# Models for Minimax Stochastic Linear Optimization Problems with Risk Aversion

Dimitris Bertsimas* Xuan Vinh Doan† Karthik Natarajan‡ Chung-Piaw Teo§

March 2008

**Abstract**

In this paper, we propose a semidefinite optimization (SDP) based model for the class of minimax two-stage stochastic linear optimization problems with risk aversion. The distribution of the second-stage random variables is assumed to be chosen from a set of multivariate distributions with known mean and second moment matrix. For the minimax stochastic problem with random objective, we provide a tight polynomial time solvable SDP formulation. For the minimax stochastic problem with random right-hand side, the problem is shown to be NP-hard in general. When the number of extreme points in the dual polytope of the second-stage stochastic problem is bounded by a function which is polynomial in the dimension, the problem can be solved in polynomial time. Explicit constructions of the worst case distributions for the minimax problems are provided. Applications in a production-transportation problem and a single facility minimax distance problem are provided to demonstrate our approach. In our computational experiments, the performance of minimax solutions is close to that of data-driven solutions under the multivariate normal distribution and is better under extremal distributions. The minimax solutions thus guarantee to hedge against these worst possible distributions while also providing a natural distribution to stress test stochastic optimization problems under distributional ambiguity.

*Boeing Professor of Operations Research, Sloan School of Management, co-director of the Operations Research Center, Massachusetts Institute of Technology, E40-147, Cambridge, MA 02139-4307, dbertsim@mit.edu.

†Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, vanxuan@mit.edu.

‡Department of Mathematics, National University of Singapore, Singapore 117543, matkbn@nus.edu.sg. The research of the author was partially supported by Singapore-MIT Alliance, NUS Risk Management Institute and NUS startup grants R-146-050-070-133 & R146-050-070-101.

§Department of Decision Sciences, Business School, National University of Singapore, Singapore 117591, bizteocp@nus.edu.sg.

# 1 Introduction

Consider a two-stage stochastic linear optimization problem with recourse:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{c}'\boldsymbol{x} + \mathbb{E}_P\left[\mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x})\right] \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned}$$

where the recourse function is given as:

$$\begin{aligned} \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) = \min_{\boldsymbol{w}} \quad & \tilde{\boldsymbol{q}}'\boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{W}\boldsymbol{w} = \tilde{\boldsymbol{h}} - \boldsymbol{T}\boldsymbol{x},\ \boldsymbol{w} \geq \boldsymbol{0}. \end{aligned}$$

The first-stage decision $\boldsymbol{x}$ is chosen from the set $X := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$ before the exact value of the random parameters $\tilde{\boldsymbol{\xi}} = (\tilde{\boldsymbol{q}}, \tilde{\boldsymbol{h}})$ are known. After the random parameters are realized, the second-stage (or recourse) decision $\boldsymbol{w}$ is chosen from the set $X(\boldsymbol{x}) := \{\boldsymbol{w} \in \mathbb{R}^p : \boldsymbol{W}\boldsymbol{w} = \tilde{\boldsymbol{h}} - \boldsymbol{T}\boldsymbol{x}, \boldsymbol{w} \geq \boldsymbol{0}\}$. The goal is to minimize the sum of the first-stage cost $\boldsymbol{c}'\boldsymbol{x}$ and the expected second-stage cost $\mathbb{E}_P\left[\mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x})\right]$ when the random parameters are sampled from the probability distribution $P$.

The traditional model for stochastic optimization considers complete knowledge of the probability distribution of the random parameters. Using a scenario-based formulation; say $(\boldsymbol{q}_1, \boldsymbol{h}_1), \ldots, (\boldsymbol{q}_N, \boldsymbol{h}_N)$ occurs with probabilities $p_1, \ldots, p_N$, the stochastic optimization problem is solved as

$$\begin{aligned} \min_{\boldsymbol{x}, \boldsymbol{w}_t} \quad & \boldsymbol{c}'\boldsymbol{x} + \sum_{t=1}^{N} p_t \boldsymbol{q}_t' \boldsymbol{w}_t \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{W}\boldsymbol{w}_t = \boldsymbol{h}_t - \boldsymbol{T}\boldsymbol{x},\ \boldsymbol{w}_t \geq \boldsymbol{0}. \end{aligned}$$

This linear optimization problem forms the basis of sampling based methods for stochastic optimization. The reader is referred to Shapiro and Homem-de-Mello [20] and Kleygweyt et al. [14] for convergence results on sampling methods for stochastic optimization problems. One difficulty with using a scenario-based formulation is that one needs to estimate the probabilities for the finitely many different scenarios. Secondly, the solution can be sensitive to the choice of the probability distribution $P$. To stress test the solution, Dupacova [9] suggests using a contaminated distribution $P_\lambda = (1-\lambda)P + \lambda Q$ for $\lambda \in [0, 1]$ where $Q$ is an alternate distribution that carries out-of-sample scenarios. Minimax stochastic optimization provides an approach to address this issue of ambiguity in distributions. Instead of assuming a single distribution, one hedges against an entire class of probability distributions. The minimax stochastic optimization problem is formulated as:

$$\min_{\boldsymbol{x} \in X} \left( \boldsymbol{c}'\boldsymbol{x} + \max_{P \in \mathcal{P}} \mathbb{E}_P\left[\mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x})\right] \right),$$

where the decision $\boldsymbol{x}$ is chosen to minimize the first-stage cost and the worst case expected second-stage cost calculated over all the distributions $P \in \mathcal{P}$. This approach has been analyzed by Žáčková [24], Dupačová [8] and more recently by Shapiro and Kleywegt [21] and Shapiro and Ahmed [19]. Among others, algorithms for minimax stochastic optimization include the sample-average approximation method (see Shapiro and Kleywegt [21] and Shapiro and Ahmed [19]), subgradient-based methods (see Breton and El Hachem [4]) and cutting plane algorithms (see Riis and Anderson [17]). The set $\mathcal{P}$ is typically described by a set of known moments. Bounds on the expected second-stage cost using first moment information include the Jensen bound and the Edmunson-Madansky bound. For extensions to second moment bounds in stochastic optimization, the reader is referred to Kall and Wallace [13] and Dokov and Morton [7].

In addition to modeling ambiguity in distributions, one is often interested in incorporating risk considerations into stochastic optimzation. An approach to model the risk in the second-stage cost is to use a convex nondecreasing disutility function $\mathbb{U}(\cdot)$:

$$\min_{\boldsymbol{x} \in X} \left( \boldsymbol{c}'\boldsymbol{x} + \mathbb{E}_P \left[ \mathbb{U} \left( \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) \right) \right] \right).$$

Special instances for this problem include:

(1) Using a weighted combination of the expected mean and expected excess beyond a target $T$:

$$\min_{\boldsymbol{x} \in X} \left( \boldsymbol{c}'\boldsymbol{x} + \mathbb{E}_P \left[ \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) \right] + \alpha \mathbb{E}_P \left[ \left( \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) - T \right)^+ \right] \right),$$

where the weighting factor $\alpha \geq 0$. This formulation is convexity preserving in the first-stage variables (see Ahmed [1] and Eichorn and Römisch [10]).

(2) Using the optimized certainty equivalent (OCE) risk measure (see Ben-Tal and Teboulle [2], [3]):

$$\min_{\boldsymbol{x} \in X, v \in \Re} \left( \boldsymbol{c}'\boldsymbol{x} + v + \mathbb{E}_P \left[ \mathbb{U} \left( \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) - v \right) \right] \right),$$

with the disutility function normalized to satisfy $\mathbb{U}(0) = 0$ and $1 \in \partial \mathbb{U}(0)$ with $\partial \mathbb{U}(\cdot)$ denotes the subdifferential map of $\mathbb{U}$. For particular choices of utility functions, Ben-Tal and Teboulle [2], [3] show that the OCE risk measure can be reduced to the mean-variance formulation and the mean-Conditional Value-at-Risk formulation. Ahmed [1] shows that using the mean-variance criterion in stochastic optimization leads to NP-hard problems. This arises from the observation that the second-stage cost $\mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x})$ is nonlinear (but convex) in $\boldsymbol{x}$ while the variance operator is convex (but non-monotone). On the other hand, the mean-Conditional Value-at-Risk formulation is convexity preserving.

## Contributions and Paper Outline

In this paper, we analyze two-stage minimax stochastic linear optimization problems with the class of probability distributions described by first and second moments. We consider separate models to incorporate the randomness in the objective and right-hand side respectively. The probability distribution $P$ is assumed to belong to the class of distributions $\mathcal{P}$ specified by the mean vector $\boldsymbol{\mu}$ and second moment matrix $\boldsymbol{Q}$. In addition to ambiguity in distributions, we incorporate risk considerations into the model by using a nondecreasing piecewise linear convex disutility function $\mathbb{U}$ on the second-stage costs. The central problem we will study is

$$Z = \min_{\boldsymbol{x} \in X} \left( \boldsymbol{c}'\boldsymbol{x} + \max_{P \in \mathcal{P}} \mathbb{E}_P \left[ \mathbb{U}\left( \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) \right) \right] \right), \tag{1}$$

where the disutility function is defined as:

$$\mathbb{U}\left( \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) \right) := \max_{k=1,\ldots,K} \left( \alpha_k \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) + \beta_k \right), \tag{2}$$

with the coefficients $\alpha_k \geq 0$ for all $k$. A related minimax problem discussed in Rutenberg [18] is to incorporate the first-stage costs into the utility function

$$\min_{\boldsymbol{x} \in X} \left( \max_{P \in \mathcal{P}} \mathbb{E}_P \left[ \mathbb{U}\left( \boldsymbol{c}'\boldsymbol{x} + \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) \right) \right] \right),$$

This formulation can be handled in our model by defining $\beta_k(\boldsymbol{x}) = \alpha_k \boldsymbol{c}'\boldsymbol{x} + \beta_k$ and solving

$$\min_{\boldsymbol{x} \in X} \max_{P \in \mathcal{P}} \mathbb{E}_P \left[ \max_{k=1,\ldots,K} \left( \alpha_k \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) + \beta_k(\boldsymbol{x}) \right) \right].$$

For $K = 1$ with $\alpha_K = 1$ and $\beta_K = 0$, problem (1)-(2) reduces to the traditional risk-neutral minimax stochastic optimization problem. Throughout the paper, we make the following assumptions:

(A1) The first-stage feasible region $X := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$ is bounded and non-empty.

(A2) The recourse matrix $\boldsymbol{W}$ satisfies the complete fixed recourse condition $\{\boldsymbol{z} \in \mathbb{R}^r : \boldsymbol{W}\boldsymbol{w} = \boldsymbol{z}, \boldsymbol{w} \geq \boldsymbol{0}\} = \Re^r$.

(A3) The recourse matrix $\boldsymbol{W}$ together with $\tilde{\boldsymbol{q}}$ satisfies the condition $\{\boldsymbol{p} \in \mathbb{R}^r : \boldsymbol{W}'\boldsymbol{p} \leq \boldsymbol{q}\} \neq \emptyset$ for all $\boldsymbol{q}$.

(A4) The first and second moments $(\boldsymbol{\mu}, \boldsymbol{Q})$ of the random vector $\tilde{\boldsymbol{\xi}}$ are finite and satisfy $\boldsymbol{Q} \succ \boldsymbol{\mu}\boldsymbol{\mu}'$.

Assumptions (A1)-(A4) guarantee that the expected second-stage costs $\mathbb{E}_P \left[ \mathbb{U}\left( \mathcal{Q}(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}) \right) \right]$ is finite for all $P \in \mathcal{P}$ and the minimax risk-averse stochastic optimization problem is well-defined.

The contributions and structure of the paper are as follows:

(1) In Section 2, we propose a tight polynomial time solvable semidefinite optimization formulation for the risk-averse and risk-neutral minimax stochastic optimization model when the uncertainty is in the objective coefficients of the second-stage linear optimization problem. We provide an explicit construction for the worst case distribution of the second-stage problem. For the risk-neutral case, the second-stage bound reduces to the simple Jensen bound while for the risk-averse case it is a convex combination of Jensen bounds.

(2) In Section 3, we prove the NP-hardness of the risk-averse and risk-neutral minimax stochastic optimization model with random right-hand side of the second-stage linear optimization problem. We consider a special case in which the problem can be solved by a polynomial sized semidefinite optimization problem. We provide an explicit construction for the worst case distribution of the second-stage problem in this case.

(3) In Section 4, we report computational results for a production-transportation problem (random objective) and a single facility minimax distance problem (random right-hand side) respectively. These results show that the performance of minimax solutions is close to that of data-driven solutions under the multivariate normal distribution and it is better under extremal distributions. The construction of the worst case distribution also provides a natural distribution to stress test the solution of stochastic optimization problems.

## 2 Uncertainty in Objective

Consider the minimax stochastic problem (1) with random objective $\tilde{\boldsymbol{q}}$ and constant right-hand side $\boldsymbol{h}$. The distribution class $\mathcal{P}$ is specified by the first and second moments:

$$\mathcal{P} = \Big\{ P \ : \ \mathbb{P}[\tilde{\boldsymbol{q}} \in \Re^p] = 1, \ \mathbb{E}_P[\tilde{\boldsymbol{q}}] = \boldsymbol{\mu}, \ \mathbb{E}_P[\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}'] = \boldsymbol{Q} \Big\}. \tag{3}$$

The second-stage cost with risk aversion and objective uncertainty is then given as

$$\mathbb{U}\left(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x})\right) := \max_{k=1,\ldots,K} \left(\alpha_k \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_k\right),$$

where

$$\begin{aligned} \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) = \min\nolimits_{\boldsymbol{w}} \quad & \tilde{\boldsymbol{q}}'\boldsymbol{w} \\ \text{s.t.} \quad & \boldsymbol{W}\boldsymbol{w} = \boldsymbol{h} - \boldsymbol{T}\boldsymbol{x}, \ \boldsymbol{w} \geq \boldsymbol{0}. \end{aligned}$$

The second-stage cost $\mathbb{U}\left(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x})\right)$ is a quasi-concave function in $\tilde{\boldsymbol{q}}$ and a convex function in $\boldsymbol{x}$. This follows from observing that it is the composition of a nondecreasing convex function $\mathbb{U}(\cdot)$, and a function

$\mathcal{Q}(\cdot,\cdot)$ that is concave in $\tilde{\boldsymbol{q}}$ and convex in $\boldsymbol{x}$. A semidefinite formulation for identifying the optimal first-stage decision is developed in Section 2.1 while the extremal distribution for the second-stage problem is constructed in Section 2.2.

### 2.1 Semidefinite Optimization Formulation

The second-stage problem $\max_{P\in\mathcal{P}} \mathbb{E}_P\left[\mathbb{U}\left(\mathcal{Q}(\tilde{\boldsymbol{q}},\boldsymbol{x})\right)\right]$ in this case, is an infinite-dimensional linear optimization problem with the probability distribution $P$ or its corresponding probability density function $f$ as the decision variable:

$$
\begin{array}{rlll}
Z(\boldsymbol{x}) = \max_f & \int_{\mathbb{R}^p} \mathbb{U}\left(\mathcal{Q}(\boldsymbol{q},\boldsymbol{x})\right) f(\boldsymbol{q})\mathrm{d}\boldsymbol{q} & & \\
\text{s.t.} & \int_{\mathbb{R}^p} q_i q_j f(\boldsymbol{q})\mathrm{d}\boldsymbol{q} = Q_{ij}, & \forall\, i,j = 1,\ldots,p, & \\
& \int_{\mathbb{R}^p} q_i f(\boldsymbol{q})\mathrm{d}\boldsymbol{q} = \mu_i, & \forall\, i = 1,\ldots,p, & \\
& \int_{\mathbb{R}^p} f(\boldsymbol{q})\mathrm{d}\boldsymbol{q} = 1, & & \\
& f(\boldsymbol{q}) \geq 0, & \forall\, \boldsymbol{q}\in\mathbb{R}^p. &
\end{array}
\tag{4}
$$

Associating dual variables $\boldsymbol{Y}\in\mathbb{S}^{p\times p}$ where $\mathbb{S}^{p\times p}$ is the set of symmetric matrices of dimension $p$ and vector $\boldsymbol{y}\in\mathbb{R}^p$, and scalar $y_0\in\mathbb{R}$ with the constraints of the primal problem (4), we obtain the dual problem:

$$
\begin{array}{rl}
Z_D(\boldsymbol{x}) = \min_{\boldsymbol{Y},\boldsymbol{y},y_0} & \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\
\text{s.t.} & \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 \geq \mathbb{U}\left(\mathcal{Q}(\boldsymbol{q},\boldsymbol{x})\right), \quad \forall\,\boldsymbol{q}\in\mathbb{R}^p.
\end{array}
\tag{5}
$$

It is easy to verify that weak duality holds, namely $Z(\boldsymbol{x}) \leq Z_D(\boldsymbol{x})$. Furthermore, if the moment vector lies in the interior of the set of feasible moment vectors, then we have strong duality, namely $Z(\boldsymbol{x}) = Z_D(\boldsymbol{x})$. The reader is referred to Isii [12] for strong duality results in the moment problem. Assumption (A4) guarantees that the covariance matrix $\boldsymbol{Q}-\boldsymbol{\mu}\boldsymbol{\mu}'$ is strictly positive definite and the strong duality condition is satisfied. This result motivates us to replace the second-stage problem by its corresponding dual and attempt to solve this model. The risk-averse minimax stochastic optimization problem is then reformulated as a semidefinite optimization problem as shown in the following theorem:

**Theorem 1** *The risk-averse minimax stochastic optimization problem (1) with random objective* $\tilde{\boldsymbol{q}}$ *and* *constant right-hand side $\boldsymbol{h}$ is equivalent to the semidefinite optimization problem:*

$$
\begin{aligned}
Z_{SDP} = \min_{\boldsymbol{x},\boldsymbol{Y},\boldsymbol{y},y_0,\boldsymbol{w}_k} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\
\text{s.t.} \quad & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k\boldsymbol{w}_k}{2} \\ \frac{(\boldsymbol{y}-\alpha_k\boldsymbol{w}_k)'}{2} & y_0-\beta_k \end{pmatrix} \succeq 0, \quad && \forall\, k = 1,\dots,K, \\
& \boldsymbol{W}\boldsymbol{w}_k + \boldsymbol{T}\boldsymbol{x} = \boldsymbol{h}, && \forall\, k = 1,\dots,K, \\
& \boldsymbol{w}_k \geq \boldsymbol{0}, && \forall\, k = 1,\dots,K, \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}.
\end{aligned}
\tag{6}
$$

**Proof.** The constraints of the dual problem (5) can be written as follows:

$$(\mathcal{C}_k) : \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 \geq \alpha_k \mathcal{Q}(\boldsymbol{q},\boldsymbol{x}) + \beta_k \quad \forall\, \boldsymbol{q} \in \mathbb{R}^p,\ k = 1,\dots,K.$$

We first claim that $\boldsymbol{Y} \succeq 0$. Suppose $\boldsymbol{Y} \not\succeq 0$. Consider the eigenvector $\boldsymbol{q}_0$ of $\boldsymbol{Y}$ corresponding to the most negative eigenvalue $\lambda_0$. Define $F_k(\boldsymbol{q},\boldsymbol{x}) := \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 - \alpha_k\mathcal{Q}(\boldsymbol{q},\boldsymbol{x}) - \beta_k$ and let $\boldsymbol{w}_0 \in \arg\min_{\boldsymbol{w}\in X(\boldsymbol{x})} \boldsymbol{q}_0'\boldsymbol{w}$. The function $f(t) = F_k(t\boldsymbol{q}_0,\boldsymbol{x})$ is then a quadratic concave function in $t \in [0,+\infty)$:

$$f(t) = \lambda_0 \boldsymbol{q}_0'\boldsymbol{q}_0 t^2 + (\boldsymbol{y} - \alpha_k\boldsymbol{w}_0)'\boldsymbol{q}_0 t + y_0 - \beta_k.$$

Therefore, there exists $t_k$ such that for all $t \geq t_k$, $F_k(t\boldsymbol{q}_0,\boldsymbol{x}) < 0$. The constraint $(\mathcal{C}_k)$ is then violated (contradiction). Thus $\boldsymbol{Y} \succeq 0$.

Since we have $\mathcal{Q}(\boldsymbol{q},\boldsymbol{x}) = \min_{\boldsymbol{w}\in X(\boldsymbol{x})} \boldsymbol{q}'\boldsymbol{w}$ and $\alpha_k \geq 0$, the constraint $(\mathcal{C}_k)$ can be rewritten as follows:

$$\forall\, \boldsymbol{q} \in \mathbb{R}^p, \exists\, \boldsymbol{w}_k \in X(\boldsymbol{x}) : \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 - \alpha_k\boldsymbol{q}'\boldsymbol{w}_k - \beta_k \geq 0,$$

or equivalently

$$\inf_{\boldsymbol{q}\in\mathbb{R}^p} \sup_{\boldsymbol{w}_k\in X(\boldsymbol{x})} \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 - \alpha_k\boldsymbol{q}'\boldsymbol{w}_k - \beta_k \geq 0.$$

Since $\boldsymbol{Y} \succeq 0$, the continuous function $\boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 - \alpha_k\boldsymbol{q}'\boldsymbol{w}_k - \beta_k$ is convex in $\boldsymbol{q}$ and affine (concave) in $\boldsymbol{w}_k$. In addition, the set $X(\boldsymbol{x})$ is a bounded convex set; then, according to Sion's minimax theorem [22], we obtain the following result:

$$\inf_{\boldsymbol{q}\in\mathbb{R}^p} \sup_{\boldsymbol{w}_k\in X(\boldsymbol{x})} \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 - \alpha_k\boldsymbol{q}'\boldsymbol{w}_k - \beta_k = \sup_{\boldsymbol{w}_k\in X(\boldsymbol{x})} \inf_{\boldsymbol{q}\in\mathbb{R}^p} \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 - \alpha_k\boldsymbol{q}'\boldsymbol{w}_k - \beta_k.$$

Thus the constraint $(\mathcal{C}_k)$ is equivalent to the following constraint:

$$\exists\, \boldsymbol{w}_k \in X(\boldsymbol{x}), \forall\, \boldsymbol{q} \in \mathbb{R}^p : \boldsymbol{q}'\boldsymbol{Y}\boldsymbol{q} + \boldsymbol{q}'\boldsymbol{y} + y_0 - \alpha_k\boldsymbol{q}'\boldsymbol{w}_k - \beta_k \geq 0.$$

The equivalent matrix linear inequality constraint is

$$\exists\, \boldsymbol{w}_k \in X(\boldsymbol{x}) : \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k \boldsymbol{w}_k}{2} \\ \frac{(\boldsymbol{y}-\alpha_k \boldsymbol{w}_k)'}{2} & y_0 - \beta_k \end{pmatrix} \succeq 0.$$

The dual problem of the minimax second-stage optimization problem can be reformulated as follows:

$$\begin{aligned} Z_D(\boldsymbol{x}) = \min_{\boldsymbol{Y},\boldsymbol{y},y_0,\boldsymbol{w}_k} \quad & \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k \boldsymbol{w}_k}{2} \\ \frac{(\boldsymbol{y}-\alpha_k \boldsymbol{w}_k)'}{2} & y_0 - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, k = 1,\ldots,K, \\ & \boldsymbol{W}\boldsymbol{w}_k + \boldsymbol{T}\boldsymbol{x} = \boldsymbol{h}, \qquad \forall\, k = 1,\ldots,K, \\ & \boldsymbol{w}_k \geq \boldsymbol{0}, \qquad \forall\, k = 1,\ldots,K. \end{aligned} \tag{7}$$

By optimizing over the first-stage variables, we obtain the semidefinite optimization reformulation for our risk-averse minimax stochastic optimization problem:

$$\begin{aligned} Z_{SDP} = \min_{\boldsymbol{x},\boldsymbol{Y},\boldsymbol{y},y_0,\boldsymbol{w}_k} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\ \text{s.t.} \quad & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k \boldsymbol{w}_k}{2} \\ \frac{(\boldsymbol{y}-\alpha_k \boldsymbol{w}_k)'}{2} & y_0 - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, k = 1,\ldots,K, \\ & \boldsymbol{W}\boldsymbol{w}_k + \boldsymbol{T}\boldsymbol{x} = \boldsymbol{h}, \qquad \forall\, k = 1,\ldots,K, \\ & \boldsymbol{w}_k \geq \boldsymbol{0}, \qquad \forall\, k = 1,\ldots,K, \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}. \end{aligned}$$

With the strong duality assumption, $Z_D(\boldsymbol{x}) = Z(\boldsymbol{x})$ for all $\boldsymbol{x} \in X$. Thus $Z = Z_{SDP}$ or (6) is the equivalent semidefinite optimization formulation of our risk-averse minimax stochastic optimization problem (1) with random objective $\tilde{\boldsymbol{q}}$ and constant right-hand side $\boldsymbol{h}$. □

In related recent work, Delage and Ye [6] use an ellipsoidal algorithm to show that the minimax stochastic optimization problem

$$\min_{\boldsymbol{x}\in X} \max_{P\in\hat{\mathcal{P}}} \mathbb{E}_P \left[ \max_{k=1,\ldots,K} f_k\left(\tilde{\boldsymbol{\xi}}, \boldsymbol{x}\right) \right].$$

is polynomial time solvable under the assumptions:

(i) The set $X$ is convex and equipped with an oracle that confirms the feasibility of $\boldsymbol{x}$ or provides a separating hyperplane in polynomial time in the dimension of the problem.

(ii) For each $k$, the function $f_k(\boldsymbol{\xi}, \boldsymbol{x})$ is concave in $\boldsymbol{\xi}$ and convex in $\boldsymbol{x}$. In addition, one can in polynomial time find the value $f_k(\boldsymbol{\xi}, \boldsymbol{x})$, a subgradient of $f_k(\boldsymbol{\xi}, \boldsymbol{x})$ in $\boldsymbol{x}$ and a subgradient of $-f_k(\boldsymbol{\xi}, \boldsymbol{x})$ in $\boldsymbol{\xi}$.

(iii) The class of distributions $\hat{\mathcal{P}}$ is defined as

$$\hat{\mathcal{P}} = \left\{ P \ : \ \mathbb{P}[\tilde{\boldsymbol{q}} \in \mathcal{S}] = 1, \ (\mathbb{E}_P[\tilde{\boldsymbol{q}}] - \mu)\, \boldsymbol{\Sigma}^{-1} (\mathbb{E}_P[\tilde{\boldsymbol{q}}] - \mu) \leq \gamma_1, \ \mathbb{E}_P[(\tilde{\boldsymbol{q}} - \mu)(\tilde{\boldsymbol{q}} - \mu)' \preceq \gamma_2 \boldsymbol{\Sigma} \right\},$$

where the constants $\gamma_1, \gamma_2 \geq 0$, $\boldsymbol{\mu} \in \boldsymbol{int}(\mathcal{S})$, $\boldsymbol{\Sigma} \succ \boldsymbol{0}$ and support $\mathcal{S}$ is a convex set for which there exists an oracle that can confirm feasibility or provide a separating hyperplane in polynomial time.

Our risk-averse two-stage stochastic linear optimization problem with objective uncertainty satisfies assumptions (i) and (ii). Furthermore, for $\mathcal{S} = \Re^p$, $\gamma_1 = 0$, $\gamma_2 = 1$ and $\boldsymbol{\Sigma} = \boldsymbol{Q} + \boldsymbol{\mu}\boldsymbol{\mu}'$, the distribution class $\mathcal{P}$ is a subset of $\hat{\mathcal{P}}$:

$$\mathcal{P} \subseteq \hat{\mathcal{P}} = \left\{ P \ : \ \mathbb{P}[\tilde{\boldsymbol{q}} \in \Re^p] = 1, \ \mathbb{E}_P[\tilde{\boldsymbol{q}}] = \boldsymbol{\mu}, \ \mathbb{E}_P[\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}'] \preceq \boldsymbol{Q} \right\}.$$

In contrast to using an ellipsoidal algorithm, our formulation in Theorem 1 is based on solving a compact semidefinite optimization problem. In the next section, we generate the extremal distribution for the second-stage problem and show the connection of our results with the Jensen bound.

## 2.2 Extremal Distribution

Taking the dual of the semidefinite optimization problem in (7), we obtain:

$$\begin{array}{lll}
Z_{DD}(\boldsymbol{x}) = \max_{\boldsymbol{V}_k, \boldsymbol{v}_k, v_{k0}, \boldsymbol{p}_k} & \displaystyle\sum_{k=1}^{K} (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_k + \beta_k v_{k0} & \\
\text{s.t.} & \displaystyle\sum_{k=1}^{K} \begin{pmatrix} \boldsymbol{V}_k & \boldsymbol{v}_k \\ \boldsymbol{v}_k' & v_{k0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q} & \boldsymbol{\mu} \\ \boldsymbol{\mu}' & 1 \end{pmatrix}, & \\
& \begin{pmatrix} \boldsymbol{V}_k & \boldsymbol{v}_k \\ \boldsymbol{v}_k' & v_{k0} \end{pmatrix} \succeq 0, & \forall\, k = 1, \ldots, K, \\
& \boldsymbol{W}'\boldsymbol{p}_k \leq \alpha_k \boldsymbol{v}_k, & \forall\, k = 1, \ldots, K.
\end{array} \tag{8}$$

The interpretation of these dual variables as a set of (scaled) conditional moments allows us to construct extremal distributions that attain the second-stage optimal value $Z(\boldsymbol{x})$.

**Theorem 2** *For an arbitrary $\boldsymbol{x} \in X$, there exists a sequence of distributions in $\mathcal{P}$ that asymptotically achieves the optimal value $Z(\boldsymbol{x}) = Z_D(\boldsymbol{x}) = Z_{DD}(\boldsymbol{x})$.*

**Proof.** Using weak duality for semidefinite optimization problems, we have $Z_{DD}(\boldsymbol{x}) \leq Z_D(\boldsymbol{x})$. We first argue that $Z_{DD}(\boldsymbol{x})$ is also an upper bound of $Z(\boldsymbol{x}) = \max_{P \in \mathcal{P}} \mathbb{E}_P\left[\mathbb{U}\left(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x})\right)\right]$.

Since $\mathcal{Q}(\boldsymbol{q}, \boldsymbol{x})$ is a linear optimization problem; for each objective vector $\tilde{\boldsymbol{q}}$, we define the primal and dual optimal solutions as $\boldsymbol{w}(\tilde{\boldsymbol{q}})$ and $\boldsymbol{p}(\tilde{\boldsymbol{q}})$. For an arbitrary distribution $P \in \mathcal{P}$, we define

$$\begin{array}{rcl} v_{k0} & = & \mathbb{P}\left(k \in \arg\max_l \left(\alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l\right)\right), \\ \boldsymbol{v}_k & = & v_{k0}\mathbb{E}_P\left[\tilde{\boldsymbol{q}} \,\middle|\, k \in \arg\max_l \left(\alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l\right)\right], \\ \boldsymbol{V}_k & = & v_{k0}\mathbb{E}_P\left[\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}' \,\middle|\, k \in \arg\max_l \left(\alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l\right)\right], \\ \boldsymbol{p}_k & = & \alpha_k v_{k0}\mathbb{E}_P\left[\boldsymbol{p}(\tilde{\boldsymbol{q}}) \,\middle|\, k \in \arg\max_l \left(\alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l\right)\right]. \end{array}$$

From the definition of the variables, we have:

$$\sum_{k=1}^{K} \begin{pmatrix} \boldsymbol{V}_k & \boldsymbol{v}_k \\ \boldsymbol{v}_k' & v_{k0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q} & \boldsymbol{\mu} \\ \boldsymbol{\mu}' & 1 \end{pmatrix},$$

with the moment feasibility conditions given as

$$\begin{pmatrix} \boldsymbol{V}_k & \boldsymbol{v}_k \\ \boldsymbol{v}_k' & v_{k0} \end{pmatrix} \succeq 0, \quad \forall\, k = 1, \ldots, K.$$

For ease of exposition, we implicitly assume that the set of $\tilde{\boldsymbol{q}}$, such that $\arg\max_l \alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l$ has multiple optimal solutions has a support with measure zero. For distributions with multiple optimal solutions, one can perturb the values of the variables to satisfy the moment equality constraints while maintaining the same objective. This arises due to the continuity of the objective function at break-points.

Since $\boldsymbol{p}(\tilde{\boldsymbol{q}})$ is the dual optimal solution of the second-stage linear optimization problem, from dual feasibility we have $\boldsymbol{W}'\boldsymbol{p}(\tilde{\boldsymbol{q}}) \leq \tilde{\boldsymbol{q}}$. Taking expectations and multiplying by $\alpha_k v_{k0}$, we obtain the inequality

$$\alpha_k v_{k0} \boldsymbol{W}' \mathbb{E}_P\left[\boldsymbol{p}(\tilde{\boldsymbol{q}}) \,\middle|\, k \in \arg\max_l \left(\alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l\right)\right] \leq \alpha_k v_{k0} \mathbb{E}_P\left[\tilde{\boldsymbol{q}} \,\middle|\, k \in \arg\max_l \left(\alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l\right)\right],$$

or $\boldsymbol{W}'\boldsymbol{p}_k \leq \alpha_k \boldsymbol{v}_k$. Thus all the constraints are satisfied and we have a feasible solution of the semidefinite optimization problem defined in (8). The objective function is expressed as

$$\mathbb{E}_P\left[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}))\right] = \sum_{k=1}^{K} \mathbb{E}_P\left[\left(\alpha_k \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_k\right) \,\middle|\, k \in \arg\max_l \left(\alpha_l \mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) + \beta_l\right)\right] v_{k0},$$

or equivalently

$$\mathbb{E}_P\left[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}))\right] = \sum_{k=1}^{K} (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_k + \beta_k v_{k0},$$

since $\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) = (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}(\tilde{\boldsymbol{q}})$. This implies that $\mathbb{E}_P\left[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}))\right] \leq Z_{DD}(\boldsymbol{x})$ for all $P \in \mathcal{P}$. Thus

$$Z(\boldsymbol{x}) = \max_{P \in \mathcal{P}} \mathbb{E}_P\left[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}))\right] \leq Z_{DD}(\boldsymbol{x}).$$

We now construct a sequence of extremal distributions that attains the bound asymptotically. Consider the dual problem defined in (8) and its optimal solution $(\boldsymbol{V}_k, \boldsymbol{v}_k, v_{k0}, \boldsymbol{p}_k)_k$. We start by assuming that $v_{k0} > 0$ for all $k = 1, \ldots, K$ (note $v_{k0} \geq 0$ due to feasibility). This assumption can be relaxed as we will see later. Consider the following random vectors:

$$\tilde{\boldsymbol{q}}_k := \frac{\boldsymbol{v}_k}{v_{k0}} + \frac{\tilde{b}_k \tilde{\boldsymbol{r}}_k}{\sqrt{\epsilon}}, \quad \forall\, k = 1, \ldots, K,$$

where $\tilde{b}_k$ is a Bernoulli random variable with distribution

$$\tilde{b}_k = \begin{cases} 0, & \text{with probability } 1-\epsilon, \\ 1, & \text{with probability } \epsilon, \end{cases}$$

and $\tilde{\boldsymbol{r}}_k$ is a multivariate normal random vector, independent of $\tilde{b}_k$ with mean and covariance matrix

$$\tilde{r}_k = \mathbb{N}\left(\boldsymbol{0}, \frac{\boldsymbol{V}_k v_{k0} - \boldsymbol{v}_k \boldsymbol{v}_k'}{v_{k0}^2}\right).$$

We construct the mixed distribution $P_m(\boldsymbol{x})$ of $\tilde{\boldsymbol{q}}$ as:

$$\tilde{\boldsymbol{q}} := \tilde{\boldsymbol{q}}_k \text{ with probability } v_{k0},\ \forall\, k = 1, \ldots, K.$$

Under this mixed distribution, we have:

$$\mathbb{E}_{P_m(\boldsymbol{x})}\left[\tilde{\boldsymbol{q}}_k\right] = \frac{\boldsymbol{v}_k}{v_{k0}} + \frac{\mathbb{E}_{P_m}[\tilde{b}_k]\mathbb{E}_{P_m}\left[\tilde{\boldsymbol{r}}_k\right]}{\sqrt{\epsilon}} = \frac{\boldsymbol{v}_k}{v_{k0}},$$

and

$$\mathbb{E}_{P_m(\boldsymbol{x})}\left[\tilde{\boldsymbol{q}}_k \tilde{\boldsymbol{q}}_k'\right] = \frac{\boldsymbol{v}_k \boldsymbol{v}_k'}{v_{k0}^2} + 2\frac{\boldsymbol{v}_k \mathbb{E}_{P_m(\boldsymbol{x})}[\tilde{b}_k]\mathbb{E}_{P_m(\boldsymbol{x})}\left[\tilde{\boldsymbol{r}}_k'\right]}{v_{k0}\sqrt{\epsilon}} + \frac{\mathbb{E}_{P_m(\boldsymbol{x})}[\tilde{b}_k^2]\mathbb{E}_{P_m(\boldsymbol{x})}\left[\tilde{\boldsymbol{r}}_k \tilde{\boldsymbol{r}}_k'\right]}{\epsilon} = \frac{\boldsymbol{V}_k}{v_{k0}}.$$

Thus $\mathbb{E}_{P_m(\boldsymbol{x})}[\tilde{\boldsymbol{q}}] = \boldsymbol{\mu}$ and $\mathbb{E}_{P_m(\boldsymbol{x})}[\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}'] = \boldsymbol{Q}$ from the feasibility conditions.

Considering the expected value $Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) = \mathbb{E}_{P_m(\boldsymbol{x})}\left[\mathbb{U}\left(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x})\right)\right]$, we have:

$$Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) \geq \sum_{k=1}^{K} v_{k0} \mathbb{E}_{P_m(\boldsymbol{x})}\left[\alpha_k \mathcal{Q}(\tilde{\boldsymbol{q}}_k, \boldsymbol{x}) + \beta_k\right].$$

Conditioning based on the value of $\tilde{b}_k$, the inequality can be rewritten as follows:

$$Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) \geq \sum_{k=1}^{K}\left(v_{k0}\epsilon \mathbb{E}_{P_m(\boldsymbol{x})}\left[\alpha_k \mathcal{Q}\left(\frac{\boldsymbol{v}_k}{v_{k0}} + \frac{\tilde{\boldsymbol{r}}_k}{\sqrt{\epsilon}}, \boldsymbol{x}\right) + \beta_k\right] + v_{k0}(1-\epsilon)\mathbb{E}_{P_m(\boldsymbol{x})}\left[\alpha_k \mathcal{Q}\left(\frac{\boldsymbol{v}_k}{v_{k0}}, \boldsymbol{x}\right) + \beta_k\right]\right).$$

Since $\mathcal{Q}(\boldsymbol{q}, \boldsymbol{x})$ is a minimization linear optimization problem with the objective coefficient vector $\boldsymbol{q}$; therefore, $\mathcal{Q}(t\boldsymbol{q}, \boldsymbol{x}) = t\mathcal{Q}(\boldsymbol{q}, \boldsymbol{x})$ for all $t > 0$ and

$$\mathcal{Q}\left(\frac{\boldsymbol{v}_k}{v_{k0}} + \frac{\boldsymbol{r}_k}{\sqrt{\epsilon}}, \boldsymbol{x}\right) \geq \mathcal{Q}\left(\frac{\boldsymbol{v}_k}{v_{k0}}, \boldsymbol{x}\right) + \mathcal{Q}\left(\frac{\boldsymbol{r}_k}{\sqrt{\epsilon}}, \boldsymbol{x}\right).$$

In addition, $\alpha_k \geq 0$ and $v_{k0} > 0$ imply

$$Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) \geq \sum_{k=1}^{K} v_{k0} \mathbb{E}_{P_m(\boldsymbol{x})} \left[ \alpha_k \mathcal{Q}\left( \frac{\boldsymbol{v}_k}{v_{k0}}, \boldsymbol{x} \right) + \beta_k \right] + \sqrt{\epsilon} \sum_{k=1}^{K} v_{k0} \alpha_k \mathbb{E}_{P_m(\boldsymbol{x})} \left[ \mathcal{Q}\left( \tilde{\boldsymbol{r}}_k, \boldsymbol{x} \right) \right],$$

or

$$Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) \geq \sum_{k=1}^{K} \left( \alpha_k \mathcal{Q}\left( \boldsymbol{v}_k, \boldsymbol{x} \right) + v_{k0}\beta_k \right) + \sqrt{\epsilon} \sum_{k=1}^{K} v_{k0} \alpha_k \mathbb{E}_{P_m(\boldsymbol{x})} \left[ \mathcal{Q}\left( \tilde{\boldsymbol{r}}_k, \boldsymbol{x} \right) \right].$$

Since $\boldsymbol{p}_k$ is a dual feasible solution to the problem $\mathcal{Q}\left(\alpha_k \boldsymbol{v}_k, \boldsymbol{x}\right)$, thus $\alpha_k \mathcal{Q}\left(\boldsymbol{v}_k, \boldsymbol{x}\right) \geq (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_k$. From Jensen's inequality, we obtain $\mathbb{E}_{P_m(\boldsymbol{x})}\left[\mathcal{Q}\left(\tilde{\boldsymbol{r}}_k, \boldsymbol{x}\right)\right] \leq \mathcal{Q}\left(\mathbb{E}_{P_m(\boldsymbol{x})}\left[\tilde{\boldsymbol{r}}_k\right], \boldsymbol{x}\right) = 0$. In addition, Assumptions (A1)-(A4) implies that $\mathbb{E}_{P_m(\boldsymbol{x})}\left[\mathcal{Q}\left(\tilde{\boldsymbol{r}}_k, \boldsymbol{x}\right)\right] > -\infty$. Therefore,

$$-\infty < \sum_{k=1}^{K} \left( (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_k + \beta_k v_{k0} \right) + \sqrt{\epsilon} \sum_{k=1}^{K} v_{k0} \alpha_k \mathbb{E}_{P_m(\boldsymbol{x})} \left[ \mathcal{Q}\left( \tilde{\boldsymbol{r}}_k, \boldsymbol{x} \right) \right] \leq Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) \leq Z(\boldsymbol{x}).$$

We then have:

$$-\infty < Z_{DD}(\boldsymbol{x}) + \sqrt{\epsilon} \sum_{k=1}^{K} v_{k0} \alpha_k \mathbb{E}_{P_m(\boldsymbol{x})} \left[ \mathcal{Q}\left( \tilde{\boldsymbol{r}}_k, \boldsymbol{x} \right) \right] \leq Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) \leq Z(\boldsymbol{x}) = Z_D(\boldsymbol{x}) \leq Z_{DD}(\boldsymbol{x}).$$

Taking limit as $\epsilon \downarrow 0$, we have $\lim_{\epsilon \downarrow 0} Z_{P_m(\boldsymbol{x})}(\boldsymbol{x}) = Z(\boldsymbol{x}) = Z_D(\boldsymbol{x}) = Z_{DD}(\boldsymbol{x})$.
Consider the case where there exists a nonempty set $L \subset \{1, \ldots, K\}$ such that $v_{k0} = 0$ for all $k \in L$. Due to feasibility of a positive semidefinite matrix, we have $\boldsymbol{v}_k = \boldsymbol{0}$ for all $k \in L$ (note that $|A_{ij}| \leq \sqrt{A_{ii}A_{jj}}$ if $\boldsymbol{A} \succeq 0$), which means $\mathcal{Q}(\boldsymbol{v}_k, \boldsymbol{x}) = 0$. We claim that there is an optimal solution of the dual problem formulated in (8) such that

$$\sum_{k \notin L} \begin{pmatrix} \boldsymbol{V}_k & \boldsymbol{v}_k \\ \boldsymbol{v}_k' & v_{k0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q} & \boldsymbol{\mu} \\ \boldsymbol{\mu}' & 1 \end{pmatrix}.$$

Indeed, if $\boldsymbol{V}_L = \sum_{k \in L} \boldsymbol{V}_k \neq \boldsymbol{0}$, construct another optimal solution with $\boldsymbol{V}_k := \boldsymbol{0}$ for all $k \in L$ and $\boldsymbol{V}_k := \boldsymbol{V}_k + \boldsymbol{V}_L/(K - |L|)$ for all $k \notin L$. All feasibility constraints are still satisfied as $\boldsymbol{V}_L \succeq 0$ and $\boldsymbol{v}_k = \boldsymbol{0}$, $v_{k0} = 0$ for all $k \in L$. The objective value remains the same. Thus we obtain an optimal solution that satisfies the above condition. Since $(\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_k + v_{k0}\beta_k = 0$ for all $k \in L$; therefore, we can then construct the sequence of extremal distributions as in the previous case. $\square$

In the risk-neutral setting with $\mathbb{U}(x) = x$, the dual problem (8) has trivial solution $\boldsymbol{V}_1 = \boldsymbol{Q}$, $\boldsymbol{v}_1 = \boldsymbol{\mu}$, and $v_{10} = 1$. The second-stage bound then simplifies to

$$\max_{\boldsymbol{p}: \boldsymbol{W}'\boldsymbol{p} \leq \boldsymbol{\mu}} (\boldsymbol{h} - \boldsymbol{T}\boldsymbol{x})' \boldsymbol{p},$$

or equivalently

$$\min_{\boldsymbol{w} \in X(x)} \boldsymbol{\mu}' \boldsymbol{w}.$$

The second-stage bound thus just reduces to Jensen's bound where the uncertain objective $\tilde{\boldsymbol{q}}$ is replaced its mean $\boldsymbol{\mu}$. For the risk-averse case with $K > 1$, the second-stage cost is no longer concave but quasi-concave in $\tilde{\boldsymbol{q}}$. The second-stage bound then reduces to a convex combination of Jensen bounds for appropriately choosen means and probabilities:

$$\begin{array}{rll} \max_{\boldsymbol{V}_k, \boldsymbol{v}_k, v_{k0}} & \displaystyle\sum_{k=1}^{K} \left( \alpha_k \min_{\boldsymbol{w}_k \in X(x)} \boldsymbol{v}_k' \boldsymbol{w}_k + \beta_k v_{k0} \right) & \\ \text{s.t.} & \displaystyle\sum_{k=1}^{K} \begin{pmatrix} \boldsymbol{V}_k & \boldsymbol{v}_k \\ \boldsymbol{v}_k' & v_{k0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q} & \boldsymbol{\mu} \\ \boldsymbol{\mu}' & 1 \end{pmatrix}, & \\ & \begin{pmatrix} \boldsymbol{V}_k & \boldsymbol{v}_k \\ \boldsymbol{v}_k' & v_{k0} \end{pmatrix} \succeq 0, & \forall\, k = 1, \ldots, K. \end{array}$$

The variable $\boldsymbol{w}_k$ can be then interpreted as the optimal second-stage solution in the extremal distribution at which the $k$th piece of the utility function attains the maximum.

## 3 Uncertainty in Right-Hand Side

Consider the minimax stochastic problem (1) with random right-hand side $\tilde{\boldsymbol{h}}$ and constant objective $\boldsymbol{q}$. The distribution class $\mathcal{P}$ is specified by the first and second moments:

$$\mathcal{P} = \left\{ P \ : \ \mathbb{P}[\tilde{\boldsymbol{h}} \in \Re^r] = 1, \ \mathbb{E}_P[\tilde{\boldsymbol{h}}] = \boldsymbol{\mu}, \ \mathbb{E}_P[\tilde{\boldsymbol{h}} \tilde{\boldsymbol{h}}'] = \boldsymbol{Q} \right\}. \tag{9}$$

The second-stage cost with risk aversion and right-hand side uncertainty is then given as

$$\mathbb{U}\left( \mathcal{Q}(\tilde{\boldsymbol{h}}, \boldsymbol{x}) \right) := \max_{k=1,\ldots,K} \left( \alpha_k \mathcal{Q}(\tilde{\boldsymbol{h}}, \boldsymbol{x}) + \beta_k \right),$$

where

$$\begin{array}{rll} \mathcal{Q}(\tilde{\boldsymbol{h}}, \boldsymbol{x}) & = \max_{\boldsymbol{p}} & (\tilde{\boldsymbol{h}} - \boldsymbol{T}\boldsymbol{x})' \boldsymbol{p} \\ & \text{s.t.} & \boldsymbol{W}' \boldsymbol{p} \le \boldsymbol{q}. \end{array}$$

In this case, the second-stage cost $\mathbb{U}\left( \mathcal{Q}(\tilde{\boldsymbol{h}}, \boldsymbol{x}) \right)$ is a convex function in $\tilde{\boldsymbol{h}}$ and $\boldsymbol{x}$. We prove the NP-hardness of the general problem in Section 3.1, while proposing a semidefinite optimization formulation for a special class of problems in Section 3.2.

## 3.1 Complexity of the General Problem

The second-stage problem $\max_{P\in\mathcal{P}} \mathbb{E}_P\left[\mathbb{U}\left(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x})\right)\right]$ of the risk-averse minimax stochastic optimization problem is an infinite-dimensional linear optimization problem with the probability distribution $P$ or its corresponding probability density function $f$ as the problem variable:

$$\begin{array}{rlr} \hat{Z}(\boldsymbol{x}) = \max_f & \int_{\mathbb{R}^r} \mathbb{U}\left(\mathcal{Q}(\boldsymbol{h},\boldsymbol{x})\right) f(\boldsymbol{h})\mathrm{d}\boldsymbol{h} & \\ \text{s.t.} & \int_{\mathbb{R}^p} h_i h_j f(\boldsymbol{h})\mathrm{d}\boldsymbol{q} = Q_{ij}, & \forall\, i,j = 1,\dots,r, \\ & \int_{\mathbb{R}^p} h_i f(\boldsymbol{h})\mathrm{d}\boldsymbol{q} = \mu_i, & \forall\, i = 1,\dots,r, \\ & \int_{\mathbb{R}^p} f(\boldsymbol{h})\mathrm{d}\boldsymbol{q} = 1, & \\ & f(\boldsymbol{h}) \geq 0, & \forall\, \boldsymbol{h} \in \mathbb{R}^r. \end{array} \tag{10}$$

Under the strong duality condition, the equivalent dual problem is:

$$\begin{array}{rl} \hat{Z}_D(\boldsymbol{x}) = \min_{\boldsymbol{Y},\boldsymbol{y},y_0} & \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\ \text{s.t.} & \boldsymbol{h}'\boldsymbol{Y}\boldsymbol{h} + \boldsymbol{y}'\boldsymbol{h} + y_0 \geq \mathbb{U}\left(\mathcal{Q}(\boldsymbol{h},\boldsymbol{x})\right), \quad \forall\, \boldsymbol{h}\in\mathbb{R}^r. \end{array} \tag{11}$$

The minimax stochastic problem is equivalent to the following problem:

$$\min_{\boldsymbol{x}\in X}\left(\boldsymbol{c}'\boldsymbol{x} + \hat{Z}_D(\boldsymbol{x})\right).$$

The constraints of the dual problem defined in (11) can be rewritten as follows:

$$\boldsymbol{h}'\boldsymbol{Y}\boldsymbol{h} + \boldsymbol{y}'\boldsymbol{h} + y_0 \geq \alpha_k\mathcal{Q}(\boldsymbol{h},\boldsymbol{x}) + \beta_k \quad \forall\, \boldsymbol{h}\in\mathbb{R}^r,\ k = 1,\dots,K.$$

As $\alpha_k \geq 0$, these constraints are equivalent to:

$$\boldsymbol{h}'\boldsymbol{Y}\boldsymbol{h} + (\boldsymbol{y}-\alpha_k\boldsymbol{p})'\boldsymbol{h} + y_0 + \alpha_k\boldsymbol{p}'\boldsymbol{T}\boldsymbol{x} - \beta_k \geq 0, \quad \forall\, \boldsymbol{h}\in\mathbb{R}^r, \boldsymbol{p} : \boldsymbol{W}'\boldsymbol{p} \leq \boldsymbol{q},\ k = 1,\dots,K.$$

The dual matrix $\boldsymbol{Y} \succeq \boldsymbol{0}$, else using an argument in Theorem 1, we can scale $\boldsymbol{h}$ and find a violated constraint. Converting to a minimization problem, the dual feasibility constraints can be expressed as:

$$\min_{\boldsymbol{h},\boldsymbol{p}:\boldsymbol{W}'\boldsymbol{p}\leq\boldsymbol{q}} \quad \boldsymbol{h}'\boldsymbol{Y}\boldsymbol{h} + (\boldsymbol{y}-\alpha_k\boldsymbol{p})'\boldsymbol{h} + y_0 + \alpha_k\boldsymbol{p}'\boldsymbol{T}\boldsymbol{x} - \beta_k \quad \geq \quad 0,\ \forall\, k = 1,\dots,K. \tag{12}$$

We will show that the separation version of this problem is NP-hard.

> Separation problem ($\mathcal{S}$): Given $\{\alpha_k,\beta_k\}_k$, $\boldsymbol{T}\boldsymbol{x}$, $\boldsymbol{W}$, $\boldsymbol{q}$, $\boldsymbol{Y}\succeq\boldsymbol{0}$, $\boldsymbol{y}$ and $y_0$, check if the dual feasibility constraints in (12) are satisfied? If not, find a $k\in\{1,\dots,K\}$, $\boldsymbol{h}\in\mathbb{R}^r$ and $\boldsymbol{p}$ satisfying $\boldsymbol{W}'\boldsymbol{p}\leq\boldsymbol{q}$ such that:
>
> $$\boldsymbol{h}'\boldsymbol{Y}\boldsymbol{h} + (\boldsymbol{y}-\alpha_k\boldsymbol{p})'\boldsymbol{h} + y_0 + \alpha_k\boldsymbol{p}'\boldsymbol{T}\boldsymbol{x} - \beta_k < 0.$$

The equivalence of separation and optimization (see Grötschel [11]) then implies that dual feasibility problem and the minimax stochastic optimization problem are NP-hard.

**Theorem 3** *The risk-averse minimax stochastic optimization problem (1) with random right-hand side* $\tilde{\boldsymbol{h}}$ *and constant objective* $\boldsymbol{q}$ *is NP-hard.*

**Proof.** We provide a reduction from the decision version of the 2-norm maximization problem over a bounded polyhedral set:

> $(\mathcal{S}_1)$: Given $\boldsymbol{A}$, $\boldsymbol{b}$ with rational entries and a nonzero rational number $s$, is there a vector $\boldsymbol{p} \in \mathbb{R}^r$ such that:
> $$\boldsymbol{A}\boldsymbol{p} \leq \boldsymbol{b}, \ \sqrt{\boldsymbol{p}'\boldsymbol{p}} \geq s \ ?$$

The 2-norm maximization problem and its related decision problem $(\mathcal{S}_1)$ are shown to be NP-complete in Mangasarian and Shiau [16]. Define the parameters of $(\mathcal{S})$ as

$$K := 1, \ \beta_K := -s^2/4, \ \boldsymbol{W}' := \boldsymbol{A} \text{ and } \boldsymbol{q} := \boldsymbol{b},$$

$$\boldsymbol{Y} := \boldsymbol{I}, \ \boldsymbol{y} := \boldsymbol{0}, \ y_0 := 0, \ \alpha_K := 1 \text{ and } \boldsymbol{T}\boldsymbol{x} := \boldsymbol{0},$$

where $\boldsymbol{I}$ is the identity matrix. The problem $(\mathcal{S}_1)$ can then be answered by the question:

$$\text{Is } \max_{\boldsymbol{p}:\boldsymbol{W}'\boldsymbol{p}\leq\boldsymbol{q}} \boldsymbol{p}'\boldsymbol{p} \geq -4\beta_K \ ?$$

or equivalently:

$$\text{Is } \min_{\boldsymbol{h},\boldsymbol{p}:\boldsymbol{W}'\boldsymbol{p}\leq\boldsymbol{q}} \boldsymbol{h}'\boldsymbol{h} - \boldsymbol{p}'\boldsymbol{h} - \beta_K \leq 0 \ ?$$

since the optimal value of $\boldsymbol{h}$ is $\boldsymbol{p}/2$. Thus $(\mathcal{S}_1)$ reduces to an instance of $(\mathcal{S})$:

$$\text{Is } \min_{\boldsymbol{h},\boldsymbol{p}:\boldsymbol{W}'\boldsymbol{p}\leq\boldsymbol{q}} \boldsymbol{h}'\boldsymbol{Y}\boldsymbol{h} + (\boldsymbol{y} - \alpha_K\boldsymbol{p})'\boldsymbol{h} + y_0 + \alpha_K\boldsymbol{p}'\boldsymbol{T}\boldsymbol{x} - \beta_K \leq 0 \ ?$$

Since $(\mathcal{S}_1)$ is NP-complete, $(\mathcal{S})$ and the corresponding minimax stochastic optimization problem are NP-hard. $\square$

### 3.2 Explicitly Known Dual Extreme Points

The NP-hardness result in the previous section is due to the non-convexity of the objective function in (12) in the joint decision variables $(\boldsymbol{h}, \boldsymbol{p})$. In this section, we consider the case where the extreme points of the dual problem of the second-stage linear optimization problem are known and bounded by a function which is polynomial in the dimension of the problem. We make the following assumption:

(A5) The $N$ extreme points $\{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N\}$ of the dual feasible region $\{\boldsymbol{p} \in \mathbb{R}^r : \boldsymbol{W}'\boldsymbol{p} \leq \boldsymbol{q}\}$ are explicitly known.

We will provide the semidefinite optimization reformulation of our minimax stochastic optimization problem in the following theorem:

**Theorem 4** *Under the additional Assumption (A5), the risk-averse minimax stochastic optimization problem (1) with random right-hand side $\tilde{\boldsymbol{h}}$ and constant objective $\boldsymbol{q}$ is equivalent to the following semidefinite optimization problem:*

$$\begin{aligned}
\hat{Z}_{SDP} = \min_{\boldsymbol{x},\boldsymbol{Y},\boldsymbol{y},y_0} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\
\text{s.t.} \quad & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k\boldsymbol{p}_i}{2} \\ \frac{(\boldsymbol{y}-\alpha_k\boldsymbol{p}_i)'}{2} & y_0 + \alpha_k\boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x} - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, k = 1,\ldots,K,\ i = 1,\ldots,N, \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}.
\end{aligned} \tag{13}$$

**Proof.** Under Assumption (A5), we have:

$$\mathcal{Q}(\tilde{\boldsymbol{h}}, \boldsymbol{x}) = \max_{i=1,\ldots,N} (\tilde{\boldsymbol{h}} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_i.$$

The dual constraints can be explicitly written as follows:

$$\boldsymbol{h}'\boldsymbol{Y}\boldsymbol{h} + (\boldsymbol{y} - \alpha_k\boldsymbol{p}_i)'\boldsymbol{h} + y_0 + \alpha_k\boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x} - \beta_k \geq 0, \quad \forall\, \boldsymbol{h} \in \mathbb{R}^r,\ k = 1,\ldots,K,\ i = 1,\ldots,N.$$

These constraints can be formulated as the linear matrix inequalities:

$$\begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k\boldsymbol{p}_i}{2} \\ \frac{(\boldsymbol{y}-\alpha_k\boldsymbol{p}_i)'}{2} & y_0 + \alpha_k\boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x} - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, k = 1,\ldots,K,\ i = 1,\ldots,N.$$

Thus the dual problem of the second-stage optimization problem is rewritten as follows:

$$\begin{aligned}
\hat{Z}_D(\boldsymbol{x}) = \min_{\boldsymbol{Y},\boldsymbol{y},y_0} \quad & \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\
\text{s.t.} \quad & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k\boldsymbol{p}_i}{2} \\ \frac{(\boldsymbol{y}-\alpha_k\boldsymbol{p}_i)'}{2} & y_0 + \alpha_k\boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x} - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, k = 1,\ldots,K,\ i = 1,\ldots,N,
\end{aligned} \tag{14}$$

which provides the semidefinite formulation for the risk-averse minimax stochastic optimization problem.

$$
\begin{array}{rl}
\hat{Z}_{SDP} = \min_{\boldsymbol{x},\boldsymbol{Y},\boldsymbol{y},y_0} & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{Q}\cdot\boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\
\text{s.t.} & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y}-\alpha_k\boldsymbol{p}_i}{2} \\ \frac{(\boldsymbol{y}-\alpha_k\boldsymbol{p}_i)'}{2} & y_0 + \alpha_k\boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x} - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, k = 1,\ldots,K,\ i = 1,\ldots,N, \\
& \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}.
\end{array}
$$

From the strong duality assumption, we have $\hat{Z}_D(\boldsymbol{x}) = \hat{Z}(\boldsymbol{x})$. Thus $Z = \hat{Z}_{SDP}$ or (13) is the equivalent semidefinite optimization formulation of the minimax stochastic optimization problem (1) with random right-hand side $\tilde{\boldsymbol{h}}$ and constant objective $\boldsymbol{q}$. □

To construct the extremal distribution, we again take dual of the problem defined in (14)

$$
\begin{array}{rll}
\hat{Z}_{DD}(\boldsymbol{x}) = \max & \displaystyle\sum_{k=1}^{K}\sum_{i=1}^{N} \left(\alpha_k\boldsymbol{p}_i'\boldsymbol{v}_k^i + v_{k0}^i(\beta_k - \alpha_k\boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x})\right) & \\
\text{s.t.} & \displaystyle\sum_{k=1}^{K}\sum_{i=1}^{N} \begin{pmatrix} \boldsymbol{V}_k^i & \boldsymbol{v}_k^i \\ (\boldsymbol{v}_k^i)' & v_{k0}^i \end{pmatrix} = \begin{pmatrix} \boldsymbol{Q} & \boldsymbol{\mu} \\ \boldsymbol{\mu}' & 1 \end{pmatrix}, & \\
& \begin{pmatrix} \boldsymbol{V}_k^i & \boldsymbol{v}_k^i \\ (\boldsymbol{v}_k^i)' & v_{k0}^i \end{pmatrix} \succeq 0, & \forall\, k = 1,\ldots,K,\ i = 1,\ldots,N.
\end{array} \tag{15}
$$

We construct an extremal distribution for the second-stage problem using the following theorem:

**Theorem 5** *For an arbitrary $\boldsymbol{x} \in X$, there exists an extremal distribution in $\mathcal{P}$ that achieves the optimal value $\hat{Z}(\boldsymbol{x})$.*

**Proof.** As $\alpha_k \geq 0$, we have

$$
\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x})) = \max_{k=1,\ldots,K}\left(\alpha_k\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}) + \beta_k\right) = \max_{k,i}\left(\alpha_k(\boldsymbol{h}-\boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_i + \beta_k\right).
$$

Using weak duality for semidefinite optimization problems, we have $\hat{Z}_{DD}(\boldsymbol{x}) \leq \hat{Z}_D(\boldsymbol{x})$. We show next that $\hat{Z}_{DD}(\boldsymbol{x})$ is an upper bound of $\hat{Z}(\boldsymbol{x}) = \max_{P\in\mathcal{P}}\mathbb{E}_P[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))]$. For any distribution $P \in \mathcal{P}$, we define:

$$
\begin{aligned}
v_{k0}^i &= \mathbb{P}\left((k,i) \in \arg\max_{l,j}\left(\alpha_l(\tilde{\boldsymbol{h}}-\boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_j + \beta_l\right)\right), \\
\boldsymbol{v}_k^i &= v_{k0}^i\mathbb{E}_P\left[\tilde{\boldsymbol{h}}\,\middle|\,(k,i) \in \arg\max_{l,j}\left(\alpha_l(\tilde{\boldsymbol{h}}-\boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_j + \beta_l\right)\right], \\
\boldsymbol{V}_k^i &= v_{k0}^i\mathbb{E}_P\left[\tilde{\boldsymbol{h}}\tilde{\boldsymbol{h}}'\,\middle|\,(k,i) \in \arg\max_{l,j}\left(\alpha_l(\tilde{\boldsymbol{h}}-\boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_j + \beta_l\right)\right].
\end{aligned}
$$

The vector $(v_{k0}^i, \boldsymbol{v}_k^i, \boldsymbol{V}_k^i)_{k,i}$ is a feasible solution to the dual problem defined in (15) and the value $\mathbb{E}_P[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))]$ is calculated as follows.

$$
\mathbb{E}_P[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] = \sum_{k=1}^{K}\sum_{i=1}^{N} v_{k0}^i\mathbb{E}_P\left[\left(\alpha_k(\tilde{\boldsymbol{h}}-\boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_i + \beta_k\right)\,\middle|\,(k,i) \in \arg\max_{l,j}\left(\alpha_l(\tilde{\boldsymbol{h}}-\boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_j + \beta_l\right)\right],
$$

or

$$\mathbb{E}_P[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] = \sum_{k=1}^{K}\sum_{i=1}^{N}\left(\alpha_k \boldsymbol{p}_i'\boldsymbol{v}_k^i + v_{k0}^i(\beta_k - \alpha_k \boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x})\right).$$

Therefore, we have: $\mathbb{E}_P[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] \leq \hat{Z}_{DD}(\boldsymbol{x})$ for all $P \in \mathcal{P}$. Thus

$$\hat{Z}(\boldsymbol{x}) = \max_{P\in\mathcal{P}} \mathbb{E}_P[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] \leq \hat{Z}_{DD}(\boldsymbol{x}).$$

We now construct the extremal distribution that achieves the optimal value $\hat{Z}(\boldsymbol{x})$. Consider the optimal solution $(v_{k0}^i, \boldsymbol{v}_k^i, \boldsymbol{V}_k^i)_{k,i}$ of the dual problem defined in (15). Without loss of generality, we can again assume that $v_{k0}^i > 0$ for all $k = 1,\ldots,K$ and $i = 1,\ldots,N$ (see Theorem 2). We then construct $NK$ multivariate normal random vectors $\tilde{\boldsymbol{h}}_k^i$ with mean and covariance matrix:

$$\tilde{\boldsymbol{h}}_k^i := \mathbb{N}\left(\frac{\boldsymbol{v}_k^i}{v_{k0}^i}, \frac{\boldsymbol{V}_k^i v_{k0} - \boldsymbol{v}_k^i(\boldsymbol{v}_k^i)'}{v_{k0}^2}\right).$$

We construct a mixed distribution $P_m(\boldsymbol{x})$ of $\tilde{\boldsymbol{h}}$:

$$\tilde{\boldsymbol{h}} := \tilde{\boldsymbol{h}}_k^i \text{ with probability } v_{k0},\ \forall\, k = 1,\ldots,K,\ i = 1,\ldots,N.$$

Clearly, $\mathbb{E}_{P_m(\boldsymbol{x})}[\tilde{\boldsymbol{h}}] = \boldsymbol{\mu}$ and $\mathbb{E}_{P_m(\boldsymbol{x})}[\tilde{\boldsymbol{h}}\tilde{\boldsymbol{h}}'] = \boldsymbol{Q}$. Thus

$$\mathbb{E}_{P_m(\boldsymbol{x})}[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] = \sum_{k=1}^{K}\sum_{i=1}^{N} v_{k0}^i \mathbb{E}_{P_m(\boldsymbol{x})}\left[\max_{k',i'}\left(\alpha_{k'}(\tilde{\boldsymbol{h}} - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_{i'} + \beta_{k'}\right)\ \Big|\ \tilde{\boldsymbol{h}} = \tilde{\boldsymbol{h}}_k^i\right].$$

We then have:

$$\mathbb{E}_{P_m(\boldsymbol{x})}[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] \geq \sum_{k=1}^{K}\sum_{i=1}^{N} v_{k0}^i \mathbb{E}_{P_m(\boldsymbol{x})}\left[\alpha_k(\tilde{\boldsymbol{h}}_k^i - \boldsymbol{T}\boldsymbol{x})'\boldsymbol{p}_i + \beta_k\right].$$

By substituting the mean vectors, we obtain:

$$\mathbb{E}_{P_m(\boldsymbol{x})}[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] \geq \sum_{k=1}^{K}\sum_{i=1}^{N} v_{k0}^i \left[\alpha_k\left(\frac{\boldsymbol{v}_k^i}{v_{k0}^i} - \boldsymbol{T}\boldsymbol{x}\right)'\boldsymbol{p}_i + \beta_k\right].$$

Finally we have:

$$\mathbb{E}_{P_m(\boldsymbol{x})}[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] \geq \sum_{k=1}^{K}\sum_{i=1}^{N}\left(\alpha_k \boldsymbol{p}_i'\boldsymbol{v}_k^i + v_{k0}^i(\beta_k - \alpha_k \boldsymbol{p}_i'\boldsymbol{T}\boldsymbol{x})\right) = \hat{Z}_{DD}(\boldsymbol{x}).$$

Thus $\hat{Z}_{DD}(\boldsymbol{x}) \leq \mathbb{E}_{P_m(\boldsymbol{x})}[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] \leq \hat{Z}(\boldsymbol{x}) \leq \hat{Z}_{DD}(\boldsymbol{x})$ or $\mathbb{E}_{P_m(\boldsymbol{x})}[\mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{h}},\boldsymbol{x}))] = \hat{Z}(\boldsymbol{x}) = \hat{Z}_{DD}(\boldsymbol{x})$. It means the constructed distribution $P_m(\boldsymbol{x})$ is the extremal distribution. $\square$

# 4 Computational Results

To illustrate our approach, we consider two following problems: the production-transportation problem with random transportation costs and the single facility minimax distance problem with random customer locations. These two problems fit into the framework of two-stage stochastic linear optimization with random objective and random right-hand side respectively.

## 4.1 Production and Transportation Problem

Suppose there are $m$ facilities and $n$ customer locations. Assume that each facility has a normalized production capacity of 1. The production cost per unit at each facility $i$ is $c_i$. The demand from each customer location $j$ is $h_j$ and known beforehand. We assume that $\sum_j h_j < m$. The transportation cost between facility $i$ and customer location $j$ is $q_{ij}$. The goal is to minimize the total production and transportation cost while satisfying all the customer orders. If we define $x_i \geq 0$ to be the amount to be produced at facility $i$ and $w_{ij}$ to be the amount transported from $i$ to $j$, the deterministic production-transportation problem is formulated as follows

$$
\begin{array}{lll}
\min & \displaystyle\sum_{i=1}^{m} c_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij} w_{ij} & \\
\text{s.t.} & \displaystyle\sum_{i=1}^{m} w_{ij} = h_j, & \forall\, j, \\
& \displaystyle\sum_{j=1}^{n} w_{ij} = x_i, & \forall\, i, \\
& 0 \leq x_i \leq 1,\ w_{ij} \geq 0, & \forall\, i,\ j.
\end{array}
$$

The two-stage version of this problem is to make the production decisions now while the transportation decision will be made once the random costs $\tilde{q}_{ij}$ are realized. The minimax stochastic problem with risk aversion can then be formulated as follows

$$
\begin{array}{rll}
Z = \min & \left( \boldsymbol{c}'\boldsymbol{x} + \max\limits_{P \in \mathcal{P}} \mathbb{E}_P \left[ \mathbb{U}(\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x})) \right] \right) & \\
\text{s.t.} & 0 \leq x_i \leq 1, & \forall\, i,
\end{array}
\tag{16}
$$

where the second-stage cost $\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x})$ is given as

$$\begin{aligned}
\mathcal{Q}(\tilde{\boldsymbol{q}}, \boldsymbol{x}) = \min \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} \tilde{q}_{ij} w_{ij} \\
\text{s.t.} \quad & \sum_{i=1}^{m} w_{ij} = h_j, \quad \forall\, j, \\
& \sum_{j=1}^{n} w_{ij} = x_i, \quad \forall\, i, \\
& w_{ij} \geq 0, \qquad \forall\, i,\, j.
\end{aligned}$$

For transportation costs with known mean and second moment matrix, the risk-averse minimax stochastic optimization problem is solved as:

$$\begin{aligned}
Z_{SDP} = \min_{\boldsymbol{x}, \boldsymbol{Y}, \boldsymbol{y}, y_0, \boldsymbol{w}_k} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{Q} \cdot \boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\
\text{s.t.} \quad & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y} - \alpha_k \boldsymbol{w}_k}{2} \\ \frac{(\boldsymbol{y} - \alpha_k \boldsymbol{w}_k)'}{2} & y_0 - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, k = 1, \ldots, K, \\
& \sum_{i=1}^{m} w_{ijk} = h_j, \qquad \forall\, j,\, k, \\
& \sum_{j=1}^{n} w_{ijk} = x_i, \qquad \forall\, i,\, k, \\
& 0 \leq x_i \leq 1,\, w_{ijk} \geq 0, \qquad \forall\, i, j,\, k.
\end{aligned} \tag{17}$$

The code for this problem is developed using Matlab 7.4 with SeDuMi solver (see [23]) and YALMIP interface (Löfberg [15]).

An alternative approach using the data-driven or sample approach is to solve the linear optimization problem:

$$\begin{aligned}
Z_D = \min \quad & \sum_{i=1}^{n} c_i x_i + \frac{1}{N} \sum_{t=1}^{N} \mathbb{U}(\mathcal{Q}(\boldsymbol{q}_t, \boldsymbol{x})) \\
\text{s.t.} \quad & 0 \leq x_i \leq 1, \qquad \forall\, i,
\end{aligned} \tag{18}$$

where $\boldsymbol{q}_t \in \mathbb{R}^{mn}$, $t = 1, \ldots, N$ are sample cost data from a given distribution. We can rewrite this as a large linear optimization problem as follows:

$$
\begin{array}{rll}
Z_S = \min & \displaystyle\sum_{i=1}^{n} c_i x_i + \frac{1}{N}\sum_{t=1}^{N} z_t & \\
\text{s.t.} & z_t \geq \alpha_k \left( \displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n} q_{ijt} w_{ijt} \right) + \beta_k, & \forall\, k,\, t, \\
& \displaystyle\sum_{i=1}^{m} w_{ijt} = h_j, & \forall\, j,\, t, \\
& \displaystyle\sum_{j=1}^{n} w_{ijt} = x_i, & \forall\, i,\, t, \\
& 0 \leq x_i \leq 1,\; w_{ijt} \geq 0, & \forall\, i,\, j,\, t.
\end{array}
$$

The code for this data-driven model is developed in C with CPLEX 9.1 solver.

## Numerical Example

We generate randomly $m = 5$ facilities and $n = 20$ customer locations within the unit square. The distance $\bar{q}_{ij}$ from facility $i$ to customer location $j$ is calculated. The first and second moments $\boldsymbol{\mu}$ and $\boldsymbol{Q}$ of the random distances $\tilde{\boldsymbol{q}}$ are generated by constructing $1,000$ uniform cost vectors $\boldsymbol{q}_t$ from independent uniform distributions on intervals $[0.5\bar{q}_{ij}, 1.5\bar{q}_{ij}]$ for all $i, j$. The production cost $c_i$ is randomly generated from a uniform distribution on the interval $[0.5\bar{c}, 1.5\bar{c}]$, where $\bar{c}$ is the average transportation cost. Similarly, the demand $h_j$ is randomly generated from the uniform distribution on the interval $[0.5\frac{m}{n}, \frac{m}{n}]$ so that the constraint $\sum_j h_j < m$ is satisfied. Customer locations and warehouse sites for this instance are shown in Figure 1.

We consider two different disutility functions - the risk-neutral one, $\mathbb{U}(x) = x$, and the piecewise linear approximation of the exponential risk-averse disutility function $\mathbb{U}_e(x) = \gamma(e^{\delta x} - 1)$, where $\gamma, \delta > 0$. For this problem instance, we set $\gamma = 0.25$ and $\delta = 2$ and use an equidistant linear approximation with $K = 5$ for $\mathbb{U}_e(x)$, $x \in [0, 1]$. Both disutility functions are plotted in Figure 2.

The data-driven model is solved with $10,000$ samples $\boldsymbol{q}_t$ generated from the normal distribution $P_d$ with the given first and second moment $\boldsymbol{\mu}$ and $\boldsymbol{Q}$. Optimal solutions and total costs of this problem instance obtained from the two models are shown in Table 1. The total cost obtained from the minimax model is indeed higher than that from the data-driven model. This can be explained by the fact that the former model hedges against the worst possible distributions. We also calculate production costs and expected risk-averse transportation costs for these two models and the results are reported in Table 2. The production costs are higher under risk-averse consideration. This indeed justifies the change

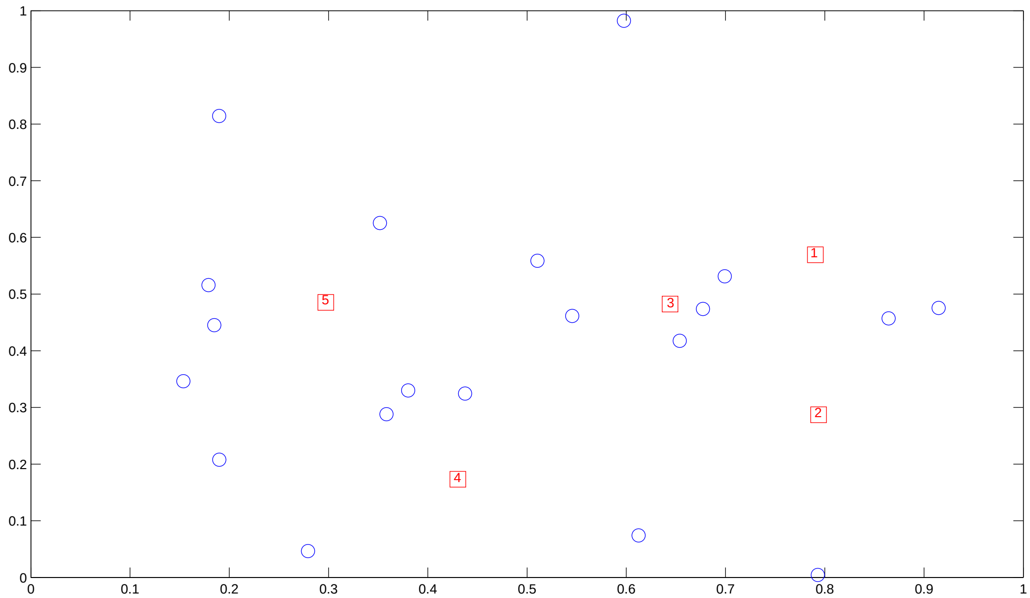


Figure 1: Customer locations (circles) and facility locations (squares)

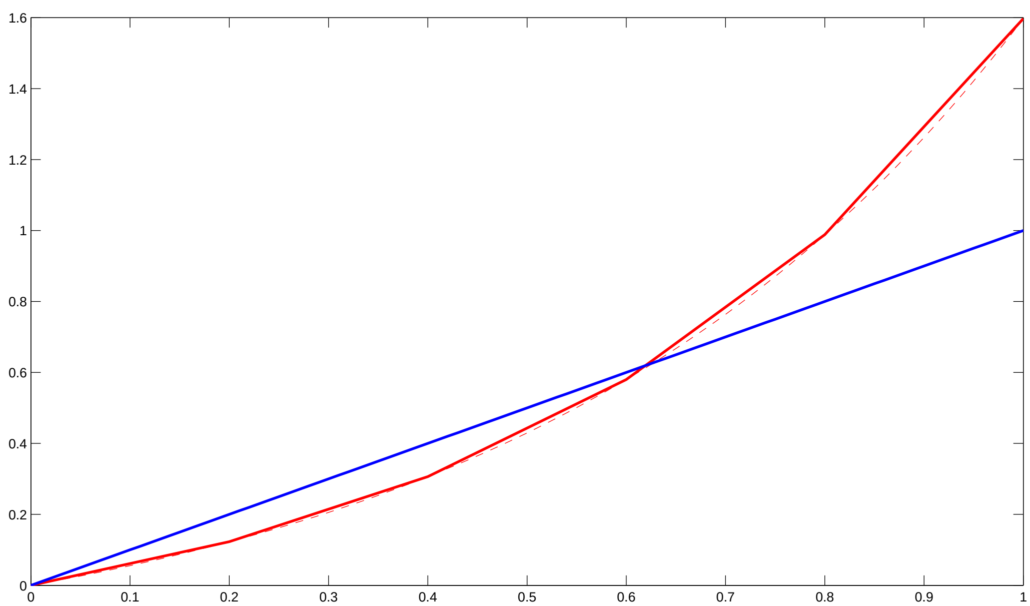


Figure 2: Approximate exponential risk-averse disutility function and risk-neutral one

in optimal first-stage solutions, which aims at reducing risk effects in the second stage (with smaller transportation cost in this case). Changes in the minimax solution are more significant than those in the data-driven one with higher relative change in the production cost.

| Disutility Function | Model | Optimal Solution | Total Cost |
|---|---|---|---|
| $\mathbb{U}(x) = x$ | Minimax | $\boldsymbol{x}_m = (0.1347; 0.6700; 0.8491; 1.0000; 1.0000)$ | 1.6089 |
| | Data-Driven | $\boldsymbol{x}_d = (0.2239; 0.5808; 0.8491; 1.0000; 1.0000)$ | 1.5668 |
| $\mathbb{U}(x) \approx 0.25(e^{2x} - 1)$ | Minimax | $\boldsymbol{x}_m = (0.5938; 0.2109; 0.8491; 1.0000; 1.0000)$ | 1.6308 |
| | Data-Driven | $\boldsymbol{x}_d = (0.3606; 0.4409; 0.8523; 1.0000; 1.0000)$ | 1.5533 |

Table 1: Optimal solutions and total costs obtained from two models under different disutility functions

| Disutility Function | Model | Production Cost | Transportation Cost | Total Cost |
|---|---|---|---|---|
| $\mathbb{U}(x) = x$ | Minimax | 0.9605 | 0.6484 | 1.6089 |
| | Data-Driven | 0.9676 | 0.5992 | 1.5668 |
| $\mathbb{U}(x) \approx 0.25(e^{2x} - 1)$ | Minimax | 0.9968 | 0.6340 | 1.6308 |
| | Data-Driven | 0.9785 | 0.5747 | 1.5533 |

Table 2: Production and risk-averse transportation costs obtained from two models under different disutility functions

Using Theorem 2, the optimal dual variables are used to construct the limiting extremal distribution $P_m(\boldsymbol{x}_m)$ for the solution $\boldsymbol{x}_m$. For the risk-neutral problem, this worst-case distribution simply reduces to a limiting one-point distribution. The Jensen's bound is obtained and with the mean transportation costs, the solution $\boldsymbol{x}_m$ performs better than the solution $\boldsymbol{x}_d$ obtained from the data-driven approach. The total cost increases from 1.6089 to 1.6101. For the risk-averse problem, the limiting extremal distribution is a discrete distribution with two positive probabilities of 0.2689 and 0.7311 for two pieces of the approximating piecewise linear risk function, $k = 3$ and $k = 4$ respectively. The total cost of 1.6308 is obtained under this distribution with the solution $\boldsymbol{x}_m$, which is indeed the maximal cost obtained from the minimax model. We can also obtain the limiting extremal distribution $P_m(\boldsymbol{x}_d)$ for the solution $\boldsymbol{x}_d$, which is again a discrete distribution. Two pieces $k = 3$ and $k = 4$ have the positive probability of 0.1939 and 0.8061 respectively while two additional pieces $k = 1$ and $k = 5$ are assigned a very small positive probability of $3.4 \times 10^{-5}$ and $2.1 \times 10^{-5}$. Under this extremal distribution, the data-driven solution $\boldsymbol{x}_d$ yields the total cost of 1.6347, which is higher than the maximal cost obtained from the minimax model.

We next stress test the quality of the stochastic optimization solution by contaminating the original probability distribution $P_d$ used in the data-driven model. We use the approach proposed in Dupacova [9] to test the quality of the solutions on the contaminated distribution

$$P_\lambda = (1-\lambda)P_d + \lambda Q,$$

for $\lambda$ varying between $[0,1]$. The distribution $Q$ is a probability distribution different from $P_d$ that one wants to test their first-stage solution against. Unfortunately, no prescription on a good choice of $Q$ is provided in Dupacova [9]. We now propose a general approach to stress-test the quality of stochastic optimization solutions:

(1) Solve the data-driven linear optimization problem arising from the distribution $P_d$ to find the optimal first-stage solution $\boldsymbol{x}_d$.

(2) Generate the extremal distribution $P_m(\boldsymbol{x}_d)$ that provides the worst case expected cost for the solution $\boldsymbol{x}_d$.

(3) Test the quality of the data-driven solution $\boldsymbol{x}_d$ on the distribution $P_\lambda = (1-\lambda)P_d + \lambda P_m(\boldsymbol{x}_d)$ as $\lambda$ is varied between $[0,1]$.

In our experiment, we compare the data-driven solution $\boldsymbol{x}_d$ and the minimax solution $\boldsymbol{x}_m$ on the contaminated distribution $P_\lambda = (1-\lambda)P_d + \lambda P_m(\boldsymbol{x}_d)$ for the risk-averse problem. For a given solution $\boldsymbol{x}$, let $z_1(\boldsymbol{x})$, $z_2(\boldsymbol{x})$ denote the production cost and the random transportation cost with respect to random cost vector $\tilde{\boldsymbol{q}}$. The total cost is $z(\boldsymbol{x}) = z_1(\boldsymbol{x}) + z_2^r(\boldsymbol{x})$, where $z_2^r(\boldsymbol{x}) = \mathbb{U}\left(z_2(\boldsymbol{x})\right)$ is the risk-averse transportation cost. For each $\lambda \in [0,1]$, we compare the minimax solution relative to the data-driven solution using the following three quantities

(1) Expectation of total cost (in %):

$$\left(\frac{\mathbb{E}_{P_\lambda}\left[z(\boldsymbol{x}_m)\right]}{\mathbb{E}_{P_\lambda}\left[z(\boldsymbol{x}_d)\right]} - 1\right) \times 100\%,$$

(2) Standard deviation of total cost (in %)

$$\left(\frac{\sqrt{\mathbb{E}_{P_\lambda}\left[z(\boldsymbol{x}_m) - \mathbb{E}_{P_\lambda}\left[z(\boldsymbol{x}_m)\right]\right]^2}}{\sqrt{\mathbb{E}_{P_\lambda}\left[z(\boldsymbol{x}_d) - \mathbb{E}_{P_\lambda}\left[z(\boldsymbol{x}_d)\right]\right]^2}} - 1\right) \times 100\%,$$

(3) Quadratic semi-deviation of total cost (in %)

$$\left( \frac{\sqrt{\mathbb{E}_{P_\lambda}\Big[\max(0, z(\boldsymbol{x}_m) - \mathbb{E}_{P_\lambda}[z(\boldsymbol{x}_m)])\Big]}}{\sqrt{\mathbb{E}_{P_\lambda}\Big[\max(0, z(\boldsymbol{x}_d) - \mathbb{E}_{P_\lambda}[z(\boldsymbol{x}_d)])\Big]}} - 1 \right) \times 100\%.$$

These measures are also applied for $z_2(\boldsymbol{x})$, the transportation cost without risk-averse consideration. When these quantities are below 0, it indicates that the minimax solution is outperforming the data-driven solution whereas when it is greater than 0, the data-driven is outperforming the minimax solution. The standard deviation is symmetric about the mean, penalizing both the upside and the downside. On the other hand, the quadratic semi-deviation penalizes only when the cost is larger than the mean value.

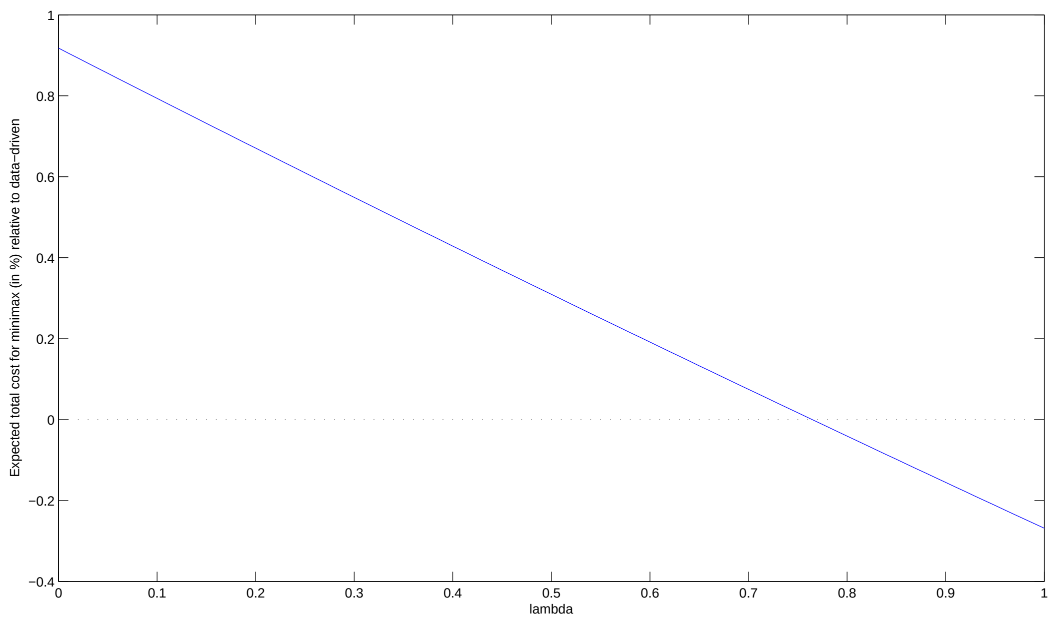


Figure 3: Relative difference in expectation of total cost of minimax and data-driven model

Figure 3 shows that the minimax solution is better than the data-driven solution in terms of total cost when $\lambda$ is large enough ($\lambda > 0.75$ in this example). If we only consider the second-stage transportation cost, the minimax solution results in smaller expected costs for all $\lambda$ and the relative differences are increased when $\lambda$ increases. This again shows that the minimax solution incurs higher production cost while maintaining smaller transportation cost to reduce the risk effects in the second stage. Figure 4 also shows that the risk-averse cost changes faster than the risk-neutral cost. The production cost $z_1(\boldsymbol{x})$

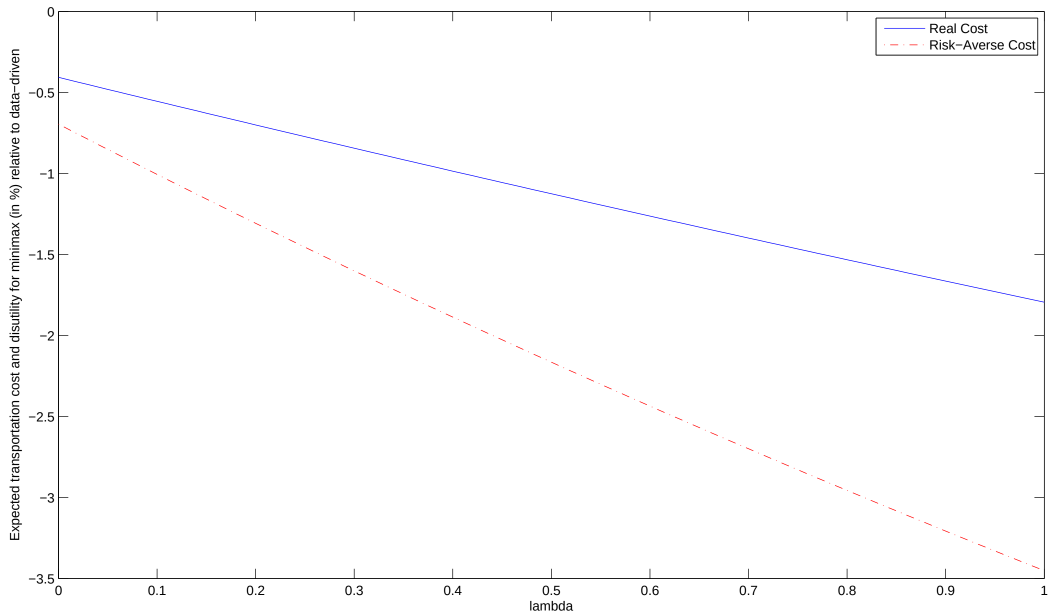


Figure 4: Relative difference in expectation of transportation costs of minimax and data-driven model

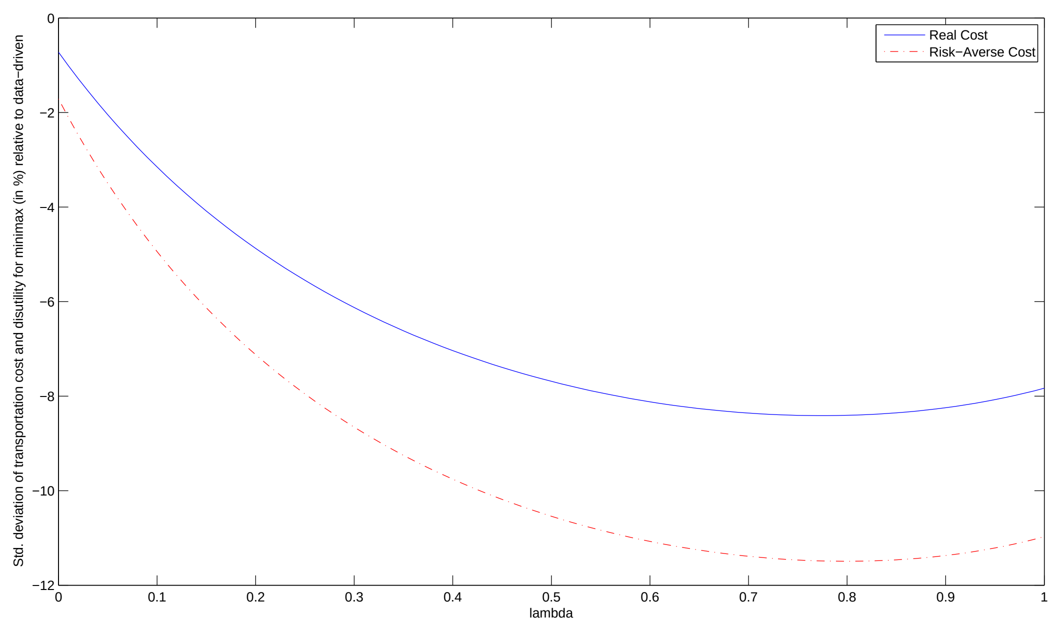


Figure 5: Relative difference in standard deviation of transportation costs of minimax and data-driven model

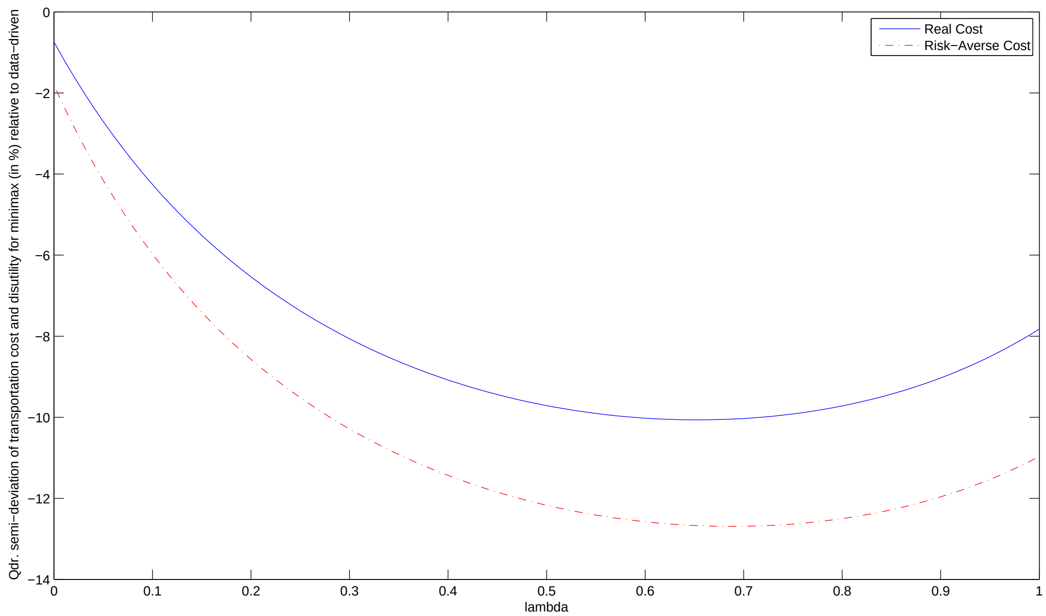


Figure 6: Relative difference in quadratic semi-deviation of transportation costs of minimax and data-driven model

is fixed for each solution $\boldsymbol{x}$; therefore, the last two measures of total cost $z(\boldsymbol{x})$ are exactly the same for those of risk-averse transportation cost $z_2^r(\boldsymbol{x})$. Figure 5 and 6 illustrates these two measures for risk-averse transportation cost and its risk-neutral counterpart. The minimax solution is clearly better than the data-driven solution in terms of standard deviation and quadratic semi-deviation for all values of $\lambda$ and the differences are more significant in the case of risk-averse cost.

## 4.2 Single Facility Minimax Distance Problem

Let $(x_1, y_1), \ldots, (x_n, y_n)$ denote $n$ customer locations on a plane. The single facility minimax distance problem is to identify a facility location $(x, y)$ that minimizes the maximum distance from the facility to the customers. Assuming a rectilinear or Manhattan distance metric, the problem is formulated as:

$$\min_{x,y} \left( \max_{i=1,\ldots,n} |x_i - x| + |y_i - y| \right).$$

This can be solved as a linear optimization problem:

$$
\begin{aligned}
\min_{x,y,z} \quad & z \\
\text{s.t.} \quad & z + x + y \geq x_i + y_i, \quad && \forall\, i, \\
& z - x - y \geq -x_i - y_i, && \forall\, i, \\
& z + x - y \geq x_i - y_i, && \forall\, i, \\
& z - x + y \geq -x_i + y_i, && \forall\, i.
\end{aligned}
$$

Carbone and Mehrez [5] studied this problem under the following stochastic model for customer locations: The coordinates $\tilde{x}_1, \tilde{y}_1, \ldots, \tilde{x}_n, \tilde{y}_n$ are assumed to be identical, pairwise independent and normally distributed random variables with mean 0 and variance 1. Under this distribution, the optimal solution to the stochastic problem:

$$\min_{x,y} \mathbb{E}\left(\max_{i=1,\ldots,n} |\tilde{x}_i - x| + |\tilde{y}_i - y|\right),$$

is just $(x, y) = (0, 0)$.

We now solve the minimax version of this problem under weaker distributional assumptions using only first and second moment information. This fits the model proposed in Section 3.2 with random right hand side. The stochastic problem for the minimax single facility distance problem can be written as follows:

$$Z = \min_{x,y} \max_{P \in \mathcal{P}} \mathbb{E}_P \left[\mathbb{U}\left(\max_{i=1,\ldots,n} |\tilde{x}_i - x| + |\tilde{y}_i - y|\right)\right],$$

where the random vector $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = (\tilde{x}_1, \tilde{y}_1, \ldots, \tilde{x}_n, \tilde{y}_n)$ has mean $\boldsymbol{\mu}$ and second moment matrix $\boldsymbol{Q}$ and $\mathbb{U}$ is the disutility function defined in (2). The equivalent semidefinite optimization problem is given as

$$
\begin{aligned}
Z_{SDP} = \min \quad & \boldsymbol{Q} \cdot \boldsymbol{Y} + \boldsymbol{\mu}'\boldsymbol{y} + y_0 \\
\text{s.t.} \quad & \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y} - \alpha_k(\boldsymbol{e}_{2i-1} + \boldsymbol{e}_{2i})}{2} \\ \frac{(\boldsymbol{y} - \alpha_k(\boldsymbol{e}_{2i-1} + \boldsymbol{e}_{2i}))'}{2} & y_0 + \alpha_k(x + y) - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, i,\, k, \\
& \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y} + \alpha_k(\boldsymbol{e}_{2i-1} + \boldsymbol{e}_{2i})}{2} \\ \frac{(\boldsymbol{y} + \alpha_k(\boldsymbol{e}_{2i-1} + \boldsymbol{e}_{2i}))'}{2} & y_0 - \alpha_k(x + y) - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, i,\, k, \\
& \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y} - \alpha_k(\boldsymbol{e}_{2i-1} - \boldsymbol{e}_{2i})}{2} \\ \frac{(\boldsymbol{y} - \alpha_k(\boldsymbol{e}_{2i-1} - \boldsymbol{e}_{2i}))'}{2} & y_0 + \alpha_k(x - y) - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, i,\, k, \\
& \begin{pmatrix} \boldsymbol{Y} & \frac{\boldsymbol{y} + \alpha_k(\boldsymbol{e}_{2i-1} - \boldsymbol{e}_{2i})}{2} \\ \frac{(\boldsymbol{y} + \alpha_k(\boldsymbol{e}_{2i-1} - \boldsymbol{e}_{2i}))'}{2} & y_0 - \alpha_k(x - y) - \beta_k \end{pmatrix} \succeq 0, \quad \forall\, i,\, k,
\end{aligned}
\tag{19}
$$

where $\boldsymbol{e}_i \in \mathbb{R}^{2n}$ denote the unit vector in $\mathbb{R}^{2n}$ with the $i$th entry having a 1 and all other entries having 0. This semidefinite optimization problem is obtained from Theorem 4 and using the fact that the set of extreme points for the dual feasible region consist of the $4n$ solutions:

$$\{\boldsymbol{e}_{2i-1}+\boldsymbol{e}_{2i}, -\boldsymbol{e}_{2i-1}-\boldsymbol{e}_{2i}, \boldsymbol{e}_{2i-1}-\boldsymbol{e}_{2i}, -\boldsymbol{e}_{2i-1}+\boldsymbol{e}_{2i}\}_{i=1,\ldots,n}.$$

The data-driven approach for this problem is solved using the formulation:

$$Z_D = \min_{x,y} \frac{1}{N}\sum_{t=1}^{N} \mathbb{U}\left(\max_{i=1,\ldots,n} |x_{it}-x|+|y_{it}-y|\right),$$

where $(x_{1t}, y_{1t}), \ldots, (x_{nt}, y_{nt})$ are location data for the samples $t = 1, \ldots, N$. This problem can be solved as the large scale linear optimization problem

$$\begin{aligned}
Z_D = \min\nolimits_{x,y,z_t} \quad & \frac{1}{N}\sum_{t=1}^{N} z_t \\
\text{s.t.} \quad & z_t + \alpha_k(x+y) \geq \alpha_k(x_{it}+y_{it}) + \beta_k \quad && \forall\, i,\, k,\, t, \\
& z_t - \alpha_k(x+y) \geq -\alpha_k(x_{it}+y_{it}) - \beta_k && \forall\, i,\, k,\, t, \\
& z_t + \alpha_k(x-y) \geq \alpha_k(x_{it}-y_{it}) + \beta_k && \forall\, i,\, k,\, t, \\
& z_t - \alpha_k(x-y) \geq -\alpha_k(x_{it}-y_{it}) + \beta_k && \forall\, i,\, k,\, t.
\end{aligned}$$

### Numerical Example

In this example, we generate $n = 20$ customer locations by randomly generating clusters within the unit square. Each customer location is perturbed from its original position by a random distance in a random direction. The first and second moments $\boldsymbol{\mu}$ and $\boldsymbol{Q}$ are estimated by performing $1,000$ such random perturbations. We first solve both the minimax and data-driven model to find the optimal facility locations $(x_m, y_m)$ and $(x_d, y_d)$ respectively. The data-driven model is solved using $10,000$ samples drawn from the normal distribution with given first and second moments. In this example, we focus on the risk-neutral case with $\mathbb{U}(x) = x$.

The optimal facility location and the expected costs are shown in Table 3. As should be, the expected maximum distance between a customer and the optimal facility is larger under the minimax model as compared to the data-driven approach. The (expected) customer locations and the optimal facility locations are plotted in Figure 7.

To compare the quality of the solutions, we plot the probability that a customer is furthest away from the optimal facility for the minimax and data-driven approach (see Figure 8). For the minimax problem, these probabilities were obtained from the optimal dual variables to the semidefinite optimization problem (19). For the data-driven approach, the probabilities were obtained through an extensive simulation using $100,000$ samples from the normal distribution. Qualitatively, these two plots look fairly

| Disutility Function | Model | Optimal Solution | Expected Maximum Distance |
|---|---|---|---|
| $\mathbb{U}(x) = x$ | Minimax | $(x_m, y_m) = (0.5975, 0.6130)$ | 0.9796 |
| | Data-Driven | $(x_d, y_d) = (0.6295, 0.5952)$ | 0.6020 |

Table 3: Optimal solutions and total costs obtained from two models for the risk-neutral case

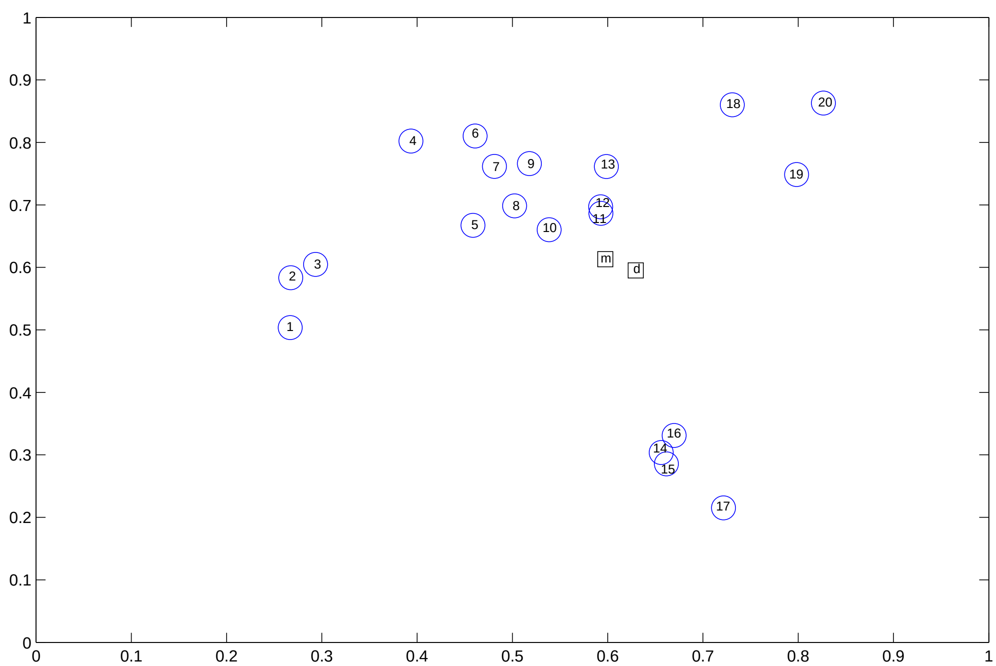


Figure 7: Facility location solutions (square) and expected customer locations (circles)

similar. In both solutions, the facilities 17, 20 and 1 (in decreasing order) have the most significant probabilities of being furthest away from the optimal facility. The worst case distribution tends to even out the probabilities that the different customers are far away from the facilities as compared to the normal distribution. For instance, the minimax solution predicts larger probabilities for facilities 5 to 16 as compared to the data-driven solution. The optimal minimax facility location thus seems to be hedging against the possibility of each customer facility moving far away from the center (extreme case). The optimal data-driven facility on the other hand seems to be hedging more against the customers that are far away from the center in an expected sense (average case). The probability distribution for the maximum distance in the two cases are provided in Figures 9 and 10. The larger distances and the discrete nature of the extremal distribution are evident as compared to the smooth normal distribution.

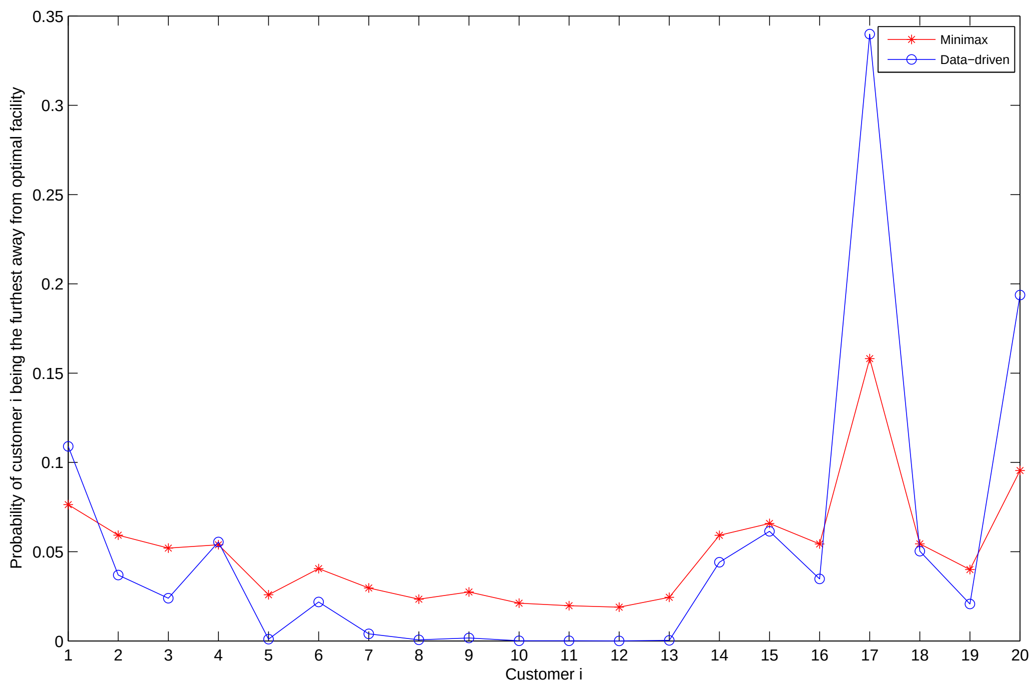


Figure 8: Probability of customers being at the maximum distance from $(x_m, y_m)$ and $(x_d, y_d)$

We next stress test the quality of the stochastic optimization solution by contaminating the original probability distribution $P_d$ used in the data-driven model. In our experiment, we compare the data-driven solution $(x_d, y_d)$ and the minimax solution $(x_m, y_m)$ on the contaminated distribution $P_\lambda = (1-\lambda)P_d + \lambda P_m(x_d, y_d)$. For a given facility location $(x, y)$, let $z(x, y)$ denote the (random) maximum

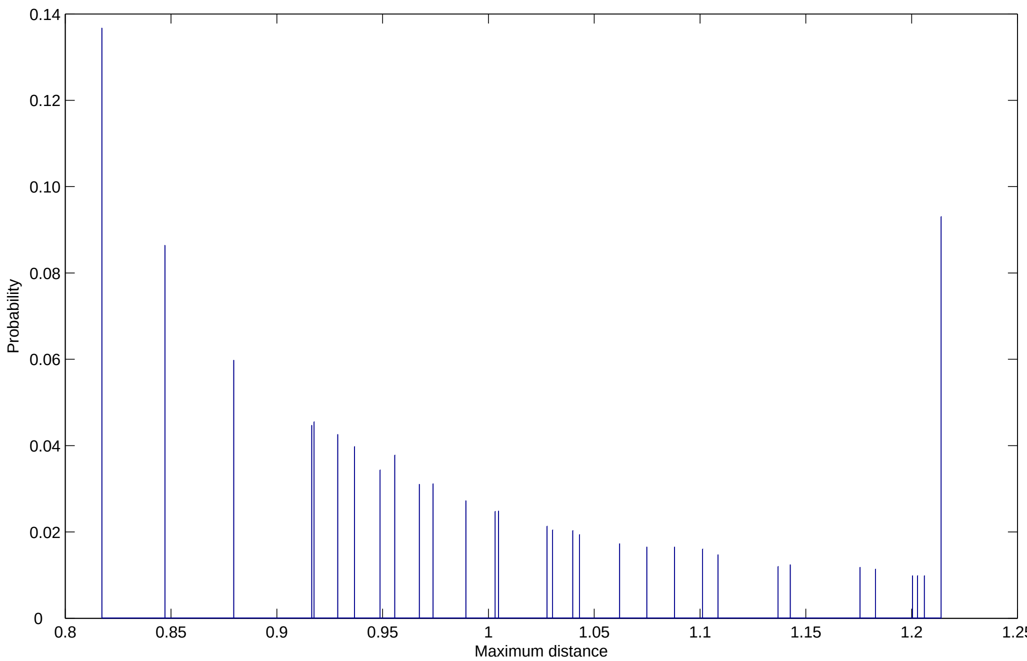


Figure 9: Distribution of maximum distances under the extremal distribution $P_m(\boldsymbol{x})$ for $(x_m, y_m)$

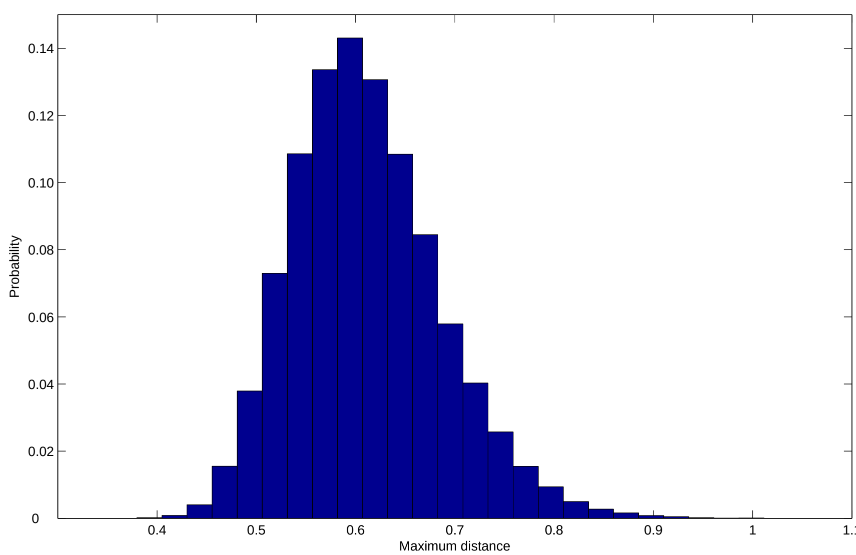


Figure 10: Distribution of maximum distances under the normal distribution for $(x_d, y_d)$

distance between the facility and customer locations:

$$z(x,y) = \max_{i=1,\dots,n} |\tilde{x}_i - x| + |\tilde{y}_i - y|.$$

For each $\lambda \in [0,1]$, we again compare the minimax solution relative to the data-driven solution using the three quantities: expectation, standard deviation and quadratic semi-deviation of max distance.

The results for different $\lambda$ are displayed in Figures 11, 12 and 13. From Figure 11, we see that for $\lambda$ closer to 0, the minimax solution has larger expected distances as compared to the data-driven solution. This should be expected, since the data-driven solution is trying to optimize the exact distribution. However as the contamination factor $\lambda$ increases (in this case beyond 0.5), the minimax solution performs better than the data-driven solution. This suggests that if there is significant uncertainty in the knowledge of the exact distribution, the minimax solution would be a better choice. The average maximum distance from the two solutions is within 2% of each other. Interestingly, again from Figures 12 and 13 it is clear that the standard deviation and the quadratic semi-deviation from the minimax solution is generally lesser than that for the data-driven solution. In our experiments this is true for all $\lambda \geq 0.05$. This is a significant benefit that the minimax solution provides as compared to the data-driven solution under contamination.

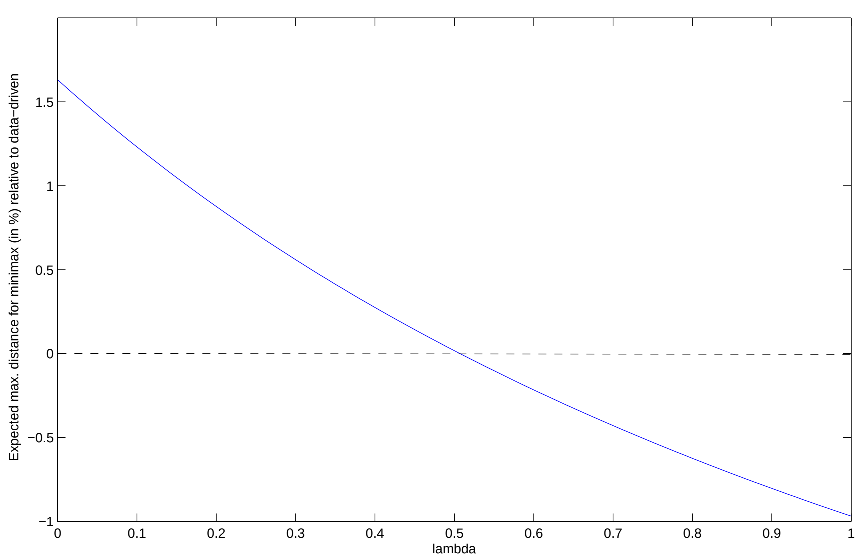


Figure 11: Relative difference in expectation of maximum distance obtained from minimax and data-driven solution

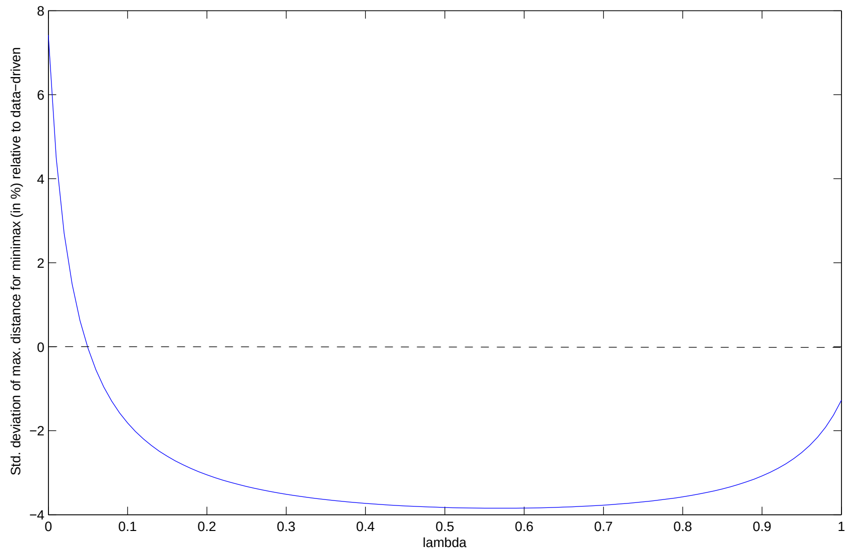


Figure 12: Relative difference in standard deviation of maximum distance obtained from minimax and data-driven solution

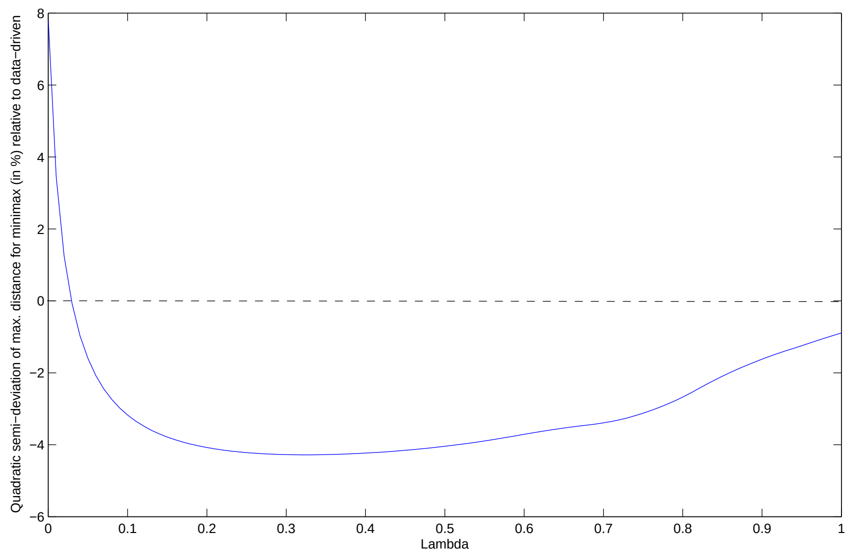


Figure 13: Relative difference in quadratic semi-deviation of maximum distance obtained from minimax and data-driven solution

# References

[1] S. Ahmed. Convexity and decomposition of mean-risk stochastic programs. *Mathematical Progamming*, 106:433–446, 2006.

[2] A. Ben-Tal and M. Teboulle. Expected utility, penalty functions and duality in stochastic nonlinear programming. *Management Science*, 32:1445–1466, 1986.

[3] A. Ben-Tal and M. Teboulle. An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance*, 17(3):449–476, 2007.

[4] M. Breton and S. El Hachem. Algorithms for the solution of stochastic dynamic minimax problems. *Computational Optimization and Applications*, 4:317–345, 1995.

[5] R. Carbone and A. Mehrez. The single facility minimax distance problem under stochastic location demand. *Management Science*, 26(1):113–115, 1980.

[6] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Submitted for publication*, 2008.

[7] S. P. Dokov and D. P. Morton. Second-order lower bounds on the expectation of a convex function. *Mathematics of Operations Research*, 30:662–677, 2005.

[8] J. Dupačová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20:73–88, 1987.

[9] J. Dupačová. Stress testing via contamination. *Coping with Uncertainty: Modeling and Policy Issues, Lecture Notes in Economics and Mathematical Systems*, 581:29–46, 2006.

[10] A. Eichorn and W. Romisch. Polyhedral risk measures in stochastic programming. *SIAM Journal on Optimization*, 16:69–95, 2005.

[11] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer, Berlin, Heidelberg, 1988.

[12] K. Isii. On the sharpness of Chebyshev-type inequalities. *Annals of the Institute of Statistical Mathematics*, 12:185–197, 1963.

[13] P. Kall and S. Wallace. *Stochastic Programming*. John Wiley & Sons, Chichester, England, 1994.

[14] A. J. Kleywegt, A. Shapiro, and T. Homem de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12:479–502, 2001.

[15] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[16] O. J. Mangasarian and T. H. Shiau. A variable-complexity norm maximization problem. *SIAM Journal on Algebraic and Discrete Methods*, 7(3):455–461, 1986.

[17] M. Riis and K. A. Andersen. Applying the minimax criterion in stochastic recourse programs. *European Journal of Operational Research*, 165(3):569–584, 2005.

[18] D. Rutenberg. Risk aversion in stochastic programming with recourse. *Operations Research*, 21(1):377–380, 1973.

[19] A. Shapiro and S. Ahmed. On a class of minimax stochastic programs. *SIAM Journal on Optimization*, 14:1237–1249, 2004.

[20] A. Shapiro and T. Homem-De-Mello. On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM Journal on Optimization*, 11:70–86, 2000.

[21] A. Shapiro and A. Kleywegt. Minimax analysis of stochastic programs. *Optimization Methods and Software*, 17:523–542, 2002.

[22] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8:171–176, 1958.

[23] J. F. Sturm. Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11-12:625–653, 1999.

[24] J. Žáčková. On minimax solutions of stochastic linear programming problems. *Casopis pro pestovani matematiky*, 91:423–430, 1966.