Costly Subjective Learning*

Youichiro Higashi, Kazuya Hyogo, and Norio Takeoka †

June 23, 2022

Abstract

In an uncertain environment, rational individuals will optimally acquire information by considering the benefits and costs. Such a model is called rational inattention and typically assumes that the cost is independent of the benefit of information. However, in several instances, the cost may be payoff-dependent in a complex manner. Moreover, from an empirical perspective, considering more general payoff-dependent costs may clarify the problems of misspecified costs. Using choice data to estimate the structure of an improperly assumed payoff-independent cost function could lead to some estimation bias or miscategorization of subjects. To seek a reasonable formulation for information acquisition costs, we investigate a choice theoretic foundation for rational inattention under possibly payoff-dependent costs, which serves as a test focusing solely on the essence of information acquisition without relying on particular features such as the payoff-independence of the cost function. This study takes a preference over menus as primitives, and generalizes the axiomatization of de Oliveira, Denti, Mihm, and Ozbek [15] for rational inattention under payoff-dependent costs.

Keywords: costly information acquisition, rational inattention, payoff-dependent costs, preference for flexibility, preference over menus.

JEL classification: D11, D81, D91.

---

*The authors are grateful to Christopher Chambers, Larry Epstein, David Dillenberger, José Heleno Faro, Kemal Ozbek, Jingyi Xue and the audience at RUD 2019 (Paris School of Economics), AMES 2019 (Xiamen University), and seminar participants at Ecole Polytechnique, Hitotsubashi University, Kyoto University, National Taiwan University, Singapore Management University, and University of Paris 1 for their helpful comments. The authors gratefully acknowledge financial support from 2018–2020 Joint Research Program of KIER (Kyoto University), and Open Research Area (ORA) for the Social Sciences 2018. This study was also supported by JSPS KAKENHI Grant Numbers JP20K01562(Y. H.), JP19K01566(K. H.), and JP15K03351(N. T.).

†Higashi is at the Faculty of Economics, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan, higash-y@okayama-u.ac.jp; Hyogo is at the Faculty of Economics, Ryukoku University, 67 Fukakusa Tsukamoto-cho, Fushimi-ku, Kyoto 612-8577, Japan, hyogo@econ.ryukoku.ac.jp; Takeoka is at the Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan, norio.takeoka@r.hit-u.ac.jp.

1 Introduction

1.1 Objective

If the information about payoff-relevant states is imperfect, an additional piece of information may be obtained by conducting information acquisition or experimentation before deciding on payoff-relevant actions. Rational individuals will optimally acquire information by considering the benefits and costs of information acquisition. Since a seminal work of Sims [33], the implications of such a hypothesis, called rational inattention, have been studied in much of the literature, both in theoretical and empirical studies. This hypothesis is also useful to examine whether individuals deliberately make probabilistic choices (Caplin and Dean [3]).

The existing literature typically assumes the agent to maximize net benefits from information minus costs. This formulation implicitly requires that the cost depends on the experiment or the information structure adopted, but is independent of the benefit from the experiment. However, in some instances, the cost of information acquisition is payoff-dependent. For example, consider a manager who makes an investment decision and consults with an expert who can access information relevant to the investment. If the manager pays a certain fraction of the value of information to the expert, the cost of information is payoff-dependent. For another example, imagine an agent who is considering how many times he should observe a probabilistic signal about payoff relevant states before choosing an action. If the marginal cost of obtaining the signal is not clear, the agent may try to maximize the benefit/cost ratio because the ratio criterion does not require the agent to estimate the marginal cost (see Gabaix, Laibson, Moloche, and Weinberg [23]).

From an empirical perspective, considering more general payoff-dependent costs may clarify the problems of using misspecified estimation models. Suppose, for example, that choice data on information acquisition are obtained from subjects for whom the true information acquisition costs are payoff-dependent. Using that data to estimate the structure of an improperly assumed payoff-independent cost function could lead to some estimation bias. Also, using the estimation results to classify cost types may lead to miscategorization of subjects. To understand what qualitative and quantitative biases may arise, it is useful to formulate general payoff-dependent cost functions.

To seek such a reasonable formulation for information acquisition costs, a main difficulty is that these costs are determined through the agent's subjective information acquisition process, which is largely unobservable to the modeler. The purpose of this paper is to investigate a choice theoretic foundation for information acquisition under payoff-dependent costs. Such a foundation is useful for at least three reasons. First, it exactly identifies the essential behavioral implications of information acquisition, whereby, we can evaluate whether the implications of a specific model of information acquisition depend solely on the essence of information acquisition or on other particular features, such as the payoff-independence of the cost function. Second, the resulting representation characterized from the behavioral foundation suggests a reasonable and testable formulation of payoff-dependent costs for information acquisition. Third, the foundation sets the stage for studying more specific models of information acquisition, such as the ratio criterion or other more complicated payoff-dependent cost functions.

1.2 Outline

We take the same setting as in de Oliveira, Denti, Mihm, and Ozbek [15] that provide an axiomatic foundation for the payoff-independent costs for information acquisition. An agent who chooses a menu F of Anscombe-Aumann acts f which associates a lottery with each state ω. After the choice of the menu, the agent may conduct an additional experiment or engage in information acquisition, which generates signals about states. The agent updates her initial prior $\overline{p}$ and makes a choice from the menu contingent upon posteriors p.

Following the literature, we interpret information acquisition as a choice of an experiment π which is a probability distribution over posteriors and whose average coincides with the prior $\overline{p}$. Then the benefit of information for the experiment π is defined as

$$b_F^u(\pi) \equiv \int \left(\max_{f \in F} \sum_{\omega} u(f(\omega))p(\omega) \right) \mathrm{d}\pi(p),$$

where u is a vNM utility index. Given the posterior p, the choice from the menu F maximizes the expected utility $\sum_{\omega} u(f(\omega))p(\omega)$. When evaluating the experiment π, the agent does not know which posterior will prevail. The benefit of information is computed as the expectation of these maximum values with respect to the distribution over posteriors.

The agent optimally chooses an experiment by considering the benefits and costs of acquiring information. The utility of any menu F is given by

$$U(F) = \max_{\pi \in \Pi(\overline{p})} W(\pi, b_F^u(\pi)) \tag{1}$$

for some aggregator function W called a net benefit function. Here, $\Pi(\overline{p})$ is the set of experiments consistent with the prior. We refer to the utility function U as a Costly Subjective Learning (CSL) Representation. We impose reasonable properties on W, which justify our interpretation of W being a net benefit of information acquisition. In this paper, we provide an axiomatic characterization of the CSL Representation.

To formulate payoff-dependent costs more explicitly, define the cost of information for the experiment π by $C(\pi, b_F^u(\pi)) \equiv b_F^u(\pi) - W(\pi, b_F^u(\pi))$. Then the CSL Representation is rewritten as

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b_F^u(\pi) - C(\pi, b_F^u(\pi))\}.$$

One special case is the payoff-independent cost characterized in de Oliveira, Denti, Mihm, and Ozbek [15], where $C(\pi, b_F^u(\pi)) = c(\pi)$ for some function c. Another is a homogeneous payoff-dependent cost given by $C(\pi, b_F^u(\pi)) = \gamma_s(\pi)|b_F^u(\pi)|$, where $s = \mathrm{sgn}(b_F^u(\pi))$.

We provide three applications with payoff-dependent costs. In the first application, we take Dewan and Neligh [16] as an example. In Dewan and Neligh [16], payoff-independent costs are estimated for each subject using experimental data about information acquisition, and subjects are categorized according their cost types. We then discuss possible categorization errors that may occur when cost types are classified assuming payoff-independent costs. We suggest additional experiments necessary for more appropriate cost classification when information acquisition costs are assumed to general payoff-dependent costs, including payoff-independent costs as special cases.

The second application is to examine the estimation bias when estimating the cost of information acquisition under a misspecified formulation. de Oliveira, Denti, Mihm, and Ozbek [15], which axiomatize payoff-independent costs, present a formula for estimating the cost of each information structure from choice data. We show that when true information acquisition costs are homogeneous, the estimation of costs under improperly assumed payoff-independence leads to incorrect predictions about the agent's information choice. We also show that when the modeler incorrectly assumes payoff-independence for an agent with a general CSL representation, costs are overestimated relative to true ones.

Finally, we consider an optimal sampling problem addressed in Cukierman [11]. An agent can obtain an additional piece of information about payoff-relevant states by observing some stochastic signals before choosing a final investment decision. The more samples obtained, the more accurate posteriors can be obtained, but at a cost proportional to the number of samples. The agent determines the optimal number of samples. Under the payoff-independent costs assumed in Cukierman [11], the optimal number of samples does not change when the mean of the prior on states improves. However, it is conceivable that the incentive to acquire information may weaken in response to such a change. We show the comparative statics that when information acquisition costs are homogeneous, the optimal sample size decreases when the mean of the prior improves.

1.3 Related literature

In the literature on preference over menus, Ergin and Sarver [22] introduce subjective optimization in the context of contemplation costs. They generalize the additive representation of Dekel, Lipman, and Rustichini [12], and characterize the additive cost function for contemplating subjective states, which is technically regarded as a counterpart of the variational preference of Maccheroni, Marinacci, and Rustichini [29].

Dillenberger, Lleras, Sadowski, and Takeoka [19] extend the framework of Dekel, Lipman, and Rustichini [12] by considering preference over menus of acts, similar to the present study. They derive a subjective information structure from preference and call their framework, subjective learning. In their model, the agent uses a common experiment for all menus. To accommodate the menu-dependent aspect of information acquisition, de Oliveira, Denti, Mihm, and Ozbek [15] generalize the subjective learning model and characterize a subjective optimization under payoff-independent costs for information acquisition. As a special case, de Oliveira [14] axiomatizes a more specific cost function, called the relative entropy, which is commonly used in the literature of rational inattention. By considering a preference over lotteries of menus consisting of acts, Pennesi [32] shows that payoff-independent cost models by de Oliveira, Denti, Mihm, and Ozbek [15] is characterized by preference for early resolution of uncertainty and other basic axioms.

By investigating preferences over pairs consisting of an action and a menu of acts, Hyogo [25] characterizes general models and payoff-independent cost models of costly information acquisition. Although information contents are subjective, it is assumed that choices of experiments are observable. Dillenberger, Krishna, and Sadowski [17, 18] consider repeated decisions of information acquisition in an infinite horizon framework of menu choice. They model information acquisition from a constrained set of experiments, which is called a constrained information model.

An alternative approach to identify costs for information acquisition is to consider a stochastic choice from menus of acts. Caplin and Dean [4] identify payoff-independent costs for information acquisition from a state-dependent stochastic choice. Caplin, Dean and Leahy [5] and Denti [13] also take a state-dependent stochastic choice as primitives, and characterize a specific class, called posterior separable costs, which includes the expected relative entropy of posterior and prior beliefs used in the literature of rational inattention. Chambers, Liu, and Rehbeck [9] take the same primitives of Caplin and Dean [4] and identify payoff-dependent information costs. Thus, their result is complementary to ours, established in the model of preference over menus.

Lin [27] provides a parsimonious model by only assuming state-independent stochastic choice, which is built on the framework of Lu [28], and characterizes payoff-independent costs. Duraj and Lin [20] take the parsimonious framework and characterize discounting costs.

Ellis [21] considers a state-dependent deterministic choice function from menus of acts and derives a cost function for partitions, which is interpreted as costly partitional learning. Aoyama [2] extends Ellis [21] by incorporating decision time as a part of primitives and derives a cost function for filtrations.

To axiomatize the CSL Representation (1), we borrow techniques from the literature of choice under ambiguity. The CSL Representation is a counterpart of the uncertain averse representation of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8], which nests two representations as special cases. One is the variational representation of Maccheroni, Marinacci, and Rustichini [29], which satisfies the property, called translation invariance, and has a parallel relationship with the payoff-independent cost model. The other is the confidence representation of Chateauneuf and Faro [10], which satisfies homotheticity and has a parallel relationship with the homogeneous payoff-dependent cost model.

2 Costly subjective learning representations

2.1 Primitives

We consider the following as primitives of the model: these primitives are exactly the same as in de Oliveira, Denti, Mihm, and Ozbek [15].

- $\Omega = \{\omega_1, ..., \omega_n\}$: the (finite) objective state space
- X: outcomes, consisting of simple lotteries on a set of deterministic prizes
- $f : \Omega \to X$: an (Anscombe-Aumann) act
- $\mathcal{F}$: the set of all acts
- $F \subset \mathcal{F}$: a nonempty finite set of acts, called a menu
- $\mathbb{F}$: the set of all menus
- Preference $\succsim$ over $\mathbb{F}$

2.2 Functional form

To begin with, we introduce the important concept of a martingale property or a Bayes plausibility (Kamenica and Gentzkow [26]) in our setting. Let $\overline{p} \in \Delta(\Omega)$ be the agent's prior belief. A probability distribution $\pi \in \Delta(\Delta(\Omega))$ is interpreted as an experiment or a signal structure about Ω. For each π, the initial prior $p^\pi \in \Delta(\Omega)$ associated with π is defined as $p^\pi(\omega) = \int_{\Delta(\Omega)} p(\omega)\, \mathrm{d}\pi(p)$ for each ω. We impose a restriction on the relationship between the prior belief and subjectively possible experiments. We say that π satisfies a martingale property if $p^\pi = \overline{p}$. That is, the initial prior associated with π exactly coincides with the agent's prior belief $\overline{p}$. The set of experiments that satisfy the martingale property is denoted by $\Pi(\overline{p}) = \{\pi \in \Delta(\Delta(\Omega)) \,|\, p^\pi = \overline{p}\}$, which is weak* closed and convex.

Given $u : X \to \mathbb{R}$ and a menu F, the benefit of information for an experiment $\pi \in \Pi(\overline{p})$ is defined as

$$b_F^u(\pi) \equiv \int_{\Delta(\Omega)} \max_{f \in F} \left(\sum_{\omega \in \Omega} u(f(\omega))p(\omega) \right) \mathrm{d}\pi(p).$$

In particular, for any singleton menu $\{f\}$ and any experiment $\pi \in \Pi(\overline{p})$, we have $b_{\{f\}}^u(\pi) = \sum_\Omega u(f(\omega))\overline{p}(\omega)$. That is, the benefit of the information exactly coincides with the expected utility of f under the prior when the agent commits to choose the act f.

To capture the trade-offs in information acquisition, we introduce the Blackwell order, which gives a partial order on $\Delta(\Delta(\Omega))$ in terms of informativeness of signals.

Definition 1 *An experiment $\pi \in \Delta(\Delta(\Omega))$ is Blackwell more informative than an experiment $\rho \in \Delta(\Delta(\Omega))$, denoted $\pi \trianglerighteq \rho$, if*

$$\int_{\Delta(\Omega)} \varphi(p)\, \mathrm{d}\pi(p) \geq \int_{\Delta(\Omega)} \varphi(p)\, \mathrm{d}\rho(p)$$

for every convex continuous function $\varphi : \Delta(\Omega) \to \mathbb{R}$.

As $\max_{f\in F}\left(\sum u(f(\omega))p(\omega)\right)$ is convex and continuous in p, we have $b_F^u(\pi) \geq b_F^u(\rho)$ for all menus F whenever π is Blackwell more informative than ρ.

There are not only benefits but also costs in acquiring information. We consider an aggregator function $W : \Pi(\overline{p}) \times \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$, which captures the net value of information, taking into account the costs for acquiring information. We say that W is linearly continuous if the map $\varphi \mapsto \sup_{\pi\in\Pi(\overline{p})} W\left(\pi, \int \varphi\, \mathrm{d}\pi\right)$ from the set of continuous functions on $\Delta(\Omega)$, denoted by $C(\Delta(\Omega))$, into $[-\infty, \infty]$, is extended-valued continuous. Let δ_a denote the Dirac measure at a.

Definition 2 *We say that $W : \Pi(\overline{p}) \times \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ is a net benefit function if*

(i) W is quasi-concave, upper semi-continuous, and linearly continuous,

(ii) for all π, $W(\pi, t)$ is non-decreasing in t,

(iii) $W(\delta_{\overline{p}}, t) = t$ for the initial prior $\overline{p}$,

(iv) for all t and $\pi, \rho \in \Pi(\overline{p})$, $\pi \trianglerighteq \rho \implies W(\pi, t) \leq W(\rho, t)$.

The function $W(\pi, t)$ captures the net benefit when an experiment π is chosen and t is the gross benefit of information. Part (i) is a technical condition to ensure a well-defined optimization problem of information acquisition. Part (ii) states that for each fixed π, the net benefit increases accordingly when the gross benefit of information increases. Part (iii) states that the gross and net benefits coincide if the prior information is used. In other words, there is no cost if there is no additional information acquisition. Part (iv) states that a more informative experiment is more costly. In fact, for each fixed level of t, its net benefit is lower under a more informative signal structure. Note also that any $\pi \in \Pi(\overline{p})$ is a mean-preserving spread of $\delta_{\overline{p}}$, and hence, $\pi \trianglerighteq \delta_{\overline{p}}$. From parts (iii) and (iv), $W(\pi, t) \leq W(\delta_{\overline{p}}, t) = t$, which means that the net benefit is always lower than the gross benefit of information t. Hence, the cost of choosing π is implicitly embodied into W.

Definition 3 *A Costly Subjective Learning (CSL) Representation is a tuple $(u, \overline{p}, W)$, where $u : X \to \mathbb{R}$ is an unbounded expected utility function with $u(X) = \mathbb{R}$, $\overline{p}$ is the initial prior, and W is a net benefit function such that $\succsim$ is represented by*

$$U(F) = \max_{\pi\in\Pi(\overline{p})} W(\pi, b_F^u(\pi)).$$

We explain special cases contained in the CSL Representation in Section 4.

3 Behavioral foundation

3.1 Basic Axioms

We provide a behavioral foundation of the CSL Representation. We start with the basic axioms that are consistent with any type of costly information acquisition.

Axiom 1 (Order) *$\succsim$ satisfies completeness and transitivity.*

For all F, G, and $\alpha \in [0,1]$, define a mixture of F and G by

$$\alpha F + (1-\alpha)G = \{\alpha f + (1-\alpha)g \,|\, f \in F,\ g \in G\} \in \mathbb{F},$$

where $\alpha f + (1-\alpha)g \in \mathcal{F}$ is defined by the state-wise mixture between f and g.

Axiom 2 (Mixture Continuity) *For all menus F, G, and H, the following sets are closed:*

$$\{\alpha \in [0,1] \,|\, \alpha F + (1-\alpha)G \succsim H\} \text{ and } \{\alpha \in [0,1] \,|\, H \succsim \alpha F + (1-\alpha)G\}.$$

Axiom 3 (Preference for Flexibility) *For all menus F and G, if $G \subset F$, then $F \succsim G$.*

This axiom states that a bigger menu is always weakly preferred.

Axiom 4 (Dominance) *For all menus F and acts g, if there exists $f \in F$ with $\{f(\omega)\} \succsim \{g(\omega)\}$ for all $\omega \in \Omega$, then $F \sim F \cup \{g\}$.*

As $F \subset F \cup \{g\}$, the latter menu is weakly preferred by preference for flexibility. However, if $\{f(\omega)\} \succsim \{g(\omega)\}$ for all $\omega \in \Omega$, f gives a more preferred lottery than g does for all states. In this sense, g is dominated by f. Irrespective of the belief the agent has on the states, g should not be chosen over f. Thus, adding g to F does not provide a strictly higher value of flexibility than F.

Axiom 5 (Two-Sided Unboundedness) *There are outcomes $x, y \in X$ with $\{x\} \succ \{y\}$ such that for all $\alpha \in (0,1)$, there are $z, z' \in X$ satisfying*

$$\{\alpha z' + (1-\alpha)y\} \succ \{x\} \succ \{y\} \succ \{\alpha z + (1-\alpha)x\}.$$

This axiom implies the unbounded range of a utility function over outcomes X.[1]

3.2 Substantive axioms for information acquisition

Based on the subjective learning model provided by Dillenberger, Lleras, Sadowski, and Takeoka [19], de Oliveira, Denti, Mihm, and Ozbek [15] argue that weakening the independence axiom is the key to accommodate costly information acquisition.[2] In the following, we introduce two weakenings of the independence axiom: Singleton Independence and Aversion to Contingent Planning. The latter is one of the axioms adopted by de Oliveira, Denti, Mihm, and Ozbek [15], while the former is implied from their other axioms.

[1]de Oliveira, Denti, Mihm, and Ozbek [15] assume One-Sided Unboundedness: There are outcomes $x, y \in X$ with $\{x\} \succ \{y\}$ such that for all $\alpha \in (0,1)$, there is $z \in X$ satisfying either $\{\alpha z + (1-\alpha)y\} \succ \{x\}$ or $\{y\} \succ \{\alpha z + (1-\alpha)x\}$. The role of Two-Sided Unboudedness is explained in the proof sketch of the theorem (see Section 3.3).

[2]For more details of their model, see Section 4.1. We discuss their work as a special case of the CSL representation.

Axiom 6 (Singleton Independence) *For all acts f, g, h, and $\alpha \in (0,1)$,*

$$\{f\} \succsim \{g\} \iff \alpha\{f\} + (1-\alpha)\{h\} \succsim \alpha\{g\} + (1-\alpha)\{h\}.$$

This axiom is to impose the independence only on singleton menus. If the agent makes a commitment to a singleton menu $\{f\}$, there is no role for information acquisition after menu choice. Thus, the commitment rankings reflect the agent's prior belief over states. Singleton Independence implies that the agent follows the subjective expected utility to evaluate acts with commitment according to his prior belief.

Formally, the next axiom requires quasi-convexity of preference.

Axiom 7 (Aversion to Contingent Planning) *For all menus F, G, and $\alpha \in (0,1)$,*

$$F \sim G \implies F \succsim \alpha F + (1-\alpha)G.$$

Note that $\alpha F + (1-\alpha)G$ is the menu of contingent plans of the form $\alpha f + (1-\alpha)g$, where $f \in F$ and $g \in G$. It is instructive to compare this menu with a randomization between F and G, denoted by $\alpha \circ F + (1-\alpha) \circ G$. These two alternatives differ in the timing of resolution of randomization. In the latter, the randomization α realizes before a choice from menus. Thus, information acquisition can be conducted contingent upon a realized menu F or G. Since $F \sim G$, $\alpha \circ F + (1-\alpha) \circ G$ should be indifferent to F. On the other hand, in the former, the randomization α is not resolved when the agent makes a choice from the menu $\alpha F + (1-\alpha)G$. Since information acquisition cannot be completely tailored for F and G, $\alpha F + (1-\alpha)G$ should be less preferred to $\alpha \circ F + (1-\alpha) \circ G$. That is, the axiom states that the agent avoids contingent planning.

We are now ready to present a representation theorem.

Theorem 1 *Preference $\succsim$ satisfies the basic axioms, Singleton Independence, and Aversion to Contingent Planning if and only if it admits a Costly Subjective Learning Representation $(u, \overline{p}, W)$. Moreover, the net benefit function W is obtained as*

$$W(\pi, t) = \inf_{\{F \,|\, b_F^u(\pi) \geq t\}} u(x_F) = \inf_{\{F \,|\, b_F^u(\pi) = t\}} u(x_F), \tag{2}$$

where $x_F \in X$ is a lottery equivalent of F satisfying $F \sim \{x_F\}$.

To obtain an intuition behind (2), note that by the CSL representation, for all menus F and all experiments π, $U(F) \geq W(\pi, b_F^u(\pi))$, and, in particular, if π^* is an optimal experiment for F, $U(F) = W(\pi^*, b_F^u(\pi^*))$. Let $t^* = b_F^u(\pi^*)$. For all menus G satisfying $b_G^u(\pi^*) \geq t^*$, since $W(\pi^*, t)$ is non-decreasing in t,

$$U(G) \geq W(\pi^*, b_G^u(\pi^*)) \geq W(\pi^*, t^*),$$

which suggests that $W(\pi^*, t^*)$ is obtained as the infimum of $u(x_G)$ among $\{G \,|\, b_G^u(\pi^*) \geq t^*\}$, as stated in (2).

The expression of W, as in (2), provides an explicit formula for eliciting the net benefit function. The vNM utility index u is elicited in a standard manner. Then, the gross benefit of information $b^u_F(\pi)$ is computed according to its definition. If a lottery equivalent of each F is elicited from the agent's preference, the net benefit function, under which costs for information acquisition is implicitly involved, can be computed according to (2). Moreover, the second equality of (2) shows that less menus are enough for recovering this cost function, which simplifies the elicitation procedure.

The next theorem shows the uniqueness property of the CSL Representation.

Theorem 2 *If there exist two CSL Representations of $\succsim$, denoted by $(u_i, \overline{p}_i, W_i)$ for $i = 1, 2$, then there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u_1 = \alpha u_2 + \beta$, $\overline{p}_1 = \overline{p}_2 = \overline{p}$, and*

$$W_1(\pi, t) = \alpha W_2\left(\pi, \frac{t-\beta}{\alpha}\right) + \beta$$

for all $(\pi, t) \in \Pi(\overline{p}) \times \mathbb{R}$.

Proof. By the uniqueness result of Anscombe and Aumann [1], there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u_1 = \alpha u_2 + \beta$. Moreover, $\overline{p}_1 = \overline{p}_2 = \overline{p}$. For any $(\pi, t) \in \Pi(\overline{p}) \times \mathbb{R}$,

$$\begin{aligned} W_1(\pi, t) &= \inf_{\{F \,|\, b^{u_1}_F(\pi) \geq t\}} u_1(x_F) = \inf_{\{F \,|\, \alpha b^{u_2}_F(\pi) + \beta \geq t\}} \alpha u_2(x_F) + \beta \\ &= \alpha \inf_{\{F \,|\, b^{u_2}_F(\pi) \geq \frac{t-\beta}{\alpha}\}} u_2(x_F) + \beta = \alpha W_2\left(\pi, \frac{t-\beta}{\alpha}\right) + \beta. \end{aligned}$$

■

3.3 Proof sketch of Theorem 1

The following is a proof sketch of the sufficiency. Using techniques known since Dekel, Lipman, and Rustichini [12], a preference over menus induces a preference over continuous functions on a suitably defined state space (or the set of posteriors in our setting) through the characterization of convex sets via support functions. Then, the desired utility representation can be obtained by applying techniques of utility representations in the choice under uncertainty. Finally, we show the martingale property of the information structures.

The first step is standard. We follow the construction of the support functions by de Oliveira, Denti, Mihm, and Ozbek [15]. Singleton Independence and the basic axioms ensures a subjective expected utility representation over $\mathcal{F}$ with an expected utility $u : X \to \mathbb{R}$ and a prior $\overline{p} \in \Delta(\Omega)$. Two-Sided Unboundedness implies the property of unbounded range $u(X) = \mathbb{R}$. This representation $U : \mathcal{F} \to \mathbb{R}$ is extended to the whole domain $\mathbb{F}$ because each menu F has its lottery equivalent $\{x_F\}$.

For any $F \in \mathbb{F}$, a support function for F is defined as, for any posterior $p \in \Delta(\Omega)$,

$$\varphi_F(p) = \max_{f \in F} \sum_{\Omega} u(f(\omega))p(\omega). \tag{3}$$

The support function identifies the menu up to indifference: $\varphi_F = \varphi_G \implies F \sim G$. Let $\Phi_{\mathbb{F}} = \{\varphi_F \,|\, F \in \mathbb{F}\} \subset C(\Delta(\Omega))$ be the set of all support functions. Given the above identification, we can induce the functional $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ by $V(\varphi_F) = U(F)$. As $\succsim$ satisfies Mixture Continuity and Aversion to Contingent Planning, we show that V is monotone, normalized, quasi-convex, and continuous, following the techniques of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8]. The functional V is extended to the set of all continuous functions $C(\Delta(\Omega))$, preserving the above properties.

To rewrite the functional $V : C(\Delta(\Omega)) \to \mathbb{R}$ as the CSL Representation, we use a duality argument for complete, monotone, and quasi-convex functionals developed by Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [6, 7]. Let $ca_+(\Delta(\Omega))$ be the set of non-negative measures on $\Delta(\Omega)$, and π denotes its generic element. For any $\varphi \in C(\Delta(\Omega))$, let $\langle \varphi, \pi \rangle = \int \varphi(p)\, \mathrm{d}\pi(p)$. If π is a probability measure and $\varphi = \varphi_F$, we have $\langle \varphi, \pi \rangle = b_F^u(\pi)$. Thus, in this subsection, we interpret π and $\langle \varphi, \pi \rangle$ an experiment and its benefit of information, respectively.

As suggested by (2), the net benefit function $W(\pi, t)$ is obtained by taking the infimum of the functional values $V(\varphi)$ restricting the benefit of information of π to be greater than t. That is, for all π and $t \in \mathbb{R}$, define the net benefit function as

$$W(\pi, t) = \inf_{\varphi \in B(\pi, t)} V(\varphi),$$

where $B(\pi, t) = \{\varphi \in C(\Delta(\Omega)) \,|\, \langle \varphi, \pi \rangle \geq t\}$.

Conversely, we show that V can be recovered from W, that is, we establish a duality between V and W.[3] A key observation is that, by the definition of B, $\varphi \in B(\pi, \langle \varphi, \pi \rangle)$ for all φ. Then, by the definition of W, $V(\varphi) \geq W(\pi, \langle \varphi, \pi \rangle)$ for all φ and π, which in turn implies

$$V(\varphi) \geq \sup_{\pi \in ca_+(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle),$$

for all φ. Thus, $V(\varphi)$ is an upper bound of net benefits $W(\pi, \langle \varphi, \pi \rangle)$ among various π. A critical step is to show that there exists π which exactly achieves $V(\varphi)$ as the supremum. Since V is continuous and quasi-convex, its strict lower contour set at φ is open and convex. By the separation hyperplane theorem, there exists some $\widetilde{\pi}$ such that $B(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle)$ has no overlap with the strict lower contour set, that is, $\varphi' \in B(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle)$ implies $V(\varphi') \geq V(\varphi)$. Therefore,

$$V(\varphi) = \inf_{\varphi' \in B(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle)} V(\varphi') = W(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle) = \max_{\pi \in ca_+(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle),$$

as desired.

As W is homogeneous of degree zero, the above maximum is achieved on $\Delta(\Delta(\Omega))$. Without loss of generality, the domain of the maximization problem can be restricted to

$$\Pi = \{\pi \in \Delta(\Delta(\Omega)) \,|\, W(\pi, t) > -\infty \text{ for some } t\}.$$

[3] The duality here is a formal analogue of the duality between direct and indirect utility functions in the consumer theory.

As $\langle \varphi_F, \pi \rangle = b^u_F(\pi)$ for all menus F, $\succsim$ is represented by

$$U(F) = V(\varphi_F) = \max_{\pi \in \Pi} W(\pi, b^u_F(\pi)). \tag{4}$$

The remaining key step is to show the martingale property, that is, $\Pi \subset \Pi(\overline{p})$, which has no counterpart in the literature on ambiguity. If this is the case, the maximization of (4) is taken on $\Pi(\overline{p})$, additionally requiring $W(\pi, \cdot) = -\infty$ for $\pi \in \Pi(\overline{p}) \setminus \Pi$. The unboundedness of u plays a key role for this step.[4] If there exists $\pi^* \in \Pi$ such that $p^{\pi^*} \neq \overline{p}$, we can find two states ω and ω' such that $p^{\pi^*}_{\omega} > \overline{p}_{\omega}$ and $p^{\pi^*}_{\omega'} < \overline{p}_{\omega'}$. Then, by unboundedness of u, we can find two acts f and $\widetilde{f}$ such that these two acts give the same expected utility in terms of $\overline{p}$, while $\widetilde{f}$ gives a strictly higher expected utility than f in terms of p^{π^*}. Since the CSL representation takes the maximum expected utility among all information structures in Π, we have $U(\{f\}) < U(\{\widetilde{f}\})$. But, since f and $\widetilde{f}$ give the same expected utility in terms of $\overline{p}$, we also have $\{f\} \sim \{\widetilde{f}\}$, which is a contradiction.

3.4 Interpersonal comparisons

Consider two agents $i = 1, 2$ having preferences $\succsim_i$ on $\mathbb{F}$. The following condition is a behavioral comparison in terms of attitude toward flexibility. The same condition is considered in Dillenberger, Lleras, Sadowski, and Takeoka [19] and de Oliveira, Denti, Mihm, and Ozbek [15].

Definition 4 *$\succsim_1$ is more averse to commitment than $\succsim_2$ if for all $F \in \mathbb{F}$ and $f \in \mathcal{F}$,*

$$F \succsim_2 \{f\} \implies F \succsim_1 \{f\}.$$

We have the following characterization:

Theorem 3 *Given two preferences $\succsim_i$, $i = 1, 2$ with Costly Subjective Learning Representations $(u_i, \overline{p}_i, W_i)$ for $i = 1, 2$, the following conditions are equivalent:*

(a) $\succsim_1$ is more averse to commitment than $\succsim_2$;

(b) there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u_1 = \alpha u_2 + \beta$, $\overline{p}_1 = \overline{p}_2 = \overline{p}$, and $W_1(\pi, t) \geq \alpha W_2(\pi, \frac{t-\beta}{\alpha}) + \beta$ for all $(\pi, t) \in \Pi(\overline{p}) \times \mathbb{R}$.

As this theorem shows, if agent 1 is more averse to commitment than agent 2, under a suitable normalization, agent 1's net benefit of information is always greater than agent 2's. In other words, information acquisition is always more costly for agent 2.

[4] de Oliveira, Denti, Mihm, and Ozbek [15]'s argument for proving the martingale property depends the payoff-independent cost, and not directly applicable for more general cases.

4 Special cases

Define the cost of information for the experiment π by $C(\pi, t) \equiv t - W(\pi, t)$. Then the CSL Representations can be rewritten as

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b_F^u(\pi) - C(\pi, b_F^u(\pi))\}.$$

In the following, we see various special cases of $C(\pi, t)$ and the corresponding representations.

4.1 Payoff-independent cost

The first special case is the rationally inattentive representation considered in de Oliveira, Denti, Mihm, and Ozbek [15], where $C(\pi, t) = c(\pi)$ for some function $c : \Pi(\overline{p}) \to [0, \infty]$. That is, preference is represented by

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b_F^u(\pi) - c(\pi)\}.$$

The function c inherits the properties induced by the net benefit function W.[5] Especially, (iii') if there is no information acquisition, there is no cost: $c(\delta_{\overline{p}}) = 0$; and (iv') a more informative experiment is more costly: $\pi' \trianglerighteq \pi \Longrightarrow c(\pi') \geq c(\pi)$.

We stress that the cost does not depend on the gross benefit of information $b_F^u(\pi)$. Therefore, in this paper, we call the above representation a payoff-independent cost representation rather than a rationally inattentive representation.

If the costs of informations are payoff-independent, mixing menus with a singleton act does not change the costs of informations. de Oliveira, Denti, Mihm, and Ozbek [15, Theorem 1] introduce the following weakening of the independence axiom and characterize the payoff-independent cost representation:

Axiom 8 (Independence of Degenerate Decisions) *For all menus F, G, all acts h, h', and $\alpha \in (0, 1)$,*

$$\alpha F + (1-\alpha)\{h\} \succsim \alpha G + (1-\alpha)\{h\} \Longrightarrow \alpha F + (1-\alpha)\{h'\} \succsim \alpha G + (1-\alpha)\{h'\}.$$

When making information acquisition decision for a contingent plan $\alpha F + (1-\alpha)\{h\}$, the agent only cares about information acquisition for αF. Hence, the information acquisition decision for $\alpha F + (1-\alpha)\{h\}$ are the same as that for $\alpha F + (1-\alpha)\{h'\}$. These contingent plans have different payoff levels depending on $\{h\}$, or $\{h'\}$. Independence of Degenerate Decisions says that these constant payoffs are irrelevant for information acquisition decision, whereby implying costs of information being payoff-independent.

[5] Note that $c(\pi) = t - W(\pi, t)$.

Corollary 1 *Suppose that preference $\succsim$ admits a CSL Representation. Then $\succsim$ satisfies Independence of Degenerate Decisions if and only if it admits a payoff-independent cost representation.*

de Oliveira, Denti, Mihm, and Ozbek [15, Claim 3] show that if $\succsim$ satisfies Aversion to Contingent Planning and Independence of Degenerate Decisions, it satisfies Singleton Independence. For the unboundedness axiom, they require only One-Sided of Unboundedness rather than Two-Sided Unboundedness.

Given the other axioms, Independence of Degenerate Decisions is equivalent to Translation Invariance; for all translations θ on X,[6]

$$F \succsim G \iff F + \theta \succsim G + \theta.$$

Since quasi-convexity and translation invariance jointly imply the convexity of the representation, the payoff-independent cost model is delivered from the convex duality applied for niveloids (Maccheroni, Marinacci, and Rustichini [29]).

From the above argument, a convex representation is more general than the payoff-independent cost model, but is nested in the CSL model. Attempts to characterize this intermediate class lead to adopting the following axiom of Mihm and Ozbek [30], which exactly ensures the convexity of the representation:

Axiom 9 (Increasing Desire for Commitment) *For any menus $F, G \in \mathbb{F}$ and lotteries $x, y \in X$, if $F \sim \{x\}$ and $G \sim \{y\}$, $\alpha\{x\} + (1-\alpha)\{y\} \succsim \alpha F + (1-\alpha)G$ for any $\alpha \in [0,1]$.*

The interpretation is that mixing menus adds to complexity and cost of information.

Though one might expect to have some specific payoff-dependent cost representation when Aversion to Contingent Planning is strengthened to Increasing Desire for Commitment in Theorem 1, the characterized model is reduced to the payoff-independent cost model. The reason is that when $u(X) = \mathbb{R}$, the convexity of the representation implies translation invariance.[7]

Corollary 2 *Preference $\succsim$ satisfies the basic axioms, Singleton Independence, and Increasing Desire for Commitment if and only if it admits a payoff-independent cost representation.*

4.2 Homogeneous payoff-dependent cost

One simple specification is a payoff-dependent cost function satisfying homogeneity, that is, for all π, t, and $\lambda > 0$, $C(\pi, \lambda t) = \lambda C(\pi, t)$. This homogeneity implies that the cost function

[6] A translation θ is defined as $\theta = x - y$ for some $x, y \in X$. Accordingly, for all acts f, $f + \theta \in \mathcal{F}$ is defined as $f(\omega) + \theta$ for all ω as long as the operation is feasible. For all menus F, $F + \theta$ is the menu given by $\{f + \theta \mid f \in F\}$.

[7] See Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8, Corollary 38] and Strzalecki [34, Theorem 3] for related results in the setting of preference over acts. Pennesi [32] uses a similar idea for preferences over $\Delta(\mathbb{F})$ and characterizes the payoff-independent cost representation.

is proportional to the gross benefit of information. Indeed, $C(\pi, t) = tC(\pi, 1)$ for $t > 0$ and $C(\pi, t) = |t|C(\pi, -1)$ for $t < 0$. For notational simplicity, denote $C(\pi, t) = \gamma_s(\pi)|t|$, where $s = \mathrm{sgn}(t)$, $\gamma_+(\pi) = C(\pi, 1)$, and $\gamma_-(\pi) = C(\pi, -1)$. Then, the CSL representation is written as

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b^u_F(\pi) - \gamma_s(\pi)|b^u_F(\pi)|\}. \tag{5}$$

The term $\gamma_s(\pi)$ represents a rate of payment for experiment π per the size of payoff. The payment rate can be different depending on whether the benefit of information is positive or negative. Moreover, the payment rate $\gamma_s(\pi)$ inherits the properties induced by the net benefit function W.[8] Especially, (iii") if there is no information acquisition, there is no cost: $\gamma_s(\delta_{\overline{p}}) = 0$ for the initial prior $\overline{p}$; and a more informative experiment is more costly: (iv") for all $\pi, \rho \in \Pi(\overline{p})$, $\pi \trianglerighteq \rho \implies \gamma_s(\pi) \geq \gamma_s(\rho)$.

A salient feature of (5) is "scale-independence". Let $x_0 \in X$ denote a lottery whose size of payoff is zero. We call x_0 a neutral outcome. A mixture $\alpha F + (1-\alpha)\{x_0\}$, simply denoted by αF, is interpreted as the menu obtained from scaling down all the acts in F by α toward the zero payoff. Given (5), it is easy to see that $U(\alpha F) = \alpha U(F)$. Thus, an optimal experiment is invariant in scale changes.

A behavioral counterpart of the scale-independence is the independence axiom imposed only when menus are mixed with the neutral outcome.

Axiom 10 (Neutral Outcome Independence) *For all menus F, G, and $\alpha \in (0,1)$,*

$$F \succsim G \Longleftrightarrow \alpha F + (1-\alpha)\{x_0\} \succsim \alpha G + (1-\alpha)\{x_0\}.$$

This axiom requires that mixing menus with a neutral outcome should not affect the optimal choice of experiments, which guarantees that a CSL Representation is homothetic.[9] Our companion paper, Higashi, Hyogo, and Takeoka [24], characterize the homogeneous payoff-dependent cost representation:

Corollary 3 *Suppose that preference $\succsim$ admits a CSL Representation. Then $\succsim$ satisfies Neutral Outcome Independence if and only if it admits a homogeneous payoff-dependent cost representation.*

Note that Corollary 2 implies that the homogeneous payoff-dependent cost representation is not convex, but quasi-convex.

4.3 Hybrid cost

We can consider a hybrid model between the payoff-independent cost and the homogeneous payoff-dependent cost models, that is, the cost function is given by $C(\pi, t) = c(\pi) + \gamma_s(\pi)|t|$.

[8] Note that $\gamma_+(\pi) = 1 - W(\pi, 1) \in [0, 1]$ and $\gamma_-(\pi) = -1 - W(\pi, -1) \geq 0$.

[9] The homogeneous payoff-dependent cost representation has a parallel relationship with the confidence representation of Chateauneuf and Faro [10], which satisfies homotheticity.

For each fixed π, this cost function is a positive affine function in the gross benefit of information. Also, this class of cost function can be viewed as a convex combination between the payoff-independent cost and the homogeneous payoff-dependent cost.

Under this specification, the CSL representation is written as

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b_F^u(\pi) - \gamma_s(\pi)|b_F^u(\pi)| - c(\pi)\}. \tag{6}$$

For an application of this hybrid cost model, see Section 5.1.

4.4 Additively separable cost

Suppose that a payoff-dependent cost is additively separable between π and t. Then, $C(\pi, t) = c(\pi) + d(t)$. Under this specification, the CSL representation is written as

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{\psi(b_F^u(\pi)) - c(\pi)\}, \tag{7}$$

where $\psi(t) = t - d(t)$.

Seemingly, the representation (7) is a generalization of the payoff-independent cost model as given in Section 4.1. By the similar argument to Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8, Proposition 12], however, when $u(X) = \mathbb{R}$, the representation (7) is nothing but the payoff-independent cost model, that is, ψ is forced to be the identity mapping.

5 Applications

In this section, we provide several examples to illustrate the difference in implications between payoff-independent and payoff-dependent cost functions in information acquisition, whereby illustrating the usefulness of general payoff-dependent cost functions.

5.1 Categorization across information costs

Dewan and Neligh [16] provide an experimental design for eliciting subjects' information choices and categorize subjects according to their information costs with assuming payoff-independent costs. They find that a considerable number of subjects do not respond to incentives. Though they conclude these subjects have a fixed information structure, our payoff-dependent cost model suggests that such non-responsiveness may be attributed to a payoff-dependent information cost. We propose an additional experimental design for distinguishing non-responsive subjects having a payoff-dependent cost and those having a fixed information cost.

5.1.1 Outline

We explain the setting of Dewan and Neligh [16]. Let Ω be the state space with $\#\Omega = n$. The agent has a uniform prior over Ω. The agent receives a reward $r > 0$ if he correctly identifies the true state, otherwise no reward is obtained. Formally, this can be regarded as a choice from the menu $F^{\times r}$ of acts f^{ω}, $\omega \in \Omega$, where $f^{\omega}(\omega) = r$ and $f^{\omega}(\omega') = 0$ for any $\omega' \neq \omega$. Note that a magnitude of r measures the strength of incentives for information acquisition about states. For convenience, such menus are referred to as *payoff scale-up*.

An experiment is captured by a probability matrix $Q = (q_{i,j})_{1 \leq i,j \leq n}$, where $q_{i,j} = P(a = \omega_j \,|\, \omega = \omega_i)$ is a conditional probability of a guess (signal) $a = \omega_j$ given the true state $\omega = \omega_i$. When the agent's guess is $a = \omega_j$, Bayes updating gives a posterior

$$P(\omega = \omega_i | a = \omega_j) = \frac{q_{i,j}\frac{1}{n}}{\sum_k q_{k,j}\frac{1}{n}} = \frac{q_{i,j}}{\sum_k q_{k,j}}. \tag{8}$$

Note that a stochastic matrix is equivalent to a distribution, denoted by $\pi_Q \in \Delta(\Delta(\Omega))$, over posteriors generated from it. Under this distribution, each posterior (8) is realized with a probability $\frac{1}{n}\sum_i q_{i,j}$.

The agent's VNM utility index is assumed to be linear. Under this setting, the gross benefit of information is obtained as

$$b^u_{F^r}(\pi_Q) = \int \max_{f \in F^r} u(f) \cdot p \, \mathrm{d}\pi_Q(p) = r\frac{1}{n}\sum_i q_{i,i} = rP, \text{ where } P = \frac{1}{n}\sum_i q_{i,i}.$$

Choosing $\pi_Q \in \Delta(\Delta(\Omega))$ is identified with choosing $P \in [0,1]$. In the uniform guess task, the agent solves

$$\max_P rP - C(P).$$

An optimal choice, denoted by $P(r)$, is called the performance function.

If C is differentiable and convex, the performance function $P(r)$ satisfies the FOC, that is, $r = C'(P(r))$. Conversely, if the $P(r)$ is obtained as observable data, then this relationship determines marginal costs of C, and C can be recovered by integrating marginal costs.

The FOC, $r = C'(P(r))$, implies that $P(r)$ is non-decreasing in r, which means that the agent (weakly) responds to incentives. This is a basic implication of rational inattention models. Depending on properties of the cost function, other implications on $P(r)$ are obtained. Responsive means that $P(r)$ is strictly increasing in r. If C is continuous and convex (called well-behavedness), then $P(r)$ is continuous. These testable implications are useful for categorizing models of rational inattention.

Dewan and Neligh [16] conduct experiments about the uniform guess task and obtain subjects' data about correctness in the uniform guess task in each incentive level r, from which $P(r)$ is inferred. They report that 86.4% of the subjects are consistent with the rational inattention model. Moreover, 60% of the subjects consistent with rational inattention are responsive to incentives.

5.1.2 Additional experimental design

Their experimental result also suggests that a considerable fraction of subjects (40% of the rationally inattentive subjects) do not respond to incentives. These subjects have a constant $P(r)$, that is, an optimal information structure is invariant among all payoff scale-up menus $F^{\times r}$. Though Dewan and Neligh [16] conclude these subjects have a fixed information structure and do not pay much attention, our payoff-dependent cost model $C(P, rP)$ suggests a further categorization among these subjects. Suppose that $C(P, rP)$ is a homogeneous payoff-dependent cost function. Since $C(P, rP) = rC(P, P)$,

$$\arg\max_P rP - C(P, rP) = \arg\max_P P - C(P, P),$$

that is, an optimal information choice $P(r)$ is independent of r. This observation means that some of the rationally inattentive subjects who are not responsive to incentives are not categorized to having a fixed information but are categorized to having homogeneous payoff-dependent costs.

To separate between these two types of subjects through observable behavior, we may conduct an additional experiment. Take some payoff $\overline{r} > 0$. Instead of payoff scale-up menus $F^{\times r}$, we consider a menu F^{+r} of acts f^ω, $\omega \in \Omega$, where $f^\omega(\omega) = \overline{r} + r$ and $f^\omega(\omega') = r$ for any $\omega' \neq \omega$. That is, the subject can obtain both a payoff $\overline{r}$ if his/her guess is correct and a payoff r irrespective of correctness of his/her guess. Thus, r serves as a minimum payoff. Such menus are referred to as *payoff translation*.

As shown below in Proposition 1, given a payoff translation menu, the homogeneous payoff-dependent cost model implies a decreasing $P(r)$, that is, the subject is more reluctant to acquire additional pieces of information if a miminum payoff r becomes larger, while the subject having a fixed information structure still exhibits a constant performance function $P(r)$.

Similarly, by using the same additional experiments, we can also separate between the responsive subjects for payoff scale-up; some of them are also responsive to payoff translation menus, while the other are not responsive to them. The former should have a general payoff-dependent cost such as the hybrid cost, as given by (6), whereas the latter is categorized into the class of the payoff-independent costs because this class of model satisfies translation invariance.

Proposition 1 *(1) In the homogeneous cost model, an optimal information choice is invariant for payoff scale-up, while negative responsive for payoff translation.*

(2) In the hybrid cost model, an optimal information choice is positive responsive for payoff scale-up and negative responsive for payoff translation.

Figure 1 summarizes the four classifications made possible by our additional experiments. Dewan and Neligh [16] separate responsive/non-responsive subjects through choice from payoff scale-up menus. By the additional experiment proposed here, each of their classification can be further subdivided to find the true proportion of subjects who do not respond to incentives and the proportion of subjects for whom hybrid costs are more fitting than payoff-independent costs.

Figure 1: Categorization of information costs

		Payoff translation F^{+r}	
		non-responsive	responsive
Payoff scale-up $F^{\times r}$	non-responsive	constant costs	homogeneous costs
	responsive	payoff-independent costs	hybrid costs

5.2 Elicitation bias from misspecification

In de Oliveira, Denti, Mihm, and Ozbek [15], the canonical cost function of a payoff-independent cost function representation is characterized by

$$c(\pi) = \sup_{F \in \mathbb{F}} \{b_F^u(\pi) - u(x_F)\} \tag{9}$$

for any $\pi \in \Pi(\overline{p})$. This formula suggests that an analyst can construct the canonical cost function of a payoff-independent cost function representation by eliciting the vNM utility index u and collecting willingness-to-pay data for $U(F) = u(x_F)$. We argue that this procedure of recovering the canonical cost function is specific to the payoff-independent cost model and has an upward bias if it is improperly applied for general payoff-dependent cost models.

First, for illustration, suppose that the agent is risk-neutral and the agent's true costs for information acquisition follow the homogeneous payoff-dependence as given by Section 4.2. For positive payoffs, the representation admits a discounted utility form $U(F) = \max_\pi \beta(\pi) b_F^u(\pi)$, where $\beta(\pi) = 1 - \gamma_+(\pi) \in [0, 1]$. Now assume that the analyst misspecifies it to be payoff-independent and tries to approximate it by the formula (9). It is common in the literature that incentives for information acquisition are measured by observing choices from payoff scale-up menus $F^{\times r}$ given as in Section 5.1.1 (see also Caplin and Dean [3] and Dewan and Neligh [16]). Then,

$$U(F^{\times r}) = \max_\pi \beta(\pi) b_{F^{\times r}}^u(\pi) = \max_\pi \beta(\pi) r b_{F^{\times 1}}^u(\pi) = rU(F^{\times 1}),$$

that is, U is homogeneous in r, and hence, we have $u(x_{F^{\times r}}) = ru(x_{F^{\times 1}})$ for all $r > 0$. If these data are improperly applied to compute the payoff-independent cost (9),

$$c(\pi) = \sup_F (b_F^u(\pi) - u(x_F)) \geq b_{F^{\times r}}^u(\pi) - u(x_{F^{\times r}}) = r(b_{F^{\times 1}}^u(\pi) - u(x_{F^{\times 1}})),$$

which implies $c(\pi) = \infty$ as $r \to \infty$. This observation suggests that the lower bound of $c(\pi)$, denoted by $\underline{c}(\pi)$, estimated through the series of elicitation $x_{F^{\times r}}$ for various r tends to be overestimated. Hence, if we apply the estimated cost function to out of sample prediction, we may have the following situation:

$$b_F^u(\pi) - \underline{c}(\pi) < b_F^u(\delta_{\overline{p}}) < \beta(\pi) b_F^u(\pi).$$

The first inequality states that in the misspecified model, acquiring information π is too costly compared with the initial prior, while the second inequality states that in the true model, the agent is willing to acquire this information.

The above observation is generalized in the following way. Suppose that an analyst misspecifies that $\succsim$ admits a payoff-independent cost representation, though the "true" representation of $\succsim$ follows a (potentially payoff-dependent) CSL Representation. We assume that the analyst elicits the vNM utility index u and collects willingness-to-pay data for $U(F) = u(x_F)$. By Theorem 1, The cost function for a CSL Representation is

$$C(\pi, t) = t - W(\pi, t) = t - \inf_{\left\{F \,|\, b_F^u(\pi) = t\right\}} u(x_F) \tag{10}$$

for any $\pi \in \Pi(\overline{p})$.

In general, $C(\pi, t)$ and $c(\pi)$ have the following relationship.

Proposition 2 *For all π,*

$$c(\pi) = \sup_t C(\pi, t).$$

Proof. Since $W(\pi, t) = \inf_{\{F \,|\, b_F^u(\pi) = t\}} u(x_F)$ by (10), we have that

$$\begin{aligned} c(\pi) &= \sup_{F \in \mathbb{F}} \{b_F^u(\pi) - u(x_F)\} = \sup_t \sup_{\{G \,|\, b_G^u(\pi) = t\}} \{b_G^u(\pi) - u(x_G)\} \\ &= \sup_t \sup_{\{G \,|\, b_G^u(\pi) = t\}} \{t - u(x_G)\} = \sup_t (t - \inf_{\{G \,|\, b_G^u(\pi) = t\}} u(x_G)) \\ &= \sup_t (t - W(\pi, t)) = \sup_t C(\pi, t). \end{aligned}$$

■

By Proposition 2, we have $c(\pi) \geq C(\pi, t)$ for all π and t. Thus, approximating $c(\pi)$ always has an upward bias compared with the true cost function $C(\pi, t)$. This suggests that costs for information acquisition may be overestimated due to improper assumptions about the payoff-independent costs.

5.3 Optimal sampling

Cukierman [11] investigates an optimal number of information acquisition before an investment decision is made, assuming payoff-independent costs for information acquisition. We adopt the same setting, assuming homogeneous costs, and examine its implications.

The state space Ω is taken to be the real line. The prior over Ω is given by a normal distribution $\omega \sim N(\mu, 1/\tau)$, where μ is the mean, and $\tau > 0$ is the precision. The signal s is correlated with ω according to a normal distribution $s \sim N(\omega, 1/\sigma)$, where $\sigma > 0$ is the precision of the signal.

The agent's payoff function is state-dependent and given by

$$u(y, \omega) = a\omega - b|\omega - y|, \ a > 0, \ b > 0.$$

A choice variable is y, which is interpreted as an investment decision. This payoff function takes its maximum at $y = \omega$, and the closer the investment decision is to the true state, the higher the payoff. Moreover, for all fixed y, higher states ω imply higher payoffs. As the payoffs change according to realization of ω, a choice of y is interpreted as a choice of act.

The agent can postpone the investment decision and instead observe signals, whereby, the prior is updated to a posterior according to Bayes' rule. If signals are observed for n times, the value of information is given by

$$b^u(n) = \int \max_y \int u(y, \omega) \, \mathrm{d}p(\omega | s_1, \cdots, s_n) \, \mathrm{d}\pi^n(s_1, \cdots, s_n), \tag{11}$$

where $p(\omega | s_1, \cdots, s_n)$ is a posterior conditional upon the realization of signals $s_1, \cdots, s_n$, and $\pi^n(s_1, \cdots, s_n)$ is an ex ante probability of the signal realization up to n. In this setting, an information structure is identified with a number of times for signal observations. A more informative signal structure is obtained by greater sample size. The set of information structures is given by $\Pi = \{\pi^n \,|\, n \geq 0\}$.

Cukierman [11] shows that (11) is written as

$$b^u(n) = a\mu - b\left(\frac{2}{\pi}\right)^{\frac{1}{2}} \left(\frac{1}{\tau + n\sigma}\right)^{\frac{1}{2}},$$

where π is the circular constant. Furthermore, to ensure that $b^u(n) > 0$ for all n, note that $b^u(n) > 0$ is equivalent to

$$n > \frac{1}{\sigma}\left(\frac{2b^2}{\pi(a\mu)^2} - \tau\right),$$

which is ensured if the right-hand side of this inequality is negative. Hence, throughout this subsection, we assume

$$\mu\tau^{\frac{1}{2}} > \frac{b}{a}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}. \tag{12}$$

We start with assuming the payoff-independent model as in Cukierman [11] and characterize an optimal simple size. The agent solves

$$\max_n \{b^u(n) - cn\},$$

where $c > 0$ is a constant marginal cost of sampling. For simplicity, let us treat n as a continuous variable. Then, the FOC is given by

$$\frac{\mathrm{d}b^u}{\mathrm{d}n}(n) = \frac{b\sigma}{2}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left(\frac{1}{\tau + n\sigma}\right)^{\frac{3}{2}} = c.$$

Clearly, if τ increases, $\frac{\mathrm{d}b^u}{\mathrm{d}n}(n)$ shifts down, and hence, the optimal sample size decreases. On the other hand, since the FOC is independent of μ, the mean of the prior has no impact on the optimal sample size. It is clear from $b^u(n)$, increasing μ means that the agent receives a higher expected payoff irrespective of the sample size. Since the reservation utility under no sampling goes up, the agent may be more reluctant to acquire signals. The payoff-independent cost model fails to capture such an intuition.

Now assume that the agent's preference is represented by the homogeneous payoff-dependent cost representation. Since $b^u(n) > 0$ for all n, (5) admits a discounted utility form, $U(F) = \max_n \beta(n)b^u_F(n)$, where $\beta(n) = 1 - \gamma_+(\pi^n) \in [0, 1]$. For simplicity, assume further that $\beta(n) = e^{-rn}$ with some $r > 0$. The agent solves an optimal sampling problem formulated as

$$\max_n e^{-rn}b^u(n).$$

From the FOC,

$$\frac{\frac{\mathrm{d}b^u}{\mathrm{d}n}(n)}{b^u(n)} = r. \tag{13}$$

Since $b^u(n) > 0$ and $\frac{\mathrm{d}^2 b^u}{\mathrm{d}n^2}(t) < 0$, $\frac{\frac{\mathrm{d}b^u}{\mathrm{d}n}(n)}{b^u(n)}$ is strictly decreasing. Thus, the SOC is satisfied, and (13) is a necessary and sufficient condition for an optimal sample size n.

Proposition 3 *Assume the homogeneous cost model and $\mu\tau^{\frac{1}{2}} > \frac{b}{a}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}$. The agent acquires more signals if either the precision, τ, or the mean, μ, of the prior over states decreases.*

The homogeneous cost model draws conclusions consistent with the intuition mentioned above. As in the payoff-independent cost model, as the accuracy of the prior decreases, the agent observes signals more often. In contrast to the payoff-independent cost model, as the mean of the prior increases, the agent observes signals less often.

Appendix

A Preliminaries

Following de Oliveira, Denti, Mihm, and Ozbek [15], we introduce some notions and mathematical preliminaries needed for the subsequent analysis. The proofs are omitted.

- $C(\Delta(\Omega))$ ($C_+(\Delta(\Omega))$): the set of all (non-negative) real-valued continuous functions over $\Delta(\Omega)$ with the supnorm
- $ca(\Delta(\Omega))$ ($ca_+(\Delta(\Omega))$): the set of all (non-negative) signed measures over $\Delta(\Omega)$ with the weak* topology
- For $\varphi \in C(\Delta(\Omega))$ and $\pi \in ca(\Delta(\Omega))$, define

$$\langle \varphi, \pi \rangle = \int_{\Delta(\Omega)} \varphi(p) \mathrm{d}\pi(p).$$

For a subset Ψ of $C(\Delta(\Omega))$, we say that a function $V : \Psi \to \mathbb{R}$ is *normalized* if $V(\alpha) = \alpha$ for each constant function $\alpha \in \Psi$; *monotone* if $V(\varphi) \geq V(\psi)$ for all $\varphi, \psi \in \Psi$ with $\varphi \geq \psi$; *convex* if $\alpha V(\varphi) + (1-\alpha)V(\psi) \geq V(\alpha\varphi + (1-\alpha)\psi)$ for all $\varphi, \psi \in \Psi$ and $\alpha \in (0,1)$; *quasi-convex* if $V(\varphi) \geq V(\alpha\varphi + (1-\alpha)\psi)$ for all $\varphi, \psi \in \Psi$ with $V(\varphi) \geq V(\psi)$ and $\alpha \in (0,1)$; *positively homogeneous* if $V(\alpha\varphi) = \alpha V(\varphi)$ for all $\varphi \in \Psi$ and $\alpha \geq 0$.

- Φ: the set of convex functions in $C(\Delta(\Omega))$
- For any expected utility function u and any menu $F \in \mathbb{F}$, φ_F is defined as in (3).
- $\Phi_{\mathbb{F}}(\Phi_{\mathcal{F}}, \Phi_X)$: the set of functions $\varphi_F(\varphi_{\{f\}}, \varphi_{\{x\}})$

Note that $u(X) = \Phi_X \subset \Phi_{\mathcal{F}} \subset \Phi_{\mathbb{F}} \subset \Phi$. Moreover, $\Phi_{\mathbb{F}}$ is convex because $\alpha\varphi_F + (1-\alpha)\varphi_G = \varphi_{\alpha F + (1-\alpha)G}$.

Assume that $u(X) = \mathbb{R}$. Then we have the following properties of $\Phi_{\mathbb{F}}$:

(i) $\Phi_{\mathbb{F}} + \mathbb{R} = \Phi_{\mathbb{F}}$

(ii) $\alpha\varphi_F \in \Phi_{\mathbb{F}}$ for every $\alpha \geq 0$

(iii) The set $\Phi_{\mathbb{F}}$ is dense in Φ.

B Proof of Theorem 1

B.1 Sufficiency

As explained in the proof sketch of Section 3.3, the first step is to induce a functional on $C(\Delta(\Omega))$ from $\succsim$ through the identification of menus F via support functions φ_F. Singleton Independence and the basic axioms ensure that there exist an expected utility function $u : X \to \mathbb{R}$ with unbounded range and a prior probability measure $\overline{p}$ over Ω such that the preference $\succsim$ over $\mathcal{F}$ is represented by the function $U : \mathcal{F} \to \mathbb{R}$ defined by

$$U(f) = \sum_{\Omega} u(f(\omega))\overline{p}(\omega).$$

de Oliveira, Denti, Mihm, and Ozbek [15, Claim 2] show that every menu F has a certainty equivalent $x_F \in X$ such that $\{x_F\} \sim F$. Hence, $U : \mathcal{F} \to \mathbb{R}$ is extended to $\mathbb{F}$ by $U(F) = U(x_F)$. Then, $U : \mathbb{F} \to \mathbb{R}$ represents $\succsim$.

Define the functional $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ by $V(\varphi_F) = U(F)$ as in de Oliveira, Denti, Mihm, and Ozbek [15]. They show that V is well-defined.

Lemma 1 *The functional $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is monotone, normalized, quasi-convex, and continuous.*

Proof. The first two properties follow from the same argument of de Oliveira, Denti, Mihm, and Ozbek [15, Claim 6].

To show quasi-convexity, we adapt the proof of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8, Lemma 56] by replacing the argument for quasi-concavity with that for quasi-convexity.

Claim 1 *$V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is quasi-convex.*

Proof. We want to show that for all F, G, $U(F) \geq U(G)$ implies $U(F) \geq U(\alpha G + (1-\alpha)F)$ for all $\alpha \in (0,1)$. If $F \sim G$, the desired result directly follows from Aversion to Contingent Planning. Then, we show that $F \succ G$ implies that $F \succsim \alpha G + (1-\alpha)F$ for all $\alpha \in (0,1)$. Suppose contrary that there exist $F \succ G$ and $\widetilde{\alpha} \in (0,1)$ such that $\widetilde{\alpha} G + (1-\widetilde{\alpha})F \succ F$. Note that $\widetilde{\alpha} \in \{\alpha \in [0,1] | \, \alpha G + (1-\alpha)F \succsim F\} \neq \emptyset$. By Mixture Continuity, this set is compact. Hence, we can find $\beta = \max\{\alpha \in [0,1] | \, \alpha G + (1-\alpha)F \succsim F\}$ and define $F_\beta = \beta G + (1-\beta)F$.

We claim that $F_\beta \sim F$. If $\beta = 1$, then $G \succsim F$, which contradicts $F \succ G$. Hence, $\beta < 1$. Now we show that $F_\beta \sim F$. Suppose contrary that $F_\beta \nsim F$, that is, $F \succ F_\beta$. Since $\{\alpha \in [0,1] | \, \alpha G + (1-\alpha)F \succ F\}$ is open, we can find an open set V such that $\beta \in V$ and $V \subset \{\alpha \in [0,1] | \, \alpha G + (1-\alpha)F \succ F\}$. Hence there exists $\beta' \in V$ such that $\beta' \in (\beta, 1)$ and $\beta' G + (1-\beta')F \succ F$. This contradicts the maximality of β. Hence, $F_\beta \sim F$.

Since $F_\beta \sim F$, Aversion to Contingent Planning implies that $F \succsim \lambda F_\beta + (1-\lambda)F$ for all $\lambda \in (0,1)$. Since $0 < \widetilde{\alpha} < \beta$, $\frac{\widetilde{\alpha}}{\beta} \in (0,1)$. Thus, $F \succsim \frac{\widetilde{\alpha}}{\beta} F_\beta + (1 - \frac{\widetilde{\alpha}}{\beta})F = \frac{\widetilde{\alpha}}{\beta}[\beta G + (1-\beta)F] + (1-\frac{\widetilde{\alpha}}{\beta})F = \widetilde{\alpha} G + (1-\widetilde{\alpha})F \succ F$, which is a contradiction. Hence, $F \succ G$ implies $F \succsim \alpha G + (1-\alpha)F$ for all $\alpha \in (0,1)$, as desired. ∎

We show continuity in a similar way to Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8, Lemma 42]. Let $\|\cdot\|$ be sup-norm. If $\{\varphi_n\}$ is a sequence in $\Phi_{\mathbb{F}}$, we write $\varphi_n \searrow \varphi$ if it is decreasing and it converges to φ in norm. The function $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is right continuous at $\varphi \in \Phi_{\mathbb{F}}$ if $\{\varphi_n\}_n \subseteq \Phi_{\mathbb{F}}$ and $\varphi_n \searrow \varphi$ implies $V(\varphi_n) \to V(\varphi)$. The function $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is upper semi-continuous if for any $\lambda > V(\varphi_F)$, there exists $\varepsilon > 0$ such that $\lambda > V(\varphi_{F'})$ for any $\varphi_{F'} \in \Phi_{\mathbb{F}}$ with $\|\varphi_F - \varphi_{F'}\| < \varepsilon$.

Similarly, if $\{\varphi_n\}$ is a sequence in $\Phi_{\mathbb{F}}$, we write $\varphi_n \nearrow \varphi$ if it is increasing and it converges to φ in norm. The function $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is left continuous at $\varphi \in \Phi_{\mathbb{F}}$ if $\{\varphi_n\}_n \subseteq \Phi_{\mathbb{F}}$ and $\varphi_n \nearrow \varphi$ implies $V(\varphi_n) \to V(\varphi)$. The function $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is lower semi-continuous if for any $\lambda < V(\varphi_F)$, there exists $\varepsilon > 0$ such that $\lambda < V(\varphi_{F'})$ for any $\varphi_{F'} \in \Phi_{\mathbb{F}}$ with $\|\varphi_F - \varphi_{F'}\| < \varepsilon$. The function V is continuous if it is upper and lower semi-continuous.

Claim 2 *$V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is continuous.*

Proof. Note that $\Phi_{\mathbb{F}}$ is convex. First, we show that V is upper semi-continuous.

Step 1: For any $F, G, H \in \mathbb{F}$, Mixture Continuity implies that the following sets are closed:

$$\begin{aligned}
\{\alpha \in [0,1] \,|\, \alpha F + (1-\alpha)G \succsim H\} &= \{\alpha \in [0,1] \,|\, U(\alpha F + (1-\alpha)G) \geq U(H)\} \\
&= \left\{\alpha \in [0,1] \,|\, V(\varphi_{\alpha F + (1-\alpha)G}) \geq V(\varphi_H)\right\} \\
&= \{\alpha \in [0,1] \,|\, V(\alpha\varphi_F + (1-\alpha)\varphi_G) \geq V(\varphi_H)\} \\
&= \{\alpha \in [0,1] \,|\, V(\alpha\varphi_F + (1-\alpha)\varphi_G) \geq \lambda\},
\end{aligned}$$

where $\lambda = V(\varphi_H)$.

Step 2: For any $\lambda \in \mathbb{R}$ and $\varphi, \varphi' \in \Phi_{\mathbb{F}}$ with $\varphi' \geq \varphi$ and $V(\varphi) < \lambda$, there exists $\alpha \in (0,1)$ such that $V(\alpha\varphi + (1-\alpha)\varphi') < \lambda$. Take such φ, φ', and λ. Suppose contrary that $V(\alpha\varphi + (1-\alpha)\varphi') \geq \lambda$ for all $\alpha \in (0,1)$. By Step 1, the set $A = \{\alpha \in [0,1] |\, V(\alpha\varphi + (1-\alpha)\varphi') \geq \lambda\}$ is closed. Since $(0,1) \subseteq A$, we have that $A = [0,1]$. This implies that $V(\varphi) \geq \lambda$, which is a contradiction.

Step 3: V is right continuous. Let $\varphi_n \searrow \overline{\varphi}$ such that $\{\varphi_n\}_{n\in\mathbb{N}} \cup \{\overline{\varphi}\} \subseteq \Phi_{\mathbb{F}}$. Monotonicity implies that $V(\varphi_n) \geq V(\varphi_{n+1}) \geq V(\overline{\varphi})$ for all $n \in \mathbb{N}$. Suppose contrary that $V(\varphi_n)$ does not converges to $V(\overline{\varphi})$, that is, there exists $\lambda \in \mathbb{R}$ such that $V(\varphi_n) \geq \lambda > V(\overline{\varphi})$ for all $n \in \mathbb{N}$. By Step 2, for each $\varphi \in \Phi_{\mathbb{F}}$ with $\varphi \geq \overline{\varphi}$, there exists $\alpha \in (0,1)$ such that $V((1-\alpha)\overline{\varphi} + \alpha\varphi) < \lambda$. Take $\varepsilon > 0$ such that $\overline{\varphi} + \varepsilon\mathbf{1} \in \Phi_{\mathbb{F}}$. Define $\varphi = \overline{\varphi} + \varepsilon\mathbf{1}$ and note that $\Phi_{\mathbb{F}} \ni (1-\alpha)\overline{\varphi} + \alpha\varphi = \overline{\varphi} - \alpha\overline{\varphi} + \alpha\overline{\varphi} + \alpha\varepsilon\mathbf{1} = \overline{\varphi} + \alpha\varepsilon\mathbf{1}$. Since $\varphi_n \searrow \overline{\varphi}$, there exists $\overline{n} \in \mathbb{N}$ such that $\varphi_n \leq \overline{\varphi} + \alpha\varepsilon\mathbf{1} = (1-\alpha)\overline{\varphi} + \alpha\varphi$ for all $n \geq \overline{n}$. Monotonicity implies that $V(\varphi_n) \leq V((1-\alpha)\overline{\varphi} + \alpha\varphi) < \lambda$ for all $n \geq \overline{n}$, which is a contradiction.

Step 4: The result. Let $\lambda \in \mathbb{R}$ and $S(V, \lambda) = \{\varphi \in \Phi_{\mathbb{F}} |\, V(\varphi_F) \geq \lambda\}$. We show that $\{\varphi_n\}_{n\in\mathbb{N}} \subseteq S(V,\lambda)$ and $\varphi_n \to \varphi \in \Phi_{\mathbb{F}}$ imply $\varphi \in S(V,\lambda)$. There exists $\overline{\varepsilon} > 0$ such that $\varphi + \varepsilon\mathbf{1} \in \Phi_{\mathbb{F}}$ for all $\varepsilon \in [0, \overline{\varepsilon}]$. Let $\varepsilon_m > 0$ be such that $\{\varepsilon_m\}_{m\in\mathbb{N}} \subseteq [0, \overline{\varepsilon}]$ and $\varepsilon_m \searrow 0$. Note that $\varphi + \varepsilon_m\mathbf{1} \in \Phi_{\mathbb{F}}$ for all $m \in \mathbb{N}$. Since $\varphi_n \to \varphi$, for all $m \in \mathbb{N}$ there exists n_m such that $\varphi + \varepsilon_m\mathbf{1} \geq \varphi_{n_m}$. Monotonicity implies that $V(\varphi + \varepsilon_m\mathbf{1}) \geq V(\varphi_{n_m}) \geq \lambda$. By right continuity, we have that $V(\varphi) = \lim_m V(\varphi + \varepsilon_m\mathbf{1}) \geq \lambda$.

Notice that the above proof goes through when $\succsim$ satisfies Mixture Continuity and V is monotone. Hence, by the symmetric argument, we can show that V is also lower semi-continuous. ■ ■

Define an extension of V to $C(\Delta(\Omega))$ by

$$V(\varphi) = \inf\{V(\varphi_F) |\, \varphi_F \in \Phi_{\mathbb{F}}, \varphi_F \geq \varphi\} \tag{14}$$

for all $\varphi \in C(\Delta(\Omega))$.

Lemma 2 *The functional $V : C(\Delta(\Omega)) \to \mathbb{R}$ is a well-defined extension of $V : \Phi_{\mathbb{F}} \to \mathbb{R}$.*

Proof. Take any $\varphi \in C(\Delta(\Omega))$. Since φ is a continuous function defined on a compact set $\Delta(\Omega)$, there exist $p^*, p_* \in \Delta(\Omega)$ such that $\varphi(p^*) \geq \varphi(p) \geq \varphi(p_*)$ for all $p \in \Delta(\Omega)$. Since $\Phi_{\mathbb{F}}$ is a cone including a constant function, $\alpha\mathbf{1} \in \Phi_{\mathbb{F}}$ for all $\alpha \in \mathbb{R}$, where $\mathbf{1} \in C(\Delta(\Omega))$ is the constant function that takes one for all coordinates. Then, for all $\alpha \geq \varphi(p^*)$, $\alpha\mathbf{1} \geq \varphi$. Therefore, $\{V(\varphi_F) | \varphi_F \in \Phi_{\mathbb{F}}, \varphi_F \geq \varphi\} \neq \emptyset$. Moreover, since $\varphi(p) \geq \varphi(p_*)$ for all p, $\varphi \geq \varphi(p_*)\mathbf{1}$. Thus, for every $\varphi_F \geq \varphi$, we have $\varphi_F \geq \varphi(p_*)\mathbf{1}$. By monotonicity of V on $\Phi_{\mathbb{F}}$, $V(\varphi_F) \geq V(\varphi(p_*)\mathbf{1}) = \varphi(p^*)$, that is, $\varphi(p^*)$ is a lower bound for the set. Thus, there exists an infimum, as desired.

To verify that this V is an extension, take any $\varphi_G \in \Phi_{\mathbb{F}}$. For all $\varphi_F \geq \varphi_G$, monotonicity of $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ implies $V(\varphi_F) \geq V(\varphi_G)$. That is, $V(\varphi_G)$ attains the infimum. Therefore, $V(\varphi_G) = \inf\{V(\varphi_F) | \varphi_F \in \Phi_{\mathbb{F}}, \varphi_F \geq \varphi_G\}$. ■

Lemma 3 *The functional $V : C(\Delta(\Omega)) \to \mathbb{R}$ is monotone, normalized, quasi-convex, and continuous.*

Proof. It is easy to see from the definition of V that V is monotone and normalized. The proofs of quasi-convexity and continuity are obtained by adopting the same argument of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8, Theorem 36].

Claim 3 *Suppose that $V : C(\Delta(\Omega)) \to \mathbb{R}$ is monotone. Then,*

$$\{\varphi \in C(\Delta(\Omega)) | V(\varphi) < \lambda\} = \{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + C_-(\Delta(\Omega)).$$

Proof. $\supseteq$: Take $\varphi = \varphi_F + \varphi_-$ such that φ_F with $V(\varphi_F) < \lambda$ and $\varphi_- \in C_-(\Delta(\Omega))$. Then, $\varphi \in C(\Delta(\Omega))$. Since $V : C(\Delta(\Omega)) \to \mathbb{R}$ is monotone, $V(\varphi) \leq V(\varphi_F) < \lambda$. Hence, $\varphi \in C(\Delta(\Omega))$ with $V(\varphi) < \lambda$.

$\subseteq$: Take $\varphi \in C(\Delta(\Omega))$ with $V(\varphi) < \lambda$. By the definition of infimum, for any $\varepsilon > 0$, there exists $\varphi_F \in \Phi_{\mathbb{F}}$ such that $\varphi_F \geq \varphi$ and $V(\varphi) + \varepsilon > V(\varphi_F)$. Fix $\varepsilon > 0$ such that $\lambda > V(\varphi) + \varepsilon$. Then, we can find $\varphi_F \in \Phi_{\mathbb{F}}$ such that $\varphi_F \geq \varphi$ and $\lambda > V(\varphi) + \varepsilon > V(\varphi_F)$. Since $\varphi_F \geq \varphi$, we have $\varphi_- = \varphi - \varphi_F \in C_-(\Delta(\Omega))$. Hence, we have $\varphi \in \{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + C_-(\Delta(\Omega))$. ■

Claim 4 *Suppose that $V : C(\Delta(\Omega)) \to \mathbb{R}$ is monotone. If $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is quasi-convex, $V : C(\Delta(\Omega)) \to \mathbb{R}$ is quasi-convex.*

Proof. We show that $\{\varphi \in C(\Delta(\Omega)) | V(\varphi) < \lambda\}$ is convex for any $\lambda \in \mathbb{R}$. Take $\varphi, \varphi' \in \{\varphi \in C(\Delta(\Omega)) | V(\varphi) < \lambda\}$. Then, by Claim 3, $\varphi = \varphi_F + \varphi_-$ and $\varphi' = \varphi_G + \varphi'_-$ such that $\varphi_F, \varphi_G \in \{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\}$ and $\varphi_-, \varphi'_- \in C_-(\Delta(\Omega))$. For any $\alpha \in (0, 1)$, $V(\alpha\varphi + (1-\alpha)\varphi') = V(\alpha(\varphi_F + \varphi_-) + (1-\alpha)(\varphi_G + \varphi'_-)) \leq V(\alpha\varphi_F + (1-\alpha)\varphi_G) < \lambda$, where the first inequality follows from monotonicity of V, and the second inequality follows from quasi-convexity of $V : \Phi_{\mathbb{F}} \to \mathbb{R}$. Hence, $\{\varphi \in C(\Delta(\Omega)) | V(\varphi) < \lambda\}$ is a convex set for any $\lambda \in \mathbb{R}$. ■

The function $V : C(\Delta(\Omega)) \to \mathbb{R}$ is upper semicontinuous if for any $\lambda > V(\varphi)$ with $\varphi \in C(\Delta(\Omega))$, there exists $\varepsilon > 0$ such that $\lambda > V(\varphi')$ for any $\varphi' \in C(\Delta(\Omega))$ with $\|\varphi - \varphi'\| < \varepsilon$. The function V is upper semi-continuous if and only if $\{\varphi' \in C(\Delta(\Omega)) | V(\varphi') < V(\varphi)\}$ is open. Similarly, the function V is lower semi-continuous if and only if $\{\varphi' \in C(\Delta(\Omega)) \,|\, V(\varphi') > V(\varphi)\}$ is open. The function V is continuous if it is upper and lower semi-continuous.

Claim 5 *The function $V : C(\Delta(\Omega)) \to \mathbb{R}$ is continuous.*

Proof. We show that the function $V : C(\Delta(\Omega)) \to \mathbb{R}$ is upper semi-continuous. This is because Claims 2 and 3 imply that $\{\varphi \in C(\Delta(\Omega)) | V(\varphi) < \lambda\} = \cup_{\varphi_- \in C_-(\Delta(\Omega))}[\{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + \{\varphi_-\}]$ is open. To show this, take $\varphi \in \cup_{\varphi_- \in C_-(\Delta(\Omega))}[\{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + \{\varphi_-\}]$. There exists $\varphi_- \in C_-(\Delta(\Omega))$ such that $\varphi \in \{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + \{\varphi_-\}$. Since $\{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + \{\varphi_-\}$ with $\varphi_- \in C_-(\Delta(\Omega))$ is open, there exists $\varepsilon > 0$ such that any $\varphi' \in C(\Delta(\Omega))$ with $\|\varphi - \varphi'\| < \varepsilon$ satisfies $\varphi' \in \{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + \{\varphi_-\} \subset \cup_{\varphi_- \in C_-(\Delta(\Omega))}[\{\varphi_F \in \Phi_{\mathbb{F}} | V(\varphi_F) < \lambda\} + \{\varphi_-\}]$.

By the symmetric argument, V is lower semi-continuous. ■ ■

For all $\pi \in ca_+(\Delta(\Omega))$ and $t \in \mathbb{R}$, define

$$\begin{aligned} B(\pi, t) &= \{\varphi \in C(\Delta(\Omega)) \,|\, \langle \varphi, \pi \rangle \geq t\}, \text{ and} \\ W(\pi, t) &= \inf_{\varphi \in B(\pi, t)} V(\varphi). \end{aligned} \tag{15}$$

Since all constant functions belong to $C(\Delta(\Omega))$, $B(\pi, t) \neq \emptyset$ for all π and t. Thus, $W(\pi, t) < \infty$ for all (π, t), but it is possible that $W(\pi, t) = -\infty$ for some (π, t).

Lemma 4 *For all $\pi \in ca_+(\Delta(\Omega))$, $t \in \mathbb{R}$, and $\alpha > 0$, the following hold:*

(1) $B(\pi, \alpha t) = \alpha B(\pi, t)$.

(2) $B(\alpha\pi, \alpha t) = B(\pi, t)$.

(3) $W(\alpha\pi, \alpha t) = W(\pi, t)$.

Proof. (1) Take any $\varphi \in B(\pi, \alpha t)$. By definition, $\langle \varphi, \pi \rangle \geq \alpha t$, which implies $\langle \varphi/\alpha, \pi \rangle \geq t$. Thus, $\varphi/\alpha \in B(\pi, t)$, or equivalently, $\varphi \in \alpha B(\pi, t)$. The converse is also true.
(2) This part follows from the definition of $B(\pi, t)$.
(3) This follows from part (2). ■

We show that V is rewritten as

$$V(\varphi) = \max_{\pi \in ca_+(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle),$$

which is a counterpart of the "uncertain averse representation"of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8] in our setting.

Lemma 5 *For all $\varphi \in C(\Delta(\Omega))$,*

$$V(\varphi) \geq \sup_{\pi \in ca_+(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle).$$

Proof. For every $\pi \in ca_+(\Delta(\Omega))$, we have $\varphi \in B(\pi, \langle \varphi, \pi \rangle)$. By the definition of W, we have $V(\varphi) \geq W(\pi, \langle \varphi, \pi \rangle)$ for any $\pi \in ca_+(\Delta(\Omega))$, and hence $V(\varphi) \geq \sup_{\pi \in ca_+(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle)$. ■

Lemma 6 *For all $\varphi \in C(\Delta(\Omega))$,*

$$V(\varphi) = \max_{\pi \in ca_+(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle).$$

Proof. We modify the proof in Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [6, Theorem 1] to our setup. We show that there exists $\widetilde{\pi} \in ca_+(\Delta(\Omega))$ such that $V(\varphi) = W(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle)$. Then, by Lemma 5, we have $V(\varphi) = \max_{\pi \in ca_+(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle)$. Let $SL(\varphi) = \{\varphi' \in C(\Delta(\Omega)) | \, V(\varphi') < V(\varphi)\} \neq \emptyset$. Since V is upper-semi continuous and quasi-convex, $SL(\varphi)$ is convex and open in $C(\Delta(\Omega))$. Since $\varphi \notin SL(\varphi)$, the separation hyperplane theorem ensures that there exists $\widetilde{\pi} \in ca(\Delta(\Omega))$ such that $\langle \varphi, \widetilde{\pi} \rangle > \langle \varphi', \widetilde{\pi} \rangle$ for all $\varphi' \in SL(\varphi)$.

We claim that this separating $\widetilde{\pi}$ belongs to $ca_+(\Delta(\Omega))$. Fix $\widetilde{\varphi} \in C_+(\Delta(\Omega))$ and $\varphi' \in SL(\varphi)$ arbitrarily. Since $\varphi' \geq \varphi' - n\widetilde{\varphi}$ for all $n \in \mathbb{N}$, the monotonicity of V implies $V(\varphi) > V(\varphi') \geq V(\varphi' - n\widetilde{\varphi})$ and, hence, $\varphi' - n\widetilde{\varphi} \in SL(\varphi)$ for all $n \in \mathbb{N}$. Then, we have that $\langle \varphi, \widetilde{\pi} \rangle > \langle \varphi', \widetilde{\pi} \rangle - n \langle \widetilde{\varphi}, \widetilde{\pi} \rangle$ for all $n \in \mathbb{N}$. Therefore, $\langle \widetilde{\varphi}, \widetilde{\pi} \rangle > \frac{1}{n} \left(\langle \varphi', \widetilde{\pi} \rangle - \langle \varphi, \widetilde{\pi} \rangle \right)$ for all $n \in \mathbb{N}$. This implies that $\langle \widetilde{\varphi}, \widetilde{\pi} \rangle \geq 0$ for any $\widetilde{\varphi} \in C_+(\Delta(\Omega))$. Since $\langle \cdot, \widetilde{\pi} \rangle$ is a positive linear functional, Riesz representation theorem implies that there exists a unique $\pi \in ca_+(\Delta(\Omega))$ representing such a positive linear functional. By the uniqueness property, we have $\pi = \widetilde{\pi} \in ca_+(\Delta(\Omega))$.

The property of the separating $\widetilde{\pi} \in ca_+(\Delta(\Omega))$ means that for all φ' with $V(\varphi') < V(\varphi)$, since $\langle \varphi, \widetilde{\pi} \rangle > \langle \varphi', \widetilde{\pi} \rangle$, we have $\varphi' \notin B(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle)$. By the contraposition, $\varphi' \in B(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle)$ implies that $V(\varphi') \geq V(\varphi)$. That is,

$$V(\varphi) = \inf_{\varphi' \in B(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle)} V(\varphi') = W(\widetilde{\pi}, \langle \varphi, \widetilde{\pi} \rangle).$$

■

By Lemmas 4 and 6, we conclude that

$$V(\varphi) = \max_{\pi \in \Delta(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle).$$

In particular, $\succsim$ is represented by

$$U(F) = V(\varphi_F) = \max_{\pi \in \Delta(\Delta(\Omega))} W(\pi, \langle \varphi_F, \pi \rangle) = \max_{\pi \in \Delta(\Delta(\Omega))} W(\pi, b^u_F(\pi)).$$

We show several properties of W.

Lemma 7 *(1) For any $\pi \in \Delta(\Delta(\Omega))$, $W(\pi, t)$ is nondecreasing in t.*

(2) $W(\pi, t)$ is quasi-concave in $(\pi, t) \in \Delta(\Delta(\Omega)) \times \mathbb{R}$.

(3) $W(\pi, t)$ is upper semi-continuous in $(\pi, t) \in \Delta(\Delta(\Omega)) \times \mathbb{R}$.

Proof. (1) Take t and t' with $t > t'$. Since $B(\pi, t) \subsetneq B(\pi, t')$, we have $W(\pi, t) \geq W(\pi, t')$.

(2) The proof follows from Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [6, Lemma 4]. Take any (π_i, t_i) for $i = 1, 2$ and $\alpha \in [0, 1]$. Let $\pi' = \alpha\pi_1 + (1-\alpha)\pi_2$ and $t' = \alpha t_1 + (1-\alpha)t_2$. Then,

$$B(\pi', t') \subset B(\pi_1, t_1) \cup B(\pi_2, t_2),$$

which implies

$$W(\pi', t') \geq \inf_{\varphi \in B(\pi_1, t_1) \cup B(\pi_2, t_2)} V(\varphi) = \min\{W(\pi_1, t_1), W(\pi_2, t_2)\},$$

which means that W is quasi-concave.

(3) The proof follows from Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [6, Lemma 5]. Take any $\overline{\pi} \in \Delta(\Delta(\Omega))$ and $\alpha, \bar{t} \in \mathbb{R}$ such that $W(\overline{\pi}, \bar{t}) < \alpha$. There exists $\varphi_0 \in C(\Delta(\Omega))$ such that $\langle \varphi_0, \overline{\pi} \rangle \geq \bar{t}$ and $V(\varphi_0) < \alpha$. The sequence $\varphi^n = \varphi_0 + \frac{1}{n}\mathbf{1}$ converges to φ_0 as $n \to \infty$. Since V is upper semi-continuous, there exists $\overline{n}$ such that $V(\varphi^{\overline{n}}) < \alpha$. Moreover,

$$\langle \varphi^{\overline{n}}, \overline{\pi} \rangle = \langle \varphi_0, \overline{\pi} \rangle + \frac{1}{\overline{n}} \langle \mathbf{1}, \overline{\pi} \rangle \geq \bar{t} + \frac{1}{\overline{n}}.$$

Note that the set

$$O = \{\pi \in \Delta(\Delta(\Omega)) \,|\, \langle \varphi^{\overline{n}}, \pi \rangle > \langle \varphi^{\overline{n}}, \overline{\pi} \rangle - \frac{1}{2\overline{n}}\}$$

is open in the topology induced by the weak* topology. It is easy to see that $O \times (-\infty, \bar{t} + \frac{1}{2\overline{n}})$ is an open neighborhood of $(\overline{\pi}, \bar{t})$. Moreover, for all $(\pi, t) \in O \times (-\infty, \bar{t} + \frac{1}{2\overline{n}})$, we have

$$\langle \varphi^{\overline{n}}, \pi \rangle > \langle \varphi^{\overline{n}}, \overline{\pi} \rangle - \frac{1}{2\overline{n}} \geq \bar{t} + \frac{1}{\overline{n}} - \frac{1}{2\overline{n}} = \bar{t} + \frac{1}{2\overline{n}} > t.$$

Hence, $W(\pi, t) \leq V(\varphi^{\overline{n}}) < \alpha$, and $W(\pi, t)$ is upper semi-continuous. ■

Lemma 8 *W is linearly continuous.*

Proof. As shown in Lemma 3, V is continuous on $C(\Delta(\Omega))$. Moreover, V is written as

$$V(\varphi) = \max_{\pi \in \Delta(\Delta(\Omega))} W(\pi, \langle \varphi, \pi \rangle),$$

which implies that W is linearly continuous. ■

Lemma 9 *For all $\pi \in \Delta(\Delta(\Omega))$ and $t \in \mathbb{R}$,*

$$W(\pi, t) = \inf_{\varphi_F \in B(\pi,t)\cap\Phi_{\mathbb{F}}} V(\varphi_F) = \inf_{\{F \,|\, b^u_F(\pi)\geq t\}} u(x_F) = \inf_{\{F \,|\, b^u_F(\pi)=t\}} u(x_F).$$

Proof. The second equality follows from definition. For the first equality, it is enough to show that

$$\inf_{\varphi \in B(\pi,t)} V(\varphi) = \inf_{\varphi_F \in B(\pi,t)\cap\Phi_{\mathbb{F}}} V(\varphi_F).$$

Since $B(\pi,t) \cap \Phi_{\mathbb{F}} \subset B(\pi,t)$, $\inf_{\varphi\in B(\pi,t)} V(\varphi) \leq \inf_{\varphi_F \in B(\pi,t)\cap\Phi_{\mathbb{F}}} V(\varphi_F)$. Hence, it is enough to show the converse.

For all $\varphi \in C(\Delta(\Omega))$, define $D(\varphi) = \{\varphi_F \in \Phi_{\mathbb{F}} \,|\, \varphi_F \geq \varphi\}$. Take any $\varepsilon > 0$. By definition of infimum, there exists $\varphi^\varepsilon \in B(\pi,t)$ such that $V(\varphi^\varepsilon) < \inf_{\varphi\in B(\pi,t)} V(\varphi) + \varepsilon$. By definition of V, $\inf_{\varphi_F \in D(\varphi^\varepsilon)} V(\varphi_F) < \inf_{\varphi\in B(\pi,t)} V(\varphi) + \varepsilon$. Again, by definition of infimum, there exists $\varphi^\varepsilon_F \in D(\varphi^\varepsilon)$ such that $V(\varphi^\varepsilon_F) < \inf_{\varphi\in B(\pi,t)} V(\varphi) + \varepsilon$. Moreover, since $\varphi^\varepsilon_F \geq \varphi^\varepsilon$ and $\varphi^\varepsilon \in B(\pi,t)$, we have $\langle \varphi^\varepsilon_F, \pi\rangle \geq \langle \varphi^\varepsilon, \pi\rangle \geq t$, that is, $\varphi^\varepsilon_F \in B(\pi,t) \cap \Phi_{\mathbb{F}}$. Consequently, for all $\varepsilon > 0$, we can find some $\varphi^\varepsilon_F \in B(\pi,t)\cap\Phi_{\mathbb{F}}$ such that

$$V(\varphi^\varepsilon_F) < \inf_{\varphi\in B(\pi,t)} V(\varphi) + \varepsilon.$$

By definition of infimum,

$$\inf_{\varphi_F \in B(\pi,t)\cap\Phi_{\mathbb{F}}} V(\varphi_F) < \inf_{\varphi\in B(\pi,t)} V(\varphi) + \varepsilon.$$

Therefore, we have the desired result as $\varepsilon \to 0$.

To show the last equality, take any menu F with $b^u_F(\pi) > t$. Let $\varepsilon = b^u_F(\pi) - t > 0$. For any act f, define f^ε as an act satisfying $u(f^\varepsilon(\omega)) = u(f(\omega)) - \varepsilon$. Since $u(X) = \mathbb{R}$, such an f^ε exists. Define $F^\varepsilon = \{f^\varepsilon \,|\, f \in F\}$. Note that Preference for Flexibility and Dominance imply $F \succsim F^\varepsilon$, that is, $u(x_F) \geq u(x_{F^\varepsilon})$. Moreover, $b^u_{F^\varepsilon}(\pi) = b^u_F(\pi) - \varepsilon = t$. We have shown that if $b^u_F(\pi) > t$, then there exists $\widetilde{F}$ satisfying $b^u_{\widetilde{F}}(\pi) = t$ and $u(x_{\widetilde{F}}) \leq u(x_F)$, which implies

$$\inf_{\{G \,|\, b^u_G(\pi)=t\}} u(x_G) \leq \inf_{\{G \,|\, b^u_G(\pi)>t\}} u(x_G).$$

Therefore,

$$\inf_{\{G \,|\, b^u_G(\pi)\geq t\}} u(x_G) = \inf_{\{G \,|\, b^u_G(\pi)=t\}} u(x_G),$$

as desired. ■

Lemma 10 $W(\delta_{\overline{p}}, t) = t$.

Proof. Take a lottery x whose value is $u(x) = t$. The representation implies

$$t = u(x) = V(\varphi_{\{x\}}) = \max_{\pi\in\Delta(\Delta(\Omega))} W(\pi, b^u_{\{x\}}(\pi)) = \max_{\pi\in\Delta(\Delta(\Omega))} W(\pi, t).$$

Thus, $t \geq W(\pi, t)$ for all $\pi \in \Delta(\Delta(\Omega))$.

It is enough to show that $W(\delta_{\overline{p}}, t) \geq t$. For all menus F,

$$\langle \varphi_F, \delta_{\overline{p}} \rangle = \max_{f \in F} \sum_{\omega} u(f(\omega))\overline{p}(\omega) = \sum_{\omega} u(f^F(\omega))\overline{p}(\omega),$$

where $f^F \in F$ is a maximizer. By monotonicity of V,

$$V(\varphi_F) \geq V(\varphi_{\{f^F\}}) = U(\{f^F\}) = \sum_{\omega} u(f^F(\omega))\overline{p}(\omega) = \langle \varphi_F, \delta_{\overline{p}} \rangle.$$

Thus, by Lemma 9,

$$W(\delta_{\overline{p}}, t) = \inf_{\varphi \in B(\delta_{\overline{p}}, t)} V(\varphi) = \inf_{\varphi_F \in B(\delta_{\overline{p}}, t) \cap \Phi_{\mathbb{F}}} V(\varphi_F) \geq \inf_{\varphi_F \in B(\delta_{\overline{p}}, t) \cap \Phi_{\mathbb{F}}} \langle \varphi_F, \delta_{\overline{p}} \rangle \geq t.$$

as desired. ■

Lemma 11 *If $\pi \trianglerighteq \rho$, $W(\pi, t) \leq W(\rho, t)$ for all t.*

Proof. If $\pi \trianglerighteq \rho$, $\langle \varphi_F, \pi \rangle \geq \langle \varphi_F, \rho \rangle$ for all menus F. Thus, $B(\rho, t) \cap \Phi_{\mathbb{F}} \subset B(\pi, t) \cap \Phi_{\mathbb{F}}$, which implies, together with Lemma 9,

$$W(\pi, t) = \inf_{\varphi_F \in B(\pi, t) \cap \Phi_{\mathbb{F}}} V(\varphi_F) \leq \inf_{\varphi_F \in B(\rho, t) \cap \Phi_{\mathbb{F}}} V(\varphi_F) = W(\rho, t).$$

■

Define

$$\Pi = \{\pi \in \Delta(\Delta(\Omega)) \,|\, W(\pi, t) > -\infty \text{ for some } t\}. \tag{16}$$

By Lemma 10, $\delta_{\overline{p}} \in \Pi$. In particular, $\Pi \neq \emptyset$. Since any $\pi \notin \Pi$ never achieves the maximum of W, the representation U is rewritten as

$$U(F) = V(\varphi_F) = \max_{\pi \in \Pi} W(\pi, \langle \varphi_F, \pi \rangle).$$

Finally, we show the Bayesian plausibility condition of Π. For all $\pi \in \Pi$, let

$$p^\pi = \int_{\Delta(\Omega)} p \,\mathrm{d}\pi(p) \in \Delta(\Omega).$$

Lemma 12 $\Pi \subset \Pi(\overline{p})$.

Proof. We show that for all $\pi \in \Pi$, $p^\pi = \overline{p}$. Seeking a contradiction, suppose that there exists $\pi^* \in \Pi$ such that $p^{\pi^*} \neq \overline{p}$. There exist ω and ω' such that $p^{\pi^*}_\omega > \overline{p}_\omega$ and $p^{\pi^*}_{\omega'} < \overline{p}_{\omega'}$.

By definition of Π, there exists some t^* such that $W(\pi^*, t^*) > -\infty$. Take any $a < W(\pi^*, t^*)$. Let f be a constant act satisfying $u(f(\omega)) = a$ for all ω. Let $\mathbf{a}$ denote the vector in $\mathbb{R}^\Omega$ which takes a value of a for all coordinates. Consider a vector

$$\widetilde{\mathbf{a}} = (a, \cdots, a, \widetilde{a}_\omega, a, \cdots, a, \widetilde{a}_{\omega'}, a, \cdots, a) \in \mathbb{R}^\Omega$$

which satisfies $\widetilde{a}_\omega = a + k$ and $\widetilde{a}_{\omega'} = a - \frac{\overline{p}_\omega}{\overline{p}_{\omega'}} k$ for some $k \in \mathbb{R}$. Since $\mathbf{a} \cdot \overline{p} = a = \widetilde{\mathbf{a}} \cdot \overline{p}$, the vector $\widetilde{\mathbf{a}}$ can be regarded as a utility act which is indifferent to a constant utility act $\mathbf{a}$ under the subjective expected utility with $\overline{p}$.

On the other hand,

$$\widetilde{\mathbf{a}} \cdot p^{\pi^*} = a + \frac{p_\omega^{\pi^*} \overline{p}_{\omega'} - p_{\omega'}^{\pi^*} \overline{p}_\omega}{\overline{p}_{\omega'}} k. \tag{17}$$

Since $p_\omega^{\pi^*} > \overline{p}_\omega$ and $p_{\omega'}^{\pi^*} < \overline{p}_{\omega'}$, the multiplier of k in (17) is positive. Since (17) is a positive linear function with respect to k, $\widetilde{\mathbf{a}} \cdot p^{\pi^*}$ varies across all the real numbers. By choosing k appropriately, we can set $\widetilde{\mathbf{a}} \cdot p^{\pi^*} = t^*$.

For this particular $\widetilde{\mathbf{a}}$, since $u(X) = \mathbb{R}$, we can find some $\widetilde{f} \in \mathcal{F}$ satisfying $u(\widetilde{f}) = \widetilde{\mathbf{a}}$. By construction, $U(\{\widetilde{f}\}) = \widetilde{\mathbf{a}} \cdot \overline{p} = a$. However, by assumption,

$$U(\{\widetilde{f}\}) = a < W(\pi^*, t^*) = W(\pi^*, \widetilde{\mathbf{a}} \cdot p^{\pi^*}) \leq \max_{\pi \in \Pi} W(\pi, \widetilde{\mathbf{a}} \cdot p^\pi) = \max_{\pi \in \Pi} W(\pi, \langle \varphi_{\{\widetilde{f}\}}, \pi \rangle),$$

which contradicts the representation. ■

By Lemma 12,

$$U(F) = V(\varphi_F) = \max_{\pi \in \Pi(\overline{p})} W(\pi, \langle \varphi_F, \pi \rangle)$$

is a CSL Representation.

B.2 Necessity

Let $\succsim$ be the preference U represents. We show that the axioms are satisfied. It is obvious that $\succsim$ is complete and transitive. Since $u(X) = \mathbb{R}$, $\succsim$ satisfies Two-Sided Unboundedness.

Mixture Continuity

Take any $F, G \in \mathbb{F}$ and $\alpha \in [0,1]$. From the representation,

$$\begin{aligned} U(\alpha F + (1-\alpha)G) &= \max_{\pi \in \Pi} W(\pi, b^u_{\alpha F + (1-\alpha)G}(\pi)) \\ &= \max_{\pi \in \Pi} W(\pi, \langle \alpha \varphi_F + (1-\alpha)\varphi_G, \pi \rangle). \end{aligned}$$

Since W is linearly continuous, $U(\alpha F + (1-\alpha)G)$ is continuous in α. Hence, $U(\alpha^n F + (1-\alpha^n)G) \to U(\alpha F + (1-\alpha)G)$ as $\alpha^n \to \alpha$, which implies Mixture Continuity of $\succsim$.

Preference for Flexibility

Take any F and G with $G \subset F$. We have $b^u_F(\pi) \geq b^u_G(\pi)$ for all π. Since $W(\pi, t)$ is non-decreasing in t,

$$U(F) = \max_{\pi \in \Pi} W(\pi, b^u_F(\pi)) \geq \max_{\pi \in \Pi} W(\pi, b^u_G(\pi)) = U(G),$$

which implies that $\succsim$ satisfies Preference for Flexibility.

Dominance

Take any F and g. Assume that there exists $f \in F$ such that $\{f(\omega)\} \succsim \{g(\omega)\}$ for all ω. Since $b^u_F(\pi) = b^u_{F\cup\{g\}}(\pi)$ for all π,

$$U(F) = \max_{\pi\in\Pi} W(\pi, b^u_F(\pi)) = \max_{\pi\in\Pi} W(\pi, b^u_{F\cup\{g\}}(\pi)) = U(F \cup \{g\}).$$

Thus, Dominance holds.

Singleton Independence

For any $f \in \mathcal{F}$ and $\pi \in \Pi$, we have

$$b^u_{\{f\}}(\pi) = \sum_\Omega u(f(\omega))p^\pi(\omega) = \sum_\Omega u(f(\omega))\overline{p}(\omega).$$

Since $\pi \trianglerighteq \delta_{\overline{p}}$ for all $\pi \in \Pi$, we have $t = W(\delta_{\overline{p}}, t) = \max_{\pi\in\Pi} W(\pi, t)$, which implies

$$U(\{f\}) = \max_{\pi\in\Pi} W(\pi, b^u_{\{f\}}(\pi)) = \max_{\pi\in\Pi} W(\pi, \sum_\Omega u(f(\omega))\overline{p}(\omega)) = \sum_\Omega u(f(\omega))\overline{p}(\omega).$$

Since $U(\{f\})$ is a subjective expected utility function, it satisfies Singleton Independence.

Aversion to Contingent Planning

The proof follows from Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [8, Lemma 47]. It suffices to show that $V(\varphi) = \sup_{\pi\in\Pi} W(\pi, \langle\varphi, \pi\rangle)$, which is defined on $\Delta(\Delta(\Omega))$, is quasi-convex. In fact, if this is the case, for any F, G, and $\alpha \in (0, 1)$,

$$\begin{aligned} U(\alpha F + (1-\alpha)G) &= \sup_{\pi\in\Pi} W(\pi, \langle\varphi_{\alpha F+(1-\alpha)G}, \pi\rangle) \\ &= \sup_{\pi\in\Pi} W(\pi, \langle\alpha\varphi_F + (1-\alpha)\varphi_G, \pi\rangle) \\ &\le \max\{\sup_{\pi\in\Pi} W(\pi, \langle\varphi_F, \pi\rangle), \sup_{\pi\in\Pi} W(\pi, \langle\varphi_G, \pi\rangle)\} \\ &= \max\{U(F), U(G)\}. \end{aligned}$$

Now take any $t \in \mathbb{R}$. We want to show that $\{\varphi \,|\, V(\varphi) \le t\}$ is convex. Define

$$L = \bigcap_{\{(\pi,t')\in\Pi\times\mathbb{R} \,|\, \{\varphi \,|\, \langle\varphi,\pi\rangle<t'\}\supset\{\varphi \,|\, V(\varphi)\le t\}\}} \{\varphi \,|\, \langle\varphi, \pi\rangle < t'\}.$$

Note that L is a convex set because it is the intersection of a family of open half spaces. Moreover, by definition, $\{\varphi \,|\, V(\varphi) \le t\} \subset L$. We will show the converse, whereby establishing $L = \{\varphi \,|\, V(\varphi) \le t\}$, and hence, $\{\varphi \,|\, V(\varphi) \le t\}$ is convex, as desired.

Take any $\overline{\varphi} \notin \{\varphi \,|\, V(\varphi) \leq t\}$. Then, $t < V(\overline{\varphi}) = \max_{\pi \in \Pi} W(\pi, \langle \overline{\varphi}, \pi \rangle)$. There exists $\overline{\pi} \in \Pi$ such that $t < W(\overline{\pi}, \langle \overline{\varphi}, \overline{\pi} \rangle)$. For any φ with $\langle \varphi, \overline{\pi} \rangle \geq \langle \overline{\varphi}, \overline{\pi} \rangle$, we have

$$t < W(\overline{\pi}, \langle \overline{\varphi}, \overline{\pi} \rangle) \leq W(\overline{\pi}, \langle \varphi, \overline{\pi} \rangle) \leq V(\varphi).$$

That is,

$$\{\varphi \,|\, \langle \varphi, \overline{\pi} \rangle \geq \langle \overline{\varphi}, \overline{\pi} \rangle\} \subset \{\varphi \,|\, V(\varphi) > t\},$$

or equivalently,

$$\{\varphi \,|\, \langle \varphi, \overline{\pi} \rangle < \langle \overline{\varphi}, \overline{\pi} \rangle\} \supset \{\varphi \,|\, V(\varphi) \leq t\}.$$

Since $\overline{\varphi} \notin \{\varphi \,|\, \langle \varphi, \overline{\pi} \rangle < \langle \overline{\varphi}, \overline{\pi} \rangle\}$, by definition of L, $\overline{\varphi} \notin L$. Therefore, $L \subset \{\varphi \,|\, V(\varphi) \leq t\}$, as desired.

C Proof of Theorem 3

(a) $\Longrightarrow$ (b): It is easy to see that if $\succsim_1$ is more averse to commitment than $\succsim_2$, then for all $f, g \in \mathcal{F}$, $\{f\} \succsim_1 \{g\}$ if and only if $\{f\} \succsim_2 \{g\}$, that is, the two preferences are identical on singleton menus. Thus, by Anscombe and Aumann [1], there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u_1 = \alpha u_2 + \beta$ and $\overline{p}_1 = \overline{p}_2$. Since $\succsim_1$ is more averse to commitment than $\succsim_2$, for any $F \in \mathbb{F}$ and $f \in \mathcal{F}$, $F \sim_2 \{f\}$ implies $F \succsim_1 \{f\}$. Let $x_F^1, x_F^2 \in X$ be $\{x_F^1\} \sim_1 F$ and $\{x_F^2\} \sim_2 F$. Hence, $\{x_F^1\} \sim_1 F \succsim_1 \{x_F^2\}$, which implies $u_1(x_F^1) \geq u_1(x_F^2)$ for any $F \in \mathbb{F}$. Since $W_i(\pi, t) = \inf_{\{F \,|\, b_F^{u_i}(\pi) \geq t\}} u_i(x_F^i)$ for $i = 1, 2$,

$$\begin{aligned} W_1(\pi, t) &= \inf_{\{F \,|\, b_F^{u_1}(\pi) \geq t\}} u_1(x_F^1) \geq \inf_{\{F \,|\, b_F^{u_1}(\pi) \geq t\}} u_1(x_F^2) = \inf_{\{F \,|\, b_F^{u_2}(\pi) \geq \frac{t-\beta}{\alpha}\}} \alpha u_2(x_F^2) + \beta \\ &= \alpha W_2\left(\pi, \frac{t-\beta}{\alpha}\right) + \beta \end{aligned}$$

for any (π, t).

(b) $\Longrightarrow$ (a): Assume that $u_1 = \alpha u_2 + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$. For any $F \in \mathbb{F}$ and $f \in \mathcal{F}$, $F \succsim_2 \{f\}$ implies that

$$\max_{\pi \in \Pi(\overline{p})} W_2(\pi, b_F^{u_2}(\pi)) \geq \max_{\pi \in \Pi(\overline{p})} W_2(\pi, b_{\{f\}}^{u_2}(\pi)) = \sum_{\omega} u_2(f(\omega)) \overline{p}(\omega).$$

Since $W_1(\pi, t) \geq \alpha W_2(\pi, \frac{t-\beta}{\alpha}) + \beta$ for all $(\pi, t) \in \Pi(\overline{p}) \times \mathbb{R}$,

$$\begin{aligned} &\max_{\pi \in \Pi(\overline{p})} W_1(\pi, b_F^{u_1}(\pi)) \geq \alpha \max_{\pi \in \Pi(\overline{p})} W_2\left(\pi, \frac{b_F^{\alpha u_2 + \beta}(\pi) - \beta}{\alpha}\right) + \beta = \alpha \max_{\pi \in \Pi(\overline{p})} W_2(\pi, b_F^{u_2}(\pi)) + \beta \\ &\geq \alpha \sum_{\omega} u_2(f(\omega)) \overline{p}(\omega) + \beta = \sum_{\omega} u_1(f(\omega)) \overline{p}(\omega). \end{aligned}$$

Hence, we have that $F \succsim_1 \{f\}$.

D Proof of Corollary 2

We define $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ as in the proof of Theorem 1. By Lemma 1, it is normalized and monotone.

Lemma 13 *Under Increasing Desire for Commitment and Singleton Independence, $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is convex.*

Proof. Let $\{x_F\}$ be a lottery equivalent of F, that is, $\{x_F\} \sim F$. The existence of a lottery equivalent is guaranteed under Order, Continuity, Monotonicity, and Dominance. By Singleton Independence, $V(\alpha\{x_F\} + (1-\alpha)\{x_G\}) = \alpha V(\{x_F\}) + (1-\alpha)V(\{x_G\}) = \alpha V(F) + (1-\alpha)V(G)$. Hence, $\alpha V(F) + (1-\alpha)V(G) \geq V(\alpha F + (1-\alpha)G)$. ■

As $\Phi_{\mathbb{F}} = \Phi_{\mathbb{F}} + \mathbb{R}$, Pennesi [32, Proposition 2] implies that $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is translation invariant. This implies that $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is monotone, normalized, convex, and translation invariant. Hence, de Oliveira, Denti, Mihm, and Ozbek [15, Claim 7] implies that $\succsim$ is represented by a payoff-independent cost representation.

E Proof of Proposition 1

(1) Since we are assuming only positive payoffs, from (5), for any menu F, the homogeneous cost model is written as

$$U(F) = \max_{\pi \in \Pi} \beta(\pi) b^u_F(\pi), \tag{18}$$

where $\beta(\pi) = 1 - \gamma_+(\pi) \in [0, 1]$. Moreover, note that

$$b^u_{F^{+r}}(\pi_Q) = \int \max_{f \in F^{+r}} u(f) \cdot p \, \mathrm{d}\pi_Q(p) = \bar{r}\frac{1}{n}\sum_i q_{i,i} + r = \bar{r}P + r, \text{ where } P = \frac{1}{n}\sum_i q_{i,i}.$$

Hence, (18) is reduced to

$$U(F^{+r}) = \max_{P \in [0,1]} \beta(P)(\bar{r}P + r).$$

A proof is a simple application of the monotone comparative statics (see Milgrom and Shannon [31]). It is enough to show that

$$f(P, r) = \beta(P)(\bar{r}P + r)$$

is single-crossing, that is, for any $P < P'$ and $r' > r$, if $f(P, r) \geq f(P', r)$, then $f(P, r') \geq f(P', r')$, and if $f(P, r) > f(P', r)$, then $f(P, r') > f(P', r')$. By rearrangement,

$$\begin{aligned} & f(P, r) \geq f(P', r) \\ \iff & \beta(P)(\bar{r}P + r) \geq \beta(P')(\bar{r}P' + r) \\ \iff & r(\beta(P) - \beta(P')) \geq \bar{r}(\beta(P')P' - \beta(P)P). \end{aligned}$$

Since $\beta(P) > \beta(P')$ by the Blackwell monotonicity of the cost function, we have $r'(\beta(P) - \beta(P')) \geq r(\beta(P) - \beta(P'))$, which in turn implies $f(P, r') \geq f(P', r')$ by the same rearrangement as above. By a similar argument, the case of strict inequality also holds.

(2) As in part (1), the hybrid cost model is written as

$$U(F) = \max_{\pi \in \Pi} \beta(\pi) b_F^u(\pi) - c(\pi).$$

In particular, for the cases of payoff scale-up and payoff translation,

$$U(F^{\times r}) = \max_{P \in [0,1]} \beta(P) r P - C(P),$$

and

$$U(F^{+r}) = \max_{P \in [0,1]} \beta(P)(\overline{r}P + r) - C(P),$$

respectively. As in part (1), it is enough to show that both $g(P, r) = \beta(P) r P - C(P)$ and $f(P, r) = \beta(P)(\overline{r}P + r) - C(P)$ are single-crossing.

For any $P < P'$ and $r' > r$,

$$\begin{aligned} & f(P, r) \geq f(P', r) \\ \iff & \beta(P)(\overline{r}P + r) - C(P) \geq \beta(P')(\overline{r}P' + r) - C(P') \\ \iff & r(\beta(P) - \beta(P')) \geq \overline{r}(\beta(P')P' - \beta(P)P) + C(P) - C(P'). \end{aligned}$$

Since $\beta(P) > \beta(P')$ by the Blackwell monotonicity of the cost function, we have $r'(\beta(P) - \beta(P')) \geq r(\beta(P) - \beta(P'))$, which in turn implies $f(P, r') \geq f(P', r')$ by the same rearrangement as above.

Similarly, for any $P' > P$ and $r' > r$,

$$\begin{aligned} & g(P', r) \geq g(P, r) \\ \iff & \beta(P') r P' - C(P') \geq \beta(P) r P - C(P) \\ \iff & r(\beta(P')P' - \beta(P)P) \geq C(P') - C(P). \end{aligned}$$

Since $C(P') > C(P)$ by the Blackwell monotonicity of the cost function, the right-hand side of the last inequality is positive, and hence, so is the left-hand side. Since $r' > r > 0$, we have $r'(\beta(P')P' - \beta(P)P) \geq r(\beta(P')P' - \beta(P)P)$, which in turn implies $g(P', r') \geq g(P, r')$ by the same rearrangement as above.

F Proof of Proposition 3

Note that the left-hand side of (13) is explicitly written as

$$\frac{\frac{b\sigma}{2}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left(\frac{1}{\tau+n\sigma}\right)^{\frac{3}{2}}}{a\mu - b\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left(\frac{1}{\tau+n\sigma}\right)^{\frac{1}{2}}}$$

or

$$f(x) = \frac{\frac{b\sigma}{2}\left(\frac{2}{\pi}\right)^{\frac{1}{2}} x^{\frac{3}{2}}}{a\mu - b\left(\frac{2}{\pi}\right)^{\frac{1}{2}} x^{\frac{1}{2}}}, \text{ where } x = \frac{1}{\tau + n\sigma}.$$

By taking a derivative,

$$f'(x) = \frac{\sigma}{2}\frac{\frac{3}{2}Ax^{\frac{1}{2}} - x}{(A - x^{\frac{1}{2}})^2},$$

where $A = a\mu/(b\left(\frac{2}{\pi}\right)^{\frac{1}{2}})$. Hence, f is strictly increasing if and only if $\frac{3}{2}Ax^{\frac{1}{2}} - x > 0$ for all x, or

$$\begin{aligned}
\frac{3}{2}Ax^{\frac{1}{2}} - x > 0 &\iff \frac{3}{2}\mu > \frac{b}{a}\left(\frac{2}{\pi}\right)^{\frac{1}{2}} x^{\frac{1}{2}} \\
&\iff \mu(\tau + n\sigma)^{\frac{1}{2}} > \frac{2}{3}\frac{b}{a}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}, \ \forall n \geq 0.
\end{aligned}$$

The last condition holds if

$$\mu\tau^{\frac{1}{2}} > \frac{2}{3}\frac{b}{a}\left(\frac{2}{\pi}\right)^{\frac{1}{2}},$$

which is ensured by the assumption. Therefore, f is strictly increasing.

Note that the FOC is $f(\frac{1}{\tau+n\sigma}) = r$. If τ decreases, $f(\frac{1}{\tau+n\sigma})$ moves upwards. By the FOC, the agent will acquire more signals. Similarly, if μ decreases, then f moves upwards. By the FOC, the agent will observe signals more often.

References

[1] Anscombe, F. J. and R. J. Aumann (1963), “A definition of subjective probability, ”*Annals of Mathematical Statistics*, 34, 199-205.

[2] Aoyama, T. (2020), “Response time and subjective filtration, ”working paper.

[3] Caplin, A., and M. Dean (2014), “Revealed preference, rational inattention, and costly information acquisition, NBER Working Paper Series 19876.

[4] Caplin, A., and M. Dean (2015), “Revealed preference, rational inattention, and costly information acquisition, ”*American Economic Review*, 105, 2183-2203.

[5] Caplin, A., M. Dean, and J. Leahy (2019), “Rationally inattentive behavior: Characterizing and generalizing Shannon entropy, ”forthcoming in *Journal of Political Economy*.

[6] Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2008), “Complete Monotone Quasiconcave Duality, ”working paper.

[7] Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011), "Complete Monotone Quasiconcave Duality, "*Mathematics of Operations Research*, 36, 321-339.

[8] Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011), "Uncertainty averse preferences, "*Journal of Economic Theory*, 146, 1275-1330.

[9] Chambers, C. P., C. Liu, and J. Rehbeck (2020), "Costly information acquisition, "*Journal of Economic Theory*, 186, 104979.

[10] Chateauneuf, A., and J. H. Faro (2009), "Ambiguity through confidence functions, "*Journal of Mathematical Economics*, 45, 535-558.

[11] Cukierman, A. (1980), "The effects of uncertainty on investment under risk neutrality with endogenous information, "*Journal of Political Economy*, 88, 462-475.

[12] Dekel, E., B. Lipman, and A. Rustichini (2001), "Representing preferences with a unique subjective state space," *Econometrica*, 69, 891-934.

[13] Denti, T. (2022), "Posterior-separable cost of information, "forthcoming in *American Economic Review*.

[14] de Oliveira, H. (2019), "Axiomatic foundations for entropic cost of attention, "working paper.

[15] de Oliveira, H., T. Denti, M. Mihm, and K. Ozbek (2017), "Rationally inattentive preferences and hidden information costs, "*Theoretical Economics*, 12, 621-654.

[16] Dewan, A., and N. Neligh (2020), "Estimating information cost functions in models of rational inattention, "*Journal of Economic Theory*, 187, 105011.

[17] Dillenberger, D., R. V. Krishna, and P. Sadowski (2020), "Subjective information choice processes, "forthcoming in *Theoretical Economics*.

[18] Dillenberger, D., R. V. Krishna, and P. Sadowski (2020), "Information-dependent utilities and beliefs, "working paper.

[19] Dillenberger, D., J. S. Lleras, P. Sadowski, and N. Takeoka (2014), "A theory of subjective learning, "*Journal of Economic Theory*, 153, 287-312.

[20] Duraj, J., and Y.-H. Lin (2021), "Costly information and random choice, "forthcoming in *Economic Theory*.

[21] Ellis, A. (2018), "Foundations for optimal inattention, "*Journal of Economic Theory*, 173, 56-94.

[22] Ergin, H., and T. Sarver (2010), "A unique costly contemplation representation, "*Econometrica*, 78, 1285-1339.

[23] Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg (2006), "Costly information acquisition: Experimental analysis of a bounded rational model, "*American Economic Review*, 96, 1043-1068.

[24] Higashi, Y., K. Hyogo, and N. Takeoka (2022), "Information acquisition under homogeneous payoff-dependent costs, "working paper.

[25] Hyogo, K. (2007), "A subjective model of experimentation, "*Journal of Economic Theory*, 133, 316-330.

[26] Kamenica, E., and M. Gentzkow (2011), "Bayesian Persuasion, "*American Economic Review*, 101, 2590-2615.

[27] Lin, Y.-H. (2022), "Stochastic choice and rational inattention, "*Journal of Economic Theory*, 202, 105450.

[28] Lu, J. (2016), "Random choice and private information, "*Econometrica*, 84, 1983-2027.

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

[30] Mihm, M., and K. Ozbek (2018), "Mood-driven choices and self-regulation, "*Journal of Economic Theory*, 176, 727-760.

[31] Milgrom, P., and C. Shannon (1994), "Monotone comparative statics, "*Econometrica*, 62, 157-180.

[32] Pennesi, D. (2015), "Costly information acquisition and the temporal resolution of uncertainty, "*Journal of Mathematical Economics*, 60, 115-122.

[33] Sims, C. A. (2003), "Implications of rational inattention, "*Journal of Monetary Economics*, 50, 665–690.

[34] Strazalecki, T. (2013), "Temporal resolution of uncertainty and recursive models of ambiguity aversion, "*Econometrica*, 81, 1039-1074.