Decision Making in Manufacturing and Services
Vol. 1 • 2007 • No. 1–2 • pp. 111–136

The Art and Science of Modeling Decision-Making Under Severe Uncertainty

Moshe Sniedovich*

Abstract. For obvious reasons, models for decision-making under severe uncertainty are austere. Simply put, there is precious little to work with under these conditions. This fact highlights the great importance of utilizing in such cases the ingredients of the mathematical model to the fullest extent, which in turn brings under the spotlight the art of mathematical modeling. In this discussion we examine some of the subtle considerations that are called for in the mathematical modeling of decision-making under severe uncertainty in general, and *worst-case analysis* in particular. As a case study we discuss the lessons learnt on this front from the *Info-Gap* experience.

Keywords: mathematical modeling, severe uncertainty, maximin, worst-case analysis, robust optimization, info-gap.

Mathematics Subject Classification: 90B50, 90C47, 91B06.

Received/Revised: 04 April 2007/30 May 2007

1. INTRODUCTION

Methodologies designed for *decision-making under severe uncertainty* are austere in the extreme because one has precious little to work with in such cases. This is vividly manifested in the simple structure of the mathematical models deployed by these methodologies.

And so given the simplicity of such models, it is not surprising that these methodologies are discussed in introductory OR/MS textbooks, eg. Markland and Sweigart (1987), Winston (1994), Ragsdale (2004), Hillier and Lieberman (2005). But from a practical point of view this means that users of such models are called upon to be imaginative and inventive if they are to make the most of them.

As we shall see, there is another side to this coin. A failure of the imagination here can lead to a failure to recognize these methodologies when they are disguised by notation and terminology.

The aim of this paper is to illustrate that the *mathematical modeling* of decision-making under severe uncertainty requires considerable subtlety. For this purpose we use *Info-Gap* (Ben-Haim, 2006) as a case study.

* Department of Mathematics and Statistics, The University of Melbourne, Australia. E-mail: moshe@unimelb.edu.au.

In the next section we briefly discuss the notion *decision problem* and then describe the two classical principles for dealing with decision-making under severe uncertainty, namely *Laplace's Principle of Insufficient Reason* and *Wald's Maximin Principle*. This is followed by a formal description of the generic *Info-Gap* model. We then discuss the mathematical modeling aspects of *Info-Gap*, showing that the generic *Info-Gap* model is an instance of *Wald's Maximin Model*. Thereafter, we analyze the implications of the local nature of the worst-case analysis deployed by *Info-Gap* and conclude that this feature of *Info-Gap* makes it thoroughly unsuitable for decision-making under severe uncertainty. The point is that *Info-Gap* does not tackle severe uncertainty – it simply takes no notice of it. A simple portfolio investment problem is subsequently presented to illustrate the points discussed in the preceding sections.

2. CLASSICAL DECISION PROBLEMS

The conventional format used in classical decision theory to describe *decision problems* is that of a *decision table* (French, 1988, Grüning and Kühn, 2005). The rows of such a table represent *actions* or *decisions* made by the *decision maker* and the column represent *external factors*. The basic assumption is that these external factors are beyond the decision maker's control.

The content of each cell of the decision table is construed as a *reward* or *payoff* to the decision maker corresponding to her decision (row) and the true value of the external factors (column).

Needless to say, the decision maker aims to maximize her reward but this might not be an easy task because she may have only limited knowledge of the true value of the external factors.

Example

Suppose that one fine morning you find the following note and four envelopes on your doorstep.

> Good morning Sir/Madam:
>
> I left on your doorstep four envelopes. Each contains a sum of money. You are welcome to open any one of these envelopes and keep the money you find there.
>
> Please note that once you open an envelope, the other three will automatically self-destruct, so think carefully about which of these envelopes you should open.
>
> To help you decide what you should do, I printed on each envelope the possible value of the amount (in Australian dollars) that you may find in it. The actual value is equal to one of these figures.
>
> Unfortunately the entire project is subject to severe uncertainty so I cannot tell you more than this.

Good luck!

Joe.

For your convenience Table 1 depicts the information Joe provided on the four envelopes.

Table 1. Easy Problem

Envelope	*Possible Amounts (Australian dollars)*
$E1$	$20, 10, 300, 786$
$E2$	$2, 4000000, 102349, 500000000, 99999999, 56435432$
$E3$	$201, 202$
$E4$	200

So what would you do Dear Sir/Madam? Which envelope would you open? And what methodology would you use to solve this problem?

One of the central issues in decision theory is the representation and quantification of the knowledge that the decision maker has about the true value of the external factors. This may vary from full knowledge to complete ignorance.

As far as terminology goes, the tradition has been to assume that the external factors are embodied in an entity called *state*, or *state of nature*, and that the true value of the state is determined by the omnipresent *Mother Nature*.

Classical decision theory (French, 1988) distinguishes between three classes of decision problems associated with the behavior of *Mother Nature*:

— Decision-making under *certainty*.
— Decision-making under *risk*.
— Decision-making under *strict uncertainty*.

Decision-making under *certainty* represents situations in which the decision maker has complete knowledge of the state selected by *Mother Nature* in response to the decision selected by the decision maker.

In decision-making under *risk* it is assumed that *Mother Nature* selects the state via a *probability distribution* and that the decision maker has full knowledge of this distribution.

The situation in decision-making under *strict uncertainty* is thoroughly different: the decision maker is totally ignorant as to how *Mother Nature* selects the state, except for knowing that the state belongs to a given set, the so called *State Space*.

Our discussion deals exclusively with decision-making under *strict uncertainty*. We shall use the terms *severe* and *strict* uncertainty interchangeably.

Now, for the purposes of our discussion it is convenient to replace the conventional *decision table* format with a simple generic model of the form $\mathfrak{M} = (\mathbb{D}, \mathbb{S}, f)$ where $\mathbb{D}$ and $\mathbb{S}$ are sets and f is a real-valued function on $\mathbb{D} \times \mathbb{S}$.

In the parlance of classical decision theory, $\mathbb{D}$ represents the *decision space*, $\mathbb{S}$ represents the *state space* and f represents the *objective function*. The conceptual framework behind this model is as follows:

— The decision maker selects a *decision* $d \in \mathbb{D}$.
— Given this decision, *Mother Nature* selects a *state* $s \in \mathbb{S}$.
— A *reward* (payoff) equal to $f(d, s)$ is awarded to the decision maker.

It is assumed that the decision maker's goal is to obtain the largest possible reward and that this guides her decision-making analysis.

The difficulty is that the reward $f(d, s)$ depends not only on the decision made by the decision maker, namely d, but also on the state selected by *Mother Nature*, namely s. This means that to make sense in this situation, the decision maker must incorporate in her decision analysis considerations pertaining to how *Mother Nature* will react to her decision.

But how can the decision maker bring this off, given that she has no inkling as to how *Mother Nature* selects her state?

3. DECISION-MAKING UNDER SEVERE UNCERTAINTY

It it clear right from the outset that fully satisfactory solutions to problems of severe uncertainty can hardly be expected. After all, the situation here is simply too amorphous for a rigorous, formal, "objective" treatment. Witness the austere structure of the model $\mathfrak{M} = (\mathbb{D}, \mathbb{S}, f)$ representing the situation: It does not provide the slightest clue as to how *Mother Nature* behaves.

In short, as far as modeling is concerned the situation is desperate, hence desperate measures are required to make our way out of it.

And to be sure, in view of the classification outlined in the preceding section, classical decision theory offers two such measures: one transforms the severe uncertainty into *risk* and one transforms it into *certainty.* These no doubt are desperate measures indeed.

Over the years these two approaches have become highly popular and famous – some would say infamous – and have gained the status of *Principles* or decision-making *Rules.*

3.1. LAPLACE'S PRINCIPLE OF INSUFFICIENT REASON (1825)

This principle argues by *symmetry*: if there is no reason to believe that any state is more/less likely than others, then assume that *all states are equally likely.* Practically speaking this means that the *state* can be regarded as a *uniformly distributed random variable.*

If we accept this assumption then the decision-making environment changes from *severe uncertainty* to *risk.* The advantage of this approach is that it transforms a difficult problem into a relatively simple one. We move from the wilderness of severe uncertainty into the kingdom of risk where statistics and probability theory reign supreme.

3.2. WALD'S MAXIMIN PRINCIPLE (1945)

This principle is far more radical (Wald 1945, 1950). It argues that if we are to play it safe, we should assume that *Mother Nature* plays "against us". In other words, the assumption is that *Mother Nature* always selects the worst state relative to the decision we make.

In the framework of the model $\mathfrak{M} = (\mathbb{D}, \mathbb{S}, f)$, this means that *Mother Nature* selects the state according to the following policy:

$$\hat{s}(d) := \arg \min_{s \in \mathbb{S}} f(d,s), \quad d \in \mathbb{D} \tag{1}$$

and therefore the reward for the decision maker generated by decision $d \in \mathbb{D}$ is as follows:

$$\begin{aligned} v(d) &:= f(d, \hat{s}(d)), \quad d \in \mathbb{D} &(2)\\ &= \min_{s \in \mathbb{S}} f(d,s) &(3) \end{aligned}$$

In the parlance of classical decision theory, $v(d)$ is the *security level* of decision d. No matter what state will actually be observed, the reward generated by decision d will be at least as large as $v(d)$.

The idea is then to select a decision $d \in \mathbb{D}$ whose security level is the largest. So the recipe is as follows:

$$\begin{aligned} v^* &:= \max_{d \in \mathbb{D}} v(d) &(4)\\ &= \max_{d \in \mathbb{D}} \min_{s \in \mathbb{S}} f(d,s) &(5) \end{aligned}$$

This principle has the attraction of transforming a decision-making problem under *severe uncertainty* into a decision-making problem under *certainty*. It is almost too good to be true!

Table 2 summarizes the results obtained by applying these two famous principles to our little 4-envelope problem in the preceding section.

Table 2. Results

Envelope	*Possible Amounts*	*Wald*	*Laplace*
$E1$	20, 10, 300, 786	10	279
$E2$	2, 4000000, 10234	2	1336745.3333 ✓
$E3$	201, 202	201 ✓	201.5
$E4$	200	200	200

Each envelope is evaluated in accordance with the two recipes. The *Wald* column selects the smallest entry in the *Possible Amounts* column, whereas the *Laplace* column computes the arithmetic average of the entries in the *Possible Amounts* column. For instance, consider the first envelope, $E1$. The super pessimistic Wald assumes that the worst value will materialize, hence the smallest item in the list $20, 10, 300, 786$, which is 10, is selected.

On the other hand, Laplace assumes that the amount in $E1$ is a uniformly distributed random variable on this very list, hence the *expected value* of the reward is equal to the arithmetic mean of the elements on the list: $\frac{1}{4}(20+10+300+786) = 279$.

In short, if you follow *Wald* you'll open the third envelope, $E3$, and if you follow *Laplace* you'll open the second envelope, $E2$.

What would you do, dear reader?

3.3. MODELING ISSUES

Details concerning obvious, and not so obvious, difficulties with these two principles can be found in French (1988).

For our purposes it suffices to note that *Wald's Maximin Model* of uncertainty is extremely conservative. It definitely does not provide a faithful representation of how we operate in reality. It may lead to exceedingly costly solutions resulting from over-protection against uncertainty.

A major difficulty with *Laplace's Principle of Insufficient Reason* is that the state space must be constructed so as to be amenable to a *uniform* probability distribution. For example, the principle cannot be applied when $\mathbb{S} = \mathbb{R}$, where $\mathbb{R}$ denotes the real line.

In our discussion we focus on the *mathematical modeling* aspects of these two principles, especially *Wald's Maximin Principle*. The rationale for this is not only the prominent role that these principles play in decision theory and the wealth of knowledge at our disposal on all aspects of these two celebrities. It is also very important to determine whether a proposed *new decision theory* is in fact one of these principles in *disguise*.

4. INFO-GAP

Info-Gap (Ben-Haim 2001, 2006) presents itself as a new theory that is radically different from all existing theories for decision-making under severe uncertainty:

> Info-gap decision theory is radically different from all current theories of decision under uncertainty. The difference originates in the modelling of uncertainty as an information gap rather than as a probability. The need for info-gap modelling and management of uncertainty arises in dealing with severe lack of information and highly unstructured uncertainty.
>
> Ben-Haim (2006, p. xii)

> In this book we concentrate on the fairly new concept of information-gap uncertainty, whose differences from more classical approaches to uncertainty are real and deep. Despite the power of classical decision theories, in many areas such as engineering, economics, management, medicine and public policy, a

need has arisen for a different format for decisions based on severely uncertain evidence.

Ben-Haim (2006, p. 11)

With these declarations in mind, we shall examine the modeling aspects of *Info-Gap* in the context of the following three fundamental questions:

Q1 Is the generic *Info-Gap* model *new*?
Q2 Is it radically *different* from the classical models of decision theory?
Q3 How well does it tackle *severe uncertainty*?

The generic *Info-Gap* model that is most relevant to such an examination consists of the following objects. We deliberately employ the standard *Info-Gap* notation and terminology (Ben-Haim 2001, 2006) where $\mathbb{R}_+$ denotes the non-negative part of $\mathbb{R}$:

— An *uncertainty region* (set), $\mathfrak{U}$.
— A *parameter* u whose true value, u°, is unknown except that $u^\circ \in \mathfrak{U}$.
— An *estimate*, $\tilde{u} \in \mathfrak{U}$, of u°.
— A parametric family of nested *regions of uncertainty*, $\mathcal{U}(\alpha, \tilde{u}) \subseteq \mathfrak{U}, \alpha \geq 0$, of varying size ($\alpha$), centered at $\tilde{u}$. It is assumed that $\mathcal{U}(0, \tilde{u}) = \{\tilde{u}\}$ and that $\mathcal{U}(\alpha, \tilde{u})$ is non-decreasing with α, namely

$$\alpha'', \alpha' \in \mathbb{R}_+,\ \alpha'' > \alpha' \Longrightarrow \mathcal{U}(\alpha', \tilde{u}) \subseteq \mathcal{U}(\alpha'', \tilde{u}) \tag{6}$$

— Set of *decisions* available to the decision maker, $\mathbb{Q}$.
— A real-valued *reward function*, R, on $\mathbb{Q} \times \mathfrak{U}$.
— A *critical reward* level, $r_c \in \mathbb{R}$.

The *decision problem* associated with this model is to determine the best decision $q \in \mathbb{Q}$ given the severe uncertainty in u and the requirement $r_c \leq R(q, u)$.

For this purpose *Info-Gap* deploys *robustness* to rank decisions, where the robustness of decision $q \in \mathbb{Q}$ is defined as follows:

$$\hat{\alpha}(q, r_c) := \max \left\{ \alpha \geq 0 : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} \tag{7}$$

That is, the robustness of a decision q is the largest value of α such that the performance requirement $r_c \leq R(q, u)$ is satisfied *for all* $u \in \mathcal{U}(\alpha, \tilde{u})$.

For simplicity it is assumed that $r_c \leq R(q, \tilde{u}), \forall q \in \mathbb{Q}$, observing that this implies that for each $q \in \mathbb{Q}$ and $\alpha \geq 0$ there is at least one $u \in \mathcal{U}(\alpha, \tilde{u})$ such that $r_c \leq R(q, \tilde{u})$. If this condition is not satisfied for a given decision $q \in \mathbb{Q}$, then this decision can be discarded at the outset.

So the decision problem posed by the *Info-Gap* model is as follows:

$$\hat{\alpha}(r_c) : = \max_{q \in \mathbb{Q}} \ \hat{\alpha}(q, r_c) \tag{8}$$

$$= \max_{q \in \mathbb{Q}} \ \max \left\{ \alpha \geq 0 : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} \tag{9}$$

In short, the recipe for an optimal decision is as follows:

$$\hat{q}(r_c) := \arg \max_{q\in\mathbb{Q}} \ \max\left\{\alpha \geq 0 : r_c \leq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right\} \tag{10}$$

observing that there could be more than one optimal decision.

The direct reference to the critical reward r_c in the notation $\hat{\alpha}(r_c)$ and $\hat{\alpha}(q, r_c)$ is an indication that the performance requirement $r_c \leq R(q, u)$ is "soft" rather than "hard". The idea is then to use *Pareto optimization* tools (Steuer, 1985) to generate the *efficient frontier* of $(\hat{\alpha}(r_c), r_c)$ pairs.

5. THE ART OF MATHEMATICAL MODELING

The most intriguing claim made in the *Info-Gap* books (Ben-Haim, 2001, 2006) is that this theory is new and radically different from all existing theories of decision-making under uncertainty. No less intriguing is the fact that these books make no mention of, let alone discuss, *Wald's Maximin Principle* and *worst-case analysis*.

Apparently the explanation for this is the view (Ben-Haim, 2005, p. 392, p. 401) that there is no worst case in an *Info-Gap* model of uncertainty and therefore *Info-Gap* is not *Maximin*.

The objective of this section is to show that by bringing mathematical modeling into play, it is possible to express the generic *Info-Gap* model (9) as a run of the mill instance of *Wald's Maximin Model* (5). As we shall see, this instance is characterized by an interesting objective function.

So, first note that in the context of *Wald's Maximin Model* it is often convenient to let the set of states available to *Mother Nature* depend on the decision selected by the decision maker. In this case, the *Maximin* model takes the following form:

$$v^* := \max_{d\in\mathbb{D}} \ \min_{s\in S(d)} \ f(d,s) \tag{11}$$

where for each $d \in \mathbb{D}$, set $S(d) \subseteq \mathbb{S}$ represents the set of decisions available to *Mother Nature* given that the decision maker selected decision d.

For instance, such a model is deployed in the analysis of the famous counterfeit coin problem (Sniedovich, 2003).

It should be noted that from a purely modeling point of view this modification is a mere technicality. By slightly modifying the objective function f we can rewrite (11) as (5), namely we can let $S(d) = \mathbb{S}, \forall d \in \mathbb{D}$ and modify f accordingly.

Next, let $\preceq$ denote the binary operation defined by:

$$a \preceq b := \begin{cases} 1, & a \leq b \\ 0, & a > b \end{cases} \quad a, b \in \mathbb{R} \tag{12}$$

and consider the real-valued function φ defined on $\mathbb{Q} \times \mathbb{R}_+ \times \mathfrak{U}$ as follows:

$$\varphi(q, \alpha, u) := \alpha \cdot (r_c \preceq R(q,u)), \quad q \in \mathbb{Q}, \alpha \geq 0, u \in \mathcal{U}(\alpha, \tilde{u}) \tag{13}$$

where $\cdot$ denotes scalar multiplication.

Then clearly, by construction $\varphi(q,\alpha,u)$ is non-decreasing with $R(q,u)$ and therefore

$$\begin{aligned}
\beta(r_c) :&= \max_{q\in\mathbb{Q},\alpha\geq 0}\ \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \varphi(q,\alpha,u) && (14)\\
&= \max_{q\in\mathbb{Q},\alpha\geq 0}\ \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha\cdot(r_c \preceq R(q,u)) && (15)\\
&= \max_{q\in\mathbb{Q},\alpha\geq 0} \alpha\cdot\left(r_c \preceq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right) && (16)\\
&= \max_{q\in\mathbb{Q}}\ \max_{\alpha\geq 0} \alpha\cdot\left(r_c \preceq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right) && (17)\\
&= \max_{q\in\mathbb{Q}} \max\left\{\alpha : r_c \leq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right\} && (18)\\
&= \hat{\alpha}(r_c) && (19)
\end{aligned}$$

In other words, utilizing φ as the objective function of the *Maximin* model we can represent the generic *Info-Gap* model compactly as follows:

$$\hat{\alpha}(r_c) := \max_{q\in\mathbb{Q},\alpha\geq 0}\ \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha\cdot(r_c \preceq R(q,u)) \tag{20}$$

Here then are side-by-side the generic classical *Maximin* model and its instance representing the generic *Info-Gap* model.

Wald's Maximin Model	*Generic Info-Gap Model*
$v^* := \max_{d\in\mathbb{D}}\ \min_{s\in S(d)} f(d,s)$	$\hat{\alpha}(r_c) = \max_{q\in\mathbb{Q},\alpha\geq 0}\ \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha\cdot(r_c \preceq R(q,u))$

So what are we to make of this?

Contrary to *Info-Gap*'s claim that it is a new theory that is radically different from all current theories for decision-making under severe uncertainty, from a *Maximin* view point *Info-Gap* is clearly one of its numerous specific instances.

To make this point crystal clear, we now repeat the above derivation in the other direction, that is we start with a standard *Maximin* model. Formally, the result is as follows:

Theorem 1. *The generic Info-Gap model is an instance of Wald's Maximin Model.*

Proof. Consider the following specific ingredients of the *Maximin* model $\mathfrak{M}^* = (\mathbb{D}^*, S^*, f^*)$, where

$$\begin{aligned}
\mathbb{D}^* &:= \mathbb{Q}\times R_+ && (21)\\
S^*(q,\alpha) &:= \mathcal{U}(\alpha,\tilde{u}), \quad q\in\mathbb{Q}, \alpha\geq 0 && (22)\\
f^*(q,\alpha,u) &:= \alpha\cdot(r_c \preceq R(q,u)) && (23)
\end{aligned}$$

The *Maximin* model associated with these specific objects is then as follows:

$$v^* := \max_{d\in\mathbb{D}^*} \min_{s\in S^*(d)} f^*(d,s) \tag{24}$$

$$= \max_{q\in\mathbb{Q},\alpha\geq 0} \min_{u\in\mathcal{U}(\alpha,\tilde{u})} f^*(q,\alpha,u) \tag{25}$$

$$= \max_{q\in\mathbb{Q},\alpha\geq 0} \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha\cdot(r_c \preceq R(q,u)) \tag{26}$$

$$= \max_{q\in\mathbb{Q},\alpha\geq 0} \alpha\cdot\left(r_c \preceq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right) \tag{27}$$

$$= \max_{q\in\mathbb{Q}} \max_{\alpha\geq 0} \alpha\cdot\left(r_c \preceq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right) \tag{28}$$

$$= \max_{q\in\mathbb{Q}} \max\left\{\alpha : r_c \leq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right\} \tag{29}$$

Clearly, this is none other than the generic *Info-Gap* model stipulated in (9). □

For the reader's convenience, Table 3 displays in full detail the elements of the *Info-Gap* model and their *Maximin* counter-parts.

Table 3. Correspondence between the generic *Info-Gap* and Maximin models

Wald's Maximin Principle	*Generic Info-Gap Model*
d	(q,α)
s	u
$\mathbb{D}$	$\mathbb{Q}\times\mathbb{R}_+$
$S(d)$	$\mathcal{U}(\alpha,\tilde{u})$
$f(d,s)$	$\alpha\cdot(r_c \preceq R(q,u))$

In words, *Info-Gap*'s generic model is an instance of *Wald's Maximin Principle* characterized by a number of particular features, the most important one being the structure of the objective function, namely $f(q,\alpha,u) = \alpha\cdot\big(r_c \preceq R(q,u)\big)$.

This formulation highlights the conflict between the decision maker, who attempts to maximize the value of α, and *Mother Nature* who attempts to minimize the value α by minimizing $R(q,u)$ within the region of uncertainty $\mathcal{U}(\alpha,\tilde{u})$ stipulated by α.

It is a typical *worst-case analysis*: each decision is evaluated by the worst outcome associated with it: *Mother Nature* selects a u in $\mathcal{U}(\alpha,\tilde{u})$ that minimizes $f(q,\alpha,u)$ over $\mathcal{U}(\alpha,\tilde{u})$. In this framework, the values of q and α are fixed, so minimizing $f(q,\alpha,u) = \alpha\cdot(r_c \preceq R(q,u))$ over $u\in\mathcal{U}(\alpha,\tilde{u})$ amounts to minimizing $R(q,u)$ over $u\in\mathcal{U}(\alpha,\tilde{u})$.

To see more clearly what is going on here, consider a given $q \in \mathbb{Q}$ and its robustness:

$$\hat{\alpha}(q, r_c) : = \max_{\alpha \geq 0} \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \cdot (r_c \preceq R(q, u)) \tag{30}$$

$$= \max_{\alpha \geq 0} \alpha \cdot \min_{u \in \mathcal{U}(\alpha, \tilde{u})} (r_c \preceq R(q, u)) \tag{31}$$

$$= \max_{\alpha \geq 0} G(\alpha) \cdot H(q, \alpha) \tag{32}$$

where:

$$G(\alpha) := \alpha, \ \ \alpha \geq 0 \tag{33}$$

$$H(q, \alpha) := \min_{u \in \mathcal{U}(\alpha, \tilde{u})} (r_c \preceq R(q, u)), \quad q \in \mathbb{Q}, \alpha \geq 0 \tag{34}$$

observing that the nesting property of the regions of uncertainty, namely (6), implies that for a given q, $H(q, \alpha)$ is a step function of α, as shown in Figure 1(a).

This implies that for a given q,

$$\beta(q, \alpha) := G(\alpha) \cdot H(q, \alpha) \tag{35}$$

consists of two linear parts: on the interval $[0, \hat{\alpha}(q, r_c)]$ this function is equal to G; and then on the interval $(\hat{\alpha}(q, r_c), \infty)$ the function is equal to 0, as shown in Figure 1(b).

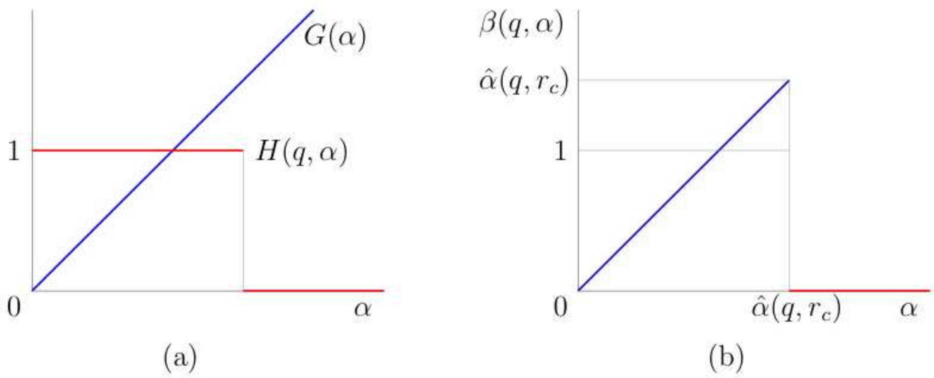

Fig. 1. $G = G(\alpha)$, $H = H(q, \alpha)$, and $\beta(q, \alpha)$ (q is fixed).

More details on the relationship between *Info-Gap* and *Maximin* and other related issues can be found in Sniedovich (2006).

In summary then, the answers to the first two questions that we raised above regarding the place and role of *Info-Gap* in decision theory are as follows:

A1 Not only is it the case that the generic *Info-Gap* model (9) is *not new*, it is a simple instance of none other than the most famous model in decision-making under severe uncertainty, namely *Wald's Maximin model*.

A2 For the very same reason, the generic *Info-Gap* model is *not radically different* from classical models for decision-making under severe uncertainty.

To formulate an answer to the third question raised above, we need to examine a peculiar feature of the generic *Info-Gap* model and its ramifications. This feature has to do with the fact that (9) is *completely oblivious* to the "size" of the total region of uncertainty $\mathfrak{U}$ in relation to the "size", $\hat{\alpha}(r_c)$, of the optimal region of uncertainty $\mathcal{U}(\hat{\alpha}(r_c), \tilde{u})$. More precisely,

Theorem 2. *The generic Info-Gap model is invariant with the total region of uncertainty* $\mathfrak{U}$*: the value of* $\hat{\alpha}(r_c)$ *does not vary with* $\mathfrak{U}$ *for all* $\mathfrak{U}$ *such that* $\mathcal{U}(\hat{\alpha}(r_c)+\varepsilon, \tilde{u}) \subseteq \mathfrak{U}$ *for some* $\varepsilon > 0$.

Proof. Let $\alpha^* := \hat{\alpha}(r_c)$ and $\mathfrak{U}^* := \mathcal{U}(\alpha^* + \varepsilon, \tilde{u})$. We have to show that α^* does not vary with $\mathfrak{U}$ for all $\mathfrak{U}$ such that $\mathfrak{U}^* \subseteq \mathfrak{U}$.

This follows immediately from the nesting property of the regions of uncertainty $\mathcal{U}(\alpha, \tilde{u}), \alpha \geq 0$ stipulated in (6) and the worst-case characteristic of robustness stipulated in the definition (7) of $\hat{\alpha}(q, \tilde{u})$. □

So the answer to the third question regarding the place and role of *Info-Gap* in decision-making under uncertainty is as follows:

A3 The generic *Info-Gap* model does not deal with the severe uncertainty aspect of the decision problem. It simply ignores it.

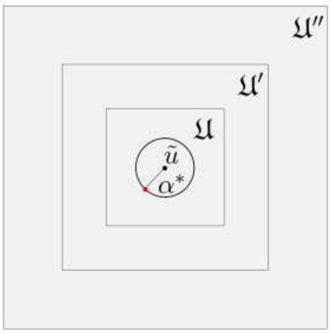

Fig. 2. Illustration of Theorem 2

This point is illustrated in Figure 2 where three regions of uncertainty are displayed, $\mathfrak{U} \subset \mathfrak{U}' \subset \mathfrak{U}''$. The same solution, α^*, is obtained for any region of uncertainty containing the set $\mathcal{U}(\alpha^* + \varepsilon, \tilde{u})$ represented by the circle.

This serious flaw in *Info-Gap*'s model of uncertainty calls for a closer examination of the mathematical structure of *Info-Gap*'s generic model (9).

6. LOCAL VS GLOBAL WORST-CASE ANALYSIS

The concept *Worst Case* plays a central role in many areas of decision theory such as decision-making under uncertainty, robust optimization, numerical complexity analysis, design of algorithms and so on.

One of the most important issues in *worst-case analysis* is the formulation of the region of uncertainty from which the worst case is selected. There are two conflicting considerations in play here:

— The need to represent the region of uncertainty as fully as possible so as to avoid precluding adverse states of nature that are relevant to the investigation.
— The need to exclude from the analysis overly-pessimistic states that will make the analysis as a whole too conservative. Hence the advice,

> If the forecaster tries to specify too many discrete forecasts, in an attempt to cover most possibilities, discrete minimax may yield too pessimistic strategies or even run into numerical, or computational, problems due to the resulting numerous scenarios. Similarly, as the upper and lower bounds on a range of forecasts get wider, to provide coverage to a wider set of possibilities, the minimax strategy may become pessimistic. Thus, scenarios have to be chosen with care, among genuinely likely values. The minimax strategy will then answer the legitimate question of what the best strategy should be, in view of the worst case.
>
> Rustem and Howe (2002, p. xiii)

One thing that the literature on worst-case analysis and robust optimization under severe uncertainty *does not* bother to take up explicitly is the danger in using *a single point estimate* to represent the region of uncertainty. On this one can only speculate but it would seem that this "omission" reflects the tacit understanding that such an idea is so alien to the basic dilemma in decision-making under severe uncertainty that it would not even be contemplated.

Be it as it may, the basic issue is this: whatever estimate we have of the true value of the state variable, under severe uncertainty this estimate is of a very poor quality and is likely to be substantially wrong. Therefore, it makes no sense to conduct the worst-case analysis *only* on the immediate region surrounding this estimate, as this region does not provide a good representation of the entire region of uncertainty.

This issue is described schematically in Figure 3. A poor estimate of the true value of the state variable (s') and its surrounding region (S') is shown inside a large region of uncertainty.

For illustrative purposes the true value of the state variable (s°) and the worst state in S' are also shown.

With this in mind, let us now re-examine the way *Info-Gap* defines the robustness of a decision q, namely:

$$\hat{\alpha}(q, r_c) := \max \left\{ \alpha \geq 0 : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} \tag{36}$$

$$= \max_{\alpha \geq 0} \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \cdot (r_c \preceq R(u, u)) \tag{37}$$

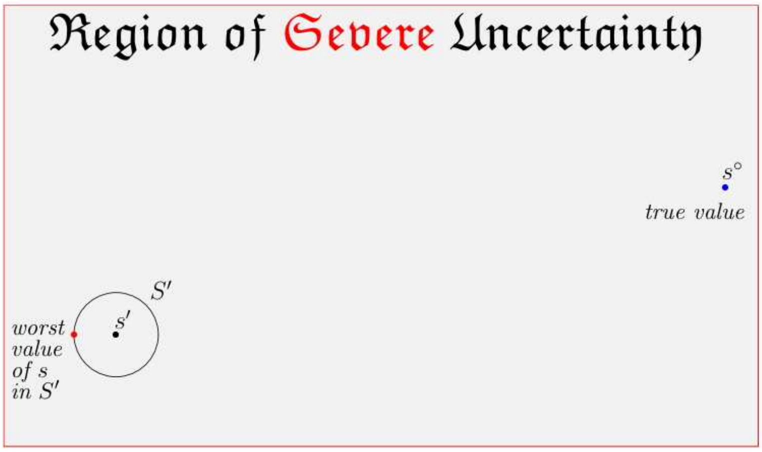

Fig. 3. Local worst-case analysis

Note that for a given (q, α) pair, the worst value of u is selected from the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$ surrounding the estimate $\tilde{u}$. This means that the worst-case analysis conducted by *Info-Gap* is *local* in nature. In fact, for a given q it covers only the region $\mathcal{U}(\hat{\alpha}(q, r_c), \tilde{u})$. This means that the largest region of uncertainty considered in the analysis is $\mathcal{U}(\hat{\alpha}(r_c), \tilde{u})$, recalling that $\hat{\alpha}(r_c) = \max_{q \in \mathbb{Q}} \hat{\alpha}(q, r_c)$.

This is shown schematically in Figure 4. The conclusion is therefore that, under severe uncertainty, there is no reason to believe that the solutions generated by *Info-Gap* are likely to be robust.

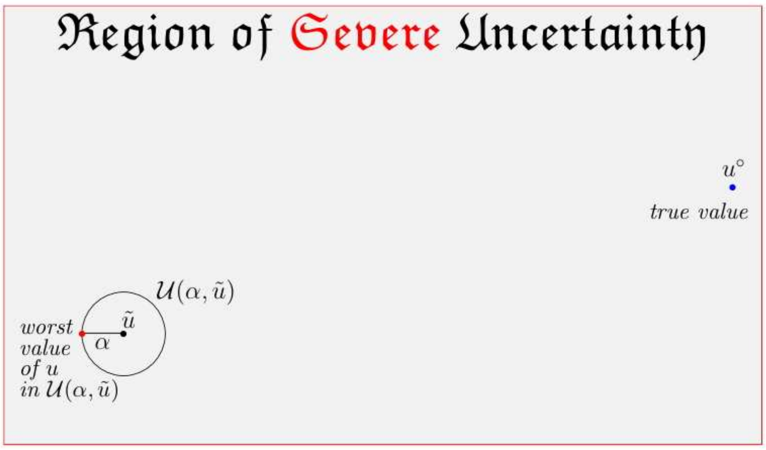

Fig. 4. Worst-case analysis à la *Info-Gap*.

For the purposes of our analysis it suffices to indicate that *Info-Gap*'s local approach to robustness is not only simplistic, but flawed. It definitely stands in sharp contrast to the much more global approach to robustness deployed in the *Robust Optimization* literature (Ben-Tal et al, 2006; Rustem and Howe, 2002; Vladimirou and Zenios, 1997; Kouvelis and Yu, 1997).

From a mathematical modeling point of view it is imperative to distinguish between local worst-case analysis of the type deployed by *Info-Gap*, and global worst-case analysis of the type formulated in *Robust Optimization*. Certainly, the notation and terminology used here should make this distinction explicit, especially for unsuspecting readers/users.

Thus, in the case of *Info-Gap*, the notation for robustness should incorporate the estimate $\tilde{u}$ so that instead of $\hat{\alpha}(r_c)$ we would write $\hat{\alpha}(r_c|\tilde{u})$ and instead of $\hat{\alpha}(q, r_c)$ we would write $\hat{\alpha}(q, r_c|\tilde{u})$.

Along the same lines, $\hat{\alpha}(q, r_c)$ should be called something like "robustness of q in the neighborhood of $\tilde{u}$, rather than just "robustness of q".

In this regard it is interesting to note that in the worst-case analysis and robust optimization literature one often finds warnings about the conservative/pessimistic nature of worst-case analysis (eg Rustem and Howe, 2002). Yet, no such issues are discussed in the *Info-Gap* literature, even though, as we have shown, *Info-Gap* is definitely a worst-case oriented methodology par excellence, albeit of a local nature.

It should be stressed that the flaw in the *Info-Gap* uncertainty model does not lie in its employment of *Wald's Maximin Principle*. Rather it lies in the use of a single point estimate and its neighborhood as an approximation of the entire region of uncertainty. Indeed, the *Principle* is used extensively in *Robust Optimization* to (properly) generate robust solutions for decision-making situations under severe uncertainty.

In summary, the local nature of the *worst-case analysis* conducted by *Info-Gap* brings to light two related but distinct points that undermine it:

— Optimality of the solution generated by the *Info-Gap* model.
— Robustness of the solution generated by the *Info-Gap* model.

Next, let $q(\tilde{u})$ denote the optimal solution generated by the *Info-Gap* model for a given value of the estimate $\tilde{u}$. Since under severe uncertainty $\tilde{u}$ is a poor estimate of the true value of u, we should be concerned about changes in $q(\tilde{u})$ resulting from changes in the value of $\tilde{u}$. For the same reason we need to be concerned about the changes in the robustness $\hat{\alpha}(q(\tilde{u}), r_c)$ as we change the value of $\tilde{u}$.

Of course, *sensitivity analysis* of this kind is also used in decision-making under risk and even in decision-making under certainty. And to be sure, such an analysis can be useful, beneficial and informative, and could serve many purposes. But the point is that in decision-making under SEVERE uncertainty these matters are crucial. Indeed, if we accept the notion that under severe uncertainty it is acceptable to base the analysis on the immediate neighborhood of some estimate $\tilde{u}$, then the distinction between decision-making under risk and decision-making under severe uncertainty is completely wiped out.

In many respects the distinction made here between local and global robustness and *worst-case analysis* is similar to the distinction made in optimization theory between local and global *optimum*. However, in optimization theory the distinction is crystal clear whereas in the context of *Info-Gap* it is not.

The difficulty is not in the fact that *an analysis* is conducted on a neighborhood of the estimate we have. The difficulty arises because the results of this analysis are construed and presented as though they were based on a methodology that takes into account the full scope of the severe uncertainty.

And to sum it up, from a modeling point of view *Info-Gap* does not take on the severe uncertainty: it simply takes no notice of it. This involves two things:

— Replacing severe uncertainty by a *poor point estimate* of the parameter of interest.
— Conducting a conventional worst-case analysis a la maximin in the *immediate neighborhood* of this *poor estimate*.

In the next section we illustrate the methodological issues discussed above in the framework of a simple portfolio investment problem.

7. ILLUSTRATIVE EXAMPLE

Consider a simplified version of the portfolio investment problem discussed in Ben-Haim (2006, pp. 70-71) where the reward function is as follows:

$$R(q,u) = \sum_{i=1}^{N} q_i u_i = q^T u \tag{38}$$

Here q_i denotes the sum invested in security i and u_i denotes the (unknown) future value of one unit of security i. The budget available for investment is Q, so the decision space is

$$\mathbb{Q} = \left\{ q \in \mathbb{R}^N_+ : \sum_{i=1}^{N} q_i = Q \right\} \tag{39}$$

To apply *Info-Gap* we have to express the region of uncertainty, $\mathfrak{U}$, in a particular way. Specifically we have to

— Determine a *nominal* value, call it $\tilde{u}$, to serve as an estimate of the true value of the parameter u, call it u°.
— Express the region of uncertainty $\mathfrak{U}$ as the limit of a parametric family $\mathcal{U}(\alpha,\tilde{u})$ where α is a scalar such that $\mathcal{U}(\alpha,\tilde{u})$ is increasing with α.

Info-Gap (Ben-Haim, 2006, pp. 70-71) suggests the following format for these sub-regions of uncertainty:

$$\mathfrak{U}(\alpha,\tilde{u}) := \{u \in \mathbb{R}^N : u = \tilde{u} + v \ , \ v^T W v \leq \alpha^2\}, \quad \alpha \geq 0 \tag{40}$$

where W is a real, symmetric, positive definite matrix.

For each value of $\alpha \geq 0$, the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$ is then an ellipsoid whose center is at the nominal point $\tilde{u}$. Note that elements of $\mathfrak{U}(\alpha, \tilde{u})$ can be negative, especially for large values of α.

By implication then, the complete region of uncertainty, $\mathfrak{U}$, is the smallest subset of $\mathbb{R}^N$ containing all these regions. This entails that $\mathfrak{U} = \mathbb{R}^N$.

At this stage we should not be concerned that these regions of uncertainty center around the nominal point $\tilde{u}$. Namely, we regard (40) as a *representation* device and $\tilde{u}$ as a *reference point.* This representation does not suggest – at this stage – that the "true" value of u is more likely to be in the neighborhood of $\tilde{u}$ than in other parts of the complete uncertainty region $\mathfrak{U}$. You'll recall that any such assertion is in blatant violation of the very concept of *severe uncertainty.*

There is another thing that we need to do to apply *Info-Gap* to the portfolio investment problem under consideration, that is we need to formulate a *robustness function* for this problem.

The idea is to devise a function stipulating how robust an investment decision (q) is. To this end *Info-Gap* regards the coefficient α as a measure of *robustness*: the larger α is, the better. Formally, the robustness function is defined as follows:

$$\hat{\alpha}(q, r_c) := \max\{\alpha \geq 0 : R(q, u) \geq r_c, \forall u \in \mathcal{U}(\alpha, \tilde{u})\} \tag{41}$$

$$= \max\left\{\alpha \geq 0 : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u)\right\} \tag{42}$$

where r_c represents a minimum critical reward (so called *minimum attractive rate of return (MARR)*). Note that this suggests the following modification of $\mathcal{U}(\alpha, \tilde{u})$:

$$\mathcal{U}(\alpha, \tilde{u}, q, r_c) := \{u \in \mathcal{U}(\alpha, \tilde{u}) : R(q, u) \geq r_c\} \tag{43}$$

By construction, $\mathcal{U}(\alpha, \tilde{u}, q, r_c)$ is the subset of $\mathcal{U}(\alpha, \tilde{u})$ whose elements satisfy the reward constraint

$$R(q, u) \geq r_c \tag{44}$$

for the specified values of q, $\tilde{u}$ and r_c.

This definition entails that

$$\hat{\alpha}(q, r_c) = \max\{\alpha \geq 0 : \mathcal{U}(\alpha, \tilde{u}, q, r_c) = \mathcal{U}(\alpha, \tilde{u})\} \tag{45}$$

To formulate a user-friendly representation of this function, observe that the smallest reward associated with u values in $\mathcal{U}(\alpha, \tilde{u})$ is

$$\underline{r}(q, \alpha) := \min_{u \in \mathfrak{U}(\alpha, \tilde{u})} q^T u, \;\; \alpha \geq 0 \tag{46}$$

$$= \min_{v}\{q^T(\tilde{u} + v) : v^T W v \leq \alpha^2\} \tag{47}$$

$$= q^T \tilde{u} + \min_{v}\{q^T v : v^T W v \leq \alpha^2\} \tag{48}$$

$$= q^T \tilde{u} - \alpha\sqrt{q^T W^{-1} q} \quad \textit{(see derivation in Appendix A.1)} \tag{49}$$

Hence, the largest value of α for which $\underline{r}(q,\alpha) \geq r_c$ is

$$\alpha^*(q,r_c) = \frac{q^T\tilde{u} - r_c}{\sqrt{q^T W^{-1} q}} \tag{50}$$

Formally, *Info-Gap* defines the robustness of q with respect to r_c as follows

$$\hat{\alpha}(q,r_c) = \begin{cases} \dfrac{q^T\tilde{u} - r_c}{\sqrt{q^T W^{-1} q}} & , \quad \dfrac{q^T\tilde{u} - r_c}{\sqrt{q^T W^{-1} q}} \geq 0 \\ 0 & , \quad \dfrac{q^T\tilde{u} - r_c}{\sqrt{q^T W^{-1} q}} < 0 \end{cases} \tag{51}$$

The optimal investment decision, $\tilde{q}$, according to *Info-Gap*, is the decision that maximizes the robustness index $\hat{\alpha}(q,r_c)$, hence (assuming that $\hat{\alpha}(q,r_c) > 0, \forall q \in \mathbb{Q}$)

$$\tilde{q}(r_c) = \arg\max_q \left\{ \hat{\alpha}(q,r_c) : \sum_{i=1}^N q_i = Q \;\; q \geq 0 \right\} \tag{52}$$

$$= \arg\max_q \left\{ \frac{q^T\tilde{u} - r_c}{\sqrt{q^T W^{-1} q}} : \sum_{i=1}^N q_i = Q, q \geq 0 \right\} \tag{53}$$

This then is the *Info-Gap* recipe for deciding on the best investment for the assumed value of the critical reward r_c.

Now, when you examine this result it seems either too good to be true or a major breakthrough in decision theory.

After all, we are dealing here with a difficult problem involving decision-making under **severe** uncertainty and yet we did not have to grapple with the central issue – *severe uncertainty*. How did we manage to do that?

Interestingly, the *Info-Gap* literature is totally oblivious to this matter. So it is instructive to consider a naive instance of the portfolio investment problem and to analyze the results generated by the *Info-Gap* model.

Our naive instance consists of only $N = 2$ assets. And to keep things simple, we assume that

$$\max\{\tilde{u}_1, \tilde{u}_2\} > r_c \tag{54}$$

$$W^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{55}$$

in which case the *Info-Gap* solution is as follows:

$$\tilde{q}(r_c, \tilde{u}, W) = \arg\max_q \left\{ \frac{q^T\tilde{u} - r_c}{\sqrt{q^T W^{-1} q}} : \sum_{i=1}^{N} q_i = Q, \quad q \geq 0 \right\} \tag{56}$$

$$= \arg\max_q \left\{ \frac{q_1\tilde{u}_1 + q_2\tilde{u}_2 - r_c}{\sqrt{q_1^2 + q_2^2}} : q_1 + q_2 = 1, \quad q \geq 0 \right\} \tag{57}$$

This boils down to a simple equivalent one-dimensional problem with $d = q_1$:

$$\tilde{d}(r_c, \tilde{u}, W) := \arg\max_{0 \leq d \leq 1} \left\{ \frac{d\tilde{u}_1 + (1-d)\tilde{u}_2 - r_c}{\sqrt{d^2 + (1-d)^2}} \right\} \tag{58}$$

$$= \max\left\{0, \min\left\{1, \frac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c}\right\}\right\} \tag{59}$$

More explicitly,

$$\tilde{d}(r_c, \tilde{u}, W) = \begin{cases} 0 & , \quad \dfrac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} < 0 \\ \dfrac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} & , \quad \textit{otherwise} \\ 1 & , \quad \dfrac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} > 1 \end{cases} \tag{60}$$

The derivation of this result is provided in Appendix A.2.

It is immediately clear that for a fixed r_c, the optimal solution $\tilde{d}(r_c, \tilde{u}, W)$ can vary widely as we change the value of the nominal point $\tilde{u}$. In fact, it can vary continuously from its lower bound (0) to its upper bound (1).

Consider for example the case where $r_c = 10$ and $\tilde{u} = (16, 14)$ in which case the objective function is as follows:

$$z(d) := \frac{16d + 14(1-d) - 10}{\sqrt{d^2 + (1-d)^2}} = \frac{2d + 4}{\sqrt{d^2 + (1-d)^2}}, \quad 0 \leq d \leq 1 \tag{61}$$

The picture depicting $z(d)$ vs d is shown in Figure 5.

The optimal decision in this case is

$$\tilde{d}(10, (16, 14), W) = \frac{16 - 10}{16 + 14 - 20} = \frac{3}{5} \tag{62}$$

and the corresponding value of the objective function is

$$z(3/5) = \frac{2(3/5) + 4}{\sqrt{(3/5)^2 + (2/5)^2}} = 2\sqrt{13} = 7.2111026 \tag{63}$$

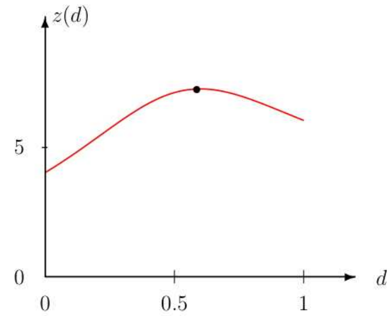

Fig. 5. Graph of the robustness $z(d)$ as a function of the decision variable d

Figure 6 illustrates how the optimal decision d changes as we change the value of the nominal point $\tilde{u}$ parametrically via a parameter $\lambda \in [10, 20]$.

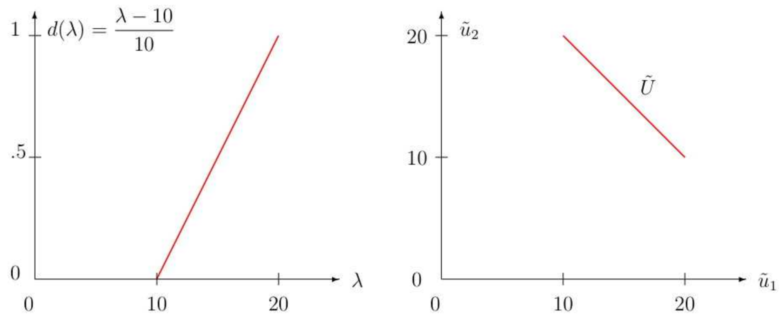

Fig. 6. Optimal d as a function of $\tilde{u} \in \tilde{U} := \{(\lambda, 30 - \lambda) : 10 \leq \lambda \leq 20\}$, $r_c = 10$

By the same token, we can see that the region of uncertainty associated with the optimal solution can be very small relative to the complete region of uncertainty, entailing that the optimal solution is robust only with respect to a relatively small region around $\tilde{u}$.

In the above example the region of uncertainty associated with the optimal decision $d = 3/5$ is

$$\mathcal{U}(\alpha, \tilde{u}) := \{u \in \mathbb{R}^2 : u = \tilde{u} + v, v^T W v \leq \alpha^2\} \tag{64}$$

$$= \left\{u \in \mathbb{R}^2 : u = (16, 14) + v, v^T W v \leq \left(2\sqrt{13}\right)^2\right\} \tag{65}$$

$$= \left\{u \in \mathbb{R}^2 : u = (16, 14) + v, v_1^2 + v_2^2 \leq 52\right\} \tag{66}$$

$$= \left\{u \in \mathbb{R}^2 : (u_1 - 16)^2 + (u_2 - 14)^2 \leq 52\right\} \tag{67}$$

This is a circle of radius $\sqrt{52}$ centered at $\tilde{u} = (16, 14)$, as shown in Figure 7. We deliberately show a large area of the total region of uncertainty to stress the local nature of the analysis. In this regard, recall that the full region of uncertainty in this case is $\mathfrak{U} = \mathbb{R}^2$.

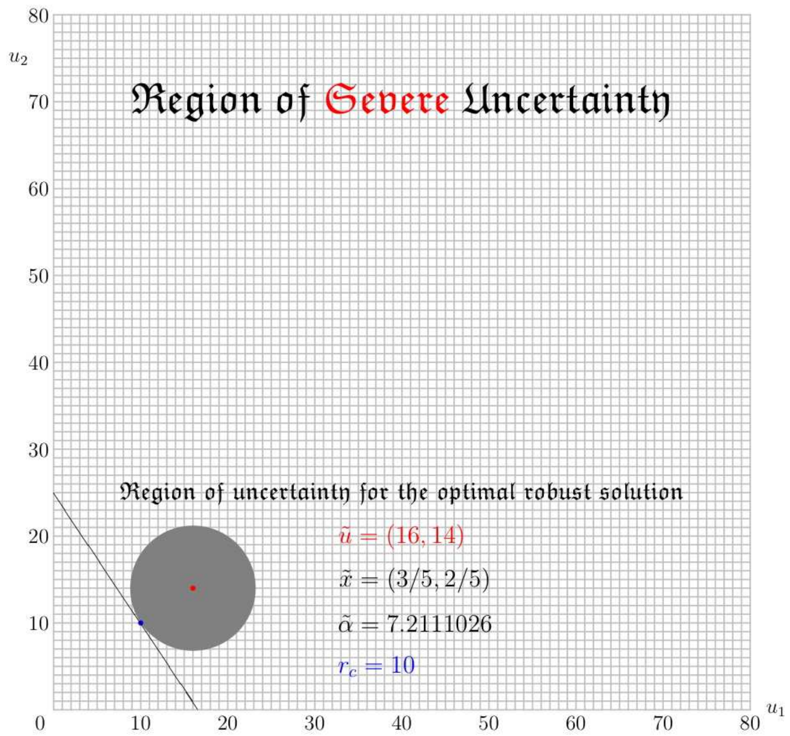

Fig. 7. The optimal uncertainty region $\mathcal{U}(\alpha, \tilde{u})$ for our example

Note that for obvious reasons we cannot display the entire (unbounded) region of uncertainty. Nevertheless, it does not take much to see that the region $\mathcal{U}(\alpha, \tilde{u})$ covered by the optimal decision generated by *Info-Gap* is exceedingly small relative to the entire region of uncertainty $\mathfrak{U} = \mathbb{R}^2$.

The implication is that in this case the *Info-Gap* analysis was confined to a minute part of the complete region of uncertainty.

Given that decisions are made here under severe uncertainty, it is advisable – if not imperative – to analyze in a similar fashion other nominal points representing other sub-regions of the region of uncertainty $\mathfrak{U}$. The question naturally arises: what will happen if we change the value of the nominal point $\tilde{u}$?

As we have noted above, if you do that you will generate new optimal solutions with corresponding new sub-regions of uncertainty.

For example, if we now test the new nominal point $\tilde{u} = (20, 20)$, the optimal solution would be

$$\tilde{d}(10, (20, 20), W) = \frac{20 - 10}{20 + 20 - 20} = \frac{1}{2} \tag{68}$$

and the corresponding value of the objective function is

$$z(1/2) = \frac{20(1/2) + 20(1/2) - 10}{\sqrt{(1/2)^2 + (1/2)^2}} = 10\sqrt{2} = 14.1421 \tag{69}$$

Observe that the robustness index for this solution is almost twice as large as the robustness index for the solution generated for $\tilde{u} = (16, 14)$. Since we are dealing here with severe uncertainty, why should we prefer the optimal solution ($d = 3/5$) generated for $\tilde{u} = (16, 14)$ to the optimal solution ($d = 1/2$) generated for $\tilde{u} = (20, 20)$? And how about other possible values for $\tilde{u}$, say $(20, 15)$ or $(15, 20)$, or whatever?

Clearly, one need hardly embark on a formal mathematical analysis of, and numerical experiments with, the *Info-Gap* methodology to demonstrate its failure to deal with decision-making under severe uncertainty. The flaw can be clearly seen pictorially.

Imagine that your complete region of uncertainty $\mathfrak{U}$ is the positive quadrant of $\mathbb{R}^2$ and suppose that, as instructed by *Info-Gap*, you resolve the severe uncertainty issue by choosing some nominal point $\tilde{u}$ in this region.

Since to determine the robustness of any decision you conduct a worst-case analysis with respect to the reward $R(q, u)$ around $\tilde{u}$, the critical value of u should not be – in general – far from $\tilde{u}$.

Hence, in general the subregion $\mathcal{U}(q, \hat{\alpha}(q, r_c))$ is expected to be considerably smaller than $\mathfrak{U}$. The picture is as shown in Figure 4.

Needless to say, there could be decision-making situations where the choice of $\tilde{u}$ can be safely justified by our knowledge and understanding of the system under consideration. In other words, the choice of the nominal value of the parameter of interest could be justified by solid data, experience, and familiarity with the problem situation under consideration.

This, of course, is a fact of life.

But such cases definitely do not fall within the category of decision-making under SEVERE uncertainty.

The bottom line is that *Info-Gap* cannot have it both ways: if it claims to be a tool for decision-making under severe uncertainty then it cannot use a model based on a single point estimate and its immediate neighborhood. If the best point estimate is so good that the recipe indeed yields robust solutions, then it cannot be claimed that we are in a situation that can be categorized as decision-making under SEVERE uncertainty.

8. CONCLUSIONS

Mathematical models for decision-making under severe uncertainty are – by necessity – austere and involve highly simplistic assumptions regarding uncertainty. For many years now, classical decision theory has offered two paradigms for this purpose: *Laplace's Principle of Insufficient Reason* and *Wald's Maximin Principle*. To be sure, these principles are far from perfect, yet they still dominate the scene.

In particular, *Wald's Maximin Principle* is used extensively in *worst-case analysis* and *robust optimization*.

As illustrated in this discussion, from a mathematical modeling point of view its deployment requires subtlety and imagination. Indeed, this is one of the lessons learned from the *Info-Gap* experience.

A. APPENDIX

A.1. PROOF OF THE RESULT USED IN EQUATION (49)

We have to show that

$$\underline{r}(q,\alpha) := q^T\tilde{u} + \min_{v}\{q^Tv : v^TWv \leq \alpha^2\} \tag{70}$$

$$= q^T\tilde{u} - \alpha\sqrt{q^TW^{-1}q} \tag{71}$$

namely that the optimal value of v is such that $q^Tv = -\alpha\sqrt{q^TW^{-1}q}$.

Observe that because q is non-negative, the constraint $v^TWv \leq \alpha^2$ is binding at the optimum. And since the minimization problem is convex, the first order optimality condition for the associated Lagrangian problem is as follows

$$\nabla_{v,\lambda}\left\{q^Tv + \lambda\left\{v^TWv - \alpha^2\right\}\right\} = \mathbf{0} \tag{72}$$

$$\left(q + \frac{1}{2}\lambda Wv \;\; v^TWv - \alpha^2\right) = (0,0) \tag{73}$$

Hence,

$$\frac{1}{2}\lambda Wv = -q \tag{74}$$

$$v^TWv = \alpha^2 \tag{75}$$

and therefore

$$\lambda = -\frac{2v^Tq}{\alpha^2} \tag{76}$$

Now, multiplying (74) on the left by q^TW^{-1} we obtain

$$\frac{1}{2}\lambda q^TW^{-1}Wv = -q^TW^{-1}q \tag{77}$$

$$\frac{1}{2}\lambda q^Tv = -q^TW^{-1}q \tag{78}$$

$$-\frac{\left(q^Tv\right)\left(v^Tq\right)}{\alpha^2} = -q^TW^{-1}q \tag{79}$$

Hence,

$$q^T v = \pm\,\alpha\sqrt{q^T W^{-1} q} \tag{80}$$

Since $\alpha \geq 0$ and we are minimizing, the optimal solution is

$$q^T v = -\alpha\sqrt{q^T W^{-1} q} \tag{81}$$

as required.

A.2. PROOF OF EQUATION (60)

We have to show that if $\max\{\tilde{u}_1, \tilde{u}_2\} \geq r_c$, then the optimal solution (d) of

$$\tilde{d}(r_c, \tilde{u}, W) := \arg\max_{0 \leq d \leq 1} \left\{ \frac{d\tilde{u}_1 + (1-d)\tilde{u}_2 - r_c}{\sqrt{d^2 + (1-d)^2}} \right\} \tag{82}$$

is

$$\tilde{d}(r_c, \tilde{u}, W) = \frac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} \tag{83}$$

Since $\max\{\tilde{u}_1, \tilde{u}_2\} \geq r_c$ implies that the optimal solution yields a non-negative numerator in (82), we have

$$\tilde{d}(r_c, \tilde{u}, W) := \arg\max_{0 \leq d \leq 1} \left\{ \frac{\{d\tilde{u}_1 + (1-d)\tilde{u}_2 - r_c\}^2}{d^2 + (1-d)^2} \right\} \tag{84}$$

The expression being maximized is pseudoconcave with d on the interval $[0, 1]$, (see Avriel, 1976, Theorem 9.6, pp. 154-5), so any stationary point in the feasible region is a global optimum.

Equating to zero the derivative (with respect to d) of the expression being maximized we obtain the two roots

$$d' = \frac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} \ ; \ d'' = \frac{\tilde{u}_2 - r_c}{\tilde{u}_2 - \tilde{u}_1} \tag{85}$$

It turns out that the second root, d'', is not relevant because it always yields 0 for the objective function, regardless of the values of $\tilde{u}$ and r_c, namely

$$\left\{ \frac{d\tilde{u}_1 + (1-d)\tilde{u}_2 - r_c}{\sqrt{d^2 + (1-d)^2}} \right\}_{d=d''} = 0 \tag{86}$$

Note that for the two end points of the feasible range of d we have

$$\left\{ \frac{d\tilde{u}_1 + (1-d)\tilde{u}_2 - r_c}{\sqrt{d^2 + (1-d)^2}} \right\}_{d=0} = \tilde{u}_2 - r_c \tag{87}$$

$$\left\{ \frac{d\tilde{u}_1 + (1-d)\tilde{u}_2 - r_c}{\sqrt{d^2 + (1-d)^2}} \right\}_{d=1} = \tilde{u}_1 - r_c \tag{88}$$

Thus, if $\max\{\tilde{u}_1, \tilde{u}_2\} \geq r_c$, then the second root, d'', cannot be strictly better than both end points.

We conclude therefore that

$$\tilde{d}(r_c, \tilde{u}, W) = \begin{cases} 0 & , \quad \dfrac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} < 0 \\ \dfrac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} & , \quad otherwise \\ 1 & , \quad \dfrac{\tilde{u}_1 - r_c}{\tilde{u}_1 + \tilde{u}_2 - 2r_c} > 1 \end{cases} \tag{89}$$

REFERENCES

1. M. Avriel, *Nonlinear programming analysis and methods,* Prentice-Hall, 1976.
2. Y. Ben-Haim, *Information Gap Decision Theory,* Academic Press, 2001.
3. Y. Ben-Haim, *Info-Gap Decision Theory,* Elsevier, 2006.
4. Y. Ben-Haim, *Value-at-risk with info-gap uncertainty,* The Journal of Risk Finance, 6 (5) (2005), 388–403.
5. A. Ben-Tal, L. El Ghaoui, A. Nemirovski, *Mathematical Programming,* Special issue on *Robust Optimization*, Volume 107 (1–2), 2006.
6. S.D. French, *Decision Theory,* Ellis Horwood, 1988.
7. R. Grünig, R. Kühn, *Successful decision-making,* Springer-Verlag. 2005.
8. S.F. Hillier, G.J. Lieberman, *Introduction to Operations Research,* 8th edition, McGraw Hill, 2005.
9. R.E. Markland, J.R. Sweigart, *Quantitative Methods: Applications to Managerial decision-making,* John Wiley, 1987.
10. P. Kouvelis, G. Yu, *Robust Discrete Optimization and Its Applications,* Kluwer, 1997.
11. C.T. Ragsdale, *Spreadsheet Modeling and Decision Analysis*, Thomson, 2004.
12. B. Rustem, M. Howe, *Algorithms for Worst-case Design and Applications to Risk Management*, Princeton University Press, 2002.
13. M. Sniedovich, *OR/MS Games: 3. The Counterfeit coin problem,* INFORMS Transactions in Education, 3 (2003), 2, 32–41.
14. M. Sniedovich, *What is wrong with Info-Gap? An Operations Research Perspective*, Working Paper MS-1-06, presented at the ASOR Mini-Conference, Dec 1, 2006, Melbourne, Australia, 2006.
15. R. Steuer, *Multiple Criteria Optimization: Theory, Computation and Application*, John Wiley, 1985.
16. H. Vladimirou, S.A. Zenios, *Stochastic Programming and Robust Optimization,* Chapter 12 in Gal T. and Greenberg H.J. (Eds) Advances in Sensitivity Analysis and Parametric Programming, Kluwer, 1997.

17. A. Wald, *Statistical decision functions which minimize the maximum risk,* The Annals of Mathematics, 46 (1945), 2, 265–280.

18. A. Wald, *Statistical Decision Functions*, John Wiley, 1950.

19. W.L. Winston, *Operations Research: Applications and Algorithms*, Duxbury Press, 1994.