DUAL SCORING

CHRISTOPHER P. CHAMBERS, PAUL J. HEALY, AND NICOLAS S. LAMBERT

Abstract. We provide a full dual characterization of the optimal solutions for the optimal choices of quasiconcave, continuous, and increasing utility functions facing proper scoring rules. The dual characterization leverages the notion of indirect utility. We use this characterization to construct a scoring rule which uniformly bounds misreports when risk-aversion is bounded above.

1. Introduction

In this note, we fully characterize the optimal decisions for individuals with convex preferences over state-contingent monetary payoffs who face proper scoring rules. In the economics tradition, a proper scoring rule is understood as an incentive device for obtaining probabilistic assessments from rational, risk-neutral agents in the fact of subjective uncertainty. An agent is hypothesized to have a probabilistic assessment over some observable set of states of the world. A *scoring rule* can be thought of a menu of state-contingent monetary payoffs, indexed by the set of possible probabilistic assessments, from which the agent is asked to choose. If the scoring rule is proper, then an agent with a particular subjective probability maximizes her welfare by choosing from the menu that state-contingent payoff associated with her subjective probability.[1] There is a long tradition of using these devices in experimental settings in order to obtain probabilistic assessments from subjects; see *e.g.* Nyarko and Schotter (2002).

Of course, it has long been recognized that in the case of subjective uncertainty, individuals may not hold probabilistic beliefs. A classical thought experiment is due to Ellsberg (1961), and is known as the Ellsberg Paradox. Ellsberg's observation, usually referred to as an instance of "ambiguity aversion," is viewed as both positively and normatively compelling by many economists. Moreover, it has

(Chambers) Department of Economics, University of California at San Diego
(Healy) Department of Economics, Ohio State University
(Lambert) Graduate School of Business, Stanford University

[1]This interpretation is firmly in line with the revealed preference tradition, whereby an agent may not necessarily perceive that she holds a probabilistic assessment, but nevertheless her behavior is consistent with such a belief. To this end, we need not ask the agent to "report" a belief, but rather to choose from a set. It is usually easier to use the language of "reporting" beliefs, however. Further to this end, there is no real reason to require that there be a unique state-contingent payoff associated with each possible probability, but it is conventional to do so, and not much is gained by weakening this assumption.

spawned a vast literature consisting of models which rationalize the choices made. This literature is too large to survey here, but one classical such model is termed the *multiple priors* model, first proposed in economics and axiomatized by Gilboa and Schmeidler (1989). This model and its underlying axiomatic development are directly inspired by the robust statistics literature (see for example Huber (1981), Proposition 2.1 of Section 10.2). Of course, there may also be reason to doubt the conclusions of the multiple priors model, at least behaviorally.[2]

Motivated by the preceding characterizations, this paper seeks to provide a full characterization of optimal behavior for individuals with "standard" economic utility functions. These are utility functions which are quasiconcave, continuous, and increase strictly when consumption in all states increases strictly. Such utility functions need not adhere to any concept of "likelihood" or probability. Instead, they are intended to represent choice behavior when an individual may not conform to any particular decision-theoretic model.

A direct precedent to this paper is Grünwald and Dawid (2004), which describes two classical approaches to the problem of robust statistics. Given is a convex and compact set of probability measures; and the goal is to "select" a probability from this set. Assume that each state-contingent monetary payoff is evaluated according to the minimal expected value according to all probabilities in the set. These authors use minimax theory to obtain a basic duality characterization. Namely, an individual seeking to maximize the minimal expected payoff across the set by choosing optimally from a menu associated with a proper scoring rule can be viewed *as if* they are minimizing a strictly convex function on the set of probabilities. Each proper scoring rule is associated with its own strictly convex function. For example, the authors observe the duality between the logarithmic scoring rule (Good, 1952) and the entropy function.[3] The result in this note directly generalizes Grünwald and Dawid (2004) in a finite states environment.

Before proceeding, we need to discuss a notion of duality which is standard in economics, but may be less familiar generally. This is the duality between a *direct* utility function and an *indirect* utility function. There are several related such dualities, but we focus on one which has recently been fruitfully exploited by Cerreia-Vioglio et al. (2011b).[4] For a given utility function U, usually understood

[2]Chambers et al. (2016) propose a "revealed preference test" for the model in a demand-based setting. With data from Hey and Pace (2014), the multiple priors model seems to perform no better than the subjective expected utility model, even when allowing subjects to exhibit risk aversion.

[3]See also Chambers (2008), who proves the same result in a much simplified environment. Chambers was unaware of Grünwald and Dawid (2004) when he published his result.

[4]This is arguably the "right" duality notion to use in our case, where we want to allow scoring rules to potentially pay negative monetary amounts. These notions have existed in economics at least since Konüs (1939) and Ville and Newman (1951-1952), which are translations of earlier foreign language works, whereas Roy (1947) arguably popularized the concept. de Finetti (1949)

as a "direct" utility over state-contingent payoffs, we define an "indirect utility" over price-wealth pairs in the natural way:

$$G(p, w) = \sup\{U(x) : p \cdot x \leq w\}.$$

That is, $G(p, w)$ asserts the maximal utility achievable by an individual with utility U when "market prices" are p, and the wealth available for expenditure is w. If p are interpreted as probabilities, then equivalently, the indirect utility gives the highest utility achievable to U by a state-contingent payoff with expectation at most w.[5]

Our main result is as follows. With any proper scoring rule, we can associate the so-called "value function," V, as the expected payoff to a risk-neutral agent with probability p who optimizes. We show that for any decision maker U who has a quasiconcave, weakly increasing, and continuous direct utility function, the unique optimal announced p for scoring rule f coincides with the unique p which minimizes $G(p, V(p))$.

As we see it, there are four main reasons why such a duality result is interesting.

- First, as an academic exercise, it directly generalizes the results of Grünwald and Dawid (2004), to a broad class of preferences over uncertain prospects.
- Second, many preference specifications in economics are defined only via their indirect utility functions. Chief among these preference classes is the *Gorman polar form* (Gorman (1961)), described below in detail and used heavily in applied modeling.
- Third, the duality presented here allows simple proofs of previously unknown results. For example, Theorem 3 below establishes that if risk aversion is not "too high," we can uniformly bound "misreports." Other such results can be similarly obtained.
- Fourth, in an avenue we have not explored heavily here; duality results are generally quite useful in comparative static exercises. Results from Cerreia-Vioglio et al. (2011b) can be used to study comparative notions of risk aversion (for example, that of Yaari (1969)), and how subjects behavior when facing different scoring rules changes when becoming more risk averse.

---

established an early duality result using such functions. The related duality studied by Shepherd (1970); Lau (1969); Cornes (1992); Weymark (1980) requires all payoffs to be nonnegative.

[5]There is a sense in which the indirect utility is related to the notion of Fenchel conjugation. To see this, observe that the Fenchel conjugate of a monotonic U can be written as: $G(y) = \inf_x x \cdot y - U(x)$, for $y \geq 0$ (not necessarily a probability). We can think of the negation of this problem, $-G(y) = \sup_x U(x) - x \cdot y$. The value of this problem is clearly equivalent to the value of the problem $\sup_{(x,t)} U(x) + t$ subject to the constraint that $t + x \cdot y \leq 0$. This is a "cardinal" version of the indirect utility when there is a "numeraire" good t with a fixed price of 1, and when total wealth is zero. When the numeraire good has a fixed price of 1, we cannot renormalize prices to sum to 1.

As a first point, to see why this directly generalizes Grünwald and Dawid (2004), observe that when $U(x) = \min_{p \in P} p \cdot x$, then $G(p, w) = w$ when $p \in P$ and $+\infty$ otherwise. In other words, the uniquely optimal announcement for such a utility function is equivalently given by $\min_{p \in P} V(p)$. While Grünwald and Dawid (2004) and Chambers (2008) each rely on versions of the minimax theorem, the proof here is simpler and directly leverages the separating hyperplane theorem.[6]

Aside from the multiple priors case, there are other examples of utility specifications where the duality is especially simple. One such example is the class of *translation invariant* utility functions. These are utility functions for which wealth effects are absent. This is operationalized by assuming that adding a dollar in each state of the world translates into an additional unit of utility. In this case, the dual minimization problem takes the form of minimizing the sum of the value function, and some convex function of probabilities, specific to the utility in question.

A particular special case of translation invariant preferences are the constant absolute risk aversion preferences. These are the unique subjective expected utility preferences which can be expressed in a translation invariant form. The convex function alluded to in the previous paragraph in this case is the *relative entropy* function, relative to the subjective probability in this case. Bickel (2007) describes the optimization problem for such individuals facing scoring rules.

A further generalization of translation invariant preferences is provided by the Gorman polar form preferences (Gorman (1961)). In our context, these are preferences for which there is some $\beta \geq 0$, $\beta \neq 0$ for which adding t units of β to consumption adds t units of utility. These preferences are highly useful in applied modeling, as they allow one to meaningfully describe a "group" of individuals as a single individual, behaving in her own best interest.

Finally, we use our main result to establish that we can elicit, to an arbitrarily high degree of precision, an individual's subjective probability when risk-aversion is known to be bounded above by some prespecified level. As far as we know, this is the first result of this type in the literature.

Section 2 presents our main result, as well as a related result stemming from Roy's identity. Section 3 presents several examples, illustrating how our result can be used. Section 4 describes a method of bounding misreports to any arbitrarily high degree of precision, starting from any proper scoring rule. Finally, section 5 concludes.

2. Optimal values for general decision models

Let Ω be a finite set of *states* and $\Delta(\Omega)$ the set of probability measures on Ω. A *scoring rule* is a mapping $f : \Delta(\Omega) \to \mathbb{R}^{\Omega}$. It is *proper* if for all $p, p' \in \Delta(\Omega)$,

[6] A proof can also be established using minimax.

$p \cdot f(p) \geq p \cdot f(p')$ and *strictly proper* if for all $p, p' \in \Delta(\Omega)$ for which $p \neq p'$, we have $p \cdot f(p) > p \cdot f(p')$. For a strictly proper scoring rule f, define the associated *value function* $V : \Delta(\Omega) \to \mathbb{R}$ by $V(p) = p \cdot f(p) = \sup_{p' \in \Delta(\Omega)} p \cdot f(p')$.[7]

We assume that $U : \mathbb{R}^\Omega$ is weakly increasing, quasiconcave, and continuous.[8] Call such a utility *standard*. The *indirect utility* is defined by $G : \Delta(\Omega) \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, given by $G(p, w) = \sup\{U(x) : p \cdot x \leq w\}$. Quite generally, we have $U(x) = \inf_{p \in \Delta(\Omega)} G(p, p \cdot x)$.[9] This is the duality we use for our results.

We emphasize that the only useful economic content of this model are the properties of increasingness and quasiconcavity of preference. Thus, this model can incorporate decision makers with the following types of utility functions; *i.e.* risk-averse expected utility:

$$U(x) = u^{-1}\left(\sum_{\omega \in \Omega} u(x_\omega) p(\omega)\right),$$

where u is increasing and concave, and p is a probability measure. This is a classical certainty-equivalent representation of a risk-averse expected utility maximizer. Of course, the representation of this preference would involve a complex G function, but there is such a function; in fact, an explicit representation of such functions appears in Cerreia-Vioglio et al. (2011b). However, the class of utility functions we can accommodate is much broader than the risk-averse expected utility class.

The following is our main result. It claims that for any standard utility and any strictly incentive compatible scoring rule, there is a unique optimal announcement $p^* \in \Delta(\Omega)$, and further, this unique announcement can be arrived at using dual techniques. Namely, it solves the problem $\min_p G(p, V(p))$. The latter problem is often easier to describe, and can allow for richer comparative statics.

Theorem 1. *Suppose that f is a continuous and strictly incentive compatible scoring rule. Then for a standard utility with indirect utility function G, there is a unique solution to the problem $\max_{p^* \in \Delta(\Omega)} U(f(p^*))$, and it is given by $\arg\min_{p^* \in \Delta(\Omega)} G(p^*, V(p^*))$, where V is the value function associated with f.*

We first sketch the idea of the proof using Figure 1. In this diagram Ω contains only two states, so $X = \mathbb{R}^2$. By strict properness, the range of the scoring rule f forms the upper boundary of a strictly convex set. The problem of maximizing U over this set results in an optimal payment vector $f(p^*)$, which is achieved by

[7]For a general, non-incentive compatible scoring rule, one can define the value function via $V(p) = \sup_{p'} p \cdot f(p')$.

[8]Weakly increasing: if $x_\omega > y_\omega$ for all $\omega \in \Omega$, then $U(x) > U(y)$.

[9]This duality is related to, but distinct from, the duality presented in Shepherd (1970); Lau (1969); Cornes (1992); Weymark (1980); which is a duality for functions on the nonnegative orthant. See Cerreia-Vioglio et al. (2011b) for details.

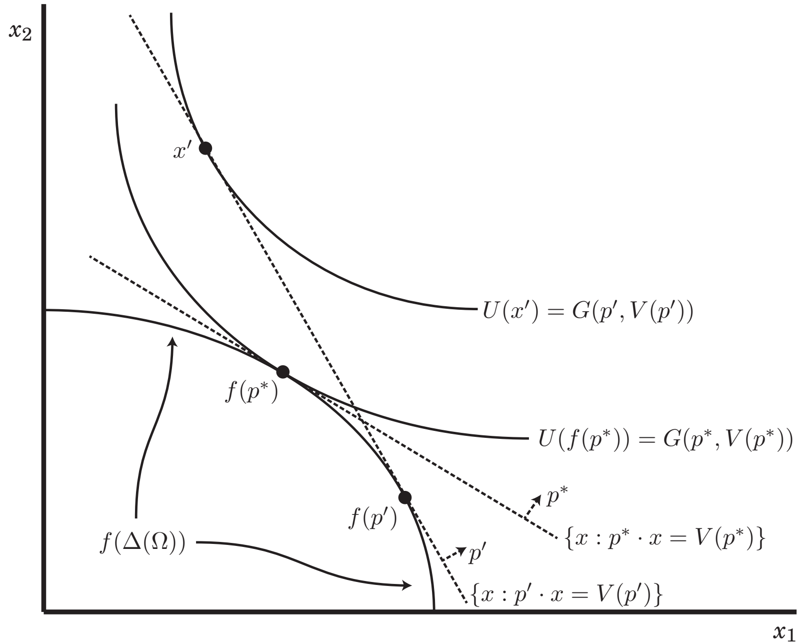

Figure 1. An illustration of the main theorem.

announcing the distribution p^*.[10] The payment of $f(p^*)$ gives an expected value (under p^*) of $V(p^*) = p^* \cdot f(p^*)$.

Now consider the 'indirect utility' maximization program: For any p we calculate the expectation of the scoring rule under p—which equals $V(p)$—and ask what point $x \in \mathbb{R}^2$ maximizes U subject to the constraint that $p \cdot x \leq V(p)$. This constraint is shown for p^* and p' as dashed lines. For p^* the maximizing point is again $f(p^*)$, which gives indirect utility $G(p^*, V(p^*)) = U(f(p^*))$. But for p' the constraint includes points strictly better than $f(p^*)$ for the decision maker, so the maximum indirect utility (which obtains at x') is $G(p', V(p')) > U(f(p^*)) = G(p^*, V(p^*))$. By the strictly convex shape of $f(\Delta(\Omega))$ this is true for any $p \neq p^*$. Thus, the original utility-maximizing point $f(p^*)$ is the unique minimizer of the indirect utility function $G(p, V(p))$.

Proof. As a first step, consider the set $K = \text{co}(f(\Delta(\Omega)))$ (the convex hull of the image of f). Observe that K is itself compact, since $\mathbb{R}^\Omega$ is finite-dimensional (Corollary 5.18 of Aliprantis and Border (1999)). We will show that there is a unique maximizer of f across the set K, and that this maximizer coincides with $\arg\max_p U(f(p))$.

---

[10]We reiterate that the decision maker need not have 'true' beliefs p^*—or any probabilistic beliefs at all. Here, p^* is interpreted only as the decision maker's optimal announcement given f.

So let $x^* \in \arg\max_{x\in K} U(x)$. Such a maximizer exists due to continuity of f and compactness of K. Let $Y = \{y : U(y) > U(x^*)\}$, which is open (by continuity) and convex (by quasiconcavity). The sets K and Y can therefore be separated by a hyperplane (Theorem 5.50 of Aliprantis and Border (1999)). Clearly, the hyperplane can be normalized to have direction in $\Delta(\Omega)$, by the fact that U is increasing. Let us call this direction p^*. Observe that the hyperplane p^* passes through x^*, as for any $\epsilon > 0$, $x^* + \epsilon(1, \ldots, 1) \in Y$. Hence, conclude that for all $x \in K$, $p^* \cdot x \leq p^* \cdot x^*$; *i.e.* x^* maximizes $p^* \cdot x$ subject to $x \in K$. Clearly $f(p^*) \in K$ satisfies this inequality. We claim that it is the unique such element of K. So, let $\hat{x} \in K$, where $\hat{x} \neq f(p^*)$. Then there are $p_1, \ldots, p_n \in \Delta(\Omega)$, not all equal to p^* and $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$ for which $\hat{x} = \sum_i \lambda_i f(p_i)$. But by strict incentive compatibility, we then obtain $p^* \cdot \hat{x} < p^* \cdot f(p^*)$, contradicting the fact that $\hat{x}$ maximizes $p^* \cdot x$ subject to $x \in K$. So $x^* = f(p^*)$. Further, $f(p^*)$ is the unique maximizer of U in K. To see this, observe that by continuity and monotonicity, the closure of Y is $\{y : U(y) \geq U(f(p^*))\}$. Hence if $x' \in \arg\max_{x\in K} U(x)$, then $p^* \cdot x' \geq p^* \cdot f(p^*)$, which we have shown to be impossible.

Since p^* separates K and Y, and again by continuity, we have that $U(y) \geq U(f(p^*))$ implies that $p^* \cdot y \geq p^* \cdot f(p^*)$. We claim that this implies $U(f(p^*)) = G(p^*, p^* \cdot f(p^*))$. To see this, suppose by means of contradiction that there is y for which $p^* \cdot y \leq p^* \cdot f(p^*)$ and $U(y) > U(f(p^*))$. By continuity of U, we conclude that there is y^* for which $p^* \cdot y^* < p^* \cdot f(p^*)$ and $U(y^*) > U(f(p^*))$, contradicting the fact that $U(y^*) \geq U(f(p^*))$ implies $p^* \cdot y^* \geq p^* \cdot U(f(p^*))$. So, $U(f(p^*)) = G(p^*, p^* \cdot f(p^*))$.

Finally, let $\hat{p} \neq p^*$. Then $\hat{p} \cdot f(p^*) < \hat{p} \cdot f(\hat{p})$, by strict incentive compatibility. Therefore, there is $\epsilon > 0$ for which $\hat{p} \cdot (f(p^*) + \epsilon(1, \ldots, 1)) < \hat{p} \cdot f(\hat{p})$, and since U is monotonic, we then conclude that $G(\hat{p}, \hat{p} \cdot f(\hat{p})) > U(f(p^*))$.

Conclude then that p^* uniquely solves $\min_p G(p, p \cdot f(p))$. □

Theorem 1 can be generalized to include scoring rules which take infinite-valued payoffs (such as the classical logarithmic scoring rule). In this case, existence of an optimal announcement is not guaranteed, but when there is such an announcement, the duality should hold.

The result actually says little more than that the announced p^* separates the convex hull of the image of payoffs of the scoring rule, and the upper contour set of the preference. But it leads to many interesting conclusions when written in this form. For example, by exploiting Roy's identity (Roy, 1947) and the fact that a scoring rule is a subdifferential of its homothetic extension, we get the following:

Corollary 2. *If U is standard, and f is a continuous and strictly incentive compatible scoring rule, then if p^* solves $p^* \in \arg\max_{p\in\Delta(\Omega)} U(f(p))$, we have $f(p^*) \in \arg\max_{\{x:p^*\cdot x\leq V(p^*)\}} U(x)$.*

The preceding corollary states that a standard decision maker will choose to announce the p^* at which her Walrasian demand (when given wealth $V(p^*)$) includes $f(p^*)$. In this sense, this result illustrates the connection between the optimal choice of a decision maker from a scoring rule, and the equilibrium price in a Robinson Crusoe economy (the output of the scoring rule playing the role of a technology here).

3. Examples

Examples follow. Most of the characterizations of the G functions here are taken directly from Cerreia-Vioglio et al. (2011b).

Example 1. [**Translation Invariant, or Variational Preferences**]

Consider the variational preferences model of Maccheroni et al. (2006); applied to our setting, these preferences are those which can be written as preferences which are translation invariant in the sense that for all $x \in \mathbb{R}^\Omega$ and all $t \in \mathbb{R}$, $U(x + t(1,\ldots,1)) = U(x) + t$. In this case, it is easy to see that we can write $G(p,w) = w + c(p)$, for some proper convex (possibly infinite-valued) lower semicontinuous function c. This model incorporates the multiple priors model in the case where $c(p)$ is the convex-analytic indicator function of a convex set of priors P.[11] Writing down our formula, we want to solve

$$\arg\min_{p^*\in\Delta(\Omega)} \left(V(p^*) + c(p^*)\right).$$

Here, c has the interpretation of a *certainty equivalent*: $c(p) - c(q)$ measures the sure amount a decision maker holding only a riskless asset would pay to move from state prices q to prices p. To understand this claim, consider a decision maker maximizing $U(x)$ subject to $p \cdot x \leq w$, where U has a certainty equivalent form (so that $U(x_1 + t, \ldots, x_n + t) = U(x) + t$). We observe that $c(p) = \max_{x:p\cdot x\leq p\cdot 0} U(x)$. Hence, measured in terms of money, $c(p)$ is the value of facing prices p when endowment is 0 (riskless). To move from prices q to p, the individual would offer to pay $c(p) - c(q)$. And this willingness to pay is independent of riskless wealth, by the variational form.

For the special case of multiple priors, $c = 0$ on the set of priors P, and is otherwise infinite. Hence, a multiple priors agent will always announce the probability in the set of priors which uniquely minimizes the value function $V(p^*)$; this is one of the main results of Grünwald and Dawid (2004). A risk neutral subjective expected

[11] In other words, it equals 0 for $p \in P$ and $+\infty$ otherwise.

utility agent with prior p^* has a c function which is infinite valued except at the point p^*, where it is equal to 0. Hence, a risk neutral subjective expected utility agent always announces their "true belief" p^*.

Example 2. [**Subjective expected utility: CARA**]

For another special case, consider the subjective expected utility maximizer with CARA utility index $u(x) = -\exp(-\alpha x)$ where $\alpha > 0$ and subjective probability π. We can consider the certainty equivalent utility representation of this preference:

$$U(x) = \alpha^{-1} \ln \left(\sum_{\omega \in \Omega} u(x_\omega) \pi(\omega) \right).$$

It turns out that the G function used in this environment has a very special form. First, it is a special case of the variational preference model (as CARA means that a preference is additively separable and translation invariant, the latter being the main characteristic of the variational model). Second, it is a special case for which the function $c(p)$ is given by a scaling of a *relative entropy* or *Kullback Leibler* function.

Formally, if q is absolutely continuous with respect to π, say that $R(q \parallel \pi) = \sum_{\omega: q(\omega)>0} \log \left(\frac{q(\omega)}{\pi(\omega)} \right) q(\omega)$, and otherwise $R(q \parallel \pi) = \infty$. CARA-preferences with parameter $\alpha > 0$ and subjective probability π have a representation with the following G function: $G(w, p) = w + \alpha^{-1} R(p \parallel \pi)$. Hence, a CARA individual will always solve the optimal scoring rule announcement problem by choosing to minimize the function $V(p^*) + \alpha^{-1} R(p^* \parallel \pi)$. Observe that $R(p^* \parallel \pi) = 0$ only when $p^* = \pi$, so that when $\alpha \to 0$, the problem tends to a risk-neutral agent with subjective probability π.

The observation that the relative entropy function leads to CARA-style preferences in this context seems to have been first made by Strzalecki (2011).

Example 3. [**CARA and multiple priors**]

One can combine CARA models with multiple prior style models. Suppose that we consider a risk-averse decision maker with ambiguity-style preferences, so that the individual has a utility index $u(x) = -\exp(-\alpha x)$ of the CARA form, and a set of priors $P \subseteq \Delta(\Omega)$. It is straightforward to establish that such an individual can be represented with a (certainty equivalent dual) indirect utility function where $c(p) = \inf_{q \in P} \alpha^{-1} R(p \parallel q)$.

Example 4. [**Gorman polar form and generalized translation invariance**]

Gorman (1961) provided necessary and sufficient conditions for utility functions to have Engel curves which are straight lines.[12] We focus on the case in which the Engel curves are parallel across different prices, since our consumption space is

[12]An *Engel curve* is the set of Walrasian demands for a fixed price as wealth varies.

unbounded both above and below. Gorman originally proposed the family in order to meaningfully talk about "representative consumers," though they also have a natural interpretation of utility functions for which wealth effects are absent with respect to some "numeraire" bundle. That is, we want there to be a numeraire bundle $\beta \in \mathbb{R}^\Omega$ for which for all $t \in \mathbb{R}$ and all $x \in \mathbb{R}^\Omega$, $U(x + \beta t) = U(x) + t$.

Gorman called his utility functions *polar form*, since they are defined in terms of indirect utility. Therefore these form a natural class where the duality result is useful.

Formally, let us take $\beta \in \mathbb{R}^\Omega$, where $\beta \geq 0$ and $\beta \neq 0$. Let $\beta_+ = \{x \in \mathbb{R}^\Omega_+ : \sum_\omega x_\omega = 1\}$. Let $c : \beta_+ \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous and convex function. Then define

$$G(p, w) = \begin{cases} \frac{w}{\beta \cdot p} + c(\frac{p}{\beta \cdot p}) \text{ if } \beta \cdot p \neq 0 \\ +\infty \text{ otherwise} \end{cases} .$$

This specifies an indirect utility in Gorman polar form. Note the obvious connection with the class of translation invariant preferences, described above. Translation invariant preferences are those which are in Gorman polar form with $\beta = (1, \ldots, 1)$.

4. Application: Bounding "misreports"

Consider the CARA model described above in example 2, in a subjective expected utility framework. Nontrivial risk attitudes lead individuals to "misreport" their true subjective belief when facing proper scoring rules. Our goal here is to establish that we can construct a proper scoring rule that is robust to such misreporting. To this end, suppose that our goal as experimentalists is to elicit the decision maker's subjective probability within some prespecified tolerance for error.

Specifically, let us define, for a given $\alpha > 0$ and probability measure π, the utility function:

$$U^\alpha_\pi(x) = -\sum_\omega \exp(-\alpha x_\omega)\pi(\{\omega\}).$$

For two measures $p, q \in \Delta(\Omega)$, we let $d(p, q) = \max_{E \subseteq \Omega} |p(E) - q(E)|$ represent the total variation metric. We use the metric d to measure the degree of misreporting.

Theorem 3. *For every $\alpha^* \geq 0$, every $\epsilon > 0$, and every proper scoring rule f, there is $\gamma > 0$ for which for all $0 \leq \alpha \leq \alpha^*$ and all $\pi \in \Delta(\Omega)$, the solution:*

$$p(\alpha, \pi, \gamma) = \arg\max U^\pi_\alpha(\gamma f(p))$$

satisfies $d(p(\alpha, \pi, \gamma), \pi) < \epsilon$.

The interpretation of this result is that, so long as we are willing to assume risk aversion is bounded above, we can always guarantee that we have elicited subjects' true subjective probabilities within some prespecified tolerance. Observe that the bound ϵ is uniform across all possible π.

Proof. Observe that we can without loss discuss maximizing any monotonic transformation of U_α^π rather than U_α^π itself. To this end, use the representation of CARA preferences described in the example. Let V be the value function associated with f, and let $R(p \parallel \pi)$ be the relative entropy function described above. By Theorem 1, it follows that we are searching for $\gamma > 0$ for which the solutions to $\arg\min_p \gamma V(p) + \alpha^{-1} R(p \parallel \pi)$ satisfy the requisite property.

Next, for any γ, if $\alpha' > \alpha$, then

$$R(p(\alpha, \pi, \gamma) \parallel \pi) \leq R(p(\alpha', \pi, \gamma) \parallel \pi)$$

simply by minimization.[13]

Similarly, if $\gamma' < \gamma$ we have $R(p(\alpha, \pi, \gamma) \parallel \pi) < R(p(\alpha, \pi, \gamma') \parallel \pi)$.

Next, observe that if $\gamma = 0$, π uniquely minimizes $\alpha^{-1} R(p \parallel \pi)$ across p, and further that $R(\pi \parallel \pi) = 0$.[14] Finally, by a Maximum Theorem-style argument, we know that $R(p(\alpha, \pi, \gamma) \parallel \pi)$ is continuous in π for each α, γ and that $R(p(\alpha, \pi, \gamma) \parallel \pi) \to 0$ as $\gamma \to 0$.[15] = Viewed as a function of π, we then observe that for any α, $R(p(\alpha, \pi, \gamma)) \to 0$ as $\gamma \to 0$; hence, we have a sequence of continuous functions which converge monotonically on a compact set; by Dini's Theorem (Berge (1997), p. 106) we may choose γ so that $R(p(\alpha^*, \pi, \gamma) \parallel \pi) < \frac{\epsilon^2}{2\ln(2)}$ for all π.

---

[13]Suppose by means of contradiction that

$$R(p(\alpha', \pi, \gamma) \parallel \pi) < R(p(\alpha, \pi, \gamma) \parallel \pi).$$

Then we have

$$(\alpha^{-1} - \alpha'^{-1}) R(p(\alpha', \pi, \gamma) \parallel \pi) < (\alpha^{-1} - \alpha'^{-1}) R(p(\alpha', \pi, \gamma) \parallel \pi)$$

and

$$\gamma V(p) + \alpha'^{-1} R(p(\alpha', \pi, \gamma) \parallel \pi) \leq \gamma V(p) + \alpha'^{-1} R(p(\alpha, \pi, \gamma) \parallel \pi),$$

so that

$$\gamma V(p) + \alpha^{-1} R(p(\alpha', \pi, \gamma) \parallel \pi) < \gamma V(p) + \alpha^{-1} R(p(\alpha, \pi, \gamma) \parallel \pi),$$

a contradiction.

[14]This does not mean that if there are no incentives, π will be the uniquely optimal choice for an individual.

[15]For the Maximum Theorem, see Berge (1997), p. 116. The result there does not apply verbatim since R can be infinite-valued. However, it is straightforward to establish that $p(\alpha, \pi, \gamma)$ is continuous in both π and γ. Thus, suppose that $(\pi_n, \gamma_n) \to (\pi^*, \gamma^*)$. Take any subsequence $p(\alpha, \pi_{n_k}, \gamma_{n_k})$, and let $p(\alpha, \pi_{n_{k_j}}, \gamma_{n_{k_j}})$ be a convergent subsequence, say to p^*. We will show that $p^* = p(\alpha, \pi^*, \gamma^*)$. So, let $\hat{p}$ be arbitrary; observe that

$$\gamma_{n_{k_j}} V(p(\alpha, \pi_{n_{k_j}}, \gamma_{n_{k_j}})) + \alpha^{-1} R(p(\alpha, \pi_{n_{k_j}}, \gamma_{n_{k_j}}) \parallel \pi_{n_{k_j}}) \leq \gamma_{n_{k_j}} V(\hat{p}) + \alpha^{-1} R(\hat{p} \parallel \pi_{n_{k_j}})$$

and take limits, using the fact that V is continuous and R continuous on its effective domain, to establish that $p(\alpha, \pi_{n_{k_j}}, \gamma_{n_{k_j}}) \to p(\alpha, \pi^*, \gamma^*)$ Since every subsequence of $p(\alpha, \pi_n, \gamma_n)$ has a sub-subsequence which converges to $p(\alpha, \pi^*, \gamma^*)$, we establish that $p(\alpha, \pi_n, \gamma_n) \to p(\alpha, \pi^*, \gamma^*)$.

Now, observe that for all $\alpha \leq \alpha^*$, we have $d(p(\alpha,\pi,\gamma),\pi) \leq \sqrt{2\ln(2)R(p(\alpha,\pi,\gamma) \parallel \pi)} \leq \epsilon$, where the first inequality follows by Pinsker's inequality and monotonicity in α, as described above (We have used the version of Pinsker's inequality from Cover and Thomas (2012), Lemma 11.6.1). □

5. Conclusion and related literature

The literature on scoring rules is vast; the first characterization of proper scoring rules in terms of subdifferentials of convex functions was provided by McCarthy (1956); see also Savage (1971) and Fang et al. (2010). Gneiting and Raftery (2007) provides a survey of the literature up to 2007.

Theorem 1 has precedence in the literature. Probably the first such result is due to Winkler and Murphy (1970), who study the quadratic scoring rule for general expected utility preferences. Kadane and Winkler (1988) calculate an explicit formula that expected utility maximizers would use when they have nontrivial risk attitudes and face a quadratic scoring rule. Grünwald and Dawid (2004) uncovers the special case of this result in the context of risk-neutral multiple priors (Chambers (2008) later derives the same result). Offerman et al. (2009) derives a related result in a general decision-theoretic model (not necessarily quasiconcave preferences) essentially with a binary state space and the quadratic scoring rule, establishing that if a subject announces a probabilistic belief of an event which is not equal to 1/2, then we can say something about that optimal announcement.[16] Our result applies to any continuous strictly incentive compatible scoring rule whatsoever; all of the classical ones as well as lesser known ones (for example Winkler (1994)). Furthermore, our result also applies to decision makers for which a meaningful concept of probability cannot even be defined.

Because of the inherent problems associated with assuming risk neutrality, a literature has arisen attempting to circumvent the issue. A classic mechanism is due to Grether (1981) and Karni (2009).[17] Lambert (2009) presents a full characterization of mechanisms of this type; see also Hossain and Okui (2013), Qu (2012), and Schlag and van der Weele (2013). Intuitively, these mechanisms make use of objective randomizing devices, which are external to the relevant state space.[18] Our main result does not speak to these models as our model has no language for discussing objective probability. However, it is likely that similar duality results can be established under suitable regularity conditions. We emphasize that the mechanisms in this literature are not designed to resolve the issue of ambiguity-averse

[16]The number 1/2 is an artifact of the symmetry of the quadratic scoring rule.

[17]This particular mechanism should be viewed as a special case of a "random decision mechanism," which is a class recently investigated by Azrieli et al. (2012) and usually attributed to Savage (1972) and Allais (1953). Probably its most popular implementation is due to Becker et al. (1964).

[18]Experimental economists usually attribute this idea to Roth and Malouf (1979).

preferences or individuals without probabilistic beliefs, but rather assume agents behave in a "probabilistically sophisticated" way (see (Machina and Schmeidler, 1992)), and so the problem of characterizing optimal choices for arbitrary agents facing such mechanisms remains open.

Bickel (2007) establishes properties of individuals with CARA-style expected utility preferences: for example, he shows that one can add a constant payoff to each action in the profile of a scoring rule without changing behavior. He attributes this to what he calls the "delta" property; something economists would call translation invariance or quasilinearity, and is characterized by the variational model. This result follows from a straightforward application of Theorem 1. Jose et al. (2008) discusses a duality related to that of Grünwald and Dawid (2004), but with a different aim. There, they want to understand how a risk-averse expected utility maximizer will "bet" against a given set of priors.

There is a recent elegant literature, pioneered by (Maccheroni et al., 2006; Cerreia-Vioglio et al., 2011b), which exploits the duality between indirect and direct utility in order to study properties of uncertainty aversion. The framework is different; namely, they work with the richer structure of Anscombe and Aumann (1963) acts. The extra mathematical structure allows for the separate study of uncertainty and risk. The duality investigated in Cerreia-Vioglio et al. (2011b) and discussed in detail in Cerreia-Vioglio et al. (2011a) is exactly the one we use here.[19] Though this work is concerned with uncertainty aversion, many of the mathematical results and characterizations appearing there apply verbatim here, and we have heavily borrowed from this work. Further, because of risk attitudes, it is often advocated that individuals be paid "in probabilities" of a good outcome, instead of monetary terms. This is based on the idea that, over purely risky prospects, individuals will likely conform to expected utility behavior (indeed this is the framework upon which the analysis of Anscombe and Aumann (1963) is built). The practical implementation of this idea in experiments is due to Roth and Malouf (1979). In such a framework, the framework of Cerreia-Vioglio et al. (2011b) is the appropriate one for studying elicitation questions. The natural counterparts of the following results hold as stated. We focus here on monetary payoffs for simplicity.

Similarly, one could restrict the domain of scoring rules to take only nonnegative values, and describe "homogeneous" utility indices U; these would correspond to G functions for which $G(p, w) = wG(p, 1)$. It is likely that similar representations for CRRA preferences could derived.

[19]Formally, they apply the duality to von Neumann-Morgenstern utils, and work with a much more general framework.

Comparative statics on the G function in terms of the "more risk averse" relation of Yaari (1969) can be provided, and exist in Cerreia-Vioglio et al. (2011b) (in the form of comparative statics on uncertainty aversion).

This paper has added to the previous papers by considering a highly general specification of utility. Most previous works assume either that preference is expected utility, or consider the special case of "risk-neutral" multiple priors (Grünwald and Dawid (2004)). In contrast, our preferences never need to reference any concept of likelihood or state-contingent utility payoffs whatsoever. We leave the question of infinite state spaces, and elicitation mechanisms utilizing objective randomization devices to future research.

References

Aliprantis, C. D. and K. C. Border (1999): *Infinite dimensional analysis, second edition*, Springer.

Allais, M. (1953): "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine," *Econometrica*, 21, 503–546.

Anscombe, F. J. and R. J. Aumann (1963): "A definition of subjective probability," *Annals of mathematical statistics*, 34, 199–205.

Azrieli, Y., C. P. Chambers, and P. J. Healy (2012): "Incentives in experiments: a theoretical analysis," Working paper.

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

Berge, C. (1997): *Topological Spaces*, Dover.

Bickel, J. E. (2007): "Some comparisons among quadratic, spherical, and logarithmic scoring rules," *Decision Analysis*, 4, 49–65.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011a): "Complete monotone quasiconcave duality," *Mathematics of Operations Research*, 36, 321–339.

——— (2011b): "Uncertainty averse preferences," *Journal of Economic Theory*, 146, 1275–1330.

Chambers, C. P. (2008): "Proper scoring rules for general decision models," *Games and Economic Behavior*, 63, 32–40.

Chambers, C. P., F. Echenique, and K. Saito (2016): "Testing theories of financial decision making," *Proceedings of the National Academy of Sciences*, 113, 4003–4008.

Cornes, R. (1992): *Duality and Modern Economics*, Cambridge University Press.

Cover, T. M. and J. A. Thomas (2012): *Elements of information theory*, John Wiley & Sons.

de Finetti, B. (1949): "Sulle stratificazioni convesse," *Annali di Matematica Pura e Applicata*, 30, 173–183.

Ellsberg, D. (1961): "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Economics*, 643–669.

Fang, F., M. B. Stinchcombe, and A. B. Whinston (2010): "Proper scoring rules with arbitrary value functions," *Journal of Mathematical Economics*, 46, 1200–1210.

Gilboa, I. and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 18, 141–153.

Gneiting, T. and A. E. Raftery (2007): "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, 102, 359–378.

Good, I. J. (1952): "Rational decisions," *Journal of the Royal Statistical Society. Series B (Methodological)*, 107–114.

Gorman, W. M. (1961): "On a class of preference fields," *Metroeconomica*, 13, 53–56.

Grether, D. (1981): "Financial incentive effects and individual decision making," Working paper.

Grünwald, P. D. and A. P. Dawid (2004): "Game theory, maximum entropy, minimum discrepancy and robust Bayesian decision theory," *Annals of Statistics*, 32, 1367–1433.

Hey, J. D. and N. Pace (2014): "The explanatory and predictive power of non two-stage-probability theories of decision making under ambiguity," *Journal of Risk and Uncertainty*, 49, 1–29.

Hossain, T. and R. Okui (2013): "The binarized scoring rule," *The Review of Economic Studies*, 80, 984–1001.

Huber, P. J. (1981): *Robust statistics*, Springer.

Jose, V. R. R., R. F. Nau, and R. L. Winkler (2008): "Scoring rules, generalized entropy, and utility maximization," *Operations research*, 56, 1146–1157.

Kadane, J. B. and R. L. Winkler (1988): "Separating probability elicitation from utilities," *Journal of the American Statistical Association*, 83, 357–363.

Karni, E. (2009): "A mechanism for eliciting probabilities," *Econometrica*, 77, 603–606.

Konüs, A. A. (1939): "The problem of the true index of the cost of living," *Econometrica*, 7, 10–29.

Lambert, N. (2009): "Probability elicitation for agents with arbitrary risk preferences," Working paper.

Lau, L. J. (1969): "Duality and the structure of utility functions," *Journal of Economic Theory*, 1, 374–396.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity aversion, robustness, and the variational representation of preferences," *Econometrica*, 74, 1447–1498.

Machina, M. J. and D. Schmeidler (1992): "A more robust definition of subjective probability," *Econometrica*, 60, 745–780.

McCarthy, J. (1956): "Measures of the value of information," *Proceedings of the National Academy of Sciences*, 42, 654–655.

Nyarko, Y. and A. Schotter (2002): "An experimental study of belief learning using elicited beliefs," *Econometrica*, 70, 971–1005.

Offerman, T., J. Sonnemans, G. Van de Kuilen, and P. P. Wakker (2009): "A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes," *The Review of Economic Studies*, 76, 1461–1489.

Qu, X. (2012): "A mechanism for eliciting a probability distribution," *Economics Letters*, 115, 399–400.

Roth, A. E. and M. W. Malouf (1979): "Game-theoretic models and the role of information in bargaining." *Psychological review*, 86, 574.

Roy, R. (1947): "La distribution du revenu entre les divers biens," *Econometrica*, 15, 205–225.

Savage, L. J. (1971): "Elicitation of personal probabilities and expectations," *Journal of the American Statistical Association*, 66, 783–801.

——— (1972): *The Foundations of Statistics*, Dover.

Schlag, K. H. and J. J. van der Weele (2013): "Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality," *Theoretical Economics Letters*, 3, 38–42.

Shepherd, R. W. (1970): *Theory of Cost and Production Functions*, Princeton University Press.

Strzalecki, T. (2011): "Axiomatic foundations of multiplier preferences," *Econometrica*, 79, 47–73.

Ville, J. and P. Newman (1951-1952): "The existence-conditions of a total utility function," *Review of Economic Studies*, 19, 123–128.

Weymark, J. A. (1980): "Duality results in demand theory," *European Economic Review*, 14, 377–395.

Winkler, R. L. (1994): "Evaluating probabilities: Asymmetric scoring rules," *Management Science*, 40, 1395–1405.

Winkler, R. L. and A. H. Murphy (1970): "Nonlinear utility and the probability score," *Journal of Applied Meteorology*, 9, 143–148.

Yaari, M. E. (1969): "Some remarks on measures of risk aversion and on their uses," *Journal of Economic theory*, 1, 315–329.