

Distributionally Robust Facility Location with Uncertain Facility Capacity and Customer Demand

Chun Cheng

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

Qinxiao Yu

Economics and Management College, Civil Aviation University of China, Tianjin 300300, China
Corresponding Author. qxyu@cauc.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

This study explores a capacitated facility location problem where facility capacity and customer demand are subject to uncertainties simultaneously. This problem decides the subset of facilities to open at the system design phase to serve customers during the operational phase. The objective is to minimize the total cost, including the first-stage location cost and the second-stage recourse cost, and guarantee the system's reliability, i.e., meeting demand as much as possible when uncertainties arise. A distributionally robust optimization (DRO) framework is utilized to model the problem. A scenario-wise ambiguity set with partial distributional information of random variables is constructed, which can capture uncertainties caused by different random events or different magnitudes of the same event type and explicitly represent the correlation between facilities' uncertain capacity and customers' uncertain demand. We apply an adaptation policy to the DRO model and reformulate it to a mixed-integer linear programming model, which is solvable by off-the-shelf solvers. Numerical results show that the scenario-wise DRO framework can provide a better trade-off between cost and service level than the stochastic programming model and the DRO model with a marginal moment-based ambiguity set, demonstrating that the proposed scenario-wise 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

History: Accepted to *Omega* on August 29, 2023.

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 sce-

nario, 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, 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, it is important for decision makers to incorporate *event-correlated* uncertainties when robustifying the supply chain system. Under event-correlated uncertainties, the impacts of a random event or the uncertainties caused by the random event are directly related to the event’s category and magnitude.

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, 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 levels. 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 nominal and disruption situations.

To tackle problems under uncertainties, three frameworks can be considered: 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 or even impossible to estimate the occurrence probability of some random events, especially natural disasters. Thus, solutions produced by stochastic models with the assumption of full knowledge about the probability distribution may lead to disappointing results in out-of-sample tests (Smith and Winkler

2006). On the other extreme, the RO approach completely ignores the probability distribution. Instead, it assumes that the uncertain parameters belong to an uncertainty set and then optimizes the performance of the *worst-case scenario* within the uncertainty set. As only the worst-case scenario is considered, the solution produced by RO will be suboptimal for other more-likely scenarios. More specifically, RO puts all its weight on the extreme scenario, and its performance would be poor or even terrible once different scenarios occur. The DRO approach, on the other hand, provides an alternative framework to combine SP and RO while attempting to address their shortcomings. With regard to knowing the probability distribution of random variables, DRO takes a middle-ground approach as opposed to the black-or-white view of the SP and RO approaches. Specifically, DRO assumes that the probability distribution of uncertain parameters belongs to a family of distributions (i.e., an ambiguity set) that shares common distributional information (e.g., mean, variance, and mean absolute deviation), which is often available and reliable (Popescu 2007). Optimization is subsequently performed to hedge against the *worst-case distribution* within the ambiguity set. Compared to the SP approach, the DRO approach does not require perfect distribution information of uncertain parameters, which is more compatible with the uncertainties considered in this study. Moreover, compared with RO, DRO can take full advantage of available data to extract statistical information on random variables to characterize randomness. As DRO considers the worst-case probability distribution rather than the single worst case, its solution is less conservative than RO's (Shang and You 2018, Shehadeh 2020, Ash et al. 2022, Wang et al. 2023); thus, it has the potential to save costs for multiple scenarios that may arise throughout a facility's life cycle.

This paper studies a capacitated facility location problem (CFLP) with both provider-side and receiver-side uncertainties, i.e., the uncertain capacity at facilities and the uncertain demand at customers. To solve the problem, the following research questions should be addressed: (1) the characterization of uncertainties to reflect the impacts of different random events on facilities' capacity and customers' demand; (2) the reformulation of the DRO model into a tractable mixed-integer linear programming (MILP) model; and (3) the quantification of trade-offs between cost and service level under different modeling frameworks. To summarize, this study makes the following contributions to the literature:

- A DRO modeling paradigm is utilized 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. Unlike RO models, it optimizes the performance under the worst-case probability distribution rather than only the worst-case scenario, reducing the conservatism of robust solutions.

- A scenario-wise ambiguity set is constructed to characterize the randomness of parameters, where the statistical information of random variables in each scenario can be estimated from historical observations, which can potentially be used in conjunction with human inputs when historical data is insufficient. The proposed ambiguity set can capture the impacts of different types of disruption events or different magnitudes of the same event type. More importantly, it can 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.
- A scenario-wise adaptation policy is applied to the second-stage recourse problem of the DRO model; namely, different recourse actions are used for different scenarios rather than a single recourse for all the cases, further alleviating the conservatism of robust solutions. We reformulate the resulting adaptive DRO model to an MILP model, which is solvable by off-the-shelf solvers.
- The proposed modeling framework is extended 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.
- Simulation tests and a case study are conducted to validate the scenario-wise DRO framework. Results show that the DRO model takes much less computing time in comparison with the SP 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 of 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, Stienen et al. 2021), 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. 2023, Kahr 2022). Dönmez et al. (2021) provide a comprehensive review of humanitarian facility location under uncertainty, where existing works are classified based on the type of facilities involved, the decision variables, the optimization objective, the modeling framework, and the solution method.

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 in 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 approach, where budgeted uncertainty sets are utilized to capture randomness. Wang et al. (2020) apply the adaptive DRO approach 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 approach to formulate the problem. Cheng et al. (2021) study a CFLP with facility disruptions and uncertain demand using a two-stage RO approach. 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.

The papers reviewed above show that some authors have begun considering the uncertain facility capacity and customer demand simultaneously. However, they often assume that these two types of uncertainty are independent, i.e., the correlation between uncertainties is neglected in their works, 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 this work differs from that of [Cheng et al. \(2021\)](#) in three aspects: First, the proposed ambiguity set can capture the dependency between facility disruption and uncertain demand, whereas [Cheng et al. \(2021\)](#) assume that the uncertain capacity of facilities and the uncertain demand of customers are independent. Second, facilities are assumed to lose their capacities completely in [Cheng et al. \(2021\)](#) once disruptions happen, whereas this work assumes that facilities can be completely or partially disrupted. This 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, the proposed model generalizes the approach considered in [Saif and Delage \(2021\)](#) by considering simultaneous supply and demand uncertainties by constructing a scenario-wise ambiguity set.

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 approach 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](#)), contract design problem in agricultural supply chains ([Zhong et al. 2023](#)), and facility location ([Shehadeh and Sanci 2021](#)), among others. In recent years, the statistical distance-based ambiguity set has gained 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 ϕ -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. 2021b), 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), Shehadeh and Sanci (2021), Perakis et al. (2023), 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), this study considers both provider-side and receiver-side uncertainty, and more importantly, the proposed 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. (2023) 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.

3. Problem Definition and Formulation

This section introduces the problem, the two-stage SP model, and the two-stage DRO model. A toy example is also provided to show the value of the scenario-wise ambiguity set.

3.1. Problem Definition

We first define the following notation that will be used throughout the paper. Symbol $[S] \triangleq \{1, \dots, S\}$ denotes the set of positive running indices up to S . Boldface lowercase and uppercase characters represent vectors and matrices with appropriate dimensions, respectively. \mathbf{a}^T is the transpose of \mathbf{a} . $\mathbf{a} \cdot \mathbf{b}$ denotes the dot product of two vectors with the same dimension. Symbol $\mathcal{P}(\mathbb{R}^I)$ denotes 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 $\tilde{\mathbf{d}} \in \mathbb{P}, \mathbb{P} \in \mathcal{P}(\Omega), \Omega \subseteq \mathbb{R}^I$ to define $\tilde{\mathbf{d}}$ as an I -dimensional random vector with support Ω and distribution \mathbb{P} .

In the CFLP, $[J]$ 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 equals to 1 if facility $j \in [J]$ is open, and 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 when 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 $\tilde{\mathbf{c}} = (\tilde{c}_1, \dots, \tilde{c}_J)^T$ and customers' uncertain demand as a random vector $\tilde{\mathbf{d}} = (\tilde{d}_1, \dots, \tilde{d}_I)^T$. The event that causes these uncertainties is represented by a random variable $\tilde{e} \in \mathbb{R}$. The event could be an earthquake, flood, or strike, i.e., \tilde{e} can represent any type of disruption event that decision-makers are concerned about. If the considered supply chain system is mainly disrupted by one type of event, \tilde{e} can denote the magnitude or the level of influence of the event, like a minor, major, or great earthquake. The joint capacity, demand, and related event is then denoted as $(\tilde{\mathbf{c}}, \tilde{\mathbf{d}}, \tilde{e}) \in \mathbb{R}^J \times \mathbb{R}^I \times \mathbb{R}$. The notation suggests that the correlation between uncertain capacity and demand is captured via the event causing these two types of uncertainties.

3.2. Two-stage Stochastic Programming Model

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

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

$$\text{s.t. } y_j \in \{0, 1\} \quad \forall j \in [J], \quad (1b)$$

where $h(\mathbf{y}, \tilde{\mathbf{c}}, \tilde{\mathbf{d}})$ is the second-stage recourse cost. Under a given location decision \mathbf{y} and a realization (\mathbf{c}, \mathbf{d}) of uncertain parameters $(\tilde{\mathbf{c}}, \tilde{\mathbf{d}})$, $h(\mathbf{y}, \tilde{\mathbf{c}}, \tilde{\mathbf{d}})$ is expressed as follows:

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

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

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

$$x_{ij} \geq 0 \quad \forall i \in [I], j \in [J], \quad (2d)$$

$$u_i \geq 0 \quad \forall i \in [I]. \quad (2e)$$

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 comprises two parts: transportation costs for serving demand and penalty costs for unmet demand. Constraints (2b) impose that the sum of the 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) assumes that \mathbb{P} is perfectly known. However, in practice, it is difficult or even impossible to obtain the true probability distribution. As an alternative, some studies (e.g., Noyan et al. (2016) and Xiang and Liu (2021)) use scenarios to represent uncertainty and associate each scenario with an occurrence probability. For our problem, suppose there are L samples of historical observations denoted as $\mathcal{L} = \{(\hat{\mathbf{c}}_1, \hat{\mathbf{d}}_1, \hat{e}_1), (\hat{\mathbf{c}}_2, \hat{\mathbf{d}}_2, \hat{e}_2), \dots, (\hat{\mathbf{c}}_L, \hat{\mathbf{d}}_L, \hat{e}_L)\}$. Under the scenario-based SP framework, each sample can be recognized as a scenario with an occurrence probability $1/L$, i.e., the empirical distribution is used to approximate the true distribution. Thus, the two-stage SP model with a scenario-based representation of uncertainty (abbreviated as SSP) is used to approximate model (1), which is as follows:

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

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

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

$$y_j \in \{0, 1\} \quad \forall j \in [J], \quad (3d)$$

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

$$u_{il} \geq 0 \quad \forall i \in [I], l \in [L], \quad (3f)$$

where \hat{d}_{il} and \hat{c}_{jl} are the i -th and j -th components of $\hat{\mathbf{d}}_l$ and $\hat{\mathbf{c}}_l$, respectively. Correspondingly, x_{ijl} and u_{il} are the recourse variables associated with scenario $l \in [L]$. Note that the information of events $\hat{e}_l, l \in [L]$, i.e., the type and the magnitude of events, is not utilized in the SSP model. Namely, the SP approach treats all the supply-demand samples as a whole and does not consider the events that have caused these uncertainties.

3.3. Two-stage Distributionally Robust Model

Formulation (3) is an MILP model, which can be solved directly by commercial solvers. Nevertheless, the empirical distribution of uncertain parameters is often unavailable, especially when dealing with 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, the DRO framework, together with a scenario-wise ambiguity set, is adopted.

3.3.1. Scenario-wise Ambiguity Set. The DRO framework assumes that the true distribution of $(\tilde{\mathbf{c}}, \tilde{\mathbf{d}}, \tilde{e}) : \mathbb{P} \in \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I \times \mathbb{R})$ 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 based on historical samples $\mathcal{L} = \{(\hat{\mathbf{c}}_1, \hat{\mathbf{d}}_1, \hat{e}_1), (\hat{\mathbf{c}}_2, \hat{\mathbf{d}}_2, \hat{e}_2), \dots, (\hat{\mathbf{c}}_L, \hat{\mathbf{d}}_L, \hat{e}_L)\}$. Let $\mathcal{E} = \{\hat{e}_1, \dots, \hat{e}_L\}$, and \mathcal{E} is subsequently partitioned 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 $\cup_{s \in [S]} \mathcal{E}_s = \mathcal{E}$. That is, (i) each scenario \mathcal{E}_s contains one or multiple realizations of \tilde{e} ; (ii) any two scenarios are exclusive; and (iii) the union of \mathcal{E}_s constitutes set \mathcal{E} . We will discuss how to do the partition later in this subsection.

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, $\mathcal{U}_s = \{l \in [L] \mid \hat{e}_l \in \mathcal{E}_s\}$. A random variable \tilde{s} taking discrete values in $[S]$ is introduced to denote the scenario $\tilde{s} = s$ associated with \mathcal{E}_s . Accordingly, using historical data, the scenario-wise ambiguity set $\mathcal{F} \subseteq \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I \times [S])$ associated with the random variable $(\tilde{\mathbf{c}}, \tilde{\mathbf{d}}, \tilde{s})$ is constructed as

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

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

$$\Omega_s = \{(\mathbf{c}, \mathbf{d}) \in \mathbb{R}^J \times \mathbb{R}^I \mid \underline{\mathbf{c}}_s \leq \mathbf{c} \leq \bar{\mathbf{c}}_s, \underline{\mathbf{d}}_s \leq \mathbf{d} \leq \bar{\mathbf{d}}_s\}.$$

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

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

the mean absolute deviations are

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

the probability is

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

and the parameters of the support set are

$$\begin{aligned} [\underline{\mathbf{c}}_s]_j &= \min_{l \in \mathcal{U}_s} \hat{c}_{jl}, & [\bar{\mathbf{c}}_s]_j &= \max_{l \in \mathcal{U}_s} \hat{c}_{jl}, & \forall j \in [J], \\ [\underline{\mathbf{d}}_s]_i &= \min_{l \in \mathcal{U}_s} \hat{d}_{il}, & [\bar{\mathbf{d}}_s]_i &= \max_{l \in \mathcal{U}_s} \hat{d}_{il}, & \forall i \in [I]. \end{aligned}$$

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

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

where the support set is defined as $\Omega = \{(\mathbf{c}, \mathbf{d}) \in \mathbb{R}^J \times \mathbb{R}^I \mid \underline{\mathbf{c}} \leq \mathbf{c} \leq \bar{\mathbf{c}}, \underline{\mathbf{d}} \leq \mathbf{d} \leq \bar{\mathbf{d}}\}$. In this case, only a single set of parameters will be derived from the entire data. The relationship between the optimal solutions under \mathcal{F} and $\bar{\mathcal{F}}$ is established 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

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

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 $\hat{e}_l^I, l \in \{m+1, \dots, L\}$ 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 $1, \dots, m$ 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 to divide samples into different scenarios. For example, the well-known and efficient K -means clustering algorithm can be implemented with a computational complexity of $\mathcal{O}(KnT)$, where K is the number of clusters, n is the number of samples, and T is the number of iterations (Jain and Dubes 1988, Jain 2010).

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, the two-stage DRO model under a scenario-wise ambiguity set (abbreviated as SDR) is constructed as

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

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

3.4. Value of the Scenario-wise Ambiguity Set

This section first presents a toy 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.

Consider a simple supply chain system with one supplier and one retailer. We assume the underlying true distribution is as follows: \tilde{e} follows a Bernoulli distribution with $\mathbb{P}(\tilde{e} = 1) = 1/2$. Conditioning on \tilde{e} , we have $\mathbb{P}(\tilde{\mathbf{c}} = \hat{\mathbf{c}}_1, \tilde{\mathbf{d}} = \hat{\mathbf{d}}_1 \mid \tilde{e} = 1) = 1$ and $\mathbb{P}(\tilde{\mathbf{c}} = \hat{\mathbf{c}}_0, \tilde{\mathbf{d}} = \hat{\mathbf{d}}_0 \mid \tilde{e} = 0) = 1$, where $\hat{\mathbf{c}}_1 \geq \hat{\mathbf{c}}_0, \hat{\mathbf{d}}_1 \leq \hat{\mathbf{d}}_0$. 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 $\mathcal{E}_0 = \{\tilde{e} = 0\}$. Correspondingly, the scenario-wise ambiguity set is constructed as

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

and the marginal moment-based ambiguity set as

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

It is observed 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. The benefits of using a scenario-wise ambiguity set are formally established 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 an 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), auxiliary random vectors $\tilde{\boldsymbol{w}} \in \mathbb{R}^J$ and $\tilde{\boldsymbol{v}} \in \mathbb{R}^I$ are first introduced to lift the ambiguity set (4) as

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

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

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

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 supreme problem in (7). Specifically, under a given location decision \mathbf{y} , \mathbb{P} is the decision variable of problem $\sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}} [h(\mathbf{y}, \tilde{\mathbf{c}}, \tilde{\mathbf{d}}, \tilde{\mathbf{s}})]$, i.e., we are choosing a distribution that maximizes the expected value of $h(\mathbf{y}, \tilde{\mathbf{c}}, \tilde{\mathbf{d}}, \tilde{\mathbf{s}})$.

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

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

$$\text{s.t. } \boldsymbol{\beta}_s, \boldsymbol{\gamma}_s \geq 0 \quad \forall s \in [S]. \quad (9b)$$

Proof. See Appendix A.2. □

Note that the model defined by (9a)–(9b) is not yet directly solvable by a commercial MILP solver in its present form, because an inner maximization problem exists. In the following, we reformulate the inner maximization problem to transform (9a) into a single minimization problem.

Scenario-wise adaptations. The scenario-wise adaptation policy introduced in Chen et al. (2020) is adopted for our problem, i.e., different recourse decisions are utilized for different scenarios. To do this, variables \mathbf{x} and \mathbf{u} are defined 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 second-stage problem $h(\mathbf{y}, \mathbf{c}, \mathbf{d}, s)$ under scenario-wise adaptations is written as

$$h(\mathbf{y}, \mathbf{c}, \mathbf{d}, s) = \min_{\mathbf{x}_s, \mathbf{u}_s} \sum_{i \in [I]} \sum_{j \in [J]} t_{ij} x_{ijs} + \sum_{i \in [I]} p_i u_{is} \quad (10a)$$

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

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

$$x_{ijs} \geq 0 \quad \forall i \in [I], j \in [J], \quad (10d)$$

$$u_{is} \geq 0 \quad \forall i \in [I]. \quad (10e)$$

Under a given first-stage decision \mathbf{y} and scenario $s \in [S]$, let O_{is} and Q_{js} be the dual variables associated with constraints (10b) and (10c), respectively. Then the dual problem of $h(\mathbf{y}, \mathbf{c}, \mathbf{d}, s)$ is

$$\max_{\mathbf{O}_s, \mathbf{Q}_s} \sum_{i \in [I]} d_{is} O_{is} + \sum_{j \in [J]} c_{js} y_j Q_{js} \quad (11a)$$

$$\text{s.t. } O_{is} + Q_{js} \leq t_{ij} \quad \forall i \in [I], j \in [J], \quad (11b)$$

$$O_{is} \leq p_i \quad \forall i \in [I], \quad (11c)$$

$$O_{is} \geq 0 \quad \forall i \in [I], \quad (11d)$$

$$Q_{js} \leq 0 \quad \forall j \in [J]. \quad (11e)$$

Now the inner maximization problem in (9a) can be equivalently reformulated as

$$\max_{(\mathbf{c}, \mathbf{w}, \mathbf{d}, \mathbf{v}) \in \Omega'_s} \left\{ q_s \left(\max_{\mathbf{O}_s, \mathbf{Q}_s} \sum_{i \in [I]} d_{is} O_{is} + \sum_{j \in [J]} c_{js} y_j Q_{js} \right) + (-\mathbf{c}^T \boldsymbol{\alpha}_s - \mathbf{w}^T \boldsymbol{\beta}_s - \mathbf{d}^T \boldsymbol{\lambda}_s - \mathbf{v}^T \boldsymbol{\gamma}_s) \right\} \quad (12a)$$

$$\text{s.t. (11b)–(11e)}. \quad (12b)$$

Regarding the inner maximization problem in (12a), only its objective function involves variables \mathbf{c} and \mathbf{d} ; thus, its optimal value is attained at $d_{is} = \bar{d}_{is}$ (due to $O_{is} \geq 0$) and $c_{js} = \underline{c}_{js}$ (due to $Q_{js} \leq 0$). Consequently, equation (12a) is rewritten as

$$q_s \left(\max_{\mathbf{O}_s, \mathbf{Q}_s} \sum_{i \in [I]} \bar{d}_{is} O_{is} + \sum_{j \in [J]} \underline{c}_{js} y_j Q_{js} \right) + \max_{(\mathbf{c}, \mathbf{w}, \mathbf{d}, \mathbf{v}) \in \Omega'_s} (-\mathbf{c}^T \boldsymbol{\alpha}_s - \mathbf{w}^T \boldsymbol{\beta}_s - \mathbf{d}^T \boldsymbol{\lambda}_s - \mathbf{v}^T \boldsymbol{\gamma}_s) \quad (13)$$

Let x_{ijs} and p_{is} be the dual variables associated with constraints (11b) and (11c), respectively. By using the dual theory, we obtain Proposition 3.

PROPOSITION 3. For any $\mathbf{y} \in \{0, 1\}^J$, $\max_{\mathbf{O}_s, \mathbf{Q}_s} \sum_{i \in [I]} \bar{d}_{is} O_{is} + \sum_{j \in [J]} \underline{c}_{js} y_j Q_{js}$ is equivalent to

$$\min_{\mathbf{x}_s, \mathbf{u}_s} \sum_{i \in [I]} \sum_{j \in [J]} t_{ij} x_{ijs} + \sum_{i \in [I]} p_i u_{is} \quad (14a)$$

$$\text{s.t. } \sum_{j \in [J]} x_{ijs} + u_{is} \geq \bar{d}_{is} \quad \forall i \in [I], \quad (14b)$$

$$\sum_{i \in [I]} x_{ijs} \leq \underline{c}_{js} y_j \quad \forall j \in [J], \quad (14c)$$

$$x_{ijs} \geq 0 \quad \forall i \in [I], j \in [J], \quad (14d)$$

$$u_{is} \geq 0 \quad \forall i \in [I]. \quad (14e)$$

PROPOSITION 4. The term $\max_{(\mathbf{c}, \mathbf{w}, \mathbf{d}, \mathbf{v}) \in \Omega'_s} -(\mathbf{c}^T \boldsymbol{\alpha}_s + \mathbf{w}^T \boldsymbol{\beta}_s + \mathbf{d}^T \boldsymbol{\lambda}_s + \mathbf{v}^T \boldsymbol{\gamma}_s)$ in (13) is equivalent to the following minimization problem

$$\min \sum_{j \in [J]} \{ \bar{c}_{js} A_{js} - \underline{c}_{js} B_{js} + \mu_{js} (D_{js} - E_{js}) \} + \sum_{i \in [I]} \{ \bar{d}_{is} F_{is} - \underline{d}_{is} G_{is} + \rho_{is} (H_{is} - K_{is}) \} \quad (15a)$$

$$\text{s.t. } A_{js} - B_{js} + D_{js} - E_{js} = -\alpha_{js} \quad \forall j \in [J], \quad (15b)$$

$$D_{js} + E_{js} = \beta_{js} \quad \forall j \in [J], \quad (15c)$$

$$F_{is} - G_{is} + H_{is} - K_{is} = -\lambda_{is} \quad \forall i \in [I], \quad (15d)$$

$$H_{is} + K_{is} = \gamma_{is} \quad \forall i \in [I], \quad (15e)$$

$$A_{js}, B_{js}, D_{js}, E_{js} \geq 0 \quad \forall j \in [J], \quad (15f)$$

$$F_{is}, G_{is}, H_{is}, K_{is} \geq 0 \quad \forall i \in [I]. \quad (15g)$$

Proof. See Appendix A.3. \square

Based on Propositions 2–4, the SDR model under the scenario-wise adaptation policy is finally reformulated to an MILP model, i.e.,

$$\begin{aligned} \min \quad & \sum_{j \in [J]} f_j y_j + \sum_{s \in [S]} \left\{ \sum_{j \in [J]} (\delta_{js} \beta_{js} + \bar{c}_{js} A_{js} - \underline{c}_{js} B_{js} + \mu_{js} (\alpha_{js} + D_{js} - E_{js})) \right. \\ & \left. + \sum_{i \in [I]} (\zeta_{is} \gamma_{is} + \bar{d}_{is} F_{is} - \underline{d}_{is} G_{is} + \rho_{is} (\lambda_{is} + H_{is} - K_{is})) + \sum_{i \in [I]} \sum_{j \in [J]} q_s t_{ij} x_{ijs} + \sum_{i \in [I]} q_s p_i u_{is} \right\} \end{aligned} \quad (16a)$$

$$\text{s.t. (9b), (14b)–(14e) for each } s \in [S], \text{ and (15b)–(15g) for each } s \in [S]. \quad (16b)$$

In the case of a non-adaptive (or static) policy, there will be $x_{ijs} = x_{ij}, u_{is} = u_i, s \in [S]$, 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 (14b) and (14c) become

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

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.

Note that the 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, parameters $\rho_s, s \in [S]$ are set to the nominal demand at customers (denoted as \mathbf{d}^n) and parameters $\zeta_s, s \in [S]$ are set to 0 in the ambiguity set (4). The bounds of demand can be set as $\underline{\mathbf{d}}_s = \bar{\mathbf{d}}_s = \mathbf{d}^n, s \in [S]$. Similarly, for distributionally robust CFLPs with receiver-side uncertainty, parameters $\mu_s, s \in [S]$ are set to facilities' nominal capacities (denoted as \mathbf{c}^n), and parameters $\delta_s, s \in [S]$ are set to 0. The bounds of capacities are set to $\underline{\mathbf{c}}_s = \bar{\mathbf{c}}_s = \mathbf{c}^n, s \in [S]$. 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, let $S = 2$. As the authors only specify the mean of demand in their ambiguity set, we can further let $\zeta_s = 0, s \in [S]$.

The scenario-wise DRO framework can also be applied 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, the capacity parameter $\mu_j, j \in [J]$ can be set to a sufficiently large number. 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, Shehadeh and Tucker 2022), 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 approach for the LIPP with simultaneous provider-side and receiver-side uncertainties. Shehadeh and Tucker (2022) use a DRO approach, together with a marginal moment-based ambiguity set and a static policy, to solve the LIPP under multiple types of uncertainties. Note that the scenario-wise DRO framework, together with the adaptation policy proposed in this work, can be directly adopted to solve the LIPP with simultaneous provider-side and receiver-side uncertainties. The corresponding DRO model is provided as follows, and the empirical insights for this problem are presented 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, the decision maker decides where to open facilities and how much inventory to pre-position at each opened facility. Binary variable y_j equals 1 if a facility is opened at site $j \in [J]$, and 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 $j \in [J]$.

The following scenario-wise ambiguity set is used to characterize distribution ambiguity

$$\mathcal{F} = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^J \times \mathbb{R}^I \times [S]) \mid \begin{array}{l} (\tilde{\mathbf{r}}, \tilde{\mathbf{d}}, \tilde{s}) \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{r}} \mid \tilde{s} = s] = \boldsymbol{\mu}_s^r \quad \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}[|\tilde{\mathbf{r}} - \boldsymbol{\mu}_s^r| \mid \tilde{s} = s] \leq \boldsymbol{\delta}_s^r \quad \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{d}} \mid \tilde{s} = s] = \boldsymbol{\mu}_s^d \quad \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}[|\tilde{\mathbf{d}} - \boldsymbol{\mu}_s^d| \mid \tilde{s} = s] \leq \boldsymbol{\delta}_s^d \quad \forall s \in [S] \\ \mathbb{P}[(\mathbf{r}, \mathbf{d}) \in \Omega_s \mid \tilde{s} = s] = 1 \quad \forall s \in [S] \\ \mathbb{P}[\tilde{s} = s] = q_s \quad \forall s \in [S] \end{array} \right\}, \quad (17)$$

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

$$\Omega_s = \{(\mathbf{r}, \mathbf{d}) \in \mathbb{R}^J \times \mathbb{R}^I \mid \underline{\mathbf{r}}_s \leq \mathbf{r} \leq \bar{\mathbf{r}}_s, \underline{\mathbf{d}}_s \leq \mathbf{d} \leq \bar{\mathbf{d}}_s\}.$$

The corresponding two-stage distributionally robust model is formulated as

$$\min_{\mathbf{y}, \mathbf{a}} \left\{ \sum_{j \in [J]} (f_j y_j + h_j a_j) + \sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}} \left[\mathcal{Q}(\mathbf{y}, \mathbf{a}, \tilde{\mathbf{r}}, \tilde{\mathbf{d}}, \tilde{s}) \right] \right\}, \quad (18a)$$

$$\text{s.t. } \sum_{j \in [J]} a_j = R, \quad (18b)$$

$$a_j \leq M_j y_j \quad \forall j \in [J], \quad (18c)$$

$$y_j \in \{0, 1\} \quad \forall j \in [J], \quad (18d)$$

$$a_j \geq 0 \quad \forall j \in [J]. \quad (18e)$$

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

$$\mathcal{Q}(\mathbf{y}, \mathbf{a}, \mathbf{r}, \mathbf{d}, s) = \min_{\mathbf{x}, \mathbf{z}^-} \sum_{i \in [I]} \sum_{j \in [J]} t_{ij} x_{ij} + \sum_{j \in [J]} \left(a_j r_j - \sum_{i \in [I]} x_{ij} \right) q_j^+ + \sum_{i \in [I]} q_i^- z_i^- \quad (19a)$$

$$\text{s.t. } \sum_{i \in [I]} x_{ij} \leq a_j r_j \quad \forall j \in [J], \quad (19b)$$

$$\sum_{j \in [J]} x_{ij} + z_i^- \geq d_i \quad \forall i \in [I], \quad (19c)$$

$$x_{ij} \geq 0 \quad \forall i \in [I], j \in [J], \quad (19d)$$

$$z_i^- \geq 0 \quad \forall i \in [I]. \quad (19e)$$

The objective function (18a) 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. Constraint (18b) means that all the available commodities are pre-positioned at facilities before the disaster. Constraints (18c) denote that decision-makers can only pre-position inventories at open facilities and that facilities' capacities must be respected. Equation (19a) refers to the second-stage objective function, which minimizes the sum of the transportation cost, the penalty cost of unused inventory, and the penalty cost of unmet demand under a given first-stage decision (\mathbf{y}, \mathbf{a}) and an uncertainty realization $(\mathbf{r}, \mathbf{d}, s)$. Constraints (19b) suggest that the total commodities delivered from a facility to customers cannot surpass the usable quantity at that facility. Constraints (19c) indicate the sum of met and unmet demand must be equal to or larger than a customer's demand. Constraints (18d), (18e), (19d), and (19e) 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. The SSP, SDR, and MDR models are compared 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 the 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. A superscript *qd* is used to denote the *quantity difference* in a performance indicator, obtained by using the value of this indicator in the SDR model to subtract the corresponding value in another model. For example, symbol $Cost_t^{qd}$ refers to the quantity difference in the expected total cost. Thus, a negative value suggests that the SDR model performs better in that corresponding indicator.

5.1. Simulation Tests

This section first introduces the instance set and then presents the results of different models.

5.1.1. Instance Set. Number-, coordinate-, and cost-related parameters are generated 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, two data sets are considered: 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 set $S = 4$ 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 $b_s^2 = 30 + 10s$. In this way, facility capacity and customer demand are connected by scenarios. Specifically, as index s increases, the mean capacity decreases while the mean demand increases. This data construction method can simulate the case that as the level of disruption increases, facilities are more seriously affected, and customers need more relief items. 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 $\{0.10, 0.15, \dots, 0.30\}$ and $\{-0.30, -0.25, \dots, -0.10, 0.10, 0.15, \dots, 0.30\}$, respectively. For a fixed value of $(I, J, \Delta^1, \Delta^2)$, 5 instances are generated, and thus there are 3000 instances in total. The parameters of the two ambiguity sets are calculated based on the training samples. The solutions of SSP are derived and evaluated only using the supply and demand information contained in the training and testing samples, respectively.

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 SSP. The MILP reformulation of SDR can be efficiently solved—all the instances are solved optimally within 2.08 seconds (refer to Table D1). The average computing time of SSP 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 SSP is more than 145 seconds longer than that of SDR.

From Table 1, it is observed that SDR opens more facilities than SSP, 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.,

Table 1 Quantity differences in average results of SDR and SSP models

No.	(J, I)	$\Delta_2 < 0$						$\Delta_2 > 0$							
		$Time^{qd}$	$\#U^{qd}$	$\#O^{qd}$	$Cost_1^{qd}$	$Cost_2^{qd}$	$Cost_t^{qd}$	$95th^{qd}$	$Time^{qd}$	$\#U^{qd}$	$\#O^{qd}$	$Cost_1^{qd}$	$Cost_2^{qd}$	$Cost_t^{qd}$	$95th^{qd}$
1	(5, 10)	-0.10	-2.20	0.50	1835.33	-1171.36	663.97	1172.29	-0.09	-3.92	0.42	1462.45	-1837.84	-375.39	-571.97
2	(10, 10)	-0.14	-0.94	0.59	1841.07	-1074.02	767.05	174.33	-0.14	-3.80	0.65	2040.29	-2781.43	-741.14	-463.48
3	(10, 20)	-0.24	-1.54	1.05	3725.36	-2166.57	1558.79	771.44	-0.24	-4.10	1.02	3625.53	-5137.28	-1511.75	-2759.90
4	(20, 20)	-1.98	-1.07	1.41	4489.12	-2307.98	2181.14	1277.30	-2.02	-4.69	1.52	4773.24	-7450.70	-2677.46	-4240.82
5	(15, 30)	-1.05	-1.48	1.59	5911.36	-3314.32	2597.04	237.86	-1.06	-4.07	1.57	5796.56	-8651.48	-2854.92	-5845.55
6	(30, 30)	-8.11	-0.91	2.19	6863.73	-3053.73	3810.00	1000.84	-7.79	-4.38	2.15	6691.20	-10529.98	-3838.78	-5698.04
7	(20, 40)	-2.45	-1.52	2.18	8054.79	-4659.87	3394.92	1429.43	-2.32	-3.90	2.03	7507.17	-11500.55	-3993.38	-5873.04
8	(40, 40)	-16.36	-0.82	2.94	9200.14	-3675.38	5524.76	3396.49	-15.66	-4.19	2.91	9291.51	-14297.34	-5005.83	-9268.72
9	(25, 50)	-4.52	-1.32	2.48	9313.50	-5322.59	3990.91	379.58	-4.15	-3.74	2.43	9054.63	-14230.03	-5175.40	-7185.63
10	(50, 50)	-28.06	-0.73	3.82	11923.04	-4350.58	7572.46	2960.65	-25.15	-4.21	3.69	11470.12	-17995.81	-6525.69	-9851.95
11	(50, 100)	-28.72	-0.98	4.20	16308.28	-8621.45	7686.83	1053.65	-30.49	-2.90	4.00	15258.74	-24850.32	-9591.57	-14678.97
12	(100, 100)	-145.82	-0.67	8.39	25811.21	-7870.88	17940.33	8014.54	-147.62	-4.53	8.43	26137.27	-41134.03	-14996.75	-27555.56

the out-of-sample demand is lower than the in-sample expected value, the expected total cost of SDR is higher than that of SSP. This is because the additional facilities opened in SDR cannot be fully utilized 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 SSP 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), SSP is expected 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. It is also observed 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 SSP. 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 SSP, it is also compared 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 $CVaR_\alpha$, is defined as

$$CVaR_\alpha(z) = \inf_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{1-\alpha} \mathbb{E}[(z - \eta)_+] \right\}, \quad (20)$$

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

$$h(\mathbf{y}, \mathbf{c}, \mathbf{d}, e) = \min \left\{ \eta + \frac{1}{1-\alpha} \sum_{l=1}^L \frac{1}{L} \left(\sum_{i \in [I]} \sum_{j \in [J]} t_{ij} x_{ijl} + \sum_{i \in [I]} p_i u_{il} - \eta \right)_+ \right\}. \quad (21)$$

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

The comparison results between the SDR and the SSP-CVaR models are presented in Table 2. Note that since some instances with $(J, I) = (100, 100)$ cannot be solved optimally within the 1800-second time limit for the SSP-CVaR model, their results are not included here. Table 2 shows that SDR produces solutions with a lower total cost than SSP-CVaR in most cases, further confirming the superiority of the scenario-wise DRO framework. When $\alpha = 0.99$, the cost difference between the two models is generally greater than that under $\alpha = 0.90$, because when α is closer to 1, the SSP-CVaR model is more conservative.

Table 2 Quantity differences in average results of SDR and SSP-CVaR models

No.	(J, I)	$\Delta_2 < 0$				$\Delta_2 > 0$			
		$\alpha = 0.90$		$\alpha = 0.99$		$\alpha = 0.90$		$\alpha = 0.99$	
		$\#U^{qd}$	Cost_t^{qd}	$\#U^{qd}$	Cost_t^{qd}	$\#U^{qd}$	Cost_t^{qd}	$\#U^{qd}$	Cost_t^{qd}
1	(5, 10)	0.80	-692.53	0.92	-1071.48	2.85	-113.48	3.76	-282.13
2	(10, 10)	0.32	-852.57	0.36	-1026.62	1.80	-26.02	2.12	-73.48
3	(10, 20)	0.57	-1825.50	0.61	-2061.01	2.23	-144.79	2.32	-189.99
4	(20, 20)	-0.06	290.58	-0.06	-28.24	-1.11	-735.77	-0.94	-766.91
5	(15, 30)	0.22	-1242.38	0.30	-2029.33	0.87	-181.33	1.49	-190.65
6	(30, 30)	-0.25	1620.21	-0.13	840.51	-1.92	-1724.02	-1.31	-1432.20
7	(20, 40)	0.17	-1645.91	0.29	-2833.50	0.88	-24.46	1.36	52.31
8	(40, 40)	-0.14	2595.31	-0.13	2361.81	-1.92	-2144.17	-1.77	-2106.42
9	(25, 50)	0.23	-2126.38	0.25	-2426.26	0.85	-89.13	0.89	-362.93
10	(50, 50)	-0.30	4736.08	-0.30	4467.76	-2.38	-3638.45	-2.37	-3874.62
11	(50, 100)	0.04	-1435.07	0.08	-2650.57	-0.08	-1884.33	0.15	-1768.99

Comparison between SDR and MDR. Table 3 provides the comparison results between the SDR and the MDR models. Note that we do not compare their computing times, because the MDR is a special case of the SDR, i.e., SDR with one scenario. As reported before, the reformulation of SDR can be optimally solved within 2.08 seconds for all the instances; therefore, the reformulation of MDR is expected to be solved in a shorter time. 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 higher costs. 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 difference 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. Note that although MDR sometimes provides lower 95th percentiles of the expected total cost than SDR when Δ_2 is positive, the differences are smaller, compared to the case of a negative Δ_2 where SDR provides better 95th percentile values. The results in this section further confirm the necessity of using SDR to model scenario-wise uncertainties and capture the correlation between facilities' uncertain capacity and customers' uncertain demand in different scenarios.

Table 3 Quantity differences in average results of SDR and MDR models

No.	(J, I)	$\Delta_2 < 0$						$\Delta_2 > 0$					
		$\#U^{qd}$	$\#O^{qd}$	$Cost_1^{qd}$	$Cost_2^{qd}$	$Cost_i^{qd}$	$95th^{qd}$	$\#U^{qd}$	$\#O^{qd}$	$Cost_1^{qd}$	$Cost_2^{qd}$	$Cost_i^{qd}$	$95th^{qd}$
1	(5, 10)	1.34	-0.92	-3469.70	937.25	-2532.45	-3200.15	5.62	-0.86	-3167.01	2414.66	-752.35	-421.93
2	(10, 10)	0.44	-0.94	-3439.88	966.37	-2473.51	-3228.28	3.67	-0.94	-3363.50	2757.35	-606.15	219.67
3	(10, 20)	0.86	-2.11	-8316.67	1578.27	-6738.40	-5842.69	5.41	-2.11	-8330.51	6482.09	-1848.42	-1376.16
4	(20, 20)	0.22	-1.64	-6105.11	1188.11	-4917.00	-5427.39	2.89	-1.50	-5676.17	4873.10	-803.07	266.48
5	(15, 30)	0.65	-3.10	-12636.13	2051.40	-10584.72	-10795.08	5.33	-3.26	-13269.33	10390.37	-2878.96	-358.22
6	(30, 30)	0.15	-1.70	-6848.16	1343.02	-5505.14	-5950.47	2.33	-1.74	-6687.25	6055.59	-631.66	1323.29
7	(20, 40)	0.58	-3.92	-15933.73	2602.47	-13331.26	-12407.67	4.78	-3.89	-16086.31	13421.26	-2665.05	2611.55
8	(40, 40)	0.14	-1.90	-8110.42	1480.56	-6629.86	-6528.33	1.79	-1.86	-7854.58	6950.54	-904.04	278.13
9	(25, 50)	0.60	-5.01	-20593.23	3296.48	-17296.75	-14281.22	4.92	-5.11	-21089.78	18044.00	-3045.77	189.69
10	(50, 50)	0.08	-2.46	-9996.84	1647.70	-8349.14	-8496.47	2.00	-2.67	-10673.30	9391.58	-1281.72	2734.10
11	(50, 100)	0.55	-9.99	-42687.25	6438.97	-36248.27	-26744.32	4.60	-9.77	-41235.20	37471.27	-3763.93	6356.96
12	(100, 100)	0.03	-2.78	-12775.30	1849.79	-10925.50	-10822.51	1.02	-2.79	-13160.05	11250.52	-1909.53	777.54

5.2. An Earthquake Case Study

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, as shown in Figure 1. The numbers next to the links are the unit transportation costs. This section uses this earthquake case study to compare models and perform sensitivity analyses based on the following reasons: (1) It has been widely used in the literature on facility location under uncertainty (e.g., Ni et al. (2018), Shehadeh and Tucker (2022), and Zhang et al. (2021a)) to explore the impact of disasters on supply chain systems; (2) Earthquakes are one of the random events that we hedge against to improve supply chain robustness; and (3) The assumptions and parameter settings in Ni et al. (2018) can be directly adopted to accommodate our problem. In the following, the performance of SDR and MDR are compared to assess the value of the scenario-wise ambiguity set. The results of SSP are also reported to provide insights with respect to the costs of the robust solutions.

In the case study, 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. Two scenarios are considered, where $s = 1$ and $s = 2$ denote a major

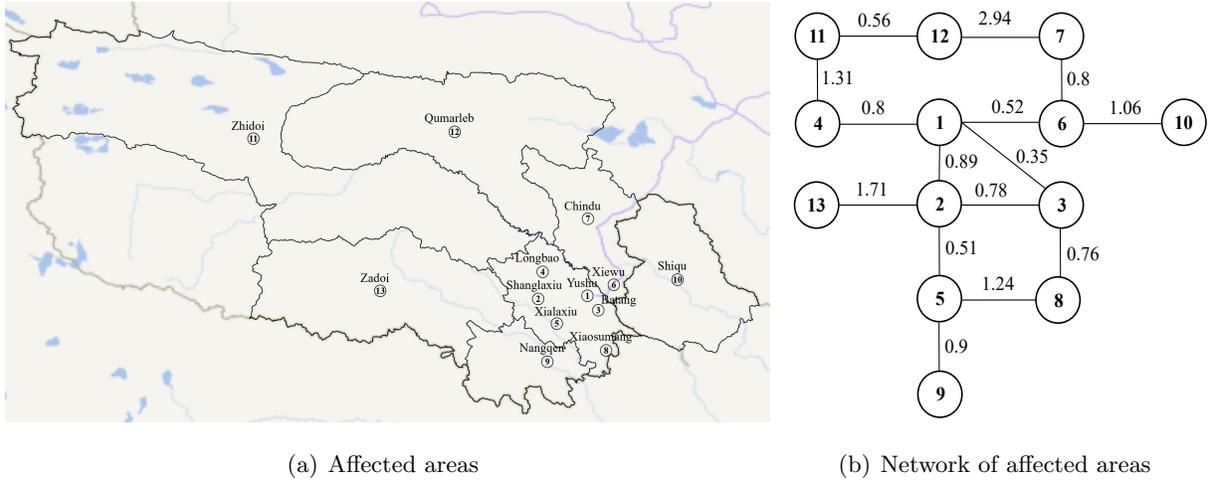


Figure 1 Network of affected areas in the earthquake case study

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 $s = 1$, the means and the standard deviations are set as in Ni et al. (2018) to denote the $M_s 7.1$ 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 $\mu_i^{ds}(1 + \zeta^1)$ and $\mu_j^{rs}(1 + \zeta^2)$, and the values of other parameters are the same as those for generating the training data. ζ^1 and ζ^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 (ζ^1, ζ^2) , 5 instances are generated, so there are 180 instances in total. For each instance, 50 samples are generated under each scenario, for a total of 100 training samples and 100 testing samples.

5.2.1. Results of Different Models. The results of different models are provided in Table 4. It shows that on average SSP (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. Regarding 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 SSP has the highest value. In addition, the cost of SSP 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. Regarding unmet demand, the out-of-sample mean, 95th percentile, and standard deviation produced by SSP are also much greater than those of the two other models.

Table 4 Results of different models for the case study

Model	Average					95th percentile		Standard deviation	
	Cost ₁	Cost ₂	Cost _t	#U	#O	Cost _t	#U	Cost _t	#U
SSP	836.11	727.27	1563.38	1.134	5.07	2659.60	7.620	471.15	2.823
SDR	1268.59	270.72	1539.32	0.001	8.11	1658.67	0.001	71.15	0.012
MDR	1389.22	212.37	1601.58	0.000	8.77	1681.38	0.000	59.86	0.000

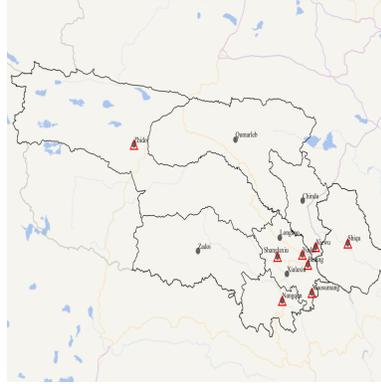
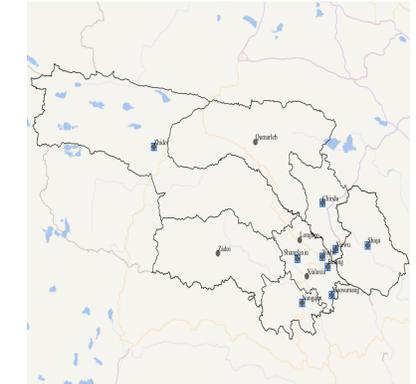
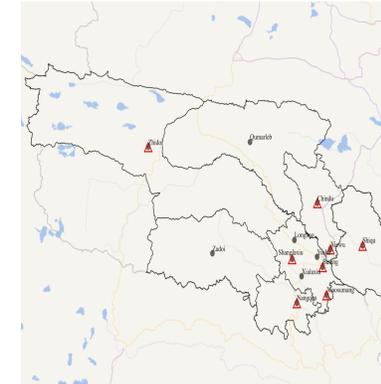
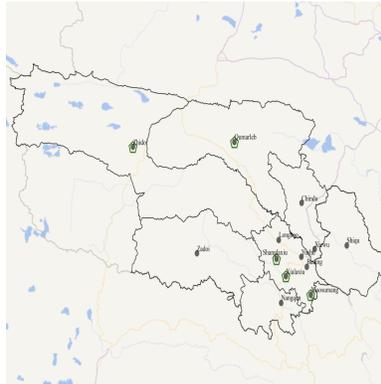
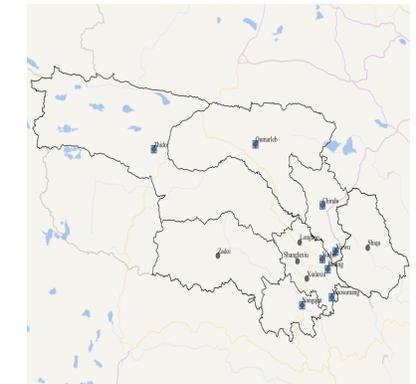
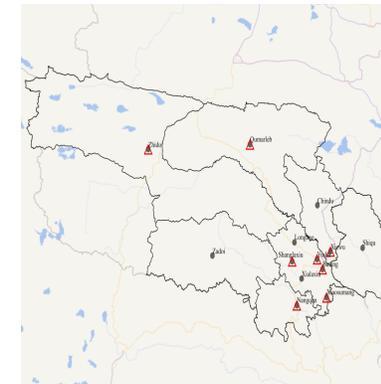
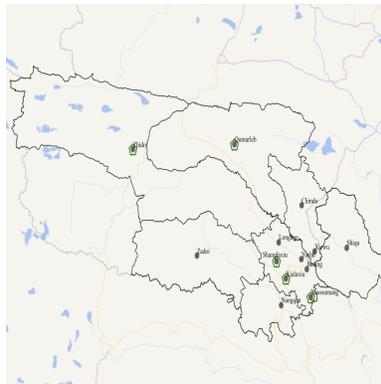
(a) SSP under $(\varsigma_1, \varsigma_2) = (-0.3, -0.3)$ (b) SDR under $(\varsigma_1, \varsigma_2) = (-0.3, -0.3)$ (c) MDR under $(\varsigma_1, \varsigma_2) = (-0.3, -0.3)$ (d) SSP under $(\varsigma_1, \varsigma_2) = (-0.3, 0.1)$ (e) SDR under $(\varsigma_1, \varsigma_2) = (-0.3, 0.1)$ (f) MDR under $(\varsigma_1, \varsigma_2) = (-0.3, 0.1)$ (g) SSP under $(\varsigma_1, \varsigma_2) = (0.3, 0.1)$ (h) SDR under $(\varsigma_1, \varsigma_2) = (0.3, 0.1)$ (i) MDR under $(\varsigma_1, \varsigma_2) = (0.3, 0.1)$ **Figure 2 Illustrations of location decisions produced by the three models for some instances**

Figure 2 illustrates the location decisions provided by the three models for some instances (the detailed results are presented in Table D3). Figures 2(b)–2(c) show that sometimes SDR and MDR may produce the same location decision. These two models also make different location decisions from two aspects: (1) Locating a different number of facilities, as shown in Figures 2(e)–2(f); and (2) Locating the same number of facilities while at different sites, as shown in Figures 2(h)–2(i).

5.2.2. Sensitivity Analyses. This section analyzes the sensitivity of solutions to three types of parameters: facilities’ fixed costs, the support set’s size, and the simultaneous variations of in-sample facility capacity and customer demand. Graphical results are reported in Figures 3–5 and detailed results are provided in Tables D4–D6 in Appendix D.

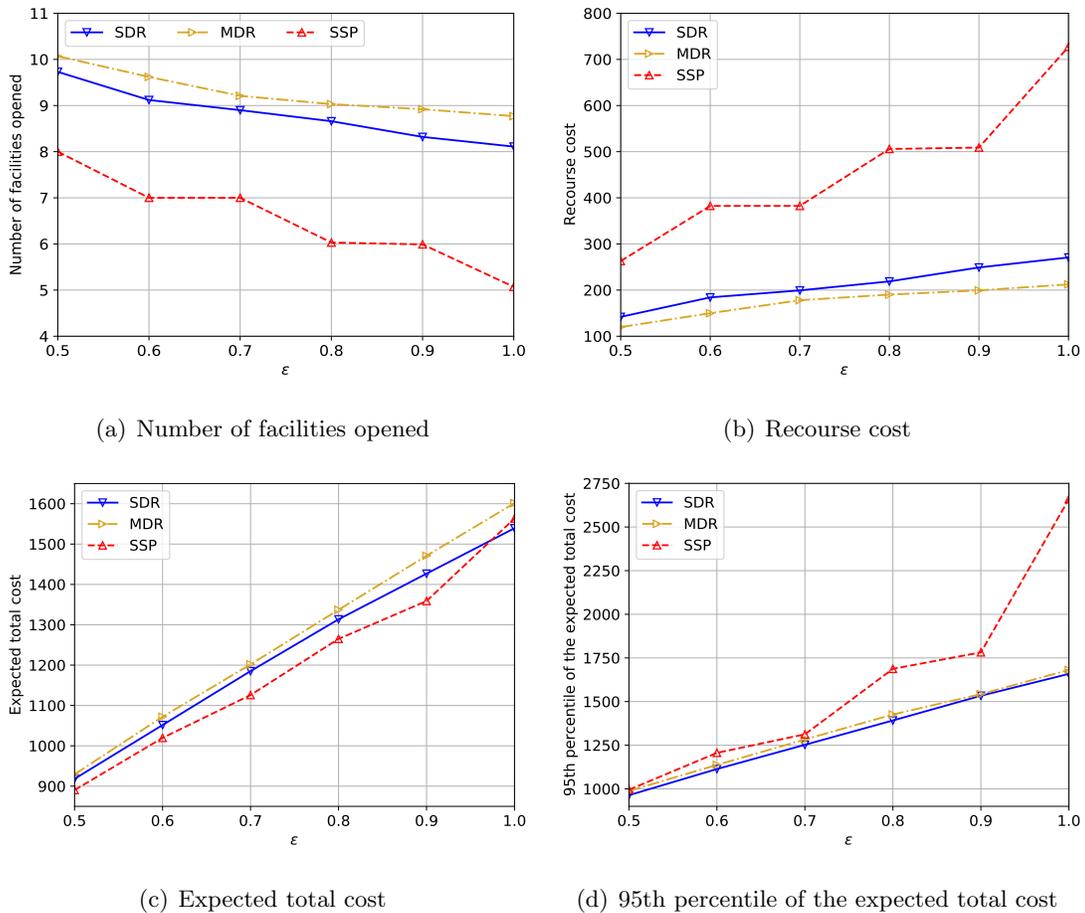


Figure 3 The impact of facilities’ fixed costs

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 3(a) show that MDR (SSP) 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 (SSP) still produces the most conservative

(optimistic) solutions. As a result, the recourse cost of SSP is much higher than those of the two other models as plotted in Figure 3(b). From Figure 3(c), it is found that when ϵ is between 0.5 and 0.9, the expected total cost of SSP is lower than those of the robust models. However, the 95th percentile of the expected total cost produced by SSP is the highest, especially when ϵ is between 0.8 and 1.0, as shown in Figure 3(d). SSP also has the largest number of unmet demands in out-of-sample tests, as shown in Table D4. Thus, we can conclude from the 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), the lower (upper) bounds of capacity and demand for each scenario are defined 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. The support set is shrunk by reducing the upper bounds as $\bar{c}'_s = (1 - \varrho)\bar{c}_s, \bar{d}'_s = (1 - \varrho)\bar{d}_s, s \in [S]$. A similar operation is also applied to the marginal moment-based ambiguity set (5) of MDR.

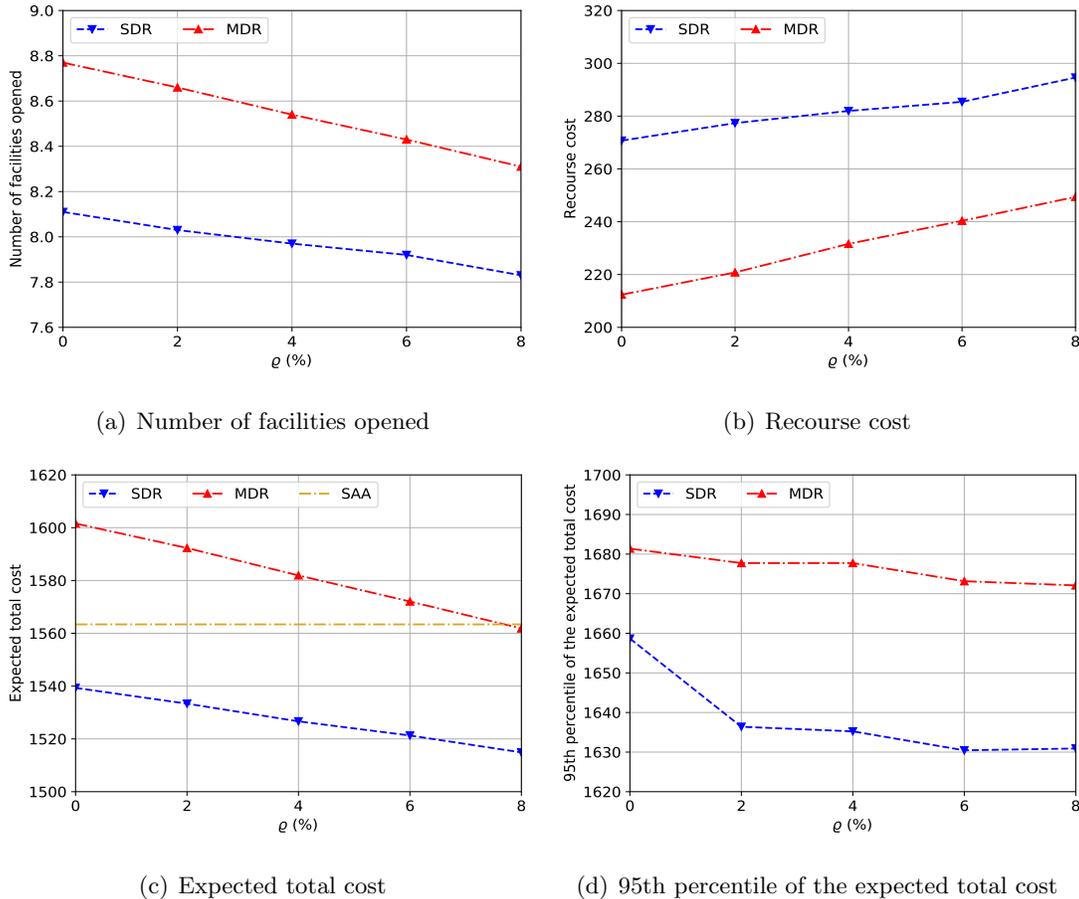


Figure 4 The impact of the support set's size

Figure 4 shows 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 4 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 D5.

Impact of in-sample facility capacity and customer demand. We simultaneously change in-sample facility capacity and customer demand and study their impacts on DRO models' costs. Specifically, we set $\mathbf{c}'_s = \nu_c \mathbf{c}_s$ and $\mathbf{d}'_s = \nu_d \mathbf{d}_s$ for all $s \in [S]$, where ν_c and ν_d take values in $\{0.90, 0.95, 1.00, 1.05, 1.10\}$. Thus, there are 25 combinations of (ν_c, ν_d) .

Figure 5(a) shows that both models have the minimal location cost under $(\nu_c, \nu_d) = (1.10, 0.90)$, i.e., under the case of the highest facility capacities and the lowest customer demands. As expected, the location cost gradually increases when ν_c decreases or ν_d increases. The MDR has a higher location cost under all the cases than the SDR. On the other hand, the additionally opened facilities by the SDR can mitigate the recourse cost in the second stage, as shown in Figure 5(b). Both models have the lowest recourse costs when $(\nu_c, \nu_d) = (0.90, 1.10)$, because they open the most facilities for this case. Figure 5(c) shows that the MDR has a higher total cost than the SDR for all the situations. This is because its location costs are much higher than those of SDR, although it has relatively lower recourse costs. In addition, it is observed that (1) when ν_d takes a larger value, e.g., $\nu_d = 1.1$, the total cost increases more quickly as ν_c decreases; and (2) when ν_c takes a smaller value, e.g., $\nu_c = 0.9$, the total cost increases more quickly as ν_d increases.

Based on the results in Sections 5.1 and 5.2, the following conclusions for the CFLP with simultaneous provider-side and receiver-side uncertainties are made: (1) SSP 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 SSP, 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) SSP 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; and (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.

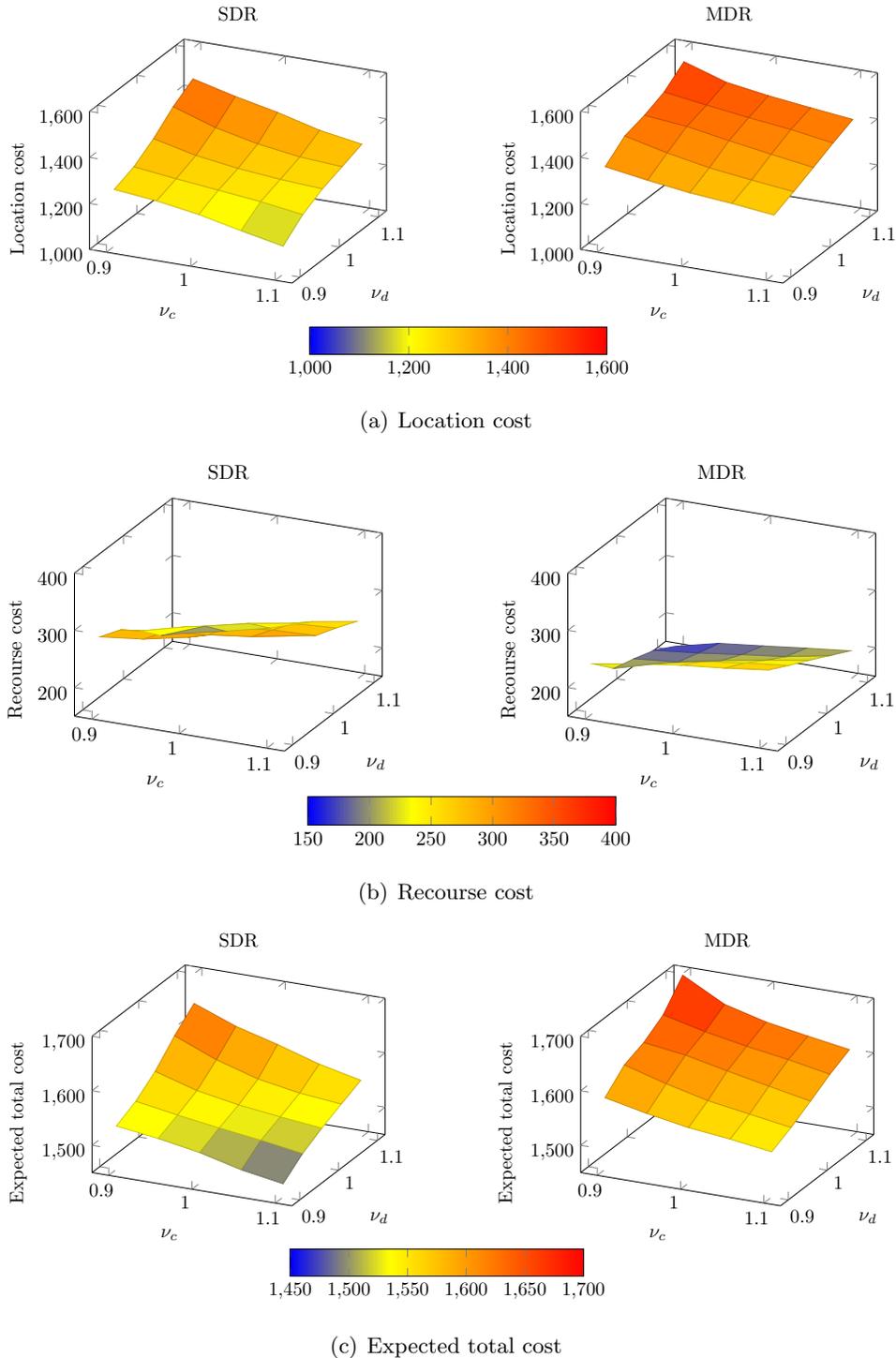


Figure 5 The simultaneous impact of in-sample facility capacity and customer demand

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, two scenarios are considered 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 $\mu_{j_1}^r, j \in [J]$ under scenario $s = 1$ as in Ni et al. (2018) and $\mu_{j_2}^r = 1.3\mu_{j_1}^r$ for scenario $s = 2$. All the related parameters are given in Table C2. For a fixed value of (R, ς^1) , 50 instances are generated, 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 5, where Cost_1^L and Cost_1^I are the first-stage location cost and inventory pre-positioning cost, respectively.

Table 5 Average results of different models for the location and inventory pre-positioning problem

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

Table 5 shows that SDR produces solutions with the lowest expected total cost and unmet demand, whereas SSP has the highest expected total cost and unmet demand. Moreover, SSP 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.

Besides the FLP and the LIPP presented in this paper, the scenario-wise DRO framework also applies to other applications that involve supply and/or demand uncertainties caused by different events. An example is the sectors that require global supply chains (e.g., the automobile and electronic industries). These companies' operations often involve multiple participants and facilities dispersed across different nations. Political instabilities in some countries may make their supply chains experience disruptions, leading to uncertainties. For example, the Russia-Ukraine war has disrupted some facilities and reduced the supply of crucial resources (e.g., nickel for batteries powering automobiles and electronics, natural gas, and oil). Thus, multinational companies that rely on these resources face supply-side uncertainty. Meanwhile, they also face demand fluctuations in local markets due to the soaring product prices caused by the disruption event. Moreover,

economic and cultural issues in some countries, such as the 2023 Canadian federal worker strike, may also expose multinational companies to uncertainties. Thus, companies involving global supply chains can use the proposed approach to consider all the disruption events of interest and make their decisions (e.g., location, production, inventory, and transportation). Besides, the scenario-wise framework also applies to local businesses with obvious supply or demand patterns under different scenarios (e.g., seasons and weather conditions), like the apparel and festival gift industries.

Furthermore, the proposed approach may also affect planners' decision-making process. Specifically, it may inspire planners to think more carefully about the reasons or events behind uncertainties and then distinguish the impacts of different events. In addition, the scenario-wise adaptation policy is more compatible with the human decision-making process, i.e., adopting different recourse actions for different events. Thus, the policy is more interpretable and likely to be adopted by practitioners. Finally, we emphasize that although the SDR does not show superior performance for all the cases, it gives better trade-offs between cost and service level in most situations. Thus, practitioners can utilize the SDR to solve most FLPs under uncertainty. On the other hand, the SSP may be preferable for cases where the probability distribution of random variables is available, or the demand decreases after disruption events.

6. Conclusions

This paper studies a capacitated facility location problem, considering facilities' uncertain capacity and customers' uncertain demand. A DRO framework is used 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. Correspondingly, a scenario-wise adaptation policy is adopted for the second-stage recourse problem to further mitigate the conservatism of robust solutions. The resulting adaptive DRO model is reformulated to an MILP model, which can be efficiently solved by off-the-shelf solvers. Simulation and case study results show that the scenario-wise DRO framework can achieve a better trade-off between cost and service level. 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. Compared with the stochastic programming model, the scenario-wise DRO model has lower values of unmet demand in out-of-sample tests, and thus it can better serve customers. The proposed modeling scheme is also extended to the location and inventory pre-positioning problem under uncertainties and shows satisfactory performance.

Future research could be conducted from two aspects: (1) The recourse variables \mathbf{x} and \mathbf{u} can be extended to scenario-wise affinely adaptive to random variables $(\tilde{\mathbf{c}}, \tilde{\mathbf{w}}, \tilde{\mathbf{d}}, \tilde{\mathbf{v}})$, i.e., \mathbf{x} and \mathbf{u} are affine functions of these variables in each scenario. This extension is expected to further reduce the

conservatism of DRO solutions and thus achieve even better performance when the demand has negative deviations from the expected values. (2) Besides reallocating customers to facilities with residual capacity, other recourse actions can be considered 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).

Acknowledgments

This research was supported by the National Natural Science Foundation of China [Grants 72101049, 72201267, 72232001], the Natural Science Foundation of Liaoning Province [Grant 2023-BS-091], the Fundamental Research Funds for the Central Universities [Grant DUT23RC(3)045], the Fundamental Research Funds for the Central Universities, Civil Aviation University of China [Grant 3122022093], and the Major Project of the National Social Science Foundation [Grant 22&ZD151]. The authors thank the editor and three anonymous referees for their helpful comments.

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.
- Ash, Cecil, Claver Diallo, Uday Venkatadri, Peter VanBerkel. 2022. Distributionally robust optimization of a canadian healthcare supply chain to enhance resilience during the covid-19 pandemic. *Computers & Industrial Engineering* **168** 108051.
- 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.
- 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.
- Dönmez, Zehranaz, Bahar Y Kara, Özlem Karsu, Francisco Saldanha-da Gama. 2021. Humanitarian facility location under uncertainty: Critical review and future prospects. *Omega* **102** 102393.
- 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.
- Jain, Anil K. 2010. Data clustering: 50 years beyond k-means **31**(8) 651–666.
- Jain, Anil K., Richard C. Dubes. 1988. *Algorithms for Clustering Data*. Prentice-Hall, Inc., USA.

- Kahr, Michael. 2022. Determining locations and layouts for parcel lockers to support supply chain viability at the last mile. *Omega* **113** 102721.
- 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. 2023. Robust pricing and production with information partitioning and adaptation. *Management Science* **69**(3) 1398–1419.

-
- 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.
- Sarf, Herbert E. 1957. A min-max solution of an inventory problem. Tech. rep., Santa Monica, Calif.
- Shang, Chao, Fengqi You. 2018. Distributionally robust optimization for planning and scheduling under uncertainty. *Computers & Chemical Engineering* **110** 53–68.
- Shehadeh, Karmel S. 2020. A distributionally robust optimization approach for a stochastic mobile facility routing and scheduling problem. *arXiv preprint arXiv:2009.10894* .
- Shehadeh, Karmel S, Ece Sancı. 2021. Distributionally robust facility location with bimodal random demand. *Computers & Operations Research* **134** 105257.
- Shehadeh, Karmel S, Emily L Tucker. 2022. Stochastic optimization models for location and inventory prepositioning of disaster relief supplies. *Transportation Research Part C: Emerging Technologies* **144** 103871.
- 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.
- Stienen, VF, JC Wagenaar, Dirk den Hertog, HA Fleuren. 2021. Optimal depot locations for humanitarian logistics service providers using robust optimization. *Omega* **104** 102494.
- 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. *IIE Transactions* **52**(10) 1122–1140.
- Wang, Shuming, Zhi Chen, Tianqi Liu. 2020. Distributionally robust hub location. *Transportation Science* **54**(5) 1189–1210.
- Wang, Weiqiao, Kai Yang, Lixing Yang, Ziyou Gao. 2023. Distributionally robust chance-constrained programming for multi-period emergency resource allocation and vehicle routing in disaster response operations. *Omega* **120** 102915.

- Wiesemann, Wolfram, Daniel Kuhn, Melvyn Sim. 2014. Distributionally robust convex optimization. *Operations Research* **62**(6) 1358–1376.
- Xiang, Xi, Changchun Liu. 2021. An expanded robust optimisation approach for the berth allocation problem considering uncertain operation time. *Omega* **103** 102444.
- 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, Yanzi Zhang, Zihao Jiao, Jing Wang. 2023. Improving relief operations via optimizing shelter location with uncertain covariates. *Transportation Research Part E: Logistics and Transportation Review* **176** 103181.
- Zhang, Yu, Zhenzhen Zhang, Andrew Lim, Melvyn Sim. 2021b. Robust data-driven vehicle routing with time windows. *Operations Research* **69**(2) 469–485.
- Zhong, Yuanguang, Ju Liu, Yong-Wu Zhou, Bin Cao, Xueliang Zheng. 2023. The role of ambiguity aversion in contract-farming supply chains: A distributionally robust approach. *Omega* **117** 102827.

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

$$\mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{c}}] = \sum_{s \in [S]} q_s \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{c}} | \tilde{s} = s] = \sum_{s \in [S]} q_s \boldsymbol{\mu}_s = \boldsymbol{\mu},$$

$$\mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{d}}] = \sum_{s \in [S]} q_s \mathbb{E}_{\mathbb{P}}[\tilde{\mathbf{d}} | \tilde{s} = s] = \sum_{s \in [S]} q_s \boldsymbol{\rho}_s = \boldsymbol{\rho}.$$

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

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

According to the law of total expectation, we have

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

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

The above inequalities suggest that $\mathbb{P} \in \mathcal{F}$ satisfy the first four constraints of set $\bar{\mathcal{F}}$. Moreover, combining $\mathbb{P}[(\mathbf{c}, \mathbf{d}) \in \Omega_s | \tilde{s} = s] = 1, s \in [S]$ and condition (iii), we have $\mathbb{P}[(\mathbf{c}, \mathbf{d}) \in \Omega] = 1$. 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 $(\tilde{\mathbf{c}}, \tilde{\mathbf{w}}, \tilde{\mathbf{d}}, \tilde{\mathbf{v}}, \tilde{s})$ from the marginal distribution $\hat{\mathbb{P}}$ of \tilde{s} supported on $[S]$ and the conditional distribution \mathbb{P}_s of $(\tilde{\mathbf{c}}, \tilde{\mathbf{w}}, \tilde{\mathbf{d}}, \tilde{\mathbf{v}})$ given $\tilde{s} = s, s \in [S]$. In this way, $\sup_{\mathbb{P} \in \mathcal{F}} \mathbb{E}_{\mathbb{P}}[h(\mathbf{y}, \mathbf{c}, \mathbf{d}, s)]$ can be rewritten as

$$\sup \sum_{s \in [S]} q_s \mathbb{E}_{\mathbb{P}_s} [h(\mathbf{y}, \tilde{\mathbf{c}}, \tilde{\mathbf{d}}, \tilde{s})] \tag{A.1a}$$

$$\text{s.t. } \mathbb{E}_{\mathbb{P}_s}[\tilde{\mathbf{c}}] = \boldsymbol{\mu}_s \quad \forall s \in [S], \tag{A.1b}$$

$$\mathbb{E}_{\mathbb{P}_s}[\tilde{\mathbf{w}}] \leq \boldsymbol{\delta}_s \quad \forall s \in [S], \tag{A.1c}$$

$$\mathbb{E}_{\mathbb{P}_s}[\tilde{\mathbf{d}}] = \boldsymbol{\rho}_s \quad \forall s \in [S], \tag{A.1d}$$

$$\mathbb{E}_{\mathbb{P}_s}[\tilde{\mathbf{v}}] \leq \boldsymbol{\zeta}_s \quad \forall s \in [S], \tag{A.1e}$$

$$\mathbb{P}_s[(\tilde{\mathbf{c}}, \tilde{\mathbf{w}}, \tilde{\mathbf{d}}, \tilde{\mathbf{v}}) \in \Omega'_s] = 1 \quad \forall s \in [S]. \tag{A.1f}$$

Let $\alpha_s \in \mathbb{R}^J$, $\beta_s \in \mathbb{R}^J$, $\lambda_s \in \mathbb{R}^I$, $\gamma_s \in \mathbb{R}^I$, 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

$$\min \sum_{s \in [S]} [(\mu_s)^T \alpha_s + (\delta_s)^T \beta_s + (\rho_s)^T \lambda_s + (\zeta_s)^T \gamma_s + \theta_s] \quad (\text{A.2a})$$

$$\text{s.t. } \mathbf{c}^T \alpha_s + \mathbf{w}^T \beta_s + \mathbf{d}^T \lambda_s + \mathbf{v}^T \gamma_s + \theta_s \geq q_s h(\mathbf{y}, \mathbf{c}, \mathbf{d}, s) \quad \forall (\mathbf{c}, \mathbf{w}, \mathbf{d}, \mathbf{v}) \in \Omega'_s, s \in [S], \quad (\text{A.2b})$$

$$\beta_s, \gamma_s \geq 0 \quad \forall s \in [S]. \quad (\text{A.2c})$$

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 $(\alpha_s, \beta_s, \lambda_s, \gamma_s, \theta_s)$, constraints (A.2b) are equivalent to

$$\theta_s \geq \max_{(\mathbf{c}, \mathbf{w}, \mathbf{d}, \mathbf{v}) \in \Omega'_s} [q_s h(\mathbf{y}, \mathbf{c}, \mathbf{d}, s) - (\mathbf{c}^T \alpha_s + \mathbf{w}^T \beta_s + \mathbf{d}^T \lambda_s + \mathbf{v}^T \gamma_s)] \quad \forall s \in [S].$$

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

A.3 Proof of Proposition 4

Under scenario $s \in [S]$ and a given value of $(\alpha_s, \beta_s, \lambda_s, \gamma_s)$, we solve the following optimization problem to reformulate the term:

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

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 (15).

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

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

Appendix C Data Related to the Earthquake Case Study

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.

Appendix D Detailed Results of Numerical Experiments

Table D3 illustrates the location decisions provided by the three models, where *Ins* is the instance number (recall that 5 instances are generated for each (ζ^1, ζ^2)) and circles are used to distinguish the decisions from the two robust models.

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^{+1}	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^{+2}	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}^+	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}^+	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.

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

Δ_2	No.	(J, I)	SDR						SSP					
			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

No.	(J, I)	$\Delta_2 < 0$					$\Delta_2 > 0$				
		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 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	SSP	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	SSP	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
4	(0.3, 0.1)	5	SSP	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

Ins is the instance number (recall that 5 instances are generated for each $(\varsigma^1, \varsigma^2)$). Circles are used to distinguish the decisions from the two robust models.

Table D4 Sensitivity analysis on facilities' fixed costs based on the case study

ϵ	SSP							SDR						MDR					
	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 D5 Sensitivity analysis on the size of the support set based on the case study

ϱ (%)	SDR							MDR					
	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	

Table D6 Sensitivity analysis on in-sample facility capacity and customer demand

ν_c	ν_d	SDR					MDR				
		Cost ₁	Cost ₂	Cost _t	#U	#O	Cost ₁	Cost ₂	Cost _t	#U	#O
0.90	0.90	1246.79	282.75	1529.54	0.00	7.99	1345.84	235.61	1581.46	0.00	8.52
0.90	0.95	1279.99	267.77	1547.76	0.00	8.19	1413.99	200.40	1614.39	0.00	8.92
0.90	1.00	1336.84	238.54	1575.38	0.00	8.52	1438.13	190.12	1628.25	0.00	9.06
0.90	1.05	1410.34	203.07	1613.41	0.00	8.92	1474.16	177.66	1651.82	0.00	9.24
0.90	1.10	1465.32	179.93	1645.25	0.00	9.21	1539.89	157.10	1696.99	0.00	9.58
0.95	0.90	1227.87	291.60	1519.47	0.00	7.88	1317.30	248.07	1565.37	0.00	8.34
0.95	0.95	1258.59	276.68	1535.28	0.00	8.06	1370.94	221.81	1592.75	0.00	8.66
0.95	1.00	1294.54	258.32	1552.86	0.00	8.27	1420.68	197.25	1617.93	0.00	8.96
0.95	1.05	1349.91	232.27	1582.18	0.00	8.59	1448.04	185.64	1633.68	0.00	9.11
0.95	1.10	1419.81	196.30	1616.10	0.00	8.97	1481.36	175.07	1656.42	0.00	9.28
1.00	0.90	1198.91	307.36	1506.27	0.01	7.71	1293.36	258.75	1552.10	0.00	8.20
1.00	0.95	1239.57	284.64	1524.21	0.00	7.94	1336.69	237.76	1574.45	0.00	8.46
1.00	1.00	1268.59	270.72	1539.32	0.00	8.11	1389.22	212.37	1601.58	0.00	8.77
1.00	1.05	1309.91	249.71	1559.62	0.00	8.34	1424.07	195.52	1619.59	0.00	8.98
1.00	1.10	1380.93	214.48	1595.41	0.00	8.75	1453.26	183.35	1636.61	0.00	9.14
1.05	0.90	1159.12	328.18	1487.30	0.02	7.48	1277.41	266.30	1543.71	0.00	8.10
1.05	0.95	1220.15	294.14	1514.29	0.00	7.82	1314.14	248.27	1562.41	0.00	8.32
1.05	1.00	1251.63	278.41	1530.04	0.00	8.01	1360.78	225.60	1586.38	0.00	8.60
1.05	1.05	1283.71	262.68	1546.39	0.00	8.18	1403.09	204.78	1607.87	0.00	8.86
1.05	1.10	1333.73	238.42	1572.15	0.00	8.47	1431.41	192.15	1623.56	0.00	9.02
1.10	0.90	1122.48	352.07	1474.55	0.02	7.26	1257.42	275.63	1533.05	0.00	7.99
1.10	0.95	1194.42	308.14	1502.56	0.01	7.67	1293.98	256.47	1550.45	0.00	8.21
1.10	1.00	1240.81	281.41	1522.22	0.00	7.93	1333.39	238.20	1571.59	0.00	8.43
1.10	1.05	1271.58	267.45	1539.03	0.00	8.11	1374.64	219.05	1593.70	0.00	8.68
1.10	1.10	1303.70	251.41	1555.11	0.00	8.29	1411.28	200.14	1611.42	0.00	8.91