Uncertain Data Envelopment Analysis

Matthias Ehrgott[a,*], Allen Holder[b], Omid Nohadani[c]

[a] *Department of Management Science, Lancaster University, Lancaster, UK.*
[b] *Department of Mathematics, Rose-Hulman Institute of Technology, Terre Haute, IN, USA.*
[c] *Department of Industrial Engineering, Northwestern University, Evanston, IL, USA.*

Abstract

Data Envelopment Analysis (DEA) is a nonparametric, data driven method to conduct relative performance measurements among a set of decision making units (DMUs). Efficiency scores are computed based on assessing input and output data for each DMU by means of linear programming. Traditionally, these data are assumed to be known precisely. We instead consider the situation in which data is uncertain, and in this case, we demonstrate that efficiency scores increase monotonically with uncertainty. This enables inefficient DMUs to leverage uncertainty to counter their assessment of being inefficient.

Using the framework of robust optimization, we propose an uncertain DEA (uDEA) model for which an optimal solution determines 1) the maximum possible efficiency score of a DMU over all permissible uncertainties, and 2) the minimal amount of uncertainty that is required to achieve this efficiency score. We show that the uDEA model is a proper generalization of traditional DEA and provide a first-order algorithm to solve the uDEA model with ellipsoidal uncertainty sets. Finally, we present a case study applying uDEA to the problem of deciding efficiency of radiotherapy treatments.

Keywords: Data Envelopment Analysis; Uncertain Data; Robust Optimization; Uncertain DEA Problem; Radiotherapy Design

1. Introduction and Motivation

Data envelopment analysis (DEA) is a well established optimization framework to conduct relative performance measurements among a group of decision

*Corresponding author, phone number +44 1524 593888
Email addresses: m.ehrgott@lancaster.ac.uk (Matthias Ehrgott), holder@rose-hulman.edu (Allen Holder), nohadani@northwestern.edu (Omid Nohadani)

making units (DMUs). There are numerous reviews of DEA, see, e.g., Cooper *et al.* (2007); Emrouznejad *et al.* (2008); Liu *et al.* (2013); Zhu (2014), and Hwang *et al.* (2016); and the concept has found a wide audience in both research and application. The principal idea is to solve an optimization problem for each DMU to identify its efficiency score relative to the other DMUs. Efficiency equates with a score of 1, and if a DMU's efficiency score is less than 1, then that DMU is outperformed no matter how it is assessed against its competitive cohort.

A DEA model is only as good as its data because DMUs are compared against each other through their assessed inputs and outputs. The importance of accurate data is thus acute in establishing a DMU's performance. However, data is often imperfect, and knowledge about the extent of uncertainty can be vague, if not obscure, as errors commonly have several compounding sources. This fact begs the question of whether or not an inefficient DMU might have been so classified because of some realization of inscrutable data, and if so, then there is a reasonable argument against its perceived under-performance. The question we consider is, what is the minimum amount of uncertainty required of the data that could render a DMU efficient?

We address uncertainty through the lens of robust optimization, which is a field of study designed to account for uncertainty in optimization problems. The preeminent theme of robust modeling is to permit a deleterious effect to the objective to better hedge against the uncertain cases that are typically ignored. Indeed, the concern of "over-optimizing" is regularly used to galvanize the use of a robust model that gives a best solution against all reasonable possibilities instead of a non-robust solution that inappropriately exaggerates the weaknesses of estimated or sampled uncertainty. Examples of this sentiment are found in antenna design (Ben-Tal and Nemirovski, 2002), inventory control (Bertsimas and Thiele, 2006), and radiotherapy design (Bertsimas *et al.*, 2010; Chue *et al.*, 2005). References for robust optimization are Ben-Tal *et al.* (2009) and Bertsimas *et al.* (2011).

Our perspective is counter to the orthodoxy that motivates robust models. The diminishing effect of the objective induced by uncertainty is inverted into a beneficial consideration in DEA due to the way efficiency scores are regularly calculated, as in (1) and (3), see Proposition 1. In particular, a DMU's efficiency score is non-decreasing as uncertainty increases. This observation suggests a keen interest in uncertainty by an inefficient DMU, as it may have a legitimate claim to efficiency modulo the imperfections of the data. As such, uncertainty might be leveraged to assert improved, if not efficient, performance within the confines of reasonable data imperfections.

Robust optimization is a relative newcomer to the task of modeling and solving optimization problems with uncertain data. Stochastic programming is the traditional stalwart, and stochastic extensions of DEA have an established literature, see Olesen and Petersen (2016) and references therein. The general relationships among robust programming, stochastic programming, and parametric and sensitivity analysis are well known (Ben-Tal *et al.*, 2009). Here, we take a fist step toward an uncertain DEA model (uDEA model), by adopting principles from robust optimization.

The real-world problem motivating our study of uDEA is that of deciding efficient radiotherapy treatments for prostate cancer (Lin *et al.*, 2013). This DEA application leads to the identification of treatments that could be improved after redesign. However, the inputs and outputs of this problem are suspect due to numerous approximations and errors. Moreover, the extent of uncertainty can only be estimated, and making definitive conclusions about a treatment's inefficiency is questionable. Instead of analyzing treatment quality with certain but imperfect data, we embrace the inherent uncertainty of this application and compare the treatments in light of their uncertain characteristics.

1.1. Contributions of this Paper

We introduce the novel concept of uDEA, a paradigm in which a DMU can select both its best data from an uncertainty set as well as its best efficiency score based on this data. The uDEA problem determines the minimum amount of uncertainty required to increase an efficiency score as much as possible. Unlike traditional DEA, the model is generally nonlinear, and indeed generally nonconvex, and we develop a first-order algorithm to solve the uncertain problem. We further provide modeling constructs and several examples to illustrate the flexibility of uDEA. We conclude with a clinical application that aids treatment evaluation and design. In particular, the application allows a planner to judge the acceptability of a treatment relative to the uncertainty of designing and delivering it.

2. General Data Envelopment Analysis

We assume the (standard) input oriented model with variable returns-to-scale from among the many DEA formulations, where the efficiency score $E^{\hat{\imath}}$ of DMU $\hat{\imath}$ is defined by solving the linear program,

$$E^{\hat{\imath}} = \min\{\theta^{\hat{\imath}} : Y\lambda - y^{\hat{\imath}} \geq 0,\ X\lambda - \theta^{\hat{\imath}} x^{\hat{\imath}} \leq 0,\ e^T\lambda = 1,\ \lambda \geq 0\}, \tag{1}$$

where e is the vector of ones. In this model, there are M outputs, indexed by m; N inputs, indexed by n; and D DMUs, indexed by i. The matrices Y and X are the nonnegative output and input matrices so that

Y_{mi} is the m-th output value for DMU i, and
X_{ni} is the n-th input value for DMU i.

The column vectors $y^{\hat{\imath}}$ and $x^{\hat{\imath}}$ are the $\hat{\imath}$-th columns of X and Y. If $E^{\hat{\imath}} = 1$, then the $\hat{\imath}$-th DMU is *efficient*, otherwise $0 \leq E^{\hat{\imath}} < 1$, and the $\hat{\imath}$-th DMU is *inefficient*.

The development that follows generally applies to numerous other formulations of DEA (Cooper *et al.*, 2007), with only straightforward adjustments being needed if the returns-to-scale assumption is altered, the orientation is changed, or if environmental constraints are added. For example, model (1) assumes variable returns-to-scale as imposed by the constraint $e^T\lambda = 1$, but this constraint could be removed to accommodate constant returns-to-scale without hindering our analysis.

We re-formulate model (1) to ease notation and development. Let

$$\bar{A}^{\hat{\imath}} = \left[\begin{array}{c|c|c} -Y & y^{\hat{\imath}} & 0 \\ \hline X & 0 & -x^{\hat{\imath}} \end{array}\right] \text{ and } B = \left[\begin{array}{c|c|c} e^T & 0 & 0 \\ \hline 0 & 1 & 0 \end{array}\right]. \tag{2}$$

The input oriented DEA model with variable returns-to-scale is thus,

$$E^{\hat{\imath}} = \min\{c^T\eta : \bar{A}^{\hat{\imath}}\eta \leq 0,\ B\eta = e,\ \eta \geq 0\}, \tag{3}$$

where $c = (0, 0, \ldots, 0, 1)^T$ and $\eta = (\lambda, 1, \theta^{\hat{\imath}})$. Let $\bar{A}^{\hat{\imath}}_k$ be the k-th row of $\bar{A}^{\hat{\imath}}$, for which k indexes the collection of outputs and inputs,

$$\{1, 2, \ldots, M, M+1, \ldots, M+N\}.$$

We require a distinction between the data of the DEA model and the data of the linear program in our forthcoming development, and to remove confusion about the term "data," we refer to X and Y as the DEA data and to $\bar{A}^{\hat{\imath}}$ and B as the linear programming instance. Importantly, the solution

$$\eta = (0, 0, \ldots, 0, 1, 0, \ldots, 0, 1, 1) = (e_{\hat{\imath}}^T, 1, 1)$$

is always feasible, where the first 1 is in the $i = \hat{\imath}$ position as noted by $e_{\hat{\imath}}$. The feasibility of this solution ensures that the efficiency score of the $\hat{\imath}$-th DMU is no greater than 1.

3. Uncertain Data Envelopment Analysis

The reliability of a DMU's efficiency score is jeopardized if the data is erroneous, which points to a desire to accommodate suspect data within a DEA application. Uncertain data fits seamlessly into the paradigm of robust linear optimization, and our overarching model adapts this robust perspective. Each constraint $\bar{A}_k^{\hat{\imath}} \eta \leq 0$ is replaced with a set of constraints

$$A_k^{\hat{\imath}} \eta \leq 0, \quad \forall A_k^{\hat{\imath}} \in \mathcal{U}_k, \tag{4}$$

which reduces to the original constraint if the uncertainty set is restricted to the singleton, $\mathcal{U}_k = \{\bar{A}_k^{\hat{\imath}}\}$.

Our goal to quantify the totality of uncertainty is aided by organizing the per constraint description of (4) into a collection that captures all inputs and outputs. Our notational convention in this regard is in Definition 1.

Definition 1. *The uncertain inputs and outputs are:*

(i) $\mathcal{U}_k$ *is an* <u>*uncertainty set*</u> *that models the possible values of the data* $A_k^{\hat{\imath}}$*. Hence each* $A_k^{\hat{\imath}} \in \mathcal{U}_k$ *is a possible row vector of input/output data for the* k*-th input/output.*

(ii) $\mathcal{U} = \{\mathcal{U}_k : k = 1, \ldots, N + M\}$ *is a* <u>*collection of uncertainty sets*</u>*, or more succinctly, a* <u>*collection of uncertainty*</u>*. Hence* $\mathcal{U}$ *contains the totality of uncertainty across all inputs and outputs.*

The DEA paradigm requires that the sets in $\mathcal{U}$ satisfy two restrictions to ensure integrity between the uncertain DEA data and the resulting linear programming instances. First, the X and Y components of each $A_k^{\hat{\imath}} \in \mathcal{U}_k$ are assumed to be nonnegative for all k. Second, we assume for each output m that the m-th element of $y^{\hat{\imath}}$ agrees with $Y_{m\hat{\imath}}$ for all $A^{\hat{\imath}}_m \in \mathcal{U}_m$. Likewise, the n-th element of $x^{\hat{\imath}}$ is assumed to agree with $X_{n\hat{\imath}}$ for all $A^{\hat{\imath}}_n \in \mathcal{U}_n$. This second restriction ensures that the uncertain input and output DEA data remains consistent with the linear programming instances. Without this restriction an input or an output of the $\hat{\imath}$-th DMU could differ in the two places it occurs in (3), making the $\hat{\imath}$-th DMU split into two DMUs for some element(s) of the uncertainty set. These restrictions essentially assume uncertainty of the DEA data, X and Y, instead of the linear programming instances, $A^{\hat{\imath}}$ and B.

Collections of uncertainty are partially ordered by set inclusion of the individual uncertainty sets. So one collection can harbor more uncertainty than another as long as the sets of the former contain those of the latter, a definition we formalize below.

Definition 2. *Consider the two collections of uncertainty,*

$$\mathcal{U}' = \{\mathcal{U}'_k : k = 1, \ldots, N+M\} \quad \textit{and} \quad \mathcal{U}'' = \{\mathcal{U}''_k : k = 1, \ldots, N+M\}.$$

We say that $\mathcal{U}''$ <u>harbors at least the uncertainty of</u> $\mathcal{U}'$, denoted by $\mathcal{U}' \trianglelefteq \mathcal{U}''$, if $\mathcal{U}'_k \subseteq \mathcal{U}''_k$ for $k = 1, \ldots, M+N$.

We define the *robust efficiency score* as the optimal value of the *robust DEA model*,

$$\mathcal{E}^{\hat{\imath}}(\mathcal{U}) \quad := \quad \min\{c^T\eta : A^{\hat{\imath}}_k\eta \le 0, \, \forall\, A^{\hat{\imath}}_k \in \mathcal{U}_k, \; \forall\, k, \, B\eta = e, \, \eta \ge 0\}, \tag{5}$$

where := indicates the definitional extension of $E^{\hat{\imath}}$ in (3) to denote the dependence on uncertain data. We inherit the feasibility of $\eta = (e^T_{\hat{\imath}}, 1, 1)^T$ in the robust DEA model from our restrictions on $\mathcal{U}$, and hence, the maximal efficiency score remains 1. Similar robust evaluations of efficiency are in Arabmaldar *et al.* (2017).

Notice that for two different collections of uncertainty $\mathcal{U}'$ and $\mathcal{U}''$, with $\mathcal{U}' \trianglelefteq \mathcal{U}''$, we have

$$\begin{aligned} \mathcal{E}^{\hat{\imath}}(\mathcal{U}') &= \min\{c^T\eta : A^{\hat{\imath}}_k\eta \le 0, \, \forall\, A^{\hat{\imath}}_k \in \mathcal{U}'_k, \; \forall\, k, \, B\eta = e, \, \eta \ge 0\} \\ &\le \min\{c^T\eta : A^{\hat{\imath}}_k\eta \le 0, \, \forall\, A^{\hat{\imath}}_k \in \mathcal{U}''_k, \; \forall\, k, \, B\eta = e, \, \eta \ge 0\} \quad = \quad \mathcal{E}^{\hat{\imath}}(\mathcal{U}''). \end{aligned}$$

The inequality is immediate because the constraints defining $\mathcal{E}^{\hat{\imath}}(\mathcal{U}'')$ are more restrictive than those defining $\mathcal{E}^{\hat{\imath}}(\mathcal{U}')$. This observation asserts that the robust efficiency score can only improve or remain the same by harboring more uncertainty, which is formally stated in following proposition.

Proposition 1. *For the collections of uncertainty $\mathcal{U}'$ and $\mathcal{U}''$, if $\mathcal{U}' \trianglelefteq \mathcal{U}''$, then $\mathcal{E}^{\hat{\imath}}(\mathcal{U}') \leq \mathcal{E}^{\hat{\imath}}(\mathcal{U}'')$.*

The nondecreasing relationship of Proposition 1 motivates our study, as it suggests that a DMU's perceived inefficiency might be rectified to efficiency if the DEA data harbors sufficient uncertainty.

Our initial question of deciding a minimal amount of uncertainty to render a DMU efficient pre-supposes initial efficiency evaluations based on dubious data. The initial data provides an inflexible, fixed, or certain data collection $\mathcal{U}^o$, where each $\mathcal{U}^o_k \in \mathcal{U}^o$ is the singleton of the nominal vector of the initial data so that $\mathcal{U}^o_k = \{\bar{A}^{\hat{\imath}}_k\}$. Prospective uncertainty must harbor at least the uncertainty of the nominal data, and we only consider $\mathcal{U}$ such that $\mathcal{U}^o \trianglelefteq \mathcal{U}$. Moreover, germane models of uncertainty might induce structural requirements on $\mathcal{U}$, and hence, the permissible collections of $\mathcal{U}$ will likely mandate adherence to restrictions appropriate to the study.

Definition 3. Ω *is the* *universe of possible collections of uncertainty.*

Although $\mathcal{U}^o \trianglelefteq \mathcal{U}$ for each $\mathcal{U} \in \Omega$ by assumption, the subset properties that define $\trianglelefteq$ do not suggest how to quantify an amount of uncertainty. However, comparative evaluations somewhat mandate that $\mathcal{U}$ be associated with a numerical value. We assume an *amount of uncertainty* to be a numerical association with $\mathcal{U}$ as defined below.

Definition 4. *An amount of uncertainty is a mapping*

$$\mathbf{m} : \Omega \to \mathbb{R}_+ : \mathcal{U} \mapsto \mathbf{m}(\mathcal{U})$$

such that

(i) $\mathbf{m}(\mathcal{U}) = 0$ *if and only if* $|\mathcal{U}_k| = 1$ *for* $k = 1, \ldots, M + N$, *and*

(ii) $\mathbf{m}(\mathcal{U}') \leq \mathbf{m}(\mathcal{U}'')$ *if* $\mathcal{U}' \trianglelefteq \mathcal{U}''$.

While $\mathbf{m}(\mathcal{U})$ intuitively hearkens to the idea of a measure or a norm on the universe of uncertain collections, such connections are only, and purposefully, allusionary. The first assumption mandates that zero uncertainty equates with the only case in which there is no uncertainty, and the second assumption imbues monotonicity of $\mathbf{m}(\mathcal{U})$. Note that these assumptions do not generally combine with Proposition 1 to establish a non-decreasing property of the robust efficiency score, and in general

$$\mathbf{m}(\mathcal{U}') \leq \mathbf{m}(\mathcal{U}'') \;\; \not\Rightarrow \;\; \mathcal{E}^{\hat{\imath}}(\mathcal{U}') \leq \mathcal{E}^{\hat{\imath}}(\mathcal{U}'')$$

even though

$$\mathcal{U}' \trianglelefteq \mathcal{U}'' \;\; \Rightarrow \;\; \mathcal{E}^{\hat{\imath}}(\mathcal{U}') \leq \mathcal{E}^{\hat{\imath}}(\mathcal{U}'').$$

Examples of $\mathbf{m}(\mathcal{U})$ are presented in Sections 4 and 5.

We suggest that if the $\hat{\imath}$-th DMU is inefficient and if the data is uncertain in Ω, then we can solve the following *uDEA problem* for the $\hat{\imath}$-th DMU,

$$\begin{aligned} \gamma^* \;&=\; \sup_{0\leq\gamma\leq 1} \Big\{\gamma : \min_{\mathcal{U}\in\Omega} \Big\{\mathbf{m}(\mathcal{U}) : \mathcal{E}^{\hat{\imath}}(\mathcal{U}) \geq \gamma\Big\}\Big\} \qquad (6)\\ &=\; \sup_{0\leq\gamma\leq 1} \Big\{\gamma : \min_{\mathcal{U}\in\Omega} \Big\{\mathbf{m}(\mathcal{U}) : \min_{\eta\geq 0} \Big\{c^T\eta : \\ &\qquad A_k^{\hat{\imath}}\eta \leq 0,\ \forall A_k^{\hat{\imath}} \in \mathcal{U}_k,\ k = 1, 2, \ldots, M+N,\ B\,\eta = e\Big\} \geq \gamma\Big\}\Big\}. \end{aligned}$$

We note that each of $\mathbf{m}(\mathcal{U})$, Ω, and γ^* can depend on $\hat{\imath}$. For instance, $\mathbf{m}(\mathcal{U})$ and Ω could be modeled differently for each $\hat{\imath}$. However, even if $\mathbf{m}(\mathcal{U})$ and Ω are consistent for all $\hat{\imath}$, the value of γ^* could still vary depending on which $\hat{\imath}$ is under study. That said, we eschew a notational dependence on $\hat{\imath}$ for $\mathbf{m}(\mathcal{U})$, Ω, and γ^* to ease notation. The use of a supremum instead of a maximum keeps the problem well defined, as there are cases in which no maximum exists although the supremum does. An example of such a case is presented in Section 5. Should γ^* be achievable, then the components of an optimal solution (γ^*, $\mathcal{U}^*$, η^*) are

- the maximum possible efficiency score γ^* over the universe of permissible uncertainty,
- a collection of uncertainty $\mathcal{U}^*$ with minimal $\mathbf{m}(\mathcal{U}^*)$ that is required to achieve the maximum possible efficiency score, and
- an optimal vector η^* that contains λ^*, which identifies an efficient target for the $\hat{\imath}$-th DMU relative to the robust DEA model (5).

An optimal solution to (6) is immediate should the $\hat{\imath}$-th DMU be efficient with the nominal data $\bar{A}$. In this case $\gamma^* = 1$, $\mathcal{U}^* = \mathcal{U}^o$, $\mathbf{m}(\mathcal{U}^*) = 0$, and $\eta^* = (e_{\hat{\imath}}, 1, 1)$. In other words, an originally efficient DMU has already identified a data instance in which it is efficient, and no further uncertainty is required. DMUs that were originally inefficient with $\bar{A}^{\hat{\imath}}$ might have instead been able to improve their efficiency score had a more beneficial data set in Ω been selected. If the efficiency score could improve to 1, then the originally perceived inefficiency could be the unfavorable byproduct of the original data rather than some assumed structural weakness.

Three observations about model (6) deserve comment. First, the model's intent is concomitant with DEA's bedrock supposition to present each DMU with the possibility of maximizing its efficiency score. Model (6) does the same, but it also allows a DMU to select a best possible collection of input and output data. Second, the objectives of increasing γ and minimizing $\mathbf{m}(\mathcal{U})$ are stated sequentially, as this seems appropriate from a DMU's perspective. The same modeling goal can be achieved with the following lexicographic optimization problem provided that the supremum can be replaced with a maximum,

$$\operatorname*{lexmin}_{\gamma \in [0,1],\, \mathcal{U} \in \Omega} \begin{pmatrix} -\gamma \\ \mathbf{m}(\mathcal{U}) \end{pmatrix} \quad \text{s.t.} \quad \mathcal{E}^{\hat{\imath}}(\mathcal{U}) \geq \gamma.$$

The lexicographic minimum assumes γ is as large as possible before it considers reductions in the amount of uncertainty. Alternate models would not necessarily assume a lexicographic ordering of the objectives and could instead seek Pareto-optimal solutions to help assess the trade-off between the maximum possible efficiency score γ and the amount of uncertainty $\mathbf{m}(\mathcal{U})$. For example, a bi-objective model seeks to maximize γ while minimizing $\mathbf{m}(\mathcal{U})$ as

$$\begin{array}{ll} \max\limits_{\gamma\in[0,1]} & \gamma \\ \min\limits_{\mathcal{U}\in\Omega} & \mathbf{m}(\mathcal{U}) \\ \text{s.t.} & \mathcal{E}^{\hat{\imath}}(\mathcal{U}) \geq \gamma, \end{array}$$

which is discussed in Example 1 and illustrated in Figure 4. Indeed, tracking the relationship between these values aids our explanatory ability and coincides with our algorithmic development in Section 7. Third, problem (6) is similar to an inverse problem because it seeks the least amount of data uncertainty to achieve a best possible efficiency score. Other inverse DEA models are found in Wei *et al.* (2000) and Zhang and Cui (2016).

DEA's dichotomy of classifying a DMU as either efficient or inefficient is altered in the uDEA problem to a DMU being either *capable* or *incapable*, as is now defined.

Definition 5. *We distinguish between three cases of a DMU under Ω:*

(i) DMU $\hat{\imath}$ is capable if $\gamma^* = \mathcal{E}^{\hat{\imath}}(\mathcal{U}) = 1$ for some $\mathcal{U} \in \Omega$.

(ii) DMU $\hat{\imath}$ is weakly incapable if $\gamma^* = 1$ but $\mathcal{E}^{\hat{\imath}}(\mathcal{U}) < 1$ for all $\mathcal{U} \in \Omega$.

(iii) DMU $\hat{\imath}$ is strongly incapable if $\gamma^* < 1$.

(iv) DMU $\hat{\imath}$ is incapable if it is either strongly or weakly incapable.

Notice that an optimal vector η^* for a capable DMU is $\eta^* = (e_{\hat{\imath}}, 1, 1)$, which follows because the $\hat{\imath}$-th DMU is its own target for the uDEA problem with an optimal $\mathcal{U}^*$. A DMU is incapable if it is inefficient for all uncertain data instances. An incapable DMU has no claim to efficiency unless it can argue for a change in Ω.

4. Configurations of Uncertainty

Both the analytical outcomes and the computational tractability of a uDEA problem rely on the type of uncertainty that is being considered and on how the amount of uncertainty is evaluated. Hence an analysis depends on the pair $(\Omega, \mathbf{m})$, which defines a *configuration.*

Definition 6. *A configuration of uncertainty, or more simply a configuration, is the pair $(\Omega, \mathbf{m})$, where Ω is a universe of possible collections of uncertainty satisfying $\mathcal{U}^o \trianglelefteq \mathcal{U}$ for all $\mathcal{U} \in \Omega$, and $\mathbf{m}$ is an amount of uncertainty.*

A configuration defines a uDEA problem by establishing the 'rules' upon which DMUs will be assessed. We mention that a configuration does not give rise to either a probability or an uncertainty space without additional assumptions, see Liu (2007), chapters 2 and 5 on probability theory and uncertainty theory, respectively.

A configuration's universe of uncertainty, along with its assessment of the amount of uncertainty, should coincide with the purpose of the specific uDEA problem, and for this reason, our development to this point has not imposed unnecessary restrictions that would have otherwise limited application within the DEA setting. However, robust problems are commonly motivated and solved with uncertainty sets of the form

$$\mathcal{U}_k = \{\bar{A}_k^{\hat{\imath}} + u^T R_k : \|u\|_p \leq 1\}, \tag{7}$$

where $\|\cdot\|_p$ is the p-norm and $\bar{A}_k^{\hat{\imath}}$ is the nominal vector of the k-th row's data.

If $p = 2$, then $\mathcal{U}_k$ in (7) is ellipsoidal. Ellipsoidal sets prevail in the literature, and they include polyhedral uncertainty (Ben-Tal and Nemirovski, 1999) and are regularly motivated stochastically (Ben-Tal *et al.*, 2009). Ellipsoids provide computational tractability since $\mathcal{E}^{\hat{\imath}}(\mathcal{U})$ is the optimal value of a second-order cone problem that can be efficiently solved. The modeling favorability and computational tractability of ellipsoids within the robust literature suggests that these desirable qualities will extend to the uDEA setting if we assume ellipsoidal uncertainty. Indeed, we review a process to model ellipsoidal uncertainty in Section 4.1, and we develop an algorithm to solve uDEA problems with ellipsoidal uncertainty in Section 7.

Uncertainty sets like those in (7) suggest an assessment of the amount of uncertainty such as

$$\mathbf{m}(\mathcal{U}) = \|\mathcal{U}\|_{p,q} := \Big\| \langle \|R_1\|_p, \ldots, \|R_{M+N}\|_p \rangle \Big\|_q,$$

which aligns with the standard $L_{p,q}$ notation associated with matrix norms. We illustrate some of these amounts of uncertainty in the forthcoming sections.

4.1. Modeling Configurations with Scenarios

While uncertainty sets like those in (7) are common, modeling the associated R_k matrices can be a hindrance. One such aid is to build the uncertainty sets by assuming stochastic data, see e.g. Ben-Tal *et al.* (2009). Here we briefly review a process that models an uncertainty set by considering random scenarios of the data. Similar developments are found in Nemirovski and Shapiro (2006) and Margellos *et al.* (2014). The result of this modeling process is a stochastic interpretation of an uncertain problem.

If we assume random data, then $A_k^{\hat{\imath}}\eta$ is a linear combination of the random inputs or outputs, where $A^{\hat{\imath}}$ is a random matrix of the form in (2). The previous notation of $\bar{A}^{\hat{\imath}}$ in (2) and (3), instead of say $A^{\hat{\imath}}$, purposefully suggests nominal data such as a mean, which is described momentarily. The stochastic theme is to replace the deterministic constraint $A_k^{\hat{\imath}}\eta \leq 0$ with the probabilistic constraint $P(A_k^{\hat{\imath}}\eta > 0) \leq \varepsilon$, where $0 \leq \varepsilon \leq 1$.

As with all stochastic modeling, what remains from a design perspective is to promote or assume distributional qualities of the random variables to express the probabilistic constraint. The most common assumption is $A_k^{\hat{\imath}}\eta \sim \mathcal{N}(\mu_k, STD_k^2)$, where the mean μ_k and the standard deviation STD_k are unknown and dependent on η. This condition is automatically satisfied if the input and output data are assumed to be independent normals, and while this stronger assumption is typical, it is questionable as a DEA premise - for instance, normality would dispute the assumption of nonnegative data. We maintain a more permissive stance by assuming that the input and output variables aggregate into an approximate normal, e.g. a truncated normal, an assumption that can, depending on the setting, gain support from the Central Limit Theorem.

Using the normality of $A_k^{\hat{\imath}}\eta$, we have that $(A_k^{\hat{\imath}}\eta - \mu^k)/STD_k$ is a standard normal, which means that

$$P(A_k^{\hat{\imath}}\eta > 0) \leq \varepsilon \quad \Leftrightarrow \quad \mu_k + STD_k\, \delta_{1-\varepsilon} \leq 0, \tag{8}$$

where $\delta_{1-\varepsilon}$ is the $1-\varepsilon$ percentile. The decision variables comprising η need to be selected so that the resulting mean μ_k and standard deviation STD_k satisfy the inequality on the right-hand side of (8).

We assume that each A_k is modeled as a collection of scenarios $(\hat{A}_k,\, p_k)$, in which each scenario is a row of the matrix $\hat{A}_k$, and the probability vector p_k defines the probabilities of the scenarios. The standard matrix expressions for the expected value and variance are then

$$\begin{aligned} \mu_k \quad &= \quad \mathrm{Exp}(A_k\eta) = \mathrm{Exp}(A_k)\eta = p_k^T \hat{A}_k \eta \;\text{ and} \\ STD_k^2 \quad &= \quad \mathrm{Var}(A_k\eta) = (\hat{A}_k\eta)^T (I - e\,p_k^T)^T P_k (I - e\,p_k^T)\hat{A}_k\eta, \end{aligned}$$

where P_k is the diagonal matrix of p_k. We set

$$R_k = \delta_{1-\varepsilon}\sqrt{P_k}(I - e\,p_k^T)\hat{A}_k \tag{9}$$

so that

$$STD_k = \sqrt{\mathrm{Var}(A_k \eta)} = \sqrt{\frac{\eta^T R_k^T R_k \eta}{\delta_{1-\varepsilon}^2}} = \frac{1}{\delta_{1-\varepsilon}} \|R_k \eta\|_2.$$

The probabilistic constraints in (8) are thus

$$p_k^T \hat{A}_k \eta + \|R_k \eta\|_2 = \bar{A}_k \eta + \|R_k \eta\|_2 \leq 0, \tag{10}$$

where $\bar{A}_k$ denotes the mean data of the scenarios. Hence, in this case the nominal data is the mean data.

The relationship between a set of uncertain constraints (4) and a second-order cone constraint with an ellipsoidal uncertainty set (10) is

$$\begin{aligned}
\bar{A}_k\, \eta + \|R_k\, \eta\|_2 \leq 0 \quad &\Leftrightarrow \quad \bar{A}_k^{\hat{\imath}}\, \eta + \max_u \{u^T R_k\, \eta : \|u\|_2 \leq 1\} \leq 0 \\
&\Leftrightarrow \quad A_k\, \eta \leq 0,\ \forall\, A_k \in \{\bar{A}_k^{\hat{\imath}} + u^T R_k : \|u\|_2 \leq 1\} \\
&\Leftrightarrow \quad A_k\, \eta \leq 0,\ \forall\, A_k \in \mathcal{U}_k, \qquad (11)
\end{aligned}$$

where the first statement follows from the definition of a matrix norm. As shown in (11), the inequalities in (10) are uncertain constraints with ellipsoidal uncertainty sets.

5. Examples

Consider the three DMUs pictured in Figure 1 and whose nominal data are listed in Table 1. DMU C is inefficient, and model (1) would scale C's input of 2 by the efficiency score of 1/2 to identify A as C's efficient target. The inefficiency of DMU C means that it has an interest in knowing if it is capable under a configuration of uncertainty. We divide the discussion into three examples with different configurations to help explore possible outcomes. This collection

- demonstrates a capable, weakly incapable, and strongly incapable DMU,
- shows that uncertainties among the DEA data can be linked, and
- motivates a numerical algorithm.

DMU	A	B	C
Output	1	3	1
Input	1	2	2

Table 1: Nominal data of a uDEA problem.

Example 1. *The uncertainty sets are of form* (7)*, where*

$$R_1 = \sigma_1 \begin{bmatrix} 0.2 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & -0.1 & 0 \end{bmatrix} \quad \textit{and} \quad R_2 = \sigma_2 \begin{bmatrix} 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & -0.1 \end{bmatrix}.$$

The multiples σ_1 and σ_2 scale the permissible uncertainty from the nominal data. We let the infinity-norm define the uncertainty sets so that

$$\begin{aligned} \mathcal{U}_1(\sigma_1) &= \{-(1,3,1,-1,0) - u^T R_1 : \|u\|_\infty \le 1\} \\ &= \{-(1+0.2\,\sigma_1 u_1,\ 3+0.2\,\sigma_1 u_2,\ 1+0.1\,\sigma_1 u_3,\ -1-0.1\,\sigma_1 u_3,\ 0) : \|u\|_\infty \le 1\}, \end{aligned}$$

and

$$\begin{aligned} \mathcal{U}_2(\sigma_2) &= \{(1,2,2,0,-2) + u^T R_2 : \|u\|_\infty \le 1\} \\ &= \{(1+0.1\,\sigma_2 u_1,\ 2+0.2\,\sigma_2 u_2,\ 2+0.1\,\sigma_2 u_3,\ 0,\ -2-0.1\,\sigma_2 u_3) : \|u\|_\infty \le 1\}. \end{aligned}$$

The collection of uncertainty for any nonnegative vector $\sigma = (\sigma_1, \sigma_2)$ is

$$\mathcal{U}(\sigma) = \{\mathcal{U}_1(\sigma_1), \mathcal{U}_2(\sigma_2)\}.$$

We further assume that the amount of uncertainty is

$$\mathbf{m}(\mathcal{U}(\sigma)) = \|\mathcal{U}(\sigma)\|_{\infty,2} = \sqrt{\|R_1\|_\infty^2 + \|R_2\|_\infty^2},$$

and that the configuration is

$$(\Omega, \mathbf{m}(\mathcal{U}(\sigma))) = (\{\mathcal{U}(\sigma) : 0 \le \sigma_1 \le 5, 0 \le \sigma_2 \le 10)\}, \mathbf{m}(\mathcal{U}(\sigma))).$$

Note that σ is bounded above to maintain nonnegative data instances.

Figure 2 illustrates how uncertainty with $\sigma_1 = \sigma_2 = 1$ alters the geometry that defines C's efficiency score. The use of the infinity-norm permits the components of u in the definitions of $\mathcal{U}_1$ and $\mathcal{U}_2$ to vary independently between 1 *and* -1*, and hence, the uDEA model includes every data collection in the shaded regions. This is a strong statement, as it means that C can independently select the input and output of all DMUs as long as the data remains in the shaded regions defined by σ.*

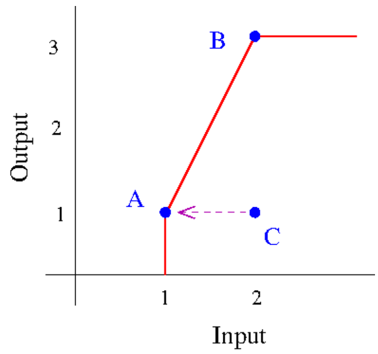

Figure 1: A simple DEA example with three DMUs.

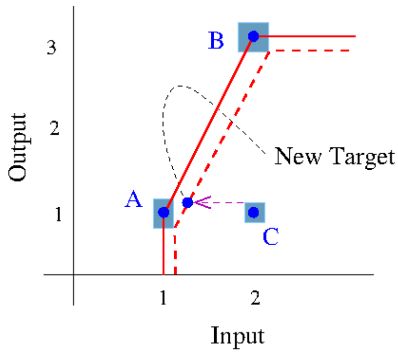

Figure 2: A simple DEA example with three DMUs and uncertainty.

Notice that the values of σ measure uncertainty per the elements of their respective R_k matrices. As an example, if σ_1 were to change from 1 to 2, then each of the shaded regions would double their vertical length, which would alter the efficient frontier and its proximity to the uncertain box around C.

From the geometry in Figure 2 it is clear that DMU C is capable under Ω; simply enlarge the shaded regions about the DMUs until the dashed line intersects the region about C. The efficiency score of C is defined by the upper-left corner of the region about C and the line through the lower-right corners of the regions about A and B, and after a few algebraic calculations, we have

$$\mathcal{E}^C(\mathcal{U}(\sigma)) = \min\left\{1,\ \frac{1 + 0.1\,\sigma_2 + 0.15\,\sigma_1\,(1 + 0.1\,\sigma_2)}{2 - 0.1\,\sigma_2}\right\}.$$

The minimum amount of uncertainty required for DMU C to have an efficiency score γ, with $0.5 \le \gamma \le 1$, is the solution to

$$\min_{\sigma}\left\{\sqrt{\|R_1\|_\infty^2 + \|R_2\|_\infty^2} : 1 + 0.1\,\sigma_2 + 0.15\,\sigma_1\,(1 + 0.1\,\sigma_2) = \gamma\,(2 - 0.1\,\sigma_2)\right\}.$$

For any efficiency score γ we can solve the constraint for σ_2 to reduce the problem to that of minimizing the objective over σ_1. The geometry of this optimization problem with $\gamma = 0.9$ is shown in Figure 3.

The trade-off between C's maximum efficiency score and the data's minimum amount of uncertainty that permits this score is the curve labeled "σ free" in Figure 4. Information about the best possible outcome for C is listed in the first row of Table 2. The minimum amount of uncertainty that renders the DMU *capable is* 0.72*. The optimal* σ *of* $(2.29, 2.80)$ *means that* C *needs to make the following claims under this configuration to refute its perceived inefficiency.*

1. *The outputs of DMUs* A *and* B *require a range of uncertainty of* $\pm 2.29 \times 0.2$*. So* A*'s output would have to be considered as an uncertain element in* $[0.54, 1.46]$ *and* B*'s as an uncertain element in* $[2.54, 3.46]$*. The output of* C *must have an uncertain range of* $\pm 2.29 \times 0.1$*, meaning that its output must be an uncertain element in* $[0.77, 1.23]$*.*
2. *The inputs of DMUs* A *and* C *must have a range of uncertainty of* $\pm 2.80 \times 0.1$*, and the range of uncertainty of* B*'s input must be* $\pm 2.80 \times 0.2$*. The resulting, and necessary, intervals of uncertainty for the inputs of* A*,* B*, and* C *are* $[0.44, 1.56]$*,* $[1.44, 2.56]$*, and* $[1.44, 2.56]$*.*

These ranges are permitted by Ω*, and hence,* C *can claim efficiency by constructing a data instance in which it is efficient.*

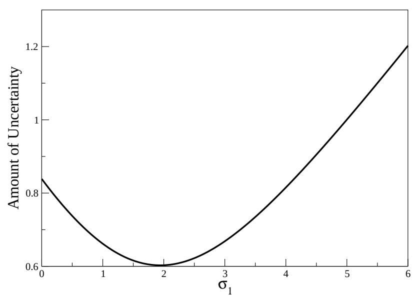

Figure 3: A graph of $\mathbf{m}(\mathcal{U})$ as a function of σ_1 with $\gamma = 0.9$.

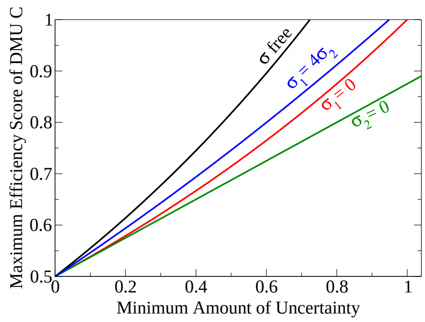

Figure 4: The minimum amount of uncertainty $\mathbf{m}(\mathcal{U}(\sigma))$ for any given maximum efficiency score γ.

Relationships among the constraints can be imposed to link the uncertainties between the inputs and outputs. Table 2 includes the results for three such cases. The first assumes $\sigma_1 = 0$ *to disallow uncertainty in the outputs. This situation might be reasonable if the DMUs are dependable even with uncertain inputs, say due to an established historical trust of achieving contractual requirements. The configuration replaces all instances of* σ_1 *with* 0*, and the regions of uncertainty in Figure 2 collapse vertically and become horizontal lines. The minimum amount*

σ restriction	Maximum Robust Efficiency Score	Minimum Amount of Uncertainty	Optimal σ
none	1	0.72	(2.29, 2.80)
$\sigma_1 = 0$	1	1.00	(0.00, 5.00)
$\sigma_2 = 0$	0.875	1.00	(5.00, 0.00)
$\sigma_1 = 4\sigma_2$	1	0.95	(4.60, 1.15)

Table 2: Example outcomes depending on an imposed relationship between the uncertainty of the inputs and outputs.

of uncertainty needed by C *to refute inefficiency increases to* 1 *with an optimal* σ *of* $(0.00, 5.00)$. *The curve labeled "*$\sigma_1 = 0$*" in Figure 4 shows this case.*

The last two cases are,

$\sigma_2 = 0$ **:** *input uncertainty is removed, which might occur if the DMUs introduce uncertainty with known inputs, and*

$\sigma_1 = 4\,\sigma_2$ **:** *input and output uncertainties are related so that output uncertainty is four times greater than input uncertainty, which means that the DMUs magnify uncertainty in the inputs.*

The curve in Figure 4 labeled "$\sigma_2 = 0$*" is for the case without input uncertainty, and the final curve is for the case with* $\sigma_1 = 4\,\sigma_2$. *DMU* C *is strongly incapable if* $\sigma_2 = 0$ *due to the necessity that* $\sigma_1 \leq 5$ *to maintain nonnegative data. If* $\sigma_1 = 4\sigma_2$, *then* C *is capable.*

Example 2. *As a second example, let* R_1 *and* R_2 *now be*

$$R_1 = \sigma_1\,[0.2,\ 0.2,\ 0.2,\ -0.2,\ 0] \text{ and } R_2 = \sigma_2\,[0.1,\ 0.1,\ 0.1,\ 0,\ -0.1],$$

with the uncertainty sets being

$$\begin{aligned}&\mathcal{U}_1(\sigma_1) = \{-(1,3,1,-1,0) - uR_1 : -1 \leq u \leq 1\}\\ &= \{-(1 + 0.2\,\sigma_1 u,\ 3 + 0.2\,\sigma_1 u,\ 1 + 0.2\,\sigma_1 u,\ -1 - 0.2\,\sigma_1 u,\ 0) : -1 \leq u \leq 1\}\end{aligned}$$

and

$$\begin{aligned}&\mathcal{U}_2(\sigma_2) = \{(1,2,2,0,-2) + uR_2 : 0 \leq u \leq 1\}\\ &= \{(1 + 0.1\,\sigma_2 u,\ 2 + 0.1\,\sigma_2 u,\ 2 + 0.1\,\sigma_2 u,\ 0,\ -2 - 0.1\,\sigma_2 u) : 0 \leq u \leq 1\}.\end{aligned}$$

Let $\Omega = \{\{\mathcal{U}_1(\sigma_1), \mathcal{U}_2(\sigma_2)\} : 0 \leq \sigma_1 \leq 5,\ 0 \leq \sigma_2\}$ *and* $\mathbf{m}(\mathcal{U}(\sigma)) = \|\mathcal{U}\|_{1,1} = \|R_1\|_1 + \|R_2\|_1$.

The interpretation of the uncertain regions surrounding the DMUs is altered from that of Example 1. First, $\mathcal{U}_2(\sigma_2)$ is directional since u is bounded below by 0, and hence, DMU C can increase, but not decrease, the inputs of all the DMUs as it seeks to improve its own efficiency score. Second, in the first example the values of u_1, u_2, and u_3 could be selected independently in $[-1, 1]$ for either of the uncertainty sets due to the infinity-norm. Moreover, each data element was adjusted by a unique component of u, and hence, the first example's configuration allowed the input and output data to be selected independently among the DMUs. Such liberal selection is not permitted here. For instance, the uncertainty set $\mathcal{U}_1(\sigma_1)$ in this example forces the outputs of all DMUs to increase or decrease by the same amount.

The efficiency score of DMU C for any σ in the second example is

$$\mathcal{E}^C(\mathcal{U}(\sigma)) = \frac{1 + 0.1\,\sigma_2}{2 + 0.1\,\sigma_2}.$$

The efficiency score in this case is independent of output uncertainty, as vertical shifts of the data leave C's efficiency score unchanged. Model (6) reduces to

$$\sup_{\gamma} \left\{ \gamma : \min_{\sigma \geq 0} \left\{ \|R_1\|_1 + \|R_2\|_1 : \frac{1 + 0.1\,\sigma_2}{2 + 0.1\,\sigma_2} \geq \gamma, \right\} \right\}.$$

We conclude that C can reach any robust efficiency score less than 1 because $\gamma \to 1$ as $\sigma_2 \to \infty$. Indeed, an optimal solution to the inner minimization for any fixed $\gamma \in [0.5, 1)$ is

$$\sigma = \left(0,\ \frac{2\gamma - 1}{0.1\,(1-\gamma)}\right) \quad \textit{and} \quad \mathbf{m}(\mathcal{U}(\sigma)) = \frac{2\,(2\gamma-1)}{(1-\gamma)}.$$

In this case DMU C is weakly incapable under Ω, as the supremum of γ is 1 but no collection of uncertainty renders DMU C efficient.

We note that if we had instead used

$$R_2 = \sigma_2 \left[\ 0.1, \quad 0.1, \quad 0.133, \quad 0, \quad -0.133\ \right],$$

then the supremum of γ would have been $0.1/0.133 \approx 0.75$. In this case DMU C would have been strongly incapable, and even an infinite amount of uncertainty would have left C inefficient.

Example 3 shows that even elementary uDEA problems can be difficult to compute, supporting the need for a numerical algorithm.

Example 3. *Let the R_k matrices for the uncertainty sets be*

$$R_1 = \sigma_1 \begin{bmatrix} 0.1 & 0.15 & 0 & -0.15 & 0 \\ 0.15 & 0.05 & 0 & -0.05 & 0 \end{bmatrix} \text{ and } R_2 = \sigma_2 \begin{bmatrix} 0.1 & 0 & -0.15 & 0 & 0.15 \\ -0.05 & 0 & 0.2 & 0 & -0.2 \end{bmatrix}.$$

Assume the uncertainty sets are the ellipsoids

$$\mathcal{U}_1(\sigma_1) = \{-(1,3,1,-1,0) - u^T R_1 : \|u\|_2 \leq 1\}$$

and

$$\mathcal{U}_2(\sigma_2) = \{(1,2,2,0,-2) + u^T R_2 : \|u\|_2 \leq 1\}.$$

Let $\mathcal{U}(\sigma) = \{\mathcal{U}_1(\sigma_1), \mathcal{U}_2(\sigma_2)\}$, $\Omega = \{\mathcal{U}(\sigma) : 0 \leq \sigma_1 \leq 6.667,\, 0 \leq \sigma_2 \leq 10\}$, and assume

$$\mathbf{m}(\mathcal{U}(\sigma)) = \|\mathcal{U}(\sigma)\|_{2,2} = \sqrt{\|R_1\|_2^2 + \|R_2\|_2^2}.$$

The third column of R_1 and the second column of R_2 being zero means that DMU C's output and DMU B's input are assumed certain. Otherwise, the relationship among the uncertain inputs and outputs is more nuanced than was the case in either of the first two examples. For instance, if $u_1 = 1$ in $\mathcal{U}_1(\sigma_1)$, then u_2 is forced to be 0, demonstrating the difference between our earlier use of the infinity-norm and the current use of the 2-norm. As such, the vectors in $\mathcal{U}_1$ and $\mathcal{U}_2$ are coupled through the selection of u, whose components are intertwined by the requirement that $\|u\|_2 \leq 1$.

Unlike the first two examples, an algebraic analysis of the third example is tedious, and we instead turn to a computational study. While we postpone algorithmic details until Section 7, the results are listed in Table 3, and the trade-off between the minimum amount of uncertainty and the maximum efficiency score is depicted in Figure 5.

The examples of this section only consider scalar versions of uncertainty in which the R_k matrices are scalar multiples of a set matrix. Nothing in the modeling framework requires such configurations. Another option would be to let the individual elements of an R_k matrix be variables themselves, possibly subject to constraints.

	$\mathcal{E}^C(\mathcal{U})$	$\mathbf{m}(\mathcal{U})$	σ
Nominal (certain)	0.5	0.0	(0, 0)
Efficient (uncertain)	1.0	1.45	(0.74, 2.70)

Table 3: What DMU C would need to assume of the data to claim efficiency.

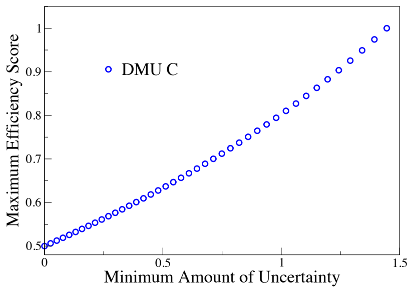

Figure 5: The trade-off between maximum γ and minimum $\mathbf{m}(\mathcal{U}) = \|\mathcal{U}\|_{2,2}$.

6. Traditional DEA as a Special Case of Uncertain DEA

A uDEA problem obviously reduces to its certain DEA progenitor if $\Omega = \{\mathcal{U}^o\}$, in which case the optimal solution satisfies

$$\mathcal{U}^* = \mathcal{U}^o, \;\; \gamma^* = \mathcal{E}^{\hat{\imath}}(\mathcal{U}^*) = E^{\hat{\imath}}, \;\; \text{and} \;\; \mathbf{m}(\mathcal{U}^*) = 0.$$

With $\Omega = \{\mathcal{U}^o\}$ the outer supremum over γ and the inner minimization over $\mathcal{U}$ are meaningless in the uncertain model (6), and the overhead of the uncertain paradigm is unwarranted with regard to solving the DEA problem. However, the traditional DEA model in (1) is essentially a parametric query that asks, how much do the inputs of the $\hat{\imath}$-th DMU need to scale (down) to reach the efficient frontier? Since the uncertainty sets of our previous examples mimic parametric scaling, a reasonable question is if model (6) can be used to solve a DEA problem outside the trivial restriction of $\Omega = \{\mathcal{U}^o\}$. In other words, can the outer supremum and inner minimization of model (6) be used to solve a traditional DEA problem? We answer this question in the affirmative in Theorem 1. The result follows by designing a configuration that holds outputs certain but that permits uncertainty in the inputs of the $\hat{\imath}$-th DMU.

Theorem 1. *Assume the configuration* $(\Omega, \mathbf{m})$*, where the uncertain collections in* Ω *are defined by the scalar* σ *so that* $0 \le \sigma \le 1$ *and* $\mathcal{U}(\sigma)$ *is comprised of*

$$\mathcal{U}_k(\sigma) = \begin{cases} \{\bar{A}_k^{\hat{\imath}}\}, & k = 1, 2, \ldots, M \\ \{\bar{A}_k^{\hat{\imath}} + \sigma\, u\, [x_k\, e_{\hat{\imath}}^T, 0, -x_k] : -1 \le u \le 1\}, & k = M+1, M+2, \ldots, M+N. \end{cases}$$

Assume the amount of uncertainty is $\mathbf{m}(\mathcal{U}(\sigma)) = \sigma$*. Then the* $\hat{\imath}$*-th DMU is capable, and the corresponding minimum amount of uncertainty is* $\sigma^* = 1 - E^{\hat{\imath}}$*, where* $E^{\hat{\imath}}$ *is the efficiency score in* (3)*.*

Proof. Assume the stated configuration $(\Omega, \mathbf{m})$. Then the robust and certain DEA models are

$$\begin{array}{lcl} \underline{\text{Robust DEA}} & & \underline{\text{Certain DEA}} \\ \begin{aligned} \mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) = \min\ & \gamma \\ \text{s.t. } Y\lambda &\ge y^{\hat{\imath}} \\ \left(X + \sigma\, u\, x^{\hat{\imath}}\, e_{\hat{\imath}}^T\right)\lambda &\le \gamma\,(1 + \sigma\, u)\, x^{\hat{\imath}}, \\ & \quad \forall\, u \in [-1, 1] \\ e^T\lambda &= 1 \\ \lambda &\ge 0 \end{aligned} & \text{and} & \begin{aligned} E^{\hat{\imath}} = \min\ & \theta^{\hat{\imath}} \\ \text{s.t. } Y\lambda &\ge y^{\hat{\imath}} \\ X\lambda &\le \theta^{\hat{\imath}}\, x^{\hat{\imath}} \\ e^T\lambda &= 1 \\ \lambda &\ge 0. \end{aligned} \end{array}$$

If the efficiency score of the certain DEA model is $E^{\hat{\imath}} = 1$, then $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(0)) = 1$, and the result is verified with $\gamma^* = 1$ and $\sigma^* = 0$.

Otherwise assume $E^{\hat{\imath}} < 1$. We first establish that $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) < 1$ for sufficiently small σ. First observe that $E^{\hat{\imath}} < 1$ guarantees a feasible $\hat{\lambda}$ to the certain DEA model such that $X\hat{\lambda} < x^{\hat{\imath}}$. Hence, for each of the input constraints we may select a positive σ_k so that

$$\max_{u \in [-1,1]} \sigma_k\, u\, x_k^{\hat{\imath}} \left(e_{\hat{\imath}}^T \hat{\lambda} - 1\right) < \left[x^{\hat{\imath}} - X\hat{\lambda}\right]_k.$$

So for any $\sigma \in [0, \min_k \sigma_k]$ we have,

$$\left(X + \sigma\, u\, x^{\hat{\imath}}\, e_{\hat{\imath}}^T\right)\hat{\lambda} < (1 + \sigma\, u)\, x^{\hat{\imath}},\ \forall\, u \in [-1, 1]. \tag{12}$$

Since $\hat{\lambda}$ inherits the other feasibility conditions of the robust DEA model from the certain DEA model, we know that $(\lambda, \gamma) = (\hat{\lambda}, 1)$ is a feasible solution to the robust DEA model for the selected σ. The strict inequality in (12) ensures that γ can be further reduced, and hence, $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) < 1$ so long as $\sigma \in [0, \min_k \sigma_k]$.

The configuration is designed to satisfy the following monotonicity property relative to the scalar σ,

$$\left.\begin{array}{rcl}\mathbf{m}(\mathcal{U}(\sigma')) = \sigma' \leq \sigma'' = \mathbf{m}(\mathcal{U}(\sigma'')) & \Leftrightarrow & \mathcal{U}(\sigma') \trianglelefteq \mathcal{U}(\sigma'') \\ & \Rightarrow & \mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma')) \leq \mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma'')),\end{array}\right\} \tag{13}$$

where the last implication follows from Proposition 1. This nondecreasing property, together with the fact that $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) < 1$ for $\sigma \in [0, \min_k \sigma_k]$, means that we need to establish that $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) \uparrow 1$ as σ increases past $\min_k \sigma_k$, which we now do.

Assume σ is such that $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) < 1$, and let $\Psi^{\hat{\imath}}(\sigma, u)$ be the optimal value of the linear program resulting from the robust DEA model with the selected σ and a fixed $u \in [-1, 1]$. Then,

$$\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) = \max_{u \in [-1,1]} \Psi^{\hat{\imath}}(\sigma, u) < 1. \tag{14}$$

The linear program of each $\Psi^{\hat{\imath}}(\sigma, u)$ is a DEA problem with $x^{\hat{\imath}}$ replaced by $(1 + \sigma\, u)\, x^{\hat{\imath}}$, and the inefficiency of (14) ensures that if (λ, γ) is an optimal solution to one of the linear programs that defines a $\Psi^{\hat{\imath}}(\sigma, u)$, then $\lambda_{\hat{\imath}} = 0$. We conclude that $e_{\hat{\imath}}^T \lambda = 0$ as long λ is an optimal solution to one of the linear programs that defines a $\Psi^{\hat{\imath}}(\sigma, u)$ as u varies in $[-1, 1]$. From (14) we know that the robust DEA model only needs to consider the optimal solutions of the linear programs that define $\Psi^{\hat{\imath}}(\sigma, u)$ for $u \in [-1, 1]$, and hence, the robust DEA model can be reduced to

$$\begin{array}{rcl}\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) = \min & \gamma & \\ \text{s.t. } Y\lambda & \geq & y^{\hat{\imath}} \\ X\lambda & \leq & \gamma\,(1 + \sigma\, u)\, x^{\hat{\imath}},\ \forall\, u \in [-1, 1] \\ e^T\lambda & = & 1 \\ \lambda & \geq & 0\end{array}$$

The most restrictive of

$$X\lambda \leq \gamma(1 + \hat{\sigma}\, u)\, x^{\hat{\imath}},\ \forall\, u \in [-1, 1]$$

occurs with $u = -1$, from which we know that the maximum of (14) is achieved at $u = -1$, i.e. $\Psi(\sigma, -1) = \mathcal{E}(\mathcal{U}(\sigma))$. Hence, if σ is such that $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) < 1$, then the robust DEA model can be further reduced to

$$
\begin{aligned}
\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) = \min\ & \gamma \\
\text{s.t. } Y\lambda &\geq y^{\hat{\imath}} \\
X\lambda &\leq \gamma\,(1-\sigma)\,x^{\hat{\imath}}, \\
e^T\lambda &= 1 \\
\lambda &\geq 0.
\end{aligned}
$$

We now have that $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma))$ is the optimal value of the certain DEA model defining $E^{\hat{\imath}}$ with θ replaced with $\gamma(1-\sigma)$. The smallest possible value of $\gamma(1-\sigma)$ is $E^{\hat{\imath}}$, making

$$\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\sigma)) = \frac{E^{\hat{\imath}}}{1-\sigma}, \tag{15}$$

provided that ratio on the right is less than 1. We conclude that

$$\gamma^* = \sup\left\{\frac{E^{\hat{\imath}}}{1-\sigma} : \sigma \in \left[0, 1-E^{\hat{\imath}}\right)\right\} = 1. \tag{16}$$

All that remains is to establish $\sigma^* = 1 - E^{\hat{\imath}}$, which is at least suggested by (16). From (15) we know that $\mathcal{E}(\mathcal{U}(\sigma)) < 1$ for all $\sigma \in [0, 1-E^{\hat{\imath}})$. Moreover, the robust DEA model has an optimal solution for all $\sigma \in [0,1]$, so $\mathcal{E}(\mathcal{U}(1-E^{\hat{\imath}}))$ exists and is between 0 and 1. The nondecreasing property in (13) further ensures that for any $\sigma \in [0, 1-E^{\hat{\imath}})$,

$$1 \geq \mathcal{E}(\mathcal{U}(1-E^{\hat{\imath}})) \geq \mathcal{E}(\mathcal{U}(\sigma)) = \frac{E^{\hat{\imath}}}{1-\sigma}.$$

Hence $\mathcal{E}(\mathcal{U}(1-E^{\hat{\imath}})) = 1$ because the right-hand side approaches 1 as $\sigma \to 1-E^{\hat{\imath}}$. The proof is complete since $\sigma = 1 - E^{\hat{\imath}}$ is the smallest value of σ achieving $\gamma^* = 1$, $\square$

We comment that the uncertainty sets of Theorem 1 were symmetric about the nominal data because $-1 \leq u \leq 1$. Nothing in the proof would have changed if we had instead used the asymmetric uncertainty sets with $-1 \leq u \leq 0$.

While Theorem 1 establishes that the uDEA paradigm subsumes traditional DEA as a modeling exercise, the computational overhead of solving a certain DEA model as a uDEA problem is difficult to support. The next section presents an algorithm to solve uDEA problems.

7. Solving Uncertain DEA Problems

Solving a uDEA problem is generally more difficult than is calculating the efficiency score of a DMU. Indeed, even if the configuration is designed to reasonably accommodate efficient calculations, computing γ^* necessitates the layering of three optimization problems, which complicates algorithm design. We restrict ourselves here to the case in which the robust DEA problem defining $\mathcal{E}^{\hat{\imath}}(\mathcal{U})$ can be efficiently solved as a second-order cone problem, i.e. we assume in our algorithmic development that

$$\mathcal{U}_k = \{\bar{A}_k^{\hat{\imath}} + u^T R_k : \|u\|_2 \leq 1\}.$$

This is the most common form of robust optimization.

The middle optimization problem seeks to minimize $\mathbf{m}(\mathcal{U})$, which we further restrict to $\mathbf{m}(\mathcal{U}) = \|\mathcal{U}\|_{2,2}$. Unfortunately, the constraint $\mathcal{E}^{\hat{\imath}}(\mathcal{U}) \geq \gamma$ is not generally convex as demonstrated by the subgraph of the function illustrated in Figure 5, and hence, the middle optimization problem introduces a loss of convexity. So even in a simplified case in which uncertainty is scaled and the efficiency score results from a quick solve of a convex problem, the uDEA model can lack convexity and challenge standard solution procedures.

We promote a first-order algorithm that relies on the sole requirement of an efficient robust DEA solver. We assume the uncertain collection $\mathcal{U}$ depends on a list of model parameters, which are arranged into the vector ψ. For example, if uncertainty is scaled as in our previous examples, then ψ would be the σ vector already used in much of our discussion. However, ψ could instead be the elements of the R_k matrices themselves. Independent of any particular ψ, the overriding goal of a uDEA problem is to calculate parameters ψ so that a DMU's efficiency score is as large as possible and so that the amount of uncertainty is as small as possible to achieve the best efficiency score.

A forward difference approximation of $\partial \mathcal{E}^{\hat{\imath}}(\psi)/\partial \psi_j$ is

$$\frac{\partial \mathcal{E}^{\hat{\imath}}}{\partial \psi_j}(\psi) \approx \frac{\mathcal{E}^{\hat{\imath}}(\psi + \delta\, e_j) - \mathcal{E}^{\hat{\imath}}(\psi)}{\delta} = h_j,$$

where δ is a reasonable perturbation for ψ_j. The existence of the partial derivative on the left is not generally guaranteed, as it depends on the configuration and its parameter vector ψ. However, independent of the actual existence of the partial derivative, the approximating finite difference can be calculated, and the vector h is an (assumed) approximate direction of steepest ascent of the robust efficiency score over the parameter space, i.e. $h \approx \nabla \mathcal{E}^{\hat{\imath}}(\psi)$ should the partial derivatives exist. The calculation of h requires a robust efficiency score for each j, making the computational burden increase with the number of parameters.

We follow h a step length of α, making the newly updated parameters $\psi + \alpha\, h$. We note that if $\mathcal{U}(\psi) \trianglelefteq \mathcal{U}(\psi + \alpha\, h)$, then the updated parameters are guaranteed to give at least the robust efficiency score of the previous parameters. For scaled uncertainty like that of the third example in Section 5, we have for any nonnegative step-size α that

$$h \geq 0 \;\Rightarrow\; \mathcal{U}(\sigma) \trianglelefteq \mathcal{U}(\sigma + \alpha\, h) \;\Rightarrow \mathcal{E}^{\hat{\imath}}(\sigma) \leq \mathcal{E}^{\hat{\imath}}(\sigma + \alpha\, h).$$

Hence, verifying the nonnegativity of h certifies that there is no loss in the robust efficiency score.

We favor small, incremental steps along h so that only minor gains in efficiency are accumulated per iteration. The rationale for small steps is that we should be able to reasonably approximate the minimum amount of uncertainty for the new gains in the robust efficiency score with a second, linear approximation. In this case we can not straightforwardly follow (an approximation of) $\nabla \mathbf{m}(\mathcal{U}(\psi + \alpha\, h))$, as doing so would decrease uncertainty without considering the gains in efficiency. Instead, we want to compute a direction d along which the efficiency is maintained but for which the directional derivative $d^T \nabla \mathbf{m}(\mathcal{U}(\psi + \alpha\, h))$ is as small as possible. Assuming $\nabla \mathcal{E}^{\hat{\imath}}$ exists, we have that these goals are satisfied by calculating

$$d \in \operatorname{argmin} \left\{ \rho^T \nabla \mathbf{m}(\mathcal{U}(\psi + \alpha\, h)) : \rho^T \nabla \mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi + \alpha\, h)) = 0,\ \|\rho\| \leq 1 \right\}.$$

Either of the gradients can again be approximated with, e.g., forward differences, but if the uncertainty sets arise from scaled versions of fixed matrices, then $\mathbf{m}(\mathcal{U}(\psi))$ can be tacitly replaced with its square to get a simple sum-of-squares for which $\nabla \mathbf{m}(\mathcal{U}(\psi + \alpha\, h))$ can be calculated directly. Once the gradients are calculated or estimated, the minimization problem identifying d has a linear objective with a single linear constraint and a single convex-quadratic constraint. The problem can be solved routinely with standard solvers.

We use a line search along d to establish a step size. The search calculates the largest β less than 1 so that

$$\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi + \alpha\, h + \beta\, d)) \geq \mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi + \alpha\, h)) - \epsilon,$$

where ϵ is some permittedly small loss in the robust efficiency score. Several line searches have been tested, including second order methods. All line searches had computational nuances that hindered their general use except for the method of bisection, which was trustworthy throughout.

Pseudocode for the numerical procedure is listed in Algorithm 1. As with all numerical methods, the algorithm relies on a set of convergence criteria and tolerances that determine the performance. The algorithm's practical ability is demonstrated in the next section.

8. A Case Study in Radiotherapy

External radiation therapy is one of the major cancer treatments along with surgery and chemotherapy, and about two thirds of all cancer patients undergo a course of radiotherapy. Radiotherapy exploits a therapeutic advantage in which cancerous cells are unable to recover as well as healthy cells from radiation damage. Moreover, radiotherapy has the advantage of delivering near conformal dose distributions to tumors with complex geometries. While radiotherapy is generally regarded as a targeted, local therapy, it is not possible to irradiate only the tumor. Therefore, the challenge in treatment planning is to achieve a high dose of radiation to the tumor while sparing surrounding organs. We refer to Bortfeld (2006) for further medical-physical details.

Cumulative dose-volume histograms (DVHs) serve in clinical practice to assess planned quality and to approximate the portion of a structure's volume that will receive a certain portion of the prescribed dose. Moreover, DVH constraints can serve as control points during the planning process (see e.g. Cambria *et al.* (2014); Dogan *et al.* (2009)), with their recommended values being determined by protocols such as ICRU-83 (2010).

Treatments are planned by iteratively adjusting delivery parameters to best adhere to contradictory goals. This process is heavily dependent on a planner's skill and experience (Nelms *et al.*, 2012). Moreover, quality assurance is often

Algorithm 1 A First-Order Algorithm for Uncertain DEA

for $\hat{\imath} = 1, 2, \ldots, D$ **do**

Calculate the noinal efficiency score, $E^{\hat{\imath}}$, of the $\hat{\imath}$-th DMU.

if $E^{\hat{\imath}} == 1$ **then**

DMU $\hat{\imath}$ is capable under Ω

else

searchFlag = True

while searchFlag **do**

Set ψ so that $\mathbf{m}(\mathcal{U}(\psi)) = 0$

Calculate h with

$$h_j = \frac{\mathcal{E}^{\hat{\imath}}(\psi + \delta\, e_j) - \mathcal{E}^{\hat{\imath}}(\psi)}{\delta}$$

Set $\psi = \psi + \alpha h$.

Recalculate h for the updated ψ.

Calculate (or approximate) $\nabla \mathbf{m}(\mathcal{U}(\psi))$

Compute direction d that solves

$$\min\{\rho^T \nabla \mathbf{m}(\mathcal{U}(\psi)) : \rho^T h = 0,\ \|\rho\| \leq 1\}$$

Use the method of bisection to search for the largest β less than 1 satisfying $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi + \beta d)) \geq \mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi)) - \varepsilon$

if $|\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi+\beta d)) - \mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi))| <$ convTol or $\mathbf{m}(\mathcal{U}(\psi+\beta d)) >$ maxUncrty **then**

searchFlag = False

else

$\psi = \psi + \beta d$.

end if

end while

if $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi + \beta d)) == 1$ and $\mathbf{m}(\mathcal{U}(\psi + \beta d)) \leq$ maxUncrty **then**

The $\hat{\imath}$-th DMU is capable under Ω.

else if $\mathcal{E}^{\hat{\imath}}(\mathcal{U}(\psi + \beta d)) == 1$ and $\mathbf{m}(\mathcal{U}(\psi + \beta d)) >$ maxUncrty **then**

The $\hat{\imath}$-th DMU is declared weakly incapable under Ω

else

The $\hat{\imath}$-th DMU is declared incapable under Ω

end if

end if

end for

done by visual inspection of DVHs, leading to judgmental interpretation and uncertain outcomes. For example, the authors of Das *et al.* (2008) observed high variability among planners and institutions, reporting that the median dose to the tumor can vary by $\pm 10\%$ of the prescribed dose across 96% of the patient population.

We consider 42 anonymized prostate IMRT treatments from Auckland Radiation Oncology. All treatments were approved for observational study based on the guidelines of the New Zealand Health and Disability Ethics Committee, see Lin *et al.* (2013). These treatments were planned with the same system[1], followed the same clinical criteria, and were delivered to patients. However, uncertainty about treatment quality remained at the end of the planning process even with these commonalities due to patient and design variations. After all, each anatomy and cancer is unique, and treatment planners tailor each treatment to an individual patient based on their personal skills. While each treatment was deemed acceptable for the specific patient at the end of the planning process, further improvement might have been possible with continued exploration of the planning parameters.

Further uncertainties caused by, for example, patient misalignment at the time of delivery affect treatment quality, causing discrepancies between the planned and delivered anatomical doses. As such, the delivered treatment is one realization amongst the uncertain possibilities that could have been delivered with the identical, but already uncertain, planned treatment. The combined uncertainties of treatment planning and delivery complicate assessment, and we show that uDEA can aid the planning process by classifying planned treatments. The classification is based on each treatment's minimum amount of uncertainty to become efficient against a competitive cohort of similar treatments. If a planned treatment seems acceptable on standard evaluative metrics and is efficient with only a small amount of uncertainty, then we gain confidence in the treatment's efficacy. If a planned treatment is either incapable or efficient with only an excessive amount of uncertainty, then the treatment should probably be re-planned.

[1]Pinnacle v9 and the SmartArc module by Philips, Netherlands

The prescribed dose to the tumor volume was 74 Gy. The protocol required that 95% of this prescribed dose be received by 99% of the tumor volume and that 99% of this dose be received by 99% of the actual prostate. The criteria for the organs at risk were that the fraction of the rectum volume that received at least 40, 60, and 70 Gy should not exceed 60%, 40%, and 10%, respectively.

Lin *et al.* (2013) used the certain DEA model in (3) to compute efficiency scores for these treatments. The data comprised of one point-wise and one averaging evaluative metric for each treatment. The point-wise D_{95} value measures the dose to 95% of the tumor volume and serves as the output of the DEA model. The averaging quantity of generalized equivalent uniform dose (gEUD) measures homogeneity of the dose delivered to the rectum and is employed as the DEA input, see (Niemierko, 1997). The input and output data were extracted using CERR (Deasy *et al.*, 2003). Lin *et al.* (2013) identified efficient and inefficient treatments based on these nominal data and suggested that the most inefficient treatments could have been re-planned with improved efficiency, which was empirically verified for several treatments. Further clinical information and the actual data are available in Lin *et al.* (2013).

The uncertainty sets for this clinical study were

$$\mathcal{U}_1 = \{\bar{A}_1^{\hat{\imath}} + \sigma_1 u^T I : \|u\|_2 \leq 1\} \text{ and } \mathcal{U}_2 = \{\bar{A}_2^{\hat{\imath}} + \sigma_1 u^T I : \|u\|_2 \leq 1\},$$

where I was the 42×42 identity. The resulting universe of possible uncertain collections was

$$\begin{aligned} \Omega &= \{\{\mathcal{U}_1(\sigma_1), \mathcal{U}_2(\sigma_2)\} : 0 \leq \sigma_1 \leq 70.875,\, 0 \leq \sigma_2 \leq 59.567\} \\ &= \{\mathcal{U}(\sigma) : 0 \leq \sigma \leq (70.875, 59.567)\}, \end{aligned}$$

where the bounds on σ_1 and σ_2 were calculated to ensure the nonnegativity of the data. The amount of uncertainty was

$$\mathbf{m}(\mathcal{U}(\sigma)) = \|\mathcal{U}(\sigma)\|_{2,2} = \sqrt{\|\sigma_1 I\|_2^2 + \|\sigma_2 I\|_2^2} = \|\sigma\|.$$

This amount of uncertainty has the same physical unit of Gray (Gy) as σ_1 and σ_2. The upper bounds on σ for this application are mathematically necessary but clinically egregious, as adjusting a treatment by 50+ Gy would radically violate the intent of treatment. These bounds could be reduced by a clinician

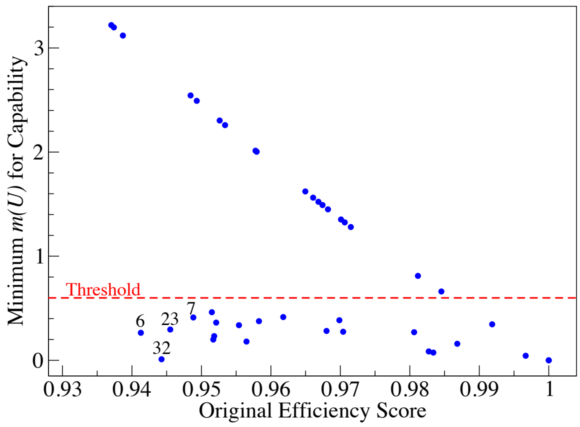

Figure 6: The solution to the uDEA problem for the IMRT data. Each symbol represents a treatment. Cases 6, 7, 23, and 32 have some of the lowest nominal efficiency scores but become efficient with only small amounts of uncertainty. The dashed line illustrates how a guide might be used as a possible threshold of acceptable uncertainty.

to better define capability. However, all treatments were capable and became efficient with less than 3.3 Gy of uncertainty.

Algorithm 1 solved each of the 42 uDEA problems associated with the treatments. The parameter vector was $\psi = \langle \sigma_1, \sigma_2 \rangle$, and the other settings were $\delta = 0.1$, $\alpha = 0.4$, and $\varepsilon = 10^{-8}$. The search direction d was calculated with the objective coefficients

$$\langle \sigma_1 \|R_1\|_2,\ \sigma_2 \|R_2\|_2 \rangle = \left(\sqrt{42}/2\right)\ \nabla \mathbf{m}^2(\mathcal{U}(\psi)).$$

The algorithm terminated once the capability of a treatment was established to within ε, i.e. $\gamma^* \geq 1 - 10^{-8}$. The code was written in Matlab, and all cases solved within a few seconds with Gurobi as the underlying solver.

Figure 6 displays the minimum amount of uncertainty required for each treatment to become efficient. The horizontal dashed line in Figure 6 illustrates how a guiding threshold could help distinguish treatments that might benefit from re-planning versus those that wouldn't. Treatments needing only small amounts of uncertainty to become efficient, i.e. those below the threshold, could be considered acceptable. Treatments needing larger amounts of uncertainty to become efficient, i.e. those above the threshold, could be re-planned. Quantifying the threshold isn't perfectly rigorous, and deciding this value would be the responsibility of the clinicians and would depend on institutional schools-of-thought.

Treatments like 6, 7, 23, and 32 deserve note. These four treatments had some of the lowest nominal efficiency scores and would have been questionable in a traditional DEA analysis. However, they only require a small amount of uncertainty to become efficient, and hence, they could have been deemed efficient within clinical discretion. Juxtapose these four treatments against those on the upper left-hand side of the graph, which also had low nominal efficiency scores but required higher amounts of uncertainty to reach efficiency. The combination of an originally low efficiency score together with a high amount of required uncertainty strongly suggests re-planning.

We end our case study with a final reminder that our analysis depends on the configuration. For instance, Figure 6 would have likely changed if the configuration had been based on a different norm or had the uncertainties been linked. The selection of a configuration is part of the uDEA modeling process, and all analysis is relative to the configuration. As with many OR modeling paradigms, the configuration should be selected to provide a meaningful analysis and computational tractability.

9. Conclusion

We investigated how DEA is affected by uncertain data. We first presented a robust DEA model that defines a robust efficiency score for known uncertainty sets. We then formally showed that an increase in the uncertainty harbored by a collection of uncertainty increases the efficiency score of a DMU. This led to the question of how much uncertainty is needed to classify a DMU as efficient. We introduced the definition of an *amount of uncertainty*, which allowed us to formulate an optimization problem that answers this question. We then discussed configurations of uncertainty from a stochastic perspective. After illustrating our concepts with simple examples, we proved that traditional DEA is a special case of uDEA for a particular configuration of uncertainty. We also provided a first-order algorithm to solve the uDEA model with ellipsoidal uncertainty sets. Finally, we presented a case study in radiotherapy to validate the relevance of uDEA in some practical applications.

We have not addressed in any detail the possible situation that configurations of uncertainty in the uDEA model depend on the DMU under assessment, which will be investigated in future work. Other questions for future research are whether and how a stochastic interpretation of uDEA opens a route to approaching the problem via simulation. The relationship between uDEA and parametric analysis will also lead to further questions.

Acknowledgment

The authors thank Emma Stubington for her thoughtful conversations on several of the topics herein. They also thank three anonymous referees for their thorough and suggestive reviews. O. N. acknowledges the support of the NSF grant CMMI-1463489.

References

Arabmaldar, A., Jablonsky, J., and Saljooghi, F. H. (2017). A new robust dea model and super-efficiency measure. *Optimization*, **66**(5), 723–736.

Ben-Tal, A. and Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, **25**(1), 1–13.

Ben-Tal, A. and Nemirovski, A. (2002). Robust optimization – methodology and applications. *Mathematical Programming*, **92**(3), 453–480.

Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009). *Robust Optimization*. Princeton University Press, Princeton.

Bertsimas, D. and Thiele, A. (2006). A robust optimization approach to inventory theory. *Operations Research*, **54**(1), 150–168.

Bertsimas, D., Nohadani, O., and Teo, K. M. (2010). Nonconvex robust optimization for problems with constraints. *INFORMS Journal on Computing*, **22**(1), 44–58.

Bertsimas, D., Brown, D. B., and Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM Review*, **53**(3), 464–501.

Bortfeld, T. (2006). Review IMRT: a review and preview. *Physics in Medicine and Biology*, **51**(13), 363–379.

Cambria, R., Cattani, F., Jereczek-Fossa, B., Pansini, F., Ciardo, D., Vigorito, S., Russo, S., Zerini, D., Cozzi, L., and Orecchia, R. (2014). Planning study to compare dynamic and rapid arc techniques for postprostatectomy radiotherapy of prostate cancer. *Strahlentherapie und Onkologie*, **190**(6), 569–574.

Chue, M., Zinchenko, Y., Henderson, S., and Sharpe, M. (2005). Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Physics in Medicine and Biology*, **50**, 5463–5477.

Cooper, W. W., Seiford, L. M., and Tone, K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-solver Software*. Springer, New York.

Das, I. J., Cheng, C.-W., Chopra, K. L., Mitra, R. K., Srivastava, S. P., and Glatstein, E. (2008). Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems. *Journal of the National Cancer Institute*, **100**(5), 300–307.

Deasy, J. O., Blanco, A. I., and Clark, V. H. (2003). CERR: A computational environment for radiotherapy research. *Medical Physics*, **30**(5), 979–985.

Dogan, N., Mihaylov, I., Wu, Y., Keall, P. J., Siebers, J. V., and Hagan, M. P. (2009). Monte Carlo dose verification of prostate patients treated with simultaneous integrated boost intensity modulated radiation therapy. *Radiation Oncology*, **4**, 18.

Emrouznejad, A., Parker, B. R., and Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. *Socio-economic Planning Sciences*, **42**, 151–157.

Hwang, S.-N., Lee, H.-S., and Zhu, J., editors (2016). *Handbook of Decision Making Using Data Envelopment Analysis*. Springer, New York.

ICRU-83 (2010). Report: Planning aims, prescription, and technical data. *J. I.C.R.U.*, **10**(1), 55–59.

Lin, K.-M., Simpson, J., Sasso, G., Raith, A., and Ehrgott, M. (2013). Quality assessment for VMAT prostate radiotherapy planning based on data envelopment analysis. *Physics in Medicine and Biology*, **58**(16), 5753.

Liu, B. (2007). Uncertainty theory. In *Uncertainty Theory*, pages 205–234. Springer.

Liu, J. S., Lu, L. Y. Y., Lu, W.-M., and Lin, B. J. Y. (2013). A survey of DEA applications. *Omega*, **41**, 893–902.

Margellos, K., Goulart, P., and Lygeros, J. (2014). On the road between robust optimization and the scenario approach for chance constrained optimization problems. *IEEE Transactions on Automatic Control*, **59**(8), 2258–2263.

Nelms, B. E., Robinson, G., Markham, J., Velasco, K., Boyd, S., Narayan, S., Wheeler, J., and Sobczak, M. L. (2012). Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Practical Radiation Oncology*, **2**(4), 296–305.

Nemirovski, A. and Shapiro, A. (2006). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, **17**(4), 969–996.

Niemierko, A. (1997). Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Medical Physics*, **24**(1), 103–110.

Olesen, O. B. and Petersen, N. C. (2016). Sochasttic data envelopment analysis. *European Journal of Operational Research*, **251**(1), 2–21.

Wei, Q., Zhang, J., and Zhang, X. (2000). An inverse dea model for inputs/outputs estimate. *European Journal of Operational Research*, **121**, 151–162.

Zhang, M. and Cui, J.-C. (2016). The extension and integration of the inverse DEA method. *Journal of the Operational Research Society*, **67**(9), 1212–1220.

Zhu, J. (2014). *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets*. SpringerSpringer, New York, 3rd edition.