Multiple Objectives Satisficing under Uncertainty *

Shao-Wei Lam† Tsan Sheng Ng‡ Melvyn Sim§ Jin-Hwa Song ¶

Dec 2010

Abstract

We propose a class of functions, called multiple objective satisficing (MOS) criteria, for evaluating the level of compliance of a set of objectives in meeting their targets collectively under uncertainty. The MOS criteria include the targets' achievement probability (success probability criterion) as a special case and also extend to situations when the probability distribution is not fully characterized. We focus on a class of MOS criteria that favors diversification, which has the potential to mitigate severe shortfalls in scenarios when an objective fails to achieve its target. Naturally, this class excludes success probability and we propose the shortfall-aware MOS criterion (S-MOS), which is diversification favoring and is a lower bound to success probability. We also show how to build tractable approximations of the S-MOS criterion. As S-MOS criterion maximization is not a convex optimization problem, we propose improvement algorithms via solving sequences of convex optimization problems. We report encouraging computational results on a blending problem in meeting specification targets even in the absence of full probability distribution description.

Keywords: satisficing, targets, multiple objectives, robust optimization

*The research is supported by the grant R-314-000-081-597 from ExxonMobil Research and Engineering Company.

†Department of Decision Sciences, NUS Business School. Email: bizlamsw@nus.edu.sg.

‡Department of Industrial & Systems Engineering, NUS. Email : isentsa@nus.edu.sg

§NUS Business School and NUS Risk Management Institute. Email: dscsimm@nus.edu.sg.

¶Corporate Strategic Research, ExxonMobil Research and Engineering Company. Email: jinhwa.song@exxonmobil.com.

1 Introduction

In this paper, we consider a decision problem where a solution is evaluated on the set of objectives that are potentially uncertain. The goal is to find the solution so that these objectives could best achieve their targets under uncertainty. The relevance of targets in managerial decision making is widespread in many business domains. Numerous research have demonstrated the importance of target setting and the multiplicity of targets (Lanzilloti, 1958; Boulding, 1952). Simon (1959) coins the term "satisfice" to describe a decision-making process involving real-world agents choosing actions that ensure that certain aspiration levels or targets will be achieved. The multiplicity of goals can also be seen as a way of hedging under uncertainty (Carter, 1971). Cyert and March (1963) describe the importance of goals, expectations and choice as the major components of organizational decision processes. Empirical research have concluded the relevance and importance of a target-based perspective in the context of financial risk management (Mao, 1970).

Uncertainty is ubiquitous in the real world and the probability of achieving targets or success probability is a natural criterion for ranking alternatives. In fact, the relationship between the success probability criterion and expected utility has been studied (Castagnoli and LiCalzi, 1996; Bordley and LiCalzi, 2000). Bordley and Kirkwood (2004) propose an approach based on the maximization of expected utility where the decision maker's utility depends on the subset of attributes that meet some targets. Along these lines of inquiry, Bordley and Pollock (2009) further present a probability maximization formulation as an alternative to the chance-constrained programming model, which is a model that has long been studied (see, for instance, Charnes et.al (1958); Miller and Wagner (1965); Prékopa (1970) and recently, Chen et al. (2010)) . One drawback of the success probability is that it does not account for the degree of a target's shortfall that might occur. In real-world situations, decision makers are not completely insensitive to the magnitudes of losses (Payne et al., 1980; Diecidue and Ven, 2008). Another issue with this criterion is the necessity for decision makers to specify the full probability distributions in the problem formulation, which may not be available in practice.

In the economic literature, the distinction between risk, where outcome frequency is known, and ambiguity, where it is not, can be retrospected to Knight (1921). Ellsberg (1961) shows convincingly by means of paradoxes that ambiguity preference cannot be reconciled with classical expected utility theory. Distributionally robust optimization or minimax stochastic programming is an approach where the decision maker is ambiguity averse and the optimal decisions are sought for the worst case probability distributions within a family of possible distributions, defined by certain properties such as their support and moments. This approach was pioneered by Žáčková (1966) and studied in many other works (see, for instance, Dupačová (1987); Shapiro and Kleywegt (2002); Bertsimas and Popescu (2005)). El Ghaoui et al. (2003) develop worst-case bounds for chance-constraints on uncertainties, when only the bounds on means and covariance matrix were available whilst Delage and Ye (2010) study distributional robust stochastic programs when the mean and covariance of the primitive uncertainties are themselves subject to uncertainty. Apart from mean and covariance, new distributional properties were proposed under this framework and unified bounds were developed for uncertainties with known support, mean, covariance and new deviation measures such as directional deviations and partitioned covariance (Chen et al., 2008; Chen and Sim , 2009; Goh and Sim, 2010; Natarajan et al., 2010). Goh and Sim (2011) introduce an algebraic modeling toolkit (ROME) to facilitate modelling of distributional robust optimization problems and we use it to implement the multiple objectives satisficing framework for situations with ambiguity.

Extending the satisficing measures of Brown and Sim (2009), we propose a class of functions, called multiple objective satisficing (MOS) criteria, for evaluating the level of compliance of a set of objectives in meeting their targets collectively under uncertainty. The MOS criteria encompass success probability as a special case and can be extended to incorporate a larger scope of uncertainty including distributional ambiguity. We also propose a class of MOS criteria that favor diversification, which has the potential to mitigate severe shortfalls in scenarios when an objective fails to achieve its target. We propose the shortfall-aware MOS criterion (S-MOS), which is diversification favoring and is an extension of the work of Chen and Sim (2009) to problems with multiple objectives. The S-MOS criterion is a lower bound to success probability and in the presence of distributional ambiguity, we present the techniques for building tractable approximations while preserving the salient properties of MOS criteria. These approximations can be implemented using the distributional robust optimization framework recently proposed by Goh and Sim (2010, 2011). Although maximizing the S-MOS criterion is not a convex optimization, we propose improvement algorithms via solving sequences of convex optimization problems. These algorithms perform remarkably well in our computational studies and we report encouraging results on a refinery blending problem in meeting specification targets even in the absence of full probability distribution description.

The structure for the rest of the paper is given here. In Section 2, we introduce the desirable properties of MOS and diversification favoring MOS criteria. Subsequently, in Section 3, we propose the shortfall-aware MOS criterion (S-MOS), which mimics the success probability criterion but is also diversification favoring. We show how the S-MOS criterion can be computed and optimized under distributional ambiguity. Finally in Section 4, we present some computational examples to compare the S-MOS criterion with success probability. The results on the performance of S-MOS under distributional ambiguity are also reported. An application of the S-MOS criterion for a problem in a refinery blending optimization is also presented.

Notations: Matrices and vectors are represented as upper and lower case boldface characters respectively. If $\boldsymbol{x}$ is a vector, we use the notation x_i to denote the ith component of the vector. Also, given a vector $\boldsymbol{x}$, we define $(y_i, \boldsymbol{x}_{-i})$ to be the vector with all components same as that in $\boldsymbol{x}$ except component i being y_i; i.e. $(y_i, \boldsymbol{x}_{-i}) = (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)$ for a vector of n components. We use $\mathbf{1}$ to denote a vector of all ones and $\mathbf{0}$ to denote the vector of all zeros. Also, we denote an uncertain parameter with a tilde, such as $\tilde{z}$. For convenience of exposition, we denote $\min\{\min_{i\in N}\{a_i\}, b\}$ where $\boldsymbol{a} \in \Re^N$ and $b \in \Re$ simply by $\min_{i\in N}\{a_i, b\}$.

2 A framework for multiple objectives satisficing

In this section, we propose a class of functions for evaluating how well the objectives meet their targets collectively under uncertainty. These functions are called multiple objective satisficing (MOS) criteria which encompass success probability criterion as a special case. We consider uncertainties that are represented by a state-space Ω and a set (sigma-algebra) $\mathcal{F}$ of events. To encompass distributional ambiguity, we do not necessarily specify a probability distribution on $(\Omega, \mathcal{F})$. Instead, we let $\mathbb{F}$ be a set of probability measures on $(\Omega, \mathcal{F})$. For a given distribution, $\mathbb{P} \in \mathbb{F}$, $\mathbb{E}_{\mathbb{P}}(\cdot)$ denotes taking expectation over the distribution $\mathbb{P}$. We use the inequality $\tilde{x} \geq 0$ to imply $\mathbb{P}(\tilde{x} \geq 0) = 1$ for all $\mathbb{P} \in \mathbb{F}$ and $\tilde{x} \ngeq 0$ to imply there exists $\mathbb{P} \in \mathbb{F}$ such that $\mathbb{P}(\tilde{x} \geq 0) < 1$. Strict inequality, such as $\tilde{x} < 0$ implies there exists an $\epsilon < 0$ such that $\mathbb{P}(\tilde{x} \leq \epsilon) = 1$ for all $\mathbb{P} \in \mathbb{F}$.

We consider n uncertain objectives and represent the outcomes as a multivariate random variable on Ω, i.e, a vector of n functions $\tilde{\boldsymbol{a}} : \Omega \mapsto \Re^n$. Let $\mathcal{Y}$ be a set of these random variables. We define the index set

$N = \{1, \ldots, n\}$. We allow targets to be uncertain and denote this as the random variable $\tilde{\boldsymbol{\tau}}$. For convenience, we suppress the notation of the targets by defining the set of *target excess* as follows

$$\mathcal{X} = \{\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{a}} - \tilde{\boldsymbol{\tau}} \ : \ \tilde{\boldsymbol{a}} \in \mathcal{Y}\}.$$

Henceforth, we assume that targets are normalized to zeros. We first propose the properties which any MOS criterion should possess.

Definition 1 *A function $\rho : \mathcal{X} \mapsto [0,1]$, is an MOS criterion if it satisfies the following properties:*

1. **Monotonicity***: If $\tilde{\boldsymbol{x}} \geq \tilde{\boldsymbol{y}}$, then $\rho(\tilde{\boldsymbol{x}}) \geq \rho(\tilde{\boldsymbol{y}})$.*
2. **Attainment Content***: $\rho(\mathbf{0}) = 1$. If there exists $i \in N$, $\tilde{x}_i \geq 0$, then $\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) = \rho((0, \tilde{\boldsymbol{x}}_{-i}))$.*
3. **Non-abandonment***: If there exists $i \in N$, $\tilde{x}_i < 0$, then $\rho(\tilde{\boldsymbol{x}}) = 0$.*
4. **Right continuity***: $\lim_{a \downarrow 0} \rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) = \rho(\tilde{\boldsymbol{x}})$.*

Monotonicity implies that if an alternative $\tilde{\boldsymbol{x}}$ is almost surely better than $\tilde{\boldsymbol{y}}$ for all distributions in the family, then it would never be less preferred. Attainment content follows from the satisficing principle. Moreover, if there is an objective that always achieves its target, the MOS criterion would be insensitive to magnitude of its over achievement above the target. Non-abandonment states that an alternative is never more preferred if there is an objective that always fails its target. Right continuity implies that if all target excess are augmented with infinitesimally small but positive amounts, the satisficing level cannot be improved in the limit. This also implies that we exclude the consideration of $\mathbb{P}(\tilde{\boldsymbol{x}} > 0)$. Note that the success probability criterion, $\mathbb{P}(\tilde{\boldsymbol{x}} \geq 0)$ is consistent with this definition of the MOS criteria. Extending the result of Brown and Sim (2009), the MOS criteria characterized in Definition 1 can be equivalently described with a dual representation as follows:

Theorem 1 *A function $\rho : \mathcal{X} \mapsto \Re$ is an MOS criterion if and only if*

$$\rho(\tilde{\boldsymbol{x}}) = \begin{cases} \sup\{k \in (0,1) : \eta_k(\tilde{\boldsymbol{x}}) \leq 0\} & \textit{if feasible} \\ 0 & \textit{otherwise,} \end{cases} \tag{1}$$

where $\{\eta_k : k \in (0,1)\}$ is a family of dual functions that are non-decreasing in k and $\eta_k : \mathcal{X} \mapsto \Re$ satisfies the following properties:

1. *If $\tilde{\boldsymbol{x}} \geq \tilde{\boldsymbol{y}}$, then $\eta_k(\tilde{\boldsymbol{x}}) \leq \eta_k(\tilde{\boldsymbol{y}})$*
2. *For all $a \in \Re$, $\eta_k(\tilde{\boldsymbol{x}} + a\mathbf{1}) = \eta_k(\tilde{\boldsymbol{x}}) - a$*
3. *If there exists $i \in N$, $\tilde{x}_i \geq 0$, then $\eta_k((0, \tilde{\boldsymbol{x}}_{-i})) = \max\{0, \eta_k(\tilde{\boldsymbol{x}})\}$*
4. *If there exists $i \in N$, $\tilde{x}_i < 0$, then $\eta_k(\tilde{\boldsymbol{x}}) > 0$*
5. *$\eta_k(\mathbf{0}) = 0$.*

For an MOS criterion, ρ, the underlying dual function for $k \in (0,1)$ is given by

$$\eta_k(\tilde{\boldsymbol{x}}) = \inf\{a : \rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) \geq k\}. \tag{2}$$

Proof : We show that ρ under representation (1) is an MOS criterion if η_k , $k \in (0,1)$, is the family of dual functions with the properties stated in Theorem 1.

Monotonicity: If $\tilde{\boldsymbol{x}} \geq \tilde{\boldsymbol{y}}$, we have by monotonicity $\eta_k(\tilde{\boldsymbol{x}}) \leq \eta_k(\tilde{\boldsymbol{y}})$ for all $k \in (0,1)$. Since η_k is non-decreasing in k, $\rho(\tilde{\boldsymbol{x}}) \geq \rho(\tilde{\boldsymbol{y}})$.

Attainment content: Since $\eta_k(\mathbf{0}) = 0$ for all $k \in (0,1)$, we have $\rho(\mathbf{0}) = 1$. If there exists $i \in N$, $\tilde{x}_i \geq 0$, we will show that $\rho((0, \tilde{\boldsymbol{x}}_{-i})) = \rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}))$ as follows:

$$\begin{aligned}
\rho((0,\tilde{\boldsymbol{x}}_{-i})) &= \begin{cases} \sup\{k \in (0,1) : \eta_k((0,\tilde{\boldsymbol{x}}_{-i})) \leq 0\} & \text{if feasible} \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \sup\{k \in (0,1) : \max\{0, \eta_k((\tilde{x}_i,\tilde{\boldsymbol{x}}_{-i}))\} \leq 0\} & \text{if feasible} \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \sup\{k \in (0,1) : \eta_k((\tilde{x}_i,\tilde{\boldsymbol{x}}_{-i})) \leq 0\} & \text{if feasible} \\ 0 & \text{otherwise} \end{cases} \\
&= \rho((\tilde{x}_i,\tilde{\boldsymbol{x}}_{-i})),
\end{aligned}$$

where the second equality follows from Property 3.

Non-abandonment: If there exists $i \in N$, $\tilde{x}_i < 0$, then $\eta_k(\tilde{\boldsymbol{x}}) > 0$ for all $k > 0$. Hence, $\rho(\tilde{\boldsymbol{x}}) = 0$ under (1).

Right continuity: For right continuity, we need to show that for any $\epsilon > 0$ there exists $a > 0$ such that $\rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) \leq \rho(\tilde{\boldsymbol{x}}) + \epsilon = \bar{\rho}$, so that $\lim_{a \downarrow 0} \rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) = \rho(\tilde{\boldsymbol{x}})$. From (1), we have $\eta_{\bar{\rho}}(\tilde{\boldsymbol{x}}) > 0$ and hence, there exists $a > 0$ such that $\eta_{\bar{\rho}} > a > 0$. By monotonicity and translation invariance of the family of functions $\{\eta_k : k \in (0,1)\}$ we have,

$$\begin{aligned}
\rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) &= \begin{cases} \sup\{k \in (0,1) : \eta_k(\tilde{\boldsymbol{x}} + a\mathbf{1}) \leq 0\} & \text{if feasible} \\ 0 & \text{otherwise,} \end{cases} \\
&= \begin{cases} \sup\{k \in (0,1) : \eta_k(\tilde{\boldsymbol{x}}) - a \leq 0\} & \text{if feasible} \\ 0 & \text{otherwise,} \end{cases} \\
&\leq \begin{cases} \sup\{k \in (0,1) : \eta_k(\tilde{\boldsymbol{x}}) < \eta_{\bar{\rho}}(\tilde{\boldsymbol{x}})\} & \text{if feasible} \\ 0 & \text{otherwise,} \end{cases} \\
&\leq \bar{\rho} = \rho(\tilde{\boldsymbol{x}}) + \epsilon.
\end{aligned}$$

We now show the converse that under representation (2), the family of functions $\{\eta_k : k \in (0,1)\}$ satisfies the properties of Theorem 1.

Property 1: This follows trivially from the monotonicity property of ρ.

Property 2: For any $c \in \Re$,

$$\begin{aligned}
\eta_k(\tilde{\boldsymbol{x}} + c\mathbf{1}) &= \inf\{a : \rho(\tilde{\boldsymbol{x}} + (a+c)\mathbf{1}) \geq k\} \\
&= \inf\{a - c : \rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) \geq k\} \\
&= \eta_k(\tilde{\boldsymbol{x}}) - c
\end{aligned}$$

Property 3: By definition, $\eta_k((0,\tilde{\boldsymbol{x}}_{-i})) = \inf\{a : \rho((0,\tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}) \geq k\}$. Observe that from the non-abandonment property of ρ,

$$\rho((0,\tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}) = 0 < k$$

for all $a < 0$, $k \in (0,1)$. Hence, we have

$$\begin{aligned}\eta_k((0,\tilde{\boldsymbol{x}}_{-i})) &= \inf\{a : \rho((0,\tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}) \geq k\} \\ &= \inf\{a : \rho((0,\tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}) \geq k, a \geq 0\} \\ &= \inf\{a : \rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) \geq k, a \geq 0\} \\ &= \max\{0, \eta_k(\tilde{\boldsymbol{x}})\},\end{aligned}$$

where the last equality follows from $\rho(\tilde{\boldsymbol{x}} + a\mathbf{1})$ being a nondecreasing in a.

Property 4: Suppose there exists $i \in N$, $\tilde{x}_i < 0$, we will show that there exists $\epsilon > 0$ such that $\eta_k(\tilde{\boldsymbol{x}}) \geq \epsilon$, or equivalently, by Property 2 shown above, $\eta_k(\tilde{\boldsymbol{x}} + \mathbf{1}\epsilon) \geq 0$. Since $\tilde{x}_i < 0$, it implies that there exists $\epsilon > 0$ such that $\tilde{x}_i + \epsilon < 0$ and consequently, $\rho(\tilde{\boldsymbol{x}} + \mathbf{1}\epsilon) = 0$ from the non-abandonment property of ρ. Therefore, since $k \in (0,1)$

$$\eta_k(\tilde{\boldsymbol{x}} + \epsilon\mathbf{1}) = \inf\{a : \rho(\tilde{\boldsymbol{x}} + \epsilon\mathbf{1} + a\mathbf{1}) \geq k\} \geq 0.$$

Property 5: For any $k \in (0,1)$, we observe that $\rho(\mathbf{0} + a\mathbf{1}) = 1 \geq k$ for all $a \geq 0$ and that $\rho(\mathbf{0} + a\mathbf{1}) = 0$ for all $a < 0$. Hence, $\eta_k(\mathbf{0}) = 0$.

We next verify that under representation (2), we can recover ρ from representation (1). For this, we consider two cases. For the first case, there exists $k \in (0,1)$, such that $\eta_k(\tilde{\boldsymbol{x}}) \leq 0$. Observe that since $\rho(\tilde{\boldsymbol{x}} + a\mathbf{1})$ is right continuous with respect to a, the infimum in (2) is achievable. Hence,

$$\begin{aligned}&\sup\{k \in (0,1) : \eta_k(\tilde{\boldsymbol{x}}) \leq 0\} \\ =\ &\sup\{k \in (0,1) : \exists a \leq 0 \text{ such that } \rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) \geq k\} \\ =\ &\sup\{\rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) : a \leq 0\} \\ =\ &\rho(\tilde{\boldsymbol{x}}).\end{aligned}$$

For the second case, there does not exists $k \in (0,1)$ such that $\eta_k(\tilde{\boldsymbol{x}}) \leq 0$, which will lead to $\rho(\tilde{\boldsymbol{x}}) = 0$ under representation (1). Indeed, this is the case, since under representation (2), this condition is the same as the non existence of $a \leq 0$, such that $\rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) \geq k$ for any $k \in (0,1)$. Since $\max\{\rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) : a \leq 0\} = \rho(\tilde{\boldsymbol{x}})$, it also means that $\rho(\tilde{\boldsymbol{x}}) \not> 0$, or equivalently, $\rho(\tilde{\boldsymbol{x}}) = 0$.

■

2.1 Diversification favoring MOS criteria

Decision maker's preference for diversification is not only prevalent but fundamentally important for the mitigation of excessive losses arising from uncertainties in decision making. We can observe the impact of diversification via a simple example. Consider two uncertain positions (or scalar-valued target excess), $\tilde{x}$ and $\tilde{y}$. The following inequality shows the impact of diversification over these two uncertain positions:

$$\inf_{\omega\in\Omega}\{\lambda x(\omega) + (1-\lambda)y(\omega)\} \geq \inf_{\omega\in\Omega}\{x(\omega), y(\omega)\} \qquad \forall\lambda \in [0,1].$$

Hence, a diversified position has the effect of cushioning the impact of bad outcomes. We now extend the characterization of MOS to include diversification preference. For the i^{th} objective, diversification preference implies that if both $\tilde{x}_i$ and $\tilde{y}_i$ are preferred over $\tilde{r}_i$, then any convex combination of $\tilde{x}_i$ and $\tilde{y}_i$ is no less preferable to $\tilde{r}_i$. In the multiple objectives setting, we consider each objective is ranked by an autonomous agent that favors diversification. Hence, the MOS criteria that favors diversification are componentwise quasi-concave functions (see also Brown and Sim (2009)). Formally, we define a sub-class of diversification favoring MOS criteria as follows:

Definition 2 *A function ρ is a diversification favoring MOS criterion if it is a componentwise quasi-concave MOS criterion, i.e., for all $i \in N$, $(\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})$, $(\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}) \in \mathcal{X}$,*

$$\rho((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \min\{\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\} \qquad \forall \lambda \in [0,1].$$

Theorem 2 *A function ρ is a diversification favoring MOS criterion if and only if the underlying function in the dual representation, η_k, $k \in (0,1)$, is also componentwise quasi-convex.*

Proof : Suppose ρ is a diversification favoring MOS criterion, we will show that for any $k \in (0,1)$, the function

$$\eta_k(\tilde{\boldsymbol{x}}) = \inf\{a \ : \ \rho(\tilde{\boldsymbol{x}} + a\mathbf{1}) \geq k\},$$

is componentwise quasi-convex. Since, $\rho(\tilde{\boldsymbol{x}} + a\mathbf{1})$ is right continuous and nondecreasing in a, the infimum in the dual representation is achievable and hence, $\rho(\tilde{\boldsymbol{x}} + \eta_k(\tilde{\boldsymbol{x}})\mathbf{1}) \geq k$. For any objective, $i \in N$, we consider $(\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}), (\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}) \in \mathcal{X}$. Let $a^* = \max\{\eta_k((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \eta_k((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\}$ and we have $\min\{\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})+a^*\mathbf{1}), \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})+a^*\mathbf{1})\} \geq k$. Observe that for any $\lambda \in [0,1]$,

$$\begin{aligned}
\eta_k((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) &= \inf\{a : \rho(\lambda(\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}) + (1-\lambda)(\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}) \geq k\} \\
&= \inf\{a : \rho(\lambda((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}) + (1-\lambda)((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}) \geq k\} \\
&\leq \inf\{a : \min\{\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1}), \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}) + a\mathbf{1})\} \geq k\} \\
&\leq a^* = \max\{\eta_k((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \eta_k((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\}.
\end{aligned}$$

where the first inequality follows from quasi-concavity of ρ. Next, we show that if η_k is componentwise quasi-convex and nondecreasing in $k \in (0,1)$, then the function

$$\rho(\tilde{\boldsymbol{x}}) = \begin{cases} \sup\{k \in (0,1) : \eta_k(\tilde{\boldsymbol{x}}) \leq 0\} & \text{if feasible} \\ 0 & \text{otherwise,} \end{cases}$$

is componentwise quasi-concave. For a given objective $i \in N$ and $(\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}), (\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}) \in \mathcal{X}$, let

$$k^* = \min\{\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\}.$$

The result is trivially true if $k^* = 0$. Suppose $k \in (0, k^*)$, we note that $\eta_k((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) \leq 0$ and $\eta_k((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \leq 0$. Since η_k is componentwise quasi-convex and is non-decreasing in k, we have

$$\begin{aligned}
\rho\big(\lambda(\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}) + (1-\lambda)(\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})\big) &= \sup\{k \in (0,1) : \eta_k\left(\lambda(\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}) + (1-\lambda)(\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})\right) \leq 0\} \\
&\geq \sup\{k \in (0,1) : \max\{\eta_k((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \eta_k((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\} \leq 0\} \\
&\geq k^* \\
&= \min\{\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\} > 0.
\end{aligned}$$

■

For the case of single objective, the function in the dual representation of the diversification favoring MOS criteria, η_k, $k \in (0,1)$ becomes a well-known normalized convex risk measure (Föllmer and Schied, 2004) with the following properties:

1. **Monotonicity**: If $\tilde{x} \geq \tilde{y}$, then $\eta_k(\tilde{x}) \leq \eta_k(\tilde{y})$.

2. **Translation invariance**: For all $a \in \Re$, $\eta_k(\tilde{x} + a) = \eta_k(\tilde{x}) - a$.

3. **Convexity**: For all $\lambda \in [0, 1]$, $\eta_k(\lambda\tilde{x} + (1 - \lambda)\tilde{y}) \leq \lambda\eta_k(\tilde{x}) + (1 - \lambda)\eta_k(\tilde{y})$.

4. **Normalization**: $\eta_k(0) = 0$.

Observe that for the case of single objective, the function Properties 3 and 4 in Theorem 1 are implied by normalization. Convexity of η_k follows directly from translation invariance and quasi-convexity as follows:

$$
\begin{aligned}
& \eta_k(\lambda\tilde{x} + (1 - \lambda)\tilde{y}) - \lambda\eta_k(\tilde{x}) - (1 - \lambda)\eta_k(\tilde{y}) \\
= \; & \eta_k\big(\lambda\tilde{x} + (1 - \lambda)\tilde{y} + \lambda\eta_k(\tilde{x}) + (1 - \lambda)\eta_k(\tilde{y})\big) && : \text{translation invariance} \\
= \; & \eta_k\big(\lambda(\tilde{x} + \eta_k(\tilde{x})) + (1 - \lambda)(\tilde{y} + \eta_k(\tilde{y}))\big) \\
\leq \; & \max\{\eta_k(\tilde{x} + \eta_k(\tilde{x})), \eta_k(\tilde{y} + \eta_k(\tilde{y}))\} && : \text{quasi-convexity} \\
= \; & 0,
\end{aligned}
$$

for all $\lambda \in [0, 1]$. The dual relationship between the convex risk measure and the satisficing criterion has been established in Brown and Sim (2009) for the single objective case. We extend this relationship to the case of multiple objectives, where the function in dual representation of the diversification favoring MOS criteria can be viewed as a multivariate extension of the convex risk measure.

3 Shortfall-aware MOS (S-MOS) criterion

The success probability, $\mathbb{P}(\tilde{\boldsymbol{x}} \geq 0)$, is a natural form of MOS criteria. However, it is well known that success probability is indifferent to the magnitude of shortfalls against the targets. Indeed, the success probability is not a diversification favoring MOS criterion. In this section, we aim to propose a new MOS criterion that approximates success probability and at the same time favors diversification.

Before introducing our criterion, we first provide the insights from the perspective of lower bounding the success probability by a family of concave utility functions. Observe that the success probability criterion can be expressed as an expectation over a step function, i.e.,

$$\mathbb{P}(\tilde{\boldsymbol{x}} \geq 0) = \mathbb{E}_{\mathbb{P}}(s(\tilde{\boldsymbol{x}}))$$

where the function $s : \Re^n \mapsto \Re$ is a step-function given by

$$s(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} \geq \mathbf{0} \\ 0 & \text{otherwise.} \end{cases}$$

From a utility perspective, the step function has the undesirable properties of being non-concave and its insensitivity to the degrees of shortfalls. On the other hand, these issues may be addressed by taking expectation over any concave nondecreasing function, $h : \Re^n \mapsto \Re$. Moreover, if the function, h is dominated by the step function, s, i.e., $h(\boldsymbol{x}) \leq s(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Re^n$, then $\mathbb{E}_{\mathbb{P}}(h(\tilde{\boldsymbol{x}}))$ is a lower bound of success probability, i.e.,

$$\mathbb{E}_{\mathbb{P}}(h(\tilde{\boldsymbol{x}})) \leq \mathbb{E}_{\mathbb{P}}(s(\tilde{\boldsymbol{x}})) = \mathbb{P}(\tilde{\boldsymbol{x}} \geq 0).$$

We consider a criterion that selects the best function within the following family of concave functions,

$$\mathcal{H} = \{h : \Re^n \mapsto \Re \mid h \text{ is concave, non-decreasing and } h(\boldsymbol{x}) \leq s(\boldsymbol{x}), \forall \boldsymbol{x} \in \Re^n\}.$$

The following result shows how we can optimize over the family of functions, $\mathcal{H}$ to obtain the tightest bound on the success probability criterion.

Proposition 1

$$\sup_{h\in\mathcal{H}} \mathbb{E}_{\mathbb{P}}(h(\tilde{\boldsymbol{x}})) = \sup_{\boldsymbol{u}\geq\boldsymbol{0}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N}\{u_i\tilde{x}_i, 1\}\right)$$

Proof : Note that for any $\boldsymbol{u}\geq\boldsymbol{0}$, the function $r(\boldsymbol{x}) = \min_{i\in N}\{u_i x_i, 1\}$ belongs to the family $\mathcal{H}$. Hence, it suffices to show that $\forall h\in\mathcal{H}$, there exists $\boldsymbol{u}\geq\boldsymbol{0}$ so that:

$$h(x) \leq \min_{i\in N}\{u_i x_i, 1\} \quad \forall \boldsymbol{x}\in\Re^n$$

or equivalently

$$\begin{aligned} h(\boldsymbol{x}) &\leq u_i x_i \quad \forall \boldsymbol{x}\in\Re^n, i\in N \\ h(\boldsymbol{x}) &\leq 1 \qquad \forall \boldsymbol{x}\in\Re^n. \end{aligned} \tag{3}$$

The last inequality is trivially true since, $h(\boldsymbol{x})\leq s(\boldsymbol{x})\leq 1$ for all $\boldsymbol{x}\in\Re^n$. Let $g_i(x_i) = \lim_{v\uparrow\infty} h\left((x_i, v\boldsymbol{1}_{-i})\right)$. Note that g_i is also concave and nondecreasing. Since h is nondecreasing, we have

$$h(\boldsymbol{x}) \leq g_i(x_i).$$

Let u_i be a sub-gradient of g_i at the origin. Since g_i is nondecreasing and concave, we have $u_i\geq 0$ and that

$$u_i\left(x_i - 0\right) \geq g_i(x_i) - g_i(0) \geq g_i(x_i) \quad \forall x_i\in\Re.$$

The last inequality follows from the observation that $g_i(0)\leq 0$. This is true since if $g_i(0) > 0$, then by concavity of g_i on domain $\Re$, there must exist $\epsilon > 0$ small enough so that $0 < g_i(0-\epsilon) = \lim_{v\uparrow\infty} h\left((-\epsilon, v\boldsymbol{1}_{-i})\right)$. This clearly contradicts $h\left((-\epsilon, v\boldsymbol{1}_{-i})\right) \leq s\left((-\epsilon, v\boldsymbol{1}_{-i})\right) = 0\ \forall v\in\Re$. Hence, combining the above, we have $h(\boldsymbol{x})\leq g_i(x_i)\leq u_i x_i$, and $h(\boldsymbol{x})\leq 1$ for all $\boldsymbol{x}\in\Re^n$. Applying the argument for each $i\in N$, the desired result follows. ■

Inspired by the bound on success probability via concave functions, we define the shortfall-aware function (SAF) as follows:

Definition 3 *The shortfall aware function (SAF), β, is defined as follows:*

$$\beta(\tilde{\boldsymbol{x}}) = \sup_{\boldsymbol{u}\geq\boldsymbol{0}} \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N}\{u_i\tilde{x}_i, 1\}\right)$$

For the case of single objective ($n = 1$), the SAF is the shortfall aspiration measure criterion proposed by Chen and Sim (2009). When compared to step utility function, which is oblivious to the magnitude of targets' shortfalls, the SAF reflects penalization against the degree of targets' shortfalls. Since, the multipliers u_i are nonnegative, whenever $\tilde{x}_i(\omega) < 0$, the level of target shortfall may contribute to a reduction in the SAF value. The SAF also reflects ambiguity aversion so that for any $\mathbb{P}\in\mathbb{F}$

$$\beta(\tilde{\boldsymbol{x}}) \leq \mathbb{P}(\tilde{\boldsymbol{x}}\geq\boldsymbol{0}).$$

Observe that since $\beta(\boldsymbol{0}) = 0$, the SAF is not an MOS criterion. However, we will show that the SAF shares similar properties to a diversification favoring MOS criterion. For convenience, we henceforth define the index set

$$N(\tilde{\boldsymbol{x}}) \;=\; \{i\in N \;:\; \tilde{x}_i \ngeq 0\}.$$

Proposition 2 *The shortfall-aware function (SAF), $\beta : \mathcal{X} \mapsto [0,1]$ satisfies the following properties:*

1. **Monotonicity**
2. **Strict attainment content**: $\beta(\mathbf{1}) = 1$. *If there exists* $i \in N$, $\tilde{x}_i > 0$, *then* $\beta((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) = \beta((1, \tilde{\boldsymbol{x}}_{-i}))$.
3. **Non-abandonment**
4. **Componentwise quasi-concave**
5. **Componentwise scale invariant**: $\beta(k_1\tilde{x}_1, \ldots, k_n\tilde{x}_n) = \beta(\tilde{\boldsymbol{x}})$ *for all* $\boldsymbol{k} > \mathbf{0}$
6. **Restricted right continuity**: $\lim_{a \downarrow 0} \beta((\tilde{x}_i + a, \tilde{\boldsymbol{x}}_{-i})) = \beta(\tilde{\boldsymbol{x}})$, *for all* $i \in N(\tilde{\boldsymbol{x}})$

Proof : Note that $\beta(\tilde{\boldsymbol{x}}) \leq 1$ and

$$\beta(\tilde{\boldsymbol{x}}) = \sup_{\boldsymbol{u} \geq \mathbf{0}} \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{u_i \tilde{x}_i, 1\} \right) \geq \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{0 \tilde{x}_i, 1\} \right) = 0.$$

Hence, $\beta(\tilde{\boldsymbol{x}}) \in [0,1]$. We next show that the following properties are satisfied. For convenience, denote the index set $N_i = N \setminus \{i\}$.

1. **Monotonicity**: This is straightforward.
2. **Strict attainment content**: Suppose $\tilde{x}_i > 0$, then there exists $u_i > 0$ such that $u_i \tilde{x}_i > 1$. Hence,

$$\begin{aligned} \beta((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) &= \sup_{\boldsymbol{u} \geq 0} \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{j \in N_i} \{u_j \tilde{x}_j, u_i \tilde{x}_i, 1\} \right) \\ &= \sup_{\boldsymbol{u} \geq 0} \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{j \in N_i} \{u_j \tilde{x}_j, 1\} \right) \\ &= \beta((1, \tilde{\boldsymbol{x}}_{-i})). \end{aligned}$$

   For the case when $\tilde{\boldsymbol{x}} > 0$, it is easy to establish that $\beta(\tilde{\boldsymbol{x}}) = 1$.

3. **Non-abandonment:** Suppose there exists $\tilde{x}_i < 0$ for some i, then

$$\beta((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) = \sup_{\boldsymbol{u} \geq 0} \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{j \in N_i} \{u_j \tilde{x}_j, u_i \tilde{x}_i, 1\} \right) \leq 0,$$

   in which the supremum is achieved at $\boldsymbol{u} = \mathbf{0}$. Hence, $\beta(\tilde{\boldsymbol{x}}) = 0$.

4. **Componentwise quasiconcavity**: Let $\beta^* = \min\{\beta((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \beta((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\}$. We will show that for any $\lambda \in [0,1]$,

$$\beta((\lambda \tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \beta^*.$$

   The result is trivial if $\lambda \in \{0,1\}$. Hence, we consider $\lambda \in (0,1)$. Observe that for any $\epsilon > 0$, there exists $\boldsymbol{u}, \boldsymbol{v} > 0$ such that

$$\inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{j \in N_i} \{u_i \tilde{x}_i, u_j \tilde{x}_j, 1\} \right) \geq \beta^* - \epsilon$$

   and

$$\inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{j \in N_i} \{v_i \tilde{y}_i, v_j \tilde{x}_j, 1\} \right) \geq \beta^* - \epsilon.$$

Let $\gamma = \lambda v_i/(u_i(1-\lambda)+\lambda v_i)$ and $p_j = \gamma u_j + (1-\gamma)v_j$ for all $j \in N$. Observe that $\gamma \in (0,1)$ and $p_j > 0$ for all $j \in N$. Moreover,

$$p_i(\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i) = \gamma u_i \tilde{x}_i + (1-\gamma)v_i\tilde{y}_i$$

and

$$p_j\tilde{z}_j = \gamma u_j\tilde{x}_j + (1-\gamma)v_j\tilde{x}_j \qquad \forall j \in N_i.$$

Hence, noting that the function $\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_j\{\tilde{x}_j\}\right)$ is concave with respect to $\tilde{\boldsymbol{x}}$, we have

$$\begin{aligned}
& \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{p_i(\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i), p_j\tilde{x}_j, 1\}\right) \\
= & \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{\gamma u_i\tilde{x}_i + (1-\gamma)v_i\tilde{y}_i, \gamma u_j\tilde{x}_j + (1-\gamma)v_j\tilde{x}_j, 1\}\right) \\
\geq & \gamma \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i\tilde{x}_i, u_j\tilde{x}_j, 1\}\right) + (1-\gamma)\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{v_i\tilde{y}_i, v_j\tilde{x}_j, 1\}\right) \\
\geq & \beta^* - \epsilon.
\end{aligned}$$

Hence, $\beta((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \beta^*$.

5. **Componentwise scale invariant**: The result is obvious since the multiplier, u_i lies in a cone.

6. **Restricted right continuity**: For any $i \in N(\tilde{\boldsymbol{x}})$ there exists $\mathbb{P} \in \mathbb{F}$ such that $\mathbb{P}(\tilde{x}_i < 0) > 0$ and so $\mathbb{P}(\tilde{x}_i + \epsilon < 0) > 0$ for some $\epsilon > 0$. Hence, for all $a \in [0,\epsilon]$

$$\begin{aligned}
\lim_{u_i\uparrow\infty}\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i(\tilde{x}_i + a), u_j\tilde{x}_j, 1\}\right) & \leq \lim_{u_i\uparrow\infty}\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min\{u_i(\tilde{x}_i + a), 1\}\right) \\
& \leq \lim_{u_i\uparrow\infty}\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min\{u_i(\tilde{x}_i + \epsilon), 1\}\right) \\
& \leq \lim_{u_i\uparrow\infty} u_i \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min\{(\tilde{x}_i + \epsilon), 1/u_i\}\right) \\
& \leq -\infty.
\end{aligned}$$

Let $\bar{u}_i$ such that $\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min\{\bar{u}_i(\tilde{x}_i + \epsilon), 1\}\right) < 0$. Since, the function

$$f(u) = \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min\{u(\tilde{x}_i + \epsilon), 1\}\right)$$

is concave in u with $f(0) = 0$, $f(\infty) = -\infty$, and $f(\bar{u}_i) < 0$, we must have $f(u) < 0$ for all $u \geq \bar{u}_i$. Hence, for all $a \in [0,\epsilon]$, $u_i \geq \bar{u}_i$

$$\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i(\tilde{x}_i + a), u_j\tilde{x}_j, 1\}\right) \leq \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min\{u_i(\tilde{x}_i + \epsilon), 1\}\right) < 0.$$

This means that the optimum solution to

$$\sup_{\boldsymbol{u}\geq 0}\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i(\tilde{x}_i + a), u_j\tilde{x}_j, 1\}\right)$$

must require $u_i \leq \bar{u}_i$ for all $a \in [0,\epsilon]$. Hence, for any $i \in N(\tilde{\boldsymbol{x}})$

$$\begin{aligned}
\beta((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) & \leq \lim_{a\downarrow 0}\beta((\tilde{x}_i + a, \tilde{\boldsymbol{x}}_{-i})) \\
& = \lim_{a\downarrow 0}\sup_{\boldsymbol{u}\geq 0}\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i(\tilde{x}_i + a), u_j\tilde{x}_j, 1\}\right) \\
& \leq \lim_{a\downarrow 0}\sup_{\boldsymbol{u}\geq 0}\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i\tilde{x}_i + a\bar{u}_i, u_j\tilde{x}_j, 1\}\right) \\
& \leq \lim_{a\downarrow 0}\left(\sup_{\boldsymbol{u}\geq 0}\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i\tilde{x}_i, u_j\tilde{x}_j, 1\}\right) + a\bar{u}_i\right) \\
& = \beta((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})).
\end{aligned}$$

■

We next show that although SAF is not an MOS criterion, it can be easily transformed to an diversification favoring MOS criterion in the following result.

Theorem 3 *Suppose a function $\beta : \mathcal{X} \mapsto [0,1]$ that satisfies the following properties in Proposition 2, then the function $\rho : \mathcal{X} \mapsto [0,1]$*

$$\rho(\tilde{\boldsymbol{x}}) = \beta(\hat{\boldsymbol{x}})$$

where we replace the 'tilde' with 'hat' to denote

$$\hat{x}_i = \begin{cases} \tilde{x}_i & \text{if } i \in N(\tilde{\boldsymbol{x}}) \\ 1 & \text{otherwise,} \end{cases}$$

is a diversification favoring MOS criterion.

Proof :

1. **Monotonicity**: Observe that if $\tilde{\boldsymbol{y}} \geq \tilde{\boldsymbol{x}}$, then $N(\tilde{\boldsymbol{y}}) \subseteq N(\tilde{\boldsymbol{x}})$ and the results follows from monotonicity and strict attainment content properties of β.

2. **Attainment content**: This trivially true from the definition of $N(\tilde{\boldsymbol{x}})$.

3. **Non-abandonment**: If there exists $i \in N$, $\tilde{x}_i < 0$, then $i \in N(\tilde{\boldsymbol{x}})$ and the result follows from the non-abandonment property of β.

4. **Right-continuity**: Observe that there exists some small $\epsilon > 0$ such that the index set $N(\tilde{\boldsymbol{x}} + a\mathbf{1}) = N(\tilde{\boldsymbol{x}})$ for all $a \in [0, \epsilon]$. Hence, right continuity follows from the restricted right continuity property of β.

5. **Componentwise quasi-concavity**: We need to show that

$$\rho((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \min\{\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\}$$

for all $\lambda \in [0,1]$, which is trivial for $\lambda \in \{0,1\}$. We hence consider $\lambda \in (0,1)$. Suppose, $\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i \geq 0$, then

$$\rho((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) = \beta((1, \hat{\boldsymbol{x}}_{-i}))$$

which is at least as large as $\beta((\hat{x}_i, \hat{\boldsymbol{x}}_{-i}))$ or $\beta((\hat{y}_i, \hat{\boldsymbol{x}}_{-i}))$. Suppose, $\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i \ngeq 0$, then we consider the two possible cases. The first case being $\tilde{x}_i, \tilde{y}_i \ngeq 0$, we have by quasiconcavity of β

$$\begin{aligned} & \rho((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \\ = \; & \beta((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \hat{\boldsymbol{x}}_{-i})) \\ \geq \; & \min\{\beta((\tilde{x}_i, \hat{\boldsymbol{x}}_{-i})), \beta((\tilde{y}_i, \hat{\boldsymbol{x}}_{-i}))\} \\ = \; & \min\{\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\}. \end{aligned}$$

The next case being $\tilde{x}_i \geq 0$ and $\tilde{y}_i \ngeq 0$, in which case

$$\rho((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) = \beta((1, \hat{\boldsymbol{x}}_{-i})) \geq \beta((\tilde{y}_i, \hat{\boldsymbol{x}}_{-i})) = \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})),$$

and it suffices to show that $\rho((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))$. Indeed, since $\tilde{x}_i \geq 0$, by monotonicity of β,

$$
\begin{aligned}
& \rho((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \\
= \; & \beta((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \hat{\boldsymbol{x}}_{-i})) \\
\geq \; & \beta(((1-\lambda)\tilde{y}_i, \hat{\boldsymbol{x}}_{-i})) \\
= \; & \beta((\tilde{y}_i, \hat{\boldsymbol{x}}_{-i})) \\
= \; & \rho((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))
\end{aligned}
$$

where the second last equality is due to the componentwise scale invariant property of β.

■

In view of Theorem 3, we define the shortfall-aware MOS (S-MOS) criterion as follows:

Definition 4 *The shortfall aware MOS (S-MOS) criterion, α, is defined as follows:*

$$
\alpha(\tilde{\boldsymbol{x}}) = \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N(\tilde{\boldsymbol{x}})} \{u_i \tilde{x}_i, 1\} \right)
$$

3.1 Tractable robust approximation of S-MOS criterion

If probability distributions of the uncertainty are known, computing the SAF requires the evaluation of expectation operators, which involves multidimensional integration and is generally computationally intractable; see for instance, Nemirovski and Shapiro (2006). Nevertheless, optimizing the parameters, $\boldsymbol{u}$ in SAF can easily be posed as a standard stochastic optimization problem, which we can approximate using sample average approximation (SAA). However, under distributional ambiguity, computing the value of SAF requires the ability to evaluate the worst-case expectation over a family of distributions and then optimize it over the parameters, $\boldsymbol{u}$. The aim is to provide a tractable approximation of SAF under distributional ambiguity and we focus on the case where the uncertain target excess are affinely dependent on m given random variables or factors, $(\tilde{z}_1, \ldots, \tilde{z}_m)$ on Ω, i.e,

$$
\mathcal{X} \subseteq \mathcal{A} = \left\{ \tilde{\boldsymbol{a}} : \exists a_i^j \in \Re, i \in N, j \in M \mid \tilde{a}_i(\omega) = a_i^0 + \sum_{j \in M} a_i^j \tilde{z}_j(\omega) \quad \forall \omega \in \Omega, i \in N, j \in M \right\},
$$

where $M = \{1, \ldots, m\}$. The factor model of uncertainty is common in robust optimization literature in which the descriptive statistics of these factors are specified and form the basis for characterizing the family of distributions, $\mathbb{F}$; see for instance, Ben Tal and Nemirovski (1998); Chen et al. (2008); Goh and Sim (2010). In this section, we will focus on $\tilde{\boldsymbol{z}}$ being described by its support $\mathcal{W}$, mean $\boldsymbol{\mu}$ and covariance, $\boldsymbol{\Sigma}$ so that the family of distributions is given as follows:

$$
\mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathcal{W}) = \{\mathbb{P} : \mathbb{E}_{\mathbb{P}}(\tilde{\boldsymbol{z}}) = \boldsymbol{\mu}, \mathbb{E}_{\mathbb{P}}(\tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}') = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}', \mathbb{P}(\tilde{\boldsymbol{z}} \in \mathcal{W}) = 1\}.
$$

For computational reasons, we assume that the support $\mathcal{W}$ is a tractable conic representable, closed, convex and bounded set with non empty interior as proposed by Ben Tal and Nemirovski (1998). In order to compute the index set $N(\tilde{\boldsymbol{x}})$ as required for S-MOS, we assume the mean $\boldsymbol{\mu}$ and covariance, $\boldsymbol{\Sigma}$ do not further constrain the support set $\mathcal{W}$ so that for every $\boldsymbol{z} \in \mathcal{W}$ there exists $\mathbb{P} \in \mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathcal{W})$ such that $\mathbb{P}(\tilde{\boldsymbol{z}} = \boldsymbol{z}) > 0$. Hence, $y + \boldsymbol{x}'\tilde{\boldsymbol{z}} \geq 0$ if and only if $y + \min\{\boldsymbol{x}'\boldsymbol{z} \;:\; \boldsymbol{z} \in \mathcal{W}\} \geq 0$, which can easily be computed under our assumption of support set, $\mathcal{W}$.

Under the factor model of uncertainty, the computational tractability for evaluating the worst-case expectation over a family of distributions depends on the specification of distributional ambiguity. Unfortunately, for any $\tilde{\boldsymbol{a}} \in \mathcal{A}$, $b \in \Re$ the problem

$$\inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N}\{\tilde{a}_i, b\}\right) \tag{4}$$

is generally NP-hard, as in the case when the family of distributions $\mathbb{F}$ is given by $\mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathcal{W})$; see Murty and Kabadi (1987). Nevertheless, there are tractable examples as follows:

Proposition 3 *Let*

$$\mathbb{F}(\boldsymbol{\mu}, \mathcal{W}) = \{\mathbb{P} : \mathbb{E}_{\mathbb{P}}(\tilde{\boldsymbol{z}}) = \boldsymbol{\mu}, \mathbb{P}(\{\tilde{\boldsymbol{z}} \in \mathcal{W}\}) = 1\}$$

and

$$\mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \{\mathbb{P} : \mathbb{E}_{\mathbb{P}}(\tilde{\boldsymbol{z}}) = \boldsymbol{\mu}, \mathbb{E}_{\mathbb{P}}(\tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}') = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}'\}.$$

Then, for any $\tilde{\boldsymbol{a}} \in \mathcal{A}$, i.e., $\tilde{a}_i = a_i^0 + \boldsymbol{a}_i'\tilde{\boldsymbol{z}}$ for some $a_i^0 \in \Re$, $\boldsymbol{a}_i \in \Re^m$, $i \in N$, we have

$$\pi_A(\tilde{\boldsymbol{a}}, b) = \inf_{\mathbb{P}\in\mathbb{F}(\boldsymbol{\mu},\mathcal{W})} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N}\{\tilde{a}_i, b\}\right)$$

$$\pi_B(\tilde{\boldsymbol{a}}, b) = \inf_{\mathbb{P}\in\mathbb{F}(\boldsymbol{\mu},\boldsymbol{\Sigma})} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N}\{\tilde{a}_i, b\}\right)$$

where the functions $\pi_A : \mathcal{A} \times \Re \mapsto \Re$ and $\pi_B : \mathcal{A} \times \Re \mapsto \Re$ are given by:

$$\begin{aligned}
\pi_A(\tilde{\boldsymbol{a}}, b) = \sup \quad & v + \boldsymbol{\mu}'\boldsymbol{r} \\
\text{s.t.} \quad & v + \max_{\boldsymbol{z}\in\mathcal{W}} (\boldsymbol{r} - \boldsymbol{a}_i)'\boldsymbol{z} \le a_i^0 \quad i \in N \\
& v + \max_{\boldsymbol{z}\in\mathcal{W}} \boldsymbol{r}'\boldsymbol{z} \le b \\
& v \in \Re, \boldsymbol{r} \in \Re^m.
\end{aligned}$$

and

$$\begin{aligned}
\pi_B(\tilde{\boldsymbol{a}}, b) = \sup \quad & s + \boldsymbol{t}'\boldsymbol{\mu} + \text{tr}(\boldsymbol{S}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}')) \\
\text{s.t.} \quad & \begin{pmatrix} a_i^0 - s & \frac{1}{2}(\boldsymbol{a}_i - \boldsymbol{t})' \\ \frac{1}{2}(\boldsymbol{a}_i - \boldsymbol{t}) & -\boldsymbol{S} \end{pmatrix} \in \mathbb{S}_+^{m+1} \quad \forall i \in N \\
& \begin{pmatrix} b - s & -\frac{1}{2}\boldsymbol{t}' \\ -\frac{1}{2}\boldsymbol{t} & -\boldsymbol{S} \end{pmatrix} \in \mathbb{S}_+^{m+1} \\
& s \in \Re, \boldsymbol{S} \in \mathbb{S}^m,
\end{aligned}$$

where $\mathbb{S}^m$ (respectively, $\mathbb{S}_+^m$) denotes the set of symmetric matrices (respectively, symmetric positive semidefinite matrices) in $\Re^{m\times m}$.

Proof :

Applying the strong duality results of Isii (1963), the problem

$$\begin{aligned}
\inf_{\mathbb{P}} \quad & \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N}\{a_i^0 + \boldsymbol{a}_i'\tilde{\boldsymbol{z}}, b\}\right) \\
\text{s.t.} \quad & \mathbb{E}_{\mathbb{P}}(\tilde{\boldsymbol{z}}) = \boldsymbol{\mu} \\
& \mathbb{P}(\{\tilde{\boldsymbol{z}} \in \mathcal{W}\}) = 1
\end{aligned}$$

achieves the same optimal objective as

$$\begin{array}{rl} \sup & v + \boldsymbol{\mu}'\boldsymbol{r} \\ \text{s.t.} & v + \boldsymbol{r}'\boldsymbol{z} \geq \min_{i \in N}\{a_i^0 + \boldsymbol{a}_i'\boldsymbol{z}, b\} \quad \forall \boldsymbol{z} \in \mathcal{W} \\ & v \in \Re, \boldsymbol{r} \in \Re^m. \end{array}$$

Hence, the first result follows. Similarly, for the second result, we also apply the strong duality results to obtain the following dual problem

$$\begin{array}{rl} \sup & s + \boldsymbol{t}'\boldsymbol{\mu} + \mathrm{tr}(\boldsymbol{S}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}')) \\ \text{s.t.} & s + \boldsymbol{t}'\boldsymbol{z} + \boldsymbol{z}'\boldsymbol{S}\boldsymbol{z} \leq a_i^0 + \boldsymbol{a}_i'\boldsymbol{z} \quad \forall \boldsymbol{z} \in \Re^m \quad i \in N \\ & s + \boldsymbol{t}'\boldsymbol{z} + \boldsymbol{z}'\boldsymbol{S}\boldsymbol{z} \leq b \quad \forall \boldsymbol{z} \in \Re^m \\ & s \in \Re, \boldsymbol{t} \in \Re^m. \end{array}$$

Observe that each of the $i \in N$ constraint is equivalent to

$$\begin{pmatrix} 1 \\ \boldsymbol{z} \end{pmatrix}' \begin{pmatrix} s - a_i^0 & \frac{1}{2}(\boldsymbol{t} - \boldsymbol{a}_i)' \\ \frac{1}{2}(\boldsymbol{t} - \boldsymbol{a}_i) & \boldsymbol{S} \end{pmatrix} \begin{pmatrix} 1 \\ \boldsymbol{z} \end{pmatrix} \leq 0 \ \forall \boldsymbol{z} \in \Re^m$$

which can be trivially cast as a positive semi-definite conic constraint. Hence, the result follows. ■

Remark : The optimization problem to evaluate π_A requires the explicit formation of the constraints $v + \max_{\boldsymbol{z} \in \mathcal{W}}(\boldsymbol{r} - \boldsymbol{a}_i)'\boldsymbol{z} \leq a_i^0$, $i \in N$ and $v + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{r}'\boldsymbol{z} \leq b$ which are known as the *robust counterparts* in the context of robust optimization. Again, under our assumption of support set $\mathcal{W}$ the robust counterparts can be transformed to a small set of tractable conic constraints. We refer interested readers to Ben Tal and Nemirovski (1998); El Ghaoui et al. (2003); Bertsimas and Sim (2004).

Observe that the functions π_A and π_B are concave functions in their inputs and the corresponding SAFs are as follows:

$$\begin{array}{rl} \beta_A\left(x_1^0 + \boldsymbol{x}_1'\tilde{\boldsymbol{z}}, \ldots, x_n^0 + \boldsymbol{x}_n'\tilde{\boldsymbol{z}}\right) = \sup & v + \boldsymbol{\mu}'\boldsymbol{r} \\ \text{s.t.} & v + \max_{\boldsymbol{z} \in \mathcal{W}}(\boldsymbol{r} - u_i\boldsymbol{x}_i)'\boldsymbol{z} \leq u_i x_i^0 \quad i \in N \\ & v + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{r}'\boldsymbol{z} \leq 1 \\ & v \in \Re, \boldsymbol{r} \in \Re^m. \end{array}$$

and

$$\begin{array}{rl} \beta_B\left(x_1^0 + \boldsymbol{x}_1'\tilde{\boldsymbol{z}}, \ldots, x_n^0 + \boldsymbol{x}_n'\tilde{\boldsymbol{z}}\right) = \sup & s + \boldsymbol{t}'\boldsymbol{\mu} + \mathrm{tr}(\boldsymbol{S}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}')) \\ \text{s.t.} & \begin{pmatrix} u_i x_i^0 - s & \frac{1}{2}(u_i\boldsymbol{x}_i - \boldsymbol{t})' \\ \frac{1}{2}(u_i\boldsymbol{x}_i - \boldsymbol{t}) & -\boldsymbol{S} \end{pmatrix} \in \mathbb{S}_+^{m+1} \quad \forall i \in N \\ & \begin{pmatrix} 1 - s & -\frac{1}{2}\boldsymbol{t}' \\ -\frac{1}{2}\boldsymbol{t} & -\boldsymbol{S} \end{pmatrix} \in \mathbb{S}_+^{m+1} \\ & \boldsymbol{u} \geq \boldsymbol{0} \\ & \boldsymbol{u} \in \Re^n, s \in \Re, \boldsymbol{S} \in \mathbb{S}^m, \end{array}$$

which are tractable conic optimization problems.

Interestingly, although $\mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathcal{W})$ is the intersection of the families $\mathbb{F}(\boldsymbol{\mu}, \mathcal{W})$ and $\mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the corresponding SAF remains intractable. We therefore propose a robust approximation for this case. Our aim is to develop an approximate SAF, $\bar{\beta} : \mathcal{A} \mapsto [0,1]$ that satisfies the properties of Proposition 2, so that the $\bar{\beta}$ can be transformed to a diversification favoring MOS criterion via Theorem 3. Indeed, this depends on the properties of the function, π that we use to approximate Problem (4).

Theorem 4 *Let $\pi : \mathcal{A} \times \Re \mapsto \Re$ be a positively homogeneous, tractable and concave function such that*

1. *For all $\tilde{\boldsymbol{a}} \in \mathcal{A}$, $b \in \Re$, $\pi(\tilde{\boldsymbol{a}}, b) \leq \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{\tilde{a}_i, b\} \right)$*
2. *For all $\boldsymbol{a} \in \Re^{n+1}$, $\pi(\boldsymbol{a}) = \min_{i \in \{1,\ldots,n+1\}} \{a_i\}$.*
3. *For all $\tilde{\boldsymbol{a}} \geq \boldsymbol{0}$, $\tilde{\boldsymbol{a}} \in \mathcal{A}$ and $b \geq 0$, $\pi(\tilde{\boldsymbol{a}}, b) \geq 0$*

The following function, $\bar{\beta} : \mathcal{A} \mapsto \Re$

$$\bar{\beta}(\tilde{\boldsymbol{x}}) = \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \pi(u_1 \tilde{x}_1, \ldots, u_n \tilde{x}_n, 1)$$

satisfies the properties of Proposition 2. Moreover, for all $\tilde{\boldsymbol{x}} \in \mathcal{A}$,

$$\bar{\beta}(\tilde{\boldsymbol{x}}) \leq \beta(\tilde{\boldsymbol{x}}).$$

Proof : The first property of the π function, leads to bound $\bar{\beta}(\tilde{\boldsymbol{x}}) \leq \beta(\tilde{\boldsymbol{x}})$ for all $\tilde{\boldsymbol{x}} \in \mathcal{A}$. We go through the properties of Proposition 2 as follows:

1. **Monotonicity**: For any $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}} \in \mathcal{A}$, such that $\tilde{\boldsymbol{x}} \geq \tilde{\boldsymbol{y}}$, we will show that $\bar{\beta}(\tilde{\boldsymbol{x}}) \geq \bar{\beta}(\tilde{\boldsymbol{y}})$. For this, we show that the function π abides by monotonicity, where $\pi(\tilde{\boldsymbol{x}}, a) \geq \pi(\tilde{\boldsymbol{y}}, b)$ for $\tilde{\boldsymbol{x}} \geq \tilde{\boldsymbol{y}}$, $a \geq b$ as follows:

$$\begin{aligned} \pi(\tilde{\boldsymbol{x}}, a) &= \pi(\tilde{\boldsymbol{y}} + (\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{y}}), b + (a - b)) \\ &\geq \pi(\tilde{\boldsymbol{y}}, b) + \underbrace{\pi((\tilde{\boldsymbol{x}} - \tilde{\boldsymbol{y}}), (a - b))}_{\geq 0} \\ &\geq \pi(\tilde{\boldsymbol{y}}, b). \end{aligned}$$

   The first inequality is due to the fact that π is concave and positively homogeneous, hence, it is also superadditive. The second inequality is due to Property 3.

2. **Strict attainment content**: We first show that for $\tilde{\boldsymbol{x}} \in \mathcal{A}$, if there exists $\tilde{x}_i > 0$, then $\bar{\beta}((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) = \bar{\beta}((1, \tilde{\boldsymbol{x}}_{-i}))$. Suppose $\tilde{x}_i > 0$, $i \in N$, then there exists $\tilde{x}_i \geq \epsilon > 0$ for some ϵ. By monotonicity, we have the following:

$$\sup_{\boldsymbol{u} \geq 0} \lim_{r \uparrow \infty} \pi((u_i r, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) \geq \sup_{\boldsymbol{u} \geq 0} \pi((u_i \tilde{x}_i, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) \geq \sup_{\boldsymbol{u} \geq 0} \pi((u_i \epsilon, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1)$$

   where "$\circ$" represents the elementwise (or Hadamard) product. Observe in the above that the upper and lower bound for $\sup_{\boldsymbol{u} \geq 0} \pi((u_i \tilde{x}_i, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1)$ are equivalent for any $\epsilon > 0$, i.e.

$$\sup_{\boldsymbol{u} \geq 0} \lim_{r \uparrow \infty} \pi((u_i r, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) = \sup_{\boldsymbol{u} \geq 0} \pi((u_i \epsilon, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1)$$

   This implies that $\bar{\beta}(\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i}) = \bar{\beta}(\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})$ for all $\tilde{y}_i > 0$. For the case when $\tilde{\boldsymbol{x}} > 0$, we observe that

$$\sup_{\boldsymbol{u} \geq 0} \lim_{r \uparrow \infty} \pi(\boldsymbol{u} r, 1) \geq \sup_{\boldsymbol{u} \geq 0} \pi(u_1 \tilde{x}_1, \ldots, u_n \tilde{x}_n, 1) \geq \sup_{\boldsymbol{u} \geq 0} \pi(\boldsymbol{u} \epsilon, 1)$$

for some ϵ satisfying $\tilde{\boldsymbol{x}} \geq \epsilon \mathbf{1} > \mathbf{0}$. Since $\pi(\boldsymbol{u}r, 1) = \min_{i\in N}\{u_i r, 1\}$ and $\pi(\boldsymbol{u}\epsilon, 1) = \min_{i\in N}\{u_i \epsilon, 1\}$ and that $\boldsymbol{u} \geq \mathbf{0}$ can be arbitrarily large, we must have

$$\bar{\beta}(\tilde{\boldsymbol{x}}) = \sup_{\boldsymbol{u}\geq 0} \pi(u_1\tilde{x}_1, \ldots, u_n\tilde{x}_n, 1) = 1.$$

Hence, $\bar{\beta}(\mathbf{1}) = 1$.

3. **Non-abandonment**: For any $\tilde{\boldsymbol{x}} \in \mathcal{A}$ such that $\tilde{x}_i < 0$,, we have from Property 1

$$\bar{\beta}(\tilde{\boldsymbol{x}}) \leq \sup_{\boldsymbol{u}\geq\mathbf{0}} \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i\tilde{x}_i, u_j\tilde{x}_j, 1\}\right) \leq \sup_{\boldsymbol{u}\geq\mathbf{0}} \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i 0, u_j\tilde{x}_j, 1\}\right) = 0.$$

Moreover,

$$\bar{\beta}(\tilde{\boldsymbol{x}}) = \sup_{\boldsymbol{u}\geq\mathbf{0}} \pi(u_1\tilde{x}_1, \ldots, u_n\tilde{x}_n, 1) \geq \pi(0\tilde{x}_1, \ldots, 0\tilde{x}_n, 1) = \min\{0, 1\} = 0.$$

Hence, $\bar{\beta}(\tilde{\boldsymbol{x}}) = 0$.

4. **Componentwise quasiconcavity**: Let $\bar{\beta}^* = \min\{\bar{\beta}((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})), \bar{\beta}((\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i}))\}$. We will show that for any $\lambda \in [0, 1]$,

$$\bar{\beta}((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \bar{\beta}^*$$

This is clearly the case for $\lambda \in \{0, 1\}$. We will focus on $\lambda \in (0, 1)$. Note that $\bar{\beta}((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \bar{\beta}^*$ means that for any $\epsilon > 0$, there exists $\boldsymbol{u}, \boldsymbol{v} > 0$ such that

$$\pi((u_i\tilde{x}_i, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) \geq \bar{\beta}^* - \epsilon.$$

and

$$\pi((v_i\tilde{y}_i, \boldsymbol{v}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) \geq \bar{\beta}^* - \epsilon.$$

Let $\gamma = \lambda v_i/(u_i(1-\lambda) + \lambda v_i)$ and $p_j = \gamma u_j + (1-\gamma)v_j$ for all $j \in N$. Observe that $\gamma \in (0, 1)$ and $p_j > 0$ for all $j \in N$. Moreover,

$$p_i(\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i) = \gamma u_i\tilde{x}_i + (1-\gamma)v_i\tilde{y}_i$$

and

$$p_j\tilde{z}_j = \gamma u_j\tilde{z}_j + (1-\gamma)v_j\tilde{z}_j \qquad \forall j \in N_i.$$

Using these transformation of variables and noting that π being a concave function, we have

$$\begin{aligned} & \pi((p_i(\lambda\tilde{x}_i + (1-\lambda)v_j\tilde{y}_i, \boldsymbol{p}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) \\ = \; & \pi((\gamma u_i\tilde{x}_i + (1-\gamma)v_i\tilde{y}_i, \gamma\boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i} + (1-\gamma)\boldsymbol{v}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) \\ \geq \; & \gamma\pi((u_i\tilde{x}_i, \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) + (1-\gamma)\pi((v_i\tilde{y}_i, \boldsymbol{v}_{-i} \circ \tilde{\boldsymbol{x}}_{-i}), 1) \\ \geq \; & \bar{\beta}^* - \epsilon. \end{aligned}$$

Hence, $\bar{\beta}((\lambda\tilde{x}_i + (1-\lambda)\tilde{y}_i, \tilde{\boldsymbol{x}}_{-i})) \geq \bar{\beta}^*$.

5. **Componentwise scale invariant**: The result is obvious since the multiplier, u_i lies in a cone.

6. **Restricted right continuity**: For any $i \in N(\tilde{\boldsymbol{x}})$ there exists $\mathbb{P} \in \mathbb{F}$ such that $\mathbb{P}(\tilde{x}_i < 0) > 0$ and so $\mathbb{P}(\tilde{x}_i + \epsilon < 0) > 0$ for some $\epsilon > 0$. Hence, for all $a \in [0, \epsilon]$

$$\lim_{u_i\uparrow\infty} \pi((u_i(\tilde{x}_i + a), \boldsymbol{u}_{-i} \circ \tilde{\boldsymbol{x}}_{-i})) \leq \lim_{u_i\uparrow\infty} \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{j\in N_i}\{u_i(\tilde{x}_i + a), u_j\tilde{x}_j, 1\}\right) = -\infty.$$

Following the proof of right continuity of Proposition 2, there exists $\bar{u}_i > 0$ such that the optimum solution

$$\sup_{\boldsymbol{u} \geq \boldsymbol{0}} \pi(u_1\tilde{x}_1, \ldots, u_n\tilde{x}_n, 1)$$

must require $u_i \leq \bar{u}_i$ for all $a \in [0, \epsilon]$. Hence, for any $i \in N(\tilde{\boldsymbol{x}})$,

$$\begin{aligned}
\bar{\beta}((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})) &\leq \lim_{a \downarrow 0} \bar{\beta}((\tilde{x}_i + a, \tilde{\boldsymbol{x}}_{-i})) \\
&= \lim_{a \downarrow 0} \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \pi(u_1\tilde{x}_1, \ldots, u_i(\tilde{x}_i + a), \ldots, u_n\tilde{x}_n) \\
&\leq \lim_{a \downarrow 0} \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \pi(u_1\tilde{x}_1, \ldots, u_i\tilde{x}_i + \bar{u}_i a, \ldots, u_n\tilde{x}_n) \\
&\leq \lim_{a \downarrow 0} \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \left(\pi(u_1\tilde{x}_1, \ldots, u_i\tilde{x}_i, \ldots, u_n\tilde{x}_n) - \pi(0, \ldots, -a\bar{u}_i, \ldots, 0)\right) \\
&= \lim_{a \downarrow 0} \sup_{\boldsymbol{u} \geq \boldsymbol{0}} \left(\pi(u_1\tilde{x}_1, \ldots, u_i\tilde{x}_i, \ldots, u_n\tilde{x}_n) + a\bar{u}_i\right) \\
&= \bar{\beta}((\tilde{x}_i, \tilde{\boldsymbol{x}}_{-i})),
\end{aligned}$$

where the last inequality is due to the superadditivity of the π function.

■

We can easily see that the functions π_A and π_B are positively homogenous, concave and satisfy Properties 1 and 2 of Theorem 4. Moreover, when the family of distribution is given by $\mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathcal{W})$, then if $\tilde{\boldsymbol{a}} \geq 0$ and $b \geq 0$, we have $\mathbb{P}(\tilde{\boldsymbol{a}} \geq \boldsymbol{0}) = 1$ for all $\mathbb{P} \in \mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathcal{W})$. Since we assume that $\boldsymbol{\Sigma}$ does not further constrain the support set $\mathcal{W}$, it also implies that $\mathbb{P}(\tilde{\boldsymbol{a}} \geq \boldsymbol{0}) = 1$ for all $\mathbb{P} \in \mathbb{F}(\boldsymbol{\mu}, \mathcal{W})$.Therefore,

$$\pi_A(\tilde{\boldsymbol{a}}, b) = \inf_{\mathbb{P} \in \mathbb{F}(\boldsymbol{\mu}, \mathcal{W})} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{\tilde{a}_i, b\} \right) \geq 0$$

and the function also satisfies Property 3. On the other hand, since π_B does not take into account of the support $\mathcal{W}$, it will not necessarily satisfy Property 3. The following result shows how we can use these functions to obtain a tighter approximation of Problem (4) while preserving the desirable properties of Theorem 4.

Theorem 5 *Let L be an index set, and the functions $\pi_l : \mathcal{A} \times \Re \mapsto \Re$, $l \in L$ are positively homogeneous, tractable and concave that satisfy Properties 1 and 2 of Theorem 4. Moreover, there exists $l^* \in L$ such that π_{l^*} also satisfy Property 3 of Theorem 4. Let the function $\pi : \mathcal{A} \times \Re \mapsto \Re$ be the supremum convolution of π_l, $l \in L$ given by*

$$\begin{aligned}
\pi(a_1^0 + \boldsymbol{a}_1'\tilde{\boldsymbol{z}}, \ldots, a_n^0 + \boldsymbol{a}_n'\tilde{\boldsymbol{z}}, b) = \sup \quad & \sum_{l \in L} \pi_l(a_{1l}^0 + \boldsymbol{a}_{1l}'\tilde{\boldsymbol{z}}, \ldots, a_{nl}^0 + \boldsymbol{a}_{nl}'\tilde{\boldsymbol{z}}, b_l) \\
\text{s.t.} \quad & \sum_{l \in L} \boldsymbol{a}_{il} = \boldsymbol{a}_i, && i \in N \\
& \sum_{l \in L} a_{il}^0 = a_i^0, && i \in N \\
& \sum_{l \in L} b_l = b \\
& b_l, a_{il}^0 \in \Re, \boldsymbol{a}_{il} \in \Re^m && i \in N, l \in L.
\end{aligned} \tag{5}$$

Then π is also a positively homogeneous, tractable and concave function that satisfies the properties in Theorem 4. Moreover, for all $l \in L$, $\tilde{\boldsymbol{a}} \in \mathcal{A}$, $b \in \Re$.

$$\pi_l(\tilde{\boldsymbol{a}}, b) \leq \pi(\tilde{\boldsymbol{a}}, b).$$

Proof : Since π is the supremum convolution of tractable, positively homogeneous and concave functions, it is also a tractable, positively homogeneous and concave function. To show that π is an improved bound, observe that for any $k \in L$,

$$\pi(\tilde{\boldsymbol{a}}, b) \geq \pi_k(a_1^0 + \boldsymbol{a}_1'\tilde{\boldsymbol{z}}, \ldots, a_n^0 + \boldsymbol{a}_n'\tilde{\boldsymbol{z}}, b) + \sum_{l \in L \setminus \{k\}} \underbrace{\pi_l(0 + \boldsymbol{0}'\tilde{\boldsymbol{z}}, \ldots, 0 + \boldsymbol{0}'\tilde{\boldsymbol{z}}, 0)}_{= \min_{i \in \{1, \ldots, n+1\}} \{0\} = 0} = \pi_k(\tilde{\boldsymbol{a}}, b).$$

Let $b_l, a_{il}^0 \in \Re, \boldsymbol{a}_{il} \in \Re^m$, $i \in N, l \in L$ be a feasible solution to Problem (5). Observe that

$$\begin{array}{rcl}\displaystyle\sum_{l \in L} \pi_l(a_{1l}^0 + \boldsymbol{a}_{1l}'\tilde{\boldsymbol{z}}, \ldots, a_{nl}^0 + \boldsymbol{a}_{nl}'\tilde{\boldsymbol{z}}, b_l) & \leq & \displaystyle\sum_{l \in L} \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{a_{il}^0 + \boldsymbol{a}_{il}'\tilde{\boldsymbol{z}}, b_l\} \right) \\ & \leq & \displaystyle\inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \left\{ \sum_{l \in L} (a_{il}^0 + \boldsymbol{a}_{il}'\tilde{\boldsymbol{z}}), \sum_{l \in L} b_l \right\} \right) \\ & = & \displaystyle\inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{a_i^0 + \boldsymbol{a}_i'\tilde{\boldsymbol{z}}, b\} \right),\end{array}$$

where the first inequality is obvious due to Property 1 of π_l in Theorem 4, and the second inequality is valid from noting that $\inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{\tilde{a}_i, b\} \right)$ is a concave and positively homogeneous function of $(\tilde{\boldsymbol{a}}, b)$ and hence, superadditive. Therefore, for all $\tilde{\boldsymbol{a}} \in \mathcal{A}$, $b \in \Re$, $l \in L$

$$\pi_l(\tilde{\boldsymbol{a}}, b) \leq \pi(\tilde{\boldsymbol{a}}, b) \leq \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in N} \{\tilde{a}_i, b\} \right). \tag{6}$$

Furthermore, for all $\boldsymbol{a} \in \Re^{n+1}$, $l \in L$, we have

$$\min_{i \in \{1, \ldots, n+1\}} \{a_i\} = \pi_l(\boldsymbol{a}) \leq \pi(\boldsymbol{a}) \leq \inf_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left(\min_{i \in \{1, \ldots, n+1\}} \{a_i\} \right) = \min_{i \in \{1, \ldots, n+1\}} \{a_i\}$$

Hence, $\pi(\boldsymbol{a}) = \min_{i \in \{1, \ldots, n+1\}} \{a_i\}$. Finally, we note that for all $\tilde{\boldsymbol{a}} \geq \boldsymbol{0}$, $\tilde{\boldsymbol{a}} \in \mathcal{A}$ and $b \geq 0$,

$$\pi(\tilde{\boldsymbol{a}}, b) \geq \pi_{l^*}(\tilde{\boldsymbol{a}}, b) \geq 0.$$

■

Using both π_A and π_B in Theorem 5, the improved bound is given explicitly as

$$\begin{array}{rlll}\pi_{AB}\left(a_1^0 + \boldsymbol{a}_1'\tilde{\boldsymbol{z}}, \ldots, a_n^0 + \boldsymbol{a}_n'\tilde{\boldsymbol{z}}, b\right) = \sup & v + \boldsymbol{\mu}'\boldsymbol{r} + s + \boldsymbol{t}'\boldsymbol{\mu} + \mathrm{tr}(\boldsymbol{S}(\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}')) & \\ \text{s.t.} & v + \max_{\boldsymbol{z} \in \mathcal{W}} (\boldsymbol{r} - \boldsymbol{a}_{iA})'\boldsymbol{z} \leq a_{iA}^0 & i \in N \\ & v + \max_{\boldsymbol{z} \in \mathcal{W}} \boldsymbol{r}'\boldsymbol{z} \leq b_A & \\ & \begin{pmatrix} a_{iB}^0 - s & \frac{1}{2}(\boldsymbol{a}_{iB} - \boldsymbol{t})' \\ \frac{1}{2}(\boldsymbol{a}_{iB} - \boldsymbol{t}) & -\boldsymbol{S} \end{pmatrix} \in \mathbb{S}_+^{m+1} & \forall i \in N(\tilde{\boldsymbol{x}}) \\ & \begin{pmatrix} b_B - s & -\frac{1}{2}\boldsymbol{t}' \\ -\frac{1}{2}\boldsymbol{t} & -\boldsymbol{S} \end{pmatrix} \in \mathbb{S}_+^{m+1} & \\ & \boldsymbol{a}_{iA} + \boldsymbol{a}_{iB} = \boldsymbol{a}_i, & i \in N \\ & a_{iA}^0 + a_{iB}^0 = a_i^0, & i \in N \\ & b_A + b_B = b & \\ & s \in \Re, \boldsymbol{S} \in \mathbb{S}^m & \\ & v \in \Re, \boldsymbol{r} \in \Re^m & \\ & b_A, b_B, a_{iA}^0, a_{iB}^0 \in \Re, \boldsymbol{a}_{iA}, \boldsymbol{a}_{iB} \in \Re^m & i \in N\end{array}$$

which is also a tractable positive semidefinite optimization problem.

3.1.1 Robust approximations of Goh and Sim (2010, 2011)

We now look at another way to obtain an approximation of Problem (4) from the perspective of distributional robust optimization as follows:

$$\begin{aligned} \pi_{\mathcal{M}}\left(a_1^0 + \boldsymbol{a}_1'\tilde{\boldsymbol{z}}, \ldots, a_n^0 + \boldsymbol{a}_n'\tilde{\boldsymbol{z}}, b\right) = \sup \quad & \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}(y(\tilde{\boldsymbol{z}})) \\ \text{s.t.} \quad & y(\tilde{\boldsymbol{z}}) \le b \\ & y(\tilde{\boldsymbol{z}}) \le a_i^0 + \boldsymbol{a}_i'\tilde{\boldsymbol{z}} \quad i \in N \\ & y \in \mathcal{M} \end{aligned} \tag{7}$$

where $\mathcal{M}$ is a space of measurable functions on $\Re^m$. Problem (7) is a well-studied problem in recent progress on robust optimization; see for instance, Ben Tal et al. (2004), Chen et al. (2008), Goh and Sim (2010) and Kuhn et al. (2010). Although it is typically an intractable problem, it is easy to verify that when y is restricted to affine functions of $\tilde{\boldsymbol{z}}$, i.e.,

$$\mathcal{L} = \{y : \Re^m \mapsto \Re \ : \ \exists v \in \Re, \boldsymbol{r} \in \Re^m \mid y(\boldsymbol{z}) = v + \boldsymbol{r}'\boldsymbol{z}\},$$

then the function $\pi_{\mathcal{L}}$ is exactly the same as π_A. When incorporating covariance of $\tilde{\boldsymbol{z}}$, Goh and Sim (2010) show that we can obtain a tighter approximation by restricting y to a space of piecewise linear functions $\mathcal{P}$, such as deflected linear decisions rule, segregated linear decisions rule or combinations of these rules. Using any of these approaches, we have for all $\tilde{\boldsymbol{a}} \in \mathcal{A}$, $b \in \Re$,

$$\pi_{\mathcal{L}}(\tilde{\boldsymbol{a}}, b) \le \pi_{\mathcal{P}}(\tilde{\boldsymbol{a}}, b) \le \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N}\{\tilde{a}_i, b\}\right).$$

Furthermore, for all $\boldsymbol{a} \in \Re^{n+1}$

$$\min_{i\in\{1,\ldots,n+1\}}\{a_i\} = \pi_{\mathcal{L}}(\boldsymbol{a}) \le \pi_{\mathcal{P}}(\boldsymbol{a}) \le \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in\{1,\ldots,n+1\}}\{a_i\}\right) = \min_{i\in\{1,\ldots,n+1\}}\{a_i\}.$$

Also, from the above, for all $\tilde{\boldsymbol{a}} \in \mathcal{A}$ and $b \in \Re$, if $\tilde{\boldsymbol{a}} \ge \boldsymbol{0}$, $b \ge 0$, then

$$\pi_{\mathcal{P}}(\tilde{\boldsymbol{a}}, b) \ge \pi_{\mathcal{L}}(\tilde{\boldsymbol{a}}, b) = \pi_A(\tilde{\boldsymbol{a}}, b) \ge 0.$$

Hence, $\pi_{\mathcal{P}}$ satisfies all the desirable properties of Theorem 4. Although, $\pi_{\mathcal{P}}$ may not be as tight as π_{AB}, the resultant model is an attractive second-order conic optimization problem, which has tremendous computational advantage over positive semidefinite optimization problem. Moreover, Goh and Sim (2010) provides a generic framework that we could exploit to extend the description of distributions beyond $\mathbb{F}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathcal{W})$. For instance, it could incorporate stochastically independent factors, directional deviations of Chen et al. (2007) and partitioned statistics of Natarajan et al. (2010), which are means of incorporating distributional asymmetry that are not captured by mean or covariance. The ROME algebraic toolbox provides a convenient platform for the implementation of these approaches.

3.2 Optimizing over S-MOS criterion

In the previous section, we show how to evaluate S-MOS criterion for a given $\tilde{\boldsymbol{x}} \in \mathcal{X}$. Our goal is to maximize the S-MOS criterion by solving the following optimization problem:

$$\begin{aligned} Z^* = \quad & \sup \quad \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N(\tilde{\boldsymbol{x}})}\{u_i\tilde{x}_i, 1\}\right) \\ & \text{s.t.} \quad \boldsymbol{u} \geq \boldsymbol{0} \\ & \qquad\ \tilde{\boldsymbol{x}} \in \mathcal{X}. \end{aligned} \tag{8}$$

For the optimization problem to be interesting, we assume that $Z^* \in (0,1)$. If $Z^* = 0$, then the targets are overly ambitious with respect to the S-MOS criterion. On the other hand if $Z^* = 1$, then there exists $\tilde{\boldsymbol{x}} \in \mathcal{X}$ such that $\tilde{\boldsymbol{x}} \geq \boldsymbol{0}$, which we can determine its existence without having to solve Problem (8).

Unfortunately, Problem (8) is not a convex optimization problem and the global optimum solution is generally hard to obtain. Our modest goal is to provide a strategy for improving the solutions by solving as a sequence of optimization subproblems. Although we can easily extend this to optimizing the approximate S-MOS based on $\bar{\beta}$ developed in the previous section, for ease of exposition, we will focus on the exact model and assume that we can obtain the optimal solutions to the subproblems as follows:

$$\begin{aligned} \alpha(\tilde{\boldsymbol{x}}) = \quad & \max \quad \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N(\tilde{\boldsymbol{x}})}\{u_i\tilde{x}_i, 1\}\right) \\ & \text{s.t.} \quad \boldsymbol{u} \geq 0, \end{aligned} \tag{9}$$

and

$$\begin{aligned} Z(\boldsymbol{u}, I) = \quad & \max \quad \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in I}\{u_i\tilde{x}_i, 1\}\right) \\ & \text{s.t.} \quad \tilde{x}_i \geq 0, \quad i \in N \setminus I \\ & \qquad\ \tilde{\boldsymbol{x}} \in \mathcal{X}. \end{aligned} \tag{10}$$

In solving Problem (9), we observe that for any $i \in N(\tilde{\boldsymbol{x}})$, we have $\tilde{x}_i \ngeq 0$. Hence, the corresponding solution, u_i must be finite. We also denote the index set $I \subseteq N$ in (10).

We next present our strategy for solving Problem (8). The algorithm assumes that we can find an initial feasible solution $\tilde{\boldsymbol{x}} \in \mathcal{X}$ such that $\alpha(\tilde{\boldsymbol{x}}) > 0$. In practice, this initial solution can be a benchmark that we would like to improve upon in terms of achieving better satisficing performance with respect to the S-MOS criterion. Alternatively, we can generate random $\boldsymbol{u}$ and solve Problem (10) with $I = N$, until a solution $\tilde{\boldsymbol{x}}$ with $\alpha(\tilde{\boldsymbol{x}}) > 0$ is found.

Algorithm 1

Input: $\tilde{\boldsymbol{x}}$: $\alpha(\tilde{\boldsymbol{x}}) > 0$.

1. Solve Problem (9) with Input $\tilde{\boldsymbol{x}}$. Obtain optimal solution $\boldsymbol{u}^*$. Set $I := N(\tilde{\boldsymbol{x}})$
2. Solve Problem (10) with Input $(\boldsymbol{u}^*, I)$. Obtain optimal solution $\tilde{\boldsymbol{x}}^*$. Set $\tilde{\boldsymbol{x}} := \tilde{\boldsymbol{x}}^*$.
3. Repeat Steps 1 and 2 until a termination criterion is met.

Theorem 6 *Given $\tilde{\boldsymbol{x}}$, let $\boldsymbol{u}^*$ be an optimal solution to Problem (9) and $\boldsymbol{x}^*$ be the optimum solution to Problem (10) in which $I = N(\tilde{\boldsymbol{x}})$ and $\boldsymbol{u} = \boldsymbol{u}^*$. Then*

$$\alpha(\tilde{\boldsymbol{x}}^*) \geq Z(\boldsymbol{u}^*, I) \geq \alpha(\tilde{\boldsymbol{x}}).$$

Proof : Observe that $\tilde{\boldsymbol{x}}$ is a feasible solution to Problem (10) in which $I = N(\tilde{\boldsymbol{x}})$ and $\boldsymbol{u} = \boldsymbol{u}^*$ and the objective is $\alpha(\tilde{\boldsymbol{x}})$. Hence,

$$Z(\boldsymbol{u}^*, I) \geq \alpha(\tilde{\boldsymbol{x}}).$$

Moreover, due to the constraints $\tilde{x}_i \geq 0$ for all $i \in N \setminus I$ in Problem (10), the corresponding optimal solution $\tilde{\boldsymbol{x}}^*$ will have $N(\tilde{\boldsymbol{x}}^*) \subseteq I$. Hence,

$$\alpha(\tilde{\boldsymbol{x}}^*) \geq \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N(\tilde{\boldsymbol{x}}^*)}\{u_i^*\tilde{x}_i^*, 1\}\right) \geq \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in I}\{u_i^*\tilde{x}_i^*, 1\}\right) = Z(\boldsymbol{u}^*, I) \geq \alpha(\tilde{\boldsymbol{x}}).$$

■

Theorem 6 suggests that in Algorithm 1, the sequence of solutions $\{\tilde{\boldsymbol{x}}^k\}$ have non-decreasing objectives, $\alpha(\tilde{\boldsymbol{x}}^k)$. Hence, $\alpha(\tilde{\boldsymbol{x}}^k) \uparrow \delta \leq 1$ in the limit. Suppose the set $\mathcal{X}$ is closed, convex, bounded and has nonempty interior, then we can replace Problem (10) with the following convex optimization problem:

$$\begin{aligned} Z_T(\boldsymbol{u}, I) = \quad & \max \quad \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in I}\{u_i\tilde{y}_i, 1\}\right) \\ & \text{s.t.} \quad \tilde{y}_i \geq 0, \quad i \in N \setminus I \\ & \qquad\; (a, \tilde{\boldsymbol{y}}) \in \text{cl}(\bar{\mathcal{X}}) \end{aligned} \tag{11}$$

where

$$\bar{\mathcal{X}} = \{(a, \tilde{\boldsymbol{y}}) \ : \ \tilde{\boldsymbol{y}}/a \in \mathcal{X}, a > 0\}$$

is a also a convex set and cl denotes closure of the set. We then consider the following solution algorithm.

Algorithm 2

Input: $\tilde{\boldsymbol{x}} : \alpha(\tilde{\boldsymbol{x}}) > 0$.

1. Solve Problem (9) with Input $\tilde{\boldsymbol{x}}$. Obtain optimal solution $\boldsymbol{u}^*$. Set $I := N(\tilde{\boldsymbol{x}})$
2. Solve Problem (11) with Input $(\boldsymbol{u}^*, I)$. Obtain optimal solution $(a^*, \tilde{\boldsymbol{y}}^*)$. Set $\tilde{\boldsymbol{x}} := \tilde{\boldsymbol{y}}^*/a^*$.
3. Repeat Steps 1 and 2 until a termination criterion is met.

Our next result suggests that Algorithm 2 has the potential to outperform Algorithm 1.

Theorem 7 *Given $(\boldsymbol{u}, I)$ such that $Z_T(\boldsymbol{u}, I) > 0$. Let $(a^*, \tilde{\boldsymbol{y}}^*)$ be the optimal solutions to Problem (11). Then $a^* > 0$ and*

$$\alpha(\tilde{\boldsymbol{y}}^*/a^*) \geq Z_T(\boldsymbol{u}, I) \geq Z(\boldsymbol{u}, I).$$

Moreover, for the case of single objective, the optimum solution of Problem (8) can be obtained directly by solving Problem (11) with inputs $u_1 = 1$ and $I = N = \{1\}$.

Proof : We first observe that Problem (11) can be reformulated as

$$\begin{aligned} Z_T(\boldsymbol{u}, I) = \quad & \max \quad \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in I}\{u_i a\tilde{x}_i, 1\}\right) \\ & \text{s.t.} \quad a\tilde{x}_i \geq 0, \quad i \in N \setminus I \\ & \qquad\; a \geq 0 \\ & \qquad\; \tilde{\boldsymbol{x}} \in \mathcal{X}, \end{aligned} \tag{12}$$

in which its optimum solution, $\tilde{\boldsymbol{x}}^\dagger = \tilde{\boldsymbol{y}}^*/a^*$, assuming $a^* > 0$. Indeed, since $Z_T(\boldsymbol{u}, I) > 0$, the solution a of Problem (12) cannot approach zero. If a approaches infinity, then we would require $\tilde{\boldsymbol{x}} \geq 0$, which we have assumed is infeasible in the set $\mathcal{X}$. Hence, at optimality, a is finite and strictly positive. Clearly, for the case of single objective, Problem (12) is exactly the same as Problem (8) in which $u_1 = 1$ and $I = N = \{1\}$. Observe that $\tilde{\boldsymbol{x}}^*$, the optimum solution of Problem (11), and $a = 1$ are feasible in Problem (12). Therefore,

$$Z_T(\boldsymbol{u}, I) \geq Z(\boldsymbol{u}, I).$$

Finally, due to the constraints $\tilde{x}_i \geq 0$ for all $i \in N \setminus I$ in Problem (12), the corresponding optimal solution, $\tilde{\boldsymbol{x}}^\dagger$ will be such that $N(\tilde{\boldsymbol{x}}^\dagger) \subseteq I$. Hence,

$$\alpha(\tilde{\boldsymbol{y}}^*/a^*) = \alpha(\tilde{\boldsymbol{x}}^\dagger) \geq \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in N(\tilde{\boldsymbol{x}}^\dagger)}\left\{u_i a^* \tilde{x}_i^\dagger, 1\right\}\right) \geq \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in I}\left\{u_i a^* \tilde{x}_i^\dagger, 1\right\}\right) = Z_T(\boldsymbol{u}, I) \geq Z(\boldsymbol{u}, I).$$

■

Remark 1 Theorems 6 and 7 implies than $\alpha(\tilde{\boldsymbol{y}}^*/a^*)$ has a higher lower bound than $\alpha(\tilde{\boldsymbol{x}}^*)$, where $\tilde{\boldsymbol{x}}^*$ is the solution of Problem (10). However, this does not necessarily imply that $\alpha(\tilde{\boldsymbol{y}}^*/a^*) \geq \alpha(\tilde{\boldsymbol{x}}^*)$. Nevertheless, our computational studies in the next section suggests that Algorithm 2 has better performance.

Remark 2 For the case of single objective, Chen and Sim (2009) propose a binary approach to obtain the optimum solution. The result in Theorem 7 shows that this can be solved as a single convex optimization problem.

For the purpose of illustration, we consider a MOS criterion optimization problem in which the feasible solution, $\boldsymbol{x}$ lies in the polytope $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$. For a given $\boldsymbol{x}$, the uncertain objectives are affinely dependent on the uncertainties and given by

$$(\tilde{\boldsymbol{p}}_1'\boldsymbol{x} + \tilde{q}_1, \ldots, \tilde{\boldsymbol{p}}_n'\boldsymbol{x} + \tilde{q}_n).$$

Let $\boldsymbol{\tau}$ be the target. The explicit formulation of Problem (11) is therefore

$$\begin{aligned} Z(\boldsymbol{u}, I) = \quad \max \quad & \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in I}\left\{u_i\left(\tilde{\boldsymbol{p}}_i'\boldsymbol{x} + \tilde{q}_i - \tau_i\right), 1\right\}\right) \\ \text{s.t.} \quad & \tilde{\boldsymbol{p}}_i'\boldsymbol{x} + \tilde{q}_i \geq \tau_i, \quad i \in N \setminus I \\ & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \end{aligned}$$

which is a convex optimization problem with respect to $\boldsymbol{x}$. Likewise, the explicit formulation of Problem (12) is as follows:

$$\begin{aligned} Z_T(\boldsymbol{u}, I) = \quad \max \quad & \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{E}_{\mathbb{P}}\left(\min_{i\in I}\left\{u_i\left(\tilde{\boldsymbol{p}}_i'\boldsymbol{y} + a\tilde{q}_i - a\tau_i\right), 1\right\}\right) \\ \text{s.t.} \quad & \tilde{\boldsymbol{p}}_i'\boldsymbol{y} + a\tilde{q}_i \geq \tau_i a, \quad i \in N \setminus I \\ & \boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{b}a \\ & a \geq 0. \end{aligned}$$

which is also a convex optimization problem with respect to the variables a and $\boldsymbol{y}$. We use these formulations in our computational studies for optimizing the S-MOS criterion. We solve the problem approximately using deflected linear decision rules (Goh and Sim, 2010) which is a straightforward implementation in ROME. For known distributions, we can use sampling average approximation (SAA), which is a standard approach in stochastic programming (see, for instance, Birge and Louveaux (1997); Ruszczynski and Shapiro (2003)).

4 Computational studies

In this section, we first present some simple computational examples to compare the S-MOS and success probability criteria. The performance of S-MOS under distributional ambiguity is also compared using the tractable approximations in the previous section. In Section 4.2, we present a computational study of the S-MOS optimization in a blending problem. The performance of optimal decisions obtained using the S-MOS criterion and its distributional robust counterpart based on Problem (7) are then compared with that obtained via the success probability criterion using simulation.

4.1 Comparing S-MOS with success probability

For purpose of illustration, we assume a two-point distribution for the target excess such that each target excess $\tilde{x}_i$ has the following distribution:

$$\mathbb{P}(\tilde{x}_i = x) = \begin{cases} \kappa_i, & x = a_i \\ 1 - \kappa_i, & x = b_i \end{cases}$$

When $a_i < 0$ the worst-case outcome of target excess i is a shortfall against the target. Also, $\kappa_i \in [0, 1]$ models the corresponding probability of shortfall. For a specified distribution modelling the uncertainties, the S-MOS criterion can be evaluated using a generated set of random samples S via sample average approximation (SAA). Under distributional ambiguity, a tractable robust approximation of S-MOS can be evaluated using deflected linear decision rules in Problem (7). In the experiments, unless otherwise specified, we consider four targets to be satisficed. A sample size $|S| = 50,000$ is used for the solving the SAA models.

Figure 1 compares the S-MOS and success probability criteria at different levels of target shortfall of one of the targets over $a_i \in [-2, 1]$. The rest of the target excess are assumed to follow two-point distributions with lower and upper supports at -0.1 and 2 respectively, and $\kappa_i = 0.1$ for all target excess. It can be observed in Figure 1 that the S-MOS criterion value decreases steadily as target shortfall level increases. A similar behavior is observed for the tractable robust S-MOS criterion. In contrast, the success probability value remains insensitive to the level of shortfall even as the target shortfall level increases. On the other hand, when the targets are always achievable, both the S-MOS and success probability criteria values are indifferent to the level of target achievement. Furthermore, we also show the stability of solutions with respect to the number of SAA samples for $|S| = 1000, 5000, 50000$. From Figure 1, it is clear that the SAA based evaluations varies significantly over the set of SAA samples generated for evaluating S-MOS and the success probability. Clearly, evaluation based on the tractable robust S-MOS on the other hand will not experience such sampling variability.

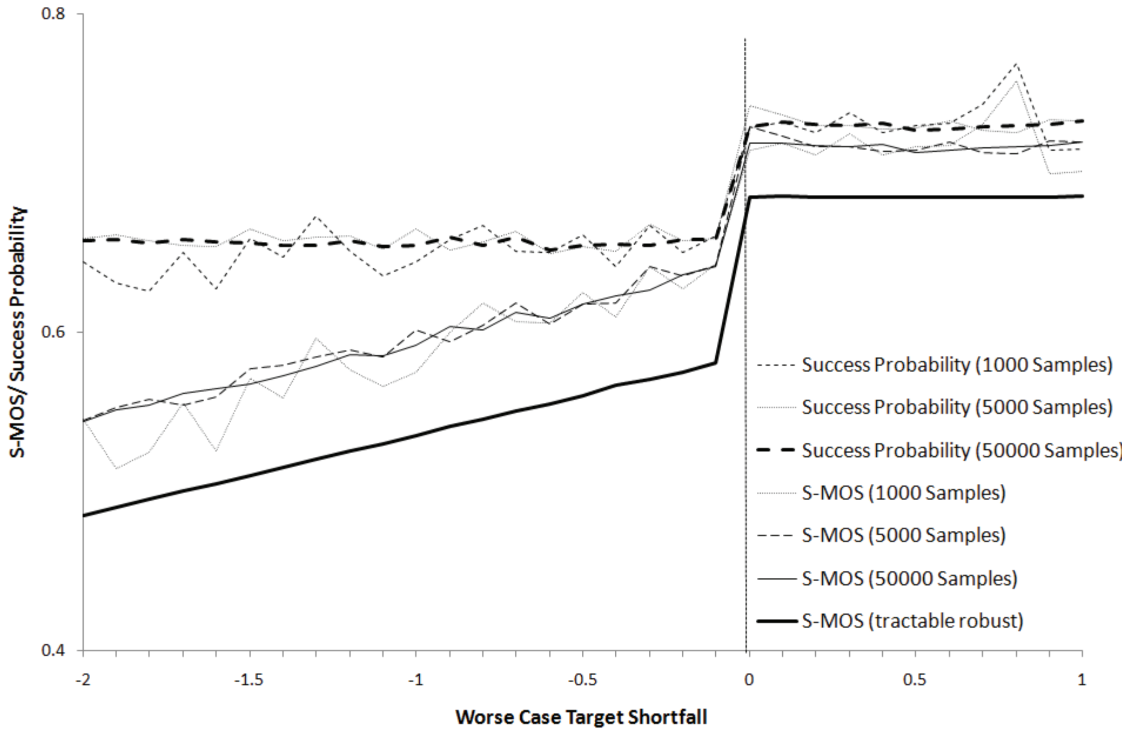

Figure 1: Characteristics of S-MOS criterion - Worst-case target shortfall a_i

Figure 2 illustrates the effect of varying the frequency of shortfalls on the S-MOS, tractable robust S-MOS and success probability. Here, we assume the worst-case shortfall for the i^{th} target excess to be $a_i = -0.5$ whilst varying κ_i for $\kappa_i \in [0.1, 1]$. The rest of the worst-case shortfall were set at $a_j = -0.1$ and $\kappa_j = 0.1$ for all $j \in \{1, \ldots, 4\} \backslash i$. For evaluating the tractable robust S-MOS, we assumed the mean, support and covariance of the two-point distributions are specified according to the same set of distributions over $\kappa_i \in [0.1, 1]$. The S-MOS criterion values provide lower bounds to the success probability for each level of shortfall frequency, and these bounds are distinguished by the level of shortfalls incurred. Naturally, the tractable robust S-MOS provides the most conservative values since it accounts for distributional ambiguity.

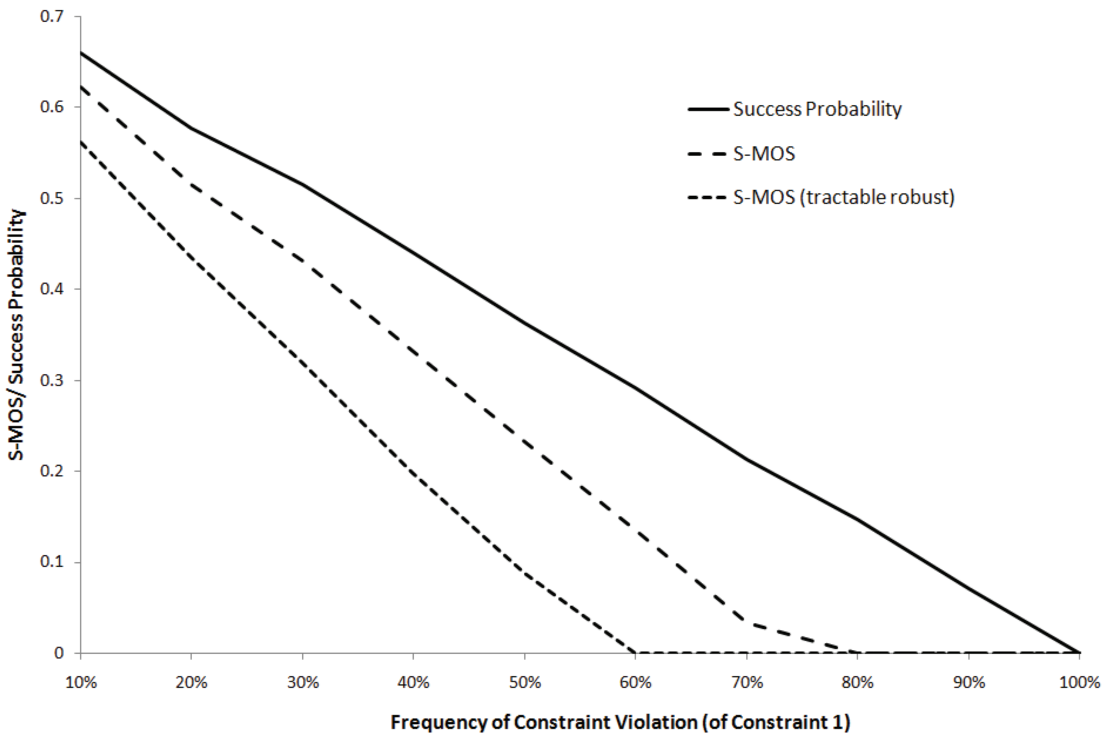

Figure 2: Characteristics of S-MOS criterion - Frequency of target shortfalls κ_i

Finally in Figure 3, the success probability and S-MOS criteria values are plotted when two of the target excess $\tilde{x}_1$ and $\tilde{x}_2$ are correlated via the scalar parameter $r \in (-1, 1)$. In particular, the two correlated target excess, $\tilde{x}_1$ and $\tilde{x}_2$, are defined as follows:

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}^{\frac{1}{2}} \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\tilde{z}_1$ and $\tilde{z}_2$ are assumed to be uniform iid random variables over $[-\sqrt{3}, \sqrt{3}]$. Hence the degree of correlation between $\tilde{x}_1$ and $\tilde{x}_2$ increases with increasing r. Figure 3 verifies that the S-MOS criteria are safe approximations to the success probability criterion. A more interesting and important observation is that both these exhibit a trend that is similar to the success probability. In this example, all three criteria are maximized when $\tilde{x}_1$ and $\tilde{x}_2$ are associated with the highest degree of correlation. These illustrations are consistent with the results established in the previous sections that S-MOS criterion offers a viable and tractable alternative to probability in the selection of decision alternatives. The implementation of S-MOS criterion for an optimization problem is shown in the following subsection.

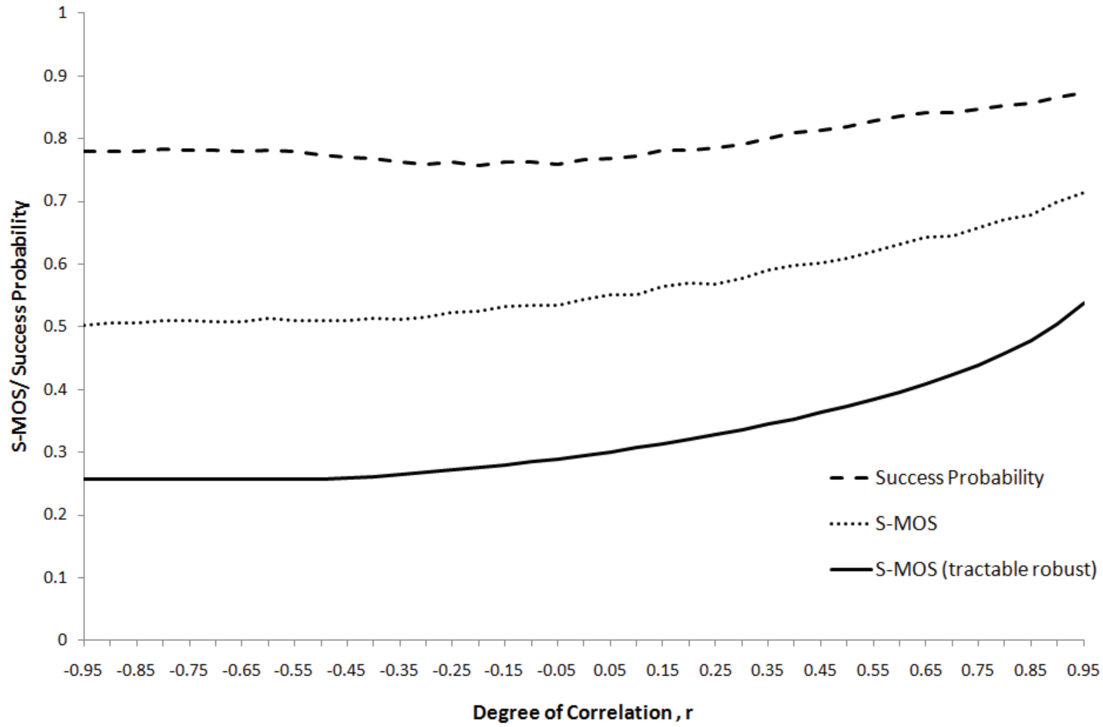

Figure 3: Characteristics of S-MOS criterion - Correlation of target excess

4.2 Example: Refinery Blending Problem

In this section, we present a computational study based on a blending problem with parameter uncertainty for refinery operations, and compare the performance of decisions obtained via solving a deterministic model and optimizing the S-MOS and success probability criteria.

Blending of raw materials and intermediate products are important steps in the synthesis of blended crude and refined petroleum products (Adhya et al. (1999)). In practice, both the feeds and the end products can have diverse quality parameters. The quality parameters for the feeds into the blending process may not be known with certainty and the end products usually have to satisfy certain pre-specified quality targets. For example, the product quality parameters may consider the levels of certain chemicals in the final blends (e.g. benzene and sulphur) and the blending process is required to produce blends that should not contain levels of such chemicals beyond stipulated thresholds. There are also minimum output levels to be met, since the output products typically form material requirements for the downstream production stages. In addition, there are also constraints on raw material availability. In our computational studies, we implemented a simplified instance of a practical refinery blending problem (see for example DeWitt et al. (1989) and Rigby et al. (1995)) possessing the aforementioned features.

The problem instance used in this case study is derived from an example in Adhya et al. (1999). In this problem instance, there are eight raw materials that can be used for blending into five output products. The blended products have to satisfy certain pre-specified quality requirements whilst the raw material quality parameters may be uncertain. The total processing costs incurred depends on the product blending options used and the production output have to meet minimum output specifications. Before describing the mathematical model, we first introduce the following notations:

Table 1: Blending problem notations

Indices	i	raw material
	j	blended product
	k	quality parameter
Variables	x_{ij}	amount of raw material i blended into product j
Coefficients	$\tilde{a}_{ik}$	k^{th} quality parameter of raw material i which can be potentially uncertain
	c_{ij}	cost of processing a unit of i^{th} raw material for j^{th} product
	b_{jk}	threshold acceptable level of quality parameter k per unit of product j
	r_i	availability of raw material i
	d_j	minimum output levels required demand of product j

For each raw material, Table 2 provides the processing cost, raw material availability and the nominal values of the two quality parameters associated with each raw material. Each of these quality parameters can be potentially uncertain over their nominal levels. For each product, Table 3 provides the minimum demand that must be satisfied and the threshold acceptable quality level for each product quality parameter. In the example, we assume that the processing cost is the raw material cost, hence $c_{ij} = c_i \ \ \forall j \in \{1, \ldots, 5\}$.

Raw Material	Raw Material	Raw Material	Raw Material Quality Parameter	
Type	Cost	Availability	Parameter 1	Parameter 2
1	15	85	1.3	1.0
2	7	85	1.7	1.6
3	4	85	1.4	1.4
4	5	85	1.7	1.3
5	6	85	1.6	2.0
6	3	85	1.4	2.0
7	5	85	1.5	1.5
8	5	85	1.2	1.5

Table 2: Raw material data

Product	Demand	Product Specifications	
		Parameter 1	Parameter 2
1	15	1.4	1.7
2	25	1.8	1.4
3	10	1.9	1.9
4	20	1.7	1.6
5	15	2.0	2.5

Table 3: Product data

Without consideration of uncertainties in the quality parameters, the blending problem can be formulated as the following deterministic linear programming model with an objective of minimizing total processing costs whilst satisfying the quality and demand requirements.

Problem LP

$$\begin{aligned} \min \quad & \sum_{j=1}^{5}\sum_{i=1}^{8} c_{ij}x_{ij} \\ \text{s.t.} \quad & b_{jk}\sum_{i=1}^{8} x_{ij} - \sum_{i=1}^{8} a_{ik}x_{ij} \geq 0 \quad j \in \{1,\dots,5\}, k \in \{1,2\} \\ & \boldsymbol{x} \in X \end{aligned} \tag{13}$$

where

$$X = \left\{ \boldsymbol{x} \in \Re^{8\times 5} \;\middle|\; \begin{array}{ll} \sum_{i=1}^{8} x_{ij} \geq d_j & j \in \{1,\dots,5\} \\ \sum_{j=1}^{5} x_{ij} \leq r_i & i \in \{1,\dots,8\} \\ \boldsymbol{x} \geq \boldsymbol{0} & \end{array} \right\}$$

The constraints in (13) defines the quality requirements on the products. The set X ensures that the production levels meet minimum specified demands and raw materials availability.

The quality parameters for the raw materials can be uncertain before the execution of the blending decisions. In this case, we can formulate an S-MOS optimization problem to locate a blending decision that satisfies the target quality requirements as well as possible under uncertainty, given a pre-specified budget τ_b for the total processing costs. For this problem, we define the target excess for the quality constraint on quality parameter k of product j as follows:

$$\tilde{p}_{jk} = b_{jk}\sum_{i=1}^{8} x_{ij} - \sum_{i=1}^{8} \tilde{a}_{ik}x_{ij}.$$

The budget τ_b can be specified based on the optimal objective value of the deterministic problem. In the case when the probability distributions of the uncertain quality parameters are known, an SAA model for the S-MOS optimization problem can be solved. For the convenience of discussing the computational results, we define this as the Problem S-MOS.

A corresponding SAA model for the optimization of success probability can be formulated as a mixed integer programming model. For this formulation, let M be a suitably large real number.

Problem SP

$$
\begin{aligned}
\max \quad & \frac{1}{|S|}\sum_{s\in S} w_s \\
\text{s.t.} \quad & b_{jk}\sum_{i=1}^{8} x_{ij} - \sum_{i=1}^{8} a_{ik}^{s} x_{ij} \geq -M(1-w_s) \quad s\in S, j\in\{1,\ldots,5\}, k\in\{1,2\} \\
& \sum_{i=1}^{8}\sum_{j=1}^{5} c_{ij}x_{ij} \leq \tau_b \\
& \mathbf{x}\in X \\
& \boldsymbol{w}\in\{0,1\}^{|S|}
\end{aligned}
\tag{14}
$$

In the above w_s is used to indicate the occurrence of the event of achieving all quality targets in realization $s \in S$. The budget constraint in the formulation ensures that the total processing costs do not exceed budget τ_b. The problem is simply to maximize the number of successes in achieving all quality targets in realization $s \in S$. It should be noted that with increased sample size, Problem SP becomes extremely difficult to solve.

In the case of distributional ambiguity, where only certain distributional information on the random parameters are available, we apply the results of Section 3.1 based on Problem (7) with deflected linear decisions rule to optimize the tractable robust approximation of S-MOS. The following model of uncertainty is assumed for the uncertain raw material quality coefficients:

$$\tilde{a}_{ik} = \bar{a}_{ik}(1+\tilde{z}_{ik}),$$

where $\bar{a}_{ik}$ represents the nominal k^{th} quality parameter of raw material i. The family of distributions is specified over the mean, support and covariance statistics as follows:

$$\mathbb{F}(\delta,\sigma) = \left\{\mathbb{P} \ : \ \mathbb{E}_{\mathbb{P}}(\boldsymbol{z}) = 0, \ \mathbb{E}_{\mathbb{P}}(\boldsymbol{z}\boldsymbol{z}') = \boldsymbol{I}\sigma^2, \ \mathbb{P}(\tilde{\boldsymbol{z}} \in [-\delta\mathbf{1}, \delta\mathbf{1}]) = 1\right\}.$$

In our computational experiments, we assumed $\tilde{z}_{ik}$ to be uniformly distributed, hence, $\sigma = \delta/\sqrt{3}$. For ease of exposition, we define this problem as the Problem DRS (Distributional Robust S-MOS).

4.2.1 Computational Experiments

In following computational studies, we compare the following characteristics of decisions derived from maximizing the S-MOS and success probability.

1. Performance of optimal solutions of Problems S-MOS, DRS and SP in terms of the probability of jointly achieving all quality specifications (success probability).

2. Ability of these optimal decisions in cushioning against extreme shortfalls.

3. Convergence characteristics of Algorithms 1 and 2.

In the computational studies, the raw material quality levels were assumed to vary over $1-3\%$ away from their nominal values shown in Table 2 (or $\delta \in [0.01, 0.03]$). Unless otherwise specified, the computations in this section assumed that the processing cost budget τ_b for the problem with uncertainties was set at 3% above the minimum processing cost achieved by Problem LP.

To solve Problems S-MOS and SP by SAA, 300 simulated samples were used. Both Problems S-MOS and SP were implemented using ROME and solved using the linear and mixed integer optimization routines in CPLEX

11.2. Algorithm 2 was implemented to solve Problem S-MOS. The maximum CPU time for solving the mixed integer Problem SP was 2500 seconds, while Problems S-MOS and DRS were both solved using Algorithm 2 in a few seconds.

The optimal objective values of Problems S-MOS, SP and DRS over the range of variations in the quality parameters are shown in Figure 4. The performance of the decisions are evaluated based on $200,000$ out-of-sample realizations of the input data. Figure 5 plots the success probability based on these out-of-sample realizations over the range of variations in the quality coefficients using optimal decisions derived from solving Problems S-MOS, SP and DRS. It is clear that, although the use of the success probability criterion through Problem SP achieved the best in-sample objective value (see Figure 4), the out-of-sample performance of the other two problems, Problems S-MOS and DRS, in terms of success probability itself, are highly comparable with that achieved using the solutions of Problem SP. In fact at higher variability levels, the solutions from Problem SP yield the lowest success probability, while the solutions of Problem DRS yield the highest success probability. This indicates that S-MOS and its tractable robust approximation not only improves computational efficiency but also the solution performance. It should also be noted that Problem SP requires much more computational time to solve. Furthermore, similar to Problem S-MOS, its solutions depend on the set of samples utilized in the SAA. Figure 1 demonstrates clearly through a simple two-point distribution that the SAA based evaluations are sensitive to the set of SAA samples generated for evaluating S-MOS and the success probability. Evaluation of Problem DRS on the other hand will not experience such sampling variability. If the actual distribution is different from what is being assumed, SAA performance might be further degraded.

Admittedly, the trends shown in Figure 5 might be different if we use a larger sample size for the computations, provided we can still solve the model. However, in practice, one rarely knows the exact distribution and such a comparison may not even be meaningful. Nonetheless, the solutions from Problem DRS not only yield the highest success probability in the computations, it is also able to handle situations where the true distribution cannot be accurately ascertained. This is the major advantage of the formulation based on Problem DRS. Specifically, despite using modest information, Problem DRS is able to produce very good results comparable to SAA-based evaluations of Problems SP and S-MOS which require large number of samples. All these were achieved with no increase in computational complexity.

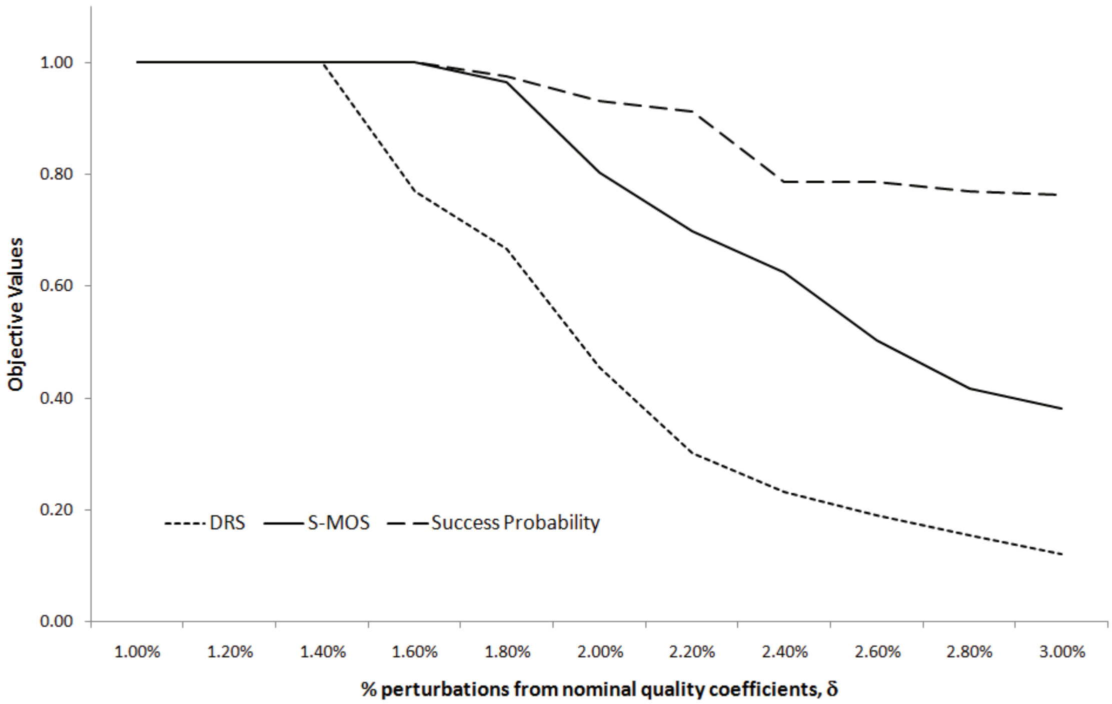

Figure 4: Objective values for the maximization of S-MOS and success probability

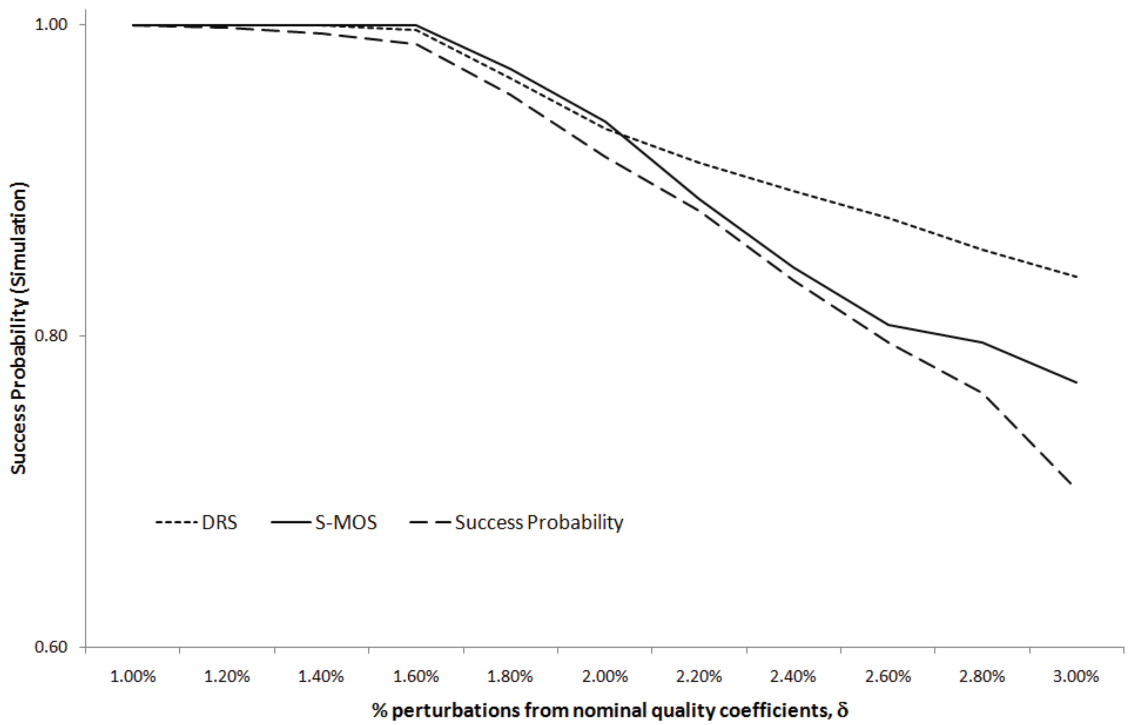

Figure 5: Probability of achieving quality specifications (simulation results)

Next, to compare the ability of the solutions in cushioning against extreme outcomes, we evaluated the 90^{th}, 95^{th} and 99^{th} percentile of the percentage violation (or shortfall) of each quality constraint. The percentage violation, $\tilde{w}_{jk}$, of a product j and quality parameter k is defined as the percentage of blend quality exceeding the specified threshold quality levels:

$$\tilde{w}_{jk} = \left(\frac{\sum_{i=1}^{8} \tilde{a}_{ijk} x_{ij}}{b_{jk} \sum_{i=1}^{8} x_{ij}} - 1 \right) \times 100\%$$

The results for Problems LP, S-MOS, SP and DRS are shown in Tables 4, 5, 6 and 7 respectively. The first rows of these tables shows the random perturbations away from the nominal values of the quality parameters shown in Table 2 over the range of $1-3\%$. The darkened cells in the tables indicate the outcomes that violate the quality threshold levels. It can be observed that the solutions from solving Problem S-MOS outperforms both the solutions of LP and SP, both in the number of violations and the corresponding percentile levels. The solution from DRS yields the smallest number of quality violations, and comparably lower percentile values. In contrast, the solution from Problem LP, while yielding the minimum cost, performed badly under uncertainty. 50% of the quality constraints were violated across the entire range of quality uncertainty considered (Table 4). Also, the solution from SP also resulted in higher degree of constraint violations, even though the optimal objective values were the higher than the SMOS criteria. The use of S-MOS and its tractable robust approximation effectively reduced the number of quality constraint violations under the range of uncertainty in the raw materials. Furthermore, the solutions seemed more stable across the range of variability studied. For instance, for the 90^{th} to 99^{th} percentile levels (Table 7), the violations always occurred for quality parameter 2 of products 1, 2 and 3. On the other hand, little such inference can be made for the solutions of Problem SP.

Table 4: Percentile of quality violations - deterministic solution (simulation results)

90th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	0.6%	0.7%	0.8%	0.9%	1.0%	1.1%	1.2%	1.3%	1.5%	1.6%	1.6%
1	2	0.8%	1.0%	1.1%	1.3%	1.4%	1.6%	1.8%	1.9%	2.1%	2.2%	2.4%
2	1	0.7%	0.8%	1.0%	1.1%	1.3%	1.4%	1.5%	1.7%	1.8%	2.0%	2.1%
2	2	0.6%	0.7%	0.8%	0.9%	1.0%	1.1%	1.2%	1.3%	1.5%	1.6%	1.7%
3	1	-19.4%	-19.2%	-19.1%	-19.0%	-18.8%	-18.7%	-18.6%	-18.5%	-18.3%	-18.2%	-18.1%
3	2	0.6%	0.7%	0.8%	0.9%	1.0%	1.1%	1.2%	1.3%	1.4%	1.5%	1.7%
4	1	-21.6%	-21.5%	-21.3%	-21.2%	-21.1%	-21.0%	-20.9%	-20.7%	-20.6%	-20.5%	-20.4%
4	2	-25.8%	-25.7%	-25.6%	-25.5%	-25.4%	-25.3%	-25.2%	-25.1%	-25.0%	-24.9%	-24.8%
5	1	-17.2%	-17.1%	-17.0%	-16.9%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.3%	-16.2%
5	2	-29.4%	-29.3%	-29.2%	-29.1%	-29.0%	-28.9%	-28.8%	-28.6%	-28.5%	-28.4%	-28.3%

95th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	0.7%	0.8%	1.0%	1.1%	1.2%	1.4%	1.5%	1.7%	1.8%	1.9%	2.0%
1	2	0.9%	1.1%	1.3%	1.4%	1.6%	1.8%	2.0%	2.2%	2.3%	2.5%	2.7%
2	1	0.8%	1.0%	1.1%	1.3%	1.4%	1.6%	1.7%	1.9%	2.1%	2.2%	2.4%
2	2	0.7%	0.8%	0.9%	1.1%	1.2%	1.4%	1.5%	1.7%	1.8%	1.9%	2.0%
3	1	-19.3%	-19.1%	-19.0%	-18.9%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%
3	2	0.7%	0.8%	1.0%	1.1%	1.2%	1.4%	1.5%	1.6%	1.8%	1.9%	2.1%
4	1	-21.5%	-21.4%	-21.2%	-21.1%	-21.0%	-20.8%	-20.7%	-20.5%	-20.4%	-20.3%	-20.1%
4	2	-25.8%	-25.6%	-25.5%	-25.4%	-25.3%	-25.2%	-25.1%	-25.0%	-24.8%	-24.7%	-24.6%
5	1	-17.1%	-16.9%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.3%	-16.1%	-16.0%	-15.9%
5	2	-29.4%	-29.2%	-29.1%	-29.0%	-28.9%	-28.7%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%

99th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	0.9%	1.0%	1.2%	1.4%	1.5%	1.7%	1.9%	2.1%	2.2%	2.4%	2.6%
1	2	1.0%	1.2%	1.4%	1.6%	1.8%	2.0%	2.2%	2.4%	2.5%	2.7%	2.9%
2	1	0.9%	1.1%	1.3%	1.4%	1.6%	1.8%	2.0%	2.2%	2.4%	2.5%	2.7%
2	2	0.9%	1.0%	1.2%	1.4%	1.6%	1.7%	1.9%	2.1%	2.2%	2.4%	2.6%
3	1	-19.2%	-19.1%	-18.9%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%	-17.6%
3	2	0.9%	1.0%	1.2%	1.4%	1.5%	1.7%	1.9%	2.1%	2.2%	2.4%	2.6%
4	1	-21.5%	-21.3%	-21.2%	-21.0%	-20.8%	-20.7%	-20.5%	-20.4%	-20.2%	-20.1%	-19.9%
4	2	-25.7%	-25.5%	-25.4%	-25.3%	-25.1%	-25.0%	-24.9%	-24.7%	-24.6%	-24.5%	-24.3%
5	1	-16.9%	-16.8%	-16.7%	-16.5%	-16.4%	-16.2%	-16.1%	-15.9%	-15.8%	-15.7%	-15.5%
5	2	-29.3%	-29.2%	-29.0%	-28.9%	-28.8%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%	-27.9%

Table 5: Percentile of quality violations - success probability solution (simulation results)

90th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-5.7%	-0.1%	-0.2%	-0.4%	-0.5%	-0.5%	-0.1%	-0.5%	0.0%	0.2%	0.1%
1	2	-0.3%	-0.2%	-0.2%	-0.4%	-0.5%	-0.1%	-0.2%	0.1%	0.2%	0.6%	0.9%
2	1	-0.1%	-0.1%	-0.1%	-0.2%	-0.4%	-0.4%	0.0%	-0.3%	-0.4%	-0.1%	-0.1%
2	2	-0.2%	-0.2%	-0.2%	-0.4%	-0.5%	-0.3%	-0.2%	-0.5%	0.1%	0.5%	0.1%
3	1	-19.4%	-19.2%	-19.1%	-19.0%	-18.9%	-18.7%	-18.6%	-18.5%	-18.3%	-18.2%	-18.1%
3	2	-0.2%	-0.5%	-0.3%	-0.6%	-0.1%	-0.7%	-0.3%	-0.1%	0.1%	-0.9%	-0.7%
4	1	-19.5%	-19.7%	-18.8%	-18.2%	-17.5%	-18.0%	-17.5%	-17.7%	-17.3%	-17.9%	-18.0%
4	2	-23.2%	-23.1%	-23.2%	-24.2%	-25.5%	-25.4%	-25.3%	-25.2%	-25.1%	-25.0%	-24.9%
5	1	-17.2%	-17.1%	-17.0%	-16.9%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.3%	-16.2%
5	2	-29.4%	-29.3%	-29.2%	-29.1%	-29.0%	-28.9%	-28.8%	-28.7%	-28.5%	-28.4%	-28.3%

95th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-5.6%	0.0%	0.0%	-0.2%	-0.3%	-0.3%	0.2%	-0.2%	0.3%	0.5%	0.5%
1	2	-0.2%	-0.1%	-0.1%	-0.2%	-0.3%	0.1%	0.1%	0.4%	0.5%	0.9%	1.2%
2	1	0.0%	0.1%	0.1%	-0.1%	-0.2%	-0.2%	0.2%	-0.1%	-0.1%	0.2%	0.2%
2	2	-0.1%	0.0%	0.0%	-0.2%	-0.3%	0.0%	0.1%	-0.2%	0.4%	0.9%	0.5%
3	1	-19.3%	-19.1%	-19.0%	-18.9%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%
3	2	-0.1%	-0.3%	-0.2%	-0.4%	0.1%	-0.5%	0.0%	0.2%	0.4%	-0.6%	-0.4%
4	1	-19.4%	-19.6%	-18.7%	-18.0%	-17.3%	-17.9%	-17.3%	-17.5%	-17.1%	-17.6%	-17.7%
4	2	-23.2%	-23.0%	-23.1%	-24.1%	-25.3%	-25.2%	-25.1%	-25.0%	-24.9%	-24.8%	-24.7%
5	1	-17.1%	-17.0%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.2%	-16.1%	-16.0%	-15.9%
5	2	-29.4%	-29.2%	-29.1%	-29.0%	-28.9%	-28.7%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%

99th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-5.5%	0.2%	0.2%	0.0%	0.0%	0.1%	0.5%	0.2%	0.7%	1.0%	1.0%
1	2	0.0%	0.0%	0.1%	0.0%	-0.1%	0.4%	0.4%	0.7%	0.8%	1.2%	1.9%
2	1	0.1%	0.2%	0.2%	0.1%	0.0%	0.0%	0.4%	0.2%	0.2%	0.5%	0.6%
2	2	0.1%	0.2%	0.2%	0.1%	0.0%	0.3%	0.4%	0.2%	0.8%	1.4%	1.0%
3	1	-19.2%	-19.1%	-18.9%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%	-17.6%
3	2	0.0%	-0.1%	0.1%	-0.1%	0.4%	-0.2%	0.3%	0.6%	0.8%	-0.1%	0.1%
4	1	-19.3%	-19.5%	-18.5%	-17.9%	-17.1%	-17.6%	-17.1%	-17.2%	-16.7%	-17.3%	-17.4%
4	2	-23.0%	-22.9%	-22.9%	-24.0%	-25.1%	-25.0%	-24.9%	-24.8%	-24.6%	-24.5%	-24.4%
5	1	-16.9%	-16.8%	-16.6%	-16.5%	-16.4%	-16.2%	-16.1%	-15.9%	-15.8%	-15.6%	-15.5%
5	2	-29.3%	-29.2%	-29.0%	-28.9%	-28.8%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%	-27.9%

Table 6: Percentile of quality violations - S-MOS solution (simulation results)

90th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-0.9%	-0.8%	-0.8%	-0.6%	-0.5%	-0.6%	-0.4%	-0.4%	-0.3%	-0.3%	-0.3%
1	2	-0.8%	-0.6%	-0.5%	-0.4%	-0.4%	-0.1%	**0.0%**	**0.2%**	**0.4%**	**0.7%**	**0.8%**
2	1	-1.0%	-0.9%	-0.8%	-0.5%	-0.5%	-0.5%	-0.4%	-0.7%	-0.5%	-0.8%	-0.8%
2	2	-0.8%	-0.8%	-0.7%	-0.6%	-0.5%	-0.4%	-0.4%	-0.2%	-0.2%	0.0%	-0.1%
3	1	-19.4%	-19.2%	-19.1%	-19.0%	-18.8%	-18.7%	-18.6%	-18.5%	-18.3%	-18.2%	-18.1%
3	2	-1.1%	-1.0%	-0.8%	-0.7%	-0.4%	-0.5%	-0.4%	-0.4%	-0.4%	-0.5%	-0.2%
4	1	-18.4%	-18.3%	-18.2%	-18.0%	-17.7%	-18.0%	-17.8%	-17.9%	-17.9%	-18.0%	-17.9%
4	2	-25.8%	-25.8%	-25.7%	-25.6%	-25.5%	-25.4%	-25.3%	-25.2%	-25.1%	-25.0%	-25.0%
5	1	-17.2%	-17.1%	-17.0%	-16.9%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.3%	-16.2%
5	2	-29.4%	-29.3%	-29.2%	-29.1%	-29.0%	-28.9%	-28.8%	-28.7%	-28.5%	-28.4%	-28.3%

95th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-0.8%	-0.7%	-0.6%	-0.4%	-0.3%	-0.3%	-0.1%	-0.2%	0.0%	**0.0%**	**0.0%**
1	2	-0.7%	-0.5%	-0.4%	-0.3%	-0.2%	**0.1%**	**0.2%**	**0.5%**	**0.7%**	**1.0%**	**1.1%**
2	1	-0.9%	-0.8%	-0.6%	-0.3%	-0.3%	-0.3%	-0.2%	-0.4%	-0.2%	-0.5%	-0.5%
2	2	-0.7%	-0.6%	-0.6%	-0.4%	-0.3%	-0.2%	-0.1%	**0.1%**	**0.1%**	**0.3%**	**0.3%**
3	1	-19.3%	-19.1%	-19.0%	-18.8%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%
3	2	-1.0%	-0.9%	-0.6%	-0.5%	-0.2%	-0.3%	-0.1%	-0.1%	-0.1%	-0.2%	**0.1%**
4	1	-18.4%	-18.2%	-18.1%	-17.9%	-17.6%	-17.8%	-17.6%	-17.7%	-17.7%	-17.8%	-17.6%
4	2	-25.8%	-25.7%	-25.6%	-25.4%	-25.3%	-25.2%	-25.1%	-25.0%	-24.9%	-24.8%	-24.7%
5	1	-17.1%	-17.0%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.2%	-16.1%	-16.0%	-15.9%
5	2	-29.4%	-29.2%	-29.1%	-29.0%	-28.9%	-28.7%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%

99th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-0.5%	-0.5%	-0.4%	-0.2%	0.0%	0.0%	**0.2%**	**0.2%**	**0.4%**	**0.5%**	**0.5%**
1	2	-0.5%	-0.4%	-0.2%	-0.1%	**0.0%**	**0.4%**	**0.5%**	**0.8%**	**1.0%**	**1.3%**	**1.5%**
2	1	-0.8%	-0.7%	-0.4%	-0.1%	-0.1%	0.0%	**0.1%**	-0.1%	**0.1%**	-0.2%	-0.1%
2	2	-0.5%	-0.4%	-0.3%	-0.1%	**0.1%**	**0.1%**	**0.3%**	**0.5%**	**0.6%**	**0.8%**	**0.8%**
3	1	-19.2%	-19.1%	-18.9%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%	-17.7%
3	2	-0.9%	-0.7%	-0.4%	-0.2%	**0.2%**	**0.0%**	**0.2%**	**0.3%**	**0.3%**	**0.3%**	**0.6%**
4	1	-18.2%	-18.1%	-17.9%	-17.7%	-17.4%	-17.5%	-17.4%	-17.4%	-17.4%	-17.5%	-17.3%
4	2	-25.7%	-25.5%	-25.4%	-25.3%	-25.1%	-25.0%	-24.9%	-24.8%	-24.6%	-24.5%	-24.4%
5	1	-16.9%	-16.8%	-16.6%	-16.5%	-16.4%	-16.2%	-16.1%	-15.9%	-15.8%	-15.6%	-15.5%
5	2	-29.3%	-29.2%	-29.0%	-28.9%	-28.8%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%	-27.9%

Table 7: Percentile of quality violations - distributional robust S-MOS solution (simulation results)

90th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-0.6%	-0.7%	-0.7%	-0.8%	-0.9%	-1.0%	-0.9%	-0.9%	-0.9%	-0.9%	-0.9%
1	2	-0.4%	-0.4%	-0.5%	-0.4%	0.0%	0.1%	0.3%	0.5%	0.7%	0.9%	1.1%
2	1	-0.6%	-0.6%	-0.5%	-0.5%	-0.6%	-0.7%	-0.7%	-0.8%	-0.9%	-1.0%	-1.0%
2	2	-0.6%	-0.6%	-0.6%	-0.7%	-0.8%	-0.8%	-0.6%	-0.6%	-0.5%	-0.3%	-0.2%
3	1	-19.6%	-19.4%	-19.1%	-19.0%	-18.8%	-18.7%	-18.6%	-18.5%	-18.3%	-18.2%	-18.1%
3	2	-0.6%	-0.6%	-0.7%	-0.8%	-0.9%	-0.6%	-0.8%	-0.7%	-0.6%	-0.6%	-0.6%
4	1	-19.1%	-18.7%	-18.2%	-18.2%	-18.4%	-18.3%	-18.4%	-18.4%	-18.4%	-18.4%	-18.4%
4	2	-25.7%	-25.7%	-25.6%	-25.6%	-25.5%	-25.4%	-25.3%	-25.2%	-25.1%	-25.1%	-25.0%
5	1	-17.1%	-17.0%	-17.0%	-16.9%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.3%	-16.2%
5	2	-29.4%	-29.3%	-29.2%	-29.1%	-29.0%	-28.9%	-28.8%	-28.7%	-28.5%	-28.4%	-28.3%

95th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-0.5%	-0.5%	-0.5%	-0.6%	-0.7%	-0.8%	-0.7%	-0.6%	-0.6%	-0.6%	-0.5%
1	2	-0.3%	-0.3%	-0.3%	-0.2%	0.1%	0.3%	0.5%	0.8%	1.0%	1.2%	1.4%
2	1	-0.5%	-0.4%	-0.4%	-0.4%	-0.4%	-0.5%	-0.5%	-0.6%	-0.6%	-0.7%	-0.7%
2	2	-0.4%	-0.5%	-0.5%	-0.5%	-0.5%	-0.6%	-0.4%	-0.3%	-0.2%	0.0%	0.1%
3	1	-19.5%	-19.3%	-19.0%	-18.8%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%
3	2	-0.4%	-0.5%	-0.5%	-0.6%	-0.7%	-0.4%	-0.5%	-0.4%	-0.3%	-0.3%	-0.3%
4	1	-19.0%	-18.6%	-18.1%	-18.0%	-18.2%	-18.2%	-18.2%	-18.2%	-18.2%	-18.2%	-18.2%
4	2	-25.6%	-25.6%	-25.5%	-25.4%	-25.3%	-25.2%	-25.1%	-25.0%	-24.9%	-24.8%	-24.7%
5	1	-17.0%	-16.9%	-16.8%	-16.7%	-16.6%	-16.5%	-16.4%	-16.2%	-16.1%	-16.0%	-15.9%
5	2	-29.3%	-29.2%	-29.1%	-29.0%	-28.9%	-28.7%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%

99th Percentile

Product	Quality Attribute	1	1.2	1.4	1.6	1.8	2	2.2	2.4	2.6	2.8	3
1	1	-0.4%	-0.3%	-0.3%	-0.3%	-0.4%	-0.4%	-0.3%	-0.3%	-0.2%	-0.1%	0.0%
1	2	-0.1%	-0.2%	-0.2%	0.0%	0.4%	0.5%	0.8%	1.0%	1.3%	1.5%	1.7%
2	1	-0.4%	-0.3%	-0.2%	-0.2%	-0.2%	-0.2%	-0.2%	-0.2%	-0.3%	-0.3%	-0.3%
2	2	-0.3%	-0.2%	-0.2%	-0.2%	-0.2%	-0.3%	0.0%	0.1%	0.3%	0.5%	0.6%
3	1	-19.5%	-19.2%	-19.0%	-18.7%	-18.6%	-18.4%	-18.3%	-18.1%	-18.0%	-17.8%	-17.6%
3	2	-0.3%	-0.3%	-0.3%	-0.3%	-0.4%	-0.1%	-0.1%	0.0%	0.1%	0.2%	0.2%
4	1	-18.9%	-18.4%	-18.0%	-17.8%	-18.1%	-17.9%	-18.0%	-17.9%	-17.9%	-17.9%	-17.9%
4	2	-25.5%	-25.5%	-25.4%	-25.3%	-25.1%	-25.0%	-24.9%	-24.8%	-24.6%	-24.5%	-24.4%
5	1	-16.9%	-16.7%	-16.6%	-16.5%	-16.4%	-16.2%	-16.1%	-15.9%	-15.8%	-15.6%	-15.5%
5	2	-29.3%	-29.2%	-29.0%	-28.9%	-28.8%	-28.6%	-28.5%	-28.4%	-28.2%	-28.1%	-27.9%

Figure 6 plots the S-MOS criterion values under various processing cost budgets τ_b at 3%, 5%, 7% and 10% above the minimum cost solution from Problem LP. While it is obvious that relaxing the cost budget improves the solution performance, the results also indicate that a small increase in operating budget can improve the ability to hedge uncertainty in the target achievement substantially. Over the range of perturbations studied ($2-7\%$), this effect appears to be most significant when the budget is set at a level of 5% above the deterministic minimum cost solution. This may provide some insights to support managerial decisions and could be used for justification of operating budget levels.

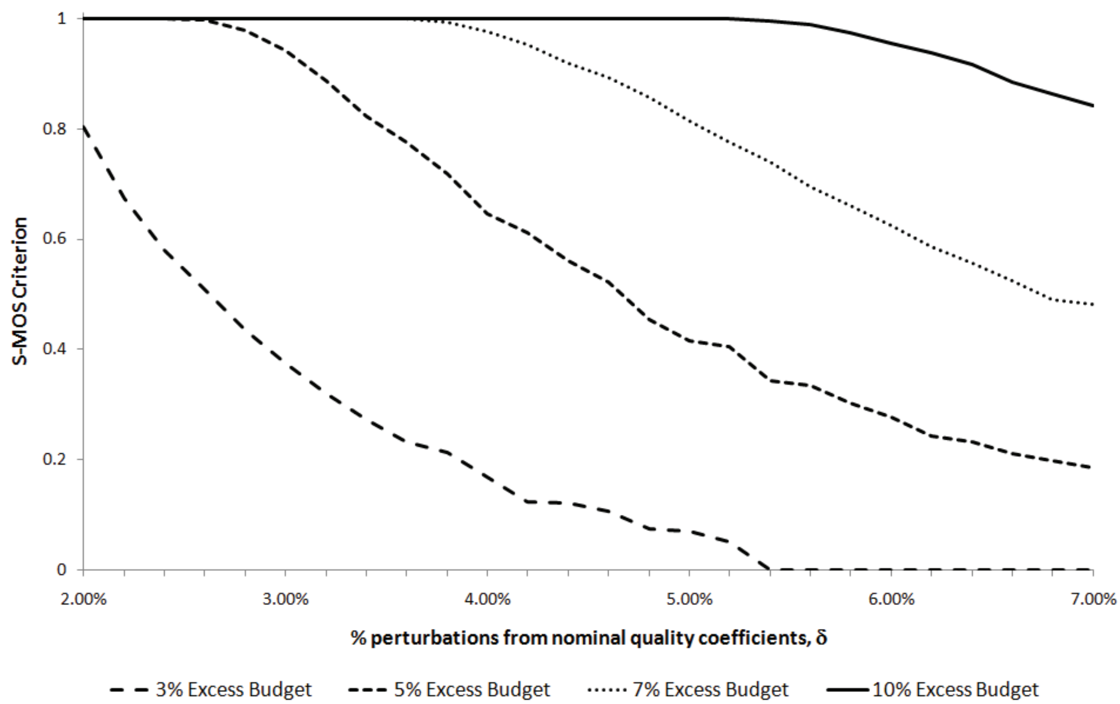

Figure 6: Sensitivity to budget requirements (Problem S-MOS)

Finally, we compare the performance of Algorithms 1 and 2, both of which depend on the initial search points. We focus on the case where a starting $\boldsymbol{u}$ vector is specified for solving Problems (10) and (11). In the following computations, we are interested in the behavior of the solution paths under different starting $\boldsymbol{u}$. The initial $\boldsymbol{u}$ are generated randomly with components iid and uniform over $[1, 2]$. For each of these starting points, we implement Algorithms 1 and 2 and track the objective values of the sample solutions at every iteration. The maximum allowed perturbations in the uncertain parameters are 2.5% away from the nominal levels. In Table 8, we present the distribution of the trajectories for solutions to Problem (9) in Algorithms 1 and 2. The first column of Table 8 indicates the range in which the objective values fall within. For example, at the end of the fifth iteration, 74% of the sample path solutions has objective values in $(0.2108, 0.2109]$ using Algorithm 1. It can be observed that all solutions from Algorithm 2 and 95% of the solutions from Algorithm 1 converged to $(0.2108, 0.2109]$. However, Algorithm 2 clearly outperforms Algorithm 1 in the number of iterations used. All the sample paths in Algorithm 2 converged after only seven iterations as compared Algorithm 1, whereas 95% of the objective values converged after fifteen iterations. It can also be observed that all the objective values converged to $(0.2107, 0.2109]$ after only five iterations using Algorithm 2, whereas only 86% of the objective values converged to this range using Algorithm 1 in the same number of iterations. For three sample paths, the objective values converged to a different solution for Algorithm 1, whereas all the sample paths converged to the same solution under Algorithm 2.

	Iteration: 1		Iteration: 2		Iteration: 3		Iteration: 5		Iteration: 7		Iteration: 15	
	Alg 1	Alg 2	Alg 1	Alg 2	Alg 1	Alg 2	Alg 1	Alg 2	Alg 1	Alg 2	Alg 1	Alg 2
(0, 0.2000]	100	100	0	0	0	0	0	0	0	0	0	0
(0.2000, 0.2100]	0	0	12	14	9	0	6	0	5	0	3	0
(0.2100, 0.2101]	0	0	0	2	0	0	0	0	0	0	0	0
(0.2101,0.2102]	0	0	1	0	3	0	0	0	0	0	0	0
(0.2102,0.2103]	0	0	2	3	0	3	0	0	1	0	0	0
(0.2103,0.2104]	0	0	9	0	2	0	3	0	0	0	0	0
(0.2104,0.2105]	0	0	10	10	0	2	3	0	0	0	0	0
(0.2105,0.2106]	0	0	13	10	5	11	0	0	1	0	0	0
(0.2106,0.2107]	0	0	19	14	12	4	2	0	5	0	0	0
(0.2107,0.2108]	0	0	27	33	50	38	12	20	8	0	2	0
(0.2108,0.2109]	0	0	7	14	19	42	74	80	80	100	95	100

Table 8: Distribution of objective values

5 Conclusion

We conclude the paper by reiterating its main contributions. In particular, a class of functions, the MOS criteria, is proposed for evaluating the level of compliance of a set of objectives in meeting their targets collectively under uncertainty. This class of MOS criteria include the success probability criterion and extends to a sub-class that favors diversification. We also propose a particular shortfall-aware MOS criterion (S-MOS) within this sub-class of MOS criteria and demonstrated, through computational studies, its ability to mitigate severe shortfalls in scenarios when an objective fails to achieve its target. Furthermore, we propose tractable approximations of the S-MOS criterion to deal with distributional ambiguity and described its implementation through the computational studies with promising results.

Acknowledgments

The authors would like to thank El-bakry, A. S. and Xu, L. of ExxonMobil Research and Engineering Company, Corporate Strategic Research, for their valuable contributions towards the development of this research. In particular, the first author would like to thank ExxonMobil Research and Engineering Company, Corporate Strategic Research, for hosting him for one month in the course of this research.

References

Adhya, N., Tawarmalani, M., Sahinidis, N. A. (1999): A Lagrangian approach to the pooling problem. *Industrial and Engineering Chemical Research*, 38: p. 1956-1972.

Ben-Tal, A., Nemirovski, A. (1998): Robust convex optimization. *Mathematics in Operations Research*, 23: p. 769-805.

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

Bertsimas, D., Popescu, I. (2005): Optimal Inequalities in Probability Theory: A Convex Optimization Approach. *SIAM Journal on Optimization*, 15(2): p. 780-804.

Bertsimas, D., Sim, M. (2004): Price of robustness. *Operations Research*, 52: p. 35-53.

Birge, J. R., Louveaux, F. (1997): *Introduction to Stochastic Programming.*, Springer, New York

Bordley, R., LiCalzi, M. (2000): Decision an analysis using targets instead of utility functions. *Decisions in Economics and Finance*, 23: p. 53-74.

Bordley, R. F., Kirkwood, C. W. (2004): Multiattribute preference analysis with performance targets. *Operations Research*, 52(6): p. 823-835.

Bordley, R. F., Pollock, S. M. (2009): A decision-analytic approach to reliability-based design optimization. *Operations Research*, 57(5): p. 1262-1270.

Boulding, K.E. (1952): Implications for general economics of more realistic theories of the firm. *American Economic Review*, 42: p. 35-44.

Brown, D. B., Sim, M. (2009): Satisficing measures for analysis of risky positions. *Management Science*, 55(1): p. 71-84.

Carter, E.E. (1971): Behavioral theory of firm and top-level corporate decisions. *Administrative Science Quarterly*, January: p. 251-273.

Castagnoli, E., LiCalzi, M. (1996): Expected utility without utility. *Theory and Decisions*, 41: p. 281-301.

Charnes, A., Cooper, W.W., Symonds G.H. (1958) Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3): p. 235-263.

Chen, X., Sim, M., Sun, P. (2007): A robust optimization perspective on stochastic programming. *Operations Research*, 55(6): p. 1058-1071.

Chen, X., Sim, M., Sun, P., Zhang, J. (2003): A linear-decision based approximation approach to stochastic programming. *Operations Research*, 56(2): p. 344-357.

Chen, W., Sim, M., Sun, J., Teo, C-P. (2010): From CVaR to uncertainty set: implications in joint chance constrained optimization. *Operations Research*, 58: p. 470-485.

Chen, W., Sim, M. (2009): Goal driven optimization. *Operations Research*, 57(2): p. 342-357.

Cyert, R.M., March, J.G. (1963): *A Behavioral Theory of the Firm.*, Prentice Hall, Englewood Cliffs,N.J.

Delage, E., Ye, Y. (2010): Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, Articles in Advance, DOI: opre.1090.0741v1-pre.1090.0741.

DeWitt, C. W., Lasdon, L. S., Waren, A. D., Brenner, D. A., Melhem, S. A. (1989): OMEGA: An improved gasoline blending system for Texaco. *Interfaces*, 19(1): p. 85-101.

Diecidue, E., van de Ven, J. (2008): Aspiration level, probability of success and expected utility. *International Economic Review*, 2(5):p. 683-700.

Dupačová, J. (1987): The minimax approach to stochastic programming and an illustrative application. *Stochastics* 20(1): p 7388.

El Ghaoui, L., Oks, M., Oustry, F. (2003): Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4): p. 543-556.

Ellsberg, D. (1961): Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, 75(4): p. 643-669.

Föllmer, H., Schied, A. (2004): *Stochastic Finance: An Introduction in Discrete Time.*, Walter de Gruyter, Berlin.

Goh, J., Sim, M. (2009): Distributionally robust optimization and its tractable approximations. *Operations Research*, 58(4): 902-917.

Goh, J., Sim, M. (2010): Robust optimization made easy with ROME. *Operations Research*, forthcoming.

Isii, K. (1963): On the sharpness of Chebyshev-type inequalities. *Annals of the Institute of Statistical Mathematics*, 12: p. 185-197.

Knight, F. (1921): *Risk, Uncertainty and Profit.*, Houghton Mifflin, Boston.

Kuhn, D., Wiesemann, W., Georghiou, A. (2010): Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming.*, Articles in Advance, DOI: 10.1007/s10107-009-0331-4.

Lanzillotti, R. F. (1958): Pricing objectives in large companies. *American Economic Review*, 48: p. 921-940.

Mao, J. C. T. (1970): Survey of capital budgeting: Theory and practice. *Journal of Finance*, 25(2): p. 349-360.

Miller, L.B., Wagner, H. (1965): Chance-constrained programming with joint constraints. *Operations Research*, 13(6): p. 930-945.

Murty, K.G., Kabadi, S.N. (1987): Some NP-complete problems in quadratic and non-linear programming. *Mathematical Programming*, 39: p. 117-129.

Natarajan, K., Sim, M., Uichanco, J, (2010): Tractable robust expected utility and risk models for portfolio optimization. *Mathematical Finance*, 20(4), 695-731.

Nemirovski, A., Shapiro, A. (2006): Convex approximations of chance constrained problems. *SIAM Journal of Optimization*, 17(4): p. 969-996.

Payne, J.W., Laughhunn, D.J., Crum, R. (1980): Translation of gambles and aspiration level effects in risky choice behavior. *Management Science*, 26: p. 1039-1060.

Prékopa, A. (1970): On probabilistic constrained programming. *Proceedings of the Princeton Symposium on Mathematical Programming*, Princeton University Press, Princeton, p. 113-138.

Rigby, B., Lasdon, L. S., Warren, A. D. (1995): The evolution of Texacos blending systems: From OMEGA to StarBlend. *Interfaces*, 25: 64-83.

Ruszczynski, A., Shapiro, A. (2003): Stochastic Programming. *Handbooks in Operations Research and Management Science*, Vol. 10, Elsevier Science, Amsterdam.

Shapiro, A., A. Kleywegt. (2002): Minimax analysis of stochastic programs. *Optimization Methods and Software*, 17(3), p. 523542.

Simon, H. A. (1959): Theories of decision making in economics and behavioral sciences. *American Economic Review*, 49(3): 253-283.

Žáčková, J. (1966): On minimax solution of stochastic linear programming problems. *Časopis pro Pěstování Matematiky*, 91, p. 423–430.