A Satisficing Model for Project Selection

Nicholas G. Hall *†

Zhuoyu Long **

Jin Qi **

Melvyn Sim ‡

* Department of Management Sciences, The Ohio State University
† Corresponding author: hall.33@osu.edu
** NUS Business School, National University of Singapore
‡ NUS Business School and Risk Management Institute, National University of Singapore

December 29, 2009
Abstract

This paper considers the problem of project selection, subject to a budget and other constraints. The return of each individual project is uncertain, and its probability distribution is only partially characterized. The criterion for project selection is based on a satisficing approach that evaluates how well the uncertain total return achieves a prespecified target return, considering the diversification preference of the decision maker. This criterion is implemented by maximizing the entropic satisficing measure. Thus, our model maximizes the value of a risk aversion parameter, while keeping the certainty equivalent of the uncertain returns under exponential utility above a given target. We allow interactions, for example synergies, between the returns of different projects. We also allow correlation between the uncertain returns of different projects. Our solution procedure solves a small sequence of subproblems via binary search and applies a cutting plane procedure to test feasibility at each subproblem. An extensive computational study shows that the satisficing model generates more effective project portfolios than any other project selection approach. We describe a simple and easily implemented greedy heuristic for the subproblem. Computational tests show that the use of this heuristic at the subproblem routinely provides close to the same return as the cutting plane approach. Finally, we identify several managerial insights from our results.

*Key words and phrases:* project management; project selection with uncertain returns; satisficing; greedy heuristic.

*OR/MS Index 1989 Subject Classification:*
Project management.
Production/scheduling: planning; approximations/heuristic; applications.
Probability: applications; entropy.
1 Introduction

A problem of central importance in project management is the selection of suitable projects. Both statistical and anecdotal evidence suggest that “doing the right projects” is a big factor in “doing projects right”. For example, well chosen projects are typically relatively easy to manage. Whereas, poorly selected projects are often doomed from the start and may even weaken the performance of other projects by draining resources away from them. At first sight, the problem of selecting projects seems quite simple. For each available project, a decision needs to be made either to accept it or to reject it. However, in many situations, practical complications prevent this decision being made independently for each project.

One complication that motivates an overall portfolio approach to project selection relates to shared resources; for example, there may be enough available budget in a particular year to fund either project A or project B, but not both. A second complication is that there may be interactions between projects; for example, the return from selecting project A may be greater if project B is also selected. A third complication is that a company may have overall priorities for its project portfolio; for example, it is often important to reduce overall risk through diversification of the selected projects. Since these complications imply the need to make selection decisions about multiple projects simultaneously, many combinations of projects need to be evaluated and compared. The most effective and efficient way of making all these comparisons accurately, while incorporating all the relevant constraints, is the use of optimization. Ideally, therefore, project selection decisions would be made using deterministic optimization models.

Herbots et al. (2007) provide a broad overview of different models for project selection. A widely used approach is application of the deterministic zero-one knapsack problem (Kellerer et al. 2004) and related models. An overview of project selection using such approaches is provided by Weber et al. (1990). Generalizations of the knapsack model (Fox et al. 1984, Dickinson et al. 2001) model interactions between the returns of different projects.

However, the use of deterministic optimization models is often inappropriate, due to uncertainty in the return from each project. This uncertainty is the principal motivation for our work. Uncertainty in the return from a project may be classified into two main categories; uncertainty regarding its technical success, and uncertainty regarding its commercial
success. The main sources of uncertainty that influence technical success include outcomes in research and development (e.g., for new products), in prototype testing (e.g., for automobile safety testing) and in regulatory approval (e.g., for new drugs). The main sources of uncertainty that influence commercial success are randomness in time to market (e.g., for seasonal products such as fashion items and toys), in the introduction of competitors’ products (e.g., for consumer electronics) and in general economic factors (e.g., a recession).

Jørgensen (1999) provides an overview of applications of financial theory to project selection. Mean-variance models (Markowitz 1959) are an important example. These models rely either on utility functions being risk averse and quadratic, or on investment returns being normally distributed. In the context of project selection, however, both these assumptions are questionable. As an alternative, Ringuest et al. (2004) study the selection of research and development projects using the mean-Gini model, an approach proposed by Shalit and Yitzhaki (1984) that, under a specific condition, is consistent with second-order stochastic dominance. However, it is not necessarily easy to compute and optimize the mean-Gini model. The authors assume that the underlying distributions are simple, discrete and known, which enables them to design an optimal enumerative approach to project portfolio selection.

Henig (1990) models the problem of selecting projects as a stochastic knapsack problem, where projects have random returns that are independent and normally distributed. He develops algorithms for the problems of (a) maximizing the probability that the return will achieve a given target, and (b) maximizing a given $\alpha$ percentile of the return, which is essentially the same as minimizing the $1 - \alpha$ value-at-risk (Föllmer and Schied 2004). For the latter problem, Atamtürk and Narayanan (2008) provide a linear characterization of the submodular mean-risk minimization objective.

Papastavrou et al. (1996) consider a dynamic version of the zero-one knapsack problem, where projects of random investment cost and return arrive over time according to a stochastic process. The objective is to determine the optimal policy for accepting or rejecting each project, so as to maximize the expected total return under risk neutrality, subject to a budget on investment cost. It is shown that the optimal policy is of threshold type. That is, a newly arrived project is accepted if and only if there is enough remaining budget to fund its cost, and further, the expected total return with the new project exceeds that without it. Similar results are independently derived by Lu et al. (1999). Kleywegt and Papastavrou
(1998) generalize the above results to allow for a waiting time cost between the acceptance of a project and the decision to terminate the selection process. Extensions to allow for random resource requirements for the projects are discussed in Kleywegt and Papastavrou (2001).

Dean et al. (2008) study a different variant of the knapsack problem, where project returns are deterministic and project costs sizes are independent random variables with arbitrary, known distributions. The objective is to identify a policy that maximizes the risk neutral expected total return of the projects that are accepted. Policies that use information about realized sizes to inform later choices are called adaptive. The authors describe fast adaptive and nonadaptive algorithms, and discuss their worst case performance.

Classical robust optimization can be used to address the project selection problem in which the uncertain returns of projects are described by an uncertainty set without distributional information; see Soyster (1973), Ben-Tal and Nemirovski (1998) and Bertsimas and Sim (2004). Kouvelis and Yu (1997) describe various decision criteria for robust selection including maximin and regret based ones. However, when applied to discrete optimization problems, these criteria can fundamentally elevate their computational complexities and render polynomially time solvable ones intractable. Bertsimas and Sim (2003) show that the adjustable polyhedral uncertainty set of Bertsimas and Sim (2004) preserves the computational complexity of the underlying discrete optimization problems. Strong formulations of the robust discrete optimization problem are characterized by Atamtürk (2007).

Liesiö et al. (2007) consider a project selection problem where each project is evaluated with respect to multiple criteria. The return of a project is given by its total weighted score on all criteria. However, it is difficult to estimate exact values for the factor weights and the project scores. Consequently, both the weights and the scores are modeled as falling within known intervals. The authors describe an algorithm to compute all nondominated portfolios of projects. They apply their model to a problem of selecting a portfolio from among 50 available road improvement projects in Finland, based on four relevant criteria.

In this paper, we consider the problem of making optimal project portfolio decisions in the presence of uncertainty about project returns. The decision maker in our work is a senior manager who makes decisions about the acceptance or rejection of the available projects. Our problem definition allows for correlation between the returns of different projects, which
is relevant in practice where they use the same resources or are subject to the same external challenges. We also allow for interaction effects, for example synergistic return, when multiple projects are selected.

Our work is related to Henig’s model of maximizing the probability of achieving a target return or aspiration level. In articulating the risk behaviors of real world managers, several descriptive studies (Lanzilloti 1958, Mao 1970, Payne et al. 1980, 1981) find that aspiration levels are a key driver of decision making. Simon (1955), who coined the term “satisfice”, argues that the main goal of most firms is not maximizing return but rather attaining a target return. In the same spirit, we adopt a satisficing approach in the project selection problem, which allows the decision maker to specify a desired aspiration level or target return. A natural satisficing model is to maximize the probability of achieving a target return. However, Diecidue and van de Ven (2005) argue against the use of success probability alone, since it does not consider the amount of shortfall relative to a target that is missed.

Our criterion for project selection is based on the satisficing approach introduced by Brown and Sim (2009), which evaluates how well the total return achieves the specified target return under uncertainty, taking into account the diversification preference of the decision maker. This criterion is implemented by maximizing the entropic satisficing measure. Our model maximizes the value of a risk aversion parameter, while keeping the certainty equivalent of the uncertain returns under exponential utility above the target. Thus, the satisficing measure identifies the most risk averse project portfolio that meets a given risk constraint with respect to a target for the total return. Our methodology for solving the problem uses binary search on the risk aversion parameter. At each value of the risk aversion parameter, we apply a cutting plane approach to approximate the objective function of the subproblem. Our computational results show that the satisficing model identifies high quality project portfolios, relative to expected utility maximization. Also, we design a simple and easily implemented heuristic for use at the subproblem. This heuristic is shown computationally to identify portfolios that are close in performance to those achieved by the cutting plane approach.

This paper is organized as follows. Section 2 describes the development of our satisficing model for project selection with uncertain returns. Section 3 describes a solution procedure for this model. In Section 4, we model ambiguity about the probability distribution that
governs uncertainty in project returns. Section 5 describes a simple heuristic for solving the satisficing model, along with a numerical example. Section 6 contains a computational study that compares the performance of our model against various alternative project selection approaches, some sensitivity analysis results, and a study of the effectiveness of the heuristic from Section 5. Finally, Section 7 provides a conclusion and some suggestions for future research.

2 Satisficing Model for Project Selection

In this section, we formally define the project selection problem to be studied, including the issues of correlation and interaction, and then describe a satisficing model for that problem.

Consider a set \( N = \{1, \ldots, n\} \) of available projects. No additional projects become available during the planning horizon. Project selection decisions are expressed as \( s = (s_1, \ldots, s_n) \), where \( s_j = 1 \) means that the \( j \)th project is accepted, and \( s_j = 0 \) means that the \( j \)th project is rejected. We denote the total return of the portfolio by \( \pi(s) \). If the returns of projects are independent of one another, then the total return of the project portfolio is simply \( \pi(s) = \sum_{j=1}^{n} r_j s_j \), where \( r_j, j = 1, \ldots, n \), is the return of project \( j \) after it is selected. However, it is common that projects interact with each other. For example, if both projects \( i \) and \( j \) are selected, a synergistic return may be earned. This synergistic return can be negative, to represent conflict between two or more projects. We allow interactions between sets of projects with any cardinality. Hence, we define a project bundle \( e \subseteq N \), and let \( r_e \) represent the total return from selecting this project bundle \( e \). We define a set \( E \) as the project bundle set, which is a minimal representation of the projects and their interactions.

**Definition 1** The project bundle set, \( E \) has the following properties:

1. \( \{i\} \in E, \forall i \in N \).

2. For all \( e_1 \in E, e_2 \in E \) where \( e_1 \cap e_2 \neq \emptyset \), \( e_1 \cup e_2 \in E \).

The first property enforces consistency with the simple case without synergistic returns. The second property shows that if two project bundles in set \( E \) share common projects, then their union, which is also a project bundle, belongs to the set \( E \). Moreover, the associated return of the union project bundle is also defined. In keeping with practical project selection situations, we assume that the cardinality of \( E \) is polynomial in size.
The total return corresponding to selection \( s \) is given by
\[
\pi(s) = \sum_{e \in \mathcal{E}} r_e \beta(e, s),
\]
where \( \beta(e, s) \) is an indicator function of bundle \( e \) from the selection \( s \) defined as
\[
\beta(e, s) = \begin{cases} 
1, & \text{if } s_j = 1, \forall j \in e, \text{ and } \nexists g \in \mathcal{E}, g \supset e \text{ such that } s_j = 1, \forall j \in g; \\
0, & \text{otherwise.}
\end{cases}
\]
The problem faced by the decision maker is the maximization of total return from the project portfolio, subject to various practical constraints. This problem can be written as
\[
\max_{s \in \mathcal{S}} \pi(s) \quad \text{s.t.} \quad s \in \mathcal{S}
\]
where
\[
\mathcal{S} = \left\{ s : A s \leq b \right\}.
\]
In the above problem, the objective requires the maximization of the total return of the portfolio. The first constraint set in \( \mathcal{S} \) requires that the selection of projects meets various deterministic restrictions. In practice, these typically include a budget constraint on total project cost, and possibly some portfolio diversification constraints and logical constraints that link the decisions to accept different projects. Examples of these and other practical constraints are discussed by Beaujon et al. (2001) and Loch et al. (2001). The second constraint ensures that each project is either accepted in full, or rejected.

We model a common practical situation where the constraints discussed above are deterministic, but the return from a selected project is uncertain. Therefore, we denote by \( \tilde{r}_e \), instead of \( r_e \) above, the uncertain return of project bundle \( e \in \mathcal{E} \). We now specify our model of uncertainty, which is defined by a state-space \( \Omega \) and a sigma-algebra \( \mathcal{F} \) of events in \( \Omega \). We further assume that each uncertain project bundle return is affinely dependent on a finite set of \( K \) bounded and independently distributed factors \( \tilde{z}_k : \Omega \rightarrow \mathbb{R}, k = 1, \ldots, K \).

We denote by \( \mathcal{V} \) the set of feasible returns, \( \tilde{r} : \Omega \rightarrow \mathbb{R} \) as follows
\[
\mathcal{V} = \left\{ \tilde{v} : \exists (v_0, v) \in \mathbb{R}^{K+1} : \tilde{v}(\omega) = v^0 + \sum_{k=1}^{K} v^k \tilde{z}_k(\omega), \ \forall \omega \in \Omega \right\}.
\]
Specifically, each uncertain project bundle return \( \tilde{r}_e \) is an element in \( \mathcal{V} \) given by
\[
\tilde{r}_e = r^0_e + \sum_{k=1}^{K} r^k_e \tilde{z}_k,
\]
where the factor coefficients $r_{e}^0, \ldots, r_{e}^k$ are given. The uncertain factors may represent the future state of the general economy or of the specific project marketplace. Returns can be correlated if there exists a subset of common factors in which the respective coefficients are nonzeros. Under the uncertainty model, the total return from selection $s$ is

$$\tilde{\pi}(s) = \sum_{e \in \mathcal{E}} \tilde{r}_e \beta(e, s) = \sum_{e \in \mathcal{E}} (r_{e}^0 + \sum_{k=1}^{K} r_{e}^k \tilde{z}_k) \beta(e, s) = \sum_{e \in \mathcal{E}} r_{e}^0 \beta(e, s) + \sum_{k=1}^{K} \left( \sum_{e \in \mathcal{E}} r_{e}^k \beta(e, s) \right) \tilde{z}_k,$$

which is also an element in $\mathcal{V}$.

We do not assume a specific probability measure on $\Omega$ in our uncertainty model. Instead of specifying a probability distribution, we permit ambiguity and assume that the true distribution, $\mathbb{P}$, lies in a family of distributions denoted by $\mathbb{F}$, which we discuss below. We use $\mathbb{E}_{\mathbb{P}}(\cdot)$ to denote the expectation over the distribution, $\mathbb{P}$. We assume the following problem

$$\sup_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}}(\exp(\alpha \tilde{z}_k))$$

is solvable for all $k = 1, \ldots, K$, and $\alpha \in \mathbb{R}$, and there exists a distribution $\mathbb{Q}$ such that $\mathbb{Q} \in \arg\sup_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}}(\exp(\alpha \tilde{z}_k))$. We further assume there exists a computationally efficient method for doing so. We show below that this assumption holds for some family of distributions $\mathbb{F}$.

The most common approach in evaluating projects under uncertainty is the maximization of expected total return (Papastavrou et al. 1996, Herbots et al. 2007). However, in practice, few managers consider return without considering risk. Classical von Neumann-Morgenstern expected utility theory can be used to evaluate projects under uncertainty. When applied to financial decision making, utility functions are usually restricted to increasing concave functions, which is equivalent to risk aversion and diversification preference. Moreover, an optimized portfolio under such a utility function is consistent with second order stochastic dominance; see for instance Levy (2006). Nevertheless, there are several issues with expected utility that limit its applicability in practice. Some of these issues are aggravated for project selection because the decision variables are discrete, as we now discuss.

1. **Parameter specification.** Expected utility is specified by a risk tolerance parameter, which is conceptually abstract and difficult to solicit accurately from a decision maker.

2. **Computational tractability.** Evaluating the expected utility over a sum of random variables involves multi-dimensional integration, which is an intractable problem; see
Nemirovski and Shapiro (2006). Moreover, assumptions such as normality that improve problem tractability limit the modeling flexibility of expected utility and lead to inconsistency in stochastic dominance. For example, the assumption of a quadratic utility function implies that the decision maker is indifferent between upside gains and downside losses, which is inconsistent with rational choices. Furthermore, the project selection problem, even in the deterministic case, is \(NP\)-hard. Hence, incorporating uncertainty requires solution of an intractable problem.

3. **Distributional ambiguity.** A particular issue in the evaluation of projects is their uniqueness, which results in a lack of historical information to elicit the actual distributions of their returns. When dealing with distributional ambiguity, it is well known from Ellsberg’s Paradox (1961) that the paradigms of expected utility theory and the subjective expected utility of Savage (1954) are inconsistent with behavioral choices. This issue can be addressed by the ambiguity averse expected utility of Gilboa and Schmeidler (1989) and by modern risk measures (Föllmer and Schied 2004). Our work extends and applies these ideas.

4. **Behavioral issues.** Despite the widespread use of expected utility, it fails to explain well known behavioral choices, even when distributions are known (Allais 1953). Some of these behavioral issues may be addressed by the more sophisticated prospect theory of Kahneman and Tversky (1979) and the rank-dependent utility model of Quiggin (1993). However, these models require computing various convex combinations of the random variables. This greatly reduces the computational tractability of the model. Hence, these approaches are difficult to apply to the project selection problem.

To address these issues collectively, Brown and Sim (2009) propose a new satisficing criterion, the *entropic satisficing measure*, that considers behavioral preferences such as diversification and ambiguity aversion. The entropic satisficing measure is defined on \(\mathcal{V}\) as follows.

**Definition 2** Given a target, \(\tau \in \mathbb{R}\), the entropic satisficing measure, \(\rho_\tau : \mathcal{V} \rightarrow [0, \infty]\) is defined by

\[
\rho_\tau(\bar{v}) = \begin{cases} 
\sup\{\alpha > 0 : C_\alpha(\bar{v}) \geq \tau\} & \text{if feasible}, \\
0 & \text{otherwise},
\end{cases}
\]

8
where the function $C_{\alpha} : \mathcal{V} \to \mathbb{R}$ is defined as

$$C_{\alpha}(\tilde{v}) = \inf_{P \in \mathcal{F}} \left( -\frac{1}{\alpha} \ln \left( \mathbb{E}_{P} \left( \exp \left( -\alpha \tilde{v} \right) \right) \right) \right) = -\frac{1}{\alpha} \ln \left( \sup_{P \in \mathcal{F}} \mathbb{E}_{P} \left( \exp \left( -\alpha \tilde{v} \right) \right) \right).$$

To understand the satisficing properties of $\rho_\tau$, observe that if $\tilde{v}$ always achieves the target, i.e., $\tilde{v}(\omega) \geq \tau$ for all $\omega \in \Omega$, then $C_{\alpha}(\tilde{v}) \geq \tau$ for all $\alpha > 0$, and hence $\rho_\tau(\tilde{v}) = \infty$. On the other hand, if $\tilde{v}$ never achieves the target, then $C_{\alpha}(\tilde{v}) < \tau$ for all $\alpha > 0$, and hence $\rho_\tau(\tilde{v}) = 0$.

The function $C_{\alpha}(\tilde{v})$ is the certainty equivalent of the uncertain return $\tilde{v}$ under Gilboa and Schmeidler’s (1989) worst-case exponential utility, and the negative of $C_{\alpha}(\tilde{v})$ is also a convex risk measure (Föllmer and Schied 2004). Moreover, if $\tilde{v}$ is normally distributed, then $C_{\alpha}(\tilde{v})$ becomes the Markowitz mean-variance measure of uncertain returns given by

$$C_{\alpha}(\tilde{v}) = \mu_{\tilde{v}} - \frac{\alpha}{2} \sigma_{\tilde{v}}^2,$$

where $\mu_{\tilde{v}}$ and $\sigma_{\tilde{v}}^2$ are the mean and variance of $\tilde{v}$, respectively.

**Lemma 1** For any $\tilde{v} \in \mathcal{V}$, $C_{\alpha}(\tilde{v})$ is non-increasing in $\alpha > 0$. Moreover,

$$\lim_{\alpha \downarrow 0} C_{\alpha}(\tilde{v}) = \inf_{P \in \mathcal{F}} \mathbb{E}_{P}(\tilde{v}). \quad (1)$$

**Proof.** For all $\alpha_1 > \alpha_2 > 0$, by Jensen’s inequality,

$$C_{\alpha_1}(\tilde{v}) = \inf_{P \in \mathcal{F}} \left( -\frac{1}{\alpha_1} \ln \left( \mathbb{E}_{P} \left( \exp \left( -\alpha_1 \tilde{v} \right) \right) \right) \right) \leq \inf_{P \in \mathcal{F}} \left( -\frac{1}{\alpha_1} \ln \left( \mathbb{E}_{P} \left( \exp \left( -\alpha_2 \tilde{v} \right) \right) \right) \right) = C_{\alpha_2}(\tilde{v}).$$

Hence, to establish the equality (1), it suffices to show that $\lim_{\alpha \downarrow 0} C_{\alpha}(\tilde{v})$ is bounded below by $\inf_{P \in \mathcal{F}} \mathbb{E}_{P}(\tilde{v})$. Observe that $\tilde{v} \in \mathcal{V}$ is bounded, since it is affinely dependent of a finite set of bounded uncertain factors. Hence, there exists $v > 0$ such that $|\tilde{v}(\omega)| \leq v$ for all $\omega \in \Omega$. 
Indeed, \( \forall \mathbb{P} \in \mathcal{F}, \alpha > 0 \), by Taylor’s theorem,

\[
E_{\mathbb{P}}(\exp(-\alpha \tilde{v})) = 1 + E_{\mathbb{P}}(-\tilde{v})\alpha + \sum_{n=2}^{\infty} \frac{1}{n!} E_{\mathbb{P}}((-\tilde{v})^n) \alpha^n \\
\leq 1 - E_{\mathbb{P}}(\tilde{v})\alpha + \sum_{n=2}^{\infty} \frac{1}{n!} E_{\mathbb{P}}(v^n) \alpha^n \\
= 1 - E_{\mathbb{P}}(\tilde{v})\alpha + \exp(v\alpha) - 1 - v\alpha \\
= \exp(v\alpha) - (v + E_{\mathbb{P}}(\tilde{v}))\alpha \\
\leq \exp(v\alpha) - (v + \mu)\alpha,
\]

where \( \mu = \inf_{\mathbb{P} \in \mathcal{F}} E_{\mathbb{P}}(\tilde{v}) \). It follows that

\[
\lim_{\alpha \downarrow 0} C_{\alpha}(\tilde{v}) = \lim \inf_{\alpha \downarrow 0} \frac{1}{\alpha} \ln \left( E_{\mathbb{P}}(\exp(-\alpha \tilde{v})) \right) \\
\geq \lim_{\alpha \downarrow 0} \left( \frac{1}{\alpha} \ln \left( \exp(v\alpha) - (v + \mu)\alpha \right) \right) \\
= \lim_{\alpha \downarrow 0} \left( \frac{- \exp(v\alpha)v - (v + \mu)}{\exp(v\alpha) - (v + \mu)\alpha} \right) \\
= \mu.
\]

The second equality follows from L’Hôpital’s rule. \( \blacksquare \)

Since \( C_{\alpha}(\tilde{v}) \) is non-increasing in \( \alpha \), \( \rho_{\tau}(\tilde{v}) \) can be found via binary search on \( \alpha \). We interpret the entropic satisficing measure, \( \rho_{\tau}(\tilde{v}) \), as the highest risk aversion parameter for which the certainty equivalent of \( \tilde{v} \) meets the minimum target, \( \tau \). When \( \tau = 0 \), the reciprocal of the entropic satisficing measure is equivalent to the riskiness index of Aumann and Serrano (2008). Brown et al. (2009) show that many satisficing measures resolve classical paradoxes including those of Allais (1953) and Ellsberg (1961), while also being consistent with second order stochastic dominance.

The project selection problem under the entropic satisficing measure is given by

\[
\rho_{\tau}^* = \max_{s \in \mathcal{S}} \rho_{\tau}(\tilde{\pi}(s)) \\
\text{s.t.} \quad \rho_{\tau}(\tilde{\pi}(s)) \geq \tau \quad \text{for } s \in \mathcal{S}, \quad \alpha > 0.
\]

or equivalently,

\[
\rho_{\tau}^* = \max_{s \in \mathcal{S}} \frac{\alpha}{\alpha} \quad \text{s.t.} \quad C_{\alpha}(\tilde{\pi}(s)) \geq \tau \quad \text{for } s \in \mathcal{S}, \quad \alpha > 0.
\]
In setting the target $\tau$, we assume that there exists a feasible project selection $s \in \mathcal{S}$ for which the expected return of the project under ambiguity aversion exceeds $\tau$, i.e,

$$\inf_{P \in \mathcal{F}} \mathbb{E}_P(\tilde{\pi}(s)) > \tau.$$  

This condition implies that the target level is strictly achievable in expectation under ambiguity aversion. It tacitly implies that the decision maker is risk and ambiguity averse and hence does not set an unattainable goal. Now, from Lemma 1, there exists an $\alpha > 0$ such that

$$C_\alpha(\tilde{\pi}(s)) \geq \tau.$$  

Hence, $\rho^*_\tau > 0$.

### 3 Algorithm for the Satisficing Model

Brown and Sim (2009) show that the problem of maximizing the satisficing measure can be solved using a binary search algorithm if the following subproblem can be solved efficiently:

$$\max_{s \in \mathcal{S}} C_\alpha(\tilde{\pi}(s))$$  

subject to

$$s \in \mathcal{S}. \quad (3)$$

We propose a mixed integer programming model and a solution procedure for this problem.

Since $\tilde{\pi}(s)$ is an element in $\mathcal{V}$ and affinely depends on the independent uncertain factors $\tilde{z}_1, \ldots, \tilde{z}_K$, we can decompose $C_\alpha(\tilde{\pi}(s))$ as follows.

**Lemma 2** For all $\alpha > 0$, $C_\alpha(\tilde{\pi}(s))$ can be decomposed as

$$C_\alpha(\tilde{\pi}(s)) = \sum_{k=0}^{K} C_\alpha \left( \tilde{z}_k \sum_{e \in \mathcal{E}} r^k_e \beta(e, s) \right),$$

where we let $\tilde{z}_0 = 1$ for simplicity.
Proof. Note that \( \tilde{z}_0, \ldots, \tilde{z}_K \) are independently distributed. Therefore,

\[
C_\alpha(\tilde{\pi}(s)) = \frac{1}{\alpha} \ln \left( \sup_{P \in \mathcal{F}} \mathbb{E}_P \left( \exp \left( -\alpha \sum_{k=0}^{K} \left( \sum_{e \in \mathcal{E}} r^k_e \beta(e, s) \tilde{z}_k \right) \right) \right) \right)
\]

\[
= -\frac{1}{\alpha} \ln \left( \prod_{k=0}^{K} \sup_{P \in \mathcal{F}} \mathbb{E}_P \left( \exp \left( -\alpha \tilde{z}_k \sum_{e \in \mathcal{E}} r^k_e \beta(e, s) \right) \right) \right)
\]

\[
= \sum_{k=0}^{K} -\frac{1}{\alpha} \ln \left( \sup_{P \in \mathcal{F}} \mathbb{E}_P \left( \exp \left( -\alpha \tilde{z}_k \sum_{e \in \mathcal{E}} r^k_e \beta(e, s) \right) \right) \right)
\]

\[
= \sum_{k=0}^{K} C_\alpha \left( \tilde{z}_k \sum_{e \in \mathcal{E}} r^k_e \beta(e, s) \right). \quad \square
\]

Using the decomposition in Lemma 2, we now evaluate the computational tractability of the problem.

**Lemma 3** If there are no interactions between the projects, and their returns are independent, then the objective function in problem (3) is linear in \( s \).

Proof. In the case of no interactions, the project bundle set is the same as the available project set, i.e., \( \mathcal{E} = \mathcal{N} \), and \( \forall i \in \mathcal{N}, \beta(i, s) = s_i \).

For all \( i \in \mathcal{N} \), let \( \mathcal{I}_i = \{ k : r^k_i \neq 0, k = 1, \ldots, K \} \), i.e., \( \tilde{r}_i = r^0_i + \sum_{k \in \mathcal{I}_i} r^k_i \tilde{z}_k \). Also, \( \forall i, j \in \mathcal{N} \) and \( i \neq j \), we have \( \mathcal{I}_i \cap \mathcal{I}_j = \emptyset \), since returns are independent. From Lemma 2,

\[
C_\alpha(\tilde{\pi}(s)) = C_\alpha \left( \tilde{z}_0 \sum_{i \in \mathcal{N}} r^0_i s_i \right) + \sum_{k=1}^{K} C_\alpha \left( \tilde{z}_k \sum_{i \in \mathcal{N}} r^k_i s_i \right)
\]

\[
= \sum_{i \in \mathcal{N}} r^0_i s_i + \sum_{i \in \mathcal{N}} \sum_{k \in \mathcal{I}_i} C_\alpha \left( \tilde{z}_k r^k_i s_i \right)
\]

\[
= \sum_{i \in \mathcal{N}} \left( r^0_i s_i + \sum_{k \in \mathcal{I}_i} C_\alpha \left( \tilde{z}_k r^k_i s_i \right) \right)
\]

\[
= \sum_{i \in \mathcal{N}} \left( r^0_i + \sum_{k \in \mathcal{I}_i} C_\alpha \left( \tilde{z}_k r^k_i \right) \right) s_i,
\]

where the last equality holds since \( s_i \in \{0, 1\} \). \( \square \)

The above results show that, without interactions and with independent uncertain returns, problem (3) retains the complexity of the nominal problem, i.e. \( \max_{s \in \mathcal{S}} r' s \). Specifically, if the nominal problem is solvable in polynomial time, then problem (3) is also polynomially solvable. Furthermore, problem (2) can also be solved in polynomial time using binary search. For example, from Lemma 3, if the feasible set of \( s \) is a uniform matroid, i.e.,
\( S = \{s \in \{0, 1\}^n : \sum_{k \in \mathcal{N}} s_k = m\} \), we can solve problem (3) in \( O(n \log n) \) time by sorting \( (\nu_i^0 + \sum_{k \in \mathcal{I}_i} C_\alpha (\tilde{z}_k r^k_i)), i = 1, \ldots, n \), in non-increasing order and choosing the first \( m \) projects.

Unfortunately, as we now show, the correlation effect fundamentally increases the computational complexity of the problem. We assume the absence of interactions, in order to focus solely on the role of correlation. Hence, \( \mathcal{E} = \mathcal{N} \). If the uncertain returns are correlated with each other, then we have the following result.

**Theorem 1** If there are no interactions between the projects, then the recognition version of problem (2) with correlated returns under a uniform matroid is NP-complete.

Proof. By reduction from the following NP-Complete problem (Garey and Johnson 1979).

**Equal Cardinality Partition:** Given a finite set \( \mathcal{N} \) of even cardinality \( n \), with size \( c_k \in \mathbb{Z}^+ \) for each \( k \in \mathcal{N} \), determine if there exists a partition \( \mathcal{N}_1 \) and \( \mathcal{N}_2 \) such that \( |\mathcal{N}_1| = |\mathcal{N}_2| = n/2 \) and \( \sum_{k \in \mathcal{N}_1} c_k = \sum_{k \in \mathcal{N}_2} c_k \).

Under a uniform matroid, i.e., \( S = \{s \in \{0, 1\}^n : \sum_{k \in \mathcal{N}} s_k = m\} \), we construct an instance of problem (2) with

\[
\tau = \frac{1}{2} \sum_{k \in \mathcal{N}} c_k \\
\tilde{r}_k = \frac{2\tau}{n} + (c_k - \frac{2\tau}{n}) \tilde{z}, \quad \forall k \in \mathcal{N} \\
m = \frac{n}{2}
\] (4) (5) (6)

Equation (4) determines a specific target to achieve. Equation (5) defines a special type of uncertain return, where the return of each project is determined by a common uncertain factor \( \tilde{z} \). We assume that \( \tilde{z} \) is +1 or -1, with equal probability. Equation (6) implies that the only feasible solutions select \( n/2 \) projects. In this instance,

\[
C_\alpha(\tilde{\pi}(s)) = -\frac{1}{\alpha} \ln \left( \frac{1}{2} \exp \left( -\alpha \sum_{k \in \mathcal{N}} c_k s_k \right) + \frac{1}{2} \exp \left( -\alpha \sum_{k \in \mathcal{N}} c_k (1 - s_k) \right) \right).
\]

We prove that there exists a selection for this instance such that the objective value is infinite, if and only if there exists a solution to Equal Cardinality Partition.

(\( \Rightarrow \)) Suppose there exists a solution \( \mathcal{N}_1 \) and \( \mathcal{N}_2 \) to Equal Cardinality Partition such that \( \sum_{k \in \mathcal{N}_1} c_k = \sum_{k \in \mathcal{N}_2} c_k = \tau \) and \( |\mathcal{N}_1| = |\mathcal{N}_2| = n/2 \). Then we select all the projects in \( \mathcal{N}_1 \),
and reject all the projects in \( N_2 \), i.e., \( s_k = 1, \forall k \in N_1 \) and \( s_k = 0, \forall k \in N_2 \). Thus, \( \forall \alpha > 0, \)

\[
C_\alpha(\pi(s)) = -\frac{1}{\alpha} \ln \left( \frac{1}{2} \exp \left( -\alpha \tau \right) + \frac{1}{2} \exp \left( -\alpha \tau \right) \right) = \tau.
\]

Therefore, the objective value of this instance is infinite.

(\( \Leftarrow \)) Suppose there exists a portfolio selection \( s^* \) such that the objective value is infinite, i.e., \( \forall \alpha > 0, C_\alpha(\pi(s^*)) \geq \tau \). Since, given \( \alpha > 0 \), the function \( \exp(-\alpha x) \) is strictly convex in \( x \), we have

\[
C_\alpha(\pi(s^*)) = -\frac{1}{\alpha} \ln \left( \frac{1}{2} \exp \left( -\alpha \sum_{k \in N} c_k s_k^* \right) + \frac{1}{2} \exp \left( -\alpha \sum_{k \in N} c_k (1 - s_k^*) \right) \right)
\]
\[
\leq -\frac{1}{\alpha} \ln \left( \exp \left( -\frac{1}{2} \alpha \left( \sum_{k \in N} c_k s_k^* + \sum_{k \in N} c_k (1 - s_k^*) \right) \right) \right)
\]
\[
= \frac{1}{2} \left( \sum_{k \in N} c_k s_k^* + \sum_{k \in N} c_k (1 - s_k^*) \right)
\]
\[
= \tau,
\]

which implies that \( C_\alpha(\pi(s^*)) = \tau \). Furthermore, equality holds if and only if \( \sum_{k \in N} c_k s_k^* = \sum_{k \in N} c_k (1 - s_k^*) = \tau \). This implies that there exists a solution \( N_1 = \{ k : s_k^* = 1, \forall k \in N \} \) and \( N_2 = N \setminus N_1 \) for Equal Cardinality Partition. \( \square \)

We now provide an approach for finding an optimal solution to problem (3) where both interaction and correlation are allowed. First, we linearize the indicator function, \( \beta(e, s) \).

**Lemma 4**

\[
\sum_{e \in \mathcal{E}} r^k_e \beta(e, s) = x^t v^k, \quad k = 0, \ldots, K,
\]

where \( x \) is determined by \( s \),

\[
x_e \leq s_i, \quad e \in \mathcal{E}, \ i \in e
\]
\[
x_e + |e| \geq \sum_{i \in e} s_i + 1, \quad e \in \mathcal{E}
\]
\[
x \in \{0, 1\}^{|\mathcal{E}|};
\]

and vector \( v^k = (v^k_e)_{e \in \mathcal{E}}, v^k_e = r^k_e - \sum_{g \subseteq e, g \in \mathcal{E}} v^k_g, \ k = 0, \ldots, K. \)

Proof. Given \( s \), we note that for all \( e \in \mathcal{E}, x_e = 1 \) if and only if \( s_i = 1 \), for all \( i \in e \). Let \( \mathcal{G} = \{ e : \beta(e, s) = 1 \} \), hence \( \sum_{e \in \mathcal{E}} r^k_e \beta(e, s) = \sum_{e \in \mathcal{G}} r^k_e. \)
From the definition of \( x \), we observe that, for all \( e \in \mathcal{E} \),
\[
x_e = \begin{cases} 
1, & \text{if } \exists g \in \mathcal{G} \text{ such that } e \subseteq g; \\
0, & \text{otherwise}.
\end{cases}
\]
Therefore, for \( k = 0, \ldots, K \),
\[
x'v^k = \sum_{e \in \mathcal{E}} x_e v^k_e = \sum_{g \in \mathcal{G}} \sum_{e \in \mathcal{E}, e \subseteq g} v^k_e = \sum_{g \in \mathcal{G}} (v^k_g + \sum_{e \in \mathcal{E}, e \subseteq g} v^k_e),
\]
and using the definition of \( v^k_g \), we have
\[
x'v^k = \sum_{g \in \mathcal{G}} \left( v^k_g - \sum_{e \in \mathcal{E}, e \subseteq g} v^k_e \right) + \sum_{e \in \mathcal{E}, e \subseteq g} v^k_e = \sum_{g \in \mathcal{G}} r^k_g = \sum_{e \in \mathcal{E}} r^k_e \beta(e, s). \quad \Box
\]
From Lemmas 2 and 4, problem (3) can be reformulated as
\[
\max \quad \sum_{k=0}^K C_\alpha(\tilde{z}_k x'v^k) \\
\text{s.t.} \quad x \in \mathcal{X},
\tag{7}
\]
where \( \mathcal{X} \) is defined as
\[
\mathcal{X} = \left\{ x : \begin{array}{l}
    x_e + |e| \geq \max \left\{ \sum_{i \in e} s_i + 1, \forall i \in e \right\}, \\
    s \in S, \\
    x \in \{0, 1\}^{|\mathcal{E}|}
\end{array} \right\}.
\]
Next, we transform \( C_\alpha(\tilde{z}_k x'v^k) \) into a piecewise linear function, and solve it using a cutting plane algorithm.

**Lemma 5** For all \( x \in \mathcal{X} \),
\[
C_\alpha(x'v^k \tilde{z}_k) = \min_{u \in \mathcal{U}(v^k)} \left\{ C_\alpha(u \tilde{z}_k) + D^k_\alpha(u)(x'v^k - u) \right\}, \quad k = 0, \ldots, K,
\]
where \( D^k_\alpha(u) \) is a subgradient defined by
\[
D^k_\alpha(u) = \frac{\mathbb{E}_Q(\tilde{z}_k \exp(-\alpha u \tilde{z}_k))}{\mathbb{E}_Q(\exp(-\alpha u \tilde{z}_k))},
\]
\( Q \in \arg \sup_{F \in \mathcal{F}} \mathbb{E}_F(\exp(-\alpha u \tilde{z}_k)) \), and \( \mathcal{U}(v) = \{ x'v : x \in \mathcal{X} \} \).

**Proof.** For any \( u \in \mathcal{U}(v^k) \),
\[
C_\alpha(u \tilde{z}_k) - C_\alpha(x'v^k \tilde{z}_k) = \frac{1}{\alpha} \ln \left( \sup_{F \in \mathcal{F}} \mathbb{E}_F(\exp(-\alpha x'v^k \tilde{z}_k)) \right) - \frac{1}{\alpha} \ln \left( \sup_{F \in \mathcal{F}} \mathbb{E}_F(\exp(-\alpha u \tilde{z}_k)) \right)
\]
\[
\geq \frac{1}{\alpha} \left( \ln \mathbb{E}_Q(\exp(-\alpha x'v^k \tilde{z}_k)) - \ln \mathbb{E}_Q(\exp(-\alpha u \tilde{z}_k)) \right)
\]
\[
\geq \frac{1}{\alpha} \left( \frac{\partial}{\partial \gamma} \ln \mathbb{E}_Q(\exp(-\alpha \gamma \tilde{z}_k)) \right)_{\gamma = u} \times (x'v^k - u)
\]
\[
= -\frac{\mathbb{E}_Q(\tilde{z}_k \exp(-\alpha u \tilde{z}_k))}{\mathbb{E}_Q(\exp(-\alpha u \tilde{z}_k))} (x'v^k - u).
\]
The first inequality follows from the supremum property. The second inequality follows from the convexity of the entropic function $\ln \mathbb{E}_P (\exp(\cdot))$. Therefore, for all $u \in \mathcal{U}(v^k)$, we have

$$C_\alpha(x^k v^k z_k) \leq C_\alpha(u z_k) + D_\alpha^k(u)(x^k v^k - u).$$

Moreover, equality is achieved in (8) when $u = x^k v^k$. □

Under the assumption that $\sup_{P \in \mathcal{F}} \mathbb{E}_P (\exp(\alpha z_k))$ is solvable for all $\alpha \in \mathbb{R}$, we can calculate $C_\alpha(u z_k)$ and $D_\alpha^k(u), \forall u \in \mathcal{U}(v^k)$. Moreover, by transforming $C_\alpha(x^k v^k z_k)$ into a piecewise linear function of $(x^k v^k)$, we can still preserve the exact value at any feasible point since the problem is discrete.

Now, problem (7) can be reformulated as the mixed integer program

$$\max_{x,t} \sum_{k=0}^K t_k \quad \text{s.t.} \quad t_k \leq C_\alpha(u z_k) + D_\alpha^k(u)(x^k v^k - u), \quad k = 0, \ldots, K, u \in \mathcal{U}(v^k)$$

In some special cases, $\mathcal{U}(v^k)$ is polynomial in size, hence problem (9) can be solved efficiently. For example, if $\forall k = 0, \ldots, K, \ v^k = 1$, then $|\mathcal{U}(v^k)| \leq |\mathcal{E}| + 1$. However, in general, the size of $\mathcal{U}(v^k)$ may be exponential, which motivates the following algorithm.

**Algorithm CuttingPlane**

1. For each $k = 0, \ldots, K$, choose a subset $\mathcal{U}^k = \{x^k v^k\}$, for some $x \in \mathcal{X}$.

2. Solve the following subproblem

$$\max_{x,t} \sum_{k=0}^K t_k \quad \text{s.t.} \quad t_k \leq C_\alpha(u z_k) + D_\alpha^k(u)(x^k v^k - u), \quad k = 0, \ldots, K, u \in \mathcal{U}^k$$

and let the solution be $x^*$ and $t_k^*, k = 0, \ldots, K$.

3. If $t_k^* = C_\alpha(x^k v^k), \forall k = 0, \ldots, K$, then output the optimal value and optimal solution $x^*$, and stop.

4. For values of $k$ such that $t_k^* > C_\alpha(x^k v^k)$, add $(x^k v^k)$ into $\mathcal{U}^k$, and go to Step 2.

**Theorem 2** Algorithm CuttingPlane terminates in a finite number of steps and finds an optimal solution to the reformulated problem (9).
Proof. When Algorithm CuttingPlane terminates, it follows from Lemma 5 that
\[ t_k^* = C_\alpha(x^*v^k) \leq C_\alpha(u \tilde{z}_k) + D_\alpha^k(u)(x^*v^k - u), \quad \forall u \in U(v^k). \]
Hence, \( x^*, t_k^*, k = 0, \ldots, K \) is feasible in problem (9). Since problem (10) is a relaxation of problem (9), \( x^*, t_k^*, k = 0, \ldots, K \) is also optimal in problem (9).

Note that \( U(v^k) \) is a finite set for each \( k \). At each iteration, there exists at least one index \( k \) such that \( U^k \) increases. Therefore, Algorithm CuttingPlane terminates in a finite number of steps. \( \square \)

4 Model of Ambiguity

In this section, we describe the distributional ambiguity structure, and provide an approach to solve
\[ \sup_{P \in \mathcal{F}} \mathbb{E}_P \left( \exp((-\alpha \tilde{z})) \right). \]

The uncertain factor \( \tilde{z} \) has various possible outcomes. We assume that its outcome is based on \( J \) scenarios, such that in the \( j \)th scenario, \( \tilde{z} \) takes the value of \( \tilde{\zeta}_j \) with probability \( \tilde{p}_j \), where \( \tilde{\zeta}_j \) and \( \tilde{p}_j \) are both potentially uncertain. Additionally, \( \tilde{\zeta}_j \) is independent of one another.

For the uncertainty of probability \( \tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_J) \), the set of possible \( \tilde{p} \) vectors is denoted by \( \mathcal{P} \subseteq \{ p : p'1 = 1, p \geq 0 \} \), such that \( \tilde{p} \) follows the axiom of probability. In different applications, there can be various additional constraints with respect to \( \tilde{p} \).

Instead of complete knowledge of the distribution of \( \tilde{\zeta}_j \), we assume that its true distribution lies in some family of distributions \( \mathcal{F}_j \). The family of distributions on \( \tilde{z} \), \( \mathcal{F} \) is thus characterized by \( \mathcal{P}, \mathcal{F}_1, \ldots, \mathcal{F}_J \) such that
\[
\sup_{P \in \mathcal{F}} \mathbb{E}_P \left( \exp(-\alpha \tilde{z}) \right) = \sup_{\tilde{p} \in \mathcal{P}} \sum_{j=1}^J p_j \sup_{\mathcal{F}_j} \mathbb{E}_{P_j} \left( \exp(-\alpha \tilde{\zeta}_j) \right) = \sup_{\tilde{p} \in \mathcal{P}} \sum_{j=1}^J p_j \phi_j(\alpha),
\]
where we define
\[ \phi_j(\alpha) = \sup_{\mathcal{F}_j} \mathbb{E}_{P_j} \left( \exp(-\alpha \tilde{\zeta}_j) \right). \]
Using the above distributional ambiguity structure, we first calculate \( \phi_j(\alpha) \) under a specified family of distributions \( \mathcal{F}_j \), and then compute \( \sup_{\tilde{p} \in \mathcal{P}} \sum_{j=1}^J p_j \phi_j(\alpha) \). For example, if \( \mathcal{P} \) is a polytope, a simple linear program is sufficient to solve it.
Given a probability set $\mathcal{P}$, the estimation of the probabilities $\tilde{p}$ in real project selection problems is far from accurate, which compromises the accuracy of the model for real world problems. Kullback and Leibler (1951) introduce the KL divergence. This is an asymmetric measure of the difference between two probability distributions, the real unknown distribution $\tilde{p}$ and the reference distribution $q$. They specify that with a reference distribution $q$ and an uncertainty level $\theta$, the possible real distribution belongs to the set

$$\mathcal{P}_\theta = \left\{ p \in \Re_+^J : \sum_{j=1}^J p_j \ln \frac{p_j}{q_j} \leq \theta, \sum_{j=1}^J p_j = 1 \right\}. \quad (11)$$

It can easily be verified that when the uncertainty level is $\theta = 0$, we have $\tilde{p} = q$. Also, when $\theta \geq \max_j (-\ln q_j)$, $\tilde{p}$ can be any distribution that follows the axioms of probability.

We assume that the family of distributions $\mathbb{F}_j$ is an information set consisting of bounded support $\tilde{\zeta}_j \in [\zeta_j, \zeta_j^\ast]$, and mean support $[\mu_j, \bar{\mu}_j] \subseteq [\zeta_j, \zeta_j]$, such that

$$\mathbb{F}_j = \left\{ \mathbb{P}_j : \mathbb{P}_j(\tilde{\zeta}_j \in [\zeta_j, \zeta_j]) = 1, \mathbb{E}_{\mathbb{P}_j}(\tilde{\zeta}_j) \in [\mu_j, \bar{\mu}_j] \right\}.$$

**Theorem 3** Given the family of distributions $\mathbb{F}_j = \left\{ \mathbb{P}_j : \mathbb{P}_j(\tilde{\zeta}_j \in [\zeta_j, \zeta_j]) = 1, \mathbb{E}_{\mathbb{P}_j}(\tilde{\zeta}_j) \in [\mu_j, \bar{\mu}_j] \right\}$ and the probability feasibility set $\mathcal{P}_\theta = \{ p \in \Re_+^J : \sum_{j=1}^J p_j \ln \frac{p_j}{q_j} \leq \theta, \sum_{j=1}^J p_j = 1 \}$, we obtain for all $\alpha > 0,$

$$\phi_j(\alpha) = \sup_{\mathbb{P}_j \in \mathbb{F}_j} \mathbb{E}_{\mathbb{P}_j} \left( \exp \left( -\alpha \tilde{\zeta}_j \right) \right) = \frac{(\mu_j - \zeta_j) \exp(-\alpha \tilde{\zeta}_j) + (\zeta_j - \mu_j) \exp(-\alpha \tilde{\zeta}_j)}{\tilde{\zeta}_j - \zeta_j};$$

and $Q \in \arg \sup_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}} \left( \exp(-\alpha \tilde{z}) \right)$ such that its distribution is

$$\tilde{z} = \begin{cases} \zeta_j, & \text{with probability } \frac{\zeta_j - \mu_j}{\tilde{\zeta}_j - \zeta_j} p_j^* \forall j = 1, \ldots, J \\ \bar{\zeta}_j, & \text{with probability } \frac{\bar{\zeta}_j - \zeta_j}{\tilde{\zeta}_j - \zeta_j} p_j^* \forall j = 1, \ldots, J, \end{cases}$$

where

$$p_j^* = \frac{q_j \exp \left( \phi_j(\alpha)/\lambda^* \right)}{\sum_{i=1}^J q_i \exp \left( \phi_i(\alpha)/\lambda^* \right)}, \quad j = 1, \ldots, J,$$

and $\lambda^*$ is the optimal solution found from $\inf_{\lambda > 0} \left( \lambda \ln \left( \sum_{j=1}^J q_j \exp \left( \phi_j(\alpha)/\lambda \right) \right) \right.$ $\left. + \theta \lambda \right)$ by using a bisection algorithm.

Proof. First, we calculate $\phi_j(\alpha)$ by using a linear program. The information set of each possible scenario $\tilde{\zeta}_j$ has the same structure. Hence, we omit the subscript $j,$ and calculate
\[ \phi(\alpha) = \sup_{p \in P} \mathbb{E}_p \left( \exp(-\alpha \tilde{\zeta}) \right) \] by solving the following optimization problem:

\[
\begin{align*}
\phi(\alpha) &= \max_f \mathbb{E}_f \left( \exp(-\alpha \zeta) \right) \\
\text{s.t.} \quad &\mathbb{E}_f(1) = 1 \\
&\mathbb{E}_f(\zeta) \leq \bar{\mu} \\
&\mathbb{E}_f(\zeta) \geq \mu \\
&f(\zeta) \geq 0, \quad \forall \zeta \in [\underline{\zeta}, \bar{\zeta}].
\end{align*}
\]

We consider \( f \) to consist of infinite dimensional decision variables indexed by \([\underline{\zeta}, \bar{\zeta}]\). From duality,

\[
\phi(\alpha) = \min_{y_0, y_1, y_2} y_0 + \bar{\mu}y_1 - \mu y_2 \\
\text{s.t.} \quad y_0 + \zeta y_1 - \bar{\zeta} y_2 \geq \exp(-\alpha \zeta), \quad \forall \zeta \in [\underline{\zeta}, \bar{\zeta}] \\
y_1, y_2 \geq 0, 
\] (12)

where the first constraint is equivalent to

\[
y_0 \geq \sup_{\zeta \in [\underline{\zeta}, \bar{\zeta}]} \{\exp(-\alpha \zeta) - (y_1 - y_2)\zeta\} \\
\quad = \max \left\{\exp(-\alpha \zeta) - (y_1 - y_2)\underline{\zeta}, \exp(-\alpha \zeta) - (y_1 - y_2)\bar{\zeta}\right\}. 
\] (13)

The equality follows from the convexity of \(\exp(-\alpha \zeta) - (y_1 - y_2)\zeta\), which implies that its maximum value is achieved at an extreme point. Therefore, problem (12) can be reformulated as

\[
\phi(\alpha) = \min_{y_1, y_2 \geq 0} \max \left\{\exp(-\alpha \zeta) + (\bar{\mu} - \zeta)y_1 + (\zeta - \mu)y_2, \exp(-\alpha \bar{\zeta}) + (\bar{\mu} - \bar{\zeta})y_1 + (\bar{\zeta} - \mu)y_2\right\}. 
\] (14)

We can easily verify that the optimal \(y_1, y_2\) must equate the two terms in (14). Hence, we have

\[
y_1 - y_2 = \frac{\exp(-\alpha \zeta) - \exp(-\alpha \bar{\zeta})}{\zeta - \bar{\zeta}} < 0,
\]

and

\[
\phi(\alpha) = \min_{y_1 \geq 0} \left\{\exp(-\alpha \zeta) + (\bar{\mu} - \zeta)y_1 + (\zeta - \mu)y_2 \frac{\exp(-\alpha \zeta) - \exp(-\alpha \bar{\zeta})}{\zeta - \bar{\zeta}}\right\}. 
\]

Similarly, in the optimal solution, \(y_1 = 0\), and \(y_2\) and \(y_0\) can also be derived. The optimal distribution is

\[
\tilde{\zeta} = \begin{cases} 
\zeta, & \text{with probability } \frac{\zeta - \mu}{\zeta - \bar{\zeta}}; \\
\bar{\zeta}, & \text{with probability } \frac{\bar{\mu} - \bar{\zeta}}{\zeta - \bar{\zeta}}.
\end{cases}
\]

We now calculate \(\sup_{p \in P_\theta} \sum_{j=1}^J p_j \phi_j(\alpha)\). For the real distribution \(p \in P_\theta\) which is defined in (11), the calculation of \(p_j^*, j = 1, \ldots, J\) follows Nilim and El Ghaoui (2005). \(\Box\)

19
5 Heuristic for the Satisficing Model

In Section 3, we describe an algorithm for maximizing the satisficing measure in the project selection problem. Theorem 1 shows that, when correlations between project returns exist, this problem is intractable. However, in many situations, project selection decisions are mainly constrained by a budget. For problems with this characteristic, we now describe an efficient and intuitive way to find approximate solutions.

With a single budget constraint, the feasible project selection set is

\[ S = \left\{ s : \frac{c^\prime s}{s} \leq b \right\} \in \{0, 1\}^n. \]

Let \( e_i \) represent an \( n \) dimensional unit vector, where the \( i \)th element is one, and the others are zero. Given the above set \( S \), we describe a simple heuristic for problem (3). We then use binary search to find \( \rho^G_\tau \), where \( \rho^G_\tau = \sup\{\alpha > 0 : C^G_\alpha \geq \tau}\).

**Heuristic Greedy**

**Input:** Risk parameter \( \alpha \).

**Output:** Optimal heuristic solution \( s^G_\alpha \) and its certainty equivalent \( C^G_\alpha \).

1. Start with an empty selection, \( s := 0 \);
   Set \( \bar{c} := 0 \), to represent the total cost of the currently selected projects.

2. Let \( \mathcal{I} = \{i : i \in \mathcal{N}, s_i = 0, \bar{c} + c_i \leq b\} \). If \( \mathcal{I} = \emptyset \), then go to Step 5.

3. Find
   \[ j \in \arg \max \left\{ \frac{C_\alpha(\tilde{\pi}(s + e_j)) - C_\alpha(\tilde{\pi}(s))}{c_i}, \quad i \in \mathcal{I} \right\}. \]

4. If \( C_\alpha(\tilde{\pi}(s + e_j)) > C_\alpha(\tilde{\pi}(s)) \), then add the \( j \)th project into the portfolio, update \( s := s + e_j \) and \( \bar{c} := \bar{c} + c_j \), and go to Step 2.

5. Output \( s^G_\alpha = s \) and \( C^G_\alpha = C_\alpha(\tilde{\pi}(s)) \), and stop.

We use a simple example, which considers both interaction and correlation, to illustrate the steps of Heuristic Greedy. Six projects are available with equal cost, to develop three products A, B and C. The project data appears in Table 1. Because of the limited budget, no more than three projects can be selected. For each product, there are two available projects, but if the manager decides to implement both of them, conflicts arise. Thus, the total return is 20% less than the sum of the two individual returns. Furthermore, the uncertain return of each project depends affinely on two random factors. The first factor, for example technical
issues in the project, influences only the project itself. The second factor, for example the state of the economy, potentially influences all projects. We define the decision variables \( s \in \{0, 1\}^6 \), and project bundle set \( E = \{\{1\}, \{2\}, \{1, 2\}, \{3\}, \{4\}, \{3, 4\}, \{5\}, \{6\}, \{5, 6\}\}. Table 1 shows the returns for each project bundle.

<table>
<thead>
<tr>
<th>Product</th>
<th>Project Bundle</th>
<th>Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>( \hat{r}_{(1)} = \hat{z}_1 + \hat{z}_7 )</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>( \hat{r}_{(2)} = \hat{z}_2 + \hat{z}_7 )</td>
</tr>
<tr>
<td></td>
<td>{1,2}</td>
<td>( \hat{r}<em>{(1,2)} = 0.8(\hat{r}</em>{(1)} + \hat{r}_{(2)}) )</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>( \hat{r}_{(3)} = \hat{z}_3 + \hat{z}_7 )</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>( \hat{r}_{(4)} = \hat{z}_4 + \hat{z}_7 )</td>
</tr>
<tr>
<td></td>
<td>{3,4}</td>
<td>( \hat{r}<em>{(3,4)} = 0.8(\hat{r}</em>{(3)} + \hat{r}_{(4)}) )</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>( \hat{r}_{(5)} = \hat{z}_5 + \hat{z}_7 )</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>( \hat{r}_{(6)} = \hat{z}_6 + \hat{z}_7 )</td>
</tr>
<tr>
<td></td>
<td>{5,6}</td>
<td>( \hat{r}<em>{(5,6)} = 0.8(\hat{r}</em>{(5)} + \hat{r}_{(6)}) )</td>
</tr>
</tbody>
</table>

Table 1: Project Bundle Data in Heuristic Greedy Example.

Each random factor is independent of the others, and the value of \( \hat{z}_i, i = 1, \ldots, 7 \), is either \( \hat{z}_i \) or \( \bar{z}_i \) with equal probability, as shown in Table 2.

<table>
<thead>
<tr>
<th>( i )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{z}_i )</td>
<td>0</td>
<td>-10</td>
<td>-20</td>
<td>-30</td>
<td>-40</td>
<td>-50</td>
<td>0</td>
</tr>
<tr>
<td>( \bar{z}_i )</td>
<td>20</td>
<td>40</td>
<td>60</td>
<td>80</td>
<td>120</td>
<td>140</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 2: Factor Returns in Heuristic Greedy Example.

We let \( \alpha = 0.0073 \) as a trial value. Table 3 shows the calculations of Heuristic Greedy for the example.

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Heuristic Steps</th>
<th>Calculations</th>
<th>( j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Initialization</td>
<td>( s = (0, 0, 0, 0, 0, 0) ), ( \bar{c} = 0 ), ( C_\alpha(\bar{\pi}(s)) = 0 )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Feasible Projects</td>
<td>( \mathcal{I} = {1, 2, 3, 4, 5, 6} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( C_\alpha(\bar{\pi}(s + e_i)) )</td>
<td>14.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( i )</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.54</td>
<td>17.61</td>
</tr>
<tr>
<td>2</td>
<td>Update</td>
<td>( s = (0, 0, 0, 0, 1, 0) ), ( \bar{c} = 1 ), ( C_\alpha(\bar{\pi}(s)) = 22.49 )</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Feasible Projects</td>
<td>( \mathcal{I} = {1, 2, 3, 4, 6} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( C_\alpha(\bar{\pi}(s + e_i)) )</td>
<td>36.85</td>
<td>39.92</td>
</tr>
<tr>
<td>3</td>
<td>Update</td>
<td>( s = (0, 0, 1, 0, 1, 0) ), ( \bar{c} = 2 ), ( C_\alpha(\bar{\pi}(s)) = 41.38 )</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Feasible Projects</td>
<td>( \mathcal{I} = {1, 2, 4, 6} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( C_\alpha(\bar{\pi}(s + e_i)) )</td>
<td>55.55</td>
<td>58.62</td>
</tr>
<tr>
<td>4</td>
<td>Update</td>
<td>( s = (0, 0, 1, 0, 1, 1) ), ( \bar{c} = 3 ), ( C_\alpha(\bar{\pi}(s)) = 59.59 )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Feasible Projects</td>
<td>( \mathcal{I} = \emptyset ), stop.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Output</td>
<td>( s_1^G = s = (0, 0, 1, 0, 1, 1) ), ( C_\alpha(\bar{\pi}(s)) = 59.59 )</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Example Calculations using Heuristic Greedy.
For any given $\alpha > 0$, we use Heuristic Greedy to solve the subproblem, and then apply binary search to find the optimal value of $\alpha$ in (2). Thus, if $\tau \leq C^G_\alpha = C_\alpha(\bar{\pi}(s)) = 59.59$, the trial value of $\alpha$ is increased; otherwise, it is decreased.

**Remark 1** For the special case where project returns are independent and no interactions exist, problem (3) with a single budget constraint can be simplified to

$$\max \sum_{i \in N} \left( r^0_i + \sum_{k \in I_i} C_\alpha(\tilde{z}_k r^k_i) \right) s_i$$

s.t.

$$c's \leq b,$$
$$s \in \{0, 1\}^n.$$  

Moreover, since $\forall i \in N$ with $s_i = 0$,

$$r^0_i + \sum_{k \in I_i} C_\alpha(\tilde{z}_k r^k_i) = C_\alpha(\bar{\pi}(s + e_i)) - C_\alpha(\bar{\pi}(s)),$$

*Heuristic Greedy is the greedy rule commonly used to solve the deterministic knapsack problem (Kellerer et al. 2004).*

## 6 Computational Studies

In this section, we present computational studies of the models and solution procedures described above. Section 6.1 contains a study of the effectiveness of satisficing, relative to other solution approaches. Section 6.2 describes sensitivity analysis for satisficing with various target levels and with various levels of interaction. Section 6.3 discusses the effectiveness of our simple heuristic for satisficing. Using the guidelines of Hall and Posner (2001), (a) we generate a wide range of parameter specifications, (b) the data generated is representative of real world scenarios, and (c) the experimental design varies only the parameters that may affect the analysis.

### 6.1 Comparison with alternative selection rules

We use an example to compare our satisficing model with several alternative selection rules, including maximizing expected return, mean-variance analysis (Markowitz 1959), maximizing the probability of attaining a given target return, and maximizing expected utility.
In this example, projects are correlated and have asymmetrically distributed returns. Overall, 100 projects are available for selection, and based on the cost and available budget, no more than 50 will be selected. Thus, the feasible selection set is $\mathcal{X} = \{x \in \{0, 1\}^{100} : x'1 \leq 50\}$. The 100 available projects are grouped into ten sets according to their return profiles. We let $\tilde{r}_{ij}$ denote the uncertain return of the $j$th project in the $i$th set. The projects within the same subset are correlated such that $\tilde{r}_{ij} = \tilde{r}_{ik} = \tilde{z}_i$, $i, j, k = 1, \ldots, 10$. The distribution of each random factor $\tilde{z}_i$, for $i = 1, \ldots, 10$, is

$$
\tilde{z}_i = \begin{cases} 
10(1 + \sqrt{\frac{1-p_i}{p_i}}), & \text{with probability } p_i = \frac{1}{2}(1 + \frac{i-1}{10}), \\
10(1 - \sqrt{\frac{p_i}{1-p_i}}), & \text{with probability } 1 - p_i.
\end{cases}
$$

(15)

The returns of all the projects have the same mean, $\mu = 10$, and standard deviation, $\sigma = 10$. The difference between the projects lies in their skewness, which we define as the cube root of the third moment. We only consider projects with negative skewness, where a higher set index is associated with a more negatively skewed return. Harvey et al. (2004) present both theoretical and experimental evidence that, ceteris paribus, managers exhibit an aversion towards negatively skewed returns.

We analyze the project selection problem according to different selection rules. The results are summarized in Table 4, where $x^i$, denotes the total number of projects selected from set $i$, for $i = 1, \ldots, 10$. We compare our satisficing model with several classical selection rules.

6.1.1 Expected return

We consider maximizing expected return as the objective. Since all projects have the same expected return, the selection of any 50 projects is optimal. Although this model is easy to solve, it fails to consider risk and the skewness of the distribution of return.

6.1.2 Markowitz model

We use the Markowitz (1959) model to select projects. For any $\eta > 0$, we solve

$$
\max_{x \in \mathcal{X}} \mathbb{E}(x'\tilde{r}) - \eta x'\Sigma x,
$$

23
### Selection Rule

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected return</td>
<td>Selection of any 50 projects</td>
</tr>
<tr>
<td>Markowitz model</td>
<td>$x_M^1, x_M^2, x_M^3, x_M^4, x_M^5, x_M^6, x_M^7, x_M^8, x_M^9, x_M^{10}$</td>
</tr>
<tr>
<td>Probability of target attainment</td>
<td>Intractable; sample solutions are shown in Table 5</td>
</tr>
</tbody>
</table>

### Exponential utility

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>Target</th>
<th>S.M.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>0</td>
<td>$\infty$</td>
</tr>
<tr>
<td>0.01</td>
<td>7</td>
<td>0.02058</td>
</tr>
<tr>
<td>0.02</td>
<td>10</td>
<td>0.0436</td>
</tr>
<tr>
<td>0.04</td>
<td>10</td>
<td>0.0220</td>
</tr>
<tr>
<td>0.06</td>
<td>10</td>
<td>0.0137</td>
</tr>
<tr>
<td>0.08</td>
<td>10</td>
<td>0.0071</td>
</tr>
<tr>
<td>0.1</td>
<td>10</td>
<td>0.0015</td>
</tr>
</tbody>
</table>

### Satisficing model

<table>
<thead>
<tr>
<th>Target</th>
<th>S.M.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\leq$ 50</td>
<td>Unattainable target</td>
</tr>
<tr>
<td>$\geq$ 500</td>
<td></td>
</tr>
</tbody>
</table>

C.E. = certainty equivalent; S.M. = satisficing measure.

Table 4: Solutions from Various Selection Rules.

where $\Sigma$ is the covariance matrix. Within our experimental setting, the Markowitz model can be equivalently reformulated as

$$\max \sum_{i=1}^{10} \left(-\eta \sigma^2 (x_i - \mu_2)^2 + \frac{\mu_2^2}{4\eta^2}\right)$$

s.t. $$\sum_{i=1}^{10} x_i \leq 50$$

where $x_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}, \forall i = 1, \ldots, 10$.

Hence the optimal solution can be obtained as: $x_1^* = \cdots = x_{10}^* = x_M^*$, where $x_M^* = 5$ if $\eta \leq \frac{\mu_2}{10\sigma^2}$; otherwise $x_M^* = \left\lfloor \frac{\mu_2}{2\eta\sigma^2} \right\rfloor$, the closest integer to $\frac{\mu_2}{2\eta\sigma^2}$. This solution shows that we select an equal number of projects from each set. The model of Markowitz (1959) considers risk, but fails to distinguish between the projects based on their skewness.

### 6.1.3 Probability of target attainment

We now consider the objective of maximizing the probability of achieving a specific target return. This model is

$$\max_{x \in \mathcal{X}} \Pr(x' \bar{r} \geq \tau).$$

However, optimization under this selection rule is an intractable problem. The typical approach for solving this problem is by simulation. We can generate a sample consisting of
independent random factor scenarios. Each scenario \( i \) contains a realization of the ten random factors \( z_i^1, \ldots, z_i^{10} \) are independently generated according to (15). The corresponding return outcomes of projects are denoted by \( r^i \). Then the problem of maximizing the probability of attaining the target \( \tau \) can be formulated as

\[
\begin{align*}
\text{max} & \quad \sum_{i=1}^{M} I_i \\
\text{s.t.} & \quad I_i \geq \frac{x^i r_i - \tau}{G}, \quad \forall i = 1, \ldots, M \\
& \quad I_i \leq \frac{x^i r_i - \tau}{G} + 1, \quad \forall i = 1, \ldots, M \\
& \quad I_i \in \{0, 1\}, \quad \forall i = 1, \ldots, M \\
& \quad x \in \mathcal{X},
\end{align*}
\]

where \( G \) is a scalar which is sufficiently large that \( \frac{x^i r_i - \tau}{G} \in (-1, 1) \).

If the sample size \( M \) is large enough, then the solution from (16) can be considered as the solution that maximizes the probability of target attainment. However, since (16) is an intractable problem, it is difficult to solve if \( M \) is large. We use CPLEX to solve problem (16). When \( M = 400 \), for example, the problem can be solved within 20 seconds, whereas if \( M = 600 \) more than two hours are required. We let \( M = 2,000 \), and set a time limit of one hour for the program to run. The results are shown in Table 5.

<table>
<thead>
<tr>
<th>Target</th>
<th>( x^1 )</th>
<th>( x^2 )</th>
<th>( x^3 )</th>
<th>( x^4 )</th>
<th>( x^5 )</th>
<th>( x^6 )</th>
<th>( x^7 )</th>
<th>( x^8 )</th>
<th>( x^9 )</th>
<th>( x^{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>6</td>
<td>0</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>200</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>300</td>
<td>8</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>400</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>450</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 5: Solutions for Maximization of Probability of Target Attainment.

However, the solution from this sampling approach is not reliable, since the sample size cannot be large enough. Repeating the same solution process ten times with different random seeds provides the results shown in Table 6, which are highly inconsistent and suggest that this procedure is unreliable.

Moreover, even if the exact solution can be obtained, maximization of the probability of target attainment cannot incorporate risk, since it fails to consider the extent of gains and losses. To illustrate this point, we compare the solutions in terms of (a) expected loss, which is defined as \( \mathbb{E}((x^i r_i - \tau)^-) \), (b) conditional expected loss, which is defined as \( \mathbb{E}(-(x^i r_i - \tau)|x^i r_i - \tau < 0) \), and (c) the probability of target attainment. Since the projects
have the same expected returns, the expected returns in all the solutions shown in Table 6 are the same.

Given a selection \( x \), all these three terms above can be calculated using simulation. The results are shown in Table 6. We denote the solutions from maximization of target attainment probability by \( x_p \), and the solutions from the satisficing measure in the bottom row by \( x_s \). Table 6 shows that \( x_p \) attains the target with a higher probability than \( x_s \). However, the small increase in the probability of target attainment is achieved at the cost of a large increase in expected loss and conditional expected loss, i.e. much greater risk.

### 6.1.4 Expected exponential utility

We now consider the commonly used selection rule of maximizing expected utility, \( u(w) = 1 - e^{-\alpha w} \). As \( \alpha \) increases, aversion to negative skewness increases. When using the expected utility criterion, a standard measure of selection value is the certainty equivalent. However, selection decisions and their certainty equivalent values are very sensitive to the choice of \( \alpha \), which is difficult to elicit reliably from decision makers. Moreover, other utility functions require fixing similar parameters. The sensitivity of the results to \( \alpha \), as shown in Figure 1, implies that outcomes from using expected utility maximization are not reliable.

### 6.1.5 Satisficing model

Finally, we consider the satisficing model described in Section 2. The results in Table 4 indicate that the satisficing measure shows aversion to negatively skewed returns, and that
this aversion reduces as the target increases. To understand this, we recall that the satisficing measure favors risk averse choices; hence, it is generally averse to negative skewness. Overall, both diversification preference and skewness aversion decrease as the target increases. When the target level is low, the decision maker can afford to be more risk averse; hence, the skewness aversion is high and project selections favor the lower index projects. Alternatively, when the target level is high, skewness aversion is lower; hence, project selections favor the higher index projects. It is important to note that, by contrast with the model discussed in Section 6.1.4, the satisficing model does not require specification of a value for $\alpha$.

6.2 Sensitivity analysis

In this section, we first study sensitivity analysis of the satisficing model to target level. Then we study sensitivity analysis to the extent of interactions between the returns of the projects.

6.2.1 Target level

We conduct a study of the sensitivity of our satisficing model to various target levels. We consider a project selection problem where each project has only two possible returns, depending on its success or failure. The success probability is 0.5, independent of other projects. There are 18 available projects, as shown in Table 7. As the project number increases, the expected return, the variance of return, and the investment cost increase. The only constraint is an
The optimal selections from the satisficing model under various target levels for the certainty equivalent are shown in Table 8. We now discuss these results.

1. The results are only relevant if the target falls within a reasonable range. For example, if the target is set below 3,200, then $\rho^*_\tau$ approaches infinity, since there exists a selection for which even if all the selected projects fail, the total return still achieves the target. Hence, there is no risk. At the other extreme, if the target is set above 14,100, then no solution exists. This range is natural in real project selection problems, since most managers do not use either a target that is very easy to achieve or one that is unattainable without the assumption of excessive risk.

2. As the target increases, the optimal selection becomes less diversified, i.e. less risk averse. Thus, when $\tau = 6,000$, 14 projects are selected; when $\tau = 11,000$, 13 projects are selected; when $\tau = 11,750$, 12 projects are selected; and when $\tau = 12,500$, only 11 projects are selected. Moreover, the satisficing measure decreases as the target increases, indicating the assumption of greater risk.

3. Also, as the target increases, the satisficing selection moves towards projects with higher mean return and risk. When $\tau \geq 12,500$, most of the projects selected have high mean return and risk. Hence, the solution does not strongly consider risk.

To compare the performance of solutions obtained under different target levels, we generate a sample consisting of $M = 1,000,000$ scenarios. Each scenario contains a realization

### Table 7: Project Information for Sensitivity Analysis.

<table>
<thead>
<tr>
<th>Project</th>
<th>Return</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Success</td>
<td>Failure</td>
</tr>
<tr>
<td>1</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>2</td>
<td>600</td>
<td>450</td>
</tr>
<tr>
<td>3</td>
<td>700</td>
<td>400</td>
</tr>
<tr>
<td>4</td>
<td>800</td>
<td>350</td>
</tr>
<tr>
<td>5</td>
<td>900</td>
<td>300</td>
</tr>
<tr>
<td>6</td>
<td>1000</td>
<td>250</td>
</tr>
<tr>
<td>7</td>
<td>1100</td>
<td>200</td>
</tr>
<tr>
<td>8</td>
<td>1200</td>
<td>150</td>
</tr>
<tr>
<td>9</td>
<td>1300</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2700</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>2800</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>2900</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>3000</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>3100</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>3200</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>3300</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>3400</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>3500</td>
</tr>
</tbody>
</table>
of the 18 projects’ returns, which are independently generated as described in Table 7. We study the return profiles of the solutions that correspond to different targets. We let $x_{6,000}$ and $x_{12,500}$ denote the solutions under the targets $\tau = 6,000$ and $\tau = 12,500$, respectively. The return profiles for these two target levels are shown in Figure 2. The higher target solution, $x_{12,500}$, leads to a higher mean return. However, with respect to the probability of the total return being above 6,000, $x_{6,000}$ is more successful than $x_{12,500}$. Moreover, there is a nonzero probability that $x_{12,500}$ results in a total return less than 3,000, which is not the case for $x_{6,000}$. If the target is 12,500, then we observe that $x_{12,500}$ is better than $x_{6,000}$, since the probability to achieve 12,500 is higher and it is possible to reach a high return level that can never be achieved by $x_{6,000}$.

<table>
<thead>
<tr>
<th>Target $\alpha^* ($×$10^4$)</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\leq 3200$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>6000</td>
<td>1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14</td>
</tr>
<tr>
<td>11000</td>
<td>1.57</td>
</tr>
<tr>
<td>11750</td>
<td>3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15</td>
</tr>
<tr>
<td>12500</td>
<td>1.72</td>
</tr>
<tr>
<td>$\geq 14100$</td>
<td>Target is unattainable, hence no solution exists</td>
</tr>
</tbody>
</table>

Table 8: Satisficing Selections under Various Targets.

![Figure 2: Return Profiles under Various Targets.](image-url)
6.2.2 Interactions

We now use an example to study how project selection varies with the amount of interaction between the projects. The return profiles of the 18 available projects remain the same as in Table 7. We assume that a synergistic value is earned if specific projects are selected simultaneously. The projects are grouped into six sets, such that the $i$th set, for $i = 1, \ldots, 6$, contains project $3i−2, 3i−1$ and $3i$. In each set, if exactly two (respectively, three) projects are selected, then an additional return in proportion $\psi_2$ (resp., $\psi_3$) is earned. For example, in set 1, if projects 1 and 2 are selected and project 3 is rejected, then the total return is $(1 + \psi_2)(\tilde{r}_1 + \tilde{r}_2)$. Also, if projects 1, 2 and 3 are all selected, then the total return is $(1 + \psi_3)(\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3)$. Projects in different sets have no interactions. We fix the target level at $\tau = 17,500$, and vary the amount of interaction by varying $\psi_2$ and $\psi_3$. The results are shown in Table 9.

<table>
<thead>
<tr>
<th>$\psi_2$</th>
<th>$\psi_3$</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.10</td>
<td>1.40</td>
<td>10 11 12 13 14 15 16 17 18</td>
</tr>
<tr>
<td>1.20</td>
<td>1.60</td>
<td>3 4 5 6 7 8 9 10 11 12 13 14 15</td>
</tr>
<tr>
<td>1.30</td>
<td>1.80</td>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14</td>
</tr>
</tbody>
</table>

Table 9: Satisficing Selections under Various Interaction Levels.

The results in Table 9 show that changing the amount of interaction leads to different selections. As greater synergies become available, the project portfolio becomes more risk averse. That is, the selection diversifies towards projects with lower mean and variance. This is because greater synergies results in the target becoming relatively more attainable. Hence, it is possible to attain this target while still being relatively risk averse.

6.3 Heuristic

We now evaluate the performance of Heuristic Greedy, which is described in Section 5. We develop several test problems that are similar to those described in Section 6.1, but with various problem sizes.

Problems with size $n \in \{5, 7, 10, 15\}$ are constructed as follows. In each case, there are $n^2$ projects available for selection. The $n^2$ projects are grouped into $n$ sets, such that $\tilde{r}_{ij} = \tilde{r}_{ik} = \tilde{z}_i$, where $\tilde{r}_{ij}$ represents the random return of the $j$th project in the $i$th set, for
$i, j, k = 1, \ldots, n$. The distribution of each factor $\tilde{z}_i$ is

$$
\tilde{z}_i = \begin{cases} 
10(1 + \sqrt{\frac{1-p_i}{p_i}}), & \text{with probability } p_i = \frac{1}{2}(1 + \frac{i-1}{n}), \\
10(1 - \sqrt{\frac{p_i}{1-p_i}}), & \text{with probability } 1 - p_i.
\end{cases}
$$

For each value of $n$, we generate $M = 500$ instances of the problem with size $n$, and observe the average performance of Heuristic Greedy over these instances. The instances are generated by modifying both the target for the satisficing measure and the cost of the project. In each instance, the target $\tau$ is generated using $\tau \sim U[10, \tau']$, where $\tau' = 6n^2$. The cost of each project is generated independently from the uniform distribution $U[0.5, 3.5]$. The budget constraint for every instance is the same, such that the total cost of the selected projects is no more than $n^2$. For each instance, the optimal satisficing measure $\rho^*_\tau$ is calculated using both Algorithm CuttingPlane from Section 3 and Heuristic Greedy from Section 5. The two resulting solution values are represented by $\rho^*_\tau$ and $\rho^G_\tau$ respectively. To evaluate the relative performance of Heuristic Greedy, we record the value of $\Delta = \frac{\rho^*_\tau - \rho^G_\tau}{\rho^*_\tau}$.

The comparison of these two solutions is shown in Figure 3, where the performance of Heuristic Greedy is very similar to that of Algorithm CuttingPlane. When $n = 15$ for example, in more than 80% of the instances tested the two approaches achieve the same result, and $\Delta \leq 0.25\%$ in all 500 instances. The performance of Heuristic Greedy with $n = 15$ is shown in Table 10.

Furthermore, by comparing the heuristic performances of different problem size in Figure 3, we observe that the performance of Heuristic Greedy improves with problem size. As $n$ increases from $n = 5$ to $n = 15$, low $\Delta$ values occur more frequently in our results.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>[0, 0.01%)</th>
<th>[0.01%, 0.1%)</th>
<th>[0.1%, 0.2%)</th>
<th>[0.2%, 0.25%)</th>
<th>[0.25%, +\infty)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of instances</td>
<td>437</td>
<td>51</td>
<td>10</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 10: Heuristic Greedy Performance for Instances with $n = 15$.

7 Concluding Remarks

This paper considers the problem of selecting a portfolio of projects when the return of each project is uncertain. The problem studied is general enough to allow interactions between the different projects, and correlations between their uncertain returns. We describe a satisficing criterion for this problem. This criterion is implemented by maximizing the entropic
satisficing measure. Thus, our model maximizes the value of a risk aversion parameter, while keeping the certainty equivalent of the uncertain returns under exponential utility above a given target. The model is solved using binary search on the risk aversion value of the portfolio, with solution of the subproblems by a cutting plane procedure. We demonstrate computationally that the satisficing model identifies better project portfolios than several classical decision models, including maximization of expected return, mean-variance analysis, maximization of target attainment probability, and expected utility maximization. For project selection decisions that are constrained only by a budget, we describe a simple but highly accurate heuristic satisficing procedure.

Our results provide several insights that managers will find useful. First, it is now possible to design a satisficing project portfolio that is as risk averse as possible, subject to meeting a target certainty equivalent level. Second, this design can be achieved very accurately using a computationally efficient procedure. Third, the resulting project portfolio typically offers significant benefits over those obtained by all previously used approaches. Fourth, it is possible to balance upside potential and downside risk accurately, by adjusting the satisficing target level. Fifth, as the target level increases, the selected project portfolios become less diversified and more focused on projects that have higher mean return and variance. Sixth, higher levels of synergistic interaction between project returns result in the selection of projects that have lower mean return and variance. Finally, in project selection
situations that are constrained only by a budget, a simple spreadsheet-based procedure routinely provides almost exact satisficing portfolios.

Several opportunities exist for future research. First, in many practical projects, the initial investment cost is not predictable, and uncertainty about it can be incorporated into a satisficing model. Second, a related extension is allowing the available budget to be random. In practice, available budgets for funding projects are often uncertain. Third, the satisficing model should be applied to dynamic project selection problems. In such problems, projects with random investment cost and return become available over time. Consequently, some part of the available budget may need to be held in reserve for future opportunities. Fourth, the problem considered here can be generalized to allow for decisions about the timing of projects, in order to match resource requirements and resource availability over time. A satisficing approach can usefully be applied to this problem. Finally, it would be valuable to perform large scale behavioral experiments on project selection decisions, to determine the factors that influence how well satisficing explains those decisions in practice. We hope that our work will encourage future research in these interesting and important directions.

Acknowledgments

The work of the first author is supported by the Summer Fellowship Program and the Center for Operational Excellence of the Fisher College of Business, The Ohio State University. The work of the fourth author is supported by the Singapore - MIT Alliance, and by National University of Singapore academic research grant R-314-000-068-122.
References


