Info-Gap Economics: An Overview

Yakov Ben-Haim[1]

Contents

[0]\papers\BoE2008\ige03.tex. 22.5.2008.

[1]Yitzhak Moda'i Chair in Technology and Economics, Technion — Israel Institute of Technology, Haifa 32000 Israel, yakov@technion.ac.il.

The author is indebted to valuable comments by Maria Demertzis and Joe Moffitt.

The paper was presented in part at the Research Forum on "Risk and Uncertainty in Monetary Policy", Centre for Central Bank Studies, Bank of England, London, 9–11 June 2008.

1 Introduction

Economists focus on two great questions: "What should be done?" and "Why didn't it work?".[2]

Part I of this paper addresses the first question, by studying the application of info-gap decision theory to the formulation and evaluation of decisions in economics. We construe "decisions" broadly, to include policy decisions as well as decisions on how to model and forecast economic processes. In part I we view modelling as a decision process in which we choose variables, relations, dimensions, and so on, with the aim of subsequently predicting behavior or choosing a policy. We begin by studying policy formulation given a linear model of the economic system (section 3). We introduce the robust-satisficing decision strategy, which satisfices the outcome at an acceptable level and maximizes the immunity against Knightian uncertainty. The main contention is that robust-satisficing is different from, and often more reliable than, optimizing the outcome based on a best-estimate of the underlying process. This thesis is strengthened when a "proxy theorem" holds. A proxy theorem asserts that the info-gap robustness is a proxy for successfully achieving the policy goals: any increase in robustness also increases the probability of success. When a proxy theorem holds, the policy maker can maximize the probability of success without knowing the underlying probability distributions at all. This is particularly important in formulating policy under Knightian uncertainty, where probabilities are unknown. Two different proxy theorems are developed (sections 3.6 and 4). In section 5 we illustrate the info-gap approach to modelling by studying the problem of estimating the slope of a Phillips curve. We show that robust-satisficing estimates are more robust to uncertainty in the underlying economics than least-squares estimation of historical data.

In part II we view modelling in the more conventional "scientific" sense of understanding and describing the behavior of economic agents. The main contention is that the neo-classical axiom of rationality—maximization of utility—is not necessary for a coherent and successful understanding of economic behavior. We look at two specific examples. In section 6 we study the equity premium puzzle. By deriving an info-gap generalization of the Lucas asset pricing relations we come to an understanding of observed premia of risky assets. The explanation hinges on the hypothesis that investors need to reliably achieve satisfactory (rather than maximal) returns, which are achieved with greatest likelihood by using a robust-satisficing strategy because a proxy theorem holds. In section 7 we formulate an info-gap approach to the micro-economics of demand. We derive and study the properties of demand functions using both the robust-satisficing and the opportune-windfalling strategies. We derive two distinct demand functions and show that they display the features of the Hicksian and Walrasian demand functions.

It is widely believed that good decision-making depends on good understanding of the processes which are influenced by the decisions. This is true in part, but not entirely. Indeed, economics has begun to outgrow 19th century positivistic idealism which taught that science converges on truth, and truth enables control. Positivism would have us first study economic behavior (part II) and then develop decision strategies (part I). This may work, more or less, in science-based technology, but it is less successful in social science. In part III, section 8, we touch on some fundamental methodological issues underlying the info-gap approach to economics, by discussing a debate between Friedman and Samuelson.

Before beginning our study we present a brief intuitive discussion of info-gap decision theory in section 2. We conclude with an outline of some further lines of research in section 9. Proofs and derivations appear in the appendices.

[2]Of course sometimes policies do succeed, but that is usually a cause for celebration rather than introspection.

2 What is Info-Gap Theory?

Info-gap theory is a methodology for supporting model-based decisions under severe uncertainty (Ben-Haim, 2006). An info-gap is a disparity between what *is known,* and what *needs to be known* in order to make a comprehensive and reliable decision. An info-gap is resolved when a surprise occurs, or a new fact is uncovered, or when our knowledge and understanding change. We know very little about the substance of an info-gap. For instance, we rarely know what unusual event will delay the completion of a task. Even more strongly, we *cannot* know what is not yet discovered, such as tomorrow's news, or future scientific theories or technological inventions. The ignorance of these things are info-gaps. An info-gap is a Knightian uncertainty (Knight, 1921) since it is not characterized by a probability distribution.

Info-gap decision theory is based on three elements. The first element is an *info-gap model of uncertainty,* which is a non-probabilistic quantification of uncertainty. The uncertainty may be in the value of a parameter, such as the slope of the Phillips curve, or in a vector such as the future returns on a portfolio of investments. An info-gap may in the shape of a function, such as demand vs. price, or the shape of the tail of the probability distribution function (pdf) of extreme financial loss. An info-gap may be in the size and shape of a set of such entities, such as the set of possible pdf's or the set of possible Phillips curves. We will encounter many examples of info-gap models of uncertainty. In all cases an info-gap model is an unbounded family of nested sets of possible realizations. For instance, if the uncertain entity is a function then the info-gap model is an unbounded family of nested sets of realizations of this function. An info-gap model does not posit a worst case or most extreme uncertainty.[3]

The second element of an info-gap analysis is a *model of the system,* such as a macroeconomic model, or a capital asset pricing model. The model expresses our knowledge about the system, and may also depend on uncertain elements whose uncertainty is represented by an info-gap model of uncertainty. The system model also depends on the decisions to be made, and quantifies the outcomes of those decisions given specific realizations of the uncertainties. For instance, the model may express macroeconomic outcomes such as inflation, unemployment, growth of the GDP, and so on.

The third element of an info-gap analysis is a set of *performance requirements.* These specify values of the outcomes which the decision maker requires or aspires to achieve. These values may constitute success of the decision, or at least minimally acceptable values. For instance, inflation targeting is sometimes formulated as a range of inflation values which are acceptable. Performance requirements can embody the concept of satisficing: doing good enough or meeting critical requirements. Alternatively, the performance requirements can express windfall aspirations for better-than-anticipated outcomes. We will encounter examples of both satisficing and windfalling requirements, though satisficing requirements are the most common.

These three components—uncertainty model, system model, and performance requirements—are combined in formulating two *decision functions* which support the choice of a course of action.

The *robustness function* assesses the greatest tolerable horizon of uncertainty. The robustness function is a quantitative answer to the question: how wrong can we be in our data, models and understanding, and the action we are considering will still lead to an acceptable outcome. The robustness function is based on a satisficing performance requirement. When operating under severe uncertainty, a decision which achieves an acceptable outcome over a large range of uncertain realizations is preferable to a decision which fails to achieve an acceptable outcome even under small error. In this way the robustness function generates preferences on available decisions.

The *opportuneness function* assesses the lowest horizon of uncertainty which is necessary for better-than-anticipated outcomes to be possible (though not guaranteed). The windfalling decision maker asks: how wrong must we be in order for quite attractive outcomes to be possible? The

[3]Sometimes the family of sets is bounded by virtue of the definition of the uncertain entity. For instance, a probability must be between zero and one, so the family of nested sets of possible probability values is bounded. However, this bound does not derive from knowledge about the event whose probability is uncertain, but only from the mathematical definition of probability. Such an info-gap model is unbounded in the universe of probability values.

opportuneness function is based on windfalling rather than satisficing. When operating under severe uncertainty it is possible that best-model anticipations are overly pessimistic; the windfaller seeks to exploit the ambient uncertainty. A decision which would result in a really wonderful outcome if we err only slightly is preferred (by the windfaller) over a decision which requires great deviation in order to enable the same outcome. The opportuneness function thus generates preferences over the available decisions. These preferences may not agree with the preferences generated by the robustness function.

Info-gap theory originated in engineering and has since been applied to a wide range of disciplines (Ben-Haim, 1996, 2006). Burgman (2005) devotes a chapter to info-gap theory as a tool for biological conservation and environmental management. Regan *et al* (2005) use info-gap theory to devise a preservation program for an endangered rare species. McCarthy and Lindenmayer (2007) use info-gap theory to manage commercial timber harvesting that competes with urban water requirements. Knoke (2007) uses info-gap theory in a financial model for forest management. Carmel and Ben-Haim (2005) use info-gap theory in a theoretical study of foraging behavior of animals. Ben-Haim and Jeske (2003) use info-gap theory to explain the home-bias paradox. Ben-Haim (2006) uses info-gap theory to study the equity premium puzzle (Mehra and Prescott, 1985) and the paradoxes of Ellsberg (1961) and Allais (see Mas-Colell, Whinston and White, 1995). Akram *et al* (2006) use info-gap theory in formulating monetary policy. Fox *et al* (2007) study the choice of the size of a statistical sample when the sampling distribution is uncertain. Klir (2006) discusses the relation between info-gap models of uncertainty and a broad taxonomy of measure-theoretic models of probability, likelihood, plausibility and so on. Moffitt *et al* (2006) employ info-gap theory in designing container-inspection strategies for homeland security of shipping ports. Pierce *et al* (2006) use info-gap theory to design artificial neural networks for technological fault diagnosis. Kanno and Takewaki (2006a, b) use info-gap theory in the analysis and design of civil engineering structures. Pantelides and Ganzerli (1998) study the design of trusses and Ganzerli and Pantelides (2000) study the optimization of civil engineering structures. Ben-Haim and Laufer (1998) and Regev *et al* (2006) apply info-gap theory for managing uncertain task-times in projects. Ben-Haim and Hipel (2002) use info-gap theory in a game-theoretic study of conflict resolution.

Thus, info-gap decision theory has been used productively to model and manage circumstances of extreme uncertainty in a wide variety of different contexts and disciplines. We now proceed to study a range of economic applications.

Part I

Decision Strategies

3 Monetary Policy with a Linear Model

We now encounter our first example of an info-gap decision analysis. We have a linear economic model (the system model) whose outcome we wish to control (which defines several possible performance requirements). The coefficients of the economic model are uncertain which is represented with an info-gap model of uncertainty. We study the ideas of info-gap robustness and probability of success as tools for evaluating and selecting a policy. Specifically, we develop the following ideas:

- Formulate an info-gap model for uncertainty in a linear economic model, and derive robustness functions for one-sided and two-sided targets (sections 3.1 and 3.2).
- Prove and discuss a proposition asserting that these robustness curves cross one another (section 3.3). This implies that a policy which is optimal based on an estimated model may be less robust than other policies, which has important implications for policy selection.
- Discuss the relation between the robust-satisficing strategy and the class of strategies various known as min-max, robust control or worst case analysis (section 3.4).
- Define the probability of success (section 3.5).
- Prove a proposition showing that each one-sided robustness is a proxy for the corresponding probability of success. This has the important implication that the policy maker can maximize the probability of success, without knowing the probability distribution, by maximizing the robustness to uncertainty. This depends on the concept of 'standardization' (section 3.6.1).
- Show that the two-sided robustness is *not* usually a proxy for the probability of success, and explain why this is so. This emphasizes the motivation for decision makers to prefer one-sided targets when selecting policy (section 3.6.2).
- Discuss briefly the implications for learning by the decision maker (section 3.7).

3.1 Goals, Models, Uncertainties

Consider a linear model whose response, y, depends on model coefficients c and control variables v:

$$y = c^T v \tag{1}$$

y could be any economic variable, for instance, y might be the output gap.

The goal is to control the range of variation of y, and we consider three different specifications:

$$y \le y_{\rm c}, \quad y \ge -y_{\rm c}, \quad |y| \le y_{\rm c} \tag{2}$$

Intuitively, the decision maker would be interested in the first condition when concerned about overshoot of the output gap. Likewise, concern about under-performance would focus attention on the second condition, while the third condition treats both types of errors equivalently.

We can immediately appreciate one of the challenges facing decision makers, by noting that the goals in eq.(2) tend to conflict with one another. A strategy which moves y below and away from $y_{\rm c}$, brings y closer to $-y_{\rm c}$. As long as the analyst is concerned only with either the first or the second goal, but not both, this conflict is not problematic. However, this conflict of goals in inherent in the third strategy, which we will see to have important repercussions when focussing on the third goal.

We will consider an info-gap model for uncertainty in the model coefficients c. An info-gap model is an unbounded family of nested sets of coefficients. Two common examples are the interval-uncertainty and ellipsoidal-uncertainty models:

$$\mathcal{U}(h) = \left\{ c:\ |c_i - \widetilde{c}_i| \le s_i h,\ \forall\ i \right\}, \quad h \ge 0 \tag{3}$$

$$\mathcal{U}(h) = \left\{ c:\ (c - \widetilde{c})^T S^{-1} (c - \widetilde{c}) \le h^2 \right\}, \quad h \ge 0 \tag{4}$$

The s_i's are non-negative and S is a positive definite symmetric matrix.

3.2 Robustness

Each of the three conditions in eq.(2) generates a robustness function, which is the greatest horizon of uncertainty, h, up to which the corresponding condition is satisfied for all realizations of the model parameters:

$$\widehat{h}_{+}(v, y_{\rm c}) = \max\left\{h : \left(\max_{c\in\mathcal{U}(h)} y\right) \le y_{\rm c}\right\} \tag{5}$$

$$\widehat{h}_{-}(v, y_{\rm c}) = \max\left\{h : \left(\min_{c\in\mathcal{U}(h)} y\right) \ge -y_{\rm c}\right\} \tag{6}$$

$$\widehat{h}(v, y_{\rm c}) = \max\left\{h : \left(\max_{c\in\mathcal{U}(h)} |y|\right) \le y_{\rm c}\right\} \tag{7}$$

The robustness is a property of the control vector v and the performance requirement $y_{\rm c}$. A large value of robustness means that the performance is achieved by the control even if the model coefficients deviate greatly from their estimated values. Large robustness implies confidence in the outcome; low robustness implies lack of confidence.

Since both the upper and lower inequalities must be satisfied in the overall robustness, it is clear that:

$$\widehat{h}(v, y_{\rm c}) = \min\{\widehat{h}_{-},\ \widehat{h}_{+}\} \tag{8}$$

We define the estimated value of y as $\widetilde{y} = \widetilde{c}^T v$. Also, we will say that the info-gap model, $\mathcal{U}(h)$, is *linearly symmetric* if there is a non-negative function, $\theta(v)$, independent of the horizon of uncertainty h, such that:

$$\max_{c\in\mathcal{U}(h)} y = \widetilde{y} + h\theta(v) \quad \text{and} \quad \min_{c\in\mathcal{U}(h)} y = \widetilde{y} - h\theta(v) \tag{9}$$

A linearly symmetric info-gap model expands linearly with increasing horizon of uncertainty, h, so that the functions in eq.(9) are linear in h. The info-gap model is symmetric in the sense that the same function, $\theta(v)$, applies to both the 'min' and the 'max'.[4]

For instance, for the unbounded-interval info-gap model, eq.(3), one finds:

$$\theta(v) = s^T|v| \tag{10}$$

where $|v|$ is the vector formed by replacing each element of v by its absolute value. Similarly, for the ellipsoidal info-gap model in eq.(4):

$$\theta(v) = \sqrt{v^T S v} \tag{11}$$

One can readily show that, for linearly symmetric info-gap models:

$$\widehat{h}_{+} = \frac{y_{\rm c} - \widetilde{y}}{\theta} \tag{12}$$

$$\widehat{h}_{-} = \frac{y_{\rm c} + \widetilde{y}}{\theta} \tag{13}$$

$$\widehat{h} = \frac{y_{\rm c} - |\widetilde{y}|}{\theta} \tag{14}$$

where $\widehat{h}_{+}$, $\widehat{h}_{-}$ or $\widehat{h}$ is defined to equal zero if the corresponding expression, eq.(12), (13) or (14), is negative. Eqs.(8), (12) and (13) are equivalent to eq.(14).

[4]For what follows, the linearity is more important than the symmetry. If a different $\theta(v)$-function holds for the 'min' and the 'max', say $\theta_{-}(v)$ and $\theta_{+}(v)$, then our results remain much the same, though they are notationally a bit more cumbersome.

3.3 Crossing of Robustness Curves

The following proposition, whose proof appears in appendix A, asserts a necessary and sufficient condition for robustness curves to cross one another for different choice of the control variables v.

Proposition 1 *Robustness curves for two different control vectors cross one another if and only if the control which is estimated to be more favorable, has greater sensitivity to uncertainty.*

Given *that the info-gap model is symmetric, so that the robustness functions are specified in eqs.(12)–(14).*

Then *these robustness curves, evaluated for two different control vectors v and v', cross at a single positive value of y_c (illustrated in figs. 1–3) if and only if:*

$$\text{For } \widehat{h}_+ : \quad [\theta(v) - \theta(v')] \, [\widetilde{y}(v) - \widetilde{y}(v')] \le 0 \quad \textit{and} \quad \theta(v) \ne \theta(v') \tag{15}$$

$$\text{For } \widehat{h}_- : \quad [\theta(v) - \theta(v')] \, [\widetilde{y}(v) - \widetilde{y}(v')] \ge 0 \quad \textit{and} \quad \theta(v) \ne \theta(v') \tag{16}$$

$$\text{For } \widehat{h} : \quad [\theta(v) - \theta(v')] \, (|\widetilde{y}(v)| - |\widetilde{y}(v')|) \le 0 \quad \textit{and} \quad \theta(v) \ne \theta(v') \tag{17}$$

They cross at positive robustness if and only if the inequality, '$\le$' or '$\ge$', is strict '$<$' or '$>$'.

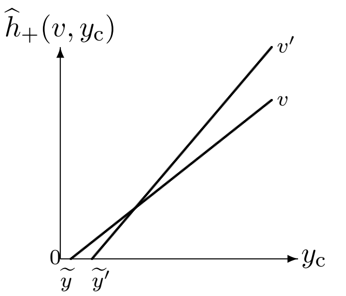

Figure 1: Crossing of upper-robustness curves, $\widehat{h}_+$. $\widetilde{y} \equiv \widetilde{y}(v)$ and $\widetilde{y}' \equiv \widetilde{y}(v')$ and $\theta(v) > \theta(v')$.

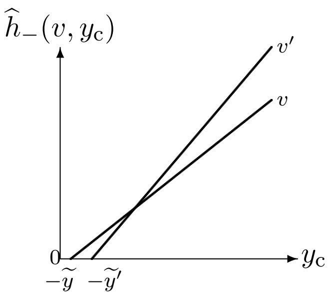

Figure 2: Crossing of lower robustness curves, $\widehat{h}_-$.

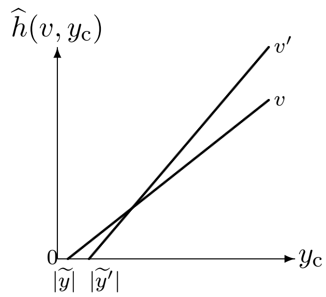

Figure 3: Crossing of total robustness curves, $\widehat{h}$.

Fig. 1 illustrates proposition 1 for $\widehat{h}_+$. (Figs. 2 and 3 illustrate the proposition for $\widehat{h}_-$ and for $\widehat{h}$.) Controls v are more favorable than v', based on the estimated model $\widetilde{c}$, since $y(v) < y(v')$, causing $\widehat{h}_+(v, y_c)$ to sprout off the y_c- axis to the left of $\widehat{h}_+(v', y_c)$. However, v is more sensitive to error as expressed by its lower robustness due to $\theta(v) > \theta(v')$. This causes $\widehat{h}_+(v, y_c)$ to have lower slope than $\widehat{h}_+(v', y_c)$. Thus the robustness curves cross.

The crossing of robustness curves has important implications. First of all, crossing of robustness curves implies the potential for reversal of preference between policies for a robust-satisficing decision maker. Let $y_\times$ denote the value of y_c at which the curves cross. If the decision maker requires $y_c < y_\times$, then v is preferred over v' because v is more robust in this range of y_c values; requiring $y_c > y_\times$ implies the reverse preference.

Second, the robustness curves of quite different policies can cross. When v and v' are very different, the preference reversal will be between substantially different policy options, one of which is likely to be substantially more aggressive than the other. That is, as the critical loss, y_c, varies, there can be large and sudden change or bifurcation in policy preference. Sometimes the more robust policy will be more aggressive, and sometimes not.

Third, by comparing eqs.(15) and (16) we see that for any pair of controls, v and v', for which the upper robustness curves *do* cross, the lower robustness curves *do not* cross, unless $\widetilde{y}(v) = \widetilde{y}(v')$ which means that the intersection is at zero robustness. The reverse is also true: if the $\widehat{h}_-$ curves cross, then the $\widehat{h}_+$ curves do not. Since crossing of robustness curves entails preference reversal, we see that policy preferences for upper and lower targets (first and second constraints in eq.(2)) can differ substantially. This is another indication that the two-sided policy target (third constraint in eq.(2)) is problematic: it involves both lower- and upper-constraints, and these constraints can induce different preferences.

3.4 Min-Max, Robust Control and Robust-Satisficing

We take a brief intermezzo to compare the robust-satisficing strategy with a class of alternatives. The terms 'min-max', 'robust control' and 'worst-case' refer to a collection of decision strategies which attempt to ameliorate a maximally adverse outcome. This can of course be formulated in a variety of ways. In one form or another, either explicitly or implicitly, a greatest level of uncertainty or a worst possible outcome is posited. Then a strategy is sought which maximally diminishes the impact of this worst outcome.

Info-gap robust-satisficing is motivated by the same perception of uncertainty which motivates the min-max class of strategies: lack of reliable probability distributions and the potential for severe and extreme events. We will see that the robust-satisficing decision will sometimes coincide with a min-max decision. On the other hand we will identify some fundamental distinctions between the min-max and the robust-satisficing strategies and we will see that they do not always lead to the same decision.

First of all, if a worst case or maximal uncertainty is unknown, then the min-max strategy cannot be implemented. That is, the min-max approach requires a specific piece of knowledge about the real world: "What is the greatest possible error of the analyst's model?". This is an *ontological* question: relating to the state of the real world. In contrast, the robust-satisficing strategy does not require knowledge of the greatest possible error of the analyst's model. The robust-satisficing strategy centers on the vulnerability of the analyst's knowledge by asking: "How wrong can the analyst be, and the decision still yields acceptable outcomes?" The answer to this question reveals nothing about how wrong the analyst in fact is or could be. The answer to this question is the info-gap robustness function, while the true maximal error may or may not exceed the info-gap robustness. This is an *epistemic* question, relating to the analyst's knowledge, positing nothing about how good that knowledge actually is. The epistemic question relates to the analyst's knowledge, while the ontological question relates to the relation between that knowledge and the state of the world. In summary, knowledge of a worst case is necessary for the min-max approach, but not necessary for the robust-satisficing approach.

The second consideration is that the min-max approaches depend on what tends to be the least reliable part of our knowledge about the uncertainty. Under Knightian uncertainty we do not know the probability distribution of the uncertain entities. We may be unsure what are typical occurrences, and the systematics of extreme events are even less clear. Nonetheless the min-max decision hinges on ameliorating what is supposed to be a worst case. This supposition may be substantially wrong, so the min-max strategy may be mis-directed.

A third point of comparison is that min-max aims to ameliorate a worst case, without worrying about whether an adequate or required outcome is achieved. This strategy is motivated by severe uncertainty which suggests that catastrophic outcomes are possible, in conjunction with a precautionary attitude which stresses preventing disaster. The robust-satisficing strategy acknowledges unbounded uncertainty, but also incorporates the outcome requirements of the analyst. The choice between the two strategies—min-max and robust-satisficing—hinges on the priorities and preferences of the analyst.

The fourth distinction between the min-max and robust-satisficing approaches is that they need not lead to the same decision, even starting with the same information.

Consider the robustness curves in fig. 1, which are reproduced in fig. 4. We must choose between two alternative decisions, v and v', whose robustness curves are shown in the figure. Suppose that the greatest possible horizon of uncertainty in the analyst's models is known to be h_{max} as shown in fig. 4. This knowledge allows us to implement the min-max strategy. The min-max choice is v' rather than v since v' has a less adverse worst possible outcome, y_{mm}.

Now suppose that the analyst requires that the outcome be no worse than the value y_{rs} where y_{rs}

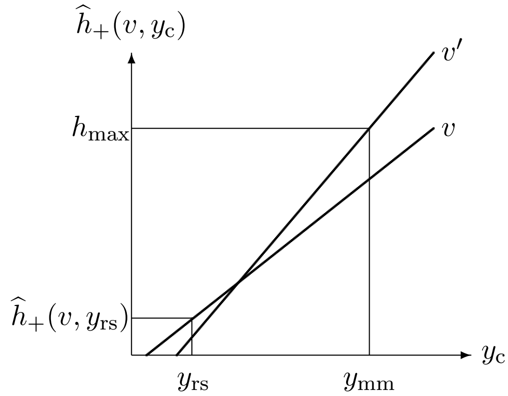

Figure 4: Crossing of upper-robustness curves, $\widehat{h}_+$. Modified from fig. 1.

is not only less than y_{mm} but also less than the value at which the robustness curves cross, as shown in fig. 4. The robust-satisficing choice, given this performance requirement, is v rather than v' since v guarantees that the outcome will be less than y_{rs} for a wider range of uncertainty than does v'. The robust-satisficing analyst acknowledges that the outcome could be as bad as y_{mm}, and could be worse with v than with v' for large uncertainty. The robust-satisficer chooses v since it guarantees the required outcome for a wider range of contingencies than does v'.

Fifth and finally, when a proxy theorem holds (as will be shown in section 3.6 to be the case for this particular problem) then the robust-satisficing choice v is in fact more likely to achieve the requirement, $y \le y_{rs}$, than is the min-max choice v'.

3.5 Probability of Success

In this section we define the probability of success for each of the three targets. In the next section we study the relation between the robustness and the probability of success.

3.5.1 Upper Robustness

Define the set of all values of the model coefficients, c, which do not violate the condition $y \le y_c$ in the upper robustness $\widehat{h}_+$:

$$\Lambda_+ = \{c: \ y \le y_c\} \tag{18}$$
$$= \left\{c: \ \frac{y - \widetilde{y}}{\theta} \le \frac{y_c - \widetilde{y}}{\theta}\right\} \tag{19}$$
$$= \left\{\zeta: \ \zeta \le \widehat{h}_+\right\} \tag{20}$$

provided that $\widehat{h}_+$ is not zero, and where we have defined the "standardized" variable $\zeta = (y - \widetilde{y})/\theta$.

If we think of c as a random vector, then y and ζ are both random variables. Let $Z(\zeta)$ denote the cumulative probability distribution (cdf) of ζ, with probability density function (pdf) $z(\zeta)$. The probability that $y \le y_c$ is the "probability of success" for which $\widehat{h}_+$ is the robustness. This probability is:

$$P_{s+} = Z(\widehat{h}_+) \tag{21}$$

3.5.2 Lower Robustness

We now do something similar for the lower robustness.

Define the set of all c-vectors which do not violate the condition $y \ge -y_c$ in the lower robustness $\widehat{h}_-$:

$$\Lambda_- = \{c: \ y \ge -y_c\} \tag{22}$$

$$= \left\{c:\ \frac{y-\widetilde{y}}{\theta} \ge -\frac{y_{\rm c}+\widetilde{y}}{\theta}\right\} \tag{23}$$

$$= \left\{\zeta:\ \zeta \ge -\widehat{h}_-\right\} \tag{24}$$

provided that $\widehat{h}_-$ is not zero.

The probability that $y \ge -y_{\rm c}$ is the probability of success for which $\widehat{h}_-$ is the robustness. This probability is:

$$P_{\rm s-} = 1 - Z(-\widehat{h}_-) \tag{25}$$

3.5.3 Two-Sided Robustness

Define Λ as the set of all c-vectors which satisfy the requirement $|y| \le y_{\rm c}$:

$$\Lambda = \{c:\ |y| \le y_{\rm c}\} \tag{26}$$

$$= \left\{c:\ -\frac{y_{\rm c}+\widetilde{y}}{\theta} \le \frac{y-\widetilde{y}}{\theta} \le \frac{y_{\rm c}-\widetilde{y}}{\theta}\right\} \tag{27}$$

$$= \left\{\zeta:\ -\widehat{h}_- \le \zeta \le \widehat{h}_+\right\} \tag{28}$$

Let $\Re$ be the set of all real numbers. The complement of Λ_- is $\Lambda_{\rm c} = \Re - \Lambda_-$. Comparing eqs.(20), (24) and (28) we find:

$$\Lambda = \Lambda_+ - \Lambda_{\rm c} \tag{29}$$

We can express the probability of success, for the two-sided target, as:

$$P_{\rm s} = Z(\Lambda) \tag{30}$$

$$= Z(\Lambda_+) - Z(\Lambda_{\rm c}) \tag{31}$$

$$= P_{\rm s+} - (1 - P_{\rm s-}) \tag{32}$$

3.6 Robustness as a Proxy for Probability of Success

We defined the goals, models and uncertainties in section 3.1. In sections 3.2 and 3.3 we defined the robustness and showed that robustness curves cross. In section 3.5 we defined the probability of success of a policy. We now use the statistical concept of 'standardization' to explore the relation between robustness and probability of success. We will show that, when the standardization condition holds, each one-sided robustness function is a proxy for the probability of success. This means that any change in the policy augments the robustness if and only if it also augments (or at least does not reduce) the probability of success. We will see that this is not true for the two-sided robustness function, which is related to the inherent conflict between the two sides in the absolute inequality which the two-sided robustness tries to control. When the proxy property holds, (as it does for the one-sided robustnesses), the policy maker can maximize the probability of success by maximizing the robustness, even if the probability distribution is entirely unknown. The *value* of the probability of success will be unknown, but maximal. This makes the robustness function a powerful tool for policy selection.

3.6.1 One-Sided Robustness

We first introduce a definition. Let x be a scalar random variable with distribution $P(x|q)$ which depends on parameters q. The distribution $P(x|q)$ varies over a class of distributions as q varies over a set of allowed values. For instance, q would contain the mean and variance of a normal distribution, or the least and greatest values of a uniform distribution on the interval $[a,\ b]$. Let ζ be a one-to-one transformation of x whose distribution $Z(\zeta)$ does not depend on the parameters q. For instance, $Z(\zeta)$ could be the standard normal distribution, or the uniform distribution on the interval $[0,\ 1]$.

$P(x|q)$ is a *standardization class,*[5] and x obeys the *standardization condition,* if $Z(\zeta)$ is the same distribution for all valid choices of q. For instance, the normal distributions form a standardization class since every normal variate, when standardized by its mean and variance, becomes a standard normal variate. Similarly the uniform distributions form a standardization class.

In our specific case, $\zeta = (y - \widetilde{y})/\theta$, defined following eq.(20), obeys the standardization condition if its cdf does not depend on the control vector v. That is, the standardization condition requires that, while y depends on the control vector v, the distribution of the standardized variable, $Z(\zeta)$, does not. Many distinct classes of distributions obey standardization conditions. We do not know which class $Z(\zeta)$ belongs to. By assuming standardization in our specific case we are assuming two things. First, we assume that the variability of $y = c^T v$ can be represented by a standardizable distribution. Second, we assume that the info-gap model for uncertainty in the model coefficients c contains enough information about the variability of c to properly standardize y. Specifically, we assume that $\widetilde{y}$ and θ—which result from the info-gap model—do in fact standardize y.

We can now state the following proposition, whose proof appears in appendix A, which asserts that each one-sided robustness, $\widehat{h}_+$ and $\widehat{h}_-$, is a proxy for the corresponding probability of success if ζ obeys standardization. That is, any change in v which augments a one-sided robustness, also augments (or at least does not reduce) the probability of satisfying the corresponding requirement. Likewise, any change in v which reduces the robustness also reduces (or at least does not augment) the probability of success.

Proposition 2 *Each one-sided robustness is a proxy for the probability of one-sided success, if the associated random variable is standardized.*

Given:

• *The linear system of eq.(1).*

• *The info-gap model is linearly symmetric, so the upper and lower robustnesses are given by eqs.(12) and (13).*

• *The random variable $\zeta = (y - \widetilde{y})/\theta$ obeys the standardization condition, so its pdf is independent of the control vector v.*

Then: *The robustness functions $\widehat{h}_+(v, y_{\mathrm{c}})$ and $\widehat{h}_-(v, y_{\mathrm{c}})$ are each a proxy for the corresponding probability of one-sided success. For each element, v_i, of the control vector:*

$$\left(\frac{\partial P_{\mathrm{s}\times}}{\partial v_i}\right)\left(\frac{\partial \widehat{h}_\times(v, y_{\mathrm{c}})}{\partial v_i}\right) \geq 0 \tag{33}$$

where "$\times$" is either "$+$" or "$-$".

3.6.2 Two-Sided Robustness

In this section we will understand why the two-sided robustness, $\widehat{h}$, is generally *not* a proxy for the probability of two-sided success, and furthermore why analysts should be motivated to choose one of the one-sided targets and to use the corresponding robustness function for policy selection. This is a continuation of the discussion of conflicting goals in sections 3.1 and 3.3.

If we assume that the standardization condition holds, then eqs.(32), (102) and (104) (from appendix A) imply:

$$\frac{\partial P_{\mathrm{s}}}{\partial v_i} = \frac{\partial P_{\mathrm{s}+}}{\partial v_i} + \frac{\partial P_{\mathrm{s}-}}{\partial v_i} \tag{34}$$

$$= z(\widehat{h}_+)\frac{\partial \widehat{h}_+}{\partial v_i} + z(-\widehat{h}_-)\frac{\partial \widehat{h}_-}{\partial v_i} \tag{35}$$

We see from eq.(8) that $\partial\widehat{h}/\partial v_i$ equals either $\partial\widehat{h}_+/\partial v_i$ or $\partial\widehat{h}_-/\partial v_i$. Thus a sufficient (though not necessary) condition for $\widehat{h}$ to be a proxy for P_{s} is that $\partial\widehat{h}_+/\partial v_i$ and $\partial\widehat{h}_-/\partial v_i$ have the same algebraic

[5]The idea of a standardization class has been used previously to derive a proxy theorem. See Ben-Haim (2006, section 11.4).

sign. However, from eqs.(12) and (13) one finds:

$$\frac{\partial \widehat{h}_-}{\partial v_i} = -\frac{\partial \widehat{h}_+}{\partial v_i} - \frac{2y_\mathrm{c}}{\theta^2} \tag{36}$$

where we note that $2y_\mathrm{c}/\theta^2 \geq 0$. Thus, if either partial derivative is positive, then the other must be negative and of greater absolute value. We conclude that the condition that both derivatives have the same sign will not always hold, and may frequently fail to hold. In fact, we can understand from eqs.(35) and (36) that $\widehat{h}$ will often *not* be a proxy for the probability of two-sided success. This is because policy-changes which improve success on one side will tend to reduce success on the other side. This motivates the analyst to decide which direction of error is more important, and to use the corresponding one-sided robustness function for policy selection. Proposition 2 asserts that each one-sided robustness *is* a proxy for the corresponding probability of success.

3.7 Learning

The proxy theorem, proposition 2, depends critically on the assumption of standardization. As explained earlier, this in turn depends on the info-gap model containing enough information about the random variability of the model coefficients c in order for ζ to be correctly standardized. (It does not require knowing the identity of the standardization class.) For instance, suppose θ in eq.(10) is the wrong normalization, and the correct normalization is eq.(11). This latter normalization would result from exactly the same analysis with the info-gap model in eq.(4) rather than eq.(3).

The validity of the proxy theorem is what assures that robust-satisficing policies are most likely to succeed. (This is particularly important since, from proposition 1, we know that robust-satisficing policies may differ from expected-model optimal strategies.) The decision maker who is able to learn how to standardize c, will be able to maximize the probability of success. This does not require learning the correct cdf of c, or even what standardization class it belongs to. It is sufficient for the non-probabilistic info-gap model to capture enough information about the variability of c to enable correct standardization of ζ. This minimal learning is a self-reenforcing process, since correct standardization leads the robust-satisficing decision maker to succeed more than any other decision maker.

4 A Proxy Theorem for Non-Linear Models

In this section we consider a situation in which we can prove a much stronger proxy theorem than proposition 2 in section 3.6, but at a cost. By assuming that the uncertainty in the model is contained in a single variable, and that the response, y, is monotonic in this variable, then one-sided robustness is a proxy for probability of success without requiring that the model be linear or that the standardization condition hold. This is of both practical and theoretical importance. Practically, when a proxy theorem holds, policy makers can choose actions which maximize the probability of achieving their goals, without actually knowing the governing probability distributions. Theoretically, the existence of a proxy theorem can explain the persistence of otherwise anomalous behavior: robust-satisficing rather than outcome-optimizing behavior persists under uncertain competition when maximization of robustness is equivalent to maximization of the probability of success. We will see an example in section 6 when we study the equity premium puzzle.

Consider a model whose response, $y(x, v)$, is a scalar-valued function which depends on a single uncertain scalar, x, as well as a vector of control variables v. For instance, this might be the linear model of eq.(1) where the uncertainty in the model coefficients $c(x)$ all derive from uncertainty in x. Or, in the linear model of eq.(1), only one of the model coefficients is uncertain and all the rest are known. However, $y(x, v)$ does not have to be linear at all.

Let $\widetilde{x}$ be our best estimate of x, and consider the following unbounded-interval info-gap model for uncertainty in x:

$$\mathcal{U}(h) = \{x : \ \widetilde{x} - h\sigma_1 \le x \le \ \widetilde{x} + h\sigma_2\}, \quad h \ge 0 \tag{37}$$

where σ_1 and σ_2 are known non-negative numbers.

The robustnesses for the three requirements in eq.(2) are defined in analogy to eqs.(5)–(7):

$$\widehat{h}_+(v, y_{\rm c}) = \max\left\{h : \left(\max_{x\in\mathcal{U}(h)} y(x, v)\right) \le y_{\rm c}\right\} \tag{38}$$

$$\widehat{h}_-(v, y_{\rm c}) = \max\left\{h : \left(\min_{x\in\mathcal{U}(h)} y(x, v)\right) \ge -y_{\rm c}\right\} \tag{39}$$

$$\widehat{h}(v, y_{\rm c}) = \max\left\{h : \left(\max_{x\in\mathcal{U}(h)} |y(x, v)|\right) \le y_{\rm c}\right\} \tag{40}$$

The probabilities of success are defined in terms of the following sets:

$$\Lambda_+ = \{x : \ y(x, v) \le y_{\rm c}\} \tag{41}$$

$$\Lambda_- = \{x : \ y(x, v) \ge -y_{\rm c}\} \tag{42}$$

$$\Lambda = \{x : \ |y(x, v)| \le y_{\rm c}\} \tag{43}$$

Let $F(x)$ denote the cdf of x. The probabilities of success, $P_{\rm s+}$, $P_{\rm s-}$ and $P_{\rm s}$, are evaluated as $F(\Lambda_+)$, $F(\Lambda_-)$ and $F(\Lambda)$, respectively.

We can now state the following proposition, whose proof appears in appendix B.

Proposition 3 *Robustness is a proxy for probability of success if the response depends monotonically on a single scalar uncertainty.*

Given:

- *The system response, $y(x, v)$, is scalar-valued and depends on a single uncertain scalar, x, and a control vector or function v.*
- *$y(x, v)$ is monotonic in x at any fixed v.*
- *The info-gap model for x is eq.(37), which does not depend on the control.*
- *Given two controls, v and v', each with finite one-sided robustness.*

Then: *One-sided robustness is a proxy for probability of success:*

$$\widehat{h}_\times(v', y_{\rm c}) < \widehat{h}_\times(v, y_{\rm c}) \quad \text{implies} \quad P_{\rm s\times}(v') \le P_{\rm s\times}(v) \tag{44}$$

where "$\times$" is either "$+$" or "$-$".

5 Modelling Uncertain Systems

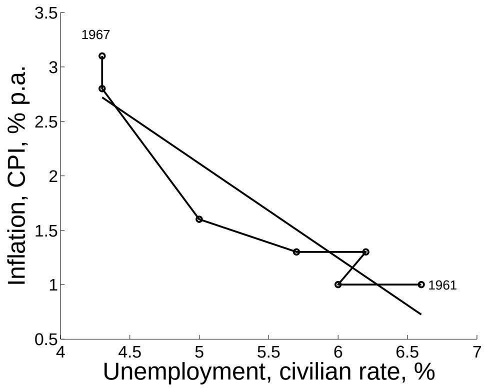

Figure 5: Inflation vs. unemployment in the US, 1961–1967.

Figure 6: Inflation vs. unemployment in the US, 1961–1993.

The trade-off between inflation and unemployment for the US, from 1961 to 1967, is roughly linear, as shown in fig. 5. The least-squares fit shown in the figure has a slope of -0.87 %CPI/%unemployment. Data from 1961 through 1993 are shown in fig. 6. If we consider the years 1980–1983 the slope of the least-squares fit is -3.34, and for the years 1985–1993 the slope is -1.08. In short, the trade-off between inflation and unemployment is often linear over moderate durations but the slope varies substantially over time, and for some periods, for instance 1970–1978, the relation between inflation and unemployment is not linear at all.

Data like this raise several questions. How should we model uncertain economic processes? How can we formulate and calibrate models from historical data to meaningfully predict or describe future behavior? How do we evaluate the confidence in historically calibrated models, when we use them for formulating policy which will be implemented in the future?

We will illustrate an info-gap approach to these questions by studying the estimation of the slope of the Phillips curve based on data from fig. 6. In section 5.1 we use the least-squares estimate of the slope and study its robustness to error in the data. In section 5.2 we explore robust-satisficing estimates—which are sub-optimal in terms of fidelity to the data—but which have acceptable fidelity and greater robustness to uncertainty in the data than the least-squares estimate.

5.1 Robustness of the Least-Squares Estimate

In this section we consider the robustness of the least-squares estimate of the slope of the Phillips curve, to uncertainty in the inflation and unemployment data.

Consider a set of observations of CPI inflation c_i and unemployment u_i, for $i = 1, \ldots, n$, such as shown in figs. 5 and 6. The least-squares estimate of the slope of a linear regression of c_i against u_i is the ratio of the covariance of c and u, to the variance of u:

$$s = \frac{\operatorname{cov}(u, c)}{\operatorname{var}(u)} \tag{45}$$

where:

$$\operatorname{cov}(u, c) \;\; = \;\; \frac{1}{n} \sum_{i=1}^{n} c_i u_i - \left(\frac{1}{n} \sum_{i=1}^{n} c_i \right) \left(\frac{1}{n} \sum_{i=1}^{n} u_i \right) \tag{46}$$

and $\operatorname{var}(u) = \operatorname{cov}(u, u)$.

The data vectors, u and c, are highly uncertain in two senses. First, estimates of past inflation and unemployment can change over time due to revision of the underlying data or modification of the definition of inflation or unemployment. Second, historical values are uncertain estimates of future values because of the possibility of structural changes in the economy. Given observed data, $\widetilde{u}$ and $\widetilde{c}$, we can estimate the slope of the Phillips curve with eq.(45). We would like to know how robust this estimate is to uncertainty in the data. How much could the data change, without unduely altering the estimated trade-off between inflation and unemployment? If the estimate is highly robust to uncertainty then we have confidence in the estimate, while low robustness suggests low confidence.

A full info-gap analysis of robustness would entail formulating an info-gap model for uncertainty in the data vectors u and c. The purpose of the present analysis is to illustrate info-gap estimation, so we will take a simpler route, and consider uncertainty only in the covariance and variance of the data.

Let us denote the estimated variance of u by $\widetilde{\sigma}^2$, and the estimated covariance of c and u by $\widetilde{\gamma}$. Consider the following fractional-error info-gap model for uncertainty in these moments:

$$\mathcal{U}(h) = \left\{ (\gamma, \sigma^2) : \ |\gamma - \widetilde{\gamma}| \le |\widetilde{\gamma}| h, \ \ |\sigma^2 - \widetilde{\sigma}^2| \le \widetilde{\sigma}^2 h, \ \ \sigma^2 \ge 0 \right\}, \quad h \ge 0 \tag{47}$$

At horizon of uncertainty h, the variance and covariance σ^2 and γ can each deviate fractionally from their estimated values by as much as h, though the variance cannot be negative. The horizon of uncertainty is unknown, so $\mathcal{U}(h)$ is an unbounded family of nested sets of moments.

Let $\widetilde{s} = \widetilde{\gamma}/\widetilde{\sigma}^2$ denote the slope based on the observed moments, while $s(\gamma, \sigma^2) = \gamma/\sigma^2$ is the slope if the moments were γ and σ^2. The robustness of the estimate $\widetilde{s}$ is the greatest horizon of uncertainty in the moments up to which the deviation between $\widetilde{s}$ and s is no greater than an acceptable value, $r_{\rm c}$:

$$\widehat{h}(\widetilde{s}, r_{\rm c}) = \max \left\{ h : \ \left(\max_{\gamma, \sigma^2 \in \mathcal{U}(h)} |s(\gamma, \sigma^2) - \widetilde{s}| \right) \le r_{\rm c} \right\} \tag{48}$$

We will consider the typical situation which is that inflation and unemployment are estimated to be anti-correlated (so $\widetilde{\gamma} < 0$), which implies that the estimated slope of the Phillips curve is negative ($\widetilde{s} < 0$). Defining $\rho = -r_{\rm c}/\widetilde{s}$, one finds the robustness to be (see appendix C):

$$\widehat{h}(\widetilde{s}, \rho) = \frac{\rho}{2 + \rho} \tag{49}$$

Figure 7: Robustness of estimated slope, $\widehat{h}(\widetilde{s}, \rho)$, vs. critical error, ρ. Eq.(49).

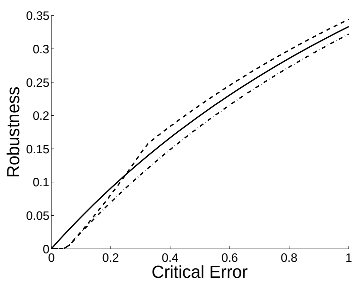

Figure 8: Robustness of estimated slope, $\widehat{h}(s_{\rm e}, \rho)$, vs. critical error, ρ. $\zeta = 1$ (solid), 1.05 (dash), 0.95 (dot-dash).

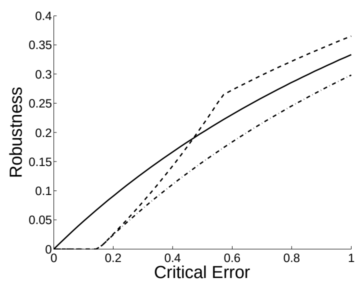

Figure 9: Robustness of estimated slope, $\widehat{h}(s_{\rm e}, \rho)$, vs. critical error, ρ. $\zeta = 1$ (solid), 1.15 (dash), 0.85 (dot-dash).

The robustness function in eq.(49) is plotted in fig. 7. From the figure we see that the robustness is zero when the error is zero: $\widehat{h}(\widetilde{s}, 0) = 0$. That is, the least-squares estimate has no immunity against error in the estimated moments. Furthermore, the slope of the curve is positive which means that positive robustness is obtained only by accepting the possibility of error in the estimate. That is, the positive slope expresses the trade-off between robustness to data-uncertainty and error in the estimate. Moving up the robustness curve, we see that an error of no more than 20% in the estimated slope, (that is, $\rho = 0.2$) is guaranteed if the moments err by no more than 9% (that is, $\widehat{h}(\widetilde{s}, 0.2) = 0.09$).

5.2 Info-Gap Estimate

Can we find a more reliable estimate than the least-squares estimate $\widetilde{s}$? Yes, but at a price. We will study a robust-satisficing estimate of the Phillips curve slope, $s_{\rm e}$, which, at positive estimation error, is more robust to data-uncertainty than the least-squares estimate $\widetilde{s}$.

Let $s_{\rm e}$ denote any estimate of the Phillips curve slope. Define the robustness of this estimate in analogy to eq.(48):

$$\widehat{h}(s_{\rm e}, r_{\rm c}) = \max\left\{h: \left(\max_{\gamma,\sigma^2\in\mathcal{U}(h)} |s(\gamma,\sigma^2) - s_{\rm e}|\right) \le r_{\rm c}\right\} \tag{50}$$

$s_{\rm e}$ is a useful and meaningful estimate of the slope if its robustness, $\widehat{h}(s_{\rm e}, r_{\rm c})$, is large at acceptable estimation error $r_{\rm c}$. When this holds, the moments can vary substantially around the observed values $\widetilde{\gamma}$ and $\widetilde{\sigma}^2$, without $s_{\rm e}$ deviating unacceptably from the true slope $s(\gamma, \sigma^2)$.

Define $\zeta = s_{\rm e}/\widetilde{s}$ and $\rho = -r_{\rm c}/\widetilde{s}$ as before, and continue to assume that $\widetilde{s} < 0$. One finds the robustness to be (see appendix C):

$$\widehat{h}(s_{\rm e}, \rho) = \begin{cases} \dfrac{\rho+\zeta-1}{\rho+\zeta+1} & \text{if } \rho^2 \ge \zeta^2 - 1 \text{ and } \rho \ge 1-\zeta \\[2ex] \dfrac{\rho-\zeta+1}{-\rho+\zeta+1} & \text{if } \rho^2 \le \zeta^2 - 1 \text{ and } \rho \ge \zeta-1 \end{cases} \tag{51}$$

$\widehat{h}(s_{\rm e}, \rho)$ is zero otherwise. Eq.(51) includes eq.(49) as a special case, which arises when $\zeta = 1$. The point to appreciate is that, when $\zeta > 1$, the robustness follows the lower line of eq.(51) (which has greater slope than the robustness curve for $\widetilde{s}$) for small ρ, and then follows the upper line of the equation for larger ρ. This causes crossing of robustness curves as illustrated by the solid and dashed lines in figs. 8 and 9. (The two lines in eq.(51) are equal when $\rho^2 = \zeta^2 - 1$.)

Figs. 8 and 9 show robustness curves for the least-squares estimate $\widetilde{s}$ (solid lines) and for two other estimates $s_{\rm e} \ne \widetilde{s}$. When $s_{\rm e} > \widetilde{s}$ (that is, $\zeta < 1$ since $s_{\rm e}$ and $\widetilde{s}$ are negative) we see that the robustness curve for $s_{\rm e}$ lies below the least-squares robustness curve, so we would never choose such a value of $s_{\rm e}$. However, when $s_{\rm e} < \widetilde{s}$ ($\zeta > 1$) the robustness curves for $s_{\rm e}$ and $\widetilde{s}$ cross at a point which we will denote $(\rho_\times, \widehat{h}_\times)$. For instance in fig. 8, for which the dashed line represents $s_{\rm e} = 1.05\widetilde{s}$, the crossing occurs when $\rho_\times = 0.25$ at a robustness of $\widehat{h}_\times = 0.11$. If we are willing to accept a fractional error in the estimated slope which is no less than 25%, then $s_{\rm e} = 1.05\widetilde{s}$ is more robust to uncertainty in the data than the least-squares estimate, $\widetilde{s}$. The robustness premium of the robust-satisficing estimate is not large (about 10% additional robustness over most of the range of $\rho > \rho_\times$) but nonetheless positive.

Fig. 9 shows robustness curves for robust-satisficing estimates which deviate more from the least-squares estimate. We see greater robustness premium for $s_{\rm e} = 1.15\widetilde{s}$ (about 18%), though the curves cross at larger error ($\rho_\times = 0.47$).

In summary, we recognize three points:

- All robustness curves have positive slope, indicating the trade-off between error and reliability. Small critical error, $r_{\rm c}$, in the estimated slope entails low robustness, $\widehat{h}(s_{\rm e}, r_{\rm c})$, to error in the data. In other words, low critical error in the slope is accompanied by low confidence in limiting the error to that value.
- The least-squares estimate of the slope, $\widetilde{s}$, has smaller mean-squared error than any other estimate, if the data are correct. That is, the robustness curve for $\widetilde{s}$ sprouts off the $r_{\rm c}$-axis further to the left than the robustness curve for any other estimate. However, $\widetilde{s}$ is less robust to data-error than any robust-satisficing estimate $s_{\rm e} < \widetilde{s}$ ($\zeta > 1$), for values of $r_{\rm c}$ in excess of some threshold. That is, the robustness curves for $\widetilde{s}$ and for $s_{\rm e}$ cross one another.

- The choice of an estimate of the Phillips curve slope, $s_{\rm e}$, depends on the acceptable level of error, $r_{\rm c}$, and on the required level of confidence in containing the error to this value as expressed by the robustness to uncertainty in the data, $\widehat{h}(s_{\rm e}, r_{\rm c})$.

Part II
Economic Behavior

6 The Equity Premium Puzzle and Robust-Satisficing

6.1 Introduction

In part I we studied strategies for decision making in economics, looking at monetary policy in section 3 and modelling and estimation in section 5. In the current section we turn to modelling and explaining economic behavior. We will consider multi-period investment decisions, derive an info-gap generalization of the Lucas asset-pricing relations, and show how this generalization points to an explanation of the equity premium puzzle (Mehra and Prescott, 1985). We will show that, for the 1-step investment decision, robustness is a proxy for the probability of successfully satisficing the utility.

Investors cannot confidently maximize their future utility from an investment because of the severe uncertainty in future asset returns. The challenge facing investors is to decide whether adequate returns are sufficiently reliable. If not, then the resources can be invested elsewhere. We model this behavior by assuming that investors choose an investment so as to **satisfice the utility** and **maximize the robustness against uncertainty** in the future returns, rather than to maximize the total discounted utility.

The crux of the matter is the treatment of uncertainty. We use an info-gap model to represent the investor's Knightian uncertainty about future returns: an unmeasurable and non-probabilistic epistemic gap between known past returns and unknown future returns. The info-gap formulation presumes that the investor knows the past returns, believes that future returns may deviate greatly from past experience, and that reliable probabilistic models of these deviations are unavailable. An info-gap model uses an unbounded family of nested sets to represent possible future payoffs. No measure functions are involved.

The investor has little confidence that past returns reflect future behavior reliably. The investor's central question, which motivates our decision model, is: how wrong can the current estimate of future returns be, without jeopardizing the attainment of a specified level of utility? The answer to this question—the robustness function—generates preferences on options, without requiring probabilistic information.

We present a multi-period two-asset implementation which provides insight into the equity premium puzzle. In section 6.2 we formulate the investment model, the info-gap model of uncertainty in future returns, the robustness function and the robust-satisficing decision strategy. In section 6.3 we derive asset-pricing relations which are the info-gap generalization of the Lucas asset-pricing model. In section 6.4 we derive an expression for the equity premium. In section 6.5 we show that the robustness is a proxy for the probability of successfully satisficing the utility. We conclude with a discussion in section 6.6.

6.2 Dynamics, Uncertainty and Robustness

We consider discrete time, $t = 0, 1, \ldots, T$. c_t is the total consumption at time step t and $u(c_t)$ is the utility from this consumption. We assume that $u(c_t)$ is continuous and that the marginal utility is positive: $u'(c_t) > 0$.

We consider two assets, one a risky stock ($i = 1$) and the other a risk-free bond ($i = 2$). The generalization to more than two assets is straightforward and would not entail any substantive alteration of our conclusions.

x_{it} is the quantity of asset i held between t and $t+1$ by a representative agent and can be either positive or negative. The holdings at time t are $x_t = (x_{1t}, x_{2t})^T$. The holdings throughout the time horizon are $x = (x_0, x_1, \ldots, x_{T-1})$. x is chosen by the investor.

p_{it} is the ex-dividend price of asset i at time t, where $p_t = (p_{1t},\, p_{2t})^T$. d_{it} is the dividend of asset i at time t. The prices throughout the time horizon are $p = (p_0,\, p_1, \ldots,\, p_{T-1})$.

$y_{it} = p_{it} + d_{it}$ is the payoff of asset i at time t, where $y_t = (y_{1t},\, y_{2t})^T$. y_0 is known, y_{1t} is risky and uncertain while y_{2t} is risk-free and known at $t = 0$ for all $t \geq 0$. The uncertain payoffs are $y_{(1)} = (y_{11},\, \ldots,\, y_{1T})$. y_{1t} can be either positive or negative. (Requiring non-negative payoff would alter our analysis only at very large robustness, which we will see is not of practical interest for understanding the equity premium puzzle.)

The budget constraint is:

$$c_t + p_t^T x_t = y_t^T x_{t-1}, \quad t = 0,\, \ldots,\, T \tag{52}$$

The initial endowment, x_{-1} and y_0, are known at time 0. The investment in the last step is zero: $x_T = 0$. Short sells are allowed but c_t cannot be negative. The choice variables are $x_0,\, \ldots,\, x_{T-1}$, which determine the consumptions through the budget constraints.

The payoff of the risky asset, y_{1t}, is uncertain for $t > 0$, and $\widetilde{y}_{1t}$ is the best known estimate of y_{1t} at time 0, where σ_{1t} is an error of this estimate. We assume that the anticipated payoff is positive: $\widetilde{y}_{1t} > 0$. y_{10} is known. The investor views $\widetilde{y}_{1t}$ as a best but rough estimate of future payoffs, without knowing how wrong this estimate will turn out to be. We use the unbounded fractional-error info-gap model to represent uncertainty in the risky-asset payoffs:

$$\mathcal{Y}(h, \widetilde{y}) = \left\{ y_{(1)} : \ |y_{1t} - \widetilde{y}_{1t}| \leq h\sigma_{1t}, \ \ t = 1,\, \ldots,\, T \right\}, \quad h \geq 0 \tag{53}$$

$\mathcal{Y}(h, \widetilde{y})$ is the set of risky payoffs y_{1t}, for $t = 1,\, \ldots,\, T$, whose fractional deviations from the anticipated payoffs $\widetilde{y}_{1t}$ are no greater than h. The fractional error, h, is unknown, so the horizon of uncertainty is unbounded. This info-gap model is not a single interval, but rather an unbounded family of nested payoff intervals.

The discounted utility up to time T is:

$$U(x, y) = \sum_{t=0}^{T} \beta^t u(c_t) \tag{54}$$

The investor desires to choose the asset holdings x throughout the time horizon so as to attain discounted utility no less than $\overline{u}$. That is, the investor wishes to satisfice the discounted utility at the value $\overline{u}$:

$$U(x, y) \geq \overline{u} \tag{55}$$

$\overline{u}$ can be thought of as a 'reservation utility'. The investment will be pursued if the investor has adequate confidence in achieving adequate reward.

The **robustness** to uncertainty in the payoffs, of holdings x with utility-aspiration $\overline{u}$, is the greatest horizon of uncertainty h up to which all payoffs yield at least the desired utility:

$$\widehat{h}(x, \overline{u}) = \max \left\{ h : \ \left(\min_{y_{(1)} \in \mathcal{Y}(h, \widetilde{y})} U(x, y) \right) \geq \overline{u} \right\} \tag{56}$$

The set of h-values in this definition is empty if $U(x, \widetilde{y}) < \overline{u}$, meaning that holdings x do not attain utility $\overline{u}$ with the anticipated payoffs $\widetilde{y}$. In this case we define $\widehat{h}(x, \overline{u}) = 0$ and we say that utility aspiration $\overline{u}$ is 'infeasible'. Any other $\overline{u}$ is 'feasible'.

We will be interested in robust-satisficing investments: those which satisfice the discounted utility and maximize the robustness:

$$\widehat{x}(\overline{u}) = \arg\max_x \widehat{h}(x, \overline{u}) \tag{57}$$

where the maximum on x is subject to the budget constraint with non-negative consumption.

6.3 Asset-Pricing Relation

Consider the time horizon $t = 0, 1, \ldots, T$. The investor has T choice vectors $x_0 = (x_{10}, x_{20})^T, \ldots, x_{T-1} = (x_{1,T-1}, x_{2,T-1})^T$. The uncertainties are the unknown risky payoffs $y_{(1)} = (y_{11}, \ldots, y_{1T})^T$. The info-gaps in these payoffs are described by $\mathcal{Y}(h, \widetilde{y})$ in eq.(53) where the horizon of uncertainty h is unknown. The risk-free payoffs $(y_{20}, \ldots, y_{2T})$ are known. Using the budget constraint in eq.(52), the discounted utility in eq.(54) is:

$$U(x, y) = \sum_{t=0}^{T} \beta^t u(y_{1t}x_{1,t-1} + y_{2t}x_{2,t-1} - p_t^T x_t) \tag{58}$$

x_{1t} can be either positive or negative. Let $\phi_t = 1$ if $x_{1t} \geq 0$ and $\phi_t = -1$ otherwise for $t \geq 0$. Define $\phi_{-1} = 0$. The marginal utility is positive, so the lowest utility up to horizon of uncertainty h occurs when the risky payoffs y_{1t} are such that $y_{1t}x_{1,t-1}$ is minimal, for $t = 1, \ldots, T$. (Recall that y_{10} is known.) The risky payoffs which minimize the utility at uncertainty h are $y_{1t} = \widetilde{y}_{1t} - h\phi_{t-1}\sigma_{1t}$. Thus the minimum in the definition of the robustness, eq.(56), is:

$$\begin{aligned} \mu(h, x, p) &= \min_{y_{(1)} \in \mathcal{Y}(h, \widetilde{y})} U(x, y) && (59) \\ &= \sum_{t=0}^{T} \beta^t u[\underbrace{\widetilde{y}_{1t}x_{1,t-1} - h\phi_{t-1}\sigma_{1t}x_{1,t-1} + y_{2t}x_{2,t-1} - p_t^T x_t}_{\widetilde{c}_t(h)}] && (60) \end{aligned}$$

which defines $\widetilde{c}_t(h)$, the lowest anticipated consumption up to uncertainty h.

The utility aspiration $\overline{u}$ is feasible if the righthand side of eq.(60) is no less than $\overline{u}$ in the absence of uncertainty ($h = 0$). Because the marginal utility is positive and $h\phi_{t-1}\widetilde{y}_{1t}x_{1,t-1}$ is also positive, the righthand side of eq.(60) decreases strictly monotonically as h increases. Hence, for any feasible aspiration, the robustness is the value of h at which the righthand side of eq.(60) equals $\overline{u}$:

$$\begin{aligned} \overline{u} &= u(\underbrace{y_0^T x_{-1} - p_0^T x_0}_{c_0}) && (61) \\ &\quad + \sum_{t=1}^{T} \beta^t u\left[\underbrace{\widetilde{y}_{1t}x_{1,t-1} - \widehat{h}(x, \overline{u})\phi_{t-1}\sigma_{1t}x_{1,t-1} + y_{2t}x_{2,t-1} - p_t^T x_t}_{\widetilde{c}_t(\widehat{h})}\right] \end{aligned}$$

If eq.(61) holds for a continuum of investments x, then its derivatives with respect to x_{1t} and x_{2t} also hold. Suppose there is a robust-satisficing investment which maximizes the robustness without constraint (the existence of such investments is studied in Ben-Haim, 2006, section 11.5.5):

$$\frac{\partial \widehat{h}}{\partial x_{it}} = 0, \quad i = 1, 2;\ t = 0, \ldots, T-1 \tag{62}$$

Differentiating eq.(61) with respect to x_{1t} (which we assume differs from zero) and x_{2t}, for $t = 0, \ldots, T-1$, and using eq.(62) yields:

$$\begin{aligned} \frac{\mathrm{d}u(c_t)}{\mathrm{d}c_t} p_{1t} &= \beta \frac{\mathrm{d}u(c_{t+1})}{\mathrm{d}c_{t+1}} (\widetilde{y}_{1,t+1} - \widehat{h}\phi_t\sigma_{1,t+1}) && (63) \\ \frac{\mathrm{d}u(c_t)}{\mathrm{d}c_t} p_{2t} &= \beta \frac{\mathrm{d}u(c_{t+1})}{\mathrm{d}c_{t+1}} y_{2,t+1} && (64) \end{aligned}$$

where $c_t = \widetilde{c}_t(\widehat{h})$ is the argument of the utility function in eq.(61). These relations are the info-gap generalizations of the first-order conditions in the Lucas asset-pricing model (Blanchard and Fischer, 1989, eq.(11), p.511). The ordinary Lucas relations result when $\widehat{h} = 0$.

Eq.(63) asserts that the risky-asset price p_{1t} depends not only on the marginal rate of substitution as in the Lucas model, but also depends on the required robustness $\widehat{h}$. At low risk aversion the marginal utility is nearly constant, so the dominant term on the righthand of eq.(63) is $1-\widehat{h}\phi_t\sigma_{1,t+1}$. When $\phi_t=+1$ (e.g. the representative agent has positive risky holdings), this means that the risky-asset price tends to decrease as the robustness which the investor requires increases. On the other hand, for short sales of the risky asset (so $\phi_t=-1$), the price tends to increase as the robustness which the investor requires increases.

The basic robustness trade-off theorem asserts that robustness decreases as aspiration increases: $\widehat{h}(x,\overline{u})$ decreases as $\overline{u}$ increases. This holds both for arbitrary investments x and for the robust-satisficing investments $\widehat{x}(\overline{u})$ in eq.(57). Furthermore, the robustness vanishes at the greatest feasible aspiration: $\widehat{h}(x,\overline{u})=0$ if $\overline{u}=U(x,\widetilde{y})$. (These results derive from the nested structure of info-gap models, e.g. $\mathcal{Y}(h,\widetilde{y})$ in eq.(53), and the definition of the robustness, eq.(56).) This trade-off between robustness and utility-aspiration is illustrated schematically in fig. 10. Fig. 11 illustrates eq.(63) for low risk aversion and $\phi_t=+1$: the risky-asset price tends to decrease as the demanded robustness increases. Consequently, combining figs. 10 and 11, the risky-asset price rises as investor aspirations increase as shown in fig. 12.

The solid dots in figs. 10–12 identify corresponding points. In fig. 10 the robustness vanishes at the maximal utility aspiration. In fig. 11 we see that zero robustness corresponds to maximal risky-asset price, which is the analog of the Lucas asset price. p_{10}^{c} represents the price when the aspiration is maximal and the robustness is zero. In fig. 12 this maximal price corresponds to maximal utility aspiration. Sub-maximal utility-aspiration (equivalently: positive robustness-aspiration) will force a drop in the price of risky assets held by the representative agent ($\phi_t=+1$).

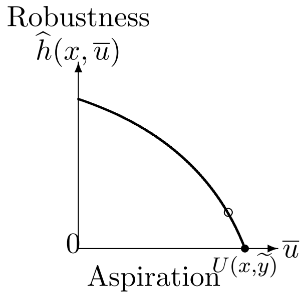

Figure 10: Trade-off of robustness against utility aspiration.

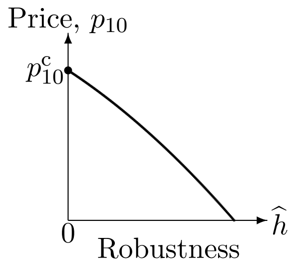

Figure 11: Price of risky asset vs. robustness, eq.(63), for $\phi_t=+1$.

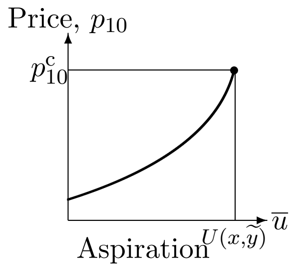

Figure 12: Price of risky asset vs. utility aspiration for $\phi_t=+1$.

6.4 Equity Premium

Define the anticipated or 'best-estimate' rates of return to the risky asset: $\widetilde{r}_{1t}=\widetilde{y}_{1t}/p_{1,t-1}$, for $t=1,\ldots,T$. Likewise the known rates of return to the risk-free asset are $r_{2t}=y_{2t}/p_{2,t-1}$. With these definitions we can write eqs.(63) and (64), for $t=0,\ldots,T-1$, as:

$$u'(c_t) = \beta u'(c_{t+1})(\widetilde{r}_{1,t+1}-\widehat{h}\phi_t\sigma_{1,t+1}/p_t) \tag{65}$$

$$u'(c_t) = \beta u'(c_{t+1})r_{2,t+1} \tag{66}$$

We will use these relations to show how the info-gap robust-satisficing decision paradigm provides insight into the equity premium puzzle. Further discussion is found in Ben-Haim (2006, section 11.5).

Subtracting eq.(66) from eq.(65) results in the info-gap generalization of a basic CAPM relation (Blanchard and Fischer, 1989, eq.(4), p.507):

$$0=\beta u'(c_{t+1})(\widetilde{r}_{1,t+1}-r_{2,t+1}-\widehat{h}\phi_t\sigma_{1,t+1}/p_t), \quad t=0,\ldots,T-1 \tag{67}$$

This relation can be re-arranged (assuming $\beta u'(c_{t+1})\neq 0$) to show that the premium for the risky asset is proportional to the required robustness:

$$\widetilde{r}_{1,t+1}-r_{2,t+1}=\widehat{h}\,\phi_t\sigma_{1,t+1}/p_t \tag{68}$$

Investors who do not require robustness to uncertainty ($\widehat{h} = 0$) also do not need a premium to attract them to have positive holdings on risky assets ($\phi_t = +1$). As the investor becomes more sensitive to Knightian uncertainty in the returns (as $\widehat{h}$ increases), a greater premium is needed to induce investment in both risky and risk-free assets.

We can now understand the essence of the equity premium puzzle. By assuming that investors aspire to maximum earnings (which underlies the Lucas pricing equations) one also (unavoidably) assumes that investors accept zero robustness (the solid dot in fig. 10). This forces the equity premium to zero as shown in eq.(68).

Investors need not be terribly sensitive to Knightian uncertainty in order to explain the usual equity premium. For example, if $\widetilde{r}_{1,t+1} = 1.07$, $r_{2,t+1} = 1.01$ and $\sigma_{1,t+1}/p_t = 0.2$, then the robustness in eq.(68) which explains this 6% premium is $\widehat{h} = 0.06/0.2 = 0.3$, or 30% of the estimated errors $\sigma_{1,t}$. This is a modest robustness in light of the posited variation of risky returns. The monotonic trade-off between robustness and aspiration for utility, illustrated in fig. 10, and the positive robustness which explains the equity premium, indicates that investors reduce their aspirations only slightly below the maximum. This is illustrated schematically by the open circle on the robustness curve of fig. 10. This reduction in aspiration, in response to Knightian uncertainty in the returns, is responsible for the observed premium for the risky asset.

We have assumed nothing about the magnitude of Arrow-Pratt risk aversion. The only assumptions concerning the utility function $u(c)$ are that it is continuous and that the marginal utility is positive. We have assumed time-separation of the discounted utility $U(x, y)$. The time horizon T is arbitrarily large (or small).

6.5 Robustness and the Probability of Success

In this section we consider a 1-step investment decision, and use proposition 3 to show that the robustness is a proxy for the probability of successfully satisficing the utility. This shows that the robust-satisficing investment strategy is at least as likely to achieve the specified goal as any other strategy. If the economic survival of the investor depends on achieving the specified goal, then the robust-satisficing investor will tend to survive when competing against investors who use other strategies.

Let $T = 1$ in the investment problem formulated in section 6.2. The single decision to be made is the choice of the risky and risk-free investments at time $t = 0$, $x_0 = (x_{10},\, x_{20})^T$. The single uncertainty is the risky payoff at time $t = 1$, y_{11}. The discounted utility, eq.(58), is:

$$U(x_0, y_{11}) = u(y_{10}x_{1,-1} + y_{20}x_{2,-1} - p_0^T x_0) + \beta u(y_{11}x_{10} + y_{21}x_{20}) \tag{69}$$

where $p_1^T x_1 = 0$ since no purchase is made at $t = 1$. Without any loss of generality we define a shifted utility as:

$$U_{\mathrm{s}}(x_0, y_{11}) = U(x_0, y_{11}) - 2\overline{u} \tag{70}$$

Now we can write the robustness of eq.(56) as:

$$\widehat{h}(x_0, \overline{u}) = \max \left\{ h : \left(\min_{y_{11} \in \mathcal{Y}(h, \widetilde{y})} U_{\mathrm{s}}(x_1, y_{11}) \right) \geq -\overline{u} \right\} \tag{71}$$

The robustness in eq.(71) is in the form of a lower robustness function in eq.(6). The discounted utility $U_{\mathrm{s}}(x_0, y_{11})$ depends on a single scalar uncertainty, the uncertain payoff y_{11}, and on the decision vector x_0. If we assume positive marginal utility, $u'(c) > 0$, then the discounted utility $U_{\mathrm{s}}(x_0, y_{11})$ increases monotonically as the uncertain payoff y_{11} increases when x_0 is fixed. The info-gap model of eq.(53) does not depend on the decision vector. There are many choices of investment for which the robustness, $\widehat{h}(x_0, \overline{u})$, is finite. Thus the conditions of proposition 3 hold. Consequently, any change in the investment which enhances the robustness also enhances (or at least does not reduce) the probability of satisficing the utility at the required value. In short, robustness for the 1-step investment is a proxy for the probability of success.

6.6 Discussion

Kocherlakota (1996, p.67) concludes his discussion of the equity premium puzzle with the comment:

> The *universality* of the equity premium tells us that, like money, the equity premium must emerge from some primitive and elementary features of asset exchange that are probably best captured through extremely stark models. With this in mind, we cannot hope to find a resolution to the equity premium puzzle by continuing in our current mode of patching the standard models of asset exchange with transactions costs here and risk aversion there. Instead, we must seek to identify what fundamental features of goods and asset markets lead to large risk adjusted price differences between stocks and bonds.

The info-gap robust-satisficing decision model offers the possibility of new insight. Three ideas are prominent.

First, risk aversion is a multi-faceted phenomenon which is not captured entirely by utility-function curvature. Especially in situations of great Knightian uncertainty, where betting is not a particularly useful concept because probabilities are unknown, risk aversion is expressed in part by the willingness to forego utility in exchange for robustness-to-failure. This trade-off is expressed by the negative slope of the robustness curve in fig. 10.

Second, the info-gap robust-satisficing decision model demonstrates that the classical axiom of utility maximization can be relaxed without losing touch with economic intuition and data. The fundamental psychological premise—that utility is desirable—does not imply that maximal utility is most desirable. In this section we have shown that realistic and sensible economic reasoning can be based on, and modelled by, the concept of satisficing the utility and maximizing the robustness with non-probabilistic info-gap models of uncertainty.

Third, we have demonstrated that 1-step investments which maximize the robustness also maximize the probability of successfully satisficing the utility. This means that the robust-satisficing strategy will tend to survive in competition with any other strategy, when survival depends on achieving specified utility outcome.

7 Demand Theory: Satisficing and Windfalling

7.1 Introduction

The aim of this section is two-fold. One goal is to continue our study of economic behavior from the perspective of info-gap theory, demonstrating the economic plausibility of info-gap reasoning, despite its fundamental deviation from the neo-classical rationality of optimization. One of the most basic economic intuitions is expressed in the relation between the price of a good and the quantity which is demanded. The fundamental and nearly ubiquitious trade-off between price and quantity is the starting point of much micro-economic theory. We will show that info-gap theory can re-create the basic properties of Hicksian and Walrasian demand curves in the context of decision under severe Knightian uncertainty.

Our second goal in this section is to introduce the second info-gap decision strategy—opportune-windfalling—which supplements the primary strategy of robust-satisficing which we have used in previous sections. We will see both strategies fitting nicely into an info-gap theory of demand: opportune-windfalling generates the Walrasian while robust-satisficing generates the Hicksian demand function.

This section is rather technical in order to prove that the info-gap demand functions have the properties characteristic of neo-classical demand functions. The main findings are these:

- **Classical choice problems** begin with a utility function which defines complete transitive preferences on the options (section 7.2.1). Then two decision problems are defined.
  - The *expenditure minimization problem* (EMP) seeks a consumption bundle which minimizes the expenditure and guarantees that the utility is no less than a critical value. A consumption bundle which satisfies the EMP is a Hicksian demand function, which obeys:
    - Price-demand trade-off: price-increases are associated with reduced demand.
    - No excess utility: the utility requirement is exactly satisfied.
    - Wealth compensation: as prices change, the Hicksian gives the demand which would arise if wealth were simultaneously adjusted to keep the utility constant.
  - The *utility maximization problem* (UMP) seeks a consumption bundle which maximizes the utility and guarantees that the expenditure is no greater than the agent's wealth. A consumption bundle which satisfies the UMP is a Walrasian demand function.

  The Hicksian and Walrasian demand functions are closely related. A consumption bundle which satisfies one of these choice problems (either EMP or UMP) also satisfies the other when the parameters (critical utility or wealth) are equivalent.

- **Info-gap choice problems** begin with an info-gap model for uncertainty in the reward resulting from a consumption bundle (section 7.2.2). Two decision strategies are studied (section 7.2.3):
  - The *robustness* of a consumption bundle is the greatest horizon of uncertainty up to which the reward resulting from this choice is no less than a critical value, which is less than the best-estimate of the reward.
  - The *opportuneness* of a consumption bundle is the lowest horizon of uncertainty which is needed to enable the reward resulting from this choice to be no less than a windfall value, which is greater than the best-estimate of the reward.

  These two decision strategies generate two choice problems (section 7.2.4):

  - The *robust-satisficing problem* (RSP) seeks a consumption bundle which minimizes the expenditure and satisfices the reward at a specified level of robustness.

    The RSP is conceptually similar to the EMP: both minimize expenditure and satisfice outcome-related entities. The consumption bundle which solves the RSP displays the same characteristics as the Hicksian demand function:

  * Price-demand trade-off: price-increases are associated with reduced demand (proposition 4, section 7.3).
  * No excess utility or robustness: the robustness requirement is exactly satisfied at the specified critical utility (lemma 1, section 7.4).
  * Wealth compensation: as prices change, the robust-satisficing consumption gives the demand which would arise if wealth were simultaneously adjusted to keep the utility constant (eq.(95), section 7.4).
- The *opportune-windfalling problem* (OWP) seeks a consumption bundle which maximizes the opportuneness for windfall, while keeping the expenditure within the agent's wealth.

  The OWP is conceptually similar to the UMP: both attempt to facilitate wonderful outcomes while constraining the expenditure to a specified budget.

The demand functions obtained by solving the RSP and the OWP are closely related. A consumption bundle which satisfies one of these choice problems (either RSP or OWP) also satisfies the other when the parameters are equivalent (propositions 5 and 6, section 7.4).

The most fundamental conclusion is that one can derive a theory of consumer demand which displays the fundamental characteristics of the neo-classical theory, without requiring utility optimization. The info-gap theory of demand is based on robust-satisficing and opportune-windfalling strategies which are motivated by the severe uncertainties which face the consumer.

Our discussion is summarized in section 7.5. All proofs appear in appendix D.

7.2 Consumer Choice Problems

In section 7.2.1 we briefly review the utility maximization and expenditure minimization problems from which demand theory arises, and discuss the intuitions which underlie these strategies of choice. In section 7.2.2 we briefly define info-gap models in the format which will be needed here. Section 7.2.3 is devoted to a discussion of the info-gap robustness and opportuneness functions. We employ these decision functions in section 7.2.4 to formulate two info-gap decision problems—robust satisficing and opportune windfalling—and explain their analogy to the traditional choice problems.

Notation: if $x \in \Re^L$, then $x > 0$ means $x_i > 0$ for all $i = 1, \ldots, L$. Also, $x \geq 0$ means $x_i \geq 0$ for all $i = 1, \ldots, L$.

7.2.1 Classical Choice Problems

Our discussion follows Mas-Colell, Whinston and Green (1995). A preference relation $\succeq$ on $X \subset \Re^L$ is **rational** if the following properties hold:

(i) **Completeness:** For all $x, y \in X$, $x \succeq y$ or $y \succeq x$ or both.

(ii) **Transitivity:** For all $x, y, z \in X$, if $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

A scalar function $u(x)$ is a **utility function** representing the preference relation $\succeq$ if, for all $x, y \in X$, $x \succeq y$ if and only if $u(x) \geq u(y)$.

Given prices $p > 0$ and utility $u > u(0)$, the **expenditure minimization problem** (EMP) is to choose a consumption bundle x for which:

$$\min_{x \geq 0} p^T x \quad \text{subject to} \quad u(x) \geq u \tag{72}$$

Given prices $p > 0$ and wealth $w > 0$, the **utility maximization problem** (UMP) is to choose x such that:

$$\max_{x \geq 0} u(x) \quad \text{subject to} \quad p^T x \leq w \tag{73}$$

A vector x which solves the EMP is a Hicksian demand function, $H(p, u)$, while a solution of the UMP is a Walrasian demand function, $x(p, w)$.

A central result in demand theory states that, if $u(x)$ is a continuous utility function representing a locally non-satiated preference relation, then if x^* is an optimal choice for the UMP with wealth $w > 0$, then it is optimal for the EMP when the required utility is $u(x^*)$. Likewise, if x^* is an optimal choice for the EMP when the required utility is $u > u(0)$, then it is optimal for the UMP with wealth $w = p^T x^*$ (Mas-Colell *et al,* Proposition 3.E.1, p.58). In otherwords, these two choice problems coalesce for appropriate choices of the parameters.

Nonetheless, different intuitions motivate these two selections of a consumption vector x. This is important for our study since it is the intuitions which we wish to preserve in the info-gap context.

The UMP is an ambitious program of maximizing satisfaction subject to a budget constraint. The UMP "squeezes the orange" as much as possible. And in fact any solution x^* satisfies Walras' law, $p^T x^* = w$, meaning that the budget is exploited to the limit (Mas-Colell *et al,* Proposition 3.D.2, pp.51–52).

The EMP, on the other hand, is motivated by (possibly cautious) satisficing. The decision maker chooses a (possibly modest) level of utility u and pinches the budget as much as possible to just achieve this level of utility. And in fact any solution x^* does exactly that: $u(x^*) = u$; there is no excess utility (Mas-Colell *et al,* Proposition 3.E.3, p.61).

7.2.2 Info-Gap Models of Uncertainty

Events are represented as vectors or vector functions f. Knowledge-deficiency is expressed at two levels by info-gap models. For fixed h the set $\mathcal{F}(h, \widetilde{f})$ represents a degree of variability of f around the centerpoint $\widetilde{f}$. The greater the value of h, the greater the range of possible variation, so h is called the *horizon of uncertainty* and expresses the information gap between what is known ($\widetilde{f}$ and the structure of the sets) and what needs to be known for an ideal solution (the exact value of f). The value of h is usually unknown, which constitutes the second level of imperfection of knowledge: the horizon of uncertainty is unbounded.

We will use a somewhat more structured form of info-gap model than used in previous sections.

Let $\Re_+$ denote the non-negative real numbers and let Ω be a Banach space in which the uncertain quantities f are defined. An info-gap model $\mathcal{F}(h, \widetilde{f})$ is a map from $\Re_+ \times \Omega$ into the power set of Ω. Info-gap models obey four axioms, only the first two of which we have used previously in this paper. *Nesting:* $\mathcal{F}(h, \widetilde{f}) \subseteq \mathcal{F}(h', \widetilde{f})$ if $h \le h'$. *Contraction:* $\mathcal{F}(0, 0)$ is the singleton set $\{0\}$. *Translation:* $\mathcal{F}(h, \widetilde{f})$ is obtained by shifting $\mathcal{F}(h, 0)$ from the origin to $\widetilde{f}$: $\mathcal{F}(h, \widetilde{f}) = \mathcal{F}(h, 0) + \widetilde{f}$. *Linear expansion:* info-gap models centered at the origin expand linearly: $\mathcal{F}(h', 0) = \frac{h'}{h}\mathcal{F}(h, 0)$ for all h, $h' > 0$. Nesting is the most characteristic of the info-gap axioms. It expresses the intuition that possibilities expand as the info-gap grows. For more discussion of these axioms see Ben-Haim (1999).

7.2.3 Robustness and Opportuneness

We now present two info-gap decision problems and explain their conceptual proximity to the intuitions underlying the classical EMP and UMP.

The decision maker will choose a commodity vector $x \in X \subseteq \Re^L$. The outcome of this choice is influenced by an unknown vector (or vector function) $f \in \Omega$ whose range of possible variation is represented by an info-gap model $\mathcal{F}(h, \widetilde{f})$, $h \ge 0$, in a Banach space Ω.

We will define two real-valued reward functions defined on (x, A) where $x \in X$ and $A \subset \Omega$. The **lower reward function** $\mathcal{R}_*(x, A)$ is monotonically decreasing on the power set of Ω, at fixed choice x:

$$A \subset B \quad \text{implies} \quad \mathcal{R}_*(x, A) \ge \mathcal{R}_*(x, B) \tag{74}$$

The **upper reward function** $\mathcal{R}^*(x, A)$ is monotonically increasing on the power set of Ω, at fixed choice x:

$$A \subset B \quad \text{implies} \quad \mathcal{R}^*(x, A) \le \mathcal{R}^*(x, B) \tag{75}$$

The monotonicity of these reward functions does not relate to the commodity vector x, or to the decision maker's preferences regarding these choices. The reward functions are monotonic in the space Ω of unknown auxiliary events f, and expresses the impact of this ambient variation on the available outcomes of a choice. Properties (74) and (75) do not establish preference relations on x.

The most common realizations of the lower and upper reward functions is in terms of least and greatest available rewards as f varies within a set $A \subset \Omega$. Let $r(x, f)$ be the reward actually realized when the choice is $x \in X$ and the unknown quantity takes the value $f \in \Omega$. Common choices of lower and upper reward functions are:

$$\mathcal{R}_*(x, A) = \min_{f \in A} r(x, f) \tag{76}$$

$$\mathcal{R}^*(x, A) = \max_{f \in A} r(x, f) \tag{77}$$

Thus $\mathcal{R}_*(x, A)$ is the least available reward, while $\mathcal{R}^*(x, A)$ is the greatest available reward, in the uncertain environment A. In this realization, $\mathcal{R}_*$ and $\mathcal{R}^*$ still have not determined a preference relation on x, both because the set A is undetermined and because the lower and upper rewards may behave differently.

More specifically, we will subsequently choose the set A as a set $\mathcal{F}(h, \widetilde{f})$ in an info-gap family of nested sets. Then, in eqs.(76) and (77), $\mathcal{R}_*[x, \mathcal{F}(h, \widetilde{f})]$ is the least accessible reward up to info-gap h, while $\mathcal{R}^*[x, \mathcal{F}(h, \widetilde{f})]$ is the greatest accessible reward up to h.

Like the classical utility function $u(x)$, the reward functions $\mathcal{R}_*(x, A)$ and $\mathcal{R}^*(x, A)$ are real-valued and represent desirable reward. The decision maker prefers more rather than less. Unlike $u(x)$, however, neither $\mathcal{R}_*(x, A)$ nor $\mathcal{R}^*(x, A)$ need be continuous or convex, (though we will sometimes assume continuity), nor do they derive from preference relations. Moreover, the decision maker can't know the values of $\mathcal{R}_*(x, A)$ and $\mathcal{R}^*(x, A)$ which will be realized in practice because the set A is unknown. Finally, while $\mathcal{R}_*$ and $\mathcal{R}^*$ entail knowledge-deficiency, they are not probabilistic but instead depend on an info-gap model.

We will now use $\mathcal{R}_*(x, A)$ and $\mathcal{R}^*(x, A)$ to define two decision functions (Ben-Haim, 2006) which generate preferences on values of x. These preferences are not unique, nor are the same preferences necessarily derived from each of the two decision functions. These preferences will vary with aspiration level, price, wealth and possibly other exogenous factors. These two decision functions are motivated by intuitions which are similar to those underlying the classical decision problems, EMP and UMP.

Let $r_{\rm c}$ be a value of reward which the decision maker strives to achieve; more reward would be better, but less than $r_{\rm c}$ would be unacceptable. The decision maker wishes to satisfice at reward level $r_{\rm c}$. The decision maker might choose or contemplate $r_{\rm c}$-values in the same way that minimal utility values u are chosen in the EMP of eq.(72). The **robustness** of choice x is the greatest level of knowledge-deficiency at which reward no less than $r_{\rm c}$ is guaranteed:

$$\widehat{h}(x, r_{\rm c}) = \max \left\{ h : \ \mathcal{R}_*[x, \mathcal{F}(h, \widetilde{f})] \geq r_{\rm c} \right\} \tag{78}$$

$\widehat{h}(x, r_{\rm c})$ is a robust satisficing decision function.

Let $r_{\rm w}$ be a large value of reward (much greater than $r_{\rm c}$) which the decision maker would be delighted to achieve; lower reward would be acceptable, but reward as large as $r_{\rm w}$ is a windfall success. The decision maker might evaluate or contemplate $r_{\rm w}$-values much as values of the maximum utility obtained in the UMP of eq.(73) is evaluated. The **opportuneness** inherent in choice x is the least level of knowledge-deficiency at which windfall can occur:

$$\widehat{\beta}(x, r_{\rm w}) = \min \left\{ h : \ \mathcal{R}^*[x, \mathcal{F}(h, \widetilde{f})] \geq r_{\rm w} \right\} \tag{79}$$

$\widehat{\beta}(x, r_{\rm w})$ is an opportune windfalling decision function.

$\widehat{h}(x, r_{\rm c})$ and $\widehat{\beta}(x, r_{\rm w})$ are **immunity functions**. $\widehat{h}(x, r_{\rm c})$ is the immunity against failure (reward less than $r_{\rm c}$). When $\widehat{h}(x, r_{\rm c})$ is large, failure can occur only at great ambient uncertainty; the decision maker is not vulnerable to pernicious uncertainty. Similarly, $\widehat{\beta}(x, r_{\rm w})$ is the immunity against windfall (reward no less than $r_{\rm w}$). When $\widehat{\beta}(x, r_{\rm w})$ is small, windfall can occur even under mundane circumstances; the decision maker is not immune to propitious uncertainty.

These considerations lead to preference rankings on the choice vector x. While "bigger is better" for robustness $\widehat{h}(x, r_c)$, "big is bad" for opportuneness $\widehat{\beta}(x, r_w)$. The preferences induced by the robust-satisficing strategy are:

$$x \succeq_r x' \quad \text{if} \quad \widehat{h}(x, r_c) \geq \widehat{h}(x', r_c) \tag{80}$$

Likewise, the preferences induced by the opportune-windfalling strategy are:

$$x \succeq_o x' \quad \text{if} \quad \widehat{\beta}(x, r_w) \leq \widehat{\beta}(x', r_w) \tag{81}$$

We note that neither $\succeq_r$ nor $\succeq_o$ is necessarily single-valued for any pair of choices x and x': the preferences may change with r_c and r_w, respectively. Moreover, $\succeq_r$ and $\succeq_o$ may rank the same pair of options differently. The preference relations $\succeq_r$ and $\succeq_o$, either alone or together, do not satisfy the rationality conditions of completeness and transitivity. They do not establish unique preferences for all pairs of available choices, and they hence do not entail transitivity of preference.

7.2.4 Info-Gap Choice Problems

We now formulate two info-gap choice problems and explain their conceptual similarity to the classical EMP and UMP.

Given prices $p > 0$, critical reward r_c, and demanded robustness $\widehat{h}_d$, the **robust satisficing problem** (RSP) is to choose a commodity vector x which satisfies:

$$\min_{x \in X} p^T x \quad \text{subject to} \quad \widehat{h}(x, r_c) \geq \widehat{h}_d \tag{82}$$

Given prices $p > 0$, windfall reward r_w, and wealth $w > 0$, the **opportune windfalling problem** (OWP) is to choose x such that:

$$\min_{x \in X} \widehat{\beta}(x, r_w) \quad \text{subject to} \quad p^T x \leq w \tag{83}$$

We now explain the intuitive similarity between robust satisficing (RSP) and expenditure minimization (EMP), and between opportune windfalling (OWP) and utility maximization (UMP).

RSP and EMP: The structural parallel between the EMP in eq.(72) and the RSP is obvious. In both, the expenditure $p^T x$ is minimized subject to a satisficing constraint on a preference-generating function: $u(x)$ or $\widehat{h}(x, r_c)$. Furthermore, the RSP also satisfices the reward by means of the inequality on the lower reward function $\mathcal{R}_*$ in eq.(78). The critical value of utility u in the EMP or the critical reward r_c and demanded robustness $\widehat{h}_d$ in the RSP can be chosen small or large at the decision maker's discretion. In both cases the decision maker is adopting a cautious or protective stance while attempting to guarantee a specified minimal level of satisfaction.

The fundamental difference between the RSP and the EMP is that the latter is based on a complete, transitive preference relation while the former is not. The EMP presumes substantially greater knowledge by the decision maker about the options and their import. The robustness function $\widehat{h}(x, r_c)$ generates a preference relation contingent upon a choice of the critical reward, r_c. Preferences may change as the decision maker's aspiration for reward, r_c, changes. The robustness function is not directly a utility, but an auxiliary evaluation of feasibility or justifiability of option x with regard to reward-aspiration r_c. Knowing a consumer's robustness function tells us less about the utility to that individual of option x, than knowing a utility function $u(x)$.

OWP and UMP: The UMP in eq.(73) is structurally and conceptually similar to the OWP once we recognize that a minimum $\widehat{\beta}$ optimizes the opportuneness for sweeping windfall success. This 'windfalling' is the info-gap analog of the classical search for maximum utility. Moreover, both $\max u(x)$ and $\min \widehat{\beta}$ are subject to the same budget constraint, $p^T x \leq w$. Furthermore, we explained in discussing eq.(73) that the UMP is an ambitious strategy, attempting to achieve the greatest utility facilitated by the available budget. There is no margin of safety in the UMP; large utility is its own protection. Similarly, the windfalling strategy aspires to facilitate reward as large as r_w, expressed by the inequality on the upper reward function $\mathcal{R}^*$ in eq.(79). In the info-gap context there may be no maximal reward even given a finite budget because of the endless potential for propitious uncertainty. What the windfalling decision maker is doing is optimizing the possibility of exploiting this potential, while attempting to facilitate great reward $r_{\rm w}$ far in excess of the critical reward $r_{\rm c}$ needed for survival.

The primary distinction between the OWP and the UMP, as in the RSP and EMP case, is that the opportuneness function $\widehat{\beta}(x, r_{\rm w})$ is not itself a utility function, but rather an auxiliary assessment of the propitiousness of option x regarding windfall aspiration $r_{\rm w}$.

Viewing the four choice-problems together, the primary difference is the symmetry in EMP/UMP which is broken in RSP/OWP by $\widehat{h}(x, r_{\rm c})$ and $\widehat{\beta}(x, r_{\rm w})$. Implications of this, together with some substantive differences between robustness and opportuneness, will arise in propositions 5 and 6.

7.3 Price-Demand Antagonism

In demand theory, $H(p, u)$ is a Hicksian demand function if it is a solution of the EMP. That is, $p^T H(p, u)$ is minimal subject to $u(x) \geq u$. If $u(x)$ is a continuous utility function representing a locally non-satiated preference relation, then $H(p, u)$ obeys the compensated demand law (Mas-Colell *et al,* Proposition 3.E.4, p.62):

$$(p - p')^T [H(p, u) - H(p', u)] \leq 0 \tag{84}$$

Let us define $\widehat{x}_{\rm r}(p, \widehat{h}_{\rm d}, r_{\rm c})$ as a consumer choice x which solves the RSP. That is, $p^T \widehat{x}_{\rm r}(p, \widehat{h}_{\rm d}, r_{\rm c})$ is minimal subject to the robustness constraint $\widehat{h}(x, r_{\rm c}) \geq \widehat{h}_{\rm d}$. $\widehat{x}_{\rm r}(p, \widehat{h}_{\rm d}, r_{\rm c})$ is the **robust satisficing demand.** In the theory we are developing here, $\widehat{x}_{\rm r}(p, \widehat{h}_{\rm d}, r_{\rm c})$ is the analog of $H(p, u)$. The analog of the classical Hicksian demand law is:

Proposition 4 Given: *$\mathcal{F}(h, \widetilde{f})$ is an info-gap model in the Banach space Ω; $\widehat{h}(x, r_{\rm c})$ is a robustness function for $\mathcal{F}(h, \widetilde{f})$ based on a lower reward function $\mathcal{R}_*(x, A)$ which is monotonically decreasing on the sets in Ω; $\widehat{x}_{\rm r}(p, \widehat{h}_{\rm d}, r_{\rm c})$ is a robust satisficing demand function for $\widehat{h}(x, r_{\rm c})$.*

Then: *for any price vectors p and p' we have:*

$$(p - p')^T [\widehat{x}_{\rm r}(p, \widehat{h}_{\rm d}, r_{\rm c}) - \widehat{x}_{\rm r}(p', \widehat{h}_{\rm d}, r_{\rm c})] \leq 0 \tag{85}$$

That is, like the Hicksian demand $H(p, u)$, price p and robust satisficing demand $\widehat{x}_{\rm r}$ vary antagonistically. Moreover, subsequent results will demonstrate that, like $H(p, u)$, the robust satisficing demand $\widehat{x}_{\rm r}(p, \widehat{h}_{\rm d}, r_{\rm c})$ is wealth-compensated.

Let us define $\widehat{x}_{\rm o}(p, w, r_{\rm w})$ as a consumer choice x which solves the OWP. That is, $\widehat{\beta}(\widehat{x}_{\rm o}, r_{\rm w})$ is a minimal value of the opportuneness function subject to the budget constraint $p^T \widehat{x}_{\rm o} \leq w$. Since the OWP is the info-gap analog of the UMP, we see that $\widehat{x}_{\rm o}(p, w, r_{\rm w})$ is the analog of the Walrasian demand function.

7.4 Wealth-Compensated Demand

The commodity vector is $x \in X \subseteq \Re^L$, and the info-gap model of uncertainty $\mathcal{F}(h, \widetilde{f})$, $h \geq 0$, is defined in a Banach space Ω.

Definition 1 *The function robustness $\widehat{h}(x, r)$ is* **continuous** *at x if, for each $\varepsilon > 0$ there is a $\delta > 0$ such that:*

$$\left| \widehat{h}(x, r) - \widehat{h}(x', r) \right| < \varepsilon \quad \text{whenever} \quad \|x - x'\| < \delta \tag{86}$$

Continuity of the opportuneness function is similarly defined.

Relations (74) and (75) define the lower and upper reward functions as monotonic in the uncertainty sets. The following property of strict monotonicity is slightly different.

Definition 2 *The lower reward function* $\mathcal{R}_*(x, A)$ *is* **strictly monotonic** *in* h *for info-gap model* $\mathcal{F}(h, \widetilde{f})$ *if, for all* $x \in X$ *and all* $\widetilde{f} \in \Omega$:

$$h < h' \quad \text{implies} \quad \mathcal{R}_*[x, \mathcal{F}(h, \widetilde{f})] > \mathcal{R}_*[x, \mathcal{F}(h', \widetilde{f})] \tag{87}$$

Definition 3 *The upper reward function* $\mathcal{R}^*(x, A)$ *is* **strictly monotonic** *in* h *for info-gap model* $\mathcal{F}(h, \widetilde{f})$ *if, for all* $x \in X$ *and all* $\widetilde{f} \in \Omega$:

$$h < h' \quad \text{implies} \quad \mathcal{R}^*[x, \mathcal{F}(h, \widetilde{f})] < \mathcal{R}^*[x, \mathcal{F}(h', \widetilde{f})] \tag{88}$$

Definition 4 *The lower reward function* $\mathcal{R}_*(x, A)$ *is* **continuous** *in* h *at* $\widetilde{f}$ *for an info-gap model* $\mathcal{F}(h, \widetilde{f})$ *if, for each* $x \in X$ *and for each* $\varepsilon > 0$ *there is a* $\delta > 0$ *such that:*

$$\left| \mathcal{R}_*(x, \mathcal{F}(h, \widetilde{f})] - \mathcal{R}_*(x, \mathcal{F}(h', \widetilde{f})] \right| < \varepsilon \quad \text{whenever} \quad |h - h'| < \delta \tag{89}$$

Continuity of an upper reward function is similarly defined.

Definition 5 *Lower and upper reward functions* $\mathcal{R}_*(x, A)$ *and* $\mathcal{R}^*(x, A)$ *are* **similarly ordered** *if, for all* $x, x' \in X$ *and for all* $A \subset \Omega$:

$$\mathcal{R}_*(x, A) < \mathcal{R}_*(x', A) \quad \text{if and only if} \quad \mathcal{R}^*(x, A) < \mathcal{R}^*(x', A) \tag{90}$$

An immunity function, $\widehat{h}(x, r_\mathrm{c})$ or $\widehat{\beta}(x, r_\mathrm{w})$, is 'non-satiated' if an arbitrarily small change in x can improve the immunity. Since "bigger is better" for robustness $\widehat{h}(x, r_\mathrm{c})$, while "big is bad" for opportuneness $\widehat{\beta}(x, r_\mathrm{w})$, the definitions of non-satiation for these immunity functions are different but symmetrical.

Definition 6 *The opportuneness function is* **non-satiated** *at* $x \in X$ *if, for every* $\varepsilon > 0$*, there is an* $x' \in X$ *such that*

$$\|x - x'\| < \varepsilon \quad \text{and} \quad \widehat{\beta}(x', r_\mathrm{w}) < \widehat{\beta}(x, r_\mathrm{w}) \tag{91}$$

Definition 7 *The robustness function is* **non-satiated** *at* $x \in X$ *if, for every* $\varepsilon > 0$*, there is an* $x' \in X$ *such that*

$$\|x - x'\| < \varepsilon \quad \text{and} \quad \widehat{h}(x', r_\mathrm{w}) > \widehat{h}(x, r_\mathrm{w}) \tag{92}$$

Lemma 1 Given: $\mathcal{R}_*(x, A)$ *is a lower reward function which is continuous in* h *at all* $\widetilde{f}$ *for an info-gap model* $\mathcal{F}(h, \widetilde{f})$; $\widehat{h}(x, r_\mathrm{c})$ *is the robustness function defined for this reward function;* $\widehat{h}(x, r_\mathrm{c})$ *is continuous at all* x.

If x *solves the RSP of eq.(82)* **then:**

$$\widehat{h}(x, r_\mathrm{c}) = \widehat{h}_\mathrm{d} \quad \text{and} \quad \mathcal{R}_*[x, \mathcal{F}(\widehat{h}(x, r_\mathrm{c}), \widetilde{f})] = r_\mathrm{c} \tag{93}$$

The RSP entails satisficing both the robustness (the inequality in eq.(82)) and the reward (the inequality in eq.(78)). Lemma 1 asserts that there is no excess robustness or reward at a solution of the RSP.

Lemma 2 Given: $\mathcal{R}^*(x, A)$ *is an upper reward function which is continuous in* h *at all* $\widetilde{f}$ *for an info-gap model* $\mathcal{F}(h, \widetilde{f})$; $\widehat{\beta}(x, r_\mathrm{w})$ *is the opportuneness function defined for this reward function;* $\widehat{\beta}(x, r_\mathrm{w})$ *is continuous at all* x.

If x *solves the OWP of eq.(83)* **and if** $\widehat{\beta}(x, r_\mathrm{w})$ *is non-satiated at this* x, **then:**

$$p^T x = w \quad \text{and} \quad \mathcal{R}^*[x, \mathcal{F}(\widehat{\beta}(x, r_\mathrm{w}), \widetilde{f})] = r_\mathrm{w} \tag{94}$$

Note that the opportuneness function is satiated if $\widehat{\beta}(x, r_{\rm w}) = 0$. Hence (94) need not hold if the minimal opportuneness function vanishes.

Lemma 2 states that if x solves the OWP then x also obeys Walras' law of complete utilization of wealth. Furthermore, the lemma asserts that there is no excess reward in the solution of the OWP: the inequality in the definition of the opportuneness function, eq.(79), is an equality at x.

Proposition 5 **Given:** *$\mathcal{R}_*(x, A)$ and $\mathcal{R}^*(x, A)$ are similarly ordered lower and upper reward functions, respectively; $\mathcal{R}_*(x, A)$ is continuous in h at all $\widetilde{f}$ for an info-gap model $\mathcal{F}(h, \widetilde{f})$; $\mathcal{R}^*(x, A)$ is strictly monotonic in h; $\widehat{h}(x, r_{\rm c})$ and $\widehat{\beta}(x, r_{\rm w})$ are robustness and opportuneness functions, respectively, defined for these reward functions; $\widehat{h}(x, r_{\rm c})$ and $\widehat{\beta}(x, r_{\rm w})$ are continuous at all x.*

If *$\widehat{x}_{\rm r}$ solves the RSP of eq.(82),* **then** *$\widehat{x}_{\rm r}$ solves the OWP of eq.(83) with $w = p^T\widehat{x}_{\rm r}$ and $r_{\rm w} = \mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]$. Moreover, $\widehat{\beta}(\widehat{x}_{\rm r}, r_{\rm w}) = \widehat{h}_{\rm d} = \widehat{h}(\widehat{x}_{\rm r}, r_{\rm c})$.*

Proposition 6 **Given:** *$\mathcal{R}_*(x, A)$ and $\mathcal{R}^*(x, A)$ are similarly ordered lower and upper reward functions, respectively; $\mathcal{R}^*(x, A)$ is continuous in h at all $\widetilde{f}$ for an info-gap model $\mathcal{F}(h, \widetilde{f})$; $\mathcal{R}_*(x, A)$ is strictly monotonic in h; $\widehat{h}(x, r_{\rm c})$ and $\widehat{\beta}(x, r_{\rm w})$ are robustness and opportuneness functions, respectively, defined for these reward functions; $\widehat{h}(x, r_{\rm c})$ and $\widehat{\beta}(x, r_{\rm w})$ are continuous at all x and $\widehat{h}(x, r_{\rm c})$ is non-satiated at all x.*

If *$\widehat{x}_{\rm o}$ solves the OWP of eq.(83)* **and if** *$\widehat{\beta}(\widehat{x}_{\rm o}, r_{\rm w})$ is non-satiated at $\widehat{x}_{\rm o}$,* **then** *$\widehat{x}_{\rm o}$ solves the RSP of eq.(82) with $\widehat{h}_{\rm d} = \widehat{\beta}(\widehat{x}_{\rm o}, r_{\rm w})$ and $r_{\rm c} = \mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(\widehat{\beta}(\widehat{x}_{\rm o}, r_{\rm w}), \widetilde{f})]$. Moreover, $p^T\widehat{x}_{\rm o} = w$ and $\widehat{h}(\widehat{x}_{\rm o}, r_{\rm c}) = \widehat{h}_{\rm d}$.*

We can summarize our results as follows. Let $w_{\rm rsp}$ denote the minimum expenditure which is obtained as the solution of the RSP, eq.(82): $w_{\rm rsp}(p, r_{\rm c}, \widehat{h}_{\rm d}) = p^T\widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d})$ since $\widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d})$ is the solution of the RSP. Proposition 5 relates $\widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d})$ to $\widehat{x}_{\rm o}(p, r_{\rm w}, w)$ (the solution of the OWP) as follows:

$$\widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d}) = \widehat{x}_{\rm o}[p, \underbrace{\mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm w}}, w_{\rm rsp}(p, r_{\rm c}, \widehat{h}_{\rm d})] \tag{95}$$

This relation explains the sense in which $\widehat{x}_{\rm r}$ is a **wealth-compensated demand.** By definition, the demand $\widehat{x}_{\rm r}$(p, $r_{\rm c}$, $\widehat{h}_{\rm d}$) is evaluated with fixed critical reward $r_{\rm c}$ and fixed demanded robustness $\widehat{h}_{\rm d}$. Moreover, eq.(95) shows that, as prices change, $\widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d})$ gives the consumption when the wealth $w_{\rm rsp}(p, r_{\rm c}, \widehat{h}_{\rm d})$, is adjusted to keep both $r_{\rm c}$ and $\widehat{h}_{\rm d}$ constant. This parallels the sense in which the Hicksian demand function is a compensated demand (Mas-Colell *et al,* p.62). This further strengthens the parallel between $\widehat{x}_{\rm r}$ and the Hicksian demand, since we already know that they both obey price-demand trade-off (proposition 4).

Let $\alpha_{\rm owp}(r_{\rm w}, w)$ denote the minimum obtained as the solution of the OWP: the lowest info-gap which facilitates windfall reward $r_{\rm w}$ given wealth w. Thus $\alpha_{\rm owp}(r_{\rm w}, w) = \widehat{\beta}(\widehat{x}_{\rm o}, r_{\rm w})$. Proposition 6 relates the solutions of the OWP and the RSP as:

$$\widehat{x}_{\rm o}(r_{\rm w}, w) = \widehat{x}_{\rm r}[\underbrace{\mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(\alpha_{\rm owp}, \widetilde{f})]}_{r_{\rm c}}, \alpha_{\rm owp}(r_{\rm w}, w)] \tag{96}$$

Propositions 5 and 6, together with lemmas 1 and 2, also assert that:

$$\widehat{h}(\widehat{x}_{\rm r}, \underbrace{\mathcal{R}_*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm c}}) = \widehat{\beta}(\widehat{x}_{\rm r}, \underbrace{\mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm w}}) \tag{97}$$

$$\widehat{h}(\widehat{x}_{\rm o}, \underbrace{\mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(\alpha_{\rm owp}, \widetilde{f})]}_{r_{\rm c}}) = \widehat{\beta}(\widehat{x}_{\rm o}, \underbrace{\mathcal{R}^*[\widehat{x}_{\rm o}, \mathcal{F}(\alpha_{\rm owp}, \widetilde{f})]}_{r_{\rm w}}) \tag{98}$$

Each of these relations establishes $r_{\rm c}$ and $r_{\rm w}$ values at which the robustness and opportuneness functions obtain the same values when evaluated at the same level of demand.

Consider for instance eq.(97). When the consumer chooses the RSP-optimal demand $\widehat{x}_{\rm r}$, then critical reward no less than $r_{\rm c}$ is guaranteed if the info-gap is no greater than $\widehat{h}(\widehat{x}_{\rm r},\ r_{\rm c})$. Eq.(97)

asserts that this is the lowest level of uncertainty at which reward as large as $r_{\rm w}$ is possible. In other words, eq.(97) identifies a situation in which survival ($r_{\rm c}$) is guaranteed while windfall ($r_{\rm w}$) is possible. A similar situation is identified by eq.(98).

Furthermore, it can be shown that $\widehat{h}(x, r_{\rm c})$ decreases as $r_{\rm c}$ increases, while $\widehat{\beta}(x, r_{\rm w})$ increases as $r_{\rm w}$ increases [8]. Thus eq.(97) specifies the greatest feasible windfall $r_{\rm w}$ at which critical reward $r_{\rm c}$ is robustly guaranteed, when demand $\widehat{x}_{\rm r}$ is chosen. Any greater value of $r_{\rm w}$ cannot be attained without increasing the info-gap above the value of robustness $\widehat{h}(\widehat{x}_{\rm r}, r_{\rm c})$. Conversely, $r_{\rm c}$ is the lowest value of critical survival-reward which can be robustly guaranteed while also enabling windfall as large as $r_{\rm w}$, when choosing $\widehat{x}_{\rm r}$. Any larger value of $r_{\rm c}$ would entail lower robustness and hence a lower feasible value of windfall reward $r_{\rm w}$. Similar conclusions apply to eq.(98) when demand function $\widehat{x}_{\rm o}$ is chosen.

The lower and upper reward functions are very commonly 'naturally ordered' (definition 8). We will explain that this property implies:

$$r_{\rm c} \le r_{\rm w} \tag{99}$$

for each of the $r_{\rm c}$-$r_{\rm w}$ pairs in eqs.(97) and (98).

Definition 8 *Lower and upper reward functions $\mathcal{R}_*(x, A)$ and $\mathcal{R}^*(x, A)$ are* **naturally ordered** *if, for all $x \in X$ and for all $f \in \Omega$:*

$$\mathcal{R}_*(x, \{f\}) \le \mathcal{R}^*(x, \{f\}) \tag{100}$$

We have not assumed natural ordering of our reward functions, but it does in fact hold for the reward functions of eqs.(76) and (77). Natural ordering together with the monotonicity properties, eqs.(74) and (75), imply relation (99) for the $r_{\rm c}$ and $r_{\rm w}$ values in eqs.(97) and (98).

7.5 Summary and Conclusion

We have developed a theory of consumer demand which preserves the phenomenological features of classical demand, without the assumptions of rational-choice theory: complete, transitive preferences. We have explained that the info-gap choice problems of robust satisficing (RSP) and opportune windfalling (OWP) correspond to the traditional problems of expenditure minimization (EMP) and utility maximization (UMP), respectively. We have demonstrated that the consumer choice resulting from the RSP, $\widehat{x}_{\rm r}$, is the analog of the Hicksian demand function, while the solution of the OWP, $\widehat{x}_{\rm o}$, is analogous to the Walrasian demand. Specifically, $\widehat{x}_{\rm r}$ is a wealth-compensated demand function which obeys Walras' law and the law of price-demand trade-off.

The workhorses of info-gap analysis are the immunity functions: the robustness function $\widehat{h}(x, r_{\rm c})$ and the opportuneness function $\widehat{\beta}(x, r_{\rm w})$. These functions are derived from an info-gap model, $\mathcal{F}(h, \widetilde{f})$, $h \ge 0$, which is an unbounded family of nested sets of events, and which expresses the unbounded domain of pernicious as well as propitious possibilities entailed by the decision maker's incomplete knowledge. The immunities also depend upon the lower and upper reward functions, $\mathcal{R}_*(x, A)$ and $\mathcal{R}^*(x, A)$, which describe anticipated reward accruing from choice x accompanied by uncertain environment A.

The robustness function $\widehat{h}(x, r_{\rm c})$ assesses the degree of robustness of choice x to pernicious ambient uncertainty, when the decision maker aspires to satisfice at the reward level $r_{\rm c}$. The opportuneness function $\widehat{\beta}(x, r_{\rm w})$ expresses the immunity of option x to the possible attainment of a great windfall reward $r_{\rm w}$. Both immunity functions depend upon aspirations, $r_{\rm c}$ and $r_{\rm w}$, which are not specified by the theory nor necessarily chosen *a priori* by the consumer. Once the aspiration levels $r_{\rm c}$ and $r_{\rm w}$ are chosen, then each immunity function generates a complete transitive preference ranking of the options. However, these preference rankings—based on robustness or opportuneness—need not be consistent with one another, nor need they remain constant as aspirations change. Since wealth, price and other contextual factors will influence consumer aspiration, info-gap theory does not presume prior knowledge of preference, nor can it predict consumer choice. Even in this context, however, we have seen that the basic results of demand theory remain intact: complementarity of the robust-satisficing and the opportune-windfalling choice problems; Hicksian- and Walrasian-like demand functions derived from these consumer-choice problems; wealth-compensated price-demand trade-off.

Part III
Conclusion

8 Positivism, F-twist, and Robust-Satisficing

The central issues of this paper— how to formulate and evaluate policy and how to model economic behavior—are methodological, and touch the hoary debate over positivism in economics. Without wishing to enter the fray (positive economics: right or wrong?), I will use the clash between Friedman (1953) and Samuelson (1963) to illuminate the methodological distinctiveness of info-gap robust-satisficing.

In a nutshell, the argument which we develop in this section is this:

- Friedman is right that good theories depend on axioms which capture an essential truth, while usually violating a messier reality.[6]
- Samuelson is right that factual inaccuracy of a theory detracts from its validity in prediction and policy formulation.[7]
- Both Samuelson and Friedman agree that economic science, like natural science, improves over time and progresses towards truth.[8]
- However, there is an inherent indeterminism in economic systems which precludes the shared belief of Samuelson and Friedman.[9]
- Hence optimization—of models or of policy outcomes—is fatuous (or serendipitious).[10] Satisficing can, sometimes, be done reliably.[11]

8.1 Friedman and Samuelson

Friedman was undoubtedly right (though rhetorical and infuriating, as usual) when he wrote (Friedman, 1953, p.14):

> Truly important and significant hypotheses will be found to have 'assumptions' that are wildly inaccurate descriptive representations of reality, and, in general, the more significant the theory, the more unrealistic the assumptions (in this sense). The reason is simple. A hypothesis is important if it 'explains' much by little, that is, if it abstracts the common and crucial elements from the mass of complex and detailed circumstances surrounding the phenomena to be explained and permits valid predictions on the basis of them alone.

A first-rate illustration is Galileo's "wildly inaccurate" Law of Inertia: a body moves at constant velocity unless acted upon by a force. This is the most blatantly counter-intuitive proposition in the history of science: no one has ever witnessed an instance of such perpetual motion. Aristotle's hypothesis is far more realistic: a body loses speed unless acted upon by a force. And yet Galileo's hypothesis is outstripped by few others for theoretical fruitfulness and predictive power. As Friedman would explain, Galileo's Law strips away the nagging nuisance of dissipative forces and cuts to the essence of material dynamics.

Samuelson summarized Friedman's position with what he referred to as the 'F-Twist' (Samuelson, 1963, p.232):

[6] "Your bait of falsehood takes this carp of truth;" Shakespeare.

[7] "Thanks to the negation sign, there are as many truths as falsehoods; we just can't always be sure which are which." Quine, (1995).

[8] "The movement of ideas toward truth may be glacial but, like a glacier, it is hard to stop." Galbraith (1986).

[9] "For fallibilism is the doctrine that our knowledge is never absolute but always swims, as it were, in a continuum of uncertainty and of indeterminacy.", Peirce (1897, 1955).

[10] "Optimization works in theory but risk management is better in practice. There is no scientific way to compute an optimal path for monetary policy." Greenspan (2005).

[11] Hence the need for "a little stodginess at the central bank." (Blinder, 1998).

> A theory is vindicated if (some of) its consequences are empirically valid to a useful degree of approximation; the (empirical) unrealism of the theory 'itself,' or of its 'assumptions,' is quite irrelevant to its validity and worth.

Samuelson (1963, p.233) is undoubtedly correct in focussing on the basic F-Twist, which

> is fundamentally wrong in thinking that unrealism in the sense of factual inaccuracy even to a tolerable degree of approximation is anything but a demerit for a theory or hypothesis (or set of hypotheses).

Samuelson presents logical and epistemological arguments. I suggest further support for Samuelson's position. We value successful tests of a theory because they add warrant to the belief that the theory will predict accurately even when no data are around. That is, when a theory passes a severe test of its predictions, we gain support for the inductive hypothesis that the theory will reliably predict in untested circumstances as well. The significance of successful tests of a theory derives from the fact—established by Hume—that empirical evidence can never verify the inductive hypothesis that what worked yesterday will work tomorrow, or that what worked in one circumstance will work in another. Predictive success strengthens our faith in the inductive hypothesis (Haack, 1993).

Thus Friedman is right that fruitful theories may rest of idealizations which are empirically violated, while Samuelson is right that disparity between axiom and observation detracts from the value of the theory.

8.2 Shackle-Popper Indeterminism

Having now established myself firmly in both Samuelson's and Friedman's camps, I will not attempt to resolve their conflict. Rather, I will demonstrate the methodological distinctiveness of info-gap robust-satisficing by disputing a point upon which they would both agree. Both men concur that economics is epistemologically identical to physical science. Other scholars, for instance Koopmans, would also agree (Koopmans, 1957 pp.134–135). Friedman is explicit (Friedman, 1953, pp.4–5). Samuelson is implicit, by applying the criterion of logical consistency which is the pride of the physicists (Samuelson, 1963, pp.233–234). Physicists yearn for universal unifying theories, and a collection of logically conflicting sub-theories (relativity, statistical mechanics, quantum theory, etc.) is viewed with discomfort.

In contrast, a number of scholars, such as Habermas (1970), have emphasized the non-nomological nature of social science. I will not review the rich literature, but only focus on what I will refer to as Shackle-Popper indeterminism. This idea, developed separately and in different ways by Shackle (1972) and Popper (1957, 1982), will lay the groundwork for the methodology of info-gap robust-satisficing.[12]

The basic idea is that the behavior of intelligent learning systems displays an element of unstructured and unpredictable indeterminism. By 'intelligence' I mean: behavior is influenced by knowledge. This is surely characteristic of humans individually and of society at large. By 'learning' I mean a process of discovery: finding out today what was unknown yesterday. One economically important example of learning is what Keynes referred to as hearing 'the news'. Finally, indeterminism arises as follows: because tomorrow's discovery is by definition unknown today, and tomorrow's behavior will be influenced by newly discovered knowledge, tomorrow's behavior is not predictable today, at least not in its entirety. Given the richness of future discovery, (or its corollary, the richness of our current ignorance), the indeterminism of future behavior is broad, deep and unstructured.

The most important domain of indeterminism, for our purposes, is mathematical modelling of social systems, in particular, economic ones. Complexity and dimensionality are severe challenges in themselves. However, here we are dealing with the limited ability of laws, derived from *past* behavior, to describe *future* behavior. Intelligent learning behavior, as we have defined it, entails an

[12] Shackle-Popper indeterminism is related to Knightian uncertainty (Knight, 1921). See also Ben-Haim (2006, chap. 12) and Ben-Haim (2007, p.157).

element of innovation which seems to explain the painful experience of social modellers. In this sense, Shackle-Popper indeterminism accounts for the partially non-nomological nature of social systems.

8.3 Methodological Implications

The point is *not* that models can never describe social or economic activity, or that there is nothing that can be called a law of behavior. Rather, models of intelligent learning systems must focus on two categories of uncertainties, which I will refer to as 'measurement' and 'epistemic' uncertainty.

Measurement uncertainty is the usual kind: noise in data, bias in samples, and so on. Quite often, though not always (Fox *et al,* 2007), measurement uncertainty can be modelled and managed with statistical tools.

Epistemic uncertainty refers to the gaps and errors in our understanding of the processes being modelled, arising in part from Shackle-Popper indeterminism. For instance, complex non-linearities are approximated with linear models, high-dimensional systems are truncated, relevant interactions are ignored or not recognized. We usually know that our understanding is deficient, and we may even be ignoring implicit knowledge for pragmatic reasons. But more profoundly, the Shackle-Popper indeterminism implies that some element of error is inevitable in describing tomorrow's behavior with today's models, and that the nature and degree of this disparity is inherently unforeseeable.

For example, the info-gap models introduced in section 3.1 represent epistemic uncertainty about the behavior of the economy. The coefficients of the model are uncertain not only due to measurement error, but due to the basic inability to entirely capture future behavior with past observation. The info-gap model of uncertainty is non-probabilistic precisely because the errors are epistemic, not statistical.

One methodological implication of Shackle-Popper indeterminism is that model error must be taken very seriously. The traditional positivistic optimism that our models are improving and even converging on the truth, is untenable. Positivism was enlisted to explain the wonderful success of the science-based technology which burst on the scene in the 19th century. The strategy of scientists such as Kelvin and Raleigh was to learn nature's laws, and then to apply them to technological enterprise. This works because Peirce was right that the study of nature does converge, asymptotically, on something that can reasonably be called the truth. As much as social scientists may wish to emulate this strategy, it will founder on the shoals of Shackle-Popper indeterminism.

One way of taking model-error seriously is to use the robust-satisficing strategy which we defined in section 3.2. For instance, $\widehat{h}_+(v, y_\mathrm{c})$ in eq.(5) is the robustness of control vector v, to errors in the economic model, when the target is satisficed to level y_c. This robustness function evaluates the policy worthiness of control v, given performance-aspiration y_c, in light of epistemic uncertainty about the economic system.

Another methodological implication, dealt with elsewhere (Ben-Haim, 2005a), touches on the process of up-dating a mathematical model based on measurements. The two horns of uncertainty—measurement and epistemic uncertainty—generate an irrevocable trade-off. A model with high fidelity to data will have low immunity to structural error in the model. Either type of uncertainty can be ameliorated only at the expense of exacerbating the other. The methodological response is again to use robust-satisficing. A model with maximal fidelity to data will have zero immunity to structural info-gaps, so it is best to satisfice the fidelity at a sub-optimal level (rather than to optimize the fidelity), in order to garner some robustness to epistemic uncertainty.

9 What Next?

We have only barely begun to apply info-gap theory to economic decisions and models. We outline here a few additional directions of study.

When considering **decision strategies** the following topics, among others, are important.

Forecasting. Forecasts which are based on sound statistical fidelity to historical data are liable to fail due to fundamental structural change in the future.[13] Shackle-Popper indeterminism mitigates against optimizing the fidelity of a forecasting model to historical data. The robust-satisficing strategy, together with a proxy theorem, can lead to more reliable predictions of future outcomes. Some work in this direction has been done (Ben-Haim, 2008) but much remains to be explored and developed.

Financial risk management. Value-at-risk and related techniques depend on knowledge of probability distributions. However, the extreme tails of these distributions are often poorly known (Hendricks, 1996). The uncertainty is not only statistical, resulting from limited historical data on extreme events. The uncertainty is also epistemic: future extreme events may result from mechanisms which have no historical manifestations. Once again we face Knightian uncertainty and Shackle-Popper indeterminism. Once again, robust-satisficing can lead to more reliable risk management strategies (Ben-Haim, 2005b).

Large-scale monetary policy formulation. To date, applications of info-gap theory to monetary policy have been based on small stylized economic models (Akram, Ben-Haim and Eitrheim, 2006; Ben-Haim, Akram and Eitrheim, 2007). The application of the info-gap robust-satisficing strategy to larger econometric models entails many challenges and opportunities.

The range of **economic behavior** which can be explored using info-gap theory is enormous, including the following topics.

Economic paradoxes. We explored an info-gap explanation of the equity premium puzzle in section 6 (see also Ben-Haim, 2006, section 11.5). This can be extended in various ways, including empirical study. An initial study of the home bias paradox (Ben-Haim and Jeske, 2003) can also be extended. Other conundrums can be studied, such as Deaton's paradox, the purchasing power parity puzzle, and others (Obstfeld and Rogoff, 1996).

Competitive markets with info-gaps. Preliminary work shows that much of the classical marginal analysis of competitive equilibrium is reproduced—in somewhat altered form—in an economic model in which firms and households robust-satisfice rather than maximize their profit or utility (Ben-Haim, 2002). This can be extended.

What are rational expectations? When an agent faces uncertainty, and wishes to achieve specified goals (like economic survival), it is *not* rational to base decisions on the best known model of the economy. Such an approach would minimize the agent's robustness to error and, when a proxy theorem applies, would also minimize the probability of achieving those goals. In order to be successful, the uncertain agent should use a robust-satisficing strategy to formulate beliefs about how the economy operates.

Learning and anticipation. Likewise, an agent who learns about the unfolding ambient reality should not try to optimize the learning algorithm in the sense of minimizing the learning time or maximizing the amount which is learned. Such strategies are based on prior knowledge about the

[13] This has been noted by several authors. In econometrics, "the [data] generating process is unknown and evolutionary" (Hendry, 1995, p.xxix). Furthermore, in "the forecasting context ... the degree of data congruence or non-congruence of a model is neither necessary nor sufficient for forecasting success or failure." (Clements and Hendry, 1999, p.xxiv). These authors develop auto-regressive forecasting models which use difference data, and non-causal variables, showing that such models can out-perform historically calibrated and causally-relevant models.

In macro-economic modelling

> there is genuine uncertainty about how good a model is, even within the sample. Moreover, since the economy is evolving, we can take it for granted that the data generation process will change in the forecast period, causing any model of it to become mis-specified over that period, and this is eventually the main problem in economic forecasting. (Bardsen *et al,* 2005, p.246)

system which is being learned, and they are maximally vulnerable to info-gaps in that knowledge. (That there are serious errors in the agent's knowledge is plausible; otherwise, what is there to learn?) Robust-satisficing strategies for learning are more reliable. These are strategies which satisfice the relevant entities (e.g. learning time) and maximize the robustness to errors in the knowledge which underlies the strategy.

10 References

1. Akram, Q.F., Yakov Ben-Haim, and Ø. Eitrheim, 2006, Managing uncertainty through robust-satisficing monetary policy. Norges Bank Working Papers, ANO 2006/10. Oslo: Norges Bank.
2. Bardsen, G., Eitrheim, Ø., Jansen, E.S. and Nymoen, R., 2005, *The Econometrics of Macroeconomic Modelling,* Oxford University Press.
3. Ben-Haim, Yakov, 1996, *Robust Reliability in the Mechanical Sciences,* Springer-Verlag, Berlin.
4. Ben-Haim, Yakov, 1999, Set-models of information-gap uncertainty: Axioms and an inference scheme, *Journal of the Franklin Institute,* 336: 1093–1117.
5. Ben-Haim, Yakov, 2002, Competitive markets with info-gaps, working paper.
6. Ben-Haim, Yakov, 2005a, Info-Gap Decision Theory For Engineering Design. Or: Why 'Good' is Preferable to 'Best', chapter 11 in *Engineering Design Reliability Handbook,* E. Nikolaides, D. Ghiocel & Surendra Singhal, eds., CRC Press.
7. Ben-Haim, Yakov, 2005b, Value at risk with info-gap uncertainty, Journal of Risk Finance, vol. 6, #5, pp.388–403.
8. Ben-Haim, Yakov, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd ed., Academic Press, London.
9. Ben-Haim, Yakov, 2007, Peirce, Haack and Info-Gaps, pp.150–164 in *Susan Haack, A Lady of Distinctions: The Philosopher Responds to Her Critics,* edited by Cornelis de Waal, Prometheus Books, New York.
10. Ben-Haim, Yakov, 2008, Info-gap forecasting and the advantage of sub-optimal models, to appear in *European Journal of Operational Research.*
11. Ben-Haim, Yakov, Q.F. Akram, and Ø. Eitrheim, 2007, Monetary policy under uncertainty: Min-max vs robust-satisficing strategies, Norges Bank Working Papers, ANO 2007/6. Oslo: Norges Bank.
12. Ben-Haim, Yakov and Keith W. Hipel, 2002, The graph model for conflict resolution with information-gap uncertainty in preferences, *Applied Mathematics and Computation, 126,* 319–340.
13. Ben-Haim, Yakov and Karsten Jeske, 2003, Home-bias in financial markets: Robust satisficing with info-gaps, Federal Reserve Bank of Atlanta, Working Paper Series, 2003-35, Dec. 2003, pdf file on the Federal Reserve Bank website. SSRN abstract and full paper at: http://ssrn.com/abstract=487585.
14. Ben-Haim, Yakov and Laufer, A., 1998, Robust reliability of projects with activity-duration uncertainty, *ASCE Journal of Construction Engineering and Management, 124,* 125–132.
15. Blanchard, Olivier Jean and Stanley Fischer, 1989, *Lectures on Macroeconomics,* MIT Press, Cambridge, MA.
16. Blinder, Alan S., 1998, *Central Banking in Theory and Practice,* Lionel Robbins Lecture, MIT Press, Cambridge, p.12.
17. Burgman, Mark A., 2005, *Risks and decisions for conservation and environmental management,* Cambridge: Cambridge University Press.
18. Carmel, Yohay and Yakov Ben-Haim, 2005, Info-gap robust-satisficing model of foraging behavior: Do foragers optimize or satisfice?, *American Naturalist, 166,* 633–641.
19. Clements, M.P., and D.F.Hendry, 1999, *Forecasting Non-Stationary Economic Time Series,* MIT Press.
20. Fox, David R., Yakov Ben-Haim, Keith R. Hayes, Michael McCarthy, Brendan Wintle, Piers Dunstan, 2007, An info-gap approach to power and sample size calculations, *Environmetrics,* vol. 18, pp.189–203.
21. Friedman, Milton, 1953, On the methodology of positive economics, In: Friedman, Milton, *Essays in Positive Economics,* University of Chicago Press, 1953.
22. Galbraith, John Kenneth, 1986, *The New Industrial State,* 4th ed., Mentor Book, p.201.

23. Greenspan, Alan, 2005, Financial Times, 27–28 August.
24. Haack, Susan, 1993, *Evidence and Inquiry: Towards Reconstruction in Epistemology.* Blackwell.
25. Habermas, Jürgen, 1970, *On the Logic of the Social Sciences,* trans. by Shierry Weber Nicholsen and Jerry A. Stark, 1990, MIT Press.
26. Hendricks, Darryll, 1996, Evaluation of value-at-risk models using historical data, *Economic Policy Review,* Federal Reserve Bank of New York, April 1996, Vol. 2, no. 1.
27. Hendry, D.F., 1995, *Dynamic Econometrics,* Oxford University Press.
28. Kanno Y. and Izuru Takewaki, 2006a, Robustness analysis of trusses with separable load and structural uncertainties, *International Journal of Solids and Structures, 43,* #9, 2646–2669.
29. Kanno, Y. and Izuru Takewaki, 2006b, Sequential semidefinite program for maximum robustness design of structures under load uncertainty, *Journal of Optimization Theory and Applications, 130,* #2, 265–287.
30. Klir, George J., 2006, *Uncertainty and information: Foundations of generalized information theory,* New York: Wiley Publishers.
31. Knight, Frank H., 1921, *Risk, Uncertainty and Profit,* Hart, Schaffner and Marx. Re-issued by Harper Torchbooks, New York, 1965.
32. Knoke, T., 2007, Mixed forests and finance—Methodological approaches, *Ecological Economics,* to appear.
33. Kocherlakota, Narayana R., 1996, The equity premium: It's still a puzzle, *Journal of Economic Literature,* 34: 42–71.
34. Koopmans, Tjalling C., *Three Essays on the State of Economic Science,* McGraw-Hill Book Co., New York.
35. Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green, 1995, *Microeconomic Theory.* Oxford University Press.
36. McCarthy, M.A. and D.B. Lindenmayer, 2007, Info-gap decision theory for assessing the management of catchments for timber production and urban water supply, *Environmental Management, 39,* #4, 553–562.
37. Mehra, Rajnish and Edward C.Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics,* 15: 145–161.
38. Moffitt, L. Joe, John K. Stranlund and Barry C. Field, 2005, Inspections to avert terrorism: Robustness under severe uncertainty, *Journal of Homeland Security and Emergency Management,* Vol.2, #3. http://www.bepress.com/jhsem/vol2/iss3/3
39. Obstfeld, Maurice and Kenneth S. Rogoff, 1996, *Foundations of International Macroeconomics,* MIT Press, Cambridge, USA.
Pantelides Chris P. and Sara Ganzerli, 1998, Design of trusses under uncertain loads using convex models, *ASCE J. Structural Engineering, 124,* #3, 318–329.
40. Peirce, Charles Sanders, 1897, *Collected Papers,* 1.170, 171-175. Reprinted in [41] pp.355-358.
41. Peirce, Charles Sanders, 1955, *Philosophical Writings of Peirce.* Justus Buchler, editor. Dover reprint of *The Philosophy of Peirce: Selected Writings,* Routledge and Kegan Paul, 1940, p.356.
42. Pierce, S. Gareth, Yakov Ben-Haim, Keith Worden and Graeme Manson, 2006, Evaluation of neural network robust reliability using information-gap theory, *IEEE Transactions on Neural Networks, 17,* #6, 1349–1361.
43. Popper, Karl R., 1957, *The Poverty of Historicism.* Harper Torchbooks edition 1964.
44. Popper, Karl, 1982, *The Open Universe: An Argument for Indeterminism.* From the Postscript to *The Logic of Scientific Discovery.* Routledge.
45. Quine, Williard V., 1995, *From Stimulus to Science.* Harvard University Press, p.67.
46. Regan, Helen M., Yakov Ben-Haim, Bill Langford, Will G. Wilson, Per Lundberg, Sandy J. Andelman, and Mark A. Burgman, 2005, Robust decision making under severe uncertainty for conservation management, *Ecological Applications, 15,* #4,: 1471–1477.

47. Regev, S., Avy Shtub and Yakov Ben-Haim, 2006, Managing project risks as knowledge gaps, *Project Management Journal, 37,* #5, 17–25.
48. Samuelson, P.A., 1963, Problems of methodology — discussion. *American Economic Review,* Papers and Proceedings of the 75th Meeting of the American Economic Association, May 1963, 53: 231–236.
49. Shackle, G.L.S., 1972, *Epistemics and Economics: A Critique of Economic Doctrines,* Transaction Publishers, 1992, originally published by Cambridge University Press.
50. Shakespeare, William, *The Tragedy of Hamlet, Prince of Denmark,* Act II, scene 1, line 69.

A Proofs for Section 3

Proof of proposition 1. The proof for $\widehat{h}_+$ derives directly from the fact that the robustness curve, $\widehat{h}_+(v, y_{\rm c})$ vs. $y_{\rm c}$, is a straight line with slope $1/\theta(v)$ and intercept $\widetilde{y}(v)$, as illustrated in fig. 1. Similar reasoning applies for $\widehat{h}_-$ and $\widehat{h}$, as illustrated in figs. 2 and 3. ■

Proof of proposition 2. Consider first the upper robustness. The standardization condition allows us to express the variation of $P_{\rm s+}$ in eq.(21) as follows:

$$\frac{\partial P_{\rm s+}}{\partial v_i} = \left.\frac{\partial Z}{\partial \zeta}\right|_{\widehat{h}_+} \frac{\partial \widehat{h}_+}{\partial v_i} \tag{101}$$

$$= z(\widehat{h}_+)\frac{\partial \widehat{h}_+}{\partial v_i} \tag{102}$$

The pdf is non-negative, so eq.(102) is equivalent to eq.(33).

Now consider the lower robustness. The standardization condition allows us to express the variation of $P_{\rm s-}$ in eq.(25) as follows:

$$\frac{\partial P_{\rm s-}}{\partial v_i} = \left.\frac{\partial Z}{\partial \zeta}\right|_{-\widehat{h}_-} \frac{\partial \widehat{h}_-}{\partial v_i} \tag{103}$$

$$= z(-\widehat{h}_-)\frac{\partial \widehat{h}_-}{\partial v_i} \tag{104}$$

The pdf is non-negative, so eq.(104) is equivalent to eq.(33). ■

B Proofs for Section 4

Proof of proposition 3. We will prove the proposition for the upper robustness, $\widehat{h}_+(v, y_{\rm c})$ and supposing that $y(x, v)$ is monotonically *in*creasing in x. The proofs if $y(x, v)$ is monotonically *de*creasing in x, and for the lower robustness, are similar and will not be elaborated.

Since $\widehat{h}_+(v, y_{\rm c}) > \widehat{h}_+(v', y_{\rm c})$ we conclude that $\widehat{h}_+(v, y_{\rm c})$ is strictly positive. Since the robustness $\widehat{h}_+(v, y_{\rm c})$ is finite, there exists an x for which $y(x, v) = y_{\rm c}$. (If $y(x, v) < y_{\rm c}$ for all x then the robustness would be infinite). Recall that $y(x, v)$ increases monotonically with x. Denote the greatest x which satisfies the inequality $y(x, v) \le y_{\rm c}$ as $x_+(v)$, defined as the maximum x satisfying:

$$y[x_+(v),\ v] = y_{\rm c} \tag{105}$$

$x_+(v)$ must be finite since $\widehat{h}_+(v, y_{\rm c})$ is finite.

Define $x_+(v')$ as in eq.(105) if it exists. If $y(x, v') > y_{\rm c}$ for all x then $x_+(v')$ for eq.(105) does not exist. However, in this case $\Lambda_+(v')$ is an empty set and $P_{\rm s+}(v') = 0$ and the proposition is true. We need continue only with the case that $x_+(v')$ exists.

From the monotonicity of $y(x, v)$, we see that Λ_+ for the control v is the one-sided interval:

$$\Lambda_+(v) = (-\infty,\ x_+(v)] \tag{106}$$

Now consider the control v', whose upper robustness is finite. The info-gap model, $\mathcal{U}(h)$ in eq.(37), when h equals $\widehat{h}_+(v', y_{\rm c})$, is a closed interval. Since $y(x, v')$ increases monotonically in x we see that the upper bound of this interval must equal $x_+(v')$, which we denote:

$$x_+(v') = {\rm UB}\,\mathcal{U}\left[\widehat{h}_+(v', y_{\rm c})\right] \tag{107}$$

where we note that the upper bound in fact belongs to the interval.

Consider the two control vectors, v and v', which, by supposition, satisfy:

$$\widehat{h}_\times(v', y_{\rm c}) < \widehat{h}_\times(v, y_{\rm c}) \tag{108}$$

The nesting axiom implies that:

$$\mathcal{U}\left[\widehat{h}_+(v', y_{\rm c})\right] \subseteq \mathcal{U}\left[\widehat{h}_+(v, y_{\rm c})\right] \tag{109}$$

This relation, with eq.(107), implies:

$$x_+(v') \in \mathcal{U}\left[\widehat{h}_+(v, y_{\rm c})\right] \tag{110}$$

This relation, with eq.(107) applied to v, implies:

$$x_+(v') \le x_+(v) \tag{111}$$

which, with eq.(106), implies:

$$\Lambda_+(v') \subseteq \Lambda_+(v) \tag{112}$$

Hence:

$$P_{\rm s+}(v') \le P_{\rm s+}(v) \tag{113}$$

which concludes the proof for the upper robustness, assuming that y is monotonically increasing in x. ■

C Derivation for Section 5

We will derive the robustness function in eq.(51), defined in eq.(50), which includes eqs.(48) and (49) as a special case. We will consider the case that $\widetilde{\gamma} < 0$.

Let $\mu(h)$ denote the inner maximum in eq.(50). Note that $\mu(h) = \infty$ if $h \ge 1$ because $s = \gamma/\sigma^2$ and σ^2 can vanish for $h \ge 1$. Thus the robustness is never greater than one, and we need to evaluate $\mu(h)$ only for $h < 1$.

This inner maximum at uncertainty h occurs when s is either maximal or minimal. That is, the inner maximum occurs either when γ is maximal (that is, negative but close to zero) and σ^2 is also maximal, or when γ is minimal (very negative) and σ^2 is also minimal. That is:

$$\begin{aligned}
\mu(h) &= \max\left[\left(\max_{\gamma,\sigma^2\in\mathcal{U}(h)} s\right) - s_{\rm e},\ s_{\rm e} - \left(\min_{\gamma,\sigma^2\in\mathcal{U}(h)} s\right)\right] && (114)\\
&= \max\left[\frac{\widetilde{\gamma} + |\widetilde{\gamma}|h}{(1+h)\widetilde{\sigma}^2} - s_{\rm e},\ s_{\rm e} - \frac{\widetilde{\gamma} - |\widetilde{\gamma}|h}{(1-h)\widetilde{\sigma}^2}\right] && (115)\\
&= \max\left[\underbrace{\frac{1-h}{1+h}\widetilde{s} - s_{\rm e}}_{\mu_+(h)},\ \underbrace{s_{\rm e} - \frac{1+h}{1-h}\widetilde{s}}_{\mu_-(h)}\right] && (116)
\end{aligned}$$

which defines the functions $\mu_-(h)$ and $\mu_+(h)$, one of which may be negative.

Define the function $\widehat{h}_+(s_{\rm e}, r_{\rm c})$ as the solution for h of $\mu_+(h) = r_{\rm c}$. That is, $\widehat{h}_+(s_{\rm e}, r_{\rm c})$ is the inverse of $\mu_+(h)$. One readily shows that $\widehat{h}_+(s_{\rm e}, r_{\rm c})$ equals the lower line in eq.(51).

Likewise, define the function $\widehat{h}_-(s_{\rm e}, r_{\rm c})$ as the solution for h of $\mu_-(h) = r_{\rm c}$. That is, $\widehat{h}_-(s_{\rm e}, r_{\rm c})$ is the inverse of $\mu_-(h)$. One readily shows that $\widehat{h}_-(s_{\rm e}, r_{\rm c})$ equals the upper line in eq.(51).

$\mu(h)$ is the inverse of $\widehat{h}(s_{\rm e}, r_{\rm c})$. That is, $\widehat{h}(s_{\rm e}, r_{\rm c})$ is the solution for h of $\mu(h) = r_{\rm c}$. Since $\mu(h)$ equals the *larger* of $\mu_+(h)$ and $\mu_-(h)$, we see that $\widehat{h}(s_{\rm e}, r_{\rm c})$ is the *smaller* of $\widehat{h}_+(s_{\rm e}, r_{\rm c})$ and $\widehat{h}_-(s_{\rm e}, r_{\rm c})$, as illustrated in fig. 13. One readily shows that $\widehat{h}_-(s_{\rm e}, r_{\rm c}) \le \widehat{h}_+(s_{\rm e}, r_{\rm c})$ if and only if $\rho^2 \ge \zeta^2 - 1$.

Finally, we must select the non-negative parts of the curves, which results in the conditions $\rho \ge 1 - \zeta$ and $\rho \ge \zeta - 1$. This completes the derivation.

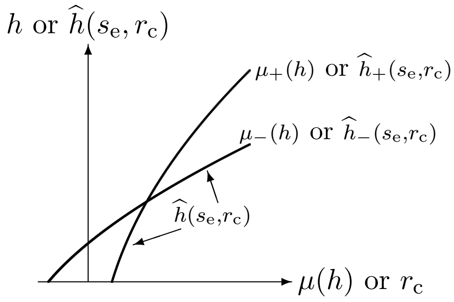

Figure 13: Illustration for appendix C.

D Proofs for Section 7

Proof of proposition 4. $\widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d})$ minimizes the expenditure with price vector p, so for any p':

$$p^T \widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d}) \le p^T \widehat{x}_{\rm r}(p', r_{\rm c}, \widehat{h}_{\rm d}) \tag{117}$$

Likewise, $\widehat{x}_{\rm r}(p', r_{\rm c}, \widehat{h}_{\rm d})$ minimizes the expenditure with price vector p', so for any p:

$$p'^T \widehat{x}_{\rm r}(p, r_{\rm c}, \widehat{h}_{\rm d}) \ge p'^T \widehat{x}_{\rm r}(p', r_{\rm c}, \widehat{h}_{\rm d}) \tag{118}$$

Subtracting (118) from (117) and re-arranging leads to the desired result, eq.(85). ■

Proof of lemma 1. (1) From the definition of the robustness function in eq.(78) and the continuity of $\mathcal{R}_*$ in h, and since x solves the RSP, we require that:

$$\mathcal{R}_*[x, \mathcal{F}(\widehat{h}(x, r_{\rm c}), \widetilde{f})] \ge r_{\rm c} \tag{119}$$

Suppose that:

$$\mathcal{R}_*[x, \mathcal{F}(\widehat{h}(x, r_{\rm c}), \widetilde{f})] > r_{\rm c} \tag{120}$$

By the info-gap axioms of nesting and linear expansion, we have: $h < h'$ implies that $\mathcal{F}(h, \widetilde{f}) \subset \mathcal{F}(h', \widetilde{f})$. Since $\mathcal{R}_*(x, A)$ is monotonically decreasing in A (eq.(74)) and continuous in h, (120) implies that there is an $h > \widehat{h}(x, r_{\rm c})$ such that:

$$\mathcal{R}_*[x, \mathcal{F}(h, \widetilde{f})] > r_{\rm c} \tag{121}$$

which contradicts the definition of $\widehat{h}(x, r_{\rm c})$. Hence

$$\mathcal{R}_*[x, \mathcal{F}(\widehat{h}(x, r_{\rm c}), \widetilde{f})] = r_{\rm c} \tag{122}$$

(2) Suppose that $\widehat{h}(x, r_{\rm c}) > \widehat{h}_{\rm d}$. Based on this supposition, and since $\widehat{h}(x, r_{\rm c})$ is continuous in x, there is an $x' \in X$ such that $p^T x' < p^T x$ and $\widehat{h}(x', r_{\rm c}) > \widehat{h}_{\rm d}$. This contradicts the supposition of the lemma that x solves the RSP. Hence there is no such x'. Consequently $\widehat{h}(x, r_{\rm c}) = \widehat{h}_{\rm d}$. ■

Proof of lemma 2. (1) From the definition of the opportuneness function in eq.(79) and the continuity of $\mathcal{R}^*$ in h, and since x solves the OWP, we require that:

$$\mathcal{R}^*[x, \mathcal{F}(\widehat{\beta}(x, r_{\rm w}), \widetilde{f})] \ge r_{\rm w} \tag{123}$$

Suppose that:

$$\mathcal{R}^*[x, \mathcal{F}(\widehat{\beta}(x, r_{\rm w}), \widetilde{f})] > r_{\rm w} \tag{124}$$

By the info-gap axioms of nesting and linear expansion, we have: $h < h'$ implies that $\mathcal{F}(h, \widetilde{f}) \subset \mathcal{F}(h', \widetilde{f})$. Since $\mathcal{R}^*(x, A)$ is monotonically increasing in A (eq.(75)) and continuous in h, (124) implies that there is an $h < \widehat{\beta}(x, r_{\rm w})$ such that:

$$\mathcal{R}^*[x, \mathcal{F}(h, \widetilde{f})] > r_{\rm w} \tag{125}$$

which contradicts the definition of $\widehat{\beta}(x, r_{\rm w})$. Hence

$$\mathcal{R}^*[x, \mathcal{F}(\widehat{\beta}(x, r_{\rm w}), \widetilde{f})] = r_{\rm w} \tag{126}$$

(2) Suppose that $p^T x < w$. Based on this supposition, and since $\widehat{\beta}(x, r_{\rm w})$ is non-satiated at x, there is an $x' \in X$ such that $\widehat{\beta}(x', r_{\rm w}) < \widehat{\beta}(x, r_{\rm w})$ and $p^T x' < w$. This contradicts the supposition of the lemma that x solves the OWP. Hence there is no such x'. Consequently $p^T x = w$. ∎

Proof of Proposition 5. $\widehat{h}_{\rm d} = \widehat{h}(\widehat{x}_{\rm r}, r_{\rm c})$ by lemma 1.

Suppose, contrary to the assertion of the proposition, that there is an $x'' \in X$ which solves the OWP with the specified w and $r_{\rm w}$, so that $p^T x'' \le w$, and for which:

$$\widehat{\beta}(x'', r_{\rm w}) < \widehat{\beta}(\widehat{x}_{\rm r}, r_{\rm w}) \tag{127}$$

which means that $\widehat{x}_{\rm r}$ does not solve the OWP. We will refer to this contradictory supposition as CS.

Given CS and the continuity of the opportuneness function $\widehat{\beta}(x, r_{\rm w})$, we see that there is an x' such that:

$$\widehat{\beta}(x', r_{\rm w}) < \widehat{\beta}(\widehat{x}_{\rm r}, r_{\rm w}) \quad \text{and} \quad p^T x' < w \tag{128}$$

That is, the opportuneness is better at x' than $\widehat{x}_{\rm r}$, and the expenditure with x' is strictly less than w.

By definition of the opportuneness function:

$$\widehat{\beta}(x', r_{\rm w}) = \min \left\{ h: \ \mathcal{R}^*[x', \mathcal{F}(h, \widetilde{f})] \ge \underbrace{\mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm w}} \right\} \tag{129}$$

and:

$$\widehat{\beta}(\widehat{x}_{\rm r}, r_{\rm w}) = \min \left\{ h: \ \mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(h, \widetilde{f})] \ge \underbrace{\mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm w}} \right\} = \widehat{h}_{\rm d} \tag{130}$$

Equality to $\widehat{h}_{\rm d}$ in eq.(130) arises as follows. It is evident from (130) that $\widehat{\beta}(\widehat{x}_{\rm r}, r_{\rm w}) \le \widehat{h}_{\rm d}$. From the nesting and linear expansion axioms of info-gap models we see that $h < \widehat{h}_{\rm d}$ implies that $\mathcal{F}(h, \widetilde{f}) \subset \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})$. Strict monotonicity of $\mathcal{R}^*(x, A)$ in h implies $\mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(h, \widetilde{f})] < \mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]$ for $h < \widehat{h}_{\rm d}$. Hence $\widehat{\beta}(\widehat{x}_{\rm r}, r_{\rm w})$ cannot be less than $\widehat{h}_{\rm d}$. Therefore $\widehat{\beta}(\widehat{x}_{\rm r}, r_{\rm w}) = \widehat{h}_{\rm d}$.

Relations (128)–(130), together with monotonicity of $\mathcal{R}^*(x, A)$ in h, imply:

$$\mathcal{R}^*[x', \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] > \mathcal{R}^*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] \tag{131}$$

By similar ordering of $\mathcal{R}_*$ and $\mathcal{R}^*$ this implies:

$$\mathcal{R}_*[x', \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] > \mathcal{R}_*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] \tag{132}$$

By definition of the robustness function:

$$\widehat{h}(x', r_{\rm c}) = \max \left\{ h: \ \mathcal{R}_*[x', \mathcal{F}(h, \widetilde{f})] \ge \underbrace{\mathcal{R}_*[\widehat{x}_{\rm r}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm c}} \right\} \tag{133}$$

The identity to $r_{\rm c}$ results from lemma 1.

Lemma 1 also implies that $\widehat{h}(\widehat{x}_{\rm r}, r_{\rm c}) = \widehat{h}_{\rm d}$. Relations (132) and (133), together with monotonicity of $\mathcal{R}_*(x, A)$ in the uncertainty sets A, relation (74), and continuity in h, imply:

$$\widehat{h}(x', r_{\rm c}) > \widehat{h}_{\rm d} = \widehat{h}(\widehat{x}_{\rm r}, r_{\rm c}) \tag{134}$$

Since $p^T x' < w$ (eq.(128)), while (by definition of w) $p^T \widehat{x}_{\rm r} = w$, we see from (134) that CS implies that $\widehat{x}_{\rm r}$ is **not** a solution of the RSP. This contradicts the condition of the proposition, so CS is false.

Hence there is no x'' satisfying relation (127). We conclude that $\widehat{x}_{\rm r}$ solves the OWP: it minimizes the opportuneness function and $p^T\widehat{x}_{\rm r} = w$. ■

Proof of Proposition 6. Lemma 2 implies $p^T\widehat{x}_{\rm o} = w$.

Suppose, contrary to the assertion of the proposition, that there is an $x'' \in X$ which solves the RSP with the specified $r_{\rm c}$ and $\widehat{h}_{\rm d}$, so that $\widehat{h}(x'', r_{\rm c}) \ge \widehat{h}_{\rm d}$ and for which:

$$p^T x'' < p^T\widehat{x}_{\rm o} \tag{135}$$

which means that $\widehat{x}_{\rm o}$ does not solve the RSP. We will refer to this contradictory supposition as CS.

Given CS and the non-satiation of the robustness function $\widehat{h}(x, r_{\rm c})$, we see that there is an x' such that:

$$p^T x' < p^T\widehat{x}_{\rm o} \quad \text{and} \quad \widehat{h}(x', r_{\rm c}) > \widehat{h}_{\rm d} \tag{136}$$

That is, the expenditure is lower at x' than $\widehat{x}_{\rm o}$, and the robustness with x' is strictly greater than $\widehat{h}_{\rm d}$.

By definition of the robustness function:

$$\widehat{h}(x', r_{\rm c}) = \max\left\{ h:\ \mathcal{R}_*[x', \mathcal{F}(h, \widetilde{f})] \ge \underbrace{\mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm c}} \right\} \tag{137}$$

and

$$\widehat{h}(\widehat{x}_{\rm o}, r_{\rm c}) = \max\left\{ h:\ \mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(h, \widetilde{f})] \ge \underbrace{\mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm c}} \right\} = \widehat{h}_{\rm d} \tag{138}$$

Equality to $\widehat{h}_{\rm d}$ in eq.(138) arises as follows. It is evident from (138) that $\widehat{h}(\widehat{x}_{\rm o}, r_{\rm c}) \ge \widehat{h}_{\rm d}$. From the nesting and linear expansion axioms of info-gap models we see that $\widehat{h}_{\rm d} < h$ implies that $\mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f}) \subset \mathcal{F}(h, \widetilde{f})$. Strict monotonicity of $\mathcal{R}_*(x, A)$ in h implies $\mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(h, \widetilde{f})] < \mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]$ for $\widehat{h}_{\rm d} < h$. Hence $\widehat{h}(\widehat{x}_{\rm o}, r_{\rm c})$ cannot be greater than $\widehat{h}_{\rm d}$. Therefore $\widehat{h}(\widehat{x}_{\rm o}, r_{\rm c}) = \widehat{h}_{\rm d}$.

Relations (136)–(138), together with the monotonicity of $\mathcal{R}_*(x, A)$ in h, imply:

$$\mathcal{R}_*[x', \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] > \mathcal{R}_*[\widehat{x}_{\rm o}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] \tag{139}$$

By similar ordering of $\mathcal{R}_*$ and $\mathcal{R}^*$:

$$\mathcal{R}^*[x', \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] > \mathcal{R}^*[\widehat{x}_{\rm o}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})] \tag{140}$$

By definition of the opportuneness function:

$$\widehat{\beta}(x', r_{\rm w}) = \min\left\{ h:\ \mathcal{R}^*[x', \mathcal{F}(h, \widetilde{f})] \ge \underbrace{\mathcal{R}^*[\widehat{x}_{\rm o}, \mathcal{F}(\widehat{h}_{\rm d}, \widetilde{f})]}_{r_{\rm w}} \right\} \tag{141}$$

The identity of $r_{\rm w}$ in (141) results from lemma 2.

By definition, $\widehat{h}_{\rm d} = \widehat{\beta}(\widehat{x}_{\rm o}, r_{\rm w})$. Relations (140) and (141), together with the monotonity of $\mathcal{R}^*(x, A)$ in the sets A, relation (75), and continuity in h, imply:

$$\widehat{\beta}(x', r_{\rm w}) < \widehat{h}_{\rm d} = \widehat{\beta}(\widehat{x}_{\rm o}, r_{\rm w}) \tag{142}$$

Since $p^T x' < p^T\widehat{x}_{\rm o}$ (from eq.(136)), and since $p^T\widehat{x}_{\rm o} = w$, (142) contradicts the supposition of the proposition that $\widehat{x}_{\rm o}$ solves the OWP. Hence there cannot be such an x' and the CS is false. Hence $\widehat{x}_{\rm o}$ solves the RSP. ■