in[I], l\in[L], \tag{3b}$$

$$\sum_{i\in[I]} x_{ijl} \leq \hat{c}_{jl} y_j \qquad \forall j\in[J], l\in[L], \tag{3c}$$

$$y_j \in \{0,1\} \qquad \forall j\in[J], \tag{3d}$$

$$x_{ijl} \geq 0 \qquad \forall i\in[I], j\in[J], l\in[L], \tag{3e}$$

$$u_{il} \geq 0 \qquad \forall i\in[I], l\in[L], \tag{3f}$$

where $\hat{d}_{il}$ and $\hat{c}_{jl}$ are the i-th and j-th components of $\hat{\boldsymbol{d}}_l$ and $\hat{\boldsymbol{c}}_l$, respectively. Correspondingly, x_{ijl} and u_{il} are the recourse variables associated with sample $l\in[L]$.

3.3. Two-stage Distributionally Robust Model

Formulation (3) is a MILP model, which can be solved directly by commercial solvers. Nevertheless, the empirical distribution of uncertain parameters is often unavailable especially when dealing with a disruption. Even in the context where such empirical information can be readily available, the solution based on the stochastic model (3) may suffer from the issue of overfitting, resulting in poor performance when the out-of-sample distribution deviates from the true distribution. Therefore, instead of specifying $\mathbb{P}$ to follow a particular distribution, we assume that it belongs to a set of distributions (i.e., the distribution is ambiguous) that shares common distributional information, as these pieces of information are often available and reliable (Popescu 2007). To tackle the distributional ambiguity in our problem, we use the DRO framework together with a scenario-wise ambiguity set.

3.3.1. Scenario-wise Ambiguity Set. The DRO framework assumes that the true distribution of $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{e}) : \mathbb{P}\in\mathcal{P}(\mathbb{R}^J\times\mathbb{R}^I\times\mathbb{R})$ belongs to an ambiguity set $\mathcal{F}\subseteq\mathcal{P}(\mathbb{R}^J\times\mathbb{R}^I\times\mathbb{R})$, which is characterized by partial distributional information estimated from historical data. The construction of $\mathcal{F}$ is key to the solution of the DRO model and its out-of-sample performance. We propose to construct a scenario-wise ambiguity set for our problem. In particular, we partition all possible realizations of $\tilde{e}$, denoted by $\mathcal{E}$, into S non-overlapping scenarios: $\mathcal{E}_s, s\in[S]$ with $\mathcal{E}_s\cap\mathcal{E}_{s'}=\emptyset$, for all $s,s'\in[S], s\neq s'$ and $\cup_{s\in[S]}\mathcal{E}_s=\mathcal{E}$. Let q_s denote the probability of scenario $s\in[S]$. By definition, we have $\mathbb{P}(\tilde{e}\in\mathcal{E}_s)=q_s$ and $\sum_{s\in[S]} q_s=1$. Let $\mathcal{U}_s, s\in[S]$ denote the index set associated with $\mathcal{E}_s$, that is, $\mathcal{U}_s=\{l\in[L]\mid \hat{e}_l\in\mathcal{E}_s\}$. We introduce a random variable $\tilde{s}$ taking discrete values in $[S]$ to

denote the scenario $\tilde{s} = s$ associated with $\mathcal{E}_s$. Accordingly, using historical data, we construct the scenario-wise ambiguity set $\mathcal{F} \subseteq \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I \times [S])$ associated with the random variable $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{s})$ as

$$\mathcal{F} = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I \times [S]) \;\middle|\; \begin{array}{ll} (\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{s}) \sim \mathbb{P} & \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{c}} \mid \tilde{s} = s\right] = \boldsymbol{\mu}_s & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{\boldsymbol{c}} - \boldsymbol{\mu}_s| \mid \tilde{s} = s\right] \leq \boldsymbol{\delta}_s & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{d}} \mid \tilde{s} = s\right] = \boldsymbol{\rho}_s & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{\boldsymbol{d}} - \boldsymbol{\rho}_s| \mid \tilde{s} = s\right] \leq \boldsymbol{\zeta}_s & \forall s \in [S] \\ \mathbb{P}\left[(\boldsymbol{c}, \boldsymbol{d}) \in \Omega_s \mid \tilde{s} = s\right] = 1 & \forall s \in [S] \\ \mathbb{P}\left[\tilde{s} = s\right] = q_s & \forall s \in [S] \end{array} \right\}, \tag{4}$$

where Ω_s is the support set associated with scenario $s \in [S]$, defined as

$$\Omega_s = \left\{ (\boldsymbol{c}, \boldsymbol{d}) \in \mathbb{R}^J \times \mathbb{R}^I \;\middle|\; \underline{\boldsymbol{c}}_s \leq \boldsymbol{c} \leq \bar{\boldsymbol{c}}_s, \underline{\boldsymbol{d}}_s \leq \boldsymbol{d} \leq \bar{\boldsymbol{d}}_s \right\}.$$

For each scenario $s \in [S]$, the mean capacity and demand are

$$\boldsymbol{\mu}_s = \frac{1}{|\mathcal{U}_s|} \sum_{l \in \mathcal{U}_s} \hat{\boldsymbol{c}}_l, \qquad \boldsymbol{\rho}_s = \frac{1}{|\mathcal{U}_s|} \sum_{l \in \mathcal{U}_s} \hat{\boldsymbol{d}}_l,$$

the mean absolute deviations are

$$\boldsymbol{\delta}_s = \frac{1}{|\mathcal{U}_s|} \sum_{l \in \mathcal{U}_s} |\hat{\boldsymbol{c}}_l - \boldsymbol{\mu}_s|, \qquad \boldsymbol{\zeta}_s = \frac{1}{|\mathcal{U}_s|} \sum_{l \in \mathcal{U}_s} |\hat{\boldsymbol{d}}_l - \boldsymbol{\rho}_s|,$$

the probability is

$$q_s = \frac{|\mathcal{U}_s|}{L},$$

and the parameters of the support set are

$$[\underline{\boldsymbol{c}}_s]_j = \min_{l \in \mathcal{U}_s} c_{jl}, \qquad [\bar{\boldsymbol{c}}_s]_j = \max_{l \in \mathcal{U}_s} c_{jl}, \quad \forall j \in [J],$$

$$[\underline{\boldsymbol{d}}_s]_i = \min_{l \in \mathcal{U}_s} d_{il}, \qquad [\bar{\boldsymbol{d}}_s]_i = \max_{l \in \mathcal{U}_s} d_{il}, \quad \forall i \in [I].$$

We observe that when $S = 1$, the scenario-wise ambiguity set (4) reduces to a marginal moment-based ambiguity set as follows:

$$\bar{\mathcal{F}} = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I) \;\middle|\; \begin{array}{l} (\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}) \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{c}}\right] = \boldsymbol{\mu} \\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{\boldsymbol{c}} - \boldsymbol{\mu}|\right] \leq \boldsymbol{\delta} \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{d}}\right] = \boldsymbol{\rho} \\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{\boldsymbol{d}} - \boldsymbol{\rho}|\right] \leq \boldsymbol{\zeta} \\ \mathbb{P}\left[(\boldsymbol{c}, \boldsymbol{d}) \in \Omega\right] = 1 \end{array} \right\}, \tag{5}$$

where the support set is defined as $\Omega = \left\{(\boldsymbol{c}, \boldsymbol{d}) \in \mathbb{R}^J \times \mathbb{R}^I \mid \underline{\boldsymbol{c}} \leq \boldsymbol{c} \leq \bar{\boldsymbol{c}}, \underline{\boldsymbol{d}} \leq \boldsymbol{d} \leq \bar{\boldsymbol{d}}\right\}$. In this case, only a single set of parameters will be derived from the entire data. We establish the relationship between the optimal solutions under $\mathcal{F}$ and $\bar{\mathcal{F}}$ in the next section. When $S = L$, each scenario encompasses exactly one sample, and the scenario-wise ambiguity set only contains the empirical distribution.

For a given set of samples $\mathcal{L}$, decision makers can flexibly construct the scenarios $\mathcal{U}_s, s \in [S]$ based on the types and the magnitudes of events. Without loss of generality, for example, the set of samples can take the following form

$$\mathcal{L} = \{(\hat{\boldsymbol{c}}_1, \hat{\boldsymbol{d}}_1, \hat{e}_1^N), \ldots, (\hat{\boldsymbol{c}}_m, \hat{\boldsymbol{d}}_m, \hat{e}_m^N), (\hat{\boldsymbol{c}}_{m+1}, \hat{\boldsymbol{d}}_{m+1}, \hat{e}_{m+1}^I), \ldots, (\hat{\boldsymbol{c}}_L, \hat{\boldsymbol{d}}_L, \hat{e}_L^I)\}, \tag{6}$$

where the subscript N represents the case where events $\hat{e}_l^N, l \in [m]$ are associated with natural disasters (like earthquakes and floods). The subscript I represents the case where events $\hat{e}_l^I, l \in \{m+1, \ldots, L\}$ are associated with industrial events (such as material shortages or strikes). Based on the types of disruption events, we can then partition the first m samples to form a scenario and the rest $L - m$ samples to form another scenario. Events of the same type can be further distinguished by their magnitudes. For example, suppose events $\hat{e}_l^N, l \in [m]$ in set (6) specifically refer to earthquakes, in this case, we can further divide samples $1, \ldots, m$ into different scenarios according to their associated magnitudes. On the other hand, if the event information is unknown, i.e., we only have historical capacity and demand samples, we can use clustering algorithms in machine learning to divide samples into different scenarios.

3.3.2. Two-stage DRO Model. Under the DRO framework, decision makers aim to minimize the sum of the first-stage location cost and the worst-case second-stage expected recourse cost over all possible distributions in $\mathcal{F}$. Thus, we construct the two-stage DRO model under a scenario-wise ambiguity set (abbreviated as SDR) as

$$\min_{\boldsymbol{y}} \left\{ \sum_{j \in [J]} f_j y_j + \sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}} \left[h(\boldsymbol{y}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{s}) \right] \right\}, \tag{7}$$

where the second-stage recourse cost relies on the realization s of the uncertain scenario $\tilde{s}$. In addition, for notational convenience, we denote the DRO model with the marginal moment-based ambiguity set $\bar{\mathcal{F}}$ as MDR.

3.4. Value of the Scenario-wise Ambiguity Set

This section first presents a small example to illustrate the value of the scenario-wise ambiguity set and then formally establishes the relationship between the optimal solutions generated by the SDR and MDR models.

We consider a simple supply chain system with one supplier and one retailer. We assume the underlying true model is as follows: $\tilde{e}$ follows a Bernoulli distribution with $\mathbb{P}(\tilde{e}=1)=1/2$. Conditioning on $\tilde{e}$, we have $\mathbb{P}(\tilde{c}=\hat{c}_1,\tilde{d}=\hat{d}_1\mid\tilde{e}=1)=1$ and $\mathbb{P}(\tilde{c}=\hat{c}_0,\tilde{d}=\hat{d}_0\mid\tilde{e}=0)=1$, where $\hat{c}_1\geq\hat{c}_0,\hat{d}_1\leq\hat{d}_0$. One can interpret, for example, $\tilde{e}=1$ and $\tilde{e}=0$ as the cases before and after a disruption event, respectively. These two cases naturally form $S=2$ scenarios, i.e., $\mathcal{E}_1=\{\tilde{e}=1\}$ and $\mathcal{E}_0=\{\tilde{e}=0\}$. We correspondingly construct the scenario-wise ambiguity set as

$$\mathcal{F}=\left\{\mathbb{P}\in\mathcal{P}(\mathbb{R}\times\mathbb{R}\times[2])\;\middle|\;\begin{array}{ll}(\tilde{c},\tilde{d},\tilde{s})\sim\mathbb{P}&\\ \mathbb{E}_{\mathbb{P}}\left[\tilde{c}\mid\tilde{s}=i\right]=\hat{c}_i & \text{for } i=0,1\\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{c}-\hat{c}_i|\mid\tilde{s}=i\right]\leq 0 & \text{for } i=0,1\\ \mathbb{E}_{\mathbb{P}}\left[\tilde{d}\mid\tilde{s}=i\right]=\hat{d}_i & \text{for } i=0,1\\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{d}-\hat{d}_i|\mid\tilde{s}=i\right]\leq 0 & \text{for } i=0,1\\ \mathbb{P}\left[(\tilde{c},\tilde{d})\in\mathbb{R}\times\mathbb{R}\mid\tilde{s}=i\right]=1 & \text{for } i=0,1\\ \mathbb{P}\left[\tilde{s}=i\right]=\frac{1}{2} & \text{for } i=0,1\end{array}\right\},$$

and the marginal moment-based ambiguity set as

$$\bar{\mathcal{F}}=\left\{\mathbb{P}\in\mathcal{P}(\mathbb{R}\times\mathbb{R})\;\middle|\;\begin{array}{l}(\tilde{c},\tilde{d})\sim\mathbb{P}\\ \mathbb{E}_{\mathbb{P}}\left[\tilde{c}\right]=\frac{\hat{c}_0+\hat{c}_1}{2}\\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{c}-\frac{\hat{c}_0+\hat{c}_1}{2}|\right]\leq\frac{|\hat{c}_0-\hat{c}_1|}{2}\\ \mathbb{E}_{\mathbb{P}}\left[\tilde{d}\right]=\frac{\hat{d}_0+\hat{d}_1}{2}\\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{d}-\frac{\hat{d}_0+\hat{d}_1}{2}|\right]\leq\frac{|\hat{d}_0-\hat{d}_1|}{2}\\ \mathbb{P}\left[(\tilde{c},\tilde{d})\in\mathbb{R}\times\mathbb{R}\right]=1\end{array}\right\}.$$

We observe that the scenario-wise ambiguity set $\mathcal{F}$ contains only a single distribution, i.e., $\mathbb{P}(\tilde{c}=\hat{c}_0,\tilde{d}=\hat{d}_0,\tilde{s}=0)=\frac{1}{2}$, $\mathbb{P}(\tilde{c}=\hat{c}_1,\tilde{d}=\hat{d}_1,\tilde{s}=1)=\frac{1}{2}$, which is the true distribution. In contrast, the marginal moment-based ambiguity set $\bar{\mathcal{F}}$ contains many possible distributions, for example $\mathbb{P}(\tilde{c}=\hat{c}_0,\tilde{d}=\hat{d}_0)=\frac{1}{2}$, $\mathbb{P}(\tilde{c}=\hat{c}_1,\tilde{d}=\hat{d}_1)=\frac{1}{2}$ and $\mathbb{P}(\tilde{c}=\frac{\hat{c}_0+\hat{c}_1}{2},\tilde{d}=\frac{\hat{d}_0+\hat{d}_1}{2})=1$, among others. We formally establish the benefits of using a scenario-wise ambiguity set in the following proposition.

PROPOSITION 1. *Let Π^{SDR} and Π^{MDR} be the optimal values of models SDR and MDR, respectively. Given ambiguity sets $\mathcal{F}$ and $\bar{\mathcal{F}}$, if (i) $\boldsymbol{\mu}=\sum_{s\in[S]}q_s\boldsymbol{\mu}_s$ and $\boldsymbol{\rho}=\sum_{s\in[S]}q_s\boldsymbol{\rho}_s$, (ii) $|\boldsymbol{\mu}-\boldsymbol{\mu}_s|\leq\boldsymbol{\delta}-\boldsymbol{\delta}_s$ and $|\boldsymbol{\rho}-\boldsymbol{\rho}_s|\leq\boldsymbol{\zeta}-\boldsymbol{\zeta}_s$ for all $s\in[S]$, and (iii) $\Omega=\cup_{s\in[S]}\Omega_s$, then $\Pi^{SDR}\leq\Pi^{MDR}$.*

Proof. See Appendix A.1. □

Proposition 1 suggests that the scenario-wise ambiguity set can produce less-conservative solutions for in-sample tests. The three conditions included in the proposition require consistency in parameter specification for both ambiguity sets, which naturally holds when these parameters are estimated from the same data set.

4. Solution Method

This section reformulates the SDR model to a MILP model and discusses an extension of the scenario-wise DRO framework.

4.1. MILP Reformulation of the DRO Model

Based on the work of Wiesemann et al. (2014), we first introduce auxiliary random vectors $\tilde{\boldsymbol{w}} \in \mathbb{R}^J$ and $\tilde{\boldsymbol{v}} \in \mathbb{R}^I$ to lift the ambiguity set (4) as

$$\mathcal{F}' = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^I \times \mathbb{R}^I \times [S]) \;\middle|\; \begin{array}{ll} (\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{d}}, \tilde{\boldsymbol{v}}, \tilde{s}) \sim \mathbb{P} & \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{c}} \mid \tilde{s} = s\right] = \boldsymbol{\mu}_s & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{w}} \mid \tilde{s} = s\right] \le \boldsymbol{\delta}_s & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{d}} \mid \tilde{s} = s\right] = \boldsymbol{\rho}_s & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{v}} \mid \tilde{s} = s\right] \le \boldsymbol{\zeta}_s & \forall s \in [S] \\ \mathbb{P}\left[(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{d}}, \tilde{\boldsymbol{v}}) \in \Omega'_s | \tilde{s} = s\right] = 1 & \forall s \in [S] \\ \mathbb{P}\left[\tilde{s} = s\right] = q_s & \forall s \in [S] \end{array} \right\}, \tag{8}$$

where Ω'_s is the lifted support set, defined as

$$\Omega'_s = \left\{ (\boldsymbol{c}, \boldsymbol{w}, \boldsymbol{d}, \boldsymbol{v}) \in \mathbb{R}^J \times \mathbb{R}^J \times \mathbb{R}^I \times \mathbb{R}^I \mid \underline{\boldsymbol{c}}_s \le \boldsymbol{c} \le \bar{\boldsymbol{c}}_s, |\boldsymbol{c} - \boldsymbol{\mu}_s| \le \boldsymbol{w}, \underline{\boldsymbol{d}}_s \le \boldsymbol{d} \le \bar{\boldsymbol{d}}_s, |\boldsymbol{d} - \boldsymbol{\rho}_s| \le \boldsymbol{v} \right\}.$$

Compared to the original ambiguity set $\mathcal{F}$, the terms inside the expectation constraints in the lifted ambiguity set $\mathcal{F}'$ are all linear, and the nonlinearities have been transferred to the support set Ω'_s. Next we reformulate the inner maximization problem in (7). Specifically, under a given location decision $\boldsymbol{y}$, $\mathbb{P}$ is the decision variable of problem $\sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[h(\boldsymbol{y}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{s})\right]$, i.e., we are choosing a distribution that maximizes the expected value of $h(\boldsymbol{y}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{s})$.

Proposition 2. *The term* $\sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[h(\boldsymbol{y}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{s})\right]$ *in (7) is equivalent to*

$$\min \sum_{s \in [S]} \Big\{ (\boldsymbol{\mu}_s)^T \boldsymbol{\alpha}_s + (\boldsymbol{\delta}_s)^T \boldsymbol{\beta}_s + (\boldsymbol{\rho}_s)^T \boldsymbol{\lambda}_s + (\boldsymbol{\zeta}_s)^T \boldsymbol{\gamma}_s + \max_{(\boldsymbol{c}, \boldsymbol{w}, \boldsymbol{d}, \boldsymbol{v}) \in \Omega'_s} \left[q_s h(\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}, s) - \left(\boldsymbol{c}^T \boldsymbol{\alpha}_s + \boldsymbol{w}^T \boldsymbol{\beta}_s + \boldsymbol{d}^T \boldsymbol{\lambda}_s + \boldsymbol{v}^T \boldsymbol{\gamma}_s \right) \right] \Big\} \tag{9a}$$

$$\text{s.t. } \boldsymbol{\beta}_s, \boldsymbol{\gamma}_s \ge 0 \qquad \forall s \in [S]. \tag{9b}$$

Proof. See Appendix A.2. □

Note that model (9) is not yet directly solvable by a commercial MILP solver in its present form for two reasons. First, $h(\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}, s)$ is a minimization problem, then a min-max-min problem is involved in (9a). Second, for the last term of (9a), we need to reformulate a min-max problem.

Scenario-wise adaptations. We utilize a scenario-wise adaptation policy introduced in Chen et al. (2020) for our problem, i.e., for each scenario we adopt different recourse decisions. To do this,

we define variables $\boldsymbol{x}$ and $\boldsymbol{u}$ as mappings of scenarios—an additional index s is featured for variables under each scenario, i.e., x_{ijs} and u_{is} for scenario $s \in [S]$. Thus, the term $\max_{(\boldsymbol{c},\boldsymbol{w},\boldsymbol{d},\boldsymbol{v})\in\Omega'_s} h(\boldsymbol{y},\boldsymbol{c},\boldsymbol{d},s)$ under scenario-wise adaptations is written as

$$\max_{(\boldsymbol{c},\boldsymbol{w},\boldsymbol{d},\boldsymbol{v})\in\Omega'_s} h(\boldsymbol{y},\boldsymbol{c},\boldsymbol{d},s) = \max_{(\boldsymbol{c},\boldsymbol{w},\boldsymbol{d},\boldsymbol{v})\in\Omega'_s} \min_{\boldsymbol{x}_s,\boldsymbol{u}_s} \sum_{i\in[I]}\sum_{j\in[J]} t_{ij}x_{ijs} + \sum_{i\in[I]} p_i u_{is} \tag{10a}$$

$$\text{s.t.} \sum_{j\in[J]} x_{ijs} + u_{is} \geq d_{is} \qquad \forall i \in [I], \tag{10b}$$

$$\sum_{i\in[I]} x_{ijs} \leq c_{js} y_j \qquad \forall j \in [J], \tag{10c}$$

$$x_{ijs} \geq 0 \qquad \forall i \in [I], j \in [J], \tag{10d}$$

$$u_{is} \geq 0 \qquad \forall i \in [I]. \tag{10e}$$

For a given scenario $s \in [S]$, model (10) finds the worst-case cost within the lifted support set Ω'_s. We can identify that under each scenario $s \in [S]$, the worst-case realization of uncertainties is attained at $d_{is} = \bar{d}_{is}, i \in [I], c_{js} = \underline{c}_{js}, j \in [J]$, thus leading to Proposition 3.

PROPOSITION 3. *For any* $\boldsymbol{y} \in \{0,1\}^J$, $\max_{(\boldsymbol{c},\boldsymbol{w},\boldsymbol{d},\boldsymbol{v})\in\Omega'_s} h(\boldsymbol{y},\boldsymbol{c},\boldsymbol{d},s)$ *is equivalent to*

$$\min_{\boldsymbol{x}_s,\boldsymbol{u}_s} \sum_{i\in[I]}\sum_{j\in[J]} t_{ij}x_{ijs} + \sum_{i\in[I]} p_i u_{is} \tag{11a}$$

$$\text{s.t.} \sum_{j\in[J]} x_{ijs} + u_{is} \geq \bar{d}_{is} \qquad \forall i \in [I], \tag{11b}$$

$$\sum_{i\in[I]} x_{ijs} \leq \underline{c}_{js} y_j \qquad \forall j \in [J], \tag{11c}$$

$$x_{ijs} \geq 0 \qquad \forall i \in [I], j \in [J], \tag{11d}$$

$$u_{is} \geq 0 \qquad \forall i \in [I]. \tag{11e}$$

In the case of a non-adaptive (or static) policy, we have $x_{ijs} = x_{ij}, u_{is} = u_i, s \in [S]$, so that the allocation decisions in the second stage do not change its solutions in response to the outcome of the scenario $\tilde{s}$. Correspondingly, constraints (11b) and (11c) become

$$\sum_{j\in[J]} x_{ij} + u_i \geq \max_{s\in[S]} \bar{d}_{is} \quad \forall i \in [I], \qquad \sum_{i\in[I]} x_{ij} \leq \min_{s\in[S]} \underline{c}_{js} y_j \quad \forall j \in [J].$$

This would be far more conservative, as the sum of satisfied and unsatisfied demand at a customer is equal to its maximal demand across all the scenarios, whereas the available capacity at a newly opened facility takes the minimal value in all the scenarios.

PROPOSITION 4. *The last term* $\max_{(\boldsymbol{c},\boldsymbol{w},\boldsymbol{d},\boldsymbol{v})\in\Omega'_s} -(\boldsymbol{c}^T\boldsymbol{\alpha}_s+\boldsymbol{w}^T\boldsymbol{\beta}_s+\boldsymbol{d}^T\boldsymbol{\lambda}_s+\boldsymbol{v}^T\boldsymbol{\gamma}_s)$ *in (9a) is equivalent to the following minimization problem*

$$\min \sum_{j\in[J]}\left\{\bar{c}_{js}A_{js}-\underline{c}_{js}B_{js}+\mu_{js}\left(D_{js}-E_{js}\right)\right\}+\sum_{i\in[I]}\left\{\bar{d}_{is}F_{is}-\underline{d}_{is}G_{is}+\rho_{is}\left(H_{is}-K_{is}\right)\right\} \tag{12a}$$

$$\text{s.t. } A_{js}-B_{js}+D_{js}-E_{js}=-\alpha_{js} \qquad \forall j\in[J], \tag{12b}$$

$$D_{js}+E_{js}=\beta_{js} \qquad \forall j\in[J], \tag{12c}$$

$$F_{is}-G_{is}+H_{is}-K_{is}=-\lambda_{is} \qquad \forall i\in[I], \tag{12d}$$

$$H_{is}+K_{is}=\gamma_{is} \qquad \forall i\in[I], \tag{12e}$$

$$A_{js},B_{js},D_{js},E_{js}\geq 0 \qquad \forall j\in[J], \tag{12f}$$

$$F_{is},G_{is},H_{is},K_{is}\geq 0 \qquad \forall i\in[I]. \tag{12g}$$

Proof. See Appendix A.3. □

Based on Propositions 2–4, we finally reformulate the SDR model to a MILP model as

$$\begin{aligned}\min \sum_{j\in[J]} f_j y_j+\sum_{s\in[S]}\Bigg\{&\sum_{j\in[J]}\left(\delta_{js}\beta_{js}+\bar{c}_{js}A_{js}-\underline{c}_{js}B_{js}+\mu_{js}\left(\alpha_{js}+D_{js}-E_{js}\right)\right)\\&+\sum_{i\in[I]}\left(\zeta_{is}\gamma_{is}+\bar{d}_{is}F_{is}-\underline{d}_{is}G_{is}+\rho_{is}\left(\lambda_{is}+H_{is}-K_{is}\right)\right)+\sum_{i\in[I]}\sum_{j\in[J]}q_s t_{ij}x_{ijs}+\sum_{i\in[I]}q_s p_i u_{is}\Bigg\}\end{aligned} \tag{13a}$$

$$\text{s.t. } (9b),(11b)\text{–}(11e) \text{ for each } s\in[S], \text{and } (12b)\text{–}(12g) \text{ for each } s\in[S]. \tag{13b}$$

We note that our scenario-wise DRO framework can be directly applied to the case of CFLPs with provider-side or receiver-side uncertainty (Baron et al. 2011, Zeng and Zhao 2013, An et al. 2014), which are the special cases of our problem. Specifically, for distributionally robust CFLPs with provider-side uncertainty, we can set parameters $\boldsymbol{\rho}_s, s\in[S]$ to the nominal demand at customers (denoted as $\boldsymbol{d}^n$) and parameters $\boldsymbol{\zeta}_s, s\in[S]$ to 0 in the ambiguity set (4). The bounds of demand can be set as $\underline{\boldsymbol{d}}_s=\bar{\boldsymbol{d}}_s=\boldsymbol{d}^n, s\in[S]$. Similarly, for distributionally robust CFLPs with receiver-side uncertainty, we can set parameters $\boldsymbol{\mu}_s, s\in[S]$ to facilities' nominal capacities (denoted as $\boldsymbol{c}^n$) and parameters $\boldsymbol{\delta}_s, s\in[S]$ to 0. The bounds of capacities are set to $\underline{\boldsymbol{c}}_s=\bar{\boldsymbol{c}}_s=\boldsymbol{c}^n, s\in[S]$. The CFLP with bimodal demand uncertainty in Shehadeh and Sanci (2021) is also a special case of our problem. In particular, to construct their ambiguity set, we set $S=2$. As the authors only specify the mean of demand in their ambiguity set, we can further let $\boldsymbol{\zeta}_s=0, s\in[S]$.

We can also apply the scenario-wise DRO framework to uncapacitated FLPs under simultaneous provider-side and receiver-side uncertainties or only one type of uncertainty (Snyder and Daskin 2005, Cui et al. 2010, Zetina et al. 2017), which are the special cases of our problem. Specifically, in the nominal disruption-free scenario, facilities are uncapacitated and can serve any customer once open; therefore, we can set a sufficiently large number for the capacity parameter $\mu_j, j\in[J]$. In a

disruption scenario, facility j will lose its service capability completely if it is disrupted, and thus $\mu_j = 0$. For customers, they have different patterns of demand in the nominal and the disruption scenarios.

4.2. An Extension of the Scenario-wise DRO Framework

To prepare for disaster threats, decision makers can simultaneously optimize the location decision, the inventory pre-positioning, and the relief delivery operations by solving a location and inventory pre-positioning problem (LIPP) (Ni et al. 2018, Velasquez et al. 2020, Shehadeha and Tucker 2020), which is an extension of the FLP presented in this paper. Under multiple disasters of different levels, or the same type of disaster with different damage levels, system parameters like the usable proportion of pre-positioned inventories, the capacity of transportation links, and the demand for emergency commodities, are subject to different levels of uncertainty. To solve the LIPP with simultaneous provider-side, in-between, and receiver-side uncertainties, Ni et al. (2018) propose a min-max robust model and use three budgeted uncertainty sets to characterize randomness. Velasquez et al. (2020) utilize a two-stage RO method for the LIPP with simultaneous provider-side and receiver-side uncertainties. Shehadeha and Tucker (2020) use a DRO method, together with a marginal moment-based ambiguity set and a static policy, to solve the LIPP under multiple types of uncertainties. We note that the scenario-wise DRO framework, together with the adaptation policy proposed in our work, can be directly adopted to solve the LIPP with simultaneous provider-side and receiver-side uncertainties. We provide the corresponding DRO model as follows and empirical insights for this problem in Section 5.3.

The LIPP is defined as follows. Consider a single-commodity network for the distribution of emergency supplies to prepare for a disaster event, which is characterized by an undirected graph containing J candidate facilities and I customers. The maximal capacity at facility $j \in [J]$ is M_j. A total amount of emergency supplies, denoted as R, is available before the disaster. In the pre-disaster phase, we decide where to open facilities and how much inventory to pre-position at each opened facility. Binary variable $y_j = 1$ if a facility is opened at site $j \in [J]$, $y_j = 0$ otherwise. The fixed cost of locating a facility at site $j \in [J]$ is f_j. Continuous variable a_j is the quantity of commodity pre-positioned at facility $j \in [J]$ if opened. For each unit of commodity handled (including purchasing, transporting, and storing) at site $j \in [J]$, there is an associated cost h_j. After a disaster, the proportion of pre-positioned inventory that is still usable at a facility and the demand at customers are subject to uncertainties, which are denoted as $\tilde{r}_j, j \in [J]$ and $\tilde{d}_i, i \in [I]$, respectively. There are two types of post-disaster (continuous) decision variables: x_{ij} is the commodity quantity transported from facility $j \in [J]$ to customer $i \in [I]$ and z_i^- is the quantity of unmet demand at customer $i \in [I]$. For these variables, the corresponding cost parameters per unit are t_{ij} and q_i^-,

respectively. In addition, a penalty cost q_j^+ is associated with each unit of unused inventory at facility $j \in [J]$.

We use the following scenario-wise ambiguity set to characterize distribution ambiguity

$$\mathcal{F} = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I \times [S]) \;\middle|\; \begin{array}{ll} (\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{d}}, \tilde{s}) \sim \mathbb{P} & \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{r}} \mid \tilde{s} = s\right] = \boldsymbol{\mu}_s^r & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{\boldsymbol{r}} - \boldsymbol{\mu}_s^r| \mid \tilde{s} = s\right] \le \boldsymbol{\delta}_s^r & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{d}} \mid \tilde{s} = s\right] = \boldsymbol{\mu}_s^d & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[|\tilde{\boldsymbol{d}} - \boldsymbol{\mu}_s^d| \mid \tilde{s} = s\right] \le \boldsymbol{\delta}_s^d & \forall s \in [S] \\ \mathbb{P}\left[(\boldsymbol{r}, \boldsymbol{d}) \in \Omega_s \mid \tilde{s} = s\right] = 1 & \forall s \in [S] \\ \mathbb{P}\left[\tilde{s} = s\right] = q_s & \forall s \in [S] \end{array} \right\}, \tag{14}$$

where Ω_s is the support set associated with scenario $s \in [S]$, defined as

$$\Omega_s = \left\{ (\boldsymbol{r}, \boldsymbol{d}) \in \mathbb{R}^J \times \mathbb{R}^I \;\middle|\; \underline{\boldsymbol{r}}_s \le \boldsymbol{r} \le \bar{\boldsymbol{r}}_s, \underline{\boldsymbol{d}}_s \le \boldsymbol{d} \le \bar{\boldsymbol{d}}_s \right\}.$$

The corresponding two-stage distributionally robust model is formulated as

$$\min_{\boldsymbol{y}, \boldsymbol{a}} \left\{ \sum_{j \in [J]} (f_j y_j + h_j a_j) + \sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}} \left[\mathcal{Q}(\boldsymbol{y}, \boldsymbol{a}, \tilde{\boldsymbol{r}}, \tilde{\boldsymbol{d}}, \tilde{s}) \right] \right\}, \tag{15a}$$

$$\text{s.t.} \sum_{j \in [J]} a_j = R, \tag{15b}$$

$$a_j \le M_j y_j \qquad \forall j \in [J], \tag{15c}$$

$$y_j \in \{0, 1\} \qquad \forall j \in [J], \tag{15d}$$

$$a_j \ge 0 \qquad \forall j \in [J]. \tag{15e}$$

Under a given first-stage decision $(\boldsymbol{y}, \boldsymbol{a})$ and a realization $(\boldsymbol{r}, \boldsymbol{d}, s)$ of uncertain parameters $(\tilde{\boldsymbol{r}}, \tilde{\boldsymbol{d}}, \tilde{s})$, the second-stage problem is defined as

$$\mathcal{Q}(\boldsymbol{y}, \boldsymbol{a}, \boldsymbol{r}, \boldsymbol{d}, s) = \min_{\boldsymbol{x}, \boldsymbol{z}^-} \sum_{i \in [I]} \sum_{j \in [J]} t_{ij} x_{ij} + \sum_{j \in [J]} \left(a_j r_j - \sum_{i \in [I]} x_{ij} \right) q_j^+ + \sum_{i \in [I]} q_i^- z_i^- \tag{16a}$$

$$\text{s.t.} \sum_{i \in [I]} x_{ij} \le a_j r_j \qquad \forall j \in [J], \tag{16b}$$

$$\sum_{j \in [J]} x_{ij} + z_i^- \ge d_i \qquad \forall i \in [I], \tag{16c}$$

$$x_{ij} \ge 0 \qquad \forall i \in [I], j \in [J], \tag{16d}$$

$$z_i^- \ge 0 \qquad \forall i \in [I]. \tag{16e}$$

Objective function (15a) minimizes the sum of the first-stage cost (including the location cost and the inventory pre-positioning cost) and the worst-case second-stage expected cost (including the

transportation cost, the penalty cost of unused inventory and unmet demand). Constraint (15b) means that all the available commodities are pre-positioned at facilities before the disaster. Constraints (15c) denote that we can only pre-position inventories at open facilities and that facilities' capacities must be respected. Constraints (16b) suggest that the total commodities delivered from a facility to customers cannot surpass the usable quantity at that facility. Constraints (16c) indicate the sum of met and unmet demand must be equal to or larger than a customer's demand. Constraints (15d), (15e), (16d), and (16e) define the type and non-negativity of variables.

5. Numerical Experiments

This section evaluates the performance of the scenario-wise DRO framework via extensive simulation tests and a case study. We compare the SAA, SDR, and MDR models in terms of cost and service level. For each model, we first solve it to generate a location decision and then evaluate its performance using out-of-sample tests. All models were coded in Python 3.8 programming language, using Gurobi 9.1.1 as the solver. The calculations were run on a personal computer with a 2 GHz Quad-Core Intel Core i5 processor and 16 GB of memory under the macOS Catalina system.

In following tables, the *Time* column reports the CPU time in seconds consumed to solve the models. $Cost_1$ is the first-stage location cost and $Cost_2$ is the expected second-stage recourse cost in out-of-sample tests. $Cost_t$ is the expected total cost, defined as the sum of $Cost_1$ and $Cost_2$. $\#U$ is the average quantity of unmet demand per customer per sample. $\#O$ is the number of facilities opened. The *percentage difference* (denoted as P-Diff) in cost is defined as $(Cost^{SDR} - Cost^{B})/Cost^{B} \times 100$, where $Cost^{SDR}$ is the cost of the SDR model and $Cost^{B}$ is the cost of the benchmark model, which can be SAA, MDR, or any other related model. The *quantity difference* (denoted as Q-Diff) is obtained by using the value of an indicator (*Time*, $\#U$, or $\#O$) in the SDR model to subtract the corresponding value in another model. Therefore, a negative value of P-Diff or Q-Diff suggests that the SDR model has better performance in that corresponding indicator.

5.1. Simulation Tests

We first introduce the instance set and then compare the results of different models.

5.1.1. Instance Set. We generate number-, coordinate-, and cost-related parameters based on the settings and assumptions widely made in the literature on facility location under uncertainty (Lei et al. 2016, Basciftci et al. 2021, Shehadeh and Sanci 2021). Specifically, the number of facilities ranges from 5 to 100 and the number of customers is between 10 and 100. In total, there are 12 combinations of (J, I). The facilities and customers are uniformly distributed on a 100 by 100 plane. Parameter t_{ij} is set to the Euclidean distance between customer $i \in [I]$ and facility $j \in [J]$. The unit penalty cost at customer $i \in [I]$ is set as $p_i = \max_{j \in [J]} t_{ij}$. The fixed cost $f_j, j \in [J]$ of facilities is a random integer in the interval $[2000, 5000]$.

To generate samples of capacity and demand, we consider two data sets: the in-sample training data to derive the location decision $y_j, j \in [J]$ and the out-of-sample testing data to evaluate its performance. As the simulation experiments in Hao et al. (2020), we consider $S = 4$ scenarios for both data sets, where each has an equal probability of $q_s = 1/4, s \in [S]$. We generate 20 samples under each scenario, for a total of 80 training samples and 80 testing samples for each instance. For each scenario, the capacity and demand samples are generated from uniform distributions with different means. For the training data, $c_{js} \sim U[a_s^1, b_s^1]$ and $d_{is} \sim U[a_s^2, b_s^2]$, where $a_s^1 = 280 - 30s$, $b_s^1 = 330 - 30s$, $a_s^2 = 10 + 10s$, and $b_s^2 = 30 + 10s$. For the testing data, $c_{js} \sim U[a_s^1(1-\Delta^1), b_s^1(1-\Delta^1)]$ and $d_{is} \sim U[a_s^2(1+\Delta^2), b_s^2(1+\Delta^2)]$, where Δ^1 and Δ^2 are perturbations used to capture the possibility that an out-of-sample distribution may deviate from the in-sample distribution. Δ^1 and Δ^2 take values from the set $\{0.10, 0.15, \ldots, 0.30\}$ and $\{-0.30, -0.25, \ldots, -0.10, 0.10, 0.15, \ldots, 0.30\}$, respectively. For a fixed value of $(I, J, \Delta^1, \Delta^2)$, we generate 5 instances and thus there are 3000 instances in total. We calculate the parameters of the two ambiguity sets based on the training samples. The solutions of SAA are also derived from the same training sample set.

5.1.2. Comparisons of Models. Tables 1–3 summarize the comparison results among different stochastic and robust models, where columns *95th* report the 95th percentiles of the expected total cost. The detailed results are reported in Tables D1 and D2 in Appendix D.

Comparison between SDR and SAA. The MILP reformulation of SDR can be efficiently solved—all the instances are solved to optimality within 2.08 seconds (refer to Table D1). The average computing time of SAA is greater than that of SDR for all the instances as shown in Table 1. In particular, when (J, I)=(100, 100), the average CPU time of SAA is more than 145 seconds longer than that of SDR.

Table 1 Comparison of average results between SDR and SAA models

		$\Delta_2 < 0$							$\Delta_2 > 0$						
		Q-Diff			P-Diff in cost				Q-Diff			P-Diff in cost			
No.	(J, I)	Time	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	95th	Time	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	95th
1	(5, 10)	-0.10	-2.20	0.50	22.79	-10.06	3.37	4.96	-0.09	-3.92	0.42	18.16	-8.97	-1.32	-1.68
2	(10, 10)	-0.14	-0.94	0.59	21.07	-11.68	4.28	0.82	-0.14	-3.80	0.65	23.16	-15.92	-2.82	-1.50
3	(10, 20)	-0.24	-1.54	1.05	22.70	-12.61	4.64	1.91	-0.24	-4.10	1.02	21.83	-14.69	-2.93	-4.42
4	(20, 20)	-1.98	-1.07	1.41	26.57	-16.45	7.05	3.50	-2.02	-4.69	1.52	28.69	-24.60	-5.71	-7.59
5	(15, 30)	-1.05	-1.48	1.59	23.79	-15.44	5.61	0.43	-1.06	-4.07	1.57	23.17	-17.82	-3.88	-6.61
6	(30, 30)	-8.11	-0.91	2.19	27.61	-17.51	9.01	2.01	-7.79	-4.38	2.15	26.90	-25.72	-5.83	-7.28
7	(20, 40)	-2.45	-1.52	2.18	24.20	-18.02	5.74	2.06	-2.32	-3.90	2.03	22.60	-18.88	-4.24	-5.10
8	(40, 40)	-16.36	-0.82	2.94	27.27	-18.20	10.24	5.34	-15.66	-4.19	2.91	27.38	-28.41	-5.94	-8.96
9	(25, 50)	-4.52	-1.32	2.48	22.28	-18.12	5.61	0.44	-4.15	-3.74	2.43	21.47	-19.54	-4.50	-5.25
10	(50, 50)	-28.06	-0.73	3.82	28.25	-19.31	11.70	3.92	-25.15	-4.21	3.69	27.05	-30.02	-6.38	-8.00
11	(50, 100)	-28.72	-0.98	4.20	18.77	-19.46	5.86	0.66	-30.49	-2.90	4.00	17.50	-19.81	-4.51	-5.69
12	(100, 100)	-145.82	-0.67	8.39	31.00	-23.37	15.34	5.95	-147.62	-4.53	8.43	31.09	-38.34	-7.84	-11.68

From Table 1, we observe that SDR opens more facilities than SAA, resulting in a higher first-stage location cost. However, these additional facilities can significantly reduce the second-stage

recourse cost and the quantity of unmet demand in out-of-sample tests. When Δ_2 is negative, i.e., the out-of-sample demand is lower than the in-sample expected value, the expected total cost of SDR is higher than that of SAA. This is because we cannot take full advantage of the additional facilities opened in SDR for the recourse problem when demand is lower than expected. In this case, the increase in the first-stage location cost surpasses the saving in the recourse cost, leading to a higher total cost. Note that the objective of SAA is to minimize the expected total cost, which is different from the robust objective. In this simulation, since the scenarios are sampled and tested using known distributions (with correlated demand), we can expect SAA to perform relatively well when the demand is lower than expected. In contrast, when Δ_2 is positive, which is the case when a firm faces surge demand such as panic buying, the expected total cost of SDR is lower. We also observe that SDR generally saves more recourse costs when Δ_2 is positive, compared to the case of a negative Δ_2. This can be explained by the fact that the additional facilities opened in SDR can be used to satisfy increased demand and thus reduce the quantity of unsatisfied demand and the penalty cost. To summarize, the solutions provided by SDR can better satisfy customer demand, compared to those of SAA. If customer demand has a positive deviation from the predicted value in disruption scenarios, the solutions of SDR can significantly help mitigate the impact of such disruption in the supply chain system.

As the objective of SDR focuses on the extreme case, besides comparing it with SAA, we also consider comparing it with a risk-averse two-stage SP model, where the conditional-value-at-risk (CVaR) is used as the risk measure. According to Rockafellar et al. (2000), Rockafellar and Uryasev (2002), and Noyan (2012), the CVaR with a confidence level α ($\alpha \in (0,1]$), denoted by CVaR_α, is defined as

$$\text{CVaR}_\alpha(z) = \inf_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\alpha} \mathbb{E}[(z-\eta)_+] \right\}, \tag{17}$$

where z is the random variable and $(z-\eta)_+ = \max\{0, z-\eta\}$. Parameter α denotes the level of conservatism of a decision maker. When α is close to 1, the measure focuses on more extreme losses. For our problem, we are interested in controlling the risk of the second-stage recourse cost. When the samples of historical observations and the corresponding probabilities are known, the second-stage cost can be calculated as

$$h(\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}, e) = \min \left\{ \eta + \frac{1}{1-\alpha} \sum_{l=1}^{L} \frac{1}{L} \left(\sum_{i \in [I]} \sum_{j \in [J]} t_{ij} x_{ijl} + \sum_{i \in [I]} p_i u_{il} - \eta \right)_+ \right\}. \tag{18}$$

Thus, the risk-averse two-stage SP model with the risk measure CVaR, denoted by SAA-CVaR, can be formulated as a MILP, which is provided in Appendix B.

The comparison results between the SDR and the SAA-CVaR models are presented in Table 2. Note that since some instances with $(J, I) = (100, 100)$ cannot be solved to optimality within the

1800-second time limit for the SAA-CVaR model, we do not include their results here. Table 2 shows that SDR produces solutions with a lower total cost than SAA-CVaR in most cases, which further confirms the superiority of the scenario-wise DRO framework. When $\alpha = 0.99$, the cost gap between the two models is generally greater than that under $\alpha = 0.90$, because when α is closer to 1, the SAA-CVaR model is more conservative.

Table 2 Comparison of average results between SDR and SAA-CVaR models

No.	(J, I)	$\Delta_2 < 0$				$\Delta_2 > 0$			
		$\alpha = 0.90$		$\alpha = 0.99$		$\alpha = 0.90$		$\alpha = 0.99$	
		Q-Diff in $\#U$	P-Diff in Cost_t	Q-Diff in $\#U$	P-Diff in Cost_t	Q-Diff in $\#U$	P-Diff in Cost_t	Q-Diff in $\#U$	P-Diff in Cost_t
1	(5, 10)	0.80	-3.29	0.92	-5.00	2.85	-0.40	3.76	-0.99
2	(10, 10)	0.32	-4.36	0.36	-5.20	1.80	-0.10	2.12	-0.29
3	(10, 20)	0.57	-4.94	0.61	-5.54	2.23	-0.29	2.32	-0.38
4	(20, 20)	-0.06	0.89	-0.06	-0.09	-1.11	-1.64	-0.94	-1.70
5	(15, 30)	0.22	-2.48	0.30	-3.98	0.87	-0.26	1.49	-0.27
6	(30, 30)	-0.25	3.64	-0.13	1.86	-1.92	-2.71	-1.31	-2.26
7	(20, 40)	0.17	-2.56	0.29	-4.33	0.88	-0.03	1.36	0.06
8	(40, 40)	-0.14	4.56	-0.13	4.14	-1.92	-2.63	-1.77	-2.59
9	(25, 50)	0.23	-2.75	0.25	-3.13	0.85	-0.08	0.89	-0.33
10	(50, 50)	-0.30	7.01	-0.30	6.59	-2.38	-3.66	-2.37	-3.89
11	(50, 100)	0.04	-1.02	0.08	-1.87	-0.08	-0.92	0.15	-0.86

Comparison between SDR and MDR. Table 3 shows that MDR provides solutions with higher expected total costs compared to SDR whether Δ_2 is negative or positive. This is because, among the three models, MDR opens the largest number of facilities, resulting in the highest first-stage location cost. When demand is lower than the expected value after a disruption, i.e., Δ_2 is negative, the solutions of MDR would be much more conservative, resulting in 20.70% higher cost when (J, I)=(50,100). However, the quantity of unmet demand in SDR is only slightly higher—the difference in unmet demand varies between the interval $[0.03, 1.34]$. When Δ_2 is positive, the gap in the total cost between MDR and SDR decreases because the additional opened facilities in MDR can be utilized to satisfy the increased demand. However, SDR still has a lower value for the expected total cost. We note that although MDR sometimes provides lower 95th percentiles of the expected total cost than SDR when Δ_2 is positive, the gaps are quite smaller, compared to the case of a negative Δ_2 where SDR provides better 95th percentile values.

5.2. An Earthquake Case Study

In this section, we use an earthquake case study to compare models and perform sensitivity analyses. In 2010, a magnitude 7.1 earthquake hit Yushu County in Qinghai Province, PR China, causing social and economic damages on a massive scale. The network of affected areas includes 13 nodes and 15 links, which is presented in Appendix C. We make assumptions and set parameters based on

Table 3 Comparison of average results between SDR and MDR models

		$\Delta_2<0$						$\Delta_2>0$					
		Q-Diff		P-Diff in cost				Q-Diff		P-Diff in cost			
No.	(J,I)	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	95th	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	95th
1	(5, 10)	1.34	-0.92	-25.97	9.83	-11.06	-11.43	5.62	-0.86	-24.97	14.87	-2.60	-1.24
2	(10, 10)	0.44	-0.94	-24.54	13.51	-11.68	-13.08	3.67	-0.94	-23.66	23.12	-2.32	0.73
3	(10, 20)	0.86	-2.11	-29.23	11.75	-16.09	-12.44	5.41	-2.11	-29.17	27.75	-3.56	-2.25
4	(20, 20)	0.22	-1.64	-22.21	11.28	-12.93	-12.56	2.89	-1.50	-20.96	27.12	-1.78	0.52
5	(15, 30)	0.65	-3.10	-29.12	12.74	-17.79	-16.21	5.33	-3.26	-30.10	35.20	-3.91	-0.43
6	(30, 30)	0.15	-1.70	-17.75	10.29	-10.66	-10.51	2.33	-1.74	-17.48	24.86	-1.01	1.86
7	(20, 40)	0.58	-3.92	-27.82	13.99	-17.57	-14.91	4.78	-3.89	-28.32	37.28	-2.87	2.45
8	(40, 40)	0.14	-1.90	-15.89	9.84	-10.03	-8.88	1.79	-1.86	-15.38	23.90	-1.13	0.30
9	(25, 50)	0.60	-5.01	-28.72	15.89	-18.71	-14.27	4.92	-5.11	-29.17	44.51	-2.70	0.15
10	(50, 50)	0.08	-2.46	-15.59	9.96	-10.35	-9.76	2.00	-2.67	-16.54	28.84	-1.32	2.47
11	(50, 100)	0.55	-9.99	-29.26	22.02	-20.70	-14.20	4.60	-9.77	-28.70	59.35	-1.82	2.68
12	(100, 100)	0.03	-2.78	-10.49	7.72	-7.49	-7.05	1.02	-2.79	-10.67	20.49	-1.07	0.37

the work of Ni et al. (2018). We note that this case study has been widely used in the literature on facility location under uncertainty (e.g., Ni et al. (2018), Shehadeha and Tucker (2020), and Zhang et al. (2021a)) to explore the impact of disasters on supply chain systems. Meanwhile, earthquakes are one of the random events that we hedge against to improve supply chain robustness in our study. Thus, we consider it is reasonable to use this earthquake case study to analyze our modeling framework and derive managerial insights. To this end, we empirically compare the performance of the two robust models, namely SDR and MDR, to assess the value of the scenario-wise ambiguity set in this context. In addition, the results based on the SAA model, which optimizes the expected total cost, are also reported to provide the insights with respect to the costs of the robust solutions (incurred in both the first and recourse stages) compared to the case when the location decisions are made based on the optimal expected cost function.

Each node is treated as both a candidate facility site and a demand site, i.e., $[J]=[I]$. The unit transportation cost t_{ij} is set to the shortest path distance between nodes i and j, which is given in Table C1. We consider two scenarios, where $s=1$ and $s=2$ denote a major and a minor earthquake, respectively. The capacity of each facility in the nominal disruption-free scenario is set to 800 as in Ni et al. (2018), and the usable portion of the capacity at facility j after disaster s is denoted by $\tilde{r}_{js}$. For the in-sample training data, the demand $\tilde{d}_{is}$ and the usable portion of capacity $\tilde{r}_{js}$ are generated from truncated normal distributions $N(\mu_i^{ds},\sigma_i^{ds},0,+\infty)$ and $N(\mu_j^{rs},\sigma_j^{rs},0,1)$, respectively. When $s=1$, we set the means and the standard deviations as in Ni et al. (2018) to denote the M_s7.1 major earthquake, i.e., for all $i\in I$, $j\in J$, $\mu_i^{d1}=100$, $\sigma_i^{d1}=10$, $\sigma_j^{r1}=0.1$, and μ_j^{r1} is given in Table C2. When $s=2$, we let $\mu_i^{d2}=70$, $\sigma_i^{d2}=10$, $\sigma_j^{r2}=0.1$, and $\mu_j^{r2}=1.3\mu_j^{r1}$, which indicate that the expected demand is smaller and the expected residual capacity is larger under a minor earthquake than under a major earthquake. For the out-of-sample testing data, the means are

set to $\mu_i^{ds}(1+\varsigma^1)$ and $\mu_j^{rs}(1+\varsigma^2)$, and the values of other parameters are the same as those for generating the training data. ς^1 and ς^2 are perturbations, which both take values from the set $\{-0.3, -0.2, -0.1, 0.1, 0.2, 0.3\}$. For a fixed value of $(\varsigma^1, \varsigma^2)$, we generate 5 instances, so there are 180 instances in total. We generate 50 samples under each scenario, for a total of 100 training samples and 100 testing samples for each instance.

5.2.1. Results of Different Models. The results of different models are provided in Table 4. It shows that on average SAA (MDR) opens the lowest (largest) number of facilities, leading to the lowest (highest) first-stage location cost and the highest (lowest) second-stage recourse cost. In terms of the expected total cost, the average result of SDR is the best among the three models. Moreover, SDR also has the lowest 95th percentile value for the expected total cost, whereas SAA has the highest value. In addition, the cost of SAA has the largest variation, with a standard deviation up to 471.15; whereas the standard deviations of SDR and MDR are 71.15 and 59.86, respectively. In terms of unmet demand, the out-of-sample mean, 95th percentile, and standard deviation produced by SAA are also much greater than those of the two other models.

Table 4 Results of different models for the case study

	Average					95th percentile		Standard deviation	
Model	Cost_1	Cost_2	Cost_t	$\#U$	$\#O$	Cost_t	$\#U$	Cost_t	$\#U$
SDR	1268.59	270.72	1539.32	0.001	8.11	1658.67	0.001	71.15	0.012
SAA	836.11	727.27	1563.38	1.134	5.07	2659.60	7.620	471.15	2.823
MDR	1389.22	212.37	1601.58	0.000	8.77	1681.38	0.000	59.86	0.000

Table 5 Illustrations of location decisions produced by the three models for the case study

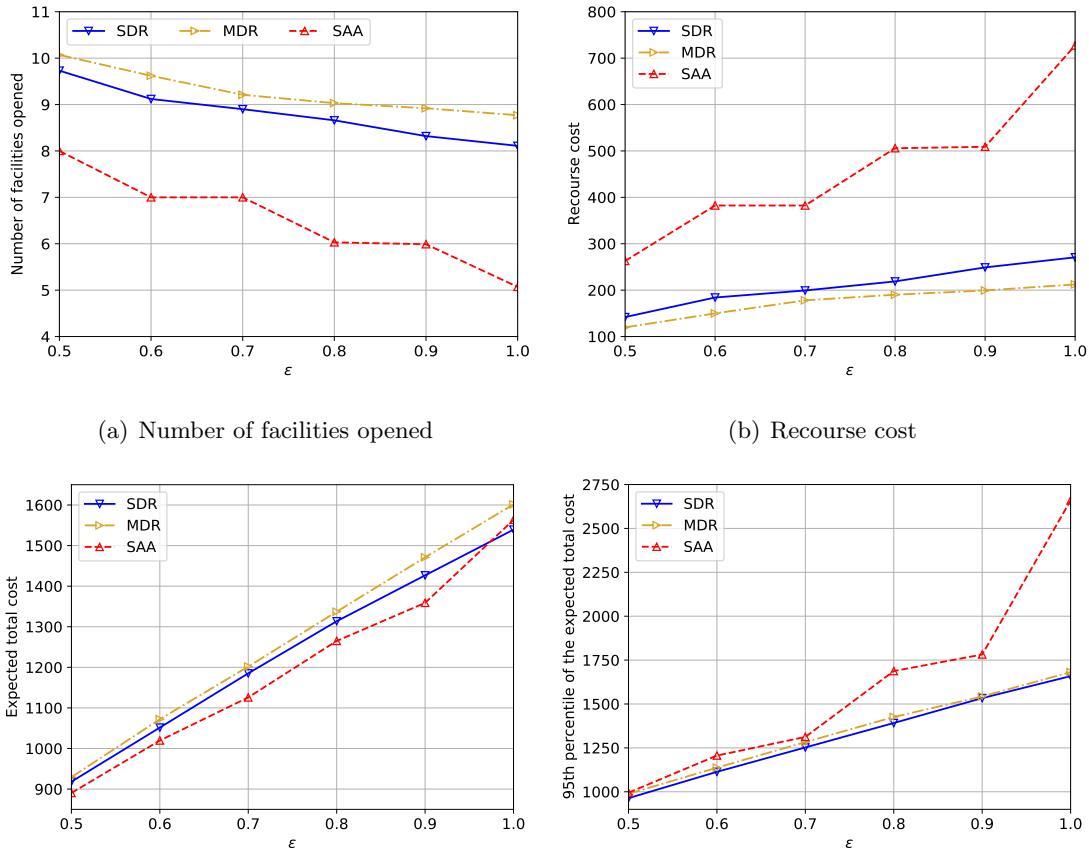
Case	$(\varsigma_1, \varsigma_2)$	Ins	Model	Facilities opened	$\#O$	Cost_1	Cost_2	Cost_t
1	(-0.3, -0.3)	1	SAA	2, 5, 8, 11, 12	5	828	400.34	1228.34
			SDR	1, 2, 3, 6, 8, 9, 10, 12	8	1269	178.98	1447.98
			MDR	1, 2, 3, 6, 8, 9, 10, 12	8	1269	178.98	1447.98
2	(-0.3, 0.1)	1	SAA	2, 5, 8, 11, 12	5	828	1433.74	2261.74
			SDR	2, 3, 6, 7, 8, 9, 10, 12	8	1233	346.64	1579.64
			MDR	①, 2, 3, 6, 7, 8, 9, 10, 12	9	1426	220.93	1646.93
3	(0.1, -0.3)	4	SAA	2, 3, 5, 8, 11, 12	6	945	309.37	1254.37
			SDR	1, 2, 3, 6, 7, 8, 9, 10, 12	9	1426	132.40	1558.40
			MDR	1, 2, 3, ⑤, 6, 7, 8, 9, 10, 12	10	1600	85.87	1685.87
4	(0.3, 0.1)	5	SAA	2, 5, 8, 11, 12	5	828	600.36	1428.36
			SDR	1, ②, 3, 6, 8, 9, 11, 12	8	1255	281.92	1536.92
			MDR	1, 3, 6, ⑦, 8, 9, 11, 12	8	1282	323.71	1605.71

Table 5 illustrates the location decisions provided by the three models, where *Ins* is the instance number (recall that 5 instances are generated for each $(\varsigma^1, \varsigma^2)$) and circles are used to distinguish

the decisions from the two robust models. We observe that sometimes SDR and MDR may produce the same location decisions, e.g., under Case 1. These two models also make different location decisions from two aspects: (i) *Locating a different number of facilities*. Under Cases 2 and 3, MDR opens one more facility than SDR. (ii) *Locating the same number of facilities but at different sites*, see Case 4.

5.2.2. Sensitivity Analyses. This section analyzes the sensitivity of solutions to two types of parameters: facilities' fixed costs and the support set's size. Graphical results are reported in Figures 1–2 and detailed results are provided in Tables D3–D4 in Appendix D.

Figure 1 The impact of facilities' fixed costs



(a) Number of facilities opened

(b) Recourse cost

(c) Expected total cost

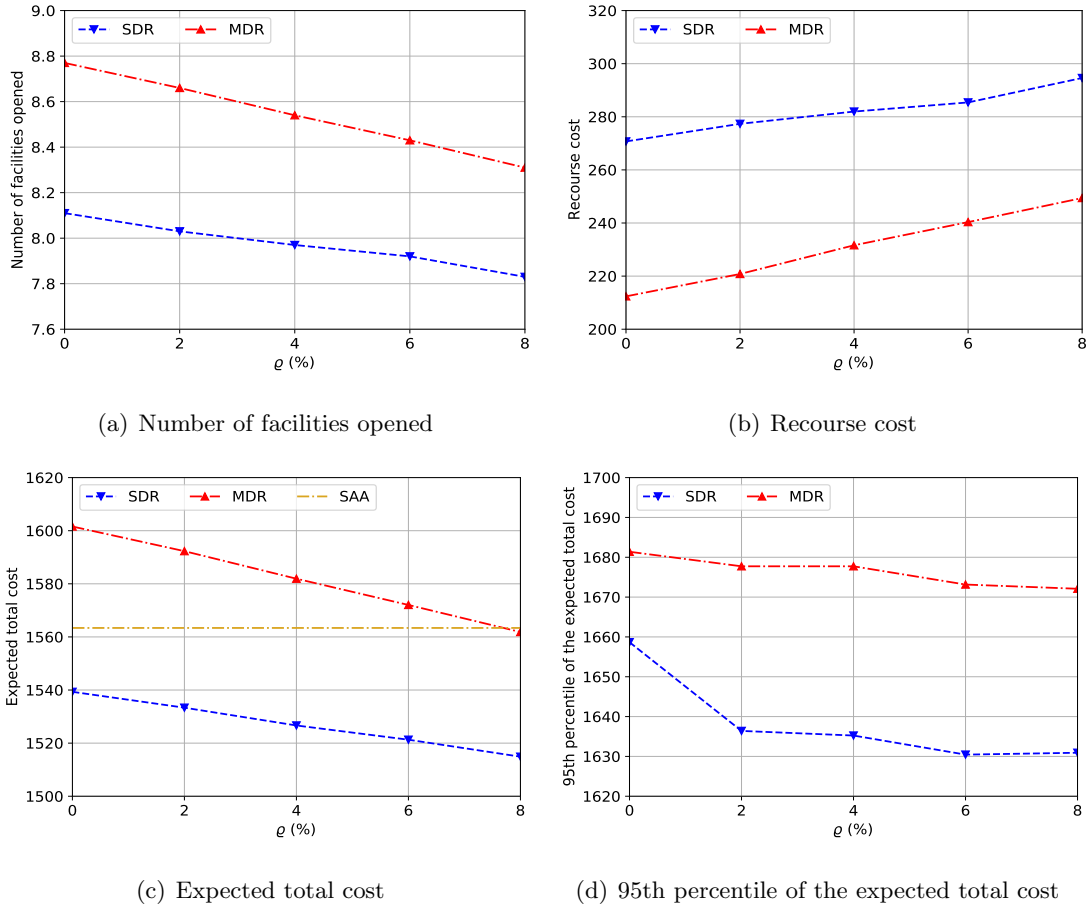
(d) 95th percentile of the expected total cost

Impact of facilities' fixed costs. We set facilities' fixed costs as $f'_j = \epsilon f_j, j \in [J]$ and vary the value of ϵ to study the impacts. Figures 1(a) show that MDR (SAA) opens the highest (lowest) number of facilities as before under a fixed value of ϵ, leading to the highest (lowest) location cost. That is to say, even when facilities are cheap, MDR (SAA) still produces the most conservative (optimistic) solutions. As a result, the recourse cost of SAA is much higher than those of the two other models as plotted in Figure 1(b). From Figure 1(c), we find that when ϵ is between 0.5

and 0.9, the expected total cost of SAA is lower than those of the robust models. However, the 95th percentile of the expected total cost produced by SAA is the highest, especially when ϵ is between 0.8 and 1.0, as shown in Figure 1(d). SAA also has the largest number of unmet demand in out-of-sample tests, as shown in Table D3. Thus, we can conclude from our analyses that SDR can still achieve a better trade-off between cost and service level even when it is inexpensive to open facilities.

Impact of the support set's size. When introducing the ambiguity set (4), we define the lower (upper) bounds of capacity and demand for each scenario as the minimal (maximal) values across all the samples within a scenario. Sometimes this setting may lead to conservative solutions when extreme cases exist in the samples. This section studies the impact of the support set's size. We shrink the support set by reducing the upper bounds as $\bar{\boldsymbol{c}}'_s = (1-\varrho)\bar{\boldsymbol{c}}_s, \bar{\boldsymbol{d}}'_s = (1-\varrho)\bar{\boldsymbol{d}}_s, s \in [S]$. A similar operation is also applied to the marginal moment-based ambiguity set (5) of MDR.

Figure 2 The impact of the support set's size



(a) Number of facilities opened

(b) Recourse cost

(c) Expected total cost

(d) 95th percentile of the expected total cost

From Figure 2, we can see that for both DRO models, narrowing the range of the support set can help reduce the number of open facilities (thus the location cost), the expected total cost, and

the 95th percentile of the expected total cost, as expected, and thus alleviate the conservatism of robust solutions. On the other hand, when we try to further reduce the upper bounds by setting $\varrho = 10\%$, SDR becomes infeasible for some instances. Thus, the parameters of the support set should be carefully decided to control the conservatism of robust solutions and also guarantee the feasibility of models. Figure 2 further shows that the recourse cost increases when the support set shrinks. However, it has almost no effect on the quantity of unmet demand as shown in Table D4.

Managerial insights for the CFLP with uncertainty. Based on the results in Sections 5.1 and 5.2, we make the following managerial conclusions for the CFLP with simultaneous provider-side and receiver-side uncertainties: (1) SAA produces relatively optimistic location decisions while MDR provides the most conservative location decisions. The location cost of SDR is between those of the two other models. (2) SDR can achieve a better trade-off between cost and service level. Specifically, compared to SAA, the solutions of SDR have lower values of unmet demand and thus a higher service level in out-of-sample tests. Moreover, SDR can also save the expected total cost when demand experiences positive deviations in out-of-samples. Compared to MDR, the solutions of SDR result in a lower expected total cost for the supply chain system, especially where demand has a negative deviation in out-of-samples. (3) SAA has significant performance variations both in terms of the total cost and the unmet demand, whereas the performance of MDR is the most stable among the three models, and SDR is in the middle. (4) The performance of SDR can be further improved by adjusting the size of the support set, e.g., shrinking the support set can help SDR find solutions with lower expected total costs.

5.3. Results for Location and Inventory Pre-Positioning Problem with Uncertainty

This section provides numerical results for the LIPP with uncertainty. As in Section 5.2, we consider two scenarios to represent a major and a minor earthquake, respectively. The total supply R can take any value in the set $\{2400, 2600, 2800, 3000\}$. The in-sample random residual inventory, in-sample random demand, and out-of-sample random demand under each scenario are generated using the same methods as in Section 5.2. For the out-of-sample random residual inventory, we set the means $\mu^r_{j1}, j \in [J]$ under scenario $s = 1$ as in Ni et al. (2018) and $\mu^r_{j2} = 1.3\mu^r_{j1}$ for scenario $s = 2$. All the related parameters are given in Table C2. For a fixed value of (R, ς^1), we generate 50 instances, so there are 1200 instances in total. We generate 50 samples under each scenario, for a total of 100 training samples and 100 testing samples for each instance. The average results are reported in Table 6, where Cost^L_1 and Cost^I_1 are the first-stage location cost and inventory pre-positioning cost, respectively.

Table 6 shows that SDR produces solutions with the lowest expected total cost and unmet demand, whereas SAA has the highest expected total cost and unmet demand. Moreover, SAA

Table 6 Average results of different models for the location and inventory pre-positioning problem

	Average							95th percentile		Standard deviation	
Model	Cost_1^L	Cost_1^I	Cost_1	Cost_2	Cost_t	$\#U$	$\#O$	Cost_t	$\#U$	Cost_t	$\#U$
SDR	616.00	7011.45	7627.45	2537.55	10165.00	8.28	3.75	12380.33	30.15	1259.04	9.66
SAA	731.54	6275.64	7007.18	3660.10	10667.28	20.75	4.63	14110.04	50.72	1986.34	15.22
MDR	501.37	6854.78	7356.14	3098.57	10454.71	14.36	3.75	16152.90	67.78	2121.55	17.68

opens the most facilities among the three models, which is different from what we have observed in the previous sections. We consider the reason is that besides location decisions, decision makers can also adjust the inventory pre-positioning decisions in the LIPP to improve supply chain robustness. That is, in the LIPP, decision makers have two types of strategies to enhance supply chain resilience, whereas they can only make location decisions in the CFLP for higher resilience. This also explains why SDR has the highest inventory pre-positioning cost and the lowest recourse cost. Finally, we emphasize that SDR outperforms MDR in both the expected total cost and the unmet demand for this case study, which further confirms the value of the scenario-wise ambiguity set for capturing event-correlated uncertainty.

6. Conclusions

This paper studies a capacitated fixed-charge location problem with both provider-side and receiver-side uncertainties. We use a DRO framework to solve the problem, where the joint distribution of uncertain parameters is assumed to lie in a scenario-wise ambiguity set, which can capture different levels of uncertainty resulting from different random events or from different magnitudes of the same event type. Correspondingly, we adopt a scenario-wise adaptation policy for the second-stage recourse problem to further mitigate the conservatism of robust solutions. The resulting adaptive DRO model is reformulated to a MILP model, which can be efficiently solved by off-the-shelf solvers. Our simulation and case study results show that the scenario-wise DRO framework can achieve a better trade-off between cost and service level for this type of problem. In particular, it provides less-conservative solutions and thus reduces the overall cost compared to the DRO model with a marginal moment-based ambiguity set. In comparison with the SAA model, the scenario-wise DRO model has lower values of unmet demand in out-of-sample tests, and thus it can better serve customers. We also discuss an extension of the proposed modeling framework. Specifically, we explicitly construct the scenario-wise ambiguity set and the distributionally robust model for the location and inventory pre-positioning problem under simultaneous provider-side and receiver-side uncertainties. We also provide numerical results for this problem, confirming that the scenario-wise ambiguity set is highly suitable for dealing with event-correlated uncertainty.

Future research could be conducted from two aspects: (i) The recourse variables $\boldsymbol{x}$ and $\boldsymbol{u}$ can be extended to scenario-wise affinely adaptive to random variables $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{d}}, \tilde{\boldsymbol{v}})$, *i.e.*, $\boldsymbol{x}$ and $\boldsymbol{u}$ are

affine functions of these variables in each scenario. This extension is expected to further reduce the conservatism of DRO solutions, and thus achieves even better performance for the case where demand has negative deviations from the expected values. (ii) Besides reallocating customers to facilities with residual capacity, we can consider other recourse actions to serve customers after a disruption event, such as goods sharing (goods are shipped from one facility to another) and subcontracting (a fraction of customer demand is satisfied by external subcontracting facilities) (Azad and Hassini 2019).

References

An, Yu, Bo Zeng, Yu Zhang, Long Zhao. 2014. Reliable p-median facility location problem: two-stage robust models and algorithms. *Transportation Research Part B: Methodological* **64** 54–72.

Atamtürk, Alper, Muhong Zhang. 2007. Two-stage robust network flow and design under demand uncertainty. *Operations Research* **55**(4) 662–673.

Azad, Nader, Elkafi Hassini. 2019. A benders decomposition method for designing reliable supply chain networks accounting for multimitigation strategies and demand losses. *Transportation Science* **53**(5) 1287–1312.

Baron, Opher, Joseph Milner, Hussein Naseraldin. 2011. Facility location: A robust optimization approach. *Production and Operations Management* **20**(5) 772–785.

Basciftci, Beste, Shabbir Ahmed, Siqian Shen. 2021. Distributionally robust facility location problem under decision-dependent stochastic demand. *European Journal of Operational Research* **292**(2) 548–561.

Ben-Tal, Aharon, Dick Den Hertog, Anja De Waegenaere, Bertrand Melenberg, Gijs Rennen. 2013. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science* **59**(2) 341–357.

Bertsimas, Dimitris, Melvyn Sim, Meilin Zhang. 2019. Adaptive distributionally robust optimization. *Management Science* **65**(2) 604–618.

Besson, Emilie Koum. 2020. Covid-19 (coronavirus): Panic buying and its impact on global health supply chains. Available at https://blogs.worldbank.org/health/covid-19-coronavirus-panic-buying-and-its-impact-global-health-supply-chains (last accessed date: January 17, 2022).

Chang, Zhiqi, Shiji Song, Yuli Zhang, Jian Ya Ding, Rui Zhang, Raymond Chiong. 2017. Distributionally robust single machine scheduling with risk aversion. *European Journal of Operational Research* **256**(1) 261–274.

Chen, Zhi, Melvyn Sim, Peng Xiong. 2020. Robust stochastic optimization made easy with RSOME. *Management Science* **66**(8) 3329–3339.

Cheng, Chun, Yossiri Adulyasak, Louis-Martin Rousseau. 2021. Robust facility location under demand uncertainty and facility disruptions. *Omega* **103** 102429.

Cheng, Chun, Yossiri Adulyasak, Louis-Martin Rousseau, Melvyn Sim. 2020. Robust drone delivery with weather information. Available at http://www.optimization-online.org/DB_FILE/2020/07/7897.pdf (last accessed date: January 17, 2022).

Chopra, Sunil, Manmohan Sodhi. 2014. Reducing the risk of supply chain disruptions. *MIT Sloan Management Review* **55**(3) 72–80.

Cui, Tingting, Yanfeng Ouyang, Zuo-Jun Max Shen. 2010. Reliable facility location design under the risk of disruptions. *Operations Research* **58**(4-part-1) 998–1011.

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

Du, Bo, Hong Zhou, Roel Leus. 2020. A two-stage robust model for a reliable p-center facility location problem. *Applied Mathematical Modelling* **77** 99–114.

Elçi, Özgün, Nilay Noyan. 2018. A chance-constrained two-stage stochastic programming model for humanitarian relief network design. *Transportation Research Part B: Methodological* **108** 55–83.

Ergun, Ozlem, Gonca Karakus, Pinar Keskinocak, Julie Swann, Monica Villarreal. 2011. Operations research to improve disaster supply chain management. *Wiley Encyclopedia of Operations Research and Management Science*. John Wiley & Sons, Inc., Hoboken, NJ.

Gao, Yuan, Zhongfeng Qin. 2016. A chance constrained programming approach for uncertain p-hub center location problem. *Computers & Industrial Engineering* **102** 10–20.

Hao, Zhaowei, Long He, Zhenyu Hu, Jun Jiang. 2020. Robust vehicle pre-allocation with uncertain covariates. *Production and Operations Management* **29**(4) 955–972.

Hou, Wenting, Rujie Zhu, Hua Wei, Hiep TranHoang. 2018. Data-driven affinely adjustable distributionally robust framework for unit commitment based on wasserstein metric. *IET Generation, Transmission & Distribution* **13**(6) 890–895.

Lei, Chao, Wei-Hua Lin, Lixin Miao. 2016. A two-stage robust optimization approach for the mobile facility fleet sizing and routing problem under uncertainty. *Computers & Operations Research* **67** 75–89.

Li, Runjie, Zheng Cui, Yong-Hong Kuo, Lianmin Zhang. 2022. Scenario-based distributionally robust optimization for the stochastic inventory routing problem Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4010328 (last accessed date: May 03, 2022).

Lu, Mengshi, Lun Ran, Zuo-Jun Max Shen. 2015. Reliable facility location design under uncertain correlated disruptions. *Manufacturing & Service Operations Management* **17**(4) 445–455.

Matthews, Logan R, Chrysanthos E Gounaris, Ioannis G Kevrekidis. 2019. Designing networks with resiliency to edge failures using two-stage robust optimization. *European Journal of Operational Research* **279**(3) 704–720.

Mazahir, Shumail, Amir Ardestani-Jaafari. 2020. Robust global sourcing under compliance legislation. *European Journal of Operational Research* **284**(1) 152–163.

Mišković, Stefan, Zorica Stanimirović, Igor Grujičić. 2017. Solving the robust two-stage capacitated facility location problem with uncertain transportation costs. *Optimization Letters* **11**(6) 1169–1184.

Mohajerin Esfahani, Peyman, Daniel Kuhn. 2018. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming* **171** 115–166.

Montgomery, Olivia. 2020. 5 types of supply chain disruption with covid-19 examples. Available at https://www.softwareadvice.com/resources/supply-chain-disruption-types (last accessed date: January 17, 2022).

Ni, Wenjun, Jia Shu, Miao Song. 2018. Location and emergency inventory pre-positioning for disaster response operations: Min-max robust model and a case study of Yushu earthquake. *Production & Operations Management* **27**(1) 160–183.

Nikoofal, Mohammad Ebrahim, Seyed Jafar Sadjadi. 2010. A robust optimization model for p-median problem with uncertain edge lengths. *The International Journal of Advanced Manufacturing Technology* **50** 391–397.

Noyan, Nilay. 2012. Risk-averse two-stage stochastic programming with an application to disaster management. *Computers & Operations Research* **39**(3) 541–559.

Noyan, Nilay, Burcu Balcik, Semih Atakan. 2016. A stochastic optimization model for designing last mile relief networks. *Transportation Science* **50**(3) 1092–1113.

Perakis, Georgia, Melvyn Sim, Qinshen Tang, Peng Xiong. 2022. Robust pricing and production with information partitioning and adaptation. *Management Science (forthcoming)*.

Popescu, Ioana. 2007. Robust mean-covariance solutions for stochastic optimization. *Operations Research* **55**(1) 98–112.

Rockafellar, R Tyrrell, Stanislav Uryasev. 2002. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* **26**(7) 1443–1471.

Rockafellar, R Tyrrell, Stanislav Uryasev, et al. 2000. Optimization of conditional value-at-risk. *Journal of Risk* **2** 21–42.

Saif, Ahmed, Erick Delage. 2021. Data-driven distributionally robust capacitated facility location problem. *European Journal of Operational Research* **291**(3) 995–1007.

Scarf, Herbert E. 1957. A min-max solution of an inventory problem. Tech. rep., Santa Monica, Calif.

Shehadeh, Karmel S, Ece Sanci. 2021. Distributionally robust facility location with bimodal random demand. *Computers & Operations Research* **134** 105257.

Shehadeha, Karmel S, Emily L Tucker. 2020. A distributionally robust optimization approach for location and inventory prepositioning of disaster relief supplies. Available at https://arxiv.org/abs/2012.05387 (last accessed date: January 17, 2022).

Shen, Zuo-Jun Max, Roger Lezhou Zhan, Jiawei Zhang. 2011. The reliable facility location problem: Formulations, heuristics, and approximation algorithms. *INFORMS Journal on Computing* **23**(3) 470–482.

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

Snyder, Lawrence V, Mark S Daskin. 2005. Reliability models for facility location: the expected failure cost case. *Transportation Science* **39**(3) 400–416.

Taherkhani, Gita, Sibel A Alumur, Mojtaba Hosseini. 2021. Robust stochastic models for profit-maximizing hub location problems. *Transportation Science* **55**(6) 1322–1350.

Velasquez, German A, Maria E Mayorga, Osman Y Özaltın. 2020. Prepositioning disaster relief supplies using robust optimization. *IISE Transactions* **52**(10) 1122–1140.

Wang, Shuming, Zhi Chen, Tianqi Liu. 2020. Distributionally robust hub location. *Transportation Science* **54**(5) 1189–1210.

Wiesemann, Wolfram, Daniel Kuhn, Melvyn Sim. 2014. Distributionally robust convex optimization. *Operations Research* **62**(6) 1358–1376.

Xie, Siyang, Yanfeng Ouyang. 2019. Reliable service systems design under the risk of network access failures. *Transportation Research Part E: Logistics and Transportation Review* **122** 1–13.

Zeng, Bo, Long Zhao. 2013. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters* **41**(5) 457–461.

Zetina, Carlos Armando, Ivan Contreras, Jean-François Cordeau, Ehsan Nikbakhsh. 2017. Robust uncapacitated hub location. *Transportation Research Part B: Methodological* **106** 393–410.

Zhang, Jianghua, Yang Liu, Guodong Yu, Zuo-Jun Shen. 2021a. Robustifying humanitarian relief systems against travel time uncertainty. *Naval Research Logistics (NRL)* **68**(7) 871–885.

Zhang, Mengling, Zihao Jiao, Wang Jing. 2021b. Improving humanitarian relief via optimizing shelter location with uncertain covariates. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3910550 (last accessed date: January 17, 2022).

Zhang, Yu, Zhenzhen Zhang, Andrew Lim, Melvyn Sim. 2021c. Robust data-driven vehicle routing with time windows. *Operations Research* **69**(2) 469–485.

Online Supplement

Appendix A Proof of Propositions

This section provides completed proofs of propositions.

A.1 Proof of Proposition 1

It is sufficient to show that for all $\mathbb{P} \in \mathcal{F}$, $\mathbb{P} \in \bar{\mathcal{F}}$ as well. If $\mathbb{P} \in \mathcal{F}$, according to the law of total expectation and condition (i), we have

$$\mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{c}}] = \sum_{s\in[S]} q_s \mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{c}} \mid \tilde{s} = s] = \sum_{s\in[S]} q_s \boldsymbol{\mu}_s = \boldsymbol{\mu},$$
$$\mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{d}}\right] = \sum_{s\in[S]} q_s \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{d}} \mid \tilde{s} = s\right] = \sum_{s\in[S]} q_s \boldsymbol{\rho}_s = \boldsymbol{\rho}.$$

From $|\boldsymbol{\mu} - \boldsymbol{\mu}_s| = |(\tilde{\boldsymbol{c}} - \boldsymbol{\mu}) - (\tilde{\boldsymbol{c}} - \boldsymbol{\mu}_s)| \geq |\tilde{\boldsymbol{c}} - \boldsymbol{\mu}| - |\tilde{\boldsymbol{c}} - \boldsymbol{\mu}_s|$ and condition (ii), we have

$$|\tilde{\boldsymbol{c}} - \boldsymbol{\mu}| - |\tilde{\boldsymbol{c}} - \boldsymbol{\mu}_s| \leq \boldsymbol{\delta} - \boldsymbol{\delta}_s \iff |\tilde{\boldsymbol{c}} - \boldsymbol{\mu}| - \boldsymbol{\delta} \leq |\tilde{\boldsymbol{c}} - \boldsymbol{\mu}_s| - \boldsymbol{\delta}_s.$$

According to the law of total expectation, we have

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}[|\tilde{\boldsymbol{c}} - \boldsymbol{\mu}|] - \boldsymbol{\delta} &= \sum_{s\in[S]} q_s \mathbb{E}_{\mathbb{P}}[|\tilde{\boldsymbol{c}} - \boldsymbol{\mu}| - \boldsymbol{\delta} \mid \tilde{s} = s] \\ &\leq \sum_{s\in[S]} q_s \mathbb{E}_{\mathbb{P}}[|\tilde{\boldsymbol{c}} - \boldsymbol{\mu}_s| - \boldsymbol{\delta}_s \mid \tilde{s} = s] \\ &\leq 0.\end{aligned}$$

Similarly, we can get $\mathbb{E}_{\mathbb{P}}\left[|\tilde{\boldsymbol{d}} - \boldsymbol{\rho}|\right] - \boldsymbol{\zeta} \leq 0$.

The above inequalities suggest that $\mathbb{P} \in \mathcal{F}$ satisfy the first four constraints of set $\bar{\mathcal{F}}$. Moreover, combining $\mathbb{P}[(\boldsymbol{c}, \boldsymbol{d}) \in \Omega_s \mid \tilde{s} = s] = 1, s \in [S]$ and condition (iii), we have $\mathbb{P}[(\boldsymbol{c}, \boldsymbol{d}) \in \Omega] = 1$. Thus, $\mathbb{P} \in \bar{\mathcal{F}}$.

A.2 Proof of Proposition 2

We apply duality theory to reformulate the *sup* term in (7). Using the law of total probability, we can construct the joint distribution $\mathbb{P}$ of $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{d}}, \tilde{\boldsymbol{v}}, \tilde{s})$ from the marginal distribution $\hat{\mathbb{P}}$ of $\tilde{s}$ supported on $[S]$ and the conditional distribution $\mathbb{P}_s$ of $(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{d}}, \tilde{\boldsymbol{v}})$ given $\tilde{s} = s, s \in [S]$. In this way, $\sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}[h(\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}, s)]$ can be rewritten as

$$\sup \sum_{s\in[S]} q_s \mathbb{E}_{\mathbb{P}_s}\left[h(\boldsymbol{y}, \tilde{\boldsymbol{c}}, \tilde{\boldsymbol{d}}, \tilde{s})\right] \tag{A.1a}$$
$$\text{s.t. } \mathbb{E}_{\mathbb{P}_s}[\tilde{\boldsymbol{c}}] = \boldsymbol{\mu}_s \qquad \forall s \in [S], \tag{A.1b}$$
$$\mathbb{E}_{\mathbb{P}_s}[\tilde{\boldsymbol{w}}] \leq \boldsymbol{\delta}_s \qquad \forall s \in [S], \tag{A.1c}$$
$$\mathbb{E}_{\mathbb{P}_s}\left[\tilde{\boldsymbol{d}}\right] = \boldsymbol{\rho}_s \qquad \forall s \in [S], \tag{A.1d}$$
$$\mathbb{E}_{\mathbb{P}_s}[\tilde{\boldsymbol{v}}] \leq \boldsymbol{\zeta}_s \qquad \forall s \in [S], \tag{A.1e}$$
$$\mathbb{P}_s\left[(\tilde{\boldsymbol{c}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{d}}, \tilde{\boldsymbol{v}}) \in \Omega'_s\right] = 1 \qquad \forall s \in [S]. \tag{A.1f}$$

Let $\boldsymbol{\alpha}_s \in \mathbb{R}^J$, $\boldsymbol{\beta}_s \in \mathbb{R}^J$, $\boldsymbol{\lambda}_s \in \mathbb{R}^I$, $\boldsymbol{\gamma}_s \in \mathbb{R}^I$, and $\theta_s \in \mathbb{R}$ be the dual variables associated with constraints (A.1b)–(A.1f), respectively. Then we can derive the dual problem of model (A.1) as

$$\min \sum_{s\in[S]} \left[(\boldsymbol{\mu}_s)^T\boldsymbol{\alpha}_s + (\boldsymbol{\delta}_s)^T\boldsymbol{\beta}_s + (\boldsymbol{\rho}_s)^T\boldsymbol{\lambda}_s + (\boldsymbol{\zeta}_s)^T\boldsymbol{\gamma}_s + \theta_s\right] \tag{A.2a}$$

$$\text{s.t. } \boldsymbol{c}^T\boldsymbol{\alpha}_s + \boldsymbol{w}^T\boldsymbol{\beta}_s + \boldsymbol{d}^T\boldsymbol{\lambda}_s + \boldsymbol{v}^T\boldsymbol{\gamma}_s + \theta_s \geq q_s h(\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}, s) \qquad \forall (\boldsymbol{c}, \boldsymbol{w}, \boldsymbol{d}, \boldsymbol{v}) \in \Omega'_s, s \in [S], \tag{A.2b}$$

$$\boldsymbol{\beta}_s, \boldsymbol{\gamma}_s \geq 0 \qquad \forall s \in [S]. \tag{A.2c}$$

Due to the feasibility and the linearity of the lifted ambiguity set (see, for instance, Mohajerin Esfahani and Kuhn (2018) and Bertsimas et al. (2019)), the strong duality condition holds. Moreover, for a fixed value of $(\boldsymbol{\alpha}_s, \boldsymbol{\beta}_s, \boldsymbol{\lambda}_s, \boldsymbol{\gamma}_s, \theta_s)$, constraints (A.2b) are equivalent to

$$\theta_s \geq \max_{(\boldsymbol{c},\boldsymbol{w},\boldsymbol{d},\boldsymbol{v})\in\Omega'_s} \left[q_s h(\boldsymbol{y}, \boldsymbol{c}, \boldsymbol{d}, s) - \left(\boldsymbol{c}^T\boldsymbol{\alpha}_s + \boldsymbol{w}^T\boldsymbol{\beta}_s + \boldsymbol{d}^T\boldsymbol{\lambda}_s + \boldsymbol{v}^T\boldsymbol{\gamma}_s\right)\right] \quad \forall s \in [S].$$

Since we are minimizing θ_s in the objective function (A.2a), the dual formulation can be further written as in the form of model (9).

A.3 Proof of Proposition 4

Under scenario $s \in [S]$ and a given value of $(\boldsymbol{\alpha}_s, \boldsymbol{\beta}_s, \boldsymbol{\lambda}_s, \boldsymbol{\gamma}_s)$, we solve the following optimization problem to reformulate the term:

$$\max \sum_{j\in[J]} (-c_j\alpha_{js} - w_j\beta_{js}) + \sum_{i\in[I]} (-d_i\lambda_{is} - v_i\gamma_{is})$$

$$\begin{aligned}
\text{s.t. } & c_j \leq \bar{c}_{js} && \forall j \in [J] && \cdots \text{dual variable}: A_{js} \in \mathbb{R},\\
& -c_j \leq -\underline{c}_{js} && \forall j \in [J] && \cdots \text{dual variable}: B_{js} \in \mathbb{R},\\
& c_j - w_j \leq \mu_{js} && \forall j \in [J] && \cdots \text{dual variable}: D_{js} \in \mathbb{R},\\
& -c_j - w_j \leq -\mu_{js} && \forall j \in [J] && \cdots \text{dual variable}: E_{js} \in \mathbb{R},\\
& d_i \leq \bar{d}_{is} && \forall i \in [I] && \cdots \text{dual variable}: F_{is} \in \mathbb{R},\\
& -d_i \leq -\underline{d}_{is} && \forall i \in [I] && \cdots \text{dual variable}: G_{is} \in \mathbb{R},\\
& d_i - v_i \leq \rho_{is} && \forall i \in [I] && \cdots \text{dual variable}: H_{is} \in \mathbb{R},\\
& -d_i - v_i \leq -\rho_{is} && \forall i \in [I] && \cdots \text{dual variable}: K_{is} \in \mathbb{R}.
\end{aligned}$$

Since the above model is feasible and the lifted support set is bounded, the strong duality condition holds. We can derive its dual problem as given in formulation (12).

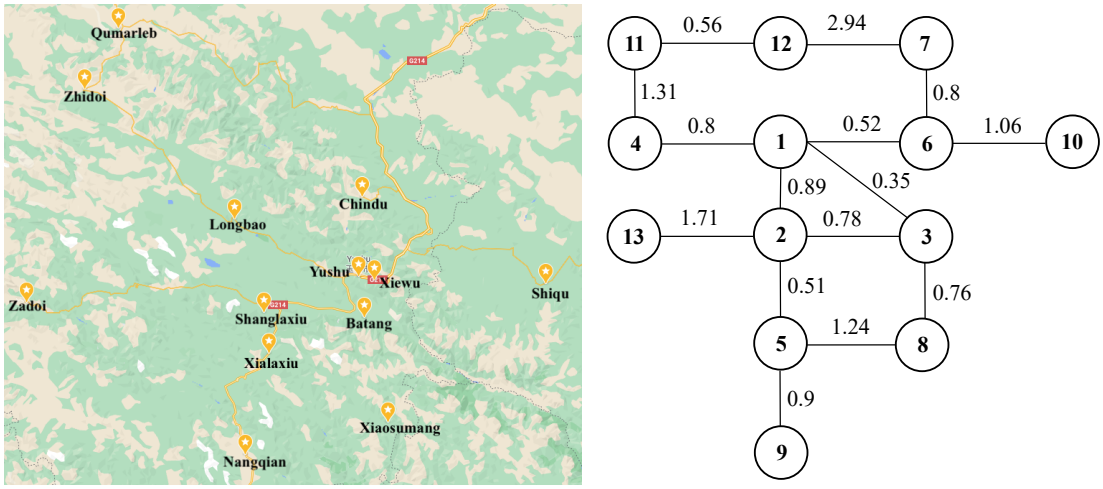
Appendix B Reformulation of the Risk-averse Stochastic Model

The risk-averse two-stage stochastic programming model with the risk measure CVaR can be reformulated as

$$
\begin{aligned}
\min \ & \sum_{j\in[J]} f_j y_j + \eta + \frac{1}{1-\alpha}\sum_{l=1}^{L}\frac{1}{L}v_l \\
\text{s.t. } & v_l \geq \sum_{i\in[I]}\sum_{j\in[J]} t_{ij}x_{ijl} + \sum_{i\in[I]} p_i u_{il} - \eta && \forall l\in[L], \\
& \sum_{j\in[J]} x_{ijl} + u_{il} \geq \hat{d}_{il} && \forall i\in[I], l\in[L], \\
& \sum_{i\in[I]} x_{ijl} \leq \hat{c}_{jl} y_j && \forall j\in[J], l\in[L], \\
& y_j \in \{0,1\} && \forall j\in[J], \\
& v_l \geq 0 && \forall l\in[L], \\
& \eta \in \mathbb{R}, \\
& x_{ijl} \geq 0 && \forall i\in[I], j\in[J], l\in[L], \\
& u_{il} \geq 0 && \forall i\in[I], l\in[L].
\end{aligned}
$$

Appendix C Data Related to the Earthquake Case Study

Figure C1 Network of affected areas in the earthquake case study



(a) Affected areas (b) Network of affected areas

For the earthquake case study, the 13 affected areas are shown in Figure C1(a). The network in the affected areas is given in Figure C1(b), where the numbers next to the links are the unit transportation costs. The unit transportation cost between any two nodes is given in Table C1, which is set to the shortest path distance between two nodes. Other parameters are given in Table

C2, where the values of f_j, p_i, h_j, q_j^+, q_i^-, μ_j^{r1}, μ_{j1}^r, $i \in [I], j \in [J]$ are directly adopted from Ni et al. (2018), and the values of μ_j^{r2} and μ_{j2}^r are set as $\mu_j^{r2} = 1.3\mu_j^{r1}, \mu_{j2}^r = 1.3\mu_{j1}^r$, $j \in [J]$ to accommodate the scenario-wise DRO framework. Note that p_i and q_i^- both denote the unit penalty cost of unmet demand, so their values are the same. However, as they are used in different models, we give their values in two different rows for notational clarity. In addition, μ_j^{r1} and μ_j^{r2} denote the means of in-sample random residual inventory. Whereas μ_{j1}^r and μ_{j2}^r represent the means of out-of-sample residual inventory used in the location and inventory pre-positioning problem.

Table C1 The unit transportation costs between nodes

Node	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.00	0.89	0.35	0.80	1.40	0.52	1.32	1.11	2.30	1.58	2.11	2.67	2.60
2	0.89	0.00	0.78	1.69	0.51	1.41	2.21	1.54	1.41	2.47	3.00	3.56	1.71
3	0.35	0.78	0.00	1.15	1.29	0.87	1.67	0.76	2.19	1.93	2.46	3.02	2.49
4	0.80	1.69	1.15	0.00	2.20	1.32	2.12	1.91	3.10	2.38	1.31	1.87	3.40
5	1.40	0.51	1.29	2.20	0.00	1.92	2.72	1.24	0.90	2.98	3.51	4.07	2.22
6	0.52	1.41	0.87	1.32	1.92	0.00	0.80	1.63	2.82	1.06	2.63	3.19	3.12
7	1.32	2.21	1.67	2.12	2.72	0.80	0.00	2.43	3.62	1.86	3.43	2.94	3.92
8	1.11	1.54	0.76	1.91	1.24	1.63	2.43	0.00	2.14	2.69	3.22	3.78	3.25
9	2.30	1.41	2.19	3.10	0.90	2.82	3.62	2.14	0.00	3.88	4.41	4.97	3.12
10	1.58	2.47	1.93	2.38	2.98	1.06	1.86	2.69	3.88	0.00	3.69	4.25	4.18
11	2.11	3.00	2.46	1.31	3.51	2.63	3.43	3.22	4.41	3.69	0.00	0.56	4.71
12	2.67	3.56	3.02	1.87	4.07	3.19	2.94	3.78	4.97	4.25	0.56	0.00	5.27
13	2.60	1.71	2.49	3.40	2.22	3.12	3.92	3.25	3.12	4.18	4.71	5.27	0.00

Table C2 Parameters of each node

Node	1	2	3	4	5	6	7	8	9	10	11	12	13
f_j	203	193	130	117	292	174	130	157	134	161	234	220	170
p_i	11.48	14.32	12.14	16.19	12.01	14.90	9.42	11.91	10.68	11.24	13.10	11.09	10.18
μ_j^{r1}	0.46	0.49	0.51	0.45	0.46	0.52	0.55	0.51	0.44	0.51	0.57	0.45	0.49
μ_j^{r2}	0.60	0.64	0.66	0.59	0.60	0.68	0.72	0.66	0.57	0.66	0.74	0.59	0.64
μ_{j1}^r	0.05	0.05	0.20	0.18	0.18	0.72	0.76	0.70	0.60	0.70	0.78	0.62	0.68
μ_{j2}^r	0.07	0.07	0.26	0.23	0.23	0.94	0.99	0.91	0.78	0.91	1.01	0.81	0.88
h_j	3.40	2.33	2.00	2.69	2.63	3.44	3.43	3.53	2.33	2.50	3.37	2.84	3.76
q_j^+	2.81	2.58	2.86	2.42	3.28	3.05	2.77	2.68	2.52	3.14	2.93	2.85	2.87
q_i^-	11.48	14.32	12.14	16.19	12.01	14.90	9.42	11.91	10.68	11.24	13.10	11.09	10.18

■ Note that the third and fourth (fifth and sixth) rows are means for generating in-sample (out-of-sample) data.

Appendix D Detailed Results of Numerical Experiments

Table D1 Detailed results of the SDR and SAA models for the simulation tests

			SDR						SAA					
Δ_2	No.	(J, I)	Time	Cost_1	Cost_2	Cost_t	$\#U$	$\#O$	Time	Cost_1	Cost_2	Cost_t	$\#U$	$\#O$
<0	1	(5, 10)	0.01	9889.70	10475.17	20364.87	1.64	2.99	0.10	8054.38	11646.52	19700.90	3.84	2.50
	2	(10, 10)	0.01	10578.13	8121.14	18699.27	0.44	3.58	0.15	8737.06	9195.17	17932.22	1.39	2.99
	3	(10, 20)	0.02	20133.52	15010.19	35143.71	0.90	6.34	0.26	16408.16	17176.76	33584.92	2.44	5.29
	4	(20, 20)	0.05	21386.78	11722.77	33109.55	0.22	7.35	2.03	16897.66	14030.75	30928.41	1.28	5.94
	5	(15, 30)	0.04	30755.88	18153.02	48908.90	0.66	9.74	1.09	24844.52	21467.34	46311.86	2.14	8.14
	6	(30, 30)	0.11	31727.10	14390.89	46117.99	0.15	11.30	8.22	24863.38	17444.62	42307.99	1.07	9.10
	7	(20, 40)	0.07	41341.84	21205.27	62547.11	0.58	13.07	2.52	33287.05	25865.14	59152.19	2.10	10.90
	8	(40, 40)	0.23	42939.48	16520.31	59459.79	0.14	15.32	16.59	33739.34	20195.68	53935.03	0.96	12.38
	9	(25, 50)	0.11	51117.29	24046.40	75163.69	0.61	16.41	4.63	41803.79	29368.99	71172.78	1.92	13.93
	10	(50, 50)	0.39	54129.70	18184.77	72314.48	0.08	19.52	28.46	42206.66	22535.35	64742.02	0.81	15.70
	11	(50, 100)	0.60	103182.52	35678.14	138860.66	0.55	33.00	29.32	86874.24	44299.58	131173.82	1.53	28.80
	12	(100, 100)	2.08	109065.76	25804.34	134870.10	0.03	40.22	147.89	83254.55	33675.22	116929.77	0.70	31.83
>0	1	(5, 10)	0.01	9515.26	18652.43	28167.70	10.22	2.92	0.10	8052.82	20490.27	28543.09	14.14	2.50
	2	(10, 10)	0.01	10849.51	14685.33	25534.84	5.52	3.63	0.15	8809.22	17466.76	26275.98	9.33	2.98
	3	(10, 20)	0.02	20230.14	29842.56	50072.71	7.96	6.33	0.25	16604.62	34979.84	51584.45	12.06	5.31
	4	(20, 20)	0.05	21410.36	22840.97	44251.33	4.50	7.50	2.07	16637.12	30291.67	46928.79	9.19	5.98
	5	(15, 30)	0.04	30812.01	39907.02	70719.03	7.51	9.67	1.10	25015.45	48558.50	73573.95	11.58	8.10
	6	(30, 30)	0.11	31567.58	30411.40	61978.97	4.40	11.26	7.90	24876.38	40941.37	65817.75	8.77	9.11
	7	(20, 40)	0.07	40724.48	49422.98	90147.46	7.15	13.07	2.38	33217.31	60923.53	94140.84	11.05	11.04
	8	(40, 40)	0.21	43222.70	36028.50	79251.19	3.90	15.43	15.87	33931.18	50325.84	84257.02	8.09	12.52
	9	(25, 50)	0.12	51218.66	58583.67	109802.33	7.12	16.33	4.26	42164.02	72813.70	114977.73	10.85	13.90
	10	(50, 50)	0.38	53868.14	41958.09	95826.22	3.87	19.33	25.53	42398.02	59953.90	102351.91	8.07	15.64
	11	(50, 100)	0.57	102438.95	100606.99	203045.94	6.77	33.22	31.06	87180.21	125457.30	212637.51	9.66	29.22
	12	(100, 100)	2.06	110219.98	66146.81	176366.79	3.19	40.21	149.68	84082.70	107280.84	191363.54	7.72	31.78

Table D2 Detailed results of the MDR model for the simulation tests

		$\Delta_2 < 0$					$\Delta_2 > 0$				
No.	(J, I)	Cost_1	Cost_2	Cost_t	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	$\#U$	$\#O$
1	(5, 10)	13359.40	9537.92	22897.32	0.30	3.91	12682.27	16237.77	28920.04	4.60	3.78
2	(10, 10)	14018.01	7154.77	21172.78	0.01	4.53	14213.02	11927.98	26140.99	1.86	4.58
3	(10, 20)	28450.19	13431.92	41882.11	0.04	8.45	28560.66	23360.47	51921.13	2.54	8.44
4	(20, 20)	27491.90	10534.66	38026.56	0.00	8.99	27086.53	17967.87	45054.40	1.62	9.00
5	(15, 30)	43392.01	16101.62	59493.62	0.01	12.84	44081.34	29516.65	73597.99	2.19	12.93
6	(30, 30)	38575.26	13047.86	51623.13	0.00	13.00	38254.82	24355.81	62610.63	2.06	13.00
7	(20, 40)	57275.57	18602.80	75878.37	0.00	16.99	56810.79	36001.72	92812.51	2.37	16.96
8	(40, 40)	51049.90	15039.74	66089.64	0.00	17.22	51077.27	29077.96	80155.23	2.11	17.30
9	(25, 50)	71710.52	20749.92	92460.44	0.00	21.42	72308.43	40539.67	112848.10	2.20	21.44
10	(50, 50)	64126.54	16537.07	80663.61	0.00	21.98	64541.43	32566.51	97107.94	1.86	22.00
11	(50, 100)	145869.77	29239.16	175108.93	0.00	42.99	143674.15	63135.72	206809.87	2.16	42.99
12	(100, 100)	121841.06	23954.55	145795.61	0.00	43.00	123380.02	54896.30	178276.32	2.17	43.00

Table D3 Sensitivity analysis on facilities' fixed costs based on the case study

	SAA						SDR						MDR					
ϵ	Cost_1	Cost_2	Cost_t	95th	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	95th	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	95th	$\#U$	$\#O$
0.5	627.50	262.85	890.35	994.89	0.28	8.00	776.23	141.95	918.19	963.65	0.00	9.73	809.57	119.84	929.41	988.16	0.00	10.07
0.6	637.20	382.43	1019.63	1206.26	0.03	7.00	867.35	184.16	1051.51	1113.55	0.00	9.12	922.18	149.87	1072.06	1137.26	0.00	9.62
0.7	743.40	382.43	1125.83	1312.46	0.03	7.00	985.83	199.34	1185.17	1252.42	0.00	8.90	1023.80	178.07	1201.88	1282.98	0.00	9.21
0.8	759.12	505.85	1264.97	1687.42	0.28	6.03	1094.71	218.81	1313.51	1391.68	0.00	8.66	1146.92	190.19	1337.11	1425.94	0.00	9.03
0.9	849.92	508.92	1358.84	1781.92	0.28	5.99	1177.66	249.02	1426.68	1533.42	0.00	8.32	1271.67	199.33	1471.01	1542.88	0.00	8.92
1.0	836.11	727.27	1563.38	2659.60	1.13	5.07	1268.59	270.72	1539.32	1658.67	0.00	8.11	1389.22	212.37	1601.58	1681.38	0.00	8.77

Table D4 Sensitivity analysis on the size of the support set based on the case study

	SDR						MDR					
ϱ (%)	Cost_1	Cost_2	Cost_t	95th	$\#U$	$\#O$	Cost_1	Cost_2	Cost_t	95th	$\#U$	$\#O$
0	1268.59	270.72	1539.32	1658.67	0.001	8.11	1389.22	212.37	1601.58	1681.38	0.000	8.77
2	1256.00	277.35	1533.35	1636.40	0.001	8.03	1371.51	220.80	1592.30	1677.73	0.000	8.66
4	1244.68	281.97	1526.65	1635.24	0.001	7.97	1350.31	231.61	1581.92	1677.73	0.000	8.54
6	1235.88	285.41	1521.29	1630.45	0.001	7.92	1331.70	240.34	1572.04	1673.15	0.000	8.43
8	1220.28	294.61	1514.89	1630.92	0.002	7.83	1312.44	249.40	1561.85	1672.09	0.000	8.31