Can Anticipatory Feelings Explain Anomalous Choices of Information Sources?*

Kfir Eliaz† and Ran Spiegler‡

August 30, 2004

Abstract

The well-being of agents is often directly affected by their beliefs, in the form of anticipatory feelings such as anxiety and hopefulness. Economists have tried to model this effect by introducing beliefs as arguments in decision makers' vNM utility function. One might expect that such a model would be capable of explaining anomalous attitudes to information that we observe in reality. We show that the model has several shortcomings in this regard, as long as Bayesian updating is retained.

*We thank Andrew Caplin, Eddie Dekel, Barton Lipman, Ronny Razin and Muhamet Yildiz for helpful discussions. Financial support from the US-Israel Binational Science Foundation, Grant No. 2002298, is gratefully acknowledged.

†Dept. of Economics, NYU. 269 Mercer St., New York, NY 10003, E-mail: kfir.eliaz@nyu.edu, URL: http://homepages.nyu.edu/~ke7/

‡School of Economics, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: rani@post.tau.ac.il. URL: http://www.tau.ac.il/~rani.

1 Introduction

In standard expected utility theory, decision makers have preferences over some set of consequences, where a consequence consists of a state of nature and an action taken by the decision maker (DM henceforth). The DM's beliefs regarding the state of nature clearly affect his decisions, but they are not part of the definition of a consequence - that is, they do not enter as arguments into his vNM utility function. In other words, the DM is "emotionally neutral" towards his beliefs. In reality, however, our well-being often seems to be *directly* affected by our beliefs. For instance, imagine that you expect the results of a medical test. Even if this information cannot lead to any change in your actions, it causes you to update your beliefs. The change in beliefs may affect your well-being by increasing or reducing your level of anxiety.

For this reason, there has been a growing sentiment among economists that the standard model of choice under uncertainty should be enriched by adding beliefs to the description of consequences, in order to capture anticipatory feelings such as anxiety or hopefulness. Specifically, economists have constructed expected-utility choice models, in which the DM's belief (in addition to his action and the state of nature) is an argument in his vNM utility function. The following quote is taken from Akerlof and Dickens (1982), probably the first paper to construct such an extended choice model:

> "...persons not only have preferences over states of the world, but also over their beliefs about the state of the world...persons have some control over their beliefs; not only are people able to exercise some choice about belief given available information, they can also manipulate their own beliefs by selecting sources of information likely to confirm 'desired' beliefs."

As Akerlof and Dickens point out, people influence their beliefs in two ways: direct choice of beliefs through 'self persuasion', and indirect choice through selection of signals. In this paper, we focus on the latter mechanism. We are concerned with the effect of agents' "preferences over their beliefs about the state of the world" on the way in which they "select sources of information". Specifically, we study the choices of signals by a DM whose behavior is governed by the above-mentioned, extended expected-utility model. At the same time, we exclude direct choice of beliefs. Our DM updates his beliefs according to Bayes' rule given available information. We pose the following question: Can an enriched expected utility model, in which a Bayesian DM's belief is an argument in his vNM utility function, explain anomalous choices of information sources that the usual expected utility model cannot explain?

To study this problem, we construct in Section 3 a model with a finite set Ω of states of nature. The DM's prior belief is a probability distribution $\mathbf{p}$ over the state space. He obtains information about the state by observing a signal $\mathbf{Q}$, which is modeled as a finite stochastic matrix (the element q_{ij} in the matrix is the probability of realization j of the signal, given that the true state is i). For each prior $\mathbf{p}$, the DM is assumed to have a complete preference ordering $\succsim_{\mathbf{p}}$ over the set of all signals. The preference profile $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ constitutes the DM's "attitude to information".

The following are examples of situations captured by the model:

- A patient prefers more accurate medical tests when he is relatively certain of being healthy, yet he avoids such tests when he is relatively certain of being ill.
- A manager consults his advisors when he is relatively sure of what their advice will be, yet avoids their advice when he is relatively unsure.
- A news reader prefers a newspaper with a left-wing bias (i.e., one that reports every piece of news that supports a leftist view, but not all the news that support a rightist view) to an objective newspaper that reports all the relevant news. He also prefers the objective newspaper to a newspaper with a right-wing bias.

These attitudes to information are anomalous, in the sense that a standard expected utility maximizer (whose beliefs do not enter into his vNM utility function) would never display them. Our intuition suggests that in these examples, the DM approaches information not as a means of making better decisions, but rather as a means of cultivating desired beliefs and suppressing undesired beliefs. The patient in the first example is afraid of learning that the state is adverse. The manager in the second example and the news reader in the third example fear the prospect of having to change their opinion.

This intuition has some systematic empirical support. Several studies in medicine and psychology have confirmed that individuals who believe that an unpleasant event is likely to occur may reject information regarding this event, even when this information may help them take actions that reduce the unpleasantness of the anticipated event. For example, Lerman *et al.* (1998) demonstrated that 46% of subjects whose blood was tested for genetic mutations refused to receive the test results despite the fact that the test results indicated whether or not these subjects were susceptible to breast cancer later in life.

How can we accommodate this intuition in the extended expected-utility model? Our DM updates his beliefs using *Bayes' rule*. Therefore, a prior $\mathbf{p}$ and a signal $\mathbf{Q}$ induce a probability distribution over the DM's posterior belief, via Bayes' formula. An attitude to information is consistent with expected utility maximization over beliefs if there is a continuous vNM utility function over posteriors $u : \Delta(\Omega) \to \mathbb{R}$, such that the DM evaluates signals by the expected utility from the induced lottery over posteriors. For example, let $\Omega = \{\omega_1, \omega_2\}$. At the prior (p_1, p_2), the DM prefers a fully informative binary signal ($q_{11} = q_{22} = 1$, $q_{12} = q_{21} = 0$) to a fully uninformative binary signal ($q_{ij} = \frac{1}{2}$ for every $i, j = 1, 2$) if $p_1 \cdot u(1, 0) + p_2 \cdot u(0, 1) > u(p_1, p_2)$. Note that u is defined only over posterior beliefs, ignoring the actions that the DM may take after observing the signal. We show that this is *without* loss of generality.

Despite its apparent intuitive appeal, the model of expected utility from posterior beliefs (with Bayesian updating) turns out to have several shortcomings as an explanation of anomalous attitudes to information. We make this argument via two sets of results.

Failure of ordinal representation. Recall Dickens and Akerlof's intuitive claim that people tend to select sources of information that are likely to confirm "desired" beliefs. It seems reasonable to suppose that people who fundamentally disagree about the relative desirability of different states would also differ in their preferences over information sources. Nevertheless, in Section 4 we show that there are non-monotonic transformations of the vNM function u, which represent the same preference profile. Two people with opposite views regarding the relative desirability of states may share the same attitude to information. Thus, the distinction between "desired" and "undesired" beliefs is not relevant to the DM's choice of information sources.

Failure to explain anomalies. We examine whether anomalous attitudes to information, such as those described in the above examples, can be explained by expected utility maximization over posterior beliefs. In Section 5 we show that despite the intuition that these types of behavior are a consequence of anticipatory feelings, they are *inconsistent* with the model.

The results are simple from a technical point of view. We believe that they are important because they demonstrate the limitations of the extended expected-utility framework as a model of anticipatory feelings. We identify three ways to proceed from these negative results, while retaining the idea of direct utility from posterior beliefs: *(i)* abandoning expected utility in favor of non-expected utility functionals; *(ii)* entering the DM's prior belief as an additional argument in his vNM utility function; *(iii)* abandoning Bayesian updating. In Section 6 we provide several arguments for the third option.

2 Related Literature

The earliest work that we know of, in which an agent's belief enters into his utility function, is Akerlof and Dickens (1982). In the works that followed their example, one can distinguish between models in which the agent's belief is a *choice variable* and models in which it is a *parameter* in his utility function. While Akerlof and Dickens (1982) falls into the former category, our paper falls into the latter. In all the models that we are aware of, DMs have *expected-utility* preferences over a space of outcomes, whose description consists of the DM's belief, his action and the true state of nature.

The former strand of the literature includes Brunnermeier and Parker (2002) and Yariv (2002), who construct multi-period choice models, in which agents directly choose their beliefs. In both models, a DM's periodic utility is a sum of two components: the usual, "physical" utility from the DM's action and the true state of nature, and a non-standard "anticipation utility" or "belief utility", which is a function of the DM's present and past beliefs. In both models, choosing to hold an incorrect belief may be advantageous because it is a "desirable" belief. However, the DM is constrained to choose an action which is optimal w.r.t his belief, such that holding an incorrect belief is costly in terms his physical utility. The two papers use very different specifications of the "non-standard" utility component and apply their models to different realms of choice behavior. In particular, Brunnermeier and Parker (2002) deal with inter-temporal consumption behavior and do not study information acquisition, whereas in Yariv (2002), the DM receives a signal every period and he can choose its accuracy level. In another paper, Eyster (2002) models DMs who choose beliefs that help them to rationalize their *past* choices.

The most important difference between these papers and the present paper is that in our model, the DM *cannot* choose beliefs directly and can only influence them indirectly through his choice of signals. Our objective is to examine the implications of belief-dependent vNM utility on attitudes to information, whereas the above-cited papers focus on the trade-off between the physical costs and emotional benefits of choosing to hold an incorrect belief.

Our paper is more directly related to the second strand in the literature, in which a DM's belief is not a choice variable, but a parameter whose value is determined in equilibrium. Caplin and Leahy (2001) develop a two-period model, in which the DM's first period utility is a function of the first period outcome and his posterior first-period probability distribution over second-period outcomes. Kőszegi (2004b) enriches this model by studying the T-period decision problem in which the DM's expectations over outcomes in each of the T periods enter as parameters in the utility function of each period t.

These authors apply their model to study an information transmission game between a physician and his patient. Caplin and Leahy (2004) study the problem faced by a physician who observes the health state of a patient (which can be either "good" or "bad") and needs to decide whether or not to credibly reveal that state. The physician wishes to maximize the patient's expected utility, knowing that the patient's posterior belief about his health state enters his utility function. There are two types of patients: one who prefers his posterior beliefs to be as close as possible to his prior, and another with the opposite preference. Assuming the physician only knows the distribution of patients' types, the authors analyze the equilibrium of the extensive game (using a solution concept developed by Geanakoplos, Pearce and Sttachetti (1989)) in which a patient first announces his type, and then the physician decides whether or not to reveal the patient's health state. Kőszegi (2004a) extends this model by allowing physicians to send non-verifiable messages, as well as by allowing patients to visit more than one physician.

Both Caplin and Leahy (2004) and Kőszegi (2004a) use a model of expected utility over beliefs to explain anomalous attitudes to information, while retaining Bayesian updating as part of the solution concept they apply. They show that the model can rationalize anomalous choices of information sources *at a given prior*. The difficulties which we identify in Section 5 arise because we examine the *entire profile* of preferences over signals - i.e., we analyze the DM's attitude to information at *all possible prior beliefs*. As the examples given in the Introduction illustrate, the reason we examine the DM's entire preference profile is that many anomalies of interest involve attitudes to information that vary with the DM's prior.

Our paper is also related to the literature that studies attitudes to temporal resolution of uncertainty, emanating from Kreps and Porteus (1978). Although our model is static, it may be possible to embed it in the Kreps-Porteus formalism. A DM who systematically prefers more (less) accurate signals in a static model, can be viewed as a DM who systematically prefers early (late) resolution of uncertainty in a multi-period model. We have no disagreement with this "embedding" argument. To the extent that preferences over beliefs can be re-interpreted as preferences over the timing of uncertainty resolution, our results apply to the Kreps-Porteus formalism. Note, however, that the original Kreps-Porteus formalism excludes prior-dependent preferences.[1]

Decision theorists have studied anomalous attitudes to information and sought to trace them to other sources than anticipatory feelings. As demonstrated by Wakker (1988), Safra and Sulganik (1995) and Grant et al. (1998, 2000), DMs' violation of the independence axiom may cause them to dislike information. In these papers, preferences are defined over *lotteries* and utility is defined over *physical* consequences; attitudes to information are deduced from these preferences. In contrast, the present paper *redefines the consequence space* to be the set of posterior beliefs, while adhering to an expected-utility functional; attitudes to information are not deduced, but serve as choice primitives.

Finally, anomalous attitudes to information have also been explained by dynamically inconsistent preferences. As shown by Carrillo and Mariotti (2000) and Bénabou and Tirole (2002), a decision maker who lacks self control may decide to avoid information today so as to discipline his future self.

3 The model

There is a finite set of states $\Omega = \{\omega_1, \ldots, \omega_n\}$. Let $\Delta(\Omega)$ denote the set of all probability distributions on the elements of Ω. A DM has a vector of prior probabilities on Ω given by $\mathbf{p} \in \Delta(\Omega)$ where p_i is the prior probability assigned to state ω_i. A *signal* is a random variable, which can take $m \geq n$ distinct values, $s_1, \ldots, s_m$. A signal is characterized by an $n \times m$ stochastic matrix of conditional probabilities denote by $\mathbf{Q} = (q_{ij})_{i=1,\ldots,n\ ;\ j=1,\ldots,m}$, where $q_{ij} \in [0, 1]$ is the probability of observing the realization s_j conditional on the state being ω_i. Of course, for every $i = 1, ..., n$, $\Sigma_{j=1,\ldots,m}\, q_{ij} = 1$. With slight abuse of terminology, we refer to a matrix $\mathbf{Q}$ as a signal and let $\mathcal{Q}$ denote the set of all signals, i.e., the set of all $n \times m$ stochastic matrices. Note that although we assume that $m \geq n$, it is always possible to replicate signals with $k < m$ realizations, by setting $m - k + 1$ rows in $\mathbf{Q}$ to be identical.

[1]In a recent paper, Caplin and Eliaz (2003) study a mechanism design problem where agents have preferences for late resolution of uncertainty. In their model, the payoff function of each agent depends on the posterior probability that the agent carries an infectious disease.

The DM's choice of signals at every prior $\mathbf{p}$ is rational. Therefore, the DM's overall attitude to information can be summarized by a profile of preference relations $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ over the set $\mathcal{Q}$. We are agnostic as to whether the DM has to take some action after observing a signal's realization. We offer a justification for this modeling choice below.

To illustrate the concept of a preference profile over signals, consider the case of $n = 2$. For ease of exposition, we write $\mathbf{Q} \geq \mathbf{R}$ for $\mathbf{Q}, \mathbf{R} \in \mathcal{Q}$ if $q_{11} \geq r_{11}$ and $q_{22} \geq r_{22}$ (and we write $\mathbf{Q} > \mathbf{R}$ if at least one of these inequalities is strict).

Example 1. An *information-seeking* DM prefers a signal $\mathbf{Q}$ to a signal $\mathbf{R}$ whenever $\mathbf{Q}$ is more accurate than $\mathbf{R}$. In particular, for every prior $\mathbf{p} \in \Delta(\Omega)$, $\mathbf{Q} > \mathbf{R}$ implies $\mathbf{Q} \succ_{\mathbf{p}} \mathbf{R}$. Similarly, an *information-averse* DM satisfies the opposite property: for every prior $\mathbf{p} \in (0, 1)^2$, $\mathbf{Q} > \mathbf{R}$ implies $\mathbf{R} \succ_{\mathbf{p}} \mathbf{Q}$.

Example 2. Consider a patient, for whom ω_1 and ω_2 stand for "being healthy" and "being ill", respectively. The patient seeks information when he is relatively sure that he healthy, yet he avoids information when he is relatively sure that he is ill. Formally, when p_1 is close to 1, $\mathbf{Q} > \mathbf{R}$ implies $\mathbf{Q} \succ_{\mathbf{p}} \mathbf{R}$; and when p_1 is close to 0, $\mathbf{Q} > \mathbf{R}$ implies $\mathbf{R} \succ_{\mathbf{p}} \mathbf{Q}$.

Example 3. A manager seeks the opinion of his employees only when he is sufficiently certain that the new information will not cause him to change his view as to which state is more probable. Formally, when $\max\{p_1, p_2\}$ is close to 1, $\mathbf{Q} > \mathbf{R}$ implies $\mathbf{Q} \succ_{\mathbf{p}} \mathbf{R}$; and when $\max\{p_1, p_2\}$ is close to $\frac{1}{2}$, $\mathbf{Q} > \mathbf{R}$ implies $\mathbf{R} \succ_{\mathbf{p}} \mathbf{Q}$.

Example 4. A student is writing a Ph.D. thesis. There are two possible states of the world: either the thesis is of high quality (state ω_1) or it is of low quality (state ω_2). He believes that the former holds with probability $p_1 > \frac{1}{2}$. The student can choose among three thesis advisors. An "objective" advisor always tells his students the truth. A "tough" advisor tells his students that their thesis is of low quality whenever it is in fact of low quality. However, even if a thesis is of high quality, he claims that it is of low quality with probability $1 - \alpha$. Finally, a "lenient" advisor tells his students that their thesis is of high quality whenever it is in fact of high quality. However, even if a thesis is of low quality, he claims that it is of high quality with probability $1 - \alpha$. Our student ranks the objective advisor below the tough advisor and above the lenient advisor, as long as α is sufficiently close to one.

Example 5. After observing the realization of a signal, the DM has to take some action. Each state is associated with some optimal action; failure to take it entails a loss. The DM chooses his action so as to maximize expected utility from physical outcomes. We refer to this DM as an EUM agent. Suppose that $p_1 \neq \frac{1}{2}$. When the probability pairs (q_{11}, q_{22}) and (r_{11}, r_{22}) are sufficiently close to $(\frac{1}{2}, \frac{1}{2})$, the DM's optimal choice of action is independent of the signals' realizations; hence, the DM is indifferent between $\mathbf{Q}$ and $\mathbf{R}$. As the signal becomes more informative, it begins to affect the DM's choice of action and his expected loss decreases. For example, suppose that there are two actions, a_1 and a_2. For every $i = 1, 2$, the utility values from taking actions a_i and a_j $(i \neq j)$ at state ω_i are 0 and -1, respectively. Thus, the DM faces a decision problem with a symmetric loss function. It follows that whenever $p_1 > \frac{1}{2}$ and $\mathbf{Q} \geq \mathbf{R}$ we have that $\mathbf{Q} \backsim_{\mathbf{p}} \mathbf{R}$ if $\frac{p_1}{1-p_1} \geq \frac{q_{22}}{1-q_{11}}$ and $\mathbf{Q} \succ_{\mathbf{p}} \mathbf{R}$ if $\frac{p_1}{1-p_1} < \frac{q_{22}}{1-q_{11}}$. Finally, $\mathbf{Q} \succ_{(\frac{1}{2}, \frac{1}{2})} \mathbf{R}$ whenever $\mathbf{Q} \geq \mathbf{R}$.

In examples 1 and 5, the DM's preference w.r.t the informativeness of signals is never reversed. Examples 2-4 display more complex attitudes to information. In examples 2 and 3, the DM's preference w.r.t the informational content of signals varies with his prior. In example 4, he finds perfect information superior to a certain class of biased signals yet inferior to another. These examples are stylized descriptions of attitudes to information that are taken from everyday experience. Introspection suggests that these attitudes may be a result of DMs' attempt to attain desired beliefs and avoid undesired beliefs. We will analyze these examples in detail in Section 5, using the model set forth in this section.

When our DM chooses a signal, he takes into account that he updates his beliefs according to Bayes' rule upon observing the signal's realization (see Section 6 for a discussion of this assumption). Thus, for every realization s_i of the random variable, a signal $\mathbf{Q}$ and a prior $\mathbf{p}$ generate a distribution of posterior probabilities on Ω. Given $\mathbf{p}$ and $\mathbf{Q}$, the posterior probability that the state is ω_i, conditional on a realization

s_j for which $\Pr(s_j|\mathbf{p},\mathbf{Q}) > 0$, is given by

$$z_i(\mathbf{p},\mathbf{Q}|s_j) = \frac{p_i q_{ij}}{\sum_{i=1}^{n} p_i q_{ij}}$$

Hence, when a DM with a prior $\mathbf{p}$ chooses a signal $\mathbf{Q}$, he effectively chooses a lottery over his posterior beliefs. The j-th element in the support of this lottery, $j = 1, ..., m$, is the posterior belief

$$\mathbf{z}(\mathbf{p},\mathbf{Q}|s_j) = (z_1(\mathbf{p},\mathbf{Q}|s_j), \ldots, z_n(\mathbf{p},\mathbf{Q}|s_j))$$

and it is drawn with probability $\sum_{i=1}^{n} p_i q_{ij}$.

We may therefore interpret choices of signals at a given prior as a revealed preference over lotteries on distributions of posterior beliefs. This raises the following question: when can we rationalize a DM's pattern of choices over signals by expected-utility preferences over lotteries on posterior beliefs?

Let $u : \Delta(\Omega) \to \mathbb{R}$ be a continuous vNM utility function over posterior beliefs. For every distribution of prior beliefs $\mathbf{p} \in \Delta(\Omega)$ and for every signal $\mathbf{Q} \in \mathcal{Q}$, define

$$U(\mathbf{p},\mathbf{Q}) \equiv \sum_{\mathbf{z}\in\Delta(\Omega)} \Pr(\mathbf{z}|\mathbf{p},\mathbf{Q}) \cdot u(\mathbf{z}) = \sum_{j} \Pr(s_j|\mathbf{p},\mathbf{Q}) \cdot u[\mathbf{z}(\mathbf{p},\mathbf{Q}|s_j)]$$

where $\Pr(s_j|\mathbf{p},\mathbf{Q}) = \sum_{i=1}^{n} p_i q_{ij}$. We interpret $U(\mathbf{p},\mathbf{Q})$ as a representation of the DM's preferences over lotteries on posterior beliefs.

Definition 1 *The function u **rationalizes** the preference profile $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ if for every $\mathbf{p} \in \Delta(\Omega)$ and every pair of signals $\mathbf{Q},\mathbf{R} \in \mathcal{Q}$,*

$$\mathbf{Q} \succsim_{\mathbf{p}} \mathbf{R} \Longleftrightarrow U(\mathbf{p},\mathbf{Q}) \geq U(\mathbf{p},\mathbf{R})$$

A preference profile $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ is said to be *rationalizable*, if there exists a continuous vNM utility function that rationalizes this profile.[2]

The natural benchmark for the model is the EUM agent (Example 5): the DM who maximizes expected utility w.r.t physical outcomes, as if he has no anticipatory feelings. Let us verify that an EUM agent can be rationalized by expected utility over posterior beliefs. Let A be the set of actions available to the DM upon observing a signal's realization. Let $v(a, \omega_i)$ denote the DM's utility from taking action a in state ω_i. For every $\mathbf{z} \in \Delta(\Omega)$, define

$$u(\mathbf{z}) = \max_{a\in A} \sum_{i=1}^{n} z_i \cdot v(a, \omega_i)$$

Thus, for every distribution of prior beliefs $\mathbf{p}$ and for every signal $\mathbf{Q}$, $U(\mathbf{p}, \mathbf{Q})$ is equal to the DM's *indirect* expected utility from physical outcomes when he uses the signal $\mathbf{Q}$. By definition, an EUM agent prefers signal $\mathbf{Q}$ to $\mathbf{R}$ if and only if $\mathbf{Q}$ yields a higher indirect expected utility than $\mathbf{R}$, i.e., $\mathbf{Q} \succsim_{\mathbf{p}} \mathbf{R}$ if and only if $U(\mathbf{p}, \mathbf{Q}) \geq U(\mathbf{p}, \mathbf{R})$. Note that u is continuous. It follows that u rationalizes the EUM agent.

Our model is a parsimonious departure from the standard expected utility framework, in the sense that it involves a minimal extension beyond the usual case which excludes anticipatory feelings. Whereas the EUM agent is rationalized by a utility function having a *particular* structure, our model imposes no structure on u except for continuity.

Why are beliefs the only argument in the utility function? In our model, agents only choose signals. We are agnostic as to whether or not they take some action after observing the signal. The consequence space consists of posterior beliefs only. Does this entail any loss of generality? We now show that it does not.

Recall that in the EUM example, the function u that rationalized the DM's attitude towards information was precisely the *indirect* expected utility function over physical outcomes. This is a general feature: the utility function over beliefs u

[2] We allow the prior to be degenerate: $p_i = 1$ for some state ω_i. In this case, Bayesian updating requires that $U(\mathbf{p}, \mathbf{Q}) = u(\mathbf{p})$ for every signal $\mathbf{Q}$.

can be viewed as the indirect utility induced by expected utility maximization over an enriched consequence space, which includes not only beliefs but also states and actions.

Consider an extended model in which the DM has to choose an action $a \in A$ after observing the realization of a signal. A consequence in this model is a triplet $(\omega_i, a, \mathbf{z})$ consisting of a state of nature ω_i, an action a and a distribution of posterior beliefs $\mathbf{z}$. Assume the DM has a preference relation over lotteries on this extended consequence such that there exists a pair of continuous functions $f : \Omega \times A \times \Delta(\Omega) \to \mathbb{R}$ with the following property: given a distribution of posterior beliefs $\mathbf{z}$, the DM weakly prefers action a to a' if, and only if

$$\sum_{i=1}^{n} z_i f(\omega_1, a, \mathbf{z}) \geq \sum_{i=1}^{n} z_i f(\omega_1, a', \mathbf{z})$$

Now, define

$$u(\mathbf{z}) = \max_{a \in A} \sum_{i=1}^{n} z_i f(\omega_1, a, \mathbf{z})$$

It is easy to see that u is continuous. A DM whose objective is to maximize the expected value of f (where the expectation is taken w.r.t $\mathbf{z}$) will choose signals so as to maximize the expectation of u. Thus, our assumption that only posterior beliefs enter the DM's vNM utility function simplifies the analysis and entails no loss of generality.

Analogy to risk attitudes. To better understand our model, it may be useful to draw the following analogy to a model of decision making under risk.[3] Consider the case of $n = 2$ (two states of nature). Think of the DM's prior belief as his "*initial wealth*" $w \in W$. Similarly, think of the DM's posterior belief as his "*final wealth*". Hence, a signal can be thought of as a fair incremental lottery, i.e., a lottery with a

[3]For a careful discussion of the analogy between attitudes to information and attitudes to risk (though not in the context of utility from beliefs), see Grant *et al.* (1998).

mean of zero (e.g., a lottery in which a gain of 1,000 and a loss of 1,000 have equal probabilities). Thus, given w, the DM can be thought of as having a preference relation $\succeq_w$ over some set of fair incremental lotteries. The questions we address in this paper are analogous to the following questions. Is the preference profile $(\succeq_w)_{w \in W}$ consistent with maximization of expected utility over the DM's *final* wealth? If so, what is the relation between the DM's attitudes to spread (exhibited by $(\succeq_w)_{w \in W}$ at different initial wealth levels w) and the risk attitudes exhibited by his vNM utility function over final wealth?

Questions similar to these were studied by decision theorists (Rothschild and Stiglitz (1970), Landsberger and Meilijson (1990), among others), whose goal was primarily to redefine risk aversion as a property of attitudes toward spread, rather than a property of preferences over (final) wealth. As far as we know, none of these studies examined the case of attitudes to spread that vary with initial wealth. Our main results will concern attitudes to information that vary with the DM's prior, and therefore they will not have a counterpart in this literature.

4 Failure of ordinal representation

Recall Akerlof and Dickens' claim that "people...manipulate their own beliefs by selecting sources of information likely to confirm 'desired' beliefs." Bénabou and Tirole (2002, p. 906) write in a similar vein: "people just *like* to think of themselves as good, able, generous, attractive, and conversely find it painful to contemplate their failures and shortcomings."

Both statements reflect the idea that a belief that the state is "good" is more desirable than the belief that the state is "bad". Intuitively, agents who differ in what constitutes a desirable belief for them will also differ in their attitudes toward information. For example, let $n = 2$, and suppose that ω_1 and ω_2 are states that support left-wing and right-wing viewpoints, respectively. Consider two news reader, whose well-being is equally affected by their beliefs, yet for one news reader (a "leftist") ω_1 is a more favorable state than ω_2, while for the other news reader (a "rightist") ω_2 is more favorable than ω_1. We expect them to read different newspapers. In particular, we suspect that they will choose differently between two partially informative newspapers, one having a right-wing bias and the other having a left-wing bias, because the two news readers have different notions of what constitutes "good news".

Our first result in this paper demonstrates that a naïve interpretation of $u(1) > u(0)$, according to which ω_1 is a "good state" and ω_2 is a "bad state", cannot be reflected in the DM's choices of signals. An agent who evaluates ω_1 as a good state and ω_2 as a bad state may be indistinguishable from an agent with the opposite evaluation.

Proposition 1 *If u rationalizes $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$, then for every vector of real numbers $(c_1, \ldots, c_n)$, the function*

$$v(\mathbf{z}) = u(\mathbf{z}) - \sum_{i=1}^{n} c_i z_i \tag{1}$$

also rationalizes $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$.

Proof. Assume that u rationalizes $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$. Define $v(\mathbf{z})$ as in (1). Let $V(\mathbf{p},\mathbf{Q}) \equiv \sum_j \Pr(s_j|\mathbf{p},\mathbf{Q}) \cdot v[\mathbf{z}(\mathbf{p},\mathbf{Q}|s_j)]$. Hence,

$$V(\mathbf{p},\mathbf{Q}) = U(\mathbf{p},\mathbf{Q}) - c\sum_{j=1}^{m}\sum_{i=1}^{n} p_i q_{ij} = U(\mathbf{p},\mathbf{Q}) - \sum_{i=1}^{n} c_i p_i$$

Thus, for any $\mathbf{Q},\mathbf{R} \in \mathcal{Q}$, $V(\mathbf{p},\mathbf{Q}) \geq V(\mathbf{p},\mathbf{R})$ if and only if $U(\mathbf{p},\mathbf{Q}) \geq U(\mathbf{p},\mathbf{R})$. It follows that v rationalizes $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$. ■

Corollary 1 *Consider a pair of posterior beliefs, $\mathbf{z}$ and $\mathbf{z}'$, with the property that $\mathbf{z}$ assigns probability one to state ω_i, while $\mathbf{z}'$ assigns probability one to state $\omega_j \neq \omega_i$. Suppose that $u(\mathbf{z}) > u(\mathbf{z}')$. Then, we can select $c_i > u(\mathbf{z}) - u(\mathbf{z}')$ and $c_j = 0$, such that the vNM function v, defined as in (1), satisfies $v(\mathbf{z}) < v(\mathbf{z}')$.*

This simple result reveals a sense in which the model fails to capture the intuition that people "select sources of information likely to confirm 'desired' beliefs", to use Akerlof and Dickens' terminology. In our model, a distribution of posterior beliefs $\mathbf{z}$ is more desirable than a distribution $\mathbf{z}'$ if $u(\mathbf{z}) > u(\mathbf{z}')$. One may expect two agents 1 and 2, for whom $u_1(1, \ldots, 0) > u_1(0, \ldots, 1)$ and $u_2(1, \ldots, 0) < u_2(0, \ldots, 1)$, to prefer different sources of information. However, Corollary 1 shows that the two agents may display the same choices between signals. Thus, the utility rankings of posterior beliefs are not entirely meaningful, as far as choices over signals are concerned.

To take an extreme case, let $n = 2$ and rewrite u as a function of z_1 (the posterior probability of ω_1) only. Suppose that $u(\cdot)$ is strictly increasing. Normalize $u(\cdot)$ such that $u(0) = 0$. Define $v(z_1) = u(z_1) - c_1 z_1$. Let $c_1 = 2 \max u'(\cdot)$. It is easy to show that v is strictly *decreasing* in z_1. Thus, every utility ranking induced by u is reversed by v. Nevertheless, u and v rationalize the same attitude towards information.

Proposition 1 is a straightforward consequence of the assumption that the DM uses Bayes' rule to update his beliefs. In fact, the only feature of Bayes' rule that is responsible for the result is the *law of iterated expectations*: i.e., the mean of $z_i(\mathbf{p}, \mathbf{Q})$ is p_i. This property implies that the DM's choices of signals can only reveal how he ranks probability distributions over posteriors whose mean is equal to his prior. In particular, no choice between signals can reveal the utility ranking between any distinct posteriors $\mathbf{z}$ and $\mathbf{z}'$.

5 Failure to explain anomalies

Our second collection of results addresses several real-life examples of anomalous attitudes to information, which intuitively seem to be a result of anticipatory feelings. We ask whether these attitudes can be rationalized by maximization of expected utility from posterior beliefs.

Examples 2 and 3 of Section 3 describe DMs whose attitudes to information vary with their prior. In Example 2, when the patient is relatively confident that he is healthy, he prefers to be fully informed; but when he is relatively confident that he is ill, he does not wish to be fully informed. In Example 3, the DM prefers to be fully informed when he is relatively confident about the true state; but when he is less confident about the true state, he wishes to remain uninformed. In both cases, although the DM's attitude to full information varies with his prior, his attitude to full information is unambiguous when his prior lies near (at least) one of the extreme points in $\Delta(\Omega)$.

The examples are expressed in terms of a two-state model. The following definition captures this property for arbitrary n. Some notation will be useful. First, let the $n \times n$ unit matrix $\mathbf{I}$ denote the fully informative signal ($q_{ij} = 1$ if $i = j$ and $q_{ij} = 0$ if $i \neq j$). Second, for every $i = 1, ..., n$, let $e_i \in \Delta(\Omega)$ denote the degenerate probability distribution that assigns probability one to ω_i. Thus, $\{e_1, ..., e_n\}$ is the set of extreme points in $\Delta(\Omega)$.

Definition 2 *A preference profile* $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ *satisfies* ***no reversal near an extreme prior (NREP)*** *if there exists* $i \in \{1, ..., n\}$ *and a neighborhood* O_i *of* e_i*, such that either* $\mathbf{I} \succsim_{\mathbf{p}} \mathbf{Q}$ *for every prior* $\mathbf{p} \in O_i$ *and every signal* $\mathbf{Q}$*, or* $\mathbf{Q} \succsim_{\mathbf{p}} \mathbf{I}$ *for every prior* $\mathbf{p} \in O_i$ *and every signal* $\mathbf{Q}$*.*

This property means that when the DM is very confident that some state of the world is true, he ranks complete information either above or below all other signals. That is, he does not display preference reversal with respect to complete information when his prior is close to one of the extreme points in $\Delta(\Omega)$. Our first result in this section shows that if a DM satisfies NREP, then he cannot satisfy any kind of preference reversal with respect to complete information.

Proposition 2 *Suppose that* $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ *satisfies NREP. Then, either* $\mathbf{I} \succsim_{\mathbf{p}} \mathbf{Q}$ *for every prior* $\mathbf{p}$ *and signal* $\mathbf{Q}$*, or* $\mathbf{Q} \succsim_{\mathbf{p}} \mathbf{I}$ *for every prior* $\mathbf{p}$ *and signal* $\mathbf{Q}$*.*

Proof. By Proposition 1, if the vNM utility function v over posteriors rationalizes $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$, then for every real vector $(c_1, ..., c_n)$, the vNM utility function u defined by $u(\mathbf{z}) = v(\mathbf{z}) - \sum_{i=1}^{n} c_i z_i$ also rationalizes $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$. Let $c_i = v(e_i)$. Then,

$u(e_i) = 0$ for every $i = 1, ..., n$. It follows that $U(\mathbf{p}, \mathbf{I}) = 0$ for every prior p. That is, the DM's indirect utility from the fully informative signal can be set to 0, independently of his prior.

Suppose that $(\succsim_{\mathbf{p}})_{\mathbf{p} \in \Delta(\Omega)}$ satisfies NREP. Without loss of generality, suppose that for every prior $\mathbf{p}$ that is sufficiently close to e_2, $\mathbf{I} \succsim_{\mathbf{p}} \mathbf{Q}$ for every signal $\mathbf{Q}$. Construct the $n \times n$ signal $\mathbf{Q}$ as follows: $q_{11} = 1$, and for every $i > 1$, $q_{ii} = 1 - q_{i1}$. Then, for every $j > 1$, $\mathbf{z}(\mathbf{p}, \mathbf{Q}|s_j) = e_j$. Now consider the posterior that results from the realization s_1. For every $i = 1, ..., n$:

$$\frac{\mathbf{z}_i(\mathbf{p}, \mathbf{Q}|s_1)}{\mathbf{z}_1(\mathbf{p}, \mathbf{Q}|s_1)} = (1 - q_{ii}) \cdot \frac{p_i}{p_1} \tag{2}$$

Let $\mathbf{x}$ be a non-extreme point in $\Delta(\Omega)$. The question is whether we can find $(q_{ii})_{i=2,...,n}$, $q_{ii} \leq 1$, such that $\mathbf{z}_i(\mathbf{p}, \mathbf{Q}|s_1) = \mathbf{x}_i$ for every $i = 1, ..., n$. Note that if p_1 is sufficiently small relative to $p_2, ..., p_n$, we can find such numbers (q_{ii}), such that the system of equations given by (2) will be satisfied. Clearly, there exists a prior $\mathbf{p}$ arbitrarily close to some e_2 such that this requirement is met. By assumption, $\mathbf{I} \succsim_{\mathbf{p}} \mathbf{Q}$ at this prior $\mathbf{p}$. Therefore, $U(\mathbf{p}, \mathbf{I}) \geq U(\mathbf{p}, \mathbf{Q})$. Because $U(\mathbf{p}, \mathbf{I}) = 0$, $U(\mathbf{p}, \mathbf{Q}) \leq 0$. Because $\mathbf{z}(\mathbf{p}, \mathbf{Q}|s_1) = \mathbf{x}$ and $\mathbf{z}(\mathbf{p}, \mathbf{Q}|s_j) = e_j$ for every $j > 1$, it follows that $u(\mathbf{x}) \leq 0$. But this means that $u(\mathbf{x}) \leq 0$ for every $\mathbf{x} \in \Delta(\Omega)$. Therefore, for every prior $\mathbf{p}$ and every signal $\mathbf{Q}$, $\mathbf{I} \succsim_{\mathbf{p}} \mathbf{Q}$.

By the same reasoning, it can be shown that if for every prior $\mathbf{p}$ that is sufficiently close to e_2, $\mathbf{Q} \succsim_{\mathbf{p}} \mathbf{I}$ for every signal $\mathbf{Q}$, then $\mathbf{Q} \succsim_{\mathbf{p}} \mathbf{I}$ for every prior $\mathbf{p}$ and every signal $\mathbf{Q}$. ■

This result states that if a DM has an unambiguous attitude to complete information when his prior is close to some extreme point, then he must display the same unambiguous attitude to complete information at *any* prior. Thus, the model is incapable of accounting for the anomalous attitudes to information described in Examples 2 and 3.

Proposition 2 does not rule out the rationalizability of the attitude to information depicted in Example 4, which also involves a complex attitude to complete information. The following is another version of the same example. An agent initially believes that the state of the world agrees with the political views of the left. The agent's decision problem is to choose which newspaper to read. A newspaper with a left-wing bias reports all the news that validate the left-wing view, but only part of the news that validate the right-wing view. A newspaper with a right-wing bias is similarly defined. The DM prefers a left-wing biased newspaper to an objective (i.e., fully informative) newspaper, and he prefers the latter to a newspaper with a right-wing bias.

The following definition formalizes the behavior described in both versions of Example 4. The example is expressed in terms of a model with two states and binary signals (i.e., $n = m = 2$). In this case, generalization to larger n and m makes little sense.

Definition 3 *Let $n = m = 2$. A preference profile $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ displays a* ***preference for type I error*** *if there exists $\alpha^* \in (0,1)$ such that whenever $p_1 > \frac{1}{2}$, we have*

$$\begin{pmatrix} \alpha & 1-\alpha \\ 0 & 1 \end{pmatrix} \succ_{\mathbf{p}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \succ_{\mathbf{p}} \begin{pmatrix} 1 & 0 \\ 1-\alpha & \alpha \end{pmatrix}$$

for all $\alpha \in [\alpha^*, 1)$.

In Example 4, the DM's preference for a type I error intuitively seems to result from an attempt to attain desired beliefs (the thesis is of high quality, the left-wing view is correct) and avoid holding undesired beliefs (the thesis is of low quality, the right-wing view is correct). The question is whether this intuition can be accommodated into the model of expected-utility maximization over posterior beliefs.

Proposition 3 *A preference profile $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$ displaying a preference for type I error is not rationalizable.*

Proof. For expositional convenience, let us use the following abbreviated notation: $p = p_1$, and a signal Q will be represented by the pair (q_{11}, q_{22}). Assume, contrary to the claim, that there exists $\alpha^* \in (0,1)$ such that for every prior $p > \frac{1}{2}$ we have $(1,\alpha) \succ_p (1,1) \succ_p (\alpha,1)$ for all $\alpha \in [\alpha^*, 1)$. By Proposition 1, we can normalize $u(0,1) = u(1,0) = 0$. Thus, $U(\mathbf{p}, (1,1)) = 0$ for every prior. Let $p > \frac{1}{2}$. Then:

$$\begin{aligned} z_1(\mathbf{p}, (1,\alpha) \mid s_1) &= \frac{p}{p + (1-p)(1-\alpha)} \\ z_1(\mathbf{p}, (1,\alpha) \mid s_2) &= 0 \end{aligned}$$

Similarly:

$$\begin{aligned} z_1(\mathbf{p}, (\alpha,1) \mid s_1) &= 1 \\ z_1(\mathbf{p}, (\alpha,1) \mid s_2) &= \frac{p(1-\alpha)}{p(1-\alpha) + (1-p)} \end{aligned}$$

By our initial assumption, $(1,\alpha) \succ_{\mathbf{p}} (1,1) \succ_{\mathbf{p}} (\alpha,1)$ for all $\alpha \in [\alpha^*, 1)$. This implies that $u[\mathbf{z}(\mathbf{p}, (1,\alpha) \mid s_1)] > 0$ and $u[\mathbf{z}(\mathbf{p}, (\alpha,1) \mid s_2)] < 0$ for all $\alpha \in (\alpha^*, 1)$. Denote $\mathbf{x}^* = \mathbf{z}(\mathbf{p}, (1,\alpha^*) \mid s_1)$. For every $\alpha > \alpha^*$, $z_1(\mathbf{p}, (1,\alpha) \mid s_1) > x_1^*$. It follows that $u(x) > 0$ whenever $x_1^* < x_1 < 1$. Observe that $\lim_{p \to 1} z_1(\mathbf{p}, (\alpha,1) \mid s_2) = 1$. Therefore, we can choose a prior $\mathbf{p}^*$ satisfying $p^* \in (x_1^*, 1)$, such that $z_1(\mathbf{p}^*, (\alpha^*,1) \mid s_2) > x_1^*$, hence $u[z_1(\mathbf{p}^*, (\alpha^*,1) \mid s_2)] > 0$, a contradiction. ■

Why does this result run contrary to our intuition? Recall that the DM in our model is Bayesian. The posteriors induced by a biased signal of the form $(1,\alpha)$ are $(0,1)$ and some (z_1, z_2) with $z_1 \in (p_1, 1)$. Reducing α has two effects. First, the probability of the interior posterior (z_1, z_2) increases. Second, z_1 itself decreases and becomes closer to p_1. Thus, even if $u(\cdot)$ is monotone in z_1, increasing the newspaper's bias has an ambiguous effect on the DM's expected utility from posteriors. Our intuitions regarding the anticipatory-feelings source of preference for type I error seem to reflect only the first effect.

Propositions 2 and 3 strongly rely on the fact that the posteriors induced by the fully informative signal **I** are always extreme points, regardless of the prior. By Proposition 1, we can set $u(e_i) = 0$ for every $i = 1, ..., n$, such that the DM's indirect expected utility from the fully informative signal is constant across all priors. This property greatly restricts the model's ability to accommodate preference reversals with respect to complete information.

The lesson from Propositions 2 and 3 is that commonly observed, anomalous attitudes to information, which seem to be a consequence of anticipatory feelings, cannot be explained by maximization of expected utility from beliefs. Note that although the model has difficulties in accounting for preference reversals with respect to full information, it is capable of rationalizing some preference profiles that exhibit aversion to information, such as in the case of Example 1.

Remark 1 *Let $\Omega = \{\omega_1, \omega_2\}$ and let u rationalize a preference profile $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$. If u is concave, then the DM is information averse. If u is convex, the DM is information seeking.*

Proof. Suppose that u is concave. Consider pair of signals **Q** and **R** with $q_{ii}, r_{ii} \in \left[\frac{1}{2}, 1\right]$. Let $\mathbf{z} \equiv \mathbf{z}(\mathbf{p}, \mathbf{Q})$ and $\mathbf{z}' \equiv \mathbf{z}(\mathbf{p}, \mathbf{R})$. In addition, denote by $z_i(s_i)$ and $z_i'(s_i)$ the posterior probability of ω_i, conditional on the realization s_i, under each of two signals. Assume w.l.o.g. that $z_1(s_1) > z_1(s_2)$ and $z_1'(s_1) > z_1'(s_2)$. If $\mathbf{Q} > \mathbf{R}$, then $z_1(s_1) > z_1'(s_1)$ and $z_1(s_2) < z_1'(s_2)$. By concavity of u, any convex combination of $u[\mathbf{z}(s_1)]$ and $u[\mathbf{z}(s_2)]$ lies below any convex combination of $u[\mathbf{z}'(s_1)]$ and $u[\mathbf{z}'(s_2)]$. Hence, $U(\mathbf{p}, \mathbf{R}) \geq U(\mathbf{p}, \mathbf{Q})$, implying that $\mathbf{R} \succsim_{\mathbf{p}} \mathbf{Q}$. The case of convex u is handled similarly. ■

Remark 1 is not an original result. Recall the analogy drawn at the end of Section 2 between our model and the more familiar literature on risk aversion and attitudes to spread. An information averse DM in our model is analogous to an agent who always prefers to reduce the spread of a fair lottery. As the works cited at the end of Section 2 show, if an agent has a concave vNM utility function over his final wealth, he has a preference for smaller spread. Remark 1 is a translation of (a special case of) these results into the present context. We conclude that the model's difficulty is not in accounting for aversion to information per se, but in accounting for attitudes to full information that vary with the prior.

6 How to proceed?

The results of Sections 4 and 5 demonstrate that the extended expected-utility model, which incorporates the DM's belief as an argument in his vNM utility function, has great difficulties in accounting for anomalous attitudes to information, as long as one assumes that the DM's inferences are Bayesian. First, ordinal utility representation fails. Two agents with opposite ordinal rankings of posterior beliefs may choose the same information sources. This result places severe limits on any "applied" model along these lines that tries to correlate differences in preferences over beliefs to differences in behavior. Second, the model turns out to be inconsistent with several real-life examples of anomalous attitudes to information, which intuitively seem to result from anticipatory feelings.

In this section, we discuss several ways to proceed from these negative results. Of course, one way to deal with the negative results is to return to the conventional assumption that people are emotionally neutral towards their beliefs, and to deny that anticipatory feelings influence people's attitudes to information. We disagree with this interpretation of our results. The notion that people sometimes avoid information because they fear the conclusions that it might imply is highly intuitive. The question is how to construct a model that captures this intuition, given the failure of the model of expected utility from posteriors coupled with Bayesian updating.

Non-Expected Utility

All the existing economic models of anticipatory feelings, described in Section 2, assume that the DM has expected-utility preferences over beliefs. We have followed this practice in this paper. However, it could be argued that some of the model's shortcomings can be overcome by abandoning expected utility in favor of familiar non-expected-utility theories. We provide two arguments to the contrary.

The most tractable non-expected utility model evaluates a probability distribution σ over posteriors in the following way: $V(\sigma) = \Sigma w(\pi) \cdot u(z)$, where the $w(\cdot)$'s are *decision weights.* Expected utility is a special case, $w(\pi) = \pi$. It is easy to show, along the same lines as Proposition 1, that if one representation (w, u) rationalizes $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\Delta(\Omega)}$, then we can find another representation (w', u'), such that u' is a non-monotonic linear representation of u. Thus, the utility representation u is not unique up to monotonic transformations, as long as we are also allowed to modify the decision weights. Because decision weights are unobservable, we see no reason to forbid modifying them. Thus, the most tractable generalization of expected utility also fails the ordinality test.

Much of the attraction of utility functionals lies in their axiomatic foundation. The reason economists sometimes prefer a non-expected utility functional to an expected utility functional is their discomfort with the Independence axiom and their wish to replace it with a weaker axiom. In our model, a ranking of signals at a given prior translates into a ranking of two distributions of posterior beliefs, via Bayes' rule. Since Bayes' rule is non-linear in $\mathbf{p}$ and $\mathbf{Q}$, linearity properties of u do not imply linearity properties of $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\mathbf{\Delta(\Omega)}}$. An independence axiom imposed on preferences over probability distributions over posterior beliefs will not translate into an independence axiom imposed on $(\succsim_{\mathbf{p}})_{\mathbf{p}\in\mathbf{\Delta(\Omega)}}$. Because familiar non-expected-utility theories are based on weakenings of the independence axiom, we conjecture that none of the familiar non-expected utility functionals can be given an elegant axiomatic characterization. In this respect, the modeling choice between expected and non-expected utility is rather arbitrary.

Prior-Dependent Utility

It might be argued that the negative results of Section 5 are due to a narrow specification of the consequence space and could be overcome by adding the DM's *prior* belief to the description of a consequence. In this way, a consequence in the choice model would be a pair consisting of a prior probability $\mathbf{p}$ and a posterior probability $\mathbf{z}$, such that $u(\mathbf{p}, \mathbf{z})$ may vary with $\mathbf{p}$. Recall that a key argument in the proofs of Propositions 2 and 3 was the property that $U(\mathbf{p}, \mathbf{I}) = 0$ for every prior $\mathbf{p}$.

This property ceases to hold when u is also a function of $\mathbf{p}$.

Note, however, that this extended model does not pass the ordinality test. Also, this extension is vulnerable to an infinite-regress argument, because today's prior is yesterday's posterior. Finally, the extension carries a great loss in explanatory power. Let $\mathbf{z}$ be some posterior belief and let $\mathbf{p}, \mathbf{p}'$ be two different prior beliefs. No choice experiment can reveal the ranking between the consequences $(\mathbf{p}, \mathbf{z})$ and $(\mathbf{p}', \mathbf{z})$. Therefore, no observable act of choice can reveal the ranking between the DM's expected utility induced by any signal at two different priors. The utility-maximization model thus becomes virtually powerless to account for choice behavior.

Non-Bayesian Updating

Throughout this paper, we assumed that the DM translates signals into probability distributions over posteriors in accordance with Bayes' rule. Our analysis relied on abstract properties of Bayes' rule (especially the law of iterated expectations). Given the negative results of Sections 4 and 5, an interesting step forward may be to study utility-over-beliefs representation of attitudes to information under *non-Bayesian* updating rules.

The following example illustrates how non-Bayesian updating can reverse some of our conclusions. Let $n = m = 2$. For notational convenience, let $p_1 = p$. Let the pair (q_{11}, q_{22}) represent the signal $\mathbf{Q}$, where $q_{11}, q_{22} \geq \frac{1}{2}$. Consider the following updating rule: *for every signal, the DM updates his beliefs as if the signal were* $(1,1)$. That is, with probability $pq_{11} + (1-p)(1-q_{22})$ his posterior belief is 1, and with probability $p(1-q_{11}) + (1-p)q_{22}$ his posterior is 0. This updating procedure reflects the DM's lack of distinction between a partially informative signal and a fully informative signal. He regards a partially informative signal as if it did not have type I and type II errors.

Unlike a Bayesian agent, our DM updates his prior beliefs even when he chooses the signal $(\frac{1}{2}, \frac{1}{2})$. Therefore, this signal ceases to be equivalent to acquiring no information. To allow for the latter possibility, define $\succsim_{\mathbf{p}}$ over the extended set $\left[\frac{1}{2}, 1\right]^2 \cup \{$ *"no signal"*$\}$, where "no signal" means that the DM does not observe any signal and therefore his posterior belief equals his prior belief.

Let $u : [0,1] \rightarrow \mathcal{R}$ be a continuous increasing function of z_1, with $u(0) = 0$ and $u(1) = 1$. It can be shown that for every $p \in (0,1)$, there exists a signal $(q_{11}, q_{22}) \in [\frac{1}{2}, 1]^2$ for which *"no signal"*$\sim_{\mathbf{p}}$ (q_{11}, q_{22}), such that $u(p) = pq_{11} + (1-p)(1-q_{22})$. The implication of this simple result is that u *passes the ordinality test*. For any pair $x, y \in (0,1)$, $u(x) \geq u(y)$ if and only if there is a prior $\mathbf{p}$ and a pair of binary signals (q_{11}, q_{22}) and (r_{11}, r_{22}), such that $x = pq_{11} + (1-p)(1-q_{22})$, $y = pr_{11} + (1-p)(1-r_{22})$ and $\mathbf{Q} \succsim_{\mathbf{p}} \mathbf{R}$. Note that because $u(p)$ is uniquely determined by the signal (q_1, q_2) for which $(q_1, q_2) \sim_{\mathbf{p}}$ *"no signal"*, u is unique up to *affine* transformations. Thus, given the above non-Bayesian updating rule, the DM's preferences over binary signals also reveal the cardinality of u.

This example illustrates how some of the difficulties presented in this paper may be surmounted when we allow for non-Bayesian updating. One may doubt this idea because it involves two non-standard assumptions: direct utility from beliefs and non-Bayesian updating. However, the linkage between non-standard effect of beliefs on welfare and non-Bayesian updating receives some justification in Compte and Postlewaite (2004), albeit in a somewhat different context. They show that if a DM's belief affects his performance, then biased inferences from past experiences can enhance his welfare.

Conclusion

The message of this paper is that incorporating decision makers' beliefs into their utility function raises non-trivial modeling problems. One would like a model of direct utility from beliefs to account for anomalous attitudes to information which are observed in reality, and seem to result from people's anticipatory feelings. The model of expected utility from beliefs with Bayesian updating turns out to be inadequate for this purpose. We have considered three ways to proceed. We find the third way, abandoning Bayesian updating, to be the most reasonable. If economists allow decision makers in their models to be emotionally non-neutral towards their beliefs, they should probably also allow them to be non-Bayesian in their inferences.

References

[1] Akerlof, G. and W. Dickens, "The Economic Consequences of Cognitive Dissonance", *American Economic Review,* LXXII (1982), 307-319.

[2] Bénabou, R. and J. Tirole, "Self-Confidence and Personal Motivation", *Quarterly Journal of Economics,* **470** (2002), 871-915.

[3] Blackwell, D., "Equivalent Comparison of Experiments", *Annals of Mathematics and Statistics* **24** (1953), 265-272.

[4] Brunnermeier, M. and J.P. Parker, "Optimal Expectations", mimeo (2002).

[5] Caplin, A. and K. Eliaz, "AIDS and Psychology: A Mechanism-Design Approach", *RAND Journal of Economics*, **34** (2003), 631-646.

[6] Caplin, A. and J. Leahy, "Psychological Expected Utility", *Quarterly Journal of Economics* **116** (2001), 55-79.

[7] Caplin, A. and J. Leahy, "The Supply of Information by a Concerned Expert", *Economic Journal*, **114** (2004), 487-505.

[8] Carrillo, J. and T. Mariotti, "Strategic Ignorance as a Self-Disciplining Device", *Review of Economic Studies,* LXVI (2000), 529-544.

[9] Compte, O. and A. Postlewaite, "Confidence Enhanced Performance", *American Economic Review*, forthcoming.

[10] Eyster, E, "Rationalizing the Past", Mimeo (2002).

[11] Geanakopolos, J., D. Pearce and E. Stachetti, "Psychological Games and Sequential Rationality", *Games and Economic Behavior* **1** (1989), 60-79.

[12] Grant, S., A. Kajii and B. Polak, "Preferences for Information and Dynamic Inconsistency", *Theory and Decision* **48** (2000), 263-286.

[13] Grant, S., A. Kajii and B. Polak, "Intrinsic Preferences for Information", *Journal of Economic Theory* **83** (1998), 233-259.

[14] Kőzsegi, B., "Health Anxiety and Patient Behavior", *Journal of Health Economics*, **22** (2003), 1073-1084.

[15] Kőzsegi, B., "Emotional Agency: The Case of Doctor-Patient Relationships", Mimeo (2004a).

[16] Kőzsegi, B., "Anticipation in Observable Behavior", Mimeo (2004b).

[17] Kreps, D. and E. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica* **46** (1978), 185-200.

[18] Landsberger, M. and I. Meilijson, "Lotteries, Insurance, and Star-Shaped Utility Function," *Journal of Economic Theory* **52** (1990), 1-17.

[19] Lerman, C., C. Hughes, S. J. Lemon, D. Main, C. Snyder, C. Durham, S. Narod and H. T. Lynch, What You Don't Know Can Hurt You: Adverse Psychological Effects in Members of BRCA1-Linked and BRCA2-Linked Families Who Decline Genetic Testing", *Journal of Clinical Oncology* **16** (1998), 1650-1654.

[20] Rothschild, M. and J. Stiglitz, "Increasing Risk: I. A Definition," *Journal of Economic Theory* **2** (1970), 225-243.

[21] Safra, Z. and E. Sulganik, "On the Nonexistence of Blackwell's Theorem-type Results With General Preference Relations", *Journal of Risk and Uncertainty* **10** (1995), 187-201.

[22] Wakker, P., "Non-expected Utility as Aversion of Information", *Journal of Behavioral Decision Making* **1** (1988), 169-175.

[23] Yariv, L., "I'll See it when I believe it - A Simple Model of Cognitive Consistency", mimeo (2002).