Disclosure to a Psychological Audience

by *Elliot Lipnowski & Laurent Mathevet*[*]

NEW YORK UNIVERSITY[†]

April 2, 2015

Abstract

Should a well-intentioned advisor always tell the whole truth? In standard economics, the answer is yes (Blackwell (1953)), but in the world of psychological preferences (Geanakoplos, Pearce, and Stacchetti (1989))—where a listener's state of mind has a direct impact on his well-being—things are not so simple. In this paper, we study how a benevolent principal should disclose information to a psychological agent. After characterizing attitudes toward information, we study optimal information disclosure. Psychological aversion to information is of particular interest. We show that, for information-averse agents, the principal may simply provide information in the form of a recommended action. Next, we study how the optimal policy changes with information-aversion. We also offer general tools of optimal disclosure. We apply our results to choices under prior-bias and cognitive dissonance; consumption-saving decisions with temptation and self-control problems (Gul and Pesendorfer (2001)); and doctor-patient relationships.

Keywords: information disclosure, psychology, information-aversion, temptation, persuasion, transparency.

JEL Classification: D03, D83, D82, D69, 118, M30.

1

1 Introduction

Choices of information disclosure—what is revealed and what is concealed—affect what people know and, in turn, the decisions that they make. In many contexts, information disclosure is intended to improve people's welfare. In public policy, for instance, regulators impose legal rules of disclosure, such as required transparency by a seller of a good to a buyer, to this effect. But should a well-intentioned advisor always tell the whole truth? In standard economics, the answer is yes, because a better informed person makes better decisions (Blackwell (1953)). However, parents do not always tell the whole truth to their children; people conceal embarrassing details about their personal lives from their family elders; doctors do not always reveal to their patients all the details of their health, when dealing with hypochondriac patients for instance; future parents do not always want to know the sex of their unborn child, even if it is pragmatic to know in advance; and many of us would not always like to know the caloric content of the food that we eat.

In this paper, we study how an informed principal should disclose information to an agent with psychological traits. It is hard to argue that our beliefs and our information affect our lives only through the actions that we take. The mere fact of knowing or not knowing something can be a source of comfort or discomfort. Bearing this in mind, how should we disclose information to a person whose state of mind has a direct impact on his well-being? This question is relevant when the principal is a trusted advisor to the agent, but also when the principal is an entity that, by law, can enforce what the informed party must disclose. Laws and regulations about public disclosure of information are widespread, ranging from administrative regulations, loans, credit cards, and sales, to insurance contracts, restaurants, doctor-patient relationships and more (see Ben-shahar and Schneider (2014)).

In their paper, "Disclosure: Psychology Changes Everything," Loewenstein, Sunstein, and Goldman (2014) argue that psychology may lead one to rethink public disclosure policies. For example, it may be counter-productive to offer detailed information to a person with limited attention, as she may ignore critical information due to cognitive limitations. Likewise, if people are nonstandard in how they process information, i.e., if they are not always Bayesian, then the regulator should take into account their biased probability judgments in the choice of disclosure policies.

At a more personal level, a well-intentioned speaker may directly take into account the psychological well-being of her interlocutor when deciding what to say. For instance, one sometimes experiences cognitive dissonance, an innate discomfort at having one's beliefs shift from an existing position: how should information disclosure best anticipate this? Other examples come to mind, in which one would like to be fully informed when the news is good, but would rather not know how bad it is otherwise. The first communication scheme that

comes to mind—reveal the news when it is good and say nothing otherwise—accomplishes little in this situation, as no news reveals bad news.

In our model, an informed principal will learn the realization of a state of nature, and she must choose what to tell the agent in each contingency. After receiving his information from the principal, the agent updates his prior beliefs and then takes an action. The principal wants to maximize the agent's ex-ante expected utility by choosing the right disclosure policy.

To study the impact of psychology on information disclosure, we must decide how to model psychology and communication. We model the former by assuming that the agent's satisfaction depends not only on the physical outcome of the situation, but also directly on his updated beliefs. In doing so, we employ the framework of psychological preferences (Geanakoplos, Pearce, and Stacchetti (1989)). This framework captures a wide range of emotions and, in our model, allows us to represent various psychological phenomena, such as purely psychological preferences, prior-bias and cognitive dissonance, temptation and self-control problems (Gul and Pesendorfer (2001)), multiplier preferences (Hansen and Sargent (2001)), and more.

We model communication by assuming that the principal can commit ex-ante, before the state is realized, to an information policy. This means that to every possible state of nature corresponds a (possibly random) message by the principal, leading the agent to update his beliefs. Since the state is distributed ex-ante according to some prior distribution, an information policy induces a distribution over updated (or posterior) beliefs and (given how the agent responds to said information) over actions. This is the methodology of random posteriors employed by Kamenica and Gentzkow (2011) (KG hereafter). Alternative methodologies are available, most notably the 'cheap talk' model of Crawford and Sobel (1982), wherein the principal cannot commit ex-ante. Here, given our focus on the impact of psychology on disclosure, we give the principal full commitment power, abstracting from strategic interactions between different (interim) principal types. Depending on the application, we see at least two reasons that such commitment power is a reasonable approximation. First, in many applications, a benevolent principal is some third-party who regulates information disclosure between a sender (e.g., a seller) and the agent, with the agent's interests in mind. Such a principal sets the disclosure policy in ignorance of the realized state, and thus faces no interim opportunity to lie. Second, in situations in which the relationship is subject to full-transparency regulations (e.g., patients have the right to access their lab test results), this legislative restriction creates commitment power. The principal makes an ex-ante decision of which information to seek and is bound by law to reveal whatever she finds.

Psychological preferences entail a number of new considerations and ask for new tools of information disclosure. First, it is important to understand the landscape of psychological preferences, for agents' relationship to information will guide disclosure. A first-order

question is: "Does the agent like information?" A classical agent is intrinsically indifferent to information. Holding his choices fixed, he derives no value from knowing more or less about the world. But if he can respond to information, then he can make better decisions and, thus, information is always a good thing (Blackwell (1953)). This observation leads us to distinguish between two general standards of information preference. On one hand, an agent may be *psychologically* information-loving [or averse] if he likes [dislikes] information for its own sake. On the other, he may be *behaviorally* information-loving [or averse] if he likes [dislikes] information, taking its instrumental value into account. To see the distinction, imagine a person will give a seminar, and consider the possibility of him finding out whether his presentation has any embarrassing typos right before he presents. If his laptop is elsewhere, he may prefer not to know about any typos, as knowing will distract him during the presentation. He is psychologically information-averse. But if he has a computer available to change the slides, he would like to know; any typos can then be changed, which is materially helpful. He is behaviorally information-loving.

We characterize these four classes of preferences in Theorem 1; and optimal information disclosure follows readily for three of them. Almost by definition, if the agent is behaviorally information-loving [-averse], it is best to tell him everything [nothing]. If an agent is psychologically information-loving, then the potential to make good use of any information only intensifies his preference for it; he is then behaviorally information-loving, so that giving full information is again optimal. More interesting is the case in which information causes some psychological distress, i.e., a psychologically information-averse agent. This case presents an intriguing dilemma, due to the tradeoff between the agent's need for information and his dislike for it. For such an agent, information is a liability that is only worth bearing if it brings enough instrumental value. A large part of the paper is devoted to information disclosure under information-aversion.

Under psychological information-aversion, and in psychological environments in general, the Revelation Principle no longer holds. In the presence of many potential disclosure policies, some of them being unreasonably cumbersome, this observation is disappointing. However, although the Revelation Principle fails in this environment, its most useful design implication remains: Theorem 2 says that (when preferences display psychological information-aversion) it is sufficient to confine attention to recommendation policies; that is, the principal can simply tell the agent what to do in an incentive-compatible fashion.[1] Recommendation policies are especially interesting because they are rather natural. In the doctor-patient example, the doctor need not tell the patient what the test statistically reveals about his illness; she may simply tell him which medical action to take.

---

[1]To be sure, this result is nontrivial. Its proof is different from typical proofs of the Revelation Principle—a point on which we elaborate in the main text—and the result is false for general psychological preferences.

Many recommendation policies can a priori be optimal for a given agent with psychological information-aversion. Can we say anything more for agents who are ranked by their information-aversion? We propose an (apparently very weak) order on information-aversion, and yet obtain a strong conclusion. When two agents are ranked according to information-aversion, their indirect utilities differ by the addition of a concave function. In this model, concavity is tantamount to disliking information everywhere, and hence one would think that a designer should respond by providing unambiguously less information to a more information-averse agent. Interestingly, this is only true in a two-state world; it fails in general due to psychological substitution effects.

We also provide general tools of optimal information disclosure to respond to the variety of psychological preferences, especially those that fall outside our classification. The concave envelope result of KG holds in our model (Theorem 3), so that we can characterize the agent's optimal welfare. But the benevolence of our design environment allows us to develop the method of "posterior covers," so that we can explicitly derive an optimal policy. An agent may not be information-loving as a whole, but he can be locally information-loving at some beliefs. Accordingly, the principal knows to be as informative as possible in those regions. A posterior cover is a collection of convex sets of posterior beliefs over each of which the agent is information-loving. By Theorem 4, the search for an optimal policy can be limited to the extreme points of a posterior cover of the indirect utility. By Proposition 3—which hinges on the assumption of aligned preferences—this problem can be reduced to that of finding posterior covers of primitives of the model. By Proposition 4, in a broad class of economic problems, all these objects can be computed explicitly.

How should we talk to someone facing a prior-bias or cognitive dissonance? To someone who has temptation and self-control problems? Or to someone who likes to hear news only when it is good? We apply our results to answer these questions. In the case of cognitive dissonance,[2] the main message is that we should communicate with such an individual by an all-or-nothing policy. Either the person can bear information, in which case we should tell him the whole truth to enable good decision-making, or he cannot, in which case we should say nothing.

In the case of temptation and self-control, the optimal policy depends on the nature of information. We study a standard consumption-saving problem where an individual has impulsive desires to consume rather than invest in some asset. When information is about the asset's return, it only appeals to the agent's rational side. Hence, full information is optimal. But when information is about (say) the weather in a vacation destination, it only appeals to the agent's impulsive side. Yet, it does not follow that no information is optimal. By

---

[2]In our model of cognitive dissonance, the agent pays a psychological penalty given by the (Euclidean) distance between his posterior belief and their prior position.

(sometimes) revealing good weather, an advisor induces the agent to consume and avoid all temptation, or to invest and face reduced temptation; partial revelation may then be optimal. We further show that the method of posterior covers can be brought to bear quite generally in the setting of temptation and self-control.

Finally, we analyze a doctor-patient relationship in which the patient, who can be either healthy or ill, has equivocal information preferences. He likes to be informed about his health when the news is good, but when it is bad enough that no useful course of action exists, he prefers not to know. He is neither information-averse nor -loving. Of course, it is suboptimal to only deliver good news, for the patient would infer that the news is bad when she hears nothing. It turns out that the doctor should recommend 'no treatment' only when the patient is healthy; but the doctor should at times deliver false positives, recommending the aggressive course of action (e.g., never smoke again) even in good health. In doing so, the doctor provides some information even when the patient does not like it locally, as the potential to provide very good news with some probability outweighs the damage of bad news.

This paper is part of a recently active literature on information design (KG; Ely, Frankel, and Kamenica (2014); Kolotilin et al. (2015); etc). The most related work, methodologically speaking, is KG. We adopt the same overall method as they do, but study a different problem and give new tools (which would be inapplicable to their setting of conflicting interests) to solve it.[3]

Our work does not belong to the cheap talk literature spanned by Crawford and Sobel (1982), because our principal can commit to a communication protocol. A notable contribution to that literature is Caplin and Leahy (2004) who study a detailed interaction between a doctor and a patient with quadratic psychological preferences. Unlike us, they allow for private information on the part of the agent. Our analysis concerns a strategically simpler interaction, but encompasses a great range of applications.

There is a conceptual connection between psychological preferences and nonstandard preferences over temporal resolution of uncertainty (e.g., Kreps and Porteus (1978)), which can be reinterpreted as a psychological effect. As such, the mathematical arguments behind our classification of preferences are similar to Grant, Kajii, and Polak (1998). However, we make a conceptual distinction between psychological and behavioral information preferences that is critical for information disclosure. Moreover, information disclosure imposes additional constraints that are absent in that literature—specifically Bayes-consistency—on the domain of the agent's preferences. Indeed, the beliefs of a Bayesian agent cannot be manipulated at will. This limits the scope of analysis for such preferences, for example when

---

[3]At the intersection of both works is but one trivial configuration: a non-psychological agent with the same preferences as the principal. In this case, full disclosure is trivially optimal.

comparing information-aversion (Section 6).

The paper is organized as follows. Section 2 introduces psychological preferences and the methodology of random posteriors. Section 3 presents the psychological agent, and studies and characterizes various attitudes towards information, namely psychological and behavioral information preferences. Section 4 concentrates on psychological information-aversion, highlighting information disclosure and proposing an order to compare information-aversion. Section 5 supplies the general tools of optimal information design. Section 6 applies these tools to various situations, including forms of prior-bias and cognitive dissonance, temptation and self-control, and equivocal information preferences.

2 The Environment

Consider an agent who must make a decision when the state of nature $\theta \in \Theta$ is uncertain.[4] Suppose that the agent has (full-support) prior $\mu \in \Delta\Theta$ and that he has access to additional information about the state. After receiving that information, the agent forms a posterior belief by updating his prior, and then he makes a decision.

2.1 Psychological Preferences

An *outcome* is an element (a, v) of $A \times \Delta\Theta$ where a is the action taken by the agent and $v \in \Delta\Theta$ denotes the agent's posterior belief at the moment when he makes his decision. We assume that the agent has a continuous utility function $u : A \times \Delta\Theta \to \mathbb{R}$ over outcomes. The true state is excluded from the utility function, without loss of generality: given any true underlying preferences $\tilde{u} : A \times \Theta \times \Delta\Theta \to \mathbb{R}$, we can define the reduced preferences u via $u(a, v) = \int_\Theta \tilde{u}(a, \theta, v) \, \mathrm{d}v(\theta)$. The reduced preferences are the only relevant information, both behaviorally and from a welfare perspective.[5] Given posterior beliefs v, the agent chooses an action $a \in A$ so as to maximize $u(a, v)$.

In the *classical case*, $\tilde{u}$ does not depend on beliefs and so $u(a, \cdot)$ is affine for every a. We do not make this assumption here. In our environment, the agent's satisfaction depends not only on the physical outcome of the situation, but also on his posterior beliefs. In the

---

[4]In this paper, all spaces are assumed nonempty compact metrizable spaces, while all maps are assumed Borel-measurable. For any space Y, we let $\Delta Y = \Delta(Y)$ denote the space of Borel probability measures on Y, endowed with the weak* topology, and so itself compact metrizable. Given $\pi \in \Delta Y$, let $\mathrm{supp}(\pi)$ denote the support of π, i.e., the smallest closed subset of Y of full π-measure.

[5]Later, when we consider questions of information design by a benevolent principal, we will also work with u rather than $\tilde{u}$. This remains without loss, but only because the designer is assumed to have full commitment power. This is in contrast to the setting of Caplin and Leahy (2004).

literature, this is known as *psychological preferences* (Gilboa and Schmeidler (1988) and Geanakoplos, Pearce, and Stacchetti (1989)).[6]

This formulation covers a wide range of phenomena in our environment. Consider the following examples where $\Theta = \{0, 1\}$ and $A = [0, 1]$, and note that we can write any posterior belief as $v = \text{Prob}(\{\theta = 1\})$:

(a) (Purely psychological agent) Let

$$u(a, v) = -\mathbb{V}(v)$$

represent an agent whose only satisfaction comes from his degree of certainty, as quantified by the variance $\mathbb{V}(v) = v(1 - v)$.

(b) (Psychological and material tensions) For $k \in \mathbb{R}$, let

$$u(a, v) = k\mathbb{V}(v) - \mathbb{E}_{\theta \sim v}\left[(\theta - a)^2\right]$$

This agent wants to guess the state but he has a psychological component based on the variance of his beliefs. When $k > 0$, the agent faces a tension between information, which he dislikes per se, and his need for it to guess the state accurately.

(c) (Prior-bias and cognitive dissonance) Let

$$u(a, v) = -|v - \mu| - \mathbb{E}_{\theta \sim v}\left[(\theta - a)^2\right]$$

represent an agent who wants to guess the state but experiences discomfort when information conflicts in any way with his prior beliefs. This functional form is reminiscent of Gabaix (2014)'s sparse-maximizer, though the economic interpretation is quite different: a distaste for dissonance, rather than a cost for considering information.

(d) (Temptation and self-control) Given two classical utility functions u_1 and u_2, and action set A, let

$$u(a, v) = \mathbb{E}_{\theta \sim v}\left\{u_1(a, \theta) - \max_{b \in A} \mathbb{E}_{\hat{\theta} \sim v}\left[u_2(b, \hat{\theta}) - u_2(a, \hat{\theta})\right]\right\},$$

be the form proposed by Gul and Pesendorfer (2001). After choosing a menu A, the agent receives some information and then chooses an action. His non-tempted self experiences utility u_1 from action a, while $\max_{b \in A} \mathbb{E}_{\theta \sim v}[u_2(b, \theta) - u_2(a, \theta)]$ is interpreted as the cost of self-control.

---

[6]In a game-theoretic context, a player's utility function could depend on the beliefs of the other players; preferences concerning fashion and gossip are examples of this phenomenon. In a dynamic context, a player's utility at some time could depend on the relationship between his current beliefs and any of his past beliefs; surprise and suspense are examples (Ely, Frankel, and Kamenica (2014)).

(e) (Ambiguity-aversion: multiplier preferences) For $k > 0$, let

$$u(a, v) = \min_{v'} \int u_c(a, \theta) dv'(\theta) + kR(v' \| v)$$

where $R(v' \| v)$ is the relative entropy of v' with respect to v.[7] These are the multiplier prefer-ences proposed by Hansen and Sargent (2001). The agent has some best guess v of the true distribution, but he does not fully trust it. Thus, he considers other probabilities v' whose plausibility decreases as the distance to v increases.

2.2 Signals and Random Posteriors

Before making his decision, the agent has access to additional information, which is given by a signal.

Definition 1. *A signal (S, σ) is a space S equipped with a map $\sigma : \Theta \to \Delta S$.*

A signal is a technology that sends a random message to the agent in each state: if the state is θ, the realized message $s \in S$ is distributed according to $\sigma(\cdot|\theta) \in \Delta S$. The signal is the only information that the agent receives about the state.

Given the signal (S, σ) and the realized message s, the agent forms posterior beliefs about the state, denoted $\beta^{S,\sigma} : S \to \Delta\Theta$.[8] Given that a signal sends a message in every state, each time leading to a posterior belief, a signal induces a distribution over posterior beliefs when seen from an ex-ante perspective. We call this distribution, which is an element of $\Delta\Delta\Theta$, a *random posterior*. The random posterior correspondence,

$$\begin{aligned}
\mathcal{R} : \Delta\Theta &\rightrightarrows \Delta\Delta\Theta \\
\mu &\longmapsto \{p \in \Delta\Delta\Theta : \mathbb{E}_{v \sim p} v = \mu\} \\
&= \left\{ p \in \Delta\Delta\Theta : \int_{\Delta\Theta} v(\hat{\Theta}) \, dp(v) = \mu(\hat{\Theta}) \text{ for every Borel } \hat{\Theta} \subseteq \Theta \right\},
\end{aligned}$$

maps every prior into the set of random posteriors whose mean equals the prior. As shown in Kamenica and Gentzkow (2011) and stated for completeness in the lemma below, Bayesian updating must yield random posteriors in $\mathcal{R}(\mu)$, and conversely, all such random posteriors can be generated by some signal. While there is room for manipulation, the posterior beliefs of a Bayesian agent must on average equal his prior.

---

[7]For any $v, v' \in \Delta\Theta$, recall that $R(v' \| v) = \sum_\theta \ln (v'(\theta)/v(\theta)) \, v'(\theta)$.

[8]In the finite case, the posterior beliefs are given by $\beta^{S,\sigma}(\theta|s) = \sigma(s|\theta)\mu(\theta) / \sum_{\hat{\theta}} \sigma(s|\hat{\theta})\mu(\hat{\theta})$ for all $\theta \in \Theta$ and $s \in S$. In general, for any Borel $\hat{\Theta} \subseteq \Theta$, the map $\beta^{S,\sigma}(\hat{\Theta}|\cdot) : S \to [0, 1]$ is the conditional (on s) expectation of $1_{\hat{\Theta}}$.

Lemma 1. *A signal* (S, σ) *induces a unique random posterior* $p^{S,\sigma} \in \mathcal{R}(\mu)$. *Conversely, given* $p \in \mathcal{R}(\mu)$, *there is a signal on* (Θ, μ) *inducing* p.

One could imagine many different signals that induce a given random posterior, but it is useful to consider the *direct signal* (S_p, σ_p) associated with p. The direct signal is one for which (*i*) every message is a posterior in $\Delta\Theta$ and (*ii*) when the agent hears message 's', his update yields a posterior belief equal to s. That is, the signal tells the agent what his posterior belief should be, and his beliefs abide. The space of messages is $S_p := \mathrm{supp}(p) \subseteq \Delta\Theta$ and, in the finite case, the signal follows law

$$\sigma_p(s|\theta) = \frac{s(\theta)p(s)}{\mu(\theta)}$$

for all messages $s \in S_p$ and states θ.[9]

3   The Psychological Agent

3.1   Information as Risk

Prospective information is a form of risk. Indeed, getting no information entails a deterministic posterior belief: the posterior equals the prior for sure. But getting some information about the state entails random posterior beliefs, because the posterior is then correlated with the realized state, which is initially uncertain. In general, when the signal is not very informative, posterior beliefs will be concentrated around the prior, while when it is informative, posterior beliefs will be more dispersed. In this context where beliefs are payoff relevant, it is appropriate to think of information preferences as a particular instance of risk preferences.

Definition 2. *Given two random posteriors,* $p, q \in \mathcal{R}(\mu)$, *$p$ is more (Blackwell-) informative than q, denoted* $p \succeq_B^\mu q$, *if* (S_q, σ_q) *is a garbling of* (S_p, σ_p), *i.e., if there is a map* $g : S_p \to \Delta(S_q)$ *such that*

$$\sigma_q(\hat{S}|\theta) = \int_{S_p} g(\hat{S}|\cdot)d\sigma_p(\cdot|\theta) \tag{1}$$

for every $\theta \in \Theta$ *and Borel* $\hat{S} \subseteq S_q$.

---

[9]In the general case,

$$\sigma_p(\hat{S}|\theta) = \int_{\hat{S}} \frac{\mathrm{d}s}{\mathrm{d}\mu}(\theta) \, \mathrm{d}p(s) \text{ for every } \theta \in \Theta \text{ and Borel } \hat{S} \subseteq S_p.$$

10

This definition based on posteriors is equivalent to the traditional definition of informativeness based on signals. Usually, we say that a signal is more informative than another if the latter is a garbling of the former (see Blackwell (1953)). But if signal (S, σ) generates random posterior p, then it is easy to verify that (S, σ) and (S_p, σ_p) are Blackwell-equivalent in the usual sense. Therefore, a signal is more informative than another in the traditional sense if and only if it induces a more informative random posterior.

The next proposition formalizes the connection between informativeness and risk.[10]

Proposition 1. *Given two random posteriors $p, q \in \mathcal{R}(\mu)$, the ranking $p \succeq_B^\mu q$ holds if and only if there is a map $r : S_q \to \Delta(S_p) \subseteq \Delta\Delta\Theta$ such that*

1. *For every Borel $S \subseteq \Delta\Theta$, $p(S) = \int_{S_q} r(S|\cdot) \, dq$ and*

2. *For every $t \in S_q$, $r(\cdot|t) \in \mathcal{R}(t)$.*

The proposition may be somewhat counterintuitive. The random posterior p represents better information than q about θ if and only if it is a *mean-preserving spread* of q. That is, greater certainty about the state ex-post entails greater risk concerning one's own beliefs ex-ante.

3.2 Love and Aversion for Information

The agent's welfare depends on his attitude toward information. In particular, we make a distinction between *psychological* and *behavioral* attitudes toward information.

Definition 3. *The agent is psychologically information-loving [resp. -averse, -neutral] if given any action $a \in A$ and any random posteriors $p, q \in \mathcal{R}(\mu)$ with $p \succeq_B^\mu q$,*

$$\int_{\Delta\Theta} u(a, \cdot) \, dp \geq [\text{resp. } \leq, =] \int_{\Delta\Theta} u(a, \cdot) \, dq.$$

An agent is psychologically information-loving [-averse] if he likes [dislikes] informativeness for its own sake, in the hypothetical event that he cannot adapt his decisions to it. That is, information is intrinsically valuable [damaging], abstracting from its instrumental value.[11]

---

[10]For finite-support random posteriors, the result follows from Grant, Kajii, and Polak (1998, Lemma A.1). Ganuza and Penalva (2010, Theorem 2) prove a related result in the special case of $\Theta \subseteq \mathbb{R}$.

[11]Grant, Kajii, and Polak (1998) define what it means to be *single-action information-loving* for an agent who makes no decision. In their model, there is no distinction to be made between psychological and behavioral information preferences.

For every posterior belief $v \in \Delta\Theta$, define the indirect utility associated with v as

$$U(v) = \max_{a \in A} u(a, v). \tag{2}$$

In the next definition, the agent is assumed to respond optimally to information, which defines his behavioral attitude towards it.

Definition 4. *The agent is behaviorally information-loving [resp. -averse, -neutral] if given any random posteriors $p, q \in \mathcal{R}(\mu)$ with $p \succeq_B^\mu q$,*

$$\int_{\Delta\Theta} U \, dp \geq [resp. \leq, =] \int_{\Delta\Theta} U \, dq \tag{3}$$

Consider the following story to see the distinction between psychological and behavioral attitudes toward information. A job market candidate has a job talk, and he may or may not find out five minutes before the presentation whether his slides have embarrassing typos. Knowing that his slides are fine makes the candidate comfortable, and learning that they are not makes him unhappy. Consider now the ex-ante difference between having his laptop and not having it when he finds out the information. If he will not have it, the person might prefer to remain ignorant: the risk of hearing bad news (and being self-conscious during the presentation) outweighs the minor benefit of potentially hearing good news. If he will have his laptop, however, he will have an opportunity to respond to the situation: there is time to fix the typos. Hence, the same person may like information when he can respond to it, and yet prefer to remain ignorant about that same information if he cannot respond to it. If so, this person intrinsically dislikes information (psychologically information-averse), but it is more than compensated by the benefit of being able to *use* it (behaviorally information-loving). A formal example will be presented after the theorem.

Theorem 1. *Psychological and behavioral information preferences are closely linked.*

1. *The agent is psychologically information-loving [resp. -averse, -neutral] if and only if $u(a, \cdot)$ is convex [resp. concave, affine] for every $a \in A$.*

2. *The agent is behaviorally information-loving [resp. -averse, -neutral] if and only if U is convex [resp. concave, affine].*

3. *If the agent is psychologically information-loving, then he is behaviorally information-loving as well.*

4. *If u is concave (and $A \subseteq \mathbb{R}^k$ is convex), then the agent is behaviorally information-averse.*

An immediate consequence of the first two points is that a classical agent is psychologically information-neutral and, in an expression of the easy direction of Blackwell's theorem, behaviorally information-loving, because the linearity of u in v implies that U is convex.

Let us use the theorem to show that for $k \in (0, 1)$ the agent described in Example (b) in Section 2.1 is psychologically information-averse but behaviorally information-loving. Given that $\mathbb{V}(v) = v(1 - v)$ and the expectation is linear, we have $\partial^2 u / \partial v^2 = -2k$. Therefore, the agent is psychologically information-loving if $k < 0$, and psychologically information-averse if $k > 0$. Clearly, the agent's optimal action is $a^*(v) = \mathbb{E}_v[\theta] = v$, and so

$$U(v) = (k - 1)\mathbb{V}(v).$$

The agent is thus behaviorally information-loving if $k < 1$, and behaviorally information-averse if $k > 1$. In particular, for $k \in (0, 1)$, the agent is psychologically information-averse, but behaviorally information-loving. He intrinsically dislikes information, but he unequivocally prefers to have it.

4 Designing Information

Imagine there is a principal who chooses what type of signal the agent will see. We assume throughout that the principal's and the agent's preferences are perfectly aligned, that is, the designer wants to maximize the agent's ex-ante expected welfare.

Definition 5. *A random posterior $p \in \mathcal{R}(\mu)$ is an optimal policy if $\int_{\Delta\Theta} U \, dp \geq \int_{\Delta\Theta} U \, dq$ for all $q \in \mathcal{R}(\mu)$. A signal (S, σ) is optimal if it induces an optimal policy.*

For purposes of characterizing optimal direct signals, the designer need only consider the *design environment* $\langle \Theta, \mu, U \rangle$. As we shall see later, the extra data A and u can be relevant for constructing reasonable indirect signals.

An optimal policy is immediate for certain classes of agents. If the agent is behaviorally information-loving, then the most informative policy is optimal: one should tell him everything. Conversely, if the agent is behaviorally information-averse, then one should tell him nothing. Finally, part 3 of Theorem 1 tells us that fully revealing the state is optimal for psychologically information-loving agents.

Things are not so simple for agents who exhibit psychological information-aversion, because they face a conflict between their innate dislike for information and their instrumental need for it. This realistic and intriguing tradeoff is the subject of the next section.

5   Psychological Information-Aversion

Under psychological information-aversion, disclosure by the principal must balance the intrinsic and the instrumental value of information. This can be accomplished by posterior policies, called direct signals in Section 2.2, by virtue of Lemma 1. Here, we consider another class of information policies, namely recommendation policies, in which each message is simply a recommended action, and ask about their ability to implement this balance. In contrast to the classical world, incentive-compatible recommendation policies are a coarse instrument in a psychological world. They enable a smaller class of (distributions over) outcomes than do posterior policies, due to the failure of the Revelation Principle in psychological settings. Still, recommendation policies are intuitive and may be considerably more practical than posterior policies, and so we would like to know what can be achieved with them. As it turns out, they will remain optimal.

Any information policy admits a corresponding recommendation policy in which the agent is simply told how he would have optimally responded to his posterior. But there is no guarantee a priori that the agent will want to follow this recommendation. The Revelation Principle (as in Myerson (1979), Aumann (1987), and KG) is the result that usually guarantees it.

The Revelation Principle says that, for *any* signal, the corresponding recommendation policy is incentive-compatible. Under psychological information-aversion, and a fortiori in general psychological environments, it is no longer true. For example, take $\Theta = A = \{0, 1\}$ and $u(a, v) = a \left[\mathbb{V}(v) - \frac{1}{9} \right]$, so that the agent is psychologically information-averse. Let $\mu = .6$ and consider a signal that generates random posterior $p = \frac{3}{8}\delta_{.1} + \frac{5}{8}\delta_{.9}$. As shown in Figure 1, whether the posterior is .1 or .9, the agent plays action 0. If, instead, the designer told the agent what he would have played, i.e., always sent the message "play $a = 0$," the agent would receive no information and his posterior would be $\mu = .6$. The agent would then play action 1 and the recommended action, $a = 0$, would not be incentive-compatible.[12]

Given the above, recommendation policies cannot generate all the distributions over actions than do posterior policies. De facto, they represent a "smaller" or more rudimentary class of policies. Even so, they are always optimal for psychologically information-averse agents, our next theorem says. This finding is also interesting because recommendation policies are practical and allow the principal to avoid posterior policies. The latter suppose that the designer sends entire posterior distributions to the agent. In a world with binary states,

---

[12]The Revelation Principle fails here because the set of posterior beliefs for which a given action is optimal may not be convex. In classical environments, if an action is optimal for certain messages, then it is without loss to replace each of the messages with a recommendation of said action. By doing so, we "pool" these messages into one and instead send that pooled message to the agent, thereby also pooling the posteriors attached to these messages. In a psychological world, the action may not remain optimal under the averaged out posteriors.
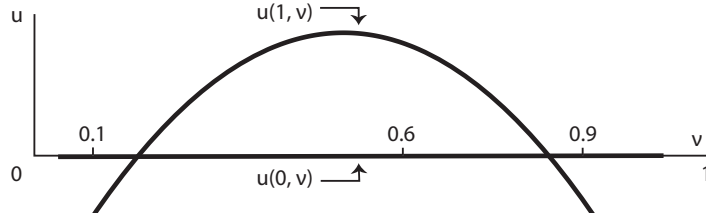
Figure 1: The set $\{v : \quad a = 0 \text{ is optimal at } v\}$ is not convex.

this can be plausible. For example, a doctor who uses a blood test to check for some disease may give a probability that the patient has the illness rather than not. When the state space is larger and more complicated, posterior policies are cumbersome and unrealistic. As it turns out, instead of telling the patient what he should believe, the doctor can simply recommend an action, such as dieting.

Theorem 2. *If the agent is psychologically information-averse, then some incentive compatible recommendation policy is optimal.*

It may be surprising at first that a recommendation policy can serve the needs of all psychologically information-averse agents—even the behaviorally information-loving ones, as in Example (b) from Section 2.1 with $k \in (0, 1)$—because passing from a policy to its associated recommendations can only "lose" information. When full revelation is optimal and the agent is psychologically information-averse, this too can be implemented by a recommendation policy; then, the optimal action necessarily reveals the state. In Example (b), the principal recommends action $a = \theta$ with probability one.

The theorem shows that for any *optimal* signal (S, σ), the corresponding recommendation policy is incentive-compatible. Intuitively, if a policy maximizes the agent's utility, then playing what is recommended should be incentive-compatible. We emphasize that this argument is specific to the benevolent design setting: the result may not hold under psychological preferences if the principal and the agent also have conflicting interests.

In general, this result breaks down for information-loving agents. Consider $A = \{\bar{a}\}$, $\theta \in \{0, 1\}$, and $u(\bar{a}, v) = -\mathbb{V}(v)$. The only recommendation is $\bar{a}$ in all states, which gives the agent no information to update his prior, and so $\mathbb{V}(v) = \mathbb{V}(\mu)$. Clearly, it would be preferable to tell the agent which state has occurred, so that $\mathbb{V}(v) = 0$.

In the case of severe psychological aversion to information, modeled by a concave u, Theorem 2 can be made precise. By part 4 of Theorem 1, a constant recommendation policy, i.e., giving no information, is optimal.

6 Comparative Information-Aversion

The value of information is a leitmotif in information economics. Does an agent benefit from more information? Blackwell (1953) answered the question unambiguously in classical environments. In psychological environments, the universe of possible preferences is so rich that there is no clear answer: different agents have a different affinity for information. Accordingly, the question is a subtler one in a psychological world. Can we compare how informative the optimal policies should be between agents who are ranked in terms of their information-aversion?

This question can have real prescriptive implications, if we find that a well-intentioned expert should provide a less informative policy to a more information-averse agent. In that case, the more information-loving agent is equipped to provide information to the other. For instance, if a doctor considers whether to tell parents the sex of their child, and the father is more information-averse than the mother, the doctor can simply provide optimal information to the mother, who will then be equipped to optimally inform the father. Likewise, if a child experiences only psychological consequences of discovery of some illness, while the parents care about the former *and* about making informed medical decisions, then a pediatrician may simply choose an optimal information policy for the parents, and leave the parents to tell their child (however much they see fit) about the diagnosis.

We compare two individuals via the following behavioral definitions. Let U_1 and U_2 be their indirect utilities, as defined in (2).

Definition 6. *Agent 2 is more (behaviorally) information-averse than agent 1 if for any p and q such that $p \geq_B^\mu q$,*

$$\int U_1 \, dq \geq [resp. \; >] \int U_1 \, dp \implies \int U_2 \, dq \geq [resp. \; >] \int U_2 \, dp.$$

An agent is more information-averse than another if any time the latter would prefer a less informative policy over a more informative one (with the same mean), so would the former agent.

This definition is qualitative rather than quantitative; it compares *when* two decision-makers prefer more/less information rather than the *degree* of such preference. This is quite different from the standard literature on comparative risk-aversion (started by Pratt (1964) and Arrow (1971), including Kihlstrom and Mirman (1974), etc.), in which risk-aversion is measured by comparison to risk-free (i.e., degenerate) lotteries, called certainty equivalents. Usually, an agent is said to be more risk-averse than another if all risky lotteries are worth no more to him than to the other individual, in risk-free terms. As such, certainty equivalents provide an objective "yardstick" to quantify risk-aversion.

In the context of information disclosure, where random posteriors can be seen as lotteries over beliefs, the only available risk-free lottery is the no-information policy. Indeed, all our lotteries must have the same mean (the prior, by Lemma 1), and hence the only risk-free lottery puts probability one on the prior. Without the tool of certainty equivalents, we have no objective yardstick. Of course, we could ask each agent which prior belief would hypothetically make him indifferent to any given policy, but this seems an unrealistic exercise.

An interesting application of the definition is to the mixture between classical and psychological preferences,

$$u_\lambda(a, \nu) = \lambda u_p(\nu) + (1 - \lambda) \int_\Theta u_c(a, \theta) \, d\nu(\theta), \tag{4}$$

where u_p is any purely psychological motive; u_c is any classical motive; and $\lambda \in [0, 1]$ measures the departure from the classical case. The more psychological (4) becomes (i.e., the larger λ), the more information-averse the agent becomes.[13] This is true regardless of the shape of the psychological component: the agent need not be psychologically information-averse.

Our definition of comparative information-aversion only speaks to how two agents rank Blackwell-comparable policies, but Blackwell *in*comparability is generic. Indeed, by Proposition 5 in the Appendix, almost all information policies are pairwise incomparable with respect to (Blackwell) informativeness. Despite being so weak, Definition 6 has strong implications for utility representation and, in two-state environments, this is enough to obtain comparative statics of optimal policies.

Proposition 2.

1. *Suppose agent 2 is more information-averse than agent 1, but 1 is not behaviorally information-loving. Then $U_2 = \gamma U_1 + C$ for some $\gamma \geq 0$ and concave $C : \Delta\Theta \to \mathbb{R}$.*

2. *In a two-state environment, if agent 2 is more information-averse than agent 1, then there exist optimal policies p_1^* and p_2^* such that $p_1^* \succeq_B^\mu p_2^*$.*

The first part of the proposition delivers a strong message: when two agents are ranked according to information-aversion, their indirect utilities differ by the addition of a concave function, which amounts to adding a motive for disliking information everywhere. From the perspective of information disclosure, one would think that a principal should respond to this by providing unambiguously less information to a more information-averse agent. But this is

---

[13]To see why, let U_λ be the indirect utility associated with u_λ. For any $\lambda'' > \lambda'$ and letting $\gamma = \lambda''/\lambda'$, we have $U_{\lambda''} = \gamma U_{\lambda'} + (1 - \gamma)U_0$. Note that $(1 - \gamma)U_0$ is concave—because U_0 is convex by Theorem 1. Notice that adding a concave function to any U yields a more information-averse agent, by Jensen's inequality.

not true in general. In Section 9.5.2 in the Appendix, we give a counterexample in which two agents are ranked by information-aversion, but due to psychological substitution effects, their optimal policies are not Blackwell-comparable. In a world of binary uncertainty, however, the principal should provide a less informative policy to a more information-averse agent. In particular, if p_i^* is uniquely optimal for $i = 1, 2$, then $p_1^* \succeq_B^\mu p_2^*$.

7 Optimal Policies: The General Case

So far, we have focused on optimal policies for special classes of psychological preferences. For information-loving agents, Theorem 1 tells us that full information is optimal. For information-averse agents, Proposition **??** tells us that it is best to say nothing when aversion is severe and, more broadly, Theorem 2 delineates a family of optimal policies for this class, a large step toward characterizing optimal policies, but not a full solution.

In general, there is a large variety of psychological preferences, many of which escape our classification. In this respect, the equivocal case—in which the agent is neither information-averse nor loving—seems especially relevant. For such cases, we need to develop general tools.

7.1 Optimal Welfare

Since any random posterior $p \in \mathcal{R}(\mu)$ can be induced by some signal, designing an optimal signal is equivalent to choosing the best random posterior in $\mathcal{R}(\mu)$. A key result in Kamenica and Gentzkow (2011) is the full characterization of the maximal ex-ante expected utility. This characterization is in terms of the concave envelope of U, defined as

$$\bar{U}(\nu) = \min\{\phi(\nu) : \quad \phi : \Delta\Theta \to \mathbb{R} \text{ affine continuous}, \phi \geq U\}.$$

This result carries over to our environment for all psychological preferences:

Theorem 3. *For any prior $\mu \in \Delta\Theta$, there is a signal that induces indirect utility $\bar{U}(\mu)$ for the agent, and no signal induces a strictly higher value.*

7.2 Optimal Policies

The concave envelope of U describes the optimal value and implicitly encodes optimal policies. Unfortunately, it is a difficult task to characterize the concave envelope of a function in general, even more so to compute it. The main issue is that the concave envelope is a

global construct: evaluating it at one point amounts to solving a *global* optimization problem (Tardella (2008)).

We approach the problem by reducing the support of the optimal policy based on local arguments, a method which will be successful in a class of problems which includes economic applications of interest.

Oftentimes, the designer can deduce from the primitives of the problem that the indirect utility U must be locally convex on various regions of $\Delta\Theta$. In every one of those regions, the agent likes (mean-preserving) spreads in beliefs, which correspond to more informative policies. As a consequence, an optimal policy need not employ beliefs in the inside of those regions, regardless of its general shape.

The central concept of our approach is that of the posterior cover.

Definition 7. *Given a function $f : \Delta\Theta \to \mathbb{R}$, an f-(posterior) cover is a family $\mathcal{C}$ of closed convex subsets of $\Delta\Theta$ such that $\bigcup \mathcal{C}$ is Borel and $f|_C$ is convex for every $C \in \mathcal{C}$.*

A posterior cover is a collection of sets of posterior beliefs over each of which some function is convex. Given a posterior cover $\mathcal{C}$, let

$$out(\mathcal{C}) = \{v \in \Delta\Theta : \quad \forall C \in \mathcal{C}, \text{ if } v \in C \text{ then } v \in ext(C)\}$$

be its set of **outer points**, or extreme points.[14] The next result explains why posterior covers and their outer points play an important role in finding an optimal policy.

Theorem 4. *If $\mathcal{C}$ is a countable U-cover, then there exists an optimal policy p such that $p(out(\mathcal{C})) = 1$.*

The search for an optimal policy can be limited to the outer points of a U-cover. For this reason, we focus on posterior covers of the indirect utility.

Theorem 4 is only useful if three conditions are met. First, given that the indirect utility is a derived object, it is important to tie the U-cover to primitives of the model. Next, computing its outer points ought to be relatively easy, to make the reduction useful in practice. Lastly, the set of outer points must be small, so that (with them in hand) solving for an optimal policy is tractable; in particular, if there are finitely many outer points, finding an optimal policy will be a linear programming problem.

First, we reduce the problem of finding a U-cover to that of finding posterior covers of primitives of the model.

---

[14] A point $v \in ext(C)$ if there is no nontrivial (of positive length) segment for which v is an interior point.

Proposition 3. *1. If C_a is a $u(a, \cdot)$-cover for every $a \in A$, then*

$$C := \bigvee_{a \in A} C_a = \left\{ \bigcap_{a \in A} C_a : \ C_a \in C_a \ \text{for all } a \in A \right\}$$

is a U-cover.

2. If u takes the form

$$u(a, v) = u_p(v) + \int_\Theta u_c(a, \cdot) \, dv \tag{5}$$

and C is a u_p-cover, then C is a U-cover.

The benevolent feature of our design problem enables us to derive a U-cover from primitives of the model. By the first part of the proposition, if we can determine the posterior cover of $u(a, \cdot)$ for each a, then it is easy to name the U-cover. The second part of the proposition presents a class of problems for which the determination of every $u(a, \cdot)$-cover comes down to deriving the posterior cover of a single primitive function. This class of problems consists of all additive preferences between a psychological and a classical component. This includes all preferences in Section 2.1 from (a) to (d). For these problems, optimal design relies on finding the posterior cover of the psychological component alone.

After connecting the problem to posterior covers of primitives, it remains to elicit these posterior covers and find their outer points. Doing so in complete generality is not practical, but we can point to a class of functions for which it can be done.

Proposition 4. *If the state space is finite and f is the pointwise minimum of a finite family of affine functions $\{f_i : \Delta\Theta \to \mathbb{R}\}_{i \in I}$, then $C := \{C_i : i \in I\}$ with*

$$C_i := \{v \in \Delta\Theta : f(v) = f_i(v)\}$$

is an f-cover.

The class of preferences described here, namely the minimum of affine functions, includes many economic situations of interest. For these problems, we can explicitly compute the U-cover. In the additive case alone (5), it is common that $u_p(v) := -\max\{f_i(v)\} = \min\{-f_i(v)\}$ with affine f_i's; this is the case for prior-bias and cognitive dissonance, and for temptation and self-control (Section 8). Then, u_p is piecewise linear; the pieces form a U-cover; and the outer points are easy to find. In Section 9.8 of the Appendix, we supplement this proposition with a hands-on characterization of the outer points.

An auxiliary benefit of the method of posterior covers is that the principal need not solve the agent's optimization problem at every conceivable posterior belief. Indeed, Proposition 3 enables us to compute a posterior cover (and its outer points) without solving the agent problem, and then Theorem 4 lets us derive an optimal policy while computing the indirect utility *only* on those outer points. This simplification—the possibility of which stems from our restriction to aligned interests—makes our lives a lot easier. For instance, we shall see in the prior-bias application that we need only solve the agent's problem at three different posterior beliefs in order to derive an optimal policy.

8 Applications

8.1 Ambiguity-Aversion: Multiplier Preferences

Since Ellsberg's paradox, there has been much evidence for ambiguity aversion, the strict preference for known risks over unknown risks. Maxmin expected utility (Gilboa and Schmeidler (1989)) and multiplier preferences (Hansen and Sargent (2001)) are two of many examples of such preferences.

How should the principal disclose information to an ambiguity-averse agent? In general, the answer does not seem to be obvious: such agents may at times prefer to reject freely available information, as shown by Siniscalchi (2011). For the special case of multiplier preferences, Theorem 1 gives an answer.

Consider an agent with multiplier preferences (case (e) in Section 2.1), who entertains beliefs v' different from his inherited posterior belief v; this doubt and the adversarial character of Nature amount to ambiguity aversion. Because relative entropy $R(v'\|v)$ is jointly convex, one can verify that $u(a, v)$ is convex in v for every a. Theorem 1 then applies: the agent is psychologically information-loving and, therefore, behaviorally information-loving. This implies that the designer should give the agent full information. With no residual uncertainty, the agent faces no negative consequences from ambiguity.

8.2 Prior-Bias and Cognitive Dissonance

The preferences given in Example (c) of Section 2.1 can represent various sources of prior-bias, including cognitive dissonance, stubbornness, reputational concerns, and others. Our methods deliver clear prescriptions for such an agent.

Consider a ruling politician who must make a decision on behalf of her country—e.g., deciding whether or not to go to war—in the face of an uncertain state $\theta \in \{0, 1\}$. She will first seek whatever information she decides and then announce publicly the country's action
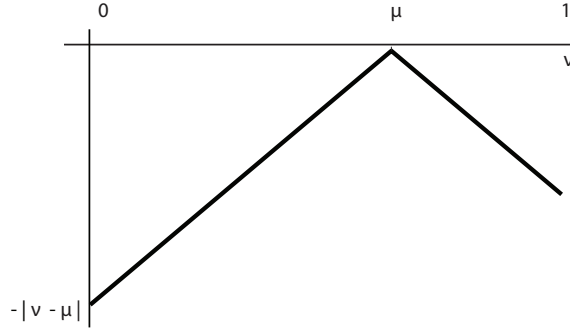
Figure 2: The Psychological Component: Concern for Reputation.

$a \in A$ and (updated) belief v that justifies it. The decision *must* serve the country's interests given whatever information is found.

On one hand, the politician wishes to make decisions that serve the country's motive $u_c(a, \theta)$. On the other, she campaigned on her deeply held beliefs $\mu \in (0, 1)$ and her political career will suffer if she is viewed as a 'flip-flopper'.[15] Publicly expressing a new belief v entails a reputational cost, $u_p(v) = -\rho |v - \mu|$ for $\rho > 0$. When the politician decides which expertise to seek, she behaves as if she has psychological preferences

$$u(a, v) = u_p(v) + \mathbb{E}_{\theta \sim v}[u_c(\theta, a)].$$

Even without any information about u_c, we can infer that the optimal information policy takes a very special form: *the politician should either seek full information or no information at all!* She finds no tradeoff worth balancing.

The psychological component u_p is affine on each member of $C := \{[0, \mu], [\mu, 1]\}$, hence C is a u_p-cover. Appealing to Proposition 3, it is also a U-cover. By Theorem 4, some optimal policy p^* is supported on $out(C) = \{0, \mu, 1\}$. By Bayesian updating, it must be that $p^* = (1 - \lambda)\delta_\mu + \lambda[(1 - \mu)\delta_0 + \mu\delta_1]$ for some $\lambda \in [0, 1]$. That is, the politician is simply tossing a λ-coin, and either seeking full information or seeking none. But in this case, she must be indifferent between the two: either full information or no information must be optimal.

More can be said from our comparative static result. If we consider increasing ρ, then the politician becomes more information-averse—because it corresponds to adding a concave function. Then, by Proposition 2, there is some cutoff $\rho^* \in \mathbb{R}_+$ such that, fixing μ, full information is optimal when $\rho < \rho^*$ and no information is optimal when $\rho > \rho^*$.

---

[15]For example, John Kerry's perceived equivocation on the Iraq war damaged his 2004 presidential campaign (for more details, see the September 19, 2004 and the June 23, 2008 issues of the New York Times).

8.3 Temptation and Self-Control

Individuals often face temptations. Empirical work by psychologists suggests the influence of temptation on consumer decision-making (e.g., Baumeister (2002)). This seems particularly true for consumption-saving decisions, as documented by Huang, Liu, and Zhu (2013). A recent literature in macroeconomics (e.g., Krusell, Kuruscu, and Smith (2010)) has asked how tax policy might be deliberately designed to alleviate the effects of temptation. In this section, we ask how information disclosure policy could do the same.

Consider a consumer who decides whether to invest in a risky asset or to consume (later, the consumer will decide whether to buy an asset, sell an asset, or to consume). The state of the world, $\theta \in \Theta$, which we will interpret below in several ways, is unknown to him, distributed according to prior $\mu \in \Delta\Theta$. Before making his decision, the agent learns additional information from an advisor. Investing draws its value from a higher consumption tomorrow but prevents today's consumption. When the agent deprives himself of consumption, it creates mental suffering.

Our consumer is a standard Gul and Pesendorfer (2001) agent. Given two classical utility functions u_r and u_t, and action set A, the consumer faces a welfare of

$$u(a, v) = \mathbb{E}_{\theta \sim v} \{u_r(a, \theta)\} - \max_{b \in A} \mathbb{E}_{\hat{\theta} \sim v} \left\{u_t(b, \hat{\theta}) - u_t(a, \hat{\theta})\right\}, \tag{6}$$

when he chooses $a \in A$ at posterior belief v. Our consumer has two "sides": a rational side u_r and a tempted side u_t. Given action a, the rational side faces expected value of $\mathbb{E}_{\theta \sim v}\{u_r(a, \theta)\}$, while exerting self-control entails a personal cost of $\max_{b \in A} \{\mathbb{E}_{\theta \sim v}[u_t(b, \theta)] - \mathbb{E}_{\theta \sim v}[u_t(a, \theta)]\}$; the psychological penalty is the value forgone by the tempted side in choosing a. The consumer makes decisions to balance these two motives.

How should a financial advisor (whose interests align with the consumer's sum welfare u) inform the consumer about θ? Giving more information is useful for rational decision making, but it might induce more temptation. We first analyze two extreme cases that bring out the main trade-offs.

If u_t is state-independent, say if θ is the rate of return of the asset, the agent is psychologically information-loving. By Theorem 1, the consumer is behaviorally information-loving too, and so the advisor optimally fully reveals the state.

If u_r is state-independent, say if θ is the weather in a vacation destination, the consumer's welfare is given by $u(a, v) = u_r(a) + \mathbb{E}_{\theta \sim v}[u_t(a, \theta)] - U_t(v)$. As the consumer is psychologically information-averse (by Theorem 1), knowing more can only harm him if his vacation choices are fixed. Yet, the financial advisor may still want to convey some information! We give the intuition in an example but, for the sake of brevity, omit the technical analysis. Say

$A = \Theta = \{0, 1\}$, $u_t(a, \theta) = -(a - \theta)^2$ and $u_r(a) = ka$ for some $k \geq 0$. If the weather in Hawaii is perfect ($\theta = 0$), the consumer is tempted to splurge on a vacation ($a = 0$); otherwise he faces no temptation. To the financially concerned side, weather is irrelevant and investing is rational regardless. It turns out that it is optimal to give information when $\mu < .5$. To see why, suppose $\mu = .4$, so that there is 40% chance of rainy weather. If the advisor said nothing, the agent would invest and face a large cost of self-control, with a 60% chance that he is missing out on perfect weather. Now, the advisor could do better: if the weather is perfect, then with probability $\frac{1}{3}$ she tells her client to take a vacation. With complementary probability, or when the weather in Hawaii is bad, she tells him to invest. This recommendation policy effectively harnesses temptation. When the consumer is told to take a vacation, he goes (and so faces no temptation); and when told to invest, he faces a reduced cost of self-control, since now it is raining in Hawaii with 50% probability.

What about the general case? In general, preferences (6) can be written as

$$u(a, \nu) = u_p(\nu) + \mathbb{E}_{\theta \sim \nu}[u_r(a, \theta) + u_t(a, \theta)],$$

where

$$u_p(\nu) = -U_t(\nu) = \min\{-\mathbb{E}_{\theta \sim \nu}[u_t(b, \theta)] : b \in A\}.$$

By linearity of expectation, u_p is a minimum of affine functions and, thus, the method of posterior covers developed in Section 7.2 is especially useful. Given a u_p-cover C, Theorem 4 and Proposition 3 together tell us that some optimal policy is supported on the outer points of C. Finding such a cover with a small set of outer points is a straightforward exercise. Let us take a concrete example.

Let $\theta \in \{\ell, h\}$ be the exchange rate of \$1 in euros tomorrow, where h (ℓ) stands for 'higher' ('lower') than today. Let $a \in \{a_1, a_2, a_3\}$ where a_i means "invest in firm i" for $i = 1, 2$, and a_3 means "buy a trip to Europe." Firm 1 uses domestic raw material to produce and mostly sells to European customers. Firm 2 imports foreign raw material from Europe and mostly sells domestically.

The consumer's rational and impulsive utilities are given in the following matrices:

	a_1	a_2	a_3
ℓ	α	$-\beta$	0
h	$-\gamma$	δ	0

Function u_r

	a_1	a_2	a_3
ℓ	0	-1	T
h	-1	0	1

Function u_t

The rational side gets no value from consuming, but likes to invest in the firm that benefits from the exchange rate. The impulsive side gets no value from investing in the right firm, but has a disutility from investing in the wrong one. Moreover, it experiences great temptation when the exchange rate is high, and temptation is governed by T when the rate is low.

From Proposition 4 and Corollary 2 (in the Appendix), we know that $C = \{C_1, C_2, C_3\}$ is a posterior cover with $C_i = \{v : \mathbb{E}_{\theta \sim v}[u_t(a_i, \theta)] \leq \mathbb{E}_{\theta \sim v}[u_t(a_j, \theta)] \ \forall j\}$ and that the outer points are contained in[16]

$$
\begin{aligned}
S &= \{0, 1\} \cup \{v : \mathbb{E}_{\theta \sim v}[u_t(a_i, \theta)] = \mathbb{E}_{\theta \sim v}[u_t(a_i, \theta)] \text{ for some } i, j\} \\
&= \left\{0, 1, \tfrac{1}{2}, \tfrac{1+T}{T}, \tfrac{T}{T-2}\right\} \cap [0, 1].
\end{aligned}
$$

From here, an optimal policy comes from solving the straightforward linear program: $\max_{p \in \mathbb{R}^S} \sum_{v \in S} p_v U(v)$ subject to $p \geq 0$, $\sum_{v \in S} p_v = 1$ and $\sum_{v \in S} p_v v = \mu$.

8.4 Equivocal Information Preferences

Consider a patient who is either healthy or ill: $\Theta = \{0, 1\}$, with 0 corresponding to good health (μ and v are probabilities of illness). The patient is psychologically information-averse, but a medical choice can be made to best match the patient's state, as in Example (b) in Sections 2.1 and 3.2,[17]

$$
u(a, v) = k\mathbb{V}(v) - \mathbb{E}_{\theta \sim v}\left[(\theta - a)^2\right]
$$

for some $k \in (0, 1)$. If $A = [0, 1]$, we know the patient is behaviorally information-loving. But what if it is impossible to respond perfectly to posterior beliefs? Suppose that if the patient is likely enough to be ill, the disease is such that there is no appropriate action to take. Then, the inability to respond to information when the news is bad enough makes the agent equivocal: she likes to be informed when the news is good, but when it is bad, she prefers not to know.

We can model this by supposing $A = [0, \bar{a}]$ for some $\bar{a} \in (0, 1)$. Our intuition might tell us that (i) when $\mu < \bar{a}$, the doctor should give some information, since it can be useful (and such use outweighs the psychological discomfort when $k < 1$), and (ii) when $\mu > \bar{a}$, the doctor should say nothing, since there is no instrumental benefit to learning anything. The first intuition is correct, but the second leads us astray.

To solve for the optimal policy, our first tool is Theorem 2: some recommendation

---

[16] If one of the fractions has zero denominator, just omit it from S.

[17] The given $u(v, a)$ is not decreasing in beliefs, which does not match the 'illness' story. Of course, we could always subtract a constant multiple of v to make it decreasing; this would have no design implications.
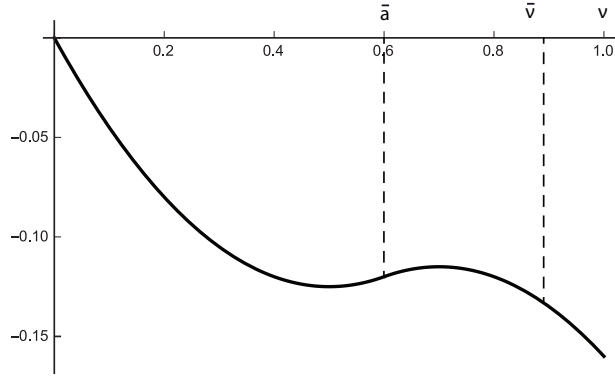
Figure 3: Equivocal Agent's Indirect Utility U when $k = 1/2$ and $\bar{a} = 3/5$.

policy is optimal, because the agent is psychologically information-averse. If she could, the agent would choose action $E_v[\theta] = v$, but given the constraint, the optimal action is $a^*(v) = \min\{v, \bar{a}\}$. That is, the agent wants her action to match her belief as well as possible. For the recommendation to be incentive-compatible, the patient's posterior belief upon hearing recommendation $a \in [0, \bar{a}]$ must be

$$\begin{cases} v = a & \text{if } a < \bar{a} \\ v \in [\bar{a}, 1] & \text{if } a = \bar{a}. \end{cases}$$

Moreover, the doctor would never use a policy that recommends actions $a \in (0, \bar{a})$, because it would always be better to give a little more information—run an additional test that will either give slightly good news or slightly bad news. This is because U is strictly convex at $a \in (0, \bar{a})$.[18] The doctor optimally recommends the patient to 'do nothing' ($a = 0$) or to take an aggressive course of action ($a = 1$; e.g., never smoke again).

For the 'nothing' option to be incentive-compatible, she must *only* recommend it when the patient is healthy. Said differently, when the patient is ill, the doctor must always recommend the aggressive action. This leaves one remaining parameter: the probability λ of making the aggressive recommendation when the patient is healthy. The value of such a policy would be (assuming $\bar{a}$ is incentive-compatible)

$$(1 - \mu)(1 - \lambda)u(0, 0) + [\mu + (1 - \mu)\lambda]u\left(\bar{a}, \frac{\mu}{\mu + (1 - \mu)\lambda}\right) = 0 + \frac{\mu}{v}u(\bar{a}, v),$$

---

[18]We could describe the patient as *locally* behaviorally information-loving at belief $a \in (0, \bar{a})$.

where $v = \frac{\mu}{\mu+(1-\mu)\lambda} \in [\mu, 1]$. Some algebra shows

$$\frac{\partial}{\partial v}\left[\frac{\mu}{v}u(\bar{a}, v)\right] = \frac{\mu}{v^2}(\bar{a}^2 - kv^2),$$

so that the patient's expected value of $\frac{\mu}{v}u(\bar{a}, v)$ is maximized at $v^* = \min\left\{1, \frac{\bar{a}}{\sqrt{k}}\right\}$. Thus, λ must be chosen so that $v = v^*$. Since $v^* > \bar{a}$, the described recommendation policy is indeed incentive-compatible.

The above work shows that the following policy is optimal:

- If the patient is healthy, the doctor suggests no treatment with probability $\frac{v^*-\mu}{v^*(1-\mu)}$. With the complementary probability, she recommends the aggressive action.

- If he is ill, she recommends the aggressive action.

Notice that, even for some beliefs $v > \bar{a}$, at which the patient is (locally) behaviorally-information-averse, the doctor optimally provides some information. This is because the potential to provide very good news with some probability—a clean bill of health, to which the patient responds with $a = 0$—outweighs the damage of news to which she cannot respond. Looking at v^*, this is particularly striking when $k \leq \bar{a}^2$, because in this case, the doctor optimally provides full information, even though the patient is not behaviorally information-loving. That is, full information is optimal, but poorly chosen partial information could be worse than saying nothing at all.

9   Appendix

9.1   Proof of Proposition 1

In both directions, the idea of the proof it to choose g (given r) or r (given g) to ensure

$$\mathrm{d}g(t|s)\,\mathrm{d}\sigma_p(s|\theta) = \mathrm{d}r(s|t)\,\mathrm{d}\sigma_q(t|\theta).$$

Everything else is just bookkeeping.

Proof. First suppose such r exists, and define $g : S_p \to \Delta(S_q)$ via

$$g(T|s) := \int_T \frac{\mathrm{d}r(\cdot|t)}{\mathrm{d}p}(s)\,\mathrm{d}q(t),$$

for any $s \in S_p$ and Borel $T \subseteq S_q$. Then for $\theta \in \Theta$ and Borel $T \subseteq S_q$, it is straightforward to verify that $\int_{S_p} g(T|s)\,\mathrm{d}\sigma_p(s|\theta) = \sigma(T|\theta)$, so that g witnesses $p \geq_B^\mu q$.

Conversely, suppose $p \geq_B^\mu$ with the map $g : S_p \to \Delta(S_q)$ as in the definition of $\geq_B^\mu$. Let

$$r(S|t) := \int_S \frac{\mathrm{d}g(\cdot|s)}{\mathrm{d}q}(t)\,\mathrm{d}p(s)$$

for $t \in S_q$ and Borel $S \subseteq S_p$. Then for Borel $S \subseteq S_p$, it is again straightforward to verify that $\int_{S_q} r(S|\cdot)\,\mathrm{d}q = p(S)$, which verifies the first condition.

Now, given Borel $T \subseteq S_q$ and $\theta \in \Theta$, one can check that

$$\int_T \frac{\mathrm{d}t}{\mathrm{d}\mu}(\theta)\,\mathrm{d}q(t) = \int_T \int_{\Delta\Theta} \frac{\mathrm{d}s}{\mathrm{d}\mu}(\theta)\,\mathrm{d}r(s|t)\,\mathrm{d}q(t),$$

so that $\frac{\mathrm{d}t}{\mathrm{d}\mu}(\theta) = \int_{\Delta\Theta} \frac{\mathrm{d}s}{\mathrm{d}\mu}(\theta)\,\mathrm{d}r(s|t)$ for a.e.-$\mu(\theta)q(t)$. From this we may show that, for any Borel $\hat{\Theta} \subseteq \Theta$,

$$\int_{\Delta\Theta} s(\hat{\Theta})\,\mathrm{d}r(s|t) = t(\hat{\Theta}),$$

which verifies the second condition. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

9.2   Proof of Theorem 1

Lemma 2. *The set $M := \{\gamma \in \Delta\Theta : \exists \epsilon > 0 \text{ s.t. } \epsilon\gamma \leq \mu\}$ is dense[19] in $\Delta\Theta$.*

---

[19]Under the w^*-topology.

Proof. First, notice that $M = \left\{ \gamma \in \Delta\Theta : \ \gamma << \mu \text{ and } \dfrac{\mathrm{d}\gamma}{\mathrm{d}\mu} \text{ is (essentially) bounded} \right\}$ is convex and extreme (i.e. a face of $\Delta\Theta$). Thus its w^*-closure $\bar{M}$ is closed (and so compact, by Banach-Alaoglu), convex, and extreme. Now, let E be the set of extreme points of $\bar{M}$. Because $\bar{M}$ is extreme, E is a subset of $ext(\Delta\Theta) = \{\delta_\theta\}_{\theta\in\Theta}$. So $E = \{\delta_\theta\}_{\theta\in\hat\Theta}$ for some $\hat\Theta \subseteq \Theta$. By Krein-Milman, $\bar{M} = \overline{co}E = \Delta(\Theta')$, where Θ' is the closure of $\hat\Theta$. Finally, notice that $\mu \in M$ implies $\Theta' \supseteq supp(\mu) = \Theta$. Thus $\bar{M} = \Delta\Theta$ as desired. $\qquad\square$

Lemma 3. *Fix a continuous function $f : \Delta\Theta \to \mathbb{R}$. Then the following are equivalent (given μ is of full support):*

1. *For all $v \in \Delta\Theta$ and $p \in \mathcal{R}(v)$, we have $\int_{\Delta\Theta} f \, dp \geq f(v)$*

2. *For all $\mu' \in \Delta\Theta$ and $p, q \in \mathcal{R}(\mu')$ with $p \succeq_B^{\mu'} q$, we have $\int_{\Delta\Theta} f \, dp \geq \int_{\Delta\Theta} f \, dq$.*

3. *For all $p, q \in \mathcal{R}(\mu)$ with $p \succeq_B^{\mu} q$, we have $\int_{\Delta\Theta} f \, dp \geq \int_{\Delta\Theta} f \, dq$.*

4. *f is convex.*

Proof. Suppose (1) holds, and consider any $\mu' \in \Delta\Theta$ and $q \in \mathcal{R}(\mu')$. If $r : S_q \to \Delta\Delta\Theta$ satisfies $r(\cdot|v) \in \mathcal{R}(v)$ for every $v \in S_q$, (1) implies $\int_{S_q} \int_{\Delta\Theta} f \, dr(\cdot|v) \, dq(v) \geq \int_{S_q} f \, dq$. Equivalently (by Proposition 1), any p more informative than q has $\int f \, dp \geq \int f \, dq$, which yields (2).

That (2) implies (3) is immediate.

Now, suppose (4) fails. That is, there exist $\gamma, \zeta, \eta \in \Delta\Theta$ and $\lambda \in (0, 1)$ such that

$$(1 - \lambda)\zeta + \lambda\eta \ = \ (1 - \lambda)\zeta + \lambda\eta.$$
$$f((1 - \lambda)\zeta + \lambda\eta) \ < \ (1 - \lambda)f(\zeta) + \lambda f(\eta).$$

Now, we want to exploit the above to construct two information-ranked random posteriors such that f has higher expectation on the less informative of the two.

To start, let us show how to do it if $\epsilon\gamma \leq \mu$ for some $\epsilon \in (0, 1)$. In this case, let

$$v := \frac{1}{1 - \epsilon}(\mu - \epsilon\gamma) \in \Delta\Theta, \quad p := (1-\epsilon)\delta_v + \epsilon(1-\lambda)\delta_\zeta + \epsilon\lambda\delta_\eta \in \mathcal{R}(\mu), \text{ and } q := (1-\epsilon)\delta_v + \epsilon\delta_\gamma \in \mathcal{R}(\mu).$$

Then $p \succeq_B^{\mu} q$, but

$$\int_{\Delta\Theta} f \, dp - \int_{\Delta\Theta} f \, dq = \epsilon\left[(1 - \lambda)f(\zeta) + \lambda f(\eta) - f(\gamma)\right] < 0,$$

as desired.

Now, notice that we need not assume that $\epsilon\gamma \leq \mu$ for some $\epsilon \in (0, 1)$. Indeed, in light of Lemma 2 and continuity of f, we can always pick $\gamma, \zeta, \eta \in \Delta\Theta$ to ensure it. So given continuous nonconvex f, we can ensure existence of $p \geq^{\mu}_B q$ with $\int_{\Delta\Theta} f \, dp < \int_{\Delta\Theta} f \, dq$. That is, (3) fails too.

Finally, notice that (4) implies (1) by Jensen's inequality. $\square$

Proof of Theorem 1.

1. This follows immediately from applying Lemma 3 to $u(a, \cdot)$ and $-u(a, \cdot)$ for each $a \in A$.

2. This follows immediately from applying Lemma 3 to U and $-U$.

3. This follows from the first two parts, and from the easy fact that a pointwise maximum of convex functions is convex.

4. This follows from part 2. and the fact that U is concave by the (convex) maximum theorem.

$\square$

9.3   Proof of Theorem 2

Proof. Suppose the agent is psychologically information-averse.
Fix some measurable selection[20] $a^* : \Delta\Theta \to A$ of the best-response correspondence $v \mapsto \arg\max_{a \in A} u(a, v)$. In particular, given any $q \in \mathcal{R}(\mu)$, $a^*|_{S_q}$ is an optimal strategy for an agent with direct signal (S_q, σ_q).

Toward a proof of the theorem, we first verify the following claim.

Claim: Given any random posterior $p \in \mathcal{R}(\mu)$, we can construct a signal (A, α_p) such that:

1. The random posterior q_p induced by (A, α_p) is less informative than p.

2. An agent who follows the recommendations of α_p performs at least as well as an agent who receives signal (S_p, σ_p) and responds optimally, i.e.

$$\int_\Theta \int_A u\left(a, \beta^{A,\alpha_p}(\cdot|a)\right) \, d\alpha_p(a|\theta) \, d\mu(\theta) \geq \int_{\Delta\Theta} U \, dp.$$

[20]One exists, by Aliprantis and Border (1999, Theorem 8.13).

To verify the claim, fix any $p \in \mathcal{R}(\mu)$, and define the map

$$
\begin{aligned}
\alpha_p : \Theta \quad &\to \quad \Delta(A) \\
\theta \quad &\mapsto \quad \alpha_p(\cdot|\theta) = \sigma_p(\cdot|\theta) \circ a^{*-1}.
\end{aligned}
$$

Then (A, α_p) is a signal with

$$
\alpha_p(\hat{A}|\theta) = \sigma_p\left(\left\{s \in S_p : \; a^*(s) \in \hat{A}\right\}\Big|\theta\right)
$$

for every $\theta \in \Theta$ and Borel $\hat{A} \subseteq A$. The signal (A, α_p) is familiar: replace each message in (S_p, σ_p) with a recommendation of the action that would have been taken.

Let $q_p \in \mathcal{R}(\mu)$ denote the random posterior induced by signal (A, α_p). Now, let us show that q_p delivers at least as high an expected value as p.

By construction,[21] $p \geq_B^\mu q$. Therefore, by Proposition 1, there is a map $r : S_q \to \Delta(S_p)$ such that for every Borel $S \subseteq \Delta\Theta$, $p(S) = \int_{S_q} r(S|\cdot) \, dq$, and for every $t \in S_q$, $r(\cdot|t) \in \mathcal{R}(t)$. Then, appealing to the definition of psychological information-aversion,

$$
\begin{aligned}
\int_A u\left(a, \beta^{A,\alpha_p}(\cdot|a)\right) \, d\alpha_p(a|\theta) &= \int_{S_p} u\left(a^*(s), \beta^{A,\alpha_p}(\cdot|a^*(s))\right) \, d\sigma_p(s|\theta) \\
&\geq \int_{S_p} \int_{S_p} u\left(a^*(s), v\right) \, dr\left(v\Big|\beta^{A,\alpha_p}(\cdot|a^*(s))\right) \, d\sigma_p(s|\theta) \\
&= \int_{S_p} \int_{S_p} U(v) \, dr\left(v\Big|\beta^{A,\alpha_p}(\cdot|a^*(s))\right) \, d\sigma_p(s|\theta) \\
&= \int_A \int_{S_p} U(v) \, dr\left(v\Big|\beta^{A,\alpha_p}(\cdot|a)\right) \, d\alpha_p(a|\theta).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\int_\Theta \int_A u\left(a, \beta^{A,\alpha_p}(\cdot|a)\right) \, d\alpha_p(a|\theta) \, d\mu(\theta) &\geq \int_\Theta \int_A \int_{S_p} U(v) \, dr\left(v\Big|\beta^{A,\alpha_p}(\cdot|a)\right) \, d\alpha_p(a|\theta) \, d\mu(\theta) \\
&= \int_{S_q} U(v) \, dr(v|t) \, dq(t) \\
&= \int_{S_p} U \, dp,
\end{aligned}
$$

---

[21] Indeed we can define g in (1) via: $g(a|s) = 1$ if $a^*(s) = a$ and 0 otherwise.

which verifies the claim.

Now, fix some optimal policy $p^* \in \mathcal{R}(\mu)$, and let $\alpha = \alpha_{p^*}$ and $q = q_{p^*}$ be as delivered by the above claim. Let the measure $Q \in \Delta(A \times \Delta\Theta)$ over recommended actions and posterior beliefs be that induced by α_p. So

$$Q(\hat{A} \times \hat{S}) = \int_\Theta \int_{\hat{A}} \mathbf{1}_{\beta^{A,\alpha}(\cdot|a) \in S} \ d\alpha(a|\theta) \ d\mu(\theta)$$

for Borel $\hat{A} \subseteq A, \hat{S} \subseteq \Delta\Theta$.

Then,[22]

$$\int_{\Delta\Theta} U \ dp \leq \int_{A \times \Delta\Theta} u \ dQ \leq \int_{\Delta\Theta} U \ dq \leq \int_{\Delta\Theta} U \ dp,$$

so that:

$$\int_{\Delta\Theta} U \ dq = \int_{\Delta\Theta} U \ dp, \text{ i.e. } q \text{ is optimal; and}$$

$$\int_{A \times \Delta\Theta} u(a, \nu) \ dQ(a, \nu) = \int_{\Delta\Theta} U \ dq = \int_{A \times \Delta\Theta} \max_{\tilde{a} \in A} u(\tilde{a}, \nu) \ dQ(a, \nu).$$

The latter point implies that $a \in \operatorname{argmax}_{\tilde{a} \in A} u(\tilde{a}, \nu)$ a.s.-$Q(a, \nu)$. In other words, the recommendation (A, α) is incentive-compatible as well. This completes the proof. □

We note that the claim in the above proof delivers something more than the result of Theorem 2. Indeed, given any finite-support random posterior p, the claim produces a constructive procedure to design an incentive-compatible recommendation policy which outperforms p. The reason is that (in the notation of the claim):

1. If $a^*|_{S_p}$ is injective, then $q_p = p$, so that (A, α_p) is an incentive-compatible recommendation policy inducing random posterior p itself.

2. Otherwise, $|S_{q_p}| < |S_p|$.

In the latter case, we can simply apply the claim to q_p. Iterating in this way—yielding a new, better policy at each stage—eventually (in fewer than $|S_p|$ stages) leads to a recommendation policy which is incentive-compatible and outperforms p.

---

[22]Indeed, the inequalities follow from the above claim, the definition of U along with the property $marg_{\Delta\Theta}Q = q$, and optimality of p, respectively.

9.4 Proof of Theorem 3

Below, we will prove the slightly stronger result. Fix any $\bar{p} \in \mathcal{R}(\mu)$, and let $S := \overline{co}(S_{\bar{p}})$ and $\Phi := \{\phi : S \to \mathbb{R} : \phi \text{ is affine continuous and } \phi \geq U|_S\}$. We will show that

$$\max_{p \in \Delta(S) \cap \mathcal{R}(\mu)} \int U \, \mathrm{d}p = cav_{\bar{p}} U(\mu),$$

where $cav_{\bar{p}} U(\nu) := \inf_{\phi \in \Phi} \phi(\nu)$.

Proof. Let $\beta : \Delta(S) \to S$ be the unique map such that $\mathbb{E}_{\nu \sim p} \nu = \beta(\nu)$. Such β is well-defined, continuous, affine, and surjective, as shown in Phelps (2001).

For every $p \in \Delta(S)$, define $\mathbb{E}U(p) = \int U \, \mathrm{d}p$. The map $\mathbb{E}U$ is then affine and continuous. For every $\nu \in S$, define $U^*(\nu) = \max_{p \in \Delta(S), \beta(p) = \nu} \mathbb{E}U(p)$, which (by Berge's theorem) is well-defined and upper-semicontinuous.

For any $\nu \in S$:

- $U^*(\nu) \geq \mathbb{E}U(\delta_\nu) = U(\nu)$. That is, an optimal policy for prior ν does at least as well as giving no information.

- For all $p \in \Delta(S)$ and all affine continuous $\phi : S \to \mathbb{R}$ with $\phi \geq U|_S$

$$\mathbb{E}U(p) = \int U \, \mathrm{d}p \leq \int \phi(s) \, \mathrm{d}p(s) = \phi\left(\int_{\Delta\Theta} s \, \mathrm{d}p(s)\right) = \phi(\beta(p)).$$

So, $U^*(\nu)$ can be no higher than $cav_{\bar{p}} U(\nu)$.

Moreover, if $\nu, \nu' \in S$ and $\lambda \in [0, 1]$, then for all $p, q \in \Delta(S)$ with $\beta(p) = \nu$ and $\beta(q) = \nu'$, we know $\beta((1 - \lambda)p + \lambda q) = (1 - \lambda)\nu + \lambda\nu'$, so that

$$U^*((1 - \lambda)\nu + \lambda\nu') \geq \mathbb{E}U((1 - \lambda)p + \lambda q) = (1 - \lambda)\mathbb{E}U(p) + \lambda\mathbb{E}U(q).$$

Optimizing over the right-hand side yields

$$U^*((1 - \lambda)\nu + \lambda\nu') \geq (1 - \lambda)U^*(\nu) + \lambda U^*(\nu').$$

That is, an optimal policy for prior $(1 - \lambda)\nu + \lambda\nu'$ does at least as well as a signal inducing (interim) posteriors from $\{\nu, \nu'\}$ followed by an optimal signal for the induced interim belief.

So far, we have established that U^* is upper-semicontinuous and concave, and $U|_S \leq U^* \leq cav_{\bar{p}} U$. Since $cav_{\bar{p}} U$ is the pointwise-lowest u.s.c. concave function above $U|_S$, it must be that $cav_{\bar{p}} U \leq U^*$, and thus $U^* = cav_{\bar{p}} U$.

$\square$

9.5 Comparative Statics: Proof of Proposition 2 and Example

9.5.1 Proof

Lemma 4. *Given $p, q \in \Delta\Delta\Theta$, the following are equivalent:*

1. *There is some $v \in \Delta\Theta$ such that $p, q \in \mathcal{R}(v)$ and $p \geq_B^v q$.*

2. *For every convex continuous $f : \Delta\Theta \to \mathbb{R}$, we have $\int_{\Delta\Theta} f \, dp \geq \int_{\Delta\Theta} f \, dq$.*

Proof. That (1) implies (2) follows from Lemma 3. Now, suppose (2) holds. The Theorem of Hardy-Littlewood-Polya-Blackwell-Stein-Sherman-Cartier reported in Phelps (2001, main Theorem, Section 15) then says that p is a mean-preserving spread of q. Then, Proposition 1 implies (1). □

Notation 1. *Given any $\mathcal{F} \cup \mathcal{G} \cup \{h\} \subseteq C(\Delta\Theta)$:*

- *Let $\succsim_{\mathcal{F}}$ be the (reflexive, transitive, continuous) binary relation on $\Delta\Delta\Theta$ given by*

$$p \succsim_{\mathcal{F}} q \iff \int_{\Delta\Theta} f \, dp \geq \int_{\Delta\Theta} f \, dq \ \forall f \in \mathcal{F}.$$

- *Let $\langle \mathcal{F} \rangle \subseteq C(\Delta\Theta)$ be the smallest closed convex cone in $C(\Delta\Theta)$ which contains $\mathcal{F}$ and all constant functions.*

- *Let $\mathcal{F} + \mathcal{G}$ denote the Minkowski sum $\{f + g : f \in \mathcal{F}, g \in \mathcal{G}\}$, which is a convex cone if $\mathcal{F}, \mathcal{G}$ are.*

- *Let $\mathbb{R}_+ h$ denote the ray $\{\alpha h : \alpha \in \mathbb{R}, \alpha \geq 0\}$.*

We now import the following representation uniqueness theorem to our setting.

Lemma 5 (Dubra, Maccheroni, and Ok (2004), Uniqueness Theorem, p. 124)**.** *Given any $\mathcal{F}, \mathcal{G} \subseteq C(\Delta\Theta)$,*

$$\succsim_{\mathcal{F}} = \succsim_{\mathcal{G}} \iff \langle \mathcal{F} \rangle = \langle \mathcal{G} \rangle.$$

As a consequence, we get the following:

Corollary 1. *Suppose $\mathcal{F} \cup \{g\} \subseteq C(\Delta\Theta)$. Then*

$$\succsim_{\mathcal{F}} \subseteq \succsim_{\{g\}} \ \text{if and only if } g \in \langle \mathcal{F} \rangle.$$

Proof. If $g \in \langle \mathcal{F} \rangle$, then $\succsim_{\{g\}} \supseteq \succsim_{\langle \mathcal{F} \rangle}$, which is equal to $\succsim_{\mathcal{F}}$ by the uniqueness theorem. If $\succsim_{\mathcal{F}} \subseteq \succsim_{\{g\}}$, then $\succsim_{\mathcal{F}} \subseteq \succsim_{\mathcal{F}} \cap \succsim_{\{g\}} = \succsim_{\mathcal{F} \cup \{g\}} \subseteq \succsim_{\mathcal{F}}$. Therefore, $\succsim_{\mathcal{F} \cup \{g\}} = \succsim_{\mathcal{F}}$. By the uniqueness theorem, $\langle \mathcal{F} \cup \{g\} \rangle = \langle \mathcal{F} \rangle$, so that $g \in \langle \mathcal{F} \rangle$. □

Lemma 6. *Suppose $\mathcal{F} \subseteq C(\Delta\Theta)$ is a closed convex cone that contains the constants and $g \in C(\Delta\Theta)$. Then either $g \in -\mathcal{F}$ or $\langle \mathcal{F} \cup \{g\}\rangle = \mathcal{F} + \mathbb{R}_+ g$.*

Proof. Suppose $\mathcal{F}$ is as described and $-g \notin \mathcal{F}$.

Since a closed convex cone must be closed under sums and nonnegative scaling, it must be that $\langle \mathcal{F} \cup \{g\}\rangle \supseteq \mathcal{F} + \mathbb{R}_+ g$. Therefore, $\langle \mathcal{F} \cup \{g\}\rangle = \langle \mathcal{F} + \mathbb{R}_+ g\rangle$, which leaves us to show $\langle \mathcal{F} + \mathbb{R}_+ g\rangle = \mathcal{F} + \mathbb{R}_+ g$. As $\mathcal{F} + \mathbb{R}_+ g$ is a convex cone which contains the constants, we need only show it is closed.

To show it, consider sequences $\{\alpha_n\}_{n=1}^\infty \subseteq \mathbb{R}_+$ and $\{f_n\}_{n=1}^\infty \subseteq \mathcal{F}$ such that $(f_n + \alpha_n g)_{n=1}^\infty$ converges to some $h \in C(\Delta\Theta)$. We wish to show that $h \in \mathcal{F} + \mathbb{R}_+ g$. By dropping to a subsequence, we may assume $(\alpha_n)_n$ converges to some $\alpha \in \mathbb{R}_+ \cup \{\infty\}$.

If α were equal to ∞, then the sequence $\left(\frac{f_n}{\alpha_n}\right)_n$ from $\mathcal{F}$ would converge to $-g$, implying (since $\mathcal{F}$ is closed) $-g \in \mathcal{F}$. Thus α is finite, so that $(f_n)_n$ converges to $f = h - \alpha g$. Then, since $\mathcal{F}$ is closed, $h = f + \alpha g \in \mathcal{F} + \mathbb{R}_+ g$. $\qquad\square$

Proof of Proposition 2, Part (1)[23]

Proof. In light of Lemma 4, that 2 is more information-averse than 1 is equivalent to the pair of conditions:

$$\succsim_{C\cup\{U_2\}} \quad \subseteq \quad \succsim_{\{U_1\}}$$
$$\succsim_{(-C)\cup\{U_1\}} \quad \subseteq \quad \succsim_{\{U_2\}},$$

where C is the set of convex continuous functions on $\Delta\Theta$; notice that C is a closed convex cone which contains the constants.

Then, applying Corollary 1 and Lemma 6 twice tells us:

- Either U_2 is concave or $U_1 \in C + \mathbb{R}_+ U_2$.

- Either U_1 is convex or $U_2 \in -C + \mathbb{R}_+ U_1$.

By hypothesis,[24] 1 is not behaviorally information-loving, i.e. U_1 is not convex. Then $U_2 \in -C + \mathbb{R}_+ U_1$, proving the proposition. $\qquad\square$

Notice: We cannot strengthen the above theorem to ensure $\gamma > 0$. Indeed, consider the $\Theta = \{0, 1\}$ world with $U_1 = H$ (entropy) and $U_2 = \mathbb{V}$ (variance). Both are strictly concave, so that in particular U_2 is more information-averse than U_1. However, any sum of a concave function and a strictly positive multiple of U_1 is non-Lipschitz (as $H'(0) = \infty$), and so is not a multiple of U_2.

---

[23]Part (2) is proven in Subsection 9.7.
[24]Note: this is the first place we've used this hypothesis

9.5.2 Example

Suppose that $\theta = (\theta_1, \theta_2) \in \{0, 1\}^2$ is distributed uniformly. Let $x_1(v)$ and $x_2(v) \in [-1, 1]$ be some well-behaved measures of information about θ_1, θ_2 respectively (the smaller x_i, the more dispersed the ith dimension of distribution v).[25] Then define U_1, U_2 via

$$U_1(v) := -\frac{1}{2}\left(x_1(v)^2 + x_2(v)^2 + x_1(v)x_2(v)\right);$$

$$U_2(v) := U_1(v) - \frac{1}{5}x_1(v).$$

Agent 2 receives extra utility from being less informed and so he is more information-averse than 1. Meanwhile, we can easily verify (via first-order conditions to optimize (x_1, x_2)) that any optimal policy for 1 will give more information concerning θ_1, but less information concerning θ_2, than will any optimal policy for 2. Therefore, the optimal policies for 1 and 2 are Blackwell incomparable, although 2 is more information-averse than 1. Intuitively, the penalty for information about θ_1 induces a substitution effect toward information about θ_2. Technically, the issue is a lack of supermodularity.

9.6   Proposition 5: Most policies are not comparable

Proposition 5. *Blackwell-incomparability is generic, i.e. the set $N_\mu = \{(p, q) \in \mathcal{R}(\mu)^2 : p \not\succeq_B^\mu q \text{ and } p \not\preceq_B^\mu q\}$ is open and dense in $\mathcal{R}(\mu)^2$.*

Proof. To show that N_μ is open and dense, it suffices to show that $\not\succeq_B^\mu$ is. Indeed, it would then be immediate that $\not\preceq_B^\mu$ is open and dense (as switching coordinates is a homeomorphism), so that their intersection N_μ is dense too.

Given Lemma 1, it is straightforward[26] to express $\succeq_B^\mu \subseteq (\Delta\Theta)^2$ as the image of a continuous map with domain $\Delta\Delta\Delta\Theta$. Therefore $\succeq_B^\mu$ is compact, making its complement an open set.

Take any $p, q \in \mathcal{R}(\mu)$. We will show existence of $\{p_\epsilon, q_\epsilon\}_{\epsilon \in (0,1)}$ such that $p_\epsilon \not\succeq_B^\mu q_\epsilon$ and $(p_\epsilon, q_\epsilon) \to (p, q)$ as $\epsilon \to 0$. For each fixed $\epsilon \in (0, 1)$, define

$$g_\epsilon : \Delta\Theta \quad \to \quad \Delta\Theta$$
$$v \quad \longmapsto \quad (1 - \epsilon)v + \epsilon\mu.$$

---

[25]For instance, let $x_i(v) := 1 - 8\mathbb{V}(\mathrm{marg}_i v)$, $i = 1, 2$. What we need is that $x_1(v), x_2(v)$ be convex $[-1, 1]$-valued functions of $\mathrm{marg}_1 v, \mathrm{marg}_2 v$ respectively, taking the priors to -1 and atomistic beliefs to 1.

[26]Indeed, let $\beta_1 : \Delta\Delta\Theta \to \Delta\Theta$ and $\beta_2 : \Delta\Delta\Delta\Theta \to \Delta\Delta\Theta$ be the maps taking each measure to its barycenter. Then define $B : \Delta\Delta\Delta\Theta \to (\Delta\Delta\Theta)^2$ via $B(Q) = (Q \circ \beta_1^{-1}, \beta_2(Q))$. Then $\succeq_B^\mu = B((\beta_1 \circ \beta_2)^{-1}(\mu))$.

Because g_ϵ is continuous and affine, its range G_ϵ is compact and convex. Define

$$p_\epsilon := p \circ g_\epsilon^{-1} \text{ and } q_\epsilon := (1 - \epsilon)q + \epsilon f,$$

where $f \in \mathcal{R}(\mu)$ is the random posterior associated with full information.[27] It is immediate, say by direct computation with the Prokhorov metric and boundedness of Θ, that $(p_\epsilon, q_\epsilon) \to (p, q)$ as $\epsilon \to 0$. Moreover,

$$\overline{co}(S_{p_\epsilon}) \subseteq \overline{co}(G_\epsilon) = G_\epsilon \subsetneq \Delta\Theta = \overline{co}(S_{q_\epsilon}).$$

In particular, $\overline{co}(S_{p_\epsilon}) \not\supseteq \overline{co}(S_{p_\epsilon})$. Then, appealing to Lemma 1, $p_\epsilon \not\succeq_B^\mu q_\epsilon$ as desired. $\qquad\square$

9.7   Proof of Proposition 2, Part (2)

Proof. First, notice that we are done if either U_1 is information-loving (in which case p_1^* can just be full information) or U_2 is information-averse (in which case p_2^* can just be no information). Thus focus on the complementary case, where Proposition 2 guarantees $\gamma U_1 - U_2$ for some $\gamma > 0$.

Next, we apply the first part of the proposition: positively scaling if necessary, it is without loss to assume $f := U_1 - U_2$ is convex.

For $i = 1, 2$, fix an optimal policy $p_i^* \in ext\mathcal{R}(\mu)$. Notice that a random posterior (in the present binary-state setting) is extreme if and only if it has support of size ≤ 2. Say $S_{p_i^*} = \{v_i^L, v_i^H\}$, where $0 \leq v_i^L \leq \mu \leq v_i^H \leq 1$. Because U_1 is continuous, we may further assume p_1^* is Blackwell-maximally optimal.

For each $z \in \{H, L\}$, we can construct some $p^z \in \mathcal{R}(v_1^z)$ such that $\int U_1 \, dp^z \geq U_1(v_1^z)$ and $p^z((v_2^L, v_2^H)) = 0$.

- Indeed, if $v_1^z \notin (v_2^L, v_2^H)$, just let p^z be the point mass on v^z. Otherwise, let p^z put weight on $S_{p_2^*}$. We know that $\int U_2 \, dp^z \geq U_2(v_1^z)$ by optimality of p_2^*, and it then follows that $\int U_1 \, dp^z \geq U_1(v_1^z)$ by convexity of f.

Replacing each v_1^z with p^z, gives a more informative optimal policy for U_1. By our assumption that p_1^* is Blackwell-maximal, it must be that we in fact haven't changed p_1^* at all; that is, $v_1^L, v_1^H \notin (v_2^L, v_2^H)$. So

$$v_1^L \leq v_2^L \leq v_2^H \leq v_1^H.$$

It follows readily that p_1^* is more informative than p_2^*. $\qquad\square$

---

[27]i.e. $f\left(\{\delta_\theta : \theta \in \hat{\Theta}\}\right) = \mu(\hat{\Theta})$ for every Borel $\hat{\Theta} \subseteq \Theta$.

9.8   On Optimal Policies: Proofs and Additional Results

Proof of Theorem 4

Proof. By continuity of Blackwell's order, there is a $\geq_B^\mu$-maximal optimal policy $p \in \mathcal{R}(\mu)$.

For any $C \in \mathcal{C}$, it must be that $p(C \setminus extC) = 0$. Indeed, Phelps (2001, Theorem 11.4) provides a measurable map $r : C \to \Delta(extC)$ with $r(\cdot|v) \in \mathcal{R}(v)$ for every $v \in C$. Then we can define $p' \in \mathcal{R}(\mu)$ via $p'(S) = p(S \setminus C) + \int_C r(S|\cdot)\, dp$ for each Borel $S \subseteq \Delta\Theta$. Then U-covering and Jensen's inequality imply $\int U\, dp' \geq \int U\, dp$, so that p' is optimal too. By construction, $p' \geq_B^\mu p$, so that (given maximality of p) the two are equal. Therefore $p(C \setminus extC) = p'(C \setminus extC) = 0$. Then, since $\mathcal{C}$ is countable,

$$p(out(\mathcal{C})) = 1 - p\left(\bigcup_{C \in \mathcal{C}}[C \setminus extC]\right) = 1.$$

$\square$

Given the above, it is useful to have results about how to find posterior covers. Below, we prove two useful propositions for doing just that.

Proof of Proposition 3

Proof. As an intersection of closed convex sets is closed convex, and a sum or supremum of convex functions is convex, the following are immediate.

1. Suppose f is the pointwise supremum of a family of functions, $f = \sup_{i \in I} f_i$. If C_i is an f_i-cover for every $i \in I$, then

$$C := \bigvee_{i \in I} C_i = \left\{\bigcap_{i \in I} C_i :\ C_i \in \mathcal{C}_i\ \forall i \in I\right\}$$

   is an f-cover.

2. Suppose f is the pointwise supremum of a family of functions, $f = \sup_{i \in I} f_i$. If C is an f_i-cover for every $i \in I$, then C is an f-cover.

3. If C is a g-cover and h is convex, then C is a $(g + h)$-cover.

The first part of Proposition 3 follows from (1), letting $I = A$ and $f_a = u(a, \cdot)$. The second part of Proposition 3 follows from (2) and (3), since $U = u_p + U_c$ for convex U_c. $\square$

We note in passing that the proof of Propostion generates a comparative statics result for posterior covers. Given Proposition 2, if U_2 is more information-averse than U_1, then any U_2-cover is a U_1-cover.

Proof of Proposition 4

Proof. By finiteness of I, the collection C covers $\Delta\Theta$. For each $i \in I$, note that $C_i = \bigcap_{j\in I}\{v \in \Delta\Theta : f_i(v) \geq f_j(v)\}$, an intersection of closed convex sets (since $\{f_j\}_{j\in I}$ are affine continuous), and so is itself closed convex. Restricted to C_i, f agrees with f_i and so is affine, and therefore convex. $\qquad\square$

Now, under the hypotheses of Proposition 4, we prove a claim that fully characterizes the set of outer points of the f-cover, reducing their computation to linear algebra.

Claim 1. *The f-cover $C = \{C_i : i \in I\}$ given by Proposition 4 satisfies $out(C) = \{v^* \in \Delta\Theta : \{v^*\} = S(v^*)\}$ where*

$$S(v^*) := \{v \in \Delta\Theta : \; supp(v) \subseteq supp(v^*)$$
$$and \; \exists J \subseteq I \; such \; that \; f_i(v) = f_j(v) = f(v) \; \forall i, j \in J\}. \quad (7)$$

Proof. Fix some $v^* \in \Delta\Theta$, for which we will show $\{v^*\} \neq S(v^*)$ if and only if $v^* \notin out(C)$.

Let us begin by supposing $\{v^*\} \neq S(v^*)$; we have to show $v^* \notin out(C)$. Since $v^* \in S(v^*)$ no matter what, there must then be some $v \in S(v^*)$ with $v \neq v^*$. We will show that $S(v^*)$ must then contain some line segment $co\{v, v'\}$ belonging to some C_i, in the interior of which lies v^*; this will then imply $v^* \notin outC$. Let $\hat{\Theta}$ be the support of v^*, and let $J := \{i \in I : \; v^* \in C_i\}$. Given that $v \in S(v^*)$, we have $v \in \Delta\hat{\Theta}$ with $f_i(v) = f_j(v) = f(v) \; \forall i, j \in J$. Now, for sufficiently small $\epsilon > 0$, we have $\epsilon(v - v^*) \leq v^*$.[28] Define $v' := v^* - \epsilon(v - v^*) \in \Delta\hat{\Theta}$. Then $f_i(v') = f_j(v') = f(v') \; \forall i, j \in J$ too and, by definition of v', we have $v^* \in co\{v, v'\}$. If $i \notin J$, then it implies $f_i(v^*) > f(v)$ because $v^* \notin C_i$. Therefore, by moving v, v' closer to v^* if necessary, we can assume $f(v) = f_j(v) < f_i(v)$ and $f(v') = f_j(v') < f_i(v')$ for any $j \in J$ and $i \notin J$. In particular, fixing some $j \in J$ yields $v, v' \in C_j$, so that v^* is not in $out(C)$.

To complete the proof, let us suppose that $v^* \notin out(C)$, or equivalently, $v^* \in C_i$ but $v^* \notin ext(C_i)$ for some $i \in I$. By definition of C_i, we have that $f_i(v^*) = f(v^*)$. The fact that $v^* \notin ext(C_i)$ implies that there is a non-trivial segment $L \subseteq C_i$ for which v^* is an interior point. It must then be that $supp(v) \subseteq supp(v^*)$ and $f_i(v) = f(v)$ for all $v \in L$. As a result, $L \subseteq S(v^*)$ so that $\{v^*\} \neq S(v^*)$, completing the proof. $\qquad\square$

---

[28]Here, $\leq$ is the usual component-wise order on $\mathbb{R}^{\hat{\Theta}}$.

Corollary 2. *Suppose* $\Theta = \{0, 1\}$*; A is finite; and for each* $a \in A$*,* $u(a, \cdot) = \min_{i \in I_a} f_{a,i}$*, where* $\{f_{a,i}\}_{i \in I_a}$ *is a finite family of distinct affine functions for each a. Then, there exists an optimal policy that puts full probability on*

$$S := \{0, 1\} \cup \bigcup_{a \in A} \{v \in [0, 1] : \quad f_{a,i}(v) = f_{a,j}(v) \text{ for some distinct } i, j \in I_a\}.$$

References

Aliprantis, Charalambos D. and Kim C. Border. 1999. *Infinite Dimensional Analysis*. Berlin: Springer, second ed.

Arrow, Kenneth J. 1971. *Essays in the Theory of Risk Bearing*. Chicago: Markham.

Aumann, Robert J. 1987. "Correlated equilibrium as an expression of Bayesian rationality." *Econometrica: Journal of the Econometric Society* :1–18.

Baumeister, Roy F. 2002. "Yielding to Temptation: Self-Control Failure, Impulsive Purchasing, and Consumer Behavior." *Journal of Consumer Research* 28:670–676.

Ben-shahar, O. and C. Schneider. 2014. "The Failure of Mandated Disclosure." John M. Ohlin Law Econ. Work. Pap. 526, Law Sch., Univ. of Chicago.

Blackwell, David. 1953. "Equivalent Comparison of Experiments." *Annals of Mathematics and Statistics* 32:265–272.

Caplin, Andrew and John Leahy. 2004. "The supply of information by a concerned expert." *Economic Journal* 114 (497):487–505.

Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6):1431–51.

Dubra, J., F. Maccheroni, and E. Ok. 2004. "Expected utility theory without the completeness axiom." *Journal of Economic Theory* 115:118–133.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica. 2014. "Suspense and Surprise." *Quarterly Journal of Economics* :Forthcoming.

Gabaix, Xavier. 2014. "A Sparsity-Based Model of Bounded Rationality." *Quarterly Journal of Economics* :Forthcoming.

Ganuza, Juan-Jose and Jose S. Penalva. 2010. "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions." *Econometrica* 78 (3):1007–1030.

Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological games and sequential rationality." *Games and Economic Behavior* 1 (1):60–79.

Gilboa, Itzhak and David Schmeidler. 1988. "Information dependent games: Can common sense be common knowledge?" *Economics Letters* 27 (3):215–221.

———. 1989. "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics* 18 (2):141–153.

Grant, Simon, Atsushi Kajii, and Ben Polak. 1998. "Intrinsic Preference for Information." *Journal of Economic Theory* 83 (2):233–259.

Gul, Faruk and Wolfgang Pesendorfer. 2001. "Temptation and Self-Control." *Econometrica* 69 (6):1403–1435.

Hansen, Lars Peter and Thomas J. Sargent. 2001. "Robust Control and Model Uncertainty." *American Economic Review* 91 (2):60–66.

Huang, Kevin, Zheng Liu, and John Q. Zhu. 2013. "Temptation and self-control: some evidence and applications." Tech. Rep. 2013-23, Federal Reserve Bank of San Francisco.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101 (6):2590–2615.

Kihlstrom, Richard E. and Leonard Jay Mirman. 1974. "Risk aversion with many commodities." *Journal of Economic Theory* 8 (3):361–388.

Kolotilin, Anton, Ming Li, Tymofiy Mylovanov, and Andriy Zapechelnyuk. 2015. "Persuasion of a Privately Informed Receiver." Working paper.

Kreps, David M. and Evan L. Porteus. 1978. "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica* 46 (1):185–200.

Krusell, Per, Burhanettin Kuruscu, and Anthony A. Smith. 2010. "Temptation and Taxation." 78 (6):2063–2084.

Loewenstein, G., Cass R. Sunstein, and Russell Goldman. 2014. "Disclosure: Psychology Changes Everything." *Annual Economic Review* 6.

Myerson, Roger. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47 (1):61–74.

Phelps, Robert R. 2001. *Lectures on Choquet's Theorem*. Berlin: Springer, second ed.

Pratt, J. W. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* 32:122–136.

Siniscalchi, Marciano. 2011. "Dynamic choice under ambiguity." *Theoretical Economics* 6 (3).

Tardella, Fabio. 2008. "On the existence of polyhedral convex envelopes." In *Frontiers In Global Optimization*, edited by C.A. Floudas and P.M. Pardalos. Boston: Kluwer, 149–188.