Decision Making under Uncertainty when Preference Information is Incomplete

Benjamin Armbruster* Erick Delage†

June 29, 2012

Abstract

We consider the problem of optimal decision making under uncertainty but assume that the decision maker's utility function is not completely known. Instead, we consider all the utilities that meet some criteria, such as preferring certain lotteries over certain other lotteries and being risk averse, s-shaped, or prudent. This extends the notion of stochastic dominance. We then give tractable formulations for such decision making problems. We formulate them as robust utility maximization problems, as optimization problems with stochastic dominance constraints, and as robust certainty equivalent maximization problems. We use a portfolio allocation problem to illustrate our results.

1 Introduction

This paper questions a key and rarely challenged assumption of decision making under uncertainty, the assumption that the decision maker can always, after a tolerable amount of introspective questioning, clearly identify the utility function that characterizes their attitude toward risk. The use of expected utility to characterize attitudes towards risk is pervasive. In large part this is due to von Neumann and Morgenstern who proved (von Neumann and Morgenstern, 1944) that any set of preferences that a decision maker may have among risky outcomes can be characterized by an expected utility measure if the preferences respect certain reasonable axioms (i.e., completeness,

*Northwestern University, armbruster@northwestern.edu
†HEC Montréal, erick.delage@hec.ca

transitivity, continuity, and independence). Specifically, there exists a utility function $u : \mathbb{R} \to \mathbb{R}$ so that among two random variables (or lotteries), W and Y, the decision maker prefers W to Y if and only if $\mathbb{E}[u(W)] \geq \mathbb{E}[u(Y)]$.

There has been much effort on determining how to choose a utility function for a decision maker, and this work plays an integral part in the design of surveys for assessing tolerance to financial risk (Grable and Lytton, 1999). The method for choosing a utility function proposed in most textbooks on decision analysis (see for instance Clemen and Reilly (1999)) is to make a set of pairwise comparisons between lotteries (often using the Becker-DeGroot-Marschak reference lotteries Becker et al. (1964)), in order to identify the value of the utility function at a discrete set of points. The utility function is then completed by naive interpolation. A more sophisticated approach assumes that the utility function has a parametric structure such as constant absolute or constant relative risk aversion. For example, if a decision maker can confirm that they are risk averse and that their preference between *any* two lotteries is invariant to the addition of *any* constant amount to all outcomes, then they have constant absolute risk aversion, and thus, their utility is of the form $u(y) = 1 - e^{-\gamma y}$. Parameters are then resolved using a small number of pairwise comparisons between lotteries.

These approaches have important shortcomings. If they do not assume a parametric form then the large or even continuous space of outcomes may require a lot of interpolation or asking the decision maker many questions. Even interpolation may not be easy because if the questions to the decision maker are binary choices between two lotteries, then their answers will not provide the value of the utility function at any point but each answer merely provides a single linear constraint on the values of the utility function on the support of these two lotteries. To justify a parametric form for the utility function, a decision maker must be able to confidently address a question about an infinite number of lottery comparisons (such as described above for utilities with constant absolute risk aversion). A more fundamental limitation is that all these procedures conclude by selecting a *single* "most likely" utility function given the evidence. They entirely disregard other plausible choices and the inherent ambiguity of that choice. In this paper, we will focus on these instances where knowledge can only be gathered using a small number of simple questions and meaningful decisions must be made even though no single utility can be unambiguously identified.

Our approach follows in spirit the line of work in the artificial intelligence literature on utility elicitation using optimization for problems that only involve a finite, although possibly large, set of outcomes. This line of work emphasizes that utility elicitation and decision analysis should be combined into a single process in order to use all the information collected about the true utility function when making a decision. In this context, Chajewska et al. (2000) represents the knowledge of the decision maker's preferences using a probability distribution over utility functions, and then judges a decision by its expected utility averaged over the distribution of utilities. For increasing their knowledge of the utility function, they use a value of information criterion to select the next question to the decision maker. In contrast to this probabilistic approach, in Boutilier et al. (2006) the authors construct the set $\mathcal{U}$ of all utility functions that do not contradict the available information. They then identify the decision that achieves minimum worst-case regret (i.e. regret experienced a posteriori once the true utility function is revealed) using a mixed integer programming approach and exploiting the assumed "generalized additive" structure of the true utility. In comparison, our paper considers uncertain real-valued outcomes and proposes formulations that are more natural for decision making and reduce to convex optimization problems.

We motivate our discussion with the following stochastic program

$$\max_{x \in \mathcal{X}} \mathbb{E}[u(h(x, \xi))] ,$$

where x is a vector of decision variables; $\mathcal{X}$ is a set of implementable decisions; $h(x, \xi)$ is a function mapping the decision x to a random return indexed by the scenario ξ; and where the expectation is over the random scenarios ξ. We assume that we have not gathered enough information to uniquely specify $u(\cdot)$. Thus we build on the theory developed by Aumann (1962) of expected utility without the completeness axiom. This theory suggests that our incomplete preferences can be characterized by a *set* of utility functions $\mathcal{U}$ (Dubra et al., 2004). This set describes our incomplete information about $u(\cdot)$ and is known to contain the true utility function. Another situation where preferences are incomplete is when groups make decisions by consensus: here $\mathcal{U}$ contains the utility functions of the group members, and two lotteries are incomparable if the group members do not agree on which is preferred.

The set $\mathcal{U}$ suggests that we face a robust optimization problem. Our approach will differ however from the typical robust optimization framework, which is robust to the possible realizations or distributions of ξ (see for example Ben-Tal and Nemirovski (1998) and Delage and Ye (2010) and references therein). Instead we are robust to the possible utilities in $\mathcal{U}$ and choose the worst-case utility function.

When the range of $h(x, \xi)$ is not restricted to a discrete set, the only existing way of dealing with ambiguity in the utility function is a stochastic program with a stochastic dominance constraint (Dentcheva and Ruszczyński, 2003),

$$\max_{x \in \mathcal{X}} \mathbb{E}[f(x, \xi)]$$
$$s.t.\ h(x, \xi) \succeq Z$$

with some objective function f. In these problems the stochastic dominance constraint, $h(x, \xi) \succeq Z$, is defined as $\mathbb{E}[u(h(x, \xi))] \geq \mathbb{E}[u(Z)]$ for all utility functions $u \in \mathcal{U}$. This constraint ensures that the random consequences of the chosen action, $h(x, \xi)$, are preferred to those of a baseline random variable Z for all utility functions in $\mathcal{U}$. For first order dominance, $\mathcal{U}$ is the set of all increasing functions, and for second order dominance, it is the set of all increasing concave functions. Some limitation of stochastic dominance constraints are 1) that they don't provide guidance with respect to choosing an objective function f; 2) that the choice of baseline Z is not clear; and 3) that the set $\mathcal{U}$ may be very large for first or second order dominance and thus the constraint may be very restrictive.

The contribution of this paper is two-fold. First, we introduce two tractable measures of performance that can naturally replace the expected utility measure when risk preferences are only partially known. In particular, we show that maximizing the worst-case certainty equivalent of the return of our decision,

$$\max_{x \in \mathcal{X}} \inf_{u \in \mathcal{U}} u^{-1}(\mathbb{E}[u(h(x, \xi))]) \,,$$

is tractable. Second, we allow more flexibility in the formulation of $\mathcal{U}$ compared to first or second order stochastic dominance constraints. In addition to imposing the rudimentary monotonicity and risk-aversion criteria on $\mathcal{U}$, we also consider s-shaped utilities and "prudent" utilities. More importantly, we incorporate information on the decision maker's risk attitudes, which can easily be obtained using risk tolerance assessment surveys. Such surveys typically involve questions such as:

> "You are on a TV game show and can choose one of the following. Which would you take? A) $1,000 in cash; B) A 50% chance at winning $5,000; C) A 25% chance at winning $10,000; D) A 5% chance at winning $100,000" (Grable and Lytton, 1999)

In particular, we will restrict the utility functions in $\mathcal{U}$ to those that make the same choice as is reported by the decision maker when presented such hypothetical gambles. This allows us to explore, numerically, how much value is added to a stochastic optimization problem when more information about the risk preferences of our decision maker can be taken into account, starting from simple knowledge of risk aversion to exact knowledge of the utility function that characterize his preferences.

In the next section we describe three formulations (including the stochastic dominance formulation) that can be used instead of maximizing expected utility when the decision maker's utility function is only known to lie inside a set $\mathcal{U}$. In section 3 we describe the sets of utilities $\mathcal{U}$ and how to optimize each formulation with these sets. We then present numerical examples involving a portfolio allocation problem in section 4 and conclude in section 5.

2 Formulations

Our work examines three formulations for decision making when you know the utility function is in some set $\mathcal{U}$. These formulations involve 1) optimizing with a stochastic dominance constraint,

$$\begin{aligned} &\max_{x\in\mathcal{X}} \ f(x) \\ &\ s.t.\ \ \mathbb{E}[u(h(x,\xi))] \geq \mathbb{E}[u(Z)] \quad \forall u \in \mathcal{U} \end{aligned} \tag{1}$$

where Z is some reference random variable; 2) maximizing the worst-case utility,

$$\max_{x\in\mathcal{X}} \ \inf_{u\in\mathcal{U}} \ \mathbb{E}[u(h(x,\xi))]; \tag{2}$$

and 3) maximizing the worst-case (or robust) certainty equivalent,

$$\max_{x\in\mathcal{X}} \ \inf_{u\in\mathcal{U}} \ \mathbb{C}_u[h(x,\xi)], \tag{3}$$

where the certainty equivalent of a lottery (i.e., random variable) X given a utility function u is typically defined as the amount for sure such that you would be indifferent between it and the lottery, that is $u^{-1}(\mathbb{E}[u(X)])$. To ensure uniqueness, we slightly modify this definition for : $\mathbb{C}_u[X] := \sup\{s : u(s) \le \mathbb{E}[u(X)]\}$. The robust certainty equivalent formulation (3) maximizes $\inf_{u \in \mathcal{U}} \mathbb{C}_u[h(x,\xi)]$, the largest amount of money we know for sure we would be willing to exchange for the lottery $h(x,\xi)$. In the context of group decision making, using the worst-case utility function means accommodating the group's least-favored member.

Since we do not know the true utility function in these formulations, any choice from $\mathcal{U}$ is as justifiable as any other. We use the worst-case utility function because of convenience: that choice turns out to make these formulations very tractable. In addition to convenience, we can also motivate the choice of utility from $\mathcal{U}$ with an analogy to Rawls' *A Theory of Justice* (1971). Rawls proposes that one imagine deciding the structure of society behind a "veil of ignorance," i.e., without knowing one's place in society. While our decision maker's choices are less weighty, their ignorance of their true utility function is somewhat analogous. Rawls then argues that this leads one to focus on the least-advantaged in society and suggests a max-min principle for allocating goods. Similarly, we focus on the least-favorable utility function using max-min formulations.

Since we seek convex formulations we will assume that the feasible set $\mathcal{X}$ is convex; the objective function f in (1) is concave; the function $h(x,\xi)$ relating the action to a random outcome is concave in x; and the utilities in $\mathcal{U}$ are concave, to ensure that the objective in (2) is concave in x (the only exception is when we discuss s-shaped utilities). For computational tractability we also assume that all the random variables have finite support. We assume that there are M scenarios for ξ, $\Omega := \{\xi_1, \ldots, \xi_M\}$ with associated probabilities $p_i := \mathbb{P}[\xi = \xi_i]$.

The key to our success is determining tractable representations of

$$\psi(x;\mathcal{U},Z) \;:=\; \inf_{u \in \mathcal{U}} \quad \mathbb{E}[u(h(x,\xi))] - \mathbb{E}[u(Z)] \tag{4}$$

where we sometimes drop the dependence on $\mathcal{U}$ and Z from out notation. Using $\psi(x;\mathcal{U},Z)$ we can write the stochastic dominance formulation (1) as

$$\max_{x \in \mathcal{X}} f(x)$$
$$s.t.\ \psi(x; \mathcal{U}, Z) \geq 0$$

and the worst-case utility formulation (2) as

$$\max_{x \in \mathcal{X}} \psi(x; \mathcal{U}, 0)$$

where we chose $Z := 0$ a.s. Unlike the other formulations, the certainty equivalent formulation is not concave but quasiconcave (see proof in appendix 6.1). Thus we can solve it using a bisection algorithm.

Remark 2.1. *We omit to study the worst-case regret formulation*

$$\min_{x \in \mathcal{X}} \sup_{u \in \mathcal{U}} \max_{x' \in \mathcal{X}} E[u(h(x', \xi))] - E[u(h(x, \xi))]$$

proposed in Boutilier et al. (2006), for two reasons. First, from a decision-theoretic point of view, minimax regret as a choice function violates the independence to irrelevant alternatives condition which is essential for rationalizing preferences (see Arrow (1959)). That condition states that our preference between decision x_1 and x_2 should not be influenced by the set of alternatives $\mathcal{X}$. Second, it is likely to be an intractable problem when $\mathcal{U}$ is a general convex set. Intuitively, the reason is that evaluating the worst-case regret associated with a fixed x reduces to solving

$$\sup_{u \in \mathcal{U}, x' \in \mathcal{X}} \int E[u(y)\delta_y(h(x', \xi)) - u(y)\delta_y(h(x, \xi))]dy$$

where $\delta_y(\cdot)$ is the Dirac measure. Unfortunately, the cross-term $u(y)\delta_y(h(x', \xi))$ prevents this from being a convex optimization problem.

3 Worst-case Utilities

The following are three common hypotheses about a decision maker's utility function.

Risk aversion A decision maker is risk averse if for any lottery X, they prefer $\mathbb{E}[X]$ for sure over the lottery X itself. This is characterized by the concavity of the utility function.

Prudence In the finance and economics literature, a prudent attitude describes the need for larger precautionary savings when facing a riskier situation. Specifically, a decision maker is said to be prudent if they are always willing to accept a lower amount of cash in exchange for a lottery when this lottery becomes "riskier", i.e. as outcomes become more dispersed while the expected return remains fixed. Prudence is a stronger condition than risk aversion. It is characterized by the convexity of the derivative of the utility function (see Leland (1968), Sandmo (1970), and Kimball (1990)) and is implied by functions with decreasing absolute risk aversion.

S-shape Prospect theory was proposed by Kahneman and Tversky (1979) to bridge the gap between normative theories of rational behavior and behavior observed by experimentalists. This theory conjectures that preferences are affected by four factors. First, outcomes are evaluated with respect to a reference point. Second, decision makers are more affected by losses than by winnings. Third, the perception of winnings or losses is diminished as they get larger. Finally, the perception of probabilities is biased (i.e. over-weighting smaller probabilities and under-weighting larger ones). These observations suggest that the decision maker is risk averse with respect to *gains* and risk seeking with respect to *losses*. Specifically, it suggests an s-shaped utility function that is concave for gains and convex for losses. As is typically done in the context of a normative study, in what follows, we will disregard the possibility of any probability assessment bias and focus on how to account for information that indicates that the utility function has this particular shape.

In what follows, we present tractable reformulation for evaluating $\psi(\cdot;\mathcal{U})$ for three different types of sets $\mathcal{U}$ that are formed from intersections of the following sets of utility functions:

$$U_2 := \{u : u' \geq 0,\ u'' \leq 0\}\,,$$

$$U_s := \{u : u' \geq 0,\ (u''(y) \leq 0\,\forall\, y \geq 0),\ (u''(y) \geq 0\,\forall\, y \leq 0)\}\,,$$

$$U_3 := \{u : u''' \geq 0\}\,,$$

$$U_n := \{u : \mathbb{E}[u(W_0)] - \mathbb{E}[u(Y_0)] = 1\}\,,$$

$$U_a := \{u : \mathbb{E}[u(W_k)] \geq \mathbb{E}[u(Y_k)],\ \forall k = 1, \ldots, K\}\,.$$

These definitions should be interpreted in the natural way when the derivatives do not exist (e.g., interpret U_2 as the set of nondecreasing, concave utilities). More formally, we consider the closure of these sets under some suitable metric on functions such as the 1-norm. Here U_2 is the set of risk-averse utilities; U_s is the set of s-shaped convex-concave utilities and is the only exception to the assumption throughout the paper that utilities are concave; U_3 is the set of prudent utilities, those with convex u'; and U_a is the set of utilities that prefer W_k to Y_k for all k. Since adding a constant to a utility or multiplying it by a positive constant results in an equivalent utility, it is often necessary to normalize utilities. There are multiple ways of normalizing utilities. Here we use U_n to specify the scaling. For example, assuming that $W_0 := 1$ and $Y_0 := 0$ a.s. enforces that $u(1) - u(0) = 1$. The choices of $\mathcal{U}$ we focus on are $\mathcal{U}^2 := U_a \cap U_n \cap U_2$, $\mathcal{U}^s := U_a \cap U_n \cap U_s$, and $\mathcal{U}^3 := U_a \cap U_n \cap U_2 \cap U_3$. These choices all incorporate U_a allowing one to tailor the problem to the specific preferences of a particular decision maker, whether he be entirely risk averse, risk-seeking over losses, or prudent. As an example, $\mathcal{U}^2$ with no specific preferences, i.e. $K = 0$, reduces to the set defining second-order dominance. We now present finite dimensional concave formulations of ψ for these choices of $\mathcal{U}$.

The notation used in the following results will refer to $\mathcal{S}$ as the joint support of all static random variables, $\mathcal{S} := \text{supp}(Z) \cup \left(\bigcup_{k=0}^{K} \text{supp}(Y_k) \cup \text{supp}(W_k)\right)$, and use y_j to denote the j-th smallest entry of $\mathcal{S}$. For clarity of exposure, scenarios in Ω will always be indexed by i, outcomes in $\mathcal{S}$ by j, and queries by k. Thus the size of our optimization problems is specified by the number of queries K, the number of scenarios M, and the size of the support $N := |\mathcal{S}|$.

3.1 Incorporating Lottery Comparisons

We first address how to account for results of K lottery comparisons for a decision maker known to be risk averse. Specifically, in this case evaluating $\psi(x;\mathcal{U}, Z)$ requires characterizing the optimal value of the infinite dimensional problem

$$\inf_{u \in \mathcal{U}^2} \ \mathbb{E}[u(h(x,\xi))] - \mathbb{E}[u(Z)] \ .$$

Our main result states that this value can be computed by solving a finite dimensional linear program involving $2(N + M)$ variables and $MN + K + M + 2N - 1$ constraints (not counting the nonnegativity constraints).

Theorem 1. *The optimal value of the linear program*

$$\begin{aligned}
\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{v},\boldsymbol{w}} \quad & \sum_i p_i(v_i h(x,\xi_i) + w_i) - \sum_j \mathbb{P}[Z = y_j]\alpha_j && \text{(5a)}\\
s.t. \quad & y_j v_i + w_i \geq \alpha_j \; \forall\, i \in \{1,\dots,M\}, j \in \{1,\dots,N\} && \text{(5b)}\\
& \sum_j \mathbb{P}[W_0 = y_j]\alpha_j - \mathbb{P}[Y_0 = y_j]\alpha_j = 1 && \text{(5c)}\\
& \sum_j \mathbb{P}[W_k = y_j]\alpha_j \geq \sum_j \mathbb{P}[Y_k = y_j]\alpha_j \qquad \forall k = 1,\dots,K && \text{(5d)}\\
& \alpha_{j+1} - \alpha_j \geq \beta_{j+1}(y_{j+1} - y_j) \; \forall\, j \in \{1,\dots,N-1\} && \text{(5e)}\\
& \alpha_{j+1} - \alpha_j \leq \beta_j(y_{j+1} - y_j) \; \forall\, j \in \{1,\dots,N-1\} && \text{(5f)}\\
& \boldsymbol{v} \geq 0\,, \quad \boldsymbol{\beta} \geq 0 && \text{(5g)}
\end{aligned}$$

equals $\psi(x;\mathcal{U}^2)$*. Furthermore, a worst-case utility function (i.e., one achieving the infimum in* (4)*) is*

$$u^*(y) = \begin{cases} \alpha_N & y \geq y_N \\ \frac{\alpha_{j+1}-\alpha_j}{y_{j+1}-y_j} y + \frac{y_{j+1}\alpha_j - y_j\alpha_{j+1}}{y_{j+1}-y_j} & y_j \leq y < y_{j+1} \; \forall\, j \in \{1,\dots,N-1\} \\ -\infty & y < y_1. \end{cases} \tag{6}$$

This is a piecewise linear function connecting the points $u(y_j) = \alpha_j$*, which equals* $-\infty$ *for* $y < y_1$ *and* α_N *for* $y \geq y_N$*.*

We present a detailed proof of this result as the ideas that are used will be reused in the proofs of Theorems 2 and 3.

Proof. We first partition the set of utility functions by their values at the points in $\mathcal{S}$, letting $U(\boldsymbol{\alpha}) := \{u : u(y_j) = \alpha_j \;\forall j\}$. Hence,

$$\psi(x;\mathcal{U}^2) = \min_{\boldsymbol{\alpha}} \; \psi(x; U(\boldsymbol{\alpha}) \cap \mathcal{U}^2)$$
$$U(\boldsymbol{\alpha}) \cap \mathcal{U}^2 \neq \emptyset.$$

Note that $U(\boldsymbol{\alpha})$ is either a subset of U_a or is disjoint from it. The same is true with respect to U_n.

Since $\mathcal{U}^2 := U_a \cap U_n \cap U_2$ it then follows that

$$\psi(x;\mathcal{U}^2) = \min_{\boldsymbol{\alpha}} \ \psi(x;U(\boldsymbol{\alpha}) \cap U_2)$$
$$U(\boldsymbol{\alpha}) \cap U_2 \neq \emptyset,\ U(\boldsymbol{\alpha}) \subseteq U_a,\ U(\boldsymbol{\alpha}) \subseteq U_n.$$

The constraint $U(\boldsymbol{\alpha}) \cap U_2 \neq \emptyset$ is represented by (5e), (5f) and $\boldsymbol{\beta} \geq 0$; $U(\boldsymbol{\alpha}) \subseteq U_n$ is by (5d); and $U(\boldsymbol{\alpha}) \subseteq U_n$ is by (5c). Note that $\mathbb{E}[u(Z)]$ is a constant, $\sum_j \mathbb{P}[Z = y_j]\alpha_j$ for $u \in U(\alpha)$. Thus, evaluating $\psi(x;U(\boldsymbol{\alpha}) \cap U_2)$ is equivalent to minimizing $\mathbb{E}[u(h(x,\xi))]$ over $u \in U(\boldsymbol{\alpha}) \cap U_2$. Among the non-decreasing concave functions in $U(\alpha)$, this is minimized by the piecewise linear function u^* in (6), which essentially forms a convex hull of the points (y_j, α_j) with the additional requirement that the function be nondecreasing. Hence when $U(\boldsymbol{\alpha}) \cap U_2 \neq \emptyset$, then the function u^* in (6) is a worst-case utility function for $\psi(x;U(\boldsymbol{\alpha}) \cap U_2)$ (i.e., achieves the infimum in (4)). Then, $\psi(x;U(\boldsymbol{\alpha}) \cap U_2) = \mathbb{E}[u^*(h(x,\xi))] - \sum_j \mathbb{P}[Z = y_j]\alpha_j$. Since u^* is concave,

$$u^*(y) = \min_{v \geq 0, w} \quad vy + w \tag{7a}$$
$$s.t. \quad vy_j + w \geq \alpha_j \ \ \forall\, j \in \{1, \ldots, N\}. \tag{7b}$$

Substituting $y = h(x, \xi)$ for every i, gives us the objective (5a) and the constraints (5b) and $\boldsymbol{v} \geq 0$. □

Remark 3.1. *An alternative way to ensure concavity of the utility functions is to replace constraints (5e), (5f), and $\boldsymbol{\beta} \geq 0$ by $\alpha_{j+1} = \alpha_j + \beta_j(y_{j+1} - y_j)$ and $\beta_{j+1} \leq \beta_j$ for all $j \in \{1, \ldots, N-1\}$, where we consider $\beta_N = 0$.*

This formulation allows us to efficiently solve the problems (1), (2), and (3). To solve (1) and (3) we look at the dual of problem (5). This allows us to write $\psi(x;\mathcal{U}^2) \geq 0$ using the dual variables

$\boldsymbol{\mu} \in \mathbb{R}^{N\times M}$, $\nu_0 \in \mathbb{R}$, $\boldsymbol{\nu} \in \mathbb{R}^K$, $\boldsymbol{\lambda}^{(1)} \in \mathbb{R}^{N-1}$, and $\boldsymbol{\lambda}^{(2)} \in \mathbb{R}^{N-1}$ and the following constraints:

$$\nu_0 \geq 0 \tag{8a}$$

$$\begin{aligned}\sum_i \mu_{i,j} - (\mathbb{P}(W_0 = y_j) - \mathbb{P}(Y_0 = y_j))\nu_0 - \sum_k (\mathbb{P}(W_k = y_j) - \mathbb{P}(Y_k = y_j))\nu_k \\ + (\lambda_j^{(1)} - \lambda_{j-1}^{(1)}) - (\lambda_j^{(2)} - \lambda_{j-1}^{(2)}) \;=\; \mathbb{P}(Z = y_j)\;\forall\, j\end{aligned} \tag{8b}$$

$$\lambda_j^{(2)}(y_{j+1} - y_j) - \lambda_{j-1}^{(1)}(y_j - y_{j-1}) \leq 0\;\forall\, j \tag{8c}$$

$$\sum_j y_j \mu_{i,j} \;\leq\; p_i h(x, \xi_i)\;\forall\, i \tag{8d}$$

$$\sum_j \mu_{i,j} \;=\; p_i\;\forall\, i \tag{8e}$$

$$\boldsymbol{\mu} \;\geq\; 0\,,\; \boldsymbol{\nu} \;\geq\; 0\,,\; \boldsymbol{\lambda}^{(1)} \;\geq\; 0\,,\; \boldsymbol{\lambda}^{(2)} \;\geq\; 0\,, \tag{8f}$$

where we consider $\lambda_0^{(1)} = \lambda_0^{(2)} = 0$. All constraints are linear in the decision variables except for (8d) which is a convex constraint in x if $h(\cdot, \xi)$ is concave. For the stochastic dominance constrained problem (1) we simply add these constraints and variables to the problem, and for the certainty equivalent problem (3) we check their feasibility. In the case of the robust utility maximization problem (2), we let $Z = 0$, take the dual formulation, and then combine the two-stages of minimization to get:

$$\begin{aligned}\max_{x\in\mathcal{X}} \inf_{u\in\mathcal{U}} \mathbb{E}[u(h(x,\xi))] = \max_{\substack{x\in\mathcal{X} \\ \boldsymbol{\mu},\nu_0,\boldsymbol{\nu},\boldsymbol{\lambda}^{(1)},\boldsymbol{\lambda}^{(2)}}} \quad & \nu_0 \\ s.t. \quad & (8b) - (8f)\end{aligned}$$

3.2 Incorporating S-Shape Information

We assume that $y = 0$ is the reference point (i.e., inflection point) for the s-shaped utility function. For simplicity we will include 0 in $\mathcal{S}$ and define the sets $\mathcal{J}^+ = \{j : y_j \geq 0\}$ and $\mathcal{J}^- = \{j : y_j \leq 0\}$. The following theorem is similar to Theorem 1. Since the proof is also similar, we defer it to the appendix 6.2.

Theorem 2. *The optimal value of the linear program*

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{v},\boldsymbol{w},\boldsymbol{s}} \quad \sum_i p_i \left(1\{h(x,\xi_i) < 0\}s_i + 1\{h(x,\xi_i) \geq 0\}(v_i h(x,\xi_i) + w_i)\right) - \sum_j \mathbb{P}[Z = y_j]\alpha_j \tag{9a}$$

$$\text{s.t.} \quad y_j v_i + w_i \geq \alpha_j \;\forall\, i \in \{1,\ldots,M\}, j \in \mathcal{J}^+ \tag{9b}$$

$$s_i \geq \beta_j(h(x,\xi_i) - y_j) + \alpha_j \;\forall\, i \in \{1,\ldots,M\}, j \in \mathcal{J}^- \tag{9c}$$

$$\sum_j \mathbb{P}[W_0 = y_j]\alpha_j - \sum_j \mathbb{P}[Y_0 = y_j]\alpha_j = 1 \tag{9d}$$

$$\sum_j \mathbb{P}[W_k = y_j]\alpha_j \geq \sum_j \mathbb{P}[Y_k = y_j]\alpha_j \qquad \forall k = 1,\ldots,K \tag{9e}$$

$$\alpha_{j+1} - \alpha_j \geq \beta_{j+1}(y_{j+1} - y_j) \;\forall\, j \in \mathcal{J}^+ \setminus \{N\} \tag{9f}$$

$$\alpha_{j+1} - \alpha_j \leq \beta_j(y_{j+1} - y_j) \;\forall\, j \in \mathcal{J}^+ \setminus \{N\} \tag{9g}$$

$$\alpha_j - \alpha_{j-1} \leq \beta_j(y_j - y_{j-1}) \;\forall\, j \in \mathcal{J}^- \tag{9h}$$

$$\alpha_j - \alpha_{j-1} \geq \beta_{j-1}(y_j - y_{j-1}) \;\forall\, j \in \mathcal{J}^- \tag{9i}$$

$$\boldsymbol{v} \geq 0\,, \quad \boldsymbol{\beta} \geq 0 \tag{9j}$$

equals $\psi(x;\mathcal{U}^s)$. Furthermore, a worst-case utility function (i.e., one achieving the infimum in (4)) is

$$u^*(y) = \begin{cases} \alpha_N & y \geq y_N \\ \frac{\alpha_{j+1}-\alpha_j}{y_{j+1}-y_j} y + \frac{y_{j+1}\alpha_j - y_j\alpha_{j+1}}{y_{j+1}-y_j} & y_j \leq y < y_{j+1} \;\forall\, j \in \mathcal{J}^+ \\ \max_{j' \in \{j,j+1\}} \beta_{j'}(y - y_{j'}) + \alpha_{j'} & y_j \leq y < y_{j+1} \;\forall\, j \in \mathcal{J}^- \\ \beta_1(y - y_1) + \alpha_1 & y < y_1 \end{cases} \tag{10}$$

This is a piecewise linear function connecting the points $u(y_j) = \alpha_j$.

Unfortunately, the general problems (1), (2), and (3) are probably hard to solve under $\mathcal{U}^s$ because even maximizing expected utility with an s-shape utility function may have multiple local maxima. Nevertheless, Theorem 2 allows us to evaluate $\psi(x;\mathcal{U}^s)$ and its derivative (using LP sensitivity analysis), despite its infinite dimensional nature. This suggests that local or global nonlinear optimization methods may be applicable.

3.3 Incorporating Prudence Information

Our results are weaker for $\mathcal{U}^3$. We will assume that $[a, b]$ contains the support of all the random variables involved in this problem. We then discretize this interval, adding values to $\mathcal{S}$ to minimize the largest $y_{j+1} - y_j$.

Theorem 3. *The optimal value of the linear program*

$$
\begin{aligned}
\hat{\psi}(x) \;:=\; \min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{v},\boldsymbol{w}} \quad & \sum_i p_i(v_i h(x,\xi_i) + w_i) - \sum_j \mathbb{P}[Z = y_j]\alpha_j && \text{(11a)}\\
s.t. \quad & y_j v_i + w_i \geq \alpha_j \;\forall\, i \in \{1,\ldots,M\}, j \in \{1,\ldots,N\} && \text{(11b)}\\
& \sum_j \mathbb{P}[W_0 = y_j]\alpha_j - \sum_j \mathbb{P}[Y_0 = y_j]\alpha_j = 1 && \text{(11c)}\\
& \sum_j \mathbb{P}[W_k = y_j]\alpha_j \geq \sum_j \mathbb{P}[Y_k = y_j]\alpha_j \qquad \forall k = 1,\ldots,K && \text{(11d)}\\
& \alpha_{j+1} - \alpha_j \geq \beta_{j+1}(y_{j+1} - y_j) \;\forall\, j \in \{1,\ldots,N-1\} && \text{(11e)}\\
& \alpha_{j+1} - \alpha_j \leq \beta_j(y_{j+1} - y_j) \;\forall\, j \in \{1,2,\ldots,N-1\} && \text{(11f)}\\
& \beta_{j+1} - \beta_j \leq \gamma_{j+1}(y_{j+1} - y_j) \;\forall\, j \in \{1,\ldots,N-1\} && \text{(11g)}\\
& \beta_{j+1} - \beta_j \geq \gamma_j(y_{j+1} - y_j) \;\forall\, j \in \{1,\ldots,N-1\} && \text{(11h)}\\
& \boldsymbol{\beta} \geq 0\,,\;\; \boldsymbol{\gamma} \leq 0,\;\; \boldsymbol{v} \geq 0 && \text{(11i)}
\end{aligned}
$$

is a lower bound for $\psi(x;\mathcal{U}^3)$. Furthermore, an approximate worst-case utility function is the piecewise linear function

$$
\hat{u}^*(y) = \begin{cases} \alpha_N & y \geq y_N \\ \frac{\alpha_{j+1}-\alpha_j}{y_{j+1}-y_j} y + \frac{y_{j+1}\alpha_j - y_j\alpha_{j+1}}{y_{j+1}-y_j} & y_j \leq y < y_{j+1} \;\forall\, j \in \{1,\ldots,N-1\} \\ -\infty & y < y_1. \end{cases} \tag{12}
$$

The proof can be found in the appendix 6.3. As the discretization becomes finer, we expect that the approximate value $\hat{\psi}(x)$ and approximate worst-case utility function $\hat{u}^*(\cdot)$ converge respectively to $\psi(x;\mathcal{U}^3)$ and its worst-case utility function.

Remark 3.2. *This theorem can be used to solve problems with third order stochastic dominance constraints because third order dominance of $h(x, \xi)$ over Z is equivalent to $\psi(x;\mathcal{U}^3) \geq 0$ with*

$K = 0$. *The existing approach for such problems uses the fact that* $h(x,\xi) \succeq_3 Z$ *is equivalent to* $\mathbb{E}[\max(0, y - h(x,\xi))^2] \leq \mathbb{E}[\max(0, y - Z)^2]$ *for all* $y \in \mathbb{R}$ *(Ogryczak and Ruszczyński, 2001). Verifying this inequality at a discrete set of points is a tractable approximation. However, Theorem 3 above, leads to a more practical approximation because 1) it only imposes linear instead of quadratic constraints; 2) it provides an inner (instead of an outer) approximation for the set of feasible* x *ensuring that dominance holds for all feasible points in the approximation; and 3) it allows us to account for additional information about the utility function.*

4 Numerical Results

In this section, we use a portfolio optimization problem to illustrate the gains that can be achieved by adopting formulations that account for the preference information that is available. In this portfolio optimization problem, we assume that there are n assets, and we let x_i be the proportion of the total budget allocated to asset i. Since we do not consider short positions, the feasible set for the vector of allocations x is the convex set $\mathcal{X} := \{x \in \mathbb{R}^n : x \geq 0,\ x \cdot 1 = 1\}$. Let ξ_i be the random weekly return of asset i. Then we let the random outcome $h(x,\xi) := x \cdot \xi$ be the return of the portfolio.

We consider two formulations. First, we consider a formulation that attempts to maximize the certainty equivalent of the constructed portfolio:

$$\max_{x \in \mathcal{X}} \ \mathbb{C}_{\bar{u}}(x \cdot \xi) \ , \tag{13}$$

where $\bar{u}$ is the utility function that would capture exactly the complete preference of our decision maker, the investor. When preference information is incomplete, i.e. only K pairwise comparisons have been made by the decision maker, the utility function is only known to lie in a set of type $\mathcal{U}^2$. Hence, one can either use this information to estimate the true utility function by some function $\hat{u}$ and solve problem (13) with $\hat{u}$ instead of $\bar{u}$, or solve the robust certainty equivalent formulation (3) with $h(x,\xi) = x \cdot \xi$. The later will effectively return a portfolio that is preferred to the bank account with largest fixed interest rate.

Alternatively, our second formulation attempts to maximize expected return of the portfolio under the constraint that this portfolio is preferred by the investor to the return of a given benchmark portfolio Z. Specifically, we are interested in solving:

$$\max_{x \in \mathcal{X}} \quad \mathbb{E}[x \cdot \xi] \tag{14a}$$

$$s.t. \quad \mathbb{E}[\bar{u}(x \cdot \xi)] \geq \mathbb{E}[\bar{u}(Z)] \tag{14b}$$

This time, in the case of incomplete preference information, while one could replace $\bar{u}$ by some estimated $\hat{u}$, we will follow the spirit of stochastic dominance, as presented in Dentcheva and Ruszczyński (2006), which suggests replacing constraint (14b) with

$$\mathbb{E}[u(x \cdot \xi)] \geq \mathbb{E}[u(Z)] \; \forall \, u \in U_2 \, .$$

Note that this approach disregards all the preference information except for the fact that the investor is risk averse. By allowing one to replace U_2 by $\mathcal{U}^2$ in problem (14), our approach corrects for this weakness.

After presenting the data used to parameterize these problem, in what follows we present empirical results that demonstrate how, in a context with incomplete preference information, decisions can improve: first, by using a worst-case analysis that accounts appropriately for this information instead of simply using an estimate $\hat{u}$ or being overly conservative through replacing $\mathcal{U}^2$ with U_2; and second, by gathering preference information that is pertinent with respect to the nature of the decision that needs to be made. Indeed, as we ask more questions, and K increases, we expect the set of potential utilities, $\mathcal{U}^2$, to shrink as our knowledge becomes better, and our portfolio performance should improve. Then in subsection 4.3.1 we describe the queries used to elicit the decision maker's preferences. We discuss the results in subsection 4.3.

4.1 Data

We gathered the weekly returns of the companies in the S&P 500 index from 3/30/1993 to 7/6/2011. We focused on the 351 companies that were continuously part of the index during this period. While not including companies that were removed from the index creates some survivorship bias, our results should remain meaningful because the absolute returns are not our focus. For each run, we randomly chose 10 companies from the pool of 351 to be our $n = 10$ assets. We considered $M = 50$ equally likely scenarios for the weekly asset returns which we choose by randomly selecting a contiguous 50-week period of historical returns for the selected companies from the data. For the stochastic dominance formulation the distribution of the benchmark return Z is given by the weekly return of the S&P 500 index during the same period.

4.2 Effectiveness of robust approach

Our first numerical study attempts to determine whether there is something to gain by accounting explicitly for available preference information in our portfolio optimization model instead of assuming more naively that the utility function takes on one of the popular shapes. In our simulation, the decision maker is risk averse and agrees with the axioms of expected utility, yet is unaware which utility function captures his risk attitude. Information about this attitude will be obtained through comparison of randomly generated pairs of lotteries (using the "Random Utility Split method" described in section 4.3.1), thus can be represented by $\mathcal{U}^2$. Although he is unaware of this, when making a comparison, the simulated decision maker acts according to the utility function $\bar{u}(y) = -20E_i(20/y) + y\exp(20/y)$, where E_i stands for the exponential integral $E_i(y) := -\int_{-x}^{\infty} \exp(-t)/t dt$. Our experiments consist of comparing four utility function selection strategies with respect to their average performance at maximizing the portfolio's certainty equivalent over a random sets of 10 companies and 50 scenarios, which are drawn as described in section 4.1.

Remark 4.1. *The function $\bar{u}(y) = -20E_i(20/y) + y\exp(20/y)$ was chosen because it has the property that $-u''(y)y^2/u'(y) = 20$. Hence, if the decision maker is only asked to compare lotteries that involve weekly returns close to 0%, then one might conclude that the absolute risk aversion of this decision maker is constant (i.e. their utility function takes the exponential form) when in fact their absolute risk aversion is decreasing and scales proportionally to $1/x^2$.*

4.2.1 Utility function selection strategies

We consider four different approaches to dealing with incomplete preference information that takes the form of a set of pairwise comparisons under the risk aversion hypothesis, i.e. $\mathcal{U}^2$.

Exponential fit This approach simply suggests approximating $u(\cdot)$ with $\hat{u}(\cdot)$ obtained by fitting an exponential utility function of the form $u_c(y) = (1 - \exp(-cy))$ to the available information. For implementation details, we refer the reader to appendix 6.4. It is interesting to note that, when a decision maker has constant absolute risk aversion, it is sufficient to identify the certainty equivalent of a single lottery to learn exactly the values that c should take. Unfortunately, here the decision maker has decreasing risk aversion, hence as $\mathcal{U}^2 \to \{\bar{u}(\cdot)\}$, the best fitted function will become unable to fit $\bar{u}(\cdot)$ exactly.

Piecewise linear fit This approach simply suggests approximating $u(\cdot)$ with $\hat{u}(\cdot)$ obtained by fitting a piecewise linear concave utility function of the form $u_{\alpha,\beta}(y) = \min_i \alpha_i y + \beta_i$ to the available information about the true utility function. Our implementation follows similar lines as used for the exponential utility function with the single exception that we enforce that $u_{\alpha,\beta}$ be in $\mathcal{U}^2$. The best fitted piecewise linear utility function does have a more complex representation: for instance, in our implementation the number of linear pieces was comparable to the size $|S|$.

Worst-case utility function This approach suggests decisions that achieve the best worst-case performance over the set of potential utility functions. See section 3.1 for implementation details.

True utility function This approach plays the role of a reference for the best performance that can be achieved in each decision context. This is done by assuming that the decision maker actually knows that their preference can be represented by the form $u(y) = -20E_i(20/y) + y\exp(20/y)$. While we argued that this situation is unlikely to occur in practice, we hope to verify that the approaches based on a piecewise linear fit or the worst-case utility function are consistent in the sense that the decisions they suggest will actually converge, as more information is obtained about the decision maker's preferences, to the decisions that should be taken if the true utility function was known.

4.2.2 Results

Table 1 presents a comparison of the 1st percentile and average of certainty equivalents achieved in 100 000 experiments when maximizing the certainty equivalent under incomplete preference information using the four utility function selection strategies that were just described. Note that the certainty equivalents which statistics are reported in this figure were evaluated using the true utility function. Since our simulations did not include a risk free option, optimal portfolios did have negative certainty equivalents on occasion in contexts were the 50 scenarios were taken from a period with a declining economy. A method might also suggest a portfolio with negative certainty equivalent if the utility function that is used to measure performance actually overestimates its expected utility.

First, we can confirm that, since they always employ utility functions that are members of $\mathcal{U}^2$ and since $\mathcal{U}^2 \to \{\bar{u}\}$, the piecewise linear and worst-case utility function approaches suggest decisions which performance converge, in terms of 1st percentile and average value, to the performance achieved knowing the true utility function. It is also as expected that making the false assumption that absolute risk aversion is constant, i.e.using an exponential utility function, can potentially lead to a significant loss in performance especially when a large quantity of information about the decision maker's risk attitude has been gathered. Indeed, the results indicate that in these experiments, after 80 queries were performed, the method that used the best fitted exponential utility function proposed portfolios that on average were equivalent to a negative guaranteed return while other methods were able to suggest portfolios that on average were equivalent to a 0.1% guaranteed weekly return investment (i.e. 5.3% annually) in terms of the decision maker's preferences. Finally, we can confirm that choosing a portfolio based on the worst-case utility function is statistically more robust, in terms of average and 1st percentile of the performance, when little preference information is available.

	Certainty equivalent (in percentage point)					
Information	5 queries		20 queries		80 queries	
Utility	1st perc.	Average	1st perc.	Average	1st perc.	Average
Exponential	-3.9 ± 1.1	-0.03 ± 0.04	-6.4 ± 1.8	-0.09 ± 0.06	-5.7 ± 0.6	-0.11 ± 0.07
Piecewiselinear	-6.7 ± 0.3	-0.57 ± 0.04	-3.8 ± 0.4	-0.08 ± 0.02	-2.0 ± 0.2	0.12 ± 0.01
Worst-case	-2.3 ± 0.3	-0.12 ± 0.02	-2.3 ± 0.3	-0.06 ± 0.02	-2.1 ± 0.2	0.08 ± 0.02
True	-1.8 ± 0.2	0.14 ± 0.01	-1.8 ± 0.2	0.14 ± 0.01	-1.8 ± 0.2	0.14 ± 0.01

Table 1: Comparison of the 1st percentile and average (with their respective 99% confidence intervals) of certainty equivalents (in percentage points) achieved in 100 000 experiments by maximizing the certainty equivalent under incomplete preference information using four utility function selection strategies. An experiment consists of randomly sampling a set of 10 companies as candidates for investment, a set of 50 return scenarios, and a set of 5, 20, or 80 answered queries.

4.3 Effectiveness of elicitation strategies

The following results shed some light on how decisions might be improved by gaining more information about the preferences of the decision maker. In particular, we compare how performance is improved as we increase the number of questions asked to the decision maker using four different elicitation strategies presented in section 4.3.1. For simplicity, in our simulation, the decision maker's true utility function over the weekly return now has a constant absolute risk aversion level of 10: $\bar{u}(y) := 1 - e^{-10y}$. Note that although the decision maker is unaware that his preferences can be represented by this function, we assume that he never contradicts the conclusions suggested by such a utility function when comparing lotteries to each other. Our experiments consist of evaluating, as the number of queries is increased, the average performance achieved by the robust approach over a random sets of 10 companies and 50 scenarios, which are drawn as described in section 4.1.

4.3.1 Elicitation strategies

We elicit information about the investor's preferences by asking them to choose the preferred among two random outcomes. For simplicity, we only consider questions that compare a certain outcome to a risky gamble with two outcomes (a.k.a. the Becker-DeGroot-Marschak reference lottery (Becker et al., 1964)). In other words, each query can be described by four values $r_1 \leq r_2 \leq r_3$ and a probability p. These four values specify the question: "Do you prefer a certain return of r_2 or a lottery where the return will be r_3 with probability p and r_1 with probability $1-p$?" If we normalize the utilities such that $u(r_1) = 0$ and $u(r_3) = 1$, then this query will identify whether $u(r_2) > p$ or not. We now describe three different schemes for sequentially choosing questions to ask the investor.

Random Utility Split This scheme lets r_1 and r_3 be the worst and best possible returns, and chooses r_2 uniformly from $[r_1, r_3]$. The scheme then seeks to reduce by half the interval $I := \{u(r_2) : u \in \mathcal{U}^2\}$ of potential utility values at r_2. Thus we choose p so that $pu(r_3)+(1-p)u(r_1)$ is the midpoint of I.

Random Relative Utility Split This scheme differs from the previous by choosing r_1 and r_3 uniformly at random from the range of potential returns and then setting $r_2 := (r_1 + r_3)/2$. Like the previous scheme we seek to reduce by half the interval $I := \{u(r_2) : u \in \mathcal{U}^2\}$, and thus, choose p so that $pu(r_3) + (1-p)u(r_1)$ is the midpoint of I.

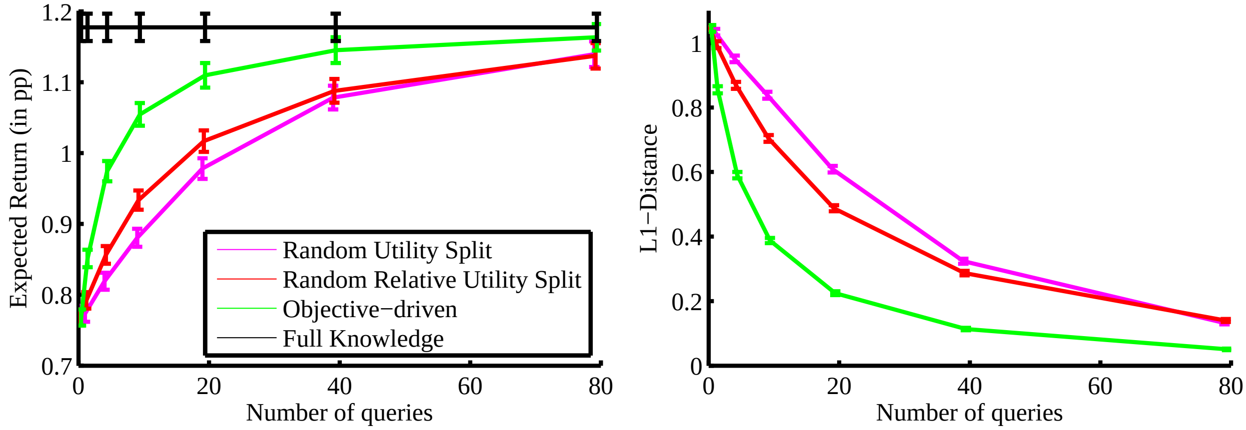
Objective-driven Relative Utility Split Unlike the previous schemes, this scheme takes the optimization objective into account and seeks to improve the optimal objective value as much as possible regardless of the answer (i.e., positive or negative) to the query. To do so, it generates 10 queries using the *Random Relative Utility Split* scheme and of those chooses the query that will give the greatest improvement in the optimal objective value for either a positive or negative answer.

4.3.2 Results

While figure 1 relates to the stochastic dominance formulation, figure 2 relates to the robust certainty equivalent formulation. Figures 1(a) and 2(a) show how our objective value improves as we gain more knowledge about the investor's preferences. Figures 1(b) and 2(b) focus on the convergence of the optimal allocation.

For both formulations, we observe that the total gain between no knowledge of preferences except risk-aversion and full knowledge is worth, on average, 0.4 percentage points of weekly return. We can also see that the improvement in performance is quick for the initial 10–20 queries. In fact for the certainty equivalent formulation, four questions chosen with the objective-driven questioning scheme increases the average certainty equivalent of the weekly return by 0.2 percentage points. After these first queries, the gains from additional information decrease. This seems to indicate that there is considerable value in using all the preference information that is available, even if minimal, thus encouraging the use of our stochastic dominance formulation instead of the one presented in Dentcheva and Ruszczyński (2006), which here would achieve the performance associated to 0 queries.

Finally, for both formulation, it is quite noticeable that the choice of questions to ask to the decision maker also has an important impact on performance: the improvement is faster for the more sophisticated objective-driven elicitation scheme than for the more simpler schemes. From figure 3, which shows graphically how the set of potential utilities $\mathcal{U}^2$ shrinks as we gather more information using the objective-driven strategy, we can finally observe that, if queries are well chosen, a manageable number of them (e.g., 40 queries) can already allow $\mathcal{U}^2$ to identify most of the behavior of the true utility function. We believe this should justify further research on what constitutes an optimal learning strategy in this context.



(a) (b)

Figure 1: Effect of increasing numbers of questions in a stochastic domination formulation: (a) presents the expected return (in percentage points) and (b) presents the L1-distance between optimal allocation with K queries and the optimal allocation with full knowledge. Shown are averages and standard errors from 1500 simulations.

5 Conclusion

In this paper we presented tractable approaches to dealing with incomplete information about the utility function. We looked at three different formulations of our aims and three different types of sets of utility functions. Particularly useful is how our formulations can incorporate information such as that the decision maker would always choose lottery A over lottery B. Particularly novel are the models involving s-shaped utilities and prudent utilities. There is little work on optimizing with such utilities despite them being common in behavioral economics and finance. Also quite innovative is the robust certainty equivalent formulation. This seems like it may have many uses in situations where one currently would use a stochastic dominance constraint. Not only does the certainty equivalent formulation have a natural interpretation, but unlike the formulation with a stochastic dominance constraint, there is no need to also define a separate objective.

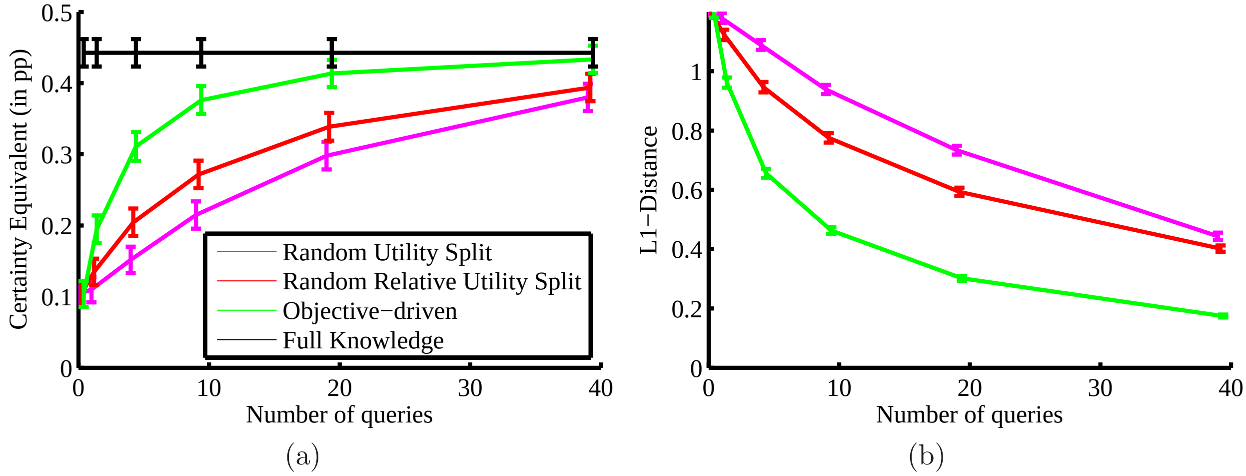


Figure 2: Effect of increasing numbers of questions in a worst-case certainty equivalent formulation: (a) presents the certainty equivalent (in percentage points) of the optimal portfolio as measured with respect to the true utility function and (b) presents the L1-distance between optimal allocation with K queries and the optimal allocation with full knowledge. Shown are averages and standard errors from 500 simulations.

References

K. Arrow. Rational choice functions and orderings. *Economica*, 26:121–127, 1959.

R. J. Aumann. Utility theory without the completeness axiom. *Econometrica*, 30(3):pp. 445–462, 1962. URL http://www.jstor.org/stable/1909888.

G. Becker, M. DeGroot, and J. Marschak. Measuring utility by a single-response sequential method. *Behavioral Science*, 9:226–232, 1964.

A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

C. Boutilier, R. Patrascu, P. Poupart, and D. Schuurmans. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence*, 170(8-9):686–713, 2006.

U. Chajewska, D. Koller, and R. Parr. Making rational decisions using adaptive utility elicitation. In *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI)*, pages 363–369. 2000.

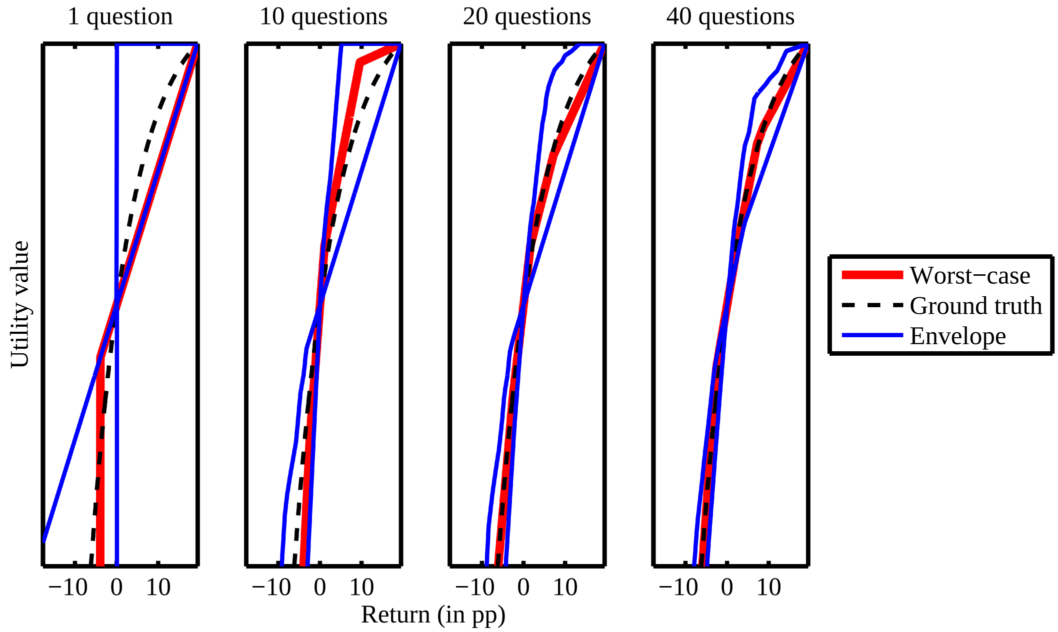


Figure 3: Evolution of the bounding envelope of utility functions in $\mathcal{U}^2$ obtained in one simulation as a growing number of queries generated using the objective-driven strategy have been answered. The worst-case utility function is obtained by solving the dominance-constrained portfolio problem.

R. Clemen and T. Reilly. *Making Hard Decisions with Decisiontools Suite*. Duxbury, 1st edition, 1999.

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

D. Dentcheva and A. Ruszczyński. Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2):548–566, 2003. URL http://link.aip.org/link/?SJE/14/548/1.

D. Dentcheva and A. Ruszczyński. Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, 30(2):433–451, 2006.

J. Dubra, F. Maccheroni, and E. Ok. Expected utility theory without the completeness axiom. *Journal of Economic Theory*, 115(1):118–133, 2004.

J. Grable and R. H. Lytton. Financial risk tolerance revisited: the development of a risk assessment instrument. *Financial Services Review*, 8(3):163–181, 1999.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

M. S. Kimball. Precautionary saving in the small and in the large. *Econometrica*, 58(1):53–73, 1990.

H. Leland. Saving and uncertainty: the precautionary demand for saving. *Quaterly Journal of Economics*, 82:465–473, 1968.

W. Ogryczak and A. Ruszczyński. On stochastic dominance and mean-semideviation models. *Math. Programming*, 89:217–232, 2001.

J. Rawls. *A Theory of Justice*. Harvard University Press, Cambridge, MA, 1971.

A. Sandmo. The effect of uncertainty on saving decisions. *The Review of Economic Studies*, 37(3):353–360, 1970.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

6 Appendix

6.1 Quasiconcavity of objective in problem (3)

For any fixed u, the certainty equivalent $\mathbb{C}_u[h(x,\xi)]$ is quasiconcave in x because $\mathbb{C}_u[h(x,\xi)] \geq t$ is equivalent to $\mathbb{E}[u(h(x,\xi))] \geq u(t)$ and the left hand side is concave in x. Thus, the pointwise infimum $\inf_{u\in\mathcal{U}} \mathbb{C}_u[h(x,\xi)]$ is quasiconcave in x with $\inf_{u\in\mathcal{U}} \mathbb{C}_u[h(x,\xi)] \geq t$ equivalent to $\psi(x;\mathcal{U},t) \geq 0$. Hence the optimum of (3) is greater than t if and only if the constraint $\psi(\cdot;\mathcal{U},t) \geq 0$ has a feasible solution $x \in \mathcal{X}$. Thus if we can bound the optimum to an interval $[t_0, t_1]$, then we can use the bisection algorithm on t to solve the problem to precision ϵ by solving $O(\log(1/\epsilon))$ feasibility problems of the type $\psi(\cdot;\mathcal{U},t) \geq 0$. In fact, any bound on $h(x,\xi)$ gives a bound on the optimum: if $t_0 \leq h(x,\xi) \leq t_1$ for all ξ and all $x \in \mathcal{X}$ then the optimum is in $[t_0, t_1]$.

6.2 Proof of theorem 2

The proof is similar to that of Theorem 1, and we start by defining

$$U(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \{u : u(y_j) = \alpha_j,\ \beta_j \in u'(y_j)\ \forall j\}.$$

Then by a similar argument,

$$\begin{aligned} \psi(x; \mathcal{U}^s) = \min_{\boldsymbol{\alpha}}\ & \psi(x; U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cap U_s) \\ & U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cap U_s \neq \emptyset,\ U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \subseteq U_a,\ U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \subseteq U_n. \end{aligned}$$

The constraint $U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cap U_s \neq \emptyset$ is represented by (9f)–(9i) and $\boldsymbol{\beta} \geq 0$; $U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \subseteq U_n$ is by (9d); and $U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \subseteq U_n$ is by (9e).

Again, since $\mathbb{E}[u(Z)]$ is a constant for $u \in U(\boldsymbol{\alpha}, \boldsymbol{\beta})$, evaluating $\psi(x; U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cap U_s)$ is equivalent to minimizing $\mathbb{E}[u(h(x, \xi))]$ over $u \in U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cap U_s$. Among the s-shaped functions in $U(\boldsymbol{\alpha}, \boldsymbol{\beta})$, this is minimized by the piecewise linear function u^* in (10). The reason u^* is minimal is for the concave portion is the same as in the proof of Theorem 1. For the convex portion, u^* is minimal because the lines $\{(y, s) : s = \beta_j(y - y_j) + \alpha_j\}$ with $j \in \mathcal{J}^-$ are supporting hyperplanes for the convex portion of any function in $U(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Thus, when $U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cap U_s \neq \emptyset$,

$$\psi(x; U(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cap U_s) = \mathbb{E}[u^*(h(x, \xi))1\{h(x, \xi) < 0\}] + \mathbb{E}[u^*(h(x, \xi))1\{h(x, \xi) \geq 0\}] - \sum_j \mathbb{P}[Z = y_j]\alpha_j.$$

Since u^* is concave for $y \geq 0$, the formulation (7) can be used for that part. When $y < 0$, u^* is convex, and thus,

$$\begin{aligned} u^*(y) = \min_s\ & s \\ s.t.\quad & s \geq \beta_j(h(x, \xi_i) - y_j) + \alpha_j\ \ \forall\, j \in \mathcal{J}^-. \end{aligned}$$

Putting these pieces together gives us the objective (9a) and the constraints (9b), $\boldsymbol{v} \geq 0$, and (9c). $\square$

6.3 Proof of theorem 3

As in the proof of Theorem 1, we define $U(\boldsymbol{\alpha}) := \{u : u(y_j) = \alpha_j \ \forall j\}$. Then by an argument similar to one used in the proof of Theorem 1,

$$\psi(x;\mathcal{U}^3) = \min_{\boldsymbol{\alpha}} \psi(x; U(\boldsymbol{\alpha}) \cap U_2 \cap U_3)$$
$$U(\boldsymbol{\alpha}) \cap U_2 \cap U_3 \neq \emptyset,\ U(\boldsymbol{\alpha}) \subseteq U_a,\ U(\boldsymbol{\alpha}) \subseteq U_n.$$

The constraint $U(\boldsymbol{\alpha}) \subseteq U_n$ is represented by (11c) and $U(\boldsymbol{\alpha}) \subseteq U_n$ is by (11d). Since we only seek a lower bound, we can represent $U(\boldsymbol{\alpha}) \cap U_2 \cap U_3 \neq \emptyset$ by the constraints (11e)–(11h), $\boldsymbol{\beta} \geq 0$, and $\boldsymbol{\gamma} \leq 0$.

Again, since $\mathbb{E}[u(Z)]$ is a constant for $u \in U(\boldsymbol{\alpha})$, evaluating $\psi(x; U(\boldsymbol{\alpha}) \cap U_2 \cap U_3)$ is equivalent to minimizing $\mathbb{E}[u(h(x,\xi))]$ over $u \in U(\boldsymbol{\alpha}) \cap U_2 \cap U_3$. Among the utilities in $U(\boldsymbol{\alpha}) \cap U_2$ (thus giving a lower bound), this is minimized by the function $\hat{u}^*$ in (12), as As in the proof of Theorem 1. Thus, we seek to minimize $\mathbb{E}[\hat{u}^*(h(x,\xi))] - \sum_j \mathbb{P}[Z = y_j]\alpha_j$. Since $\hat{u}^*$ is concave, we use the formulation (7), which then gives us the objective (11a) and the constraints (11b) and $\boldsymbol{v} \geq 0$. □

6.4 Fitting an exponential utility function to $\mathcal{U}^2$

To fit a utility function, common practice typically suggests fixing the utility value at two reference points $u(y_0) = 0$ and $u(w_0) = 1$, and using queries to locate the relative utility values achieved at a set of returns $u_j \approx u(y_j)\ \forall\, j = 1, 2, ..., J$. The "best fitted" function is then the one that maximize the following mean square error problem:

$$\begin{aligned} \min_{a,b,c} \quad & \sum_{j=1}^{J} (a(1 - \exp(-cy_j)) + b - u_j)^2 \\ s.t. \quad & a(1 - \exp(-cy_0)) = 0 \ \ \& \ \ a(1 - \exp(-cw_0)) = 1 \\ & a \geq 0\,, c \geq 0\,. \end{aligned}$$

Although non-convex, this problem is typically considered computationally feasible since it reduces to a search over the single parameter c. We adapt this procedure to the context where the preference information takes the shape of $\mathcal{U}^2$. Specifically, without loss of generality, we first let Y_0 and W_0 be certain lotteries and fixed $\mathbb{E}[u(Y_0)] = 0$ and $\mathbb{E}[u(W_0)] = 1$. Next, for a set of $\{y_j\}_{j=1}^{J}$, we can use the information in $\mathcal{U}^2$ to evaluate a range of possible utility values at each y_j. We let u_j take on the mid-value of this interval, $u_j := (\min_{u\in\mathcal{U}^2} u(y_j) + \max_{u\in\mathcal{U}^2} u(y_j))/2$, hence capturing the fact that we wish the exponential utility function pass as close as possible to the center of the intervals in which we know the function should pass. We solve the same mean square error problem to select our best fitted exponential utility function $\hat{u}(y)$. Note that this approach reduces to the method described above when $\mathcal{U}^2 = \{u(\cdot)|u(y_j) = u_j\}$. For computational reasons, our implementation used the set $\{y_j\}_{j=0}^J := \bigcup_{k=0}^K \text{supp}(Y_k) \cup \text{supp}(W_k)$, which uniformly spanned the range of possible returns.