Information Acquisition under Homogeneous Payoff-Dependent Costs*

Youichiro Higashi, Kazuya Hyogo, and Norio Takeoka †

June 23, 2022

Abstract

By generalizing a result of de Oliveira, Denti, Mihm, and Ozbek [4], Higashi, Hyogo, and Takeoka [5] provide a choice theoretic foundation for information acquisition under general payoff-dependent costs, which serves as a test focusing solely on the essence of information acquisition without relying on particular features such as the additivity of the cost function. In the present study, we axiomatize a special case where costs for information acquisition are proportional to benefits of information. When payoffs are positive, this representation is rewritten as a discounted utility form.

Keywords: costly information acquisition, rational inattention, homogeneous costs, preference for flexibility, preference over menus.

JEL classification: D11, D81, D91.

*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

Rational inattention is a model where rational agents will optimally acquire information by considering the benefits and costs of information acquisition. Since a seminal work of Sims [8], the implications of this model has been studied in much of the literature. 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, but is independent of the benefit from the experiment. However, in some instances, the cost of information acquisition may be payoff-dependent.

Higashi, Hyogo, and Takeoka [5] (henceforth HHT) axiomatize a general payoff-dependent cost function for information acquisition. To see their result formally, consider the following choice environment: let Ω be a finite set of objective states and X be a set of lotteries. A function $f : \Omega \to X$ is called an act. We consider a finite subset F of acts, called a menu, as a choice object. That is, we take preference $\succsim$ over those menus as primitive.

Suppose that the agent has an expected utility function $u : X \to \mathbb{R}$ and an initial prior $\overline{p}$ over Ω. Before choosing from a menu F, the agent may conduct an additional experiment or engage in information acquisition, which generates signals about states. The agent updates his prior and makes a choice from the menu contingent upon posteriors. Formally, information acquisition is interpreted as a choice of an information structure $\pi \in \Delta(\Delta(\Omega))$, whose prior coincides with $\overline{p}$.

Given each menu F, the benefit of information of π is defined as

$$b_F^u(\pi) \equiv \int \left(\max_{f \in F} \sum_{\omega} u(f(\omega))p(\omega) \right) \mathrm{d}\pi(p). \tag{1}$$

After choosing the information structure, the agent observes a signal and updates his prior belief to the posterior p. Given the posterior belief p, the agent chooses an act f from a menu F to maximize the expected utility. The benefit of information is computed as the expectation of these maximum values with respect to the distribution over signals, given by π.

HHT propose the following representation: A preference $\succsim$ over menus admits a *costly subjective learning (CSL)* representation if there exist an expected utility function $u : X \to \mathbb{R}$, a prior belief $\overline{p}$ over Ω, a function $W(\pi, t)$, interpreted as a net benefit function of information acquisition, such that

$$U(F) = \max_{\pi \in \Pi(\overline{p})} W(\pi, b_F^u(\pi)) \tag{2}$$

represents $\succsim$, where $\Pi(\overline{p})$ is the set of information structures consistent with the prior. This representation is given as an indirect utility function of the maximization, where the agent optimally chooses an information structure by considering the benefits and costs of acquiring information. The function W satisfies several properties, which justify an interpretation of W being a net benefit of information acquisition.

In the CSL representation, costs for information acquisition are implicitly incorporated into W. To formulate this cost explicitly, (2) is rewritten as

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b_F^u(\pi) - C(\pi, b_F^u(\pi))\}, \tag{3}$$

where $C(\pi, t) := t - W(\pi, t)$. That is, (3) is regarded as a model of a general payoff-dependent cost function. If $C(\pi, t)$ is payoff-independent and written as $c(\pi)$, the CSL model is reduced to the standard model for information acquisition,

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b_F^u(\pi) - c(\pi)\}, \tag{4}$$

which is axiomatized by de Oliveira, Denti, Mihm, and Ozbek [4].

The purpose of this paper is to axiomatize an alternative specification of the payoff-dependent cost representation (3). A preference $\succsim$ over menus admits a homogeneous cost representation if there exist a tuple $(u, \overline{p}, \gamma_s)$, where $u : X \to \mathbb{R}$ is an unbounded expected utility function with $u(X) = \mathbb{R}$, $\overline{p}$ is the initial prior, $\gamma_s : \Pi(\overline{p}) \to \mathbb{R}_+$ is a payment rate function depending on s =sign of $b_F^u(\pi)$ such that $\succsim$ is represented by

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

Moreover, $\gamma_+(\pi) \in [0, 1]$. It is easy to see that this is a special case of (3) in that $C(\pi, t) = \gamma_s(\pi)|t|$ where s is the sign of t. In this model, a cost for information acquisition depends not only on an information structure but also on the benefit of information proportionally. In particular, when payoffs are positive, (5) is reduced to the discounted utility form:

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

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

To axiomatize (5), we borrow techniques from the literature on choice under ambiguity. The CSL representation is a counterpart of the uncertain averse representation of Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio [2], which nests two representations as special cases. One is the variational representation of Maccheroni, Marinacci, and Rustichini [7], which satisfies the property, called translation invariance, and has a parallel relationship with (4). The other is the confidence representation of Chateauneuf and Faro [3], which satisfies homotheticity and has a parallel relationship with the homogeneous cost representation (5).

1.1 Motivating example

We illustrate a type of behavior that cannot be explained by the payoff-independent cost model, but can be explained by the homogeneous cost model. For simplicity, assume that the objective state space is given by $\Omega = \{\omega_1, \omega_2\}$. In this illustration, an act is a function defined on Ω, which pays a positive payoff according to a realization of states. Hence, an act is identified with (x_1, x_2), where $x_i \in \mathbb{R}$ is a payoff when ω_i is realized, and is interpreted as an investment. Payoffs are interpreted as either utils or monetary prizes, assuming the agent is risk neutral.

Imagine an agent facing a decision about investment opportunities, which are considered as menus of acts. The agent has preference over menus of acts.

Suppose that

$$\{(100, 0)\} \sim \{(0, 100)\} \sim \{(50, 50)\}.$$

The first indifference ranking implies that the agent's prior over states is given by $(\frac{1}{2}, \frac{1}{2})$. The second indifference ranking suggests that his willingness to pay to each act is 50.

Now consider a bigger menu $\{(100, 0), (0, 100)\}$, which allows the agent to postpone his investment decision in the future. Presumably, facing this menu, the agent optimally solves the problem of costly information acquisition and will make a choice from the menu contingent upon the arrival of new information. Thus, the agent will exhibit preference for flexibility such as

$$\{(60, 60)\} \sim \{(100, 0), (0, 100)\} \succ \{(100, 0)\} \sim \{(0, 100)\}.$$

The first indifference suggests that the agent's willingness to pay to the bigger menu is given as the payoff of 60. Therefore, the net benefit of optimal information acquisition compared with the prior being adopted is $60 - 50 = 10$.

Let us consider an implication of the payoff-independent cost model given in (4). As the payoff-independent cost model satisfies the property, called translation invariance, for all positive payoff m,

$$\{(60 + m, 60 + m)\} \sim \{(100 + m, m), (m, 100 + m)\}.$$

This ranking means that the agent's willingness to pay to the menu increases exactly by m after adding m throughout. Again, the net benefit of optimal information acquisition at $\{(100+m, m), (m, 100+m)\}$ compared with that of the prior is given as $(60+m)-(50+m) = 10$. As the net benefit of information acquisition is invariant with m, an optimal level of information acquisition should be invariant among all those menus.

However, a level of common payoff m may affect the incentive for costly information acquisition. When m is sufficiently large, the significance of the state-dependent payoff of 100 relative to the constant payoff m seems to diminish. This may make the agent more reluctant to acquire information. The payoff-independent cost model fails to capture the effect of such payoff changes on the incentives to acquire information.

On the other hand, if the costs for information acquisition are payoff-dependent and homogeneous, as the common payoff m increases, information acquisition costs also increase proportionally, in contrast to payoff-independent costs, which makes the agent more reluctant to acquire information. The impact on the incentives appears in preference over menus as a deviation from translation invariance, as follows:

$$\{(60 + m, 60 + m)\} \succ \{(100 + m, m), (m, 100 + m)\}$$

for large $m > 0$.

2 Homogeneous cost representation

2.1 Primitives

We consider the following as primitives of the model: these primitives are exactly the same as in HHT and de Oliveira, Denti, Mihm, and Ozbek [4].

- $\Omega = \{\omega_1, ..., \omega_n\}$: 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 non-empty finite set of acts, called a menu
- $\mathbb{F}$: the set of all menus
- Preference $\succsim$ over $\mathbb{F}$

2.2 Preliminaries

Let $\overline{p} \in \Delta(\Omega)$ be the agent's prior belief. A probability distribution $\pi \in \Delta(\Delta(\Omega))$ is interpreted as an information structure 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 information structures. We say that π satisfies a martingale property or a Bayesian plausibility constraint (Kamenica and Gentzkow [6]) if

$$p^\pi = \overline{p}. \tag{6}$$

That is, the initial prior associated with π exactly coincides with the agent's prior belief $\overline{p}$. Define

$$\Pi(\overline{p}) = \{\pi \in \Delta(\Delta(\Omega)) \,|\, p^\pi = \overline{p}\},$$

which is weak* closed and convex.

Given $u : X \to \mathbb{R}$ and a menu F, a benefit of information of $\pi \in \Pi(\overline{p})$, denoted by $b_F^u(\pi)$, is defined as in (1). In particular, for any singleton menu $\{f\}$ and $\pi \in \Pi(\overline{p})$, we have

$$b_{\{f\}}^u(\pi) = \sum_\Omega u(f(\omega))\overline{p}(\omega),$$

that is, the benefit of information exactly coincides with the expected utility of f under the prior if the agents makes a commitment.

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

Definition 1 *A signal $\pi \in \Delta(\Delta(\Omega))$ is Blackwell more informative than a signal $\rho \in \Delta(\Delta(\Omega))$, 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 $b_F^u(\pi) \geq b_F^u(\rho)$ whenever π is Blackwell more informative than ρ.

2.3 Homogeneous payoff-dependent costs

We start with the definition of our representation.

Definition 2 *A homogeneous cost representation is a tuple $(u, \overline{p}, \gamma_s)$, where $u : X \to \mathbb{R}$ is an unbounded expected utility function with $u(X) = \mathbb{R}$, $\overline{p}$ is the initial prior, γ_s is a payment rate function such that $\succsim$ is represented by*

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{b_F^u(\pi) - \gamma_s(\pi)|b_F^u(\pi)|\}. \tag{7}$$

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. We assume the payment rate function $\gamma_s(\pi)$ to satisfy the following properties: (i) γ_s is quasi-convex and lower semi-continuous, (ii) $\gamma_+(\pi) \in [0, 1]$ and $\gamma_-(\pi) \geq 0$, (iii) if there is no information acquisition, there is no cost: $\gamma_s(\delta_{\overline{p}}) = 0$ for the initial prior $\overline{p}$, and (iv) a more informative experiment is more costly: for all $\pi, \rho \in \Pi(\overline{p})$, $\pi \trianglerighteq \rho \implies \gamma_s(\pi) \geq \gamma_s(\rho)$.

One simple observation is that (7) is a special case of the CSL representation when its payoff-dependent cost $C(\pi, t)$ satisfies homogeneity of degree one, that is, for all π, t, and $\lambda > 0$,

$$C(\pi, \lambda t) = \lambda C(\pi, t).$$

This homogeneity implies that the cost function 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 $t < 0$. Thus, the payoff-dependent cost is written as $C(\pi, t) = \gamma_s(\pi)|t|$, where $s = \mathrm{sgn}(t)$, $\gamma_+(\pi) = C(\pi, 1)$, and $\gamma_-(\pi) = C(\pi, -1)$. All the properties of $\gamma_s(\pi)$ stated above are inherited from the properties of the net benefit function $W(\pi, t)$ (see HHT for more details).

2.3.1 Reduced form representation

The homogeneous cost representation admits a convenient reduced form. If $b_F^u(\pi) \geq 0$, the net benefit of π is written as $(1-\gamma_+(\pi))b_F^u(\pi) \geq 0$, while if $b_F^u(\pi) < 0$, the net benefit of π is written as $(1+\gamma_-(\pi))b_F^u(\pi) < 0$. Thus, $(1-\gamma_+)$ and $(1+\gamma_-)$ are multipliers for gross benefits of information. Define

$$\beta_+(\pi) = 1-\gamma_+(\pi), \text{ and } \beta_-(\pi) = 1+\gamma_-(\pi).$$

To rewrite the functional form in Definition 2, define

$$\Pi^+(F) = \{\pi \in \Pi(\overline{p}) \,|\, b_F^u(\pi) \geq 0\}.$$

Note that $\Pi^+(F) \neq \emptyset$ if $b_F^u(\pi) \geq 0$ for some π, and $\Pi^+(F) = \emptyset$ if $b_F^u(\pi) < 0$ for all π. Then, (7) can be rewritten as

$$U(F) = \begin{cases} \max_{\pi \in \Pi(\overline{p})} \beta_+(\pi)b_F^u(\pi) \geq 0 & \text{if } \Pi^+(F) \neq \emptyset, \\ \max_{\pi \in \Pi(\overline{p})} \beta_-(\pi)b_F^u(\pi) < 0 & \text{if } \Pi^+(F) = \emptyset. \end{cases} \tag{8}$$

In the gain frame (or $\Pi^+(F) \neq \emptyset$), $U(F)$ admits a discounted utility form. On the other hand, in the loss frame (or $\Pi^+(F) = \emptyset$), the representation is written as

$$U(F) = -\min_{\pi \in \Pi(\overline{p})} \beta_-(\pi)|b_F^u(\pi)|.$$

That is, the agent chooses an experiment to minimize losses amplified by the multipliers $\beta_-(\pi)$.

The function $\beta_+ : \Pi(\overline{p}) \to [0,1]$ is called a discounting function, and the properties of γ_+ are translatred to those of β_+ as follows:

(i) β_+ is quasi-concave and upper semi-continuous.

(ii) $\beta_+(\delta_{\overline{p}}) = 1$ for the initial prior $\overline{p}$.

(iii) For all $\pi, \rho \in \Pi(\overline{p})$, $\pi \trianglerighteq \rho \implies \beta_+(\pi) \leq \beta_+(\rho)$.

Part (i) is a technical condition to ensure a well-defined optimization problem of information acquisition. Part (ii) states that there is no cost (no discounting) if the prior information is chosen. Part (iii) states that if a more informative signal structure is chosen, then the benefit of information is more discounted.

The function $\beta_- : \Pi(\overline{p}) \to [1,\infty]$, which is applied to negative benefits or losses of information, is called a premium function.[1] The properties of γ_- are translatred to those of β_- as follows:

[1] An information structure π with $\beta_-(\pi) = \infty$ is not relevant since it is too costly and never chosen in the information acquisition stage.

(i) β_- is quasi-convex and lower semi-continuous.

(ii) $\beta_-(\delta_{\overline{p}}) = 1$ for the initial prior $\overline{p}$.

(iii) For all $\pi, \rho \in \Pi(\overline{p})$, $\pi \trianglerighteq \rho \implies \beta_-(\pi) \geq \beta_-(\rho)$.

Parts (i) and (ii) are symmetric to those of the discounting function. Part (iii) states that if a more informative signal structure is chosen, then the decision maker has to pay more premium.

3 Behavioral foundation

3.1 Axioms

The first five axioms, referred to as the basic axioms, 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 $\omega \in \Omega$, for all states, f gives a more preferred lottery than g does. 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.[2]

The next axiom is a weak form of Independence, which is imposed only on singleton menus.

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

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

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 $g \in G$. If the agent has $\alpha F + (1-\alpha)G$, the randomization α is realized after the agent makes a choice from $\alpha F + (1-\alpha)G$. Thus, information acquisition cannot be completely tailored for F and G. The axiom states that the agent avoids contingent planning.

Higashi, Hyogo, and Takeoka [5] show the following result.

Theorem 1 (Higashi, Hyogo, and Takeoka [5]) *Preference $\succsim$ satisfies the basic axioms, Singleton Independence, and Aversion to Contingent Planning if and only if it admits a CSL representation $(u, \overline{p}, W)$.*

The homogeneous cost representation is a special case of the CSL representation. We need an additional axiom for its characterization. A salient feature of (7) 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 (7), it is easy to see that $U(\alpha F) = \alpha U(F)$. A behavioral counterpart of the scale-independence is the independence axiom imposed only when menus are mixed with the neutral outcome.

Let $x_0 \in X$ be a neutral outcome, which is related to the next axiom.

Axiom 8 (Neutral Outcome Independence) *For all menus F, G, and $\alpha \in (0,1)$,*

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

This axiom requires that mixing menus with a neutral outcome should not affect the optimal choice of experiments.

[2] de Oliveira, Denti, Mihm, and Ozbek [4] 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 $\{y\} \succ \{\alpha z + (1-\alpha)x\}$.

3.2 Representation result

We are ready to state our main theorem in this paper.

Theorem 2 *Suppose that preference $\succsim$ admits a CSL Representation. Then $\succsim$ satisfies Neutral Outcome Independence if and only if it admits a homogeneous cost representation $(u, \overline{p}, \beta_+, \beta_-)$. Moreover, discounting and premium functions are obtained as*

$$\beta_+(\pi) = \inf_{\{F \in \mathbb{F} \,|\, b_F^u(\pi) > 0\}} \frac{u(x_F)}{b_F^u(\pi)}, \tag{9}$$

and

$$\beta_-(\pi) = \sup_{\{F \in \mathbb{F} \,|\, b_F^u(\pi) < 0\}} \frac{u(x_F)}{b_F^u(\pi)}, \tag{10}$$

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

One additional axiom, Neutral Outcome Independence, guarantees that a CLS Representation is homogeneous. The homogeneous payoff-dependent cost representation has a parallel relationship with the confidence representation of Chateauneuf and Faro [3], which satisfies homogeneity.

The expressions of β_+ and β_- provide an explicit formula for eliciting the discounting and premium functions. If u and x_F are elicited from the agent's preference, $\beta_+(\pi)$ and $\beta_-(\pi)$ can be computed according to formulae (9) and (10), respectively.

To obtain an intuition behind the formula (9), take any π and menu F. The homogeneous cost representation implies that if $b_F^u(\pi) > 0$,

$$U(F) = \max_{\rho \in \Pi(\overline{p})} \beta_+(\rho) b_F^u(\rho) \geq \beta_+(\pi) b_F^u(\pi),$$

which implies, for all such π,

$$\beta_+(\pi) \leq \frac{U(F)}{b_F^u(\pi)}.$$

Thus, $\beta_+(\pi)$ is a lower bound of $\frac{U(F)}{b_F^u(\pi)}$ among menus F with $b_F^u(\pi) > 0$. In particular, if π is an optimal information structure at F, the above inequality holds with equality. Hence, $\beta_+(\pi)$ can be derived as the infimum of $\frac{u(x_F)}{b_F^u(\pi)}$ among all menus F with $b_F^u(\pi) > 0$.

The symmetric argument is applicable for elicitation of β_-. Take any F such that $b_F^u(\pi) < 0$ for all π. According to the representation,

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

As $b_F^u(\pi) < 0$,

$$\beta_-(\pi) \geq \frac{U(F)}{b_F^u(\pi)},$$

where the equality holds if π is an optimal information structure at F. Hence, $\beta_-(\pi)$ can be derived as the supremum of $\frac{u(x_F)}{b_F^u(\pi)}$ among all menus F with $b_F^u(\pi) < 0$.

3.3 Uniqueness

The next theorem shows the uniqueness property of the homogeneous cost representation.

Theorem 3 *Assume that there are two homogeneous cost representations* $(u_i, \overline{p}_i, \beta_{+,i}, \beta_{-,i})$, $i = 1, 2$ *that represent the same preference* $\succsim$ *on* $\mathbb{F}$. *Then, there exists* $\alpha > 0$ *such that* $u_2 = \alpha u_1$, $\overline{p}_1 = \overline{p}_2$, $\beta_{+,1} = \beta_{+,2}$, *and* $\beta_{-,1} = \beta_{-,2}$.

As $\succsim$ is represented by a subjective expected utility over acts, the uniqueness about expected utility and prior directly follows from the uniqueness of Anscombe and Aumann [1]. As we assume that $u_i(x_0) = 0$ for all i, we have $u_2 = \alpha u_1$. Moreover, by (9) of Theorem 2, for any $\pi \in \Pi(\overline{p})$,

$$\beta_{+,2}(\pi) = \inf_{\{F \in \mathbb{F} \,|\, b_F^{u_2}(\pi) > 0\}} \frac{u_2(x_F)}{b_F^{u_2}(\pi)} = \inf_{\{F \in \mathbb{F} \,|\, b_F^{u_1}(\pi) > 0\}} \frac{\alpha u_1(x_F)}{\alpha b_F^{u_1}(\pi)} = \beta_{+,1}(\pi).$$

Similarly, by (10),

$$\beta_{-,2}(\pi) = \sup_{\{F \in \mathbb{F} \,|\, b_F^{u_2}(\pi) < 0\}} \frac{u_2(x_F)}{b_F^{u_2}(\pi)} = \sup_{\{F \in \mathbb{F} \,|\, b_F^{u_1}(\pi) < 0\}} \frac{\alpha u_1(x_F)}{\alpha b_F^{u_1}(\pi)} = \beta_{-,1}(\pi).$$

3.4 Proof sketch of Theorem 2

The following is a proof sketch of the sufficiency. Up to the middle step, we follow the construction of the CSL representation in HHT. By Singleton Independence and the basic axioms, $\succsim$ on acts is represented by a subjective expected utility with an expected utility function $u : X \to \mathbb{R}$ with $u(X) = \mathbb{R}$ and a prior $\overline{p}$ over Ω. This utility function can be extended to the whole domain $\mathbb{F}$ because each menu admits its lottery equivalent $x_F \in X$ with $\{x_F\} \sim F$. Let U denote this representation.

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

$$\varphi_F(p) = \max_{f \in F} \sum_{\Omega} u(f(\omega))p(\omega). \tag{11}$$

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 $V(\varphi_F) = U(F)$. The functional V is extended to the set of all continuous functions $C(\Delta(\Omega))$, preserving the desired properties. HHT show that V is rewritten as

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

where $W(\pi, t)$ is a net benefit function and $\langle \varphi, \pi \rangle = \int \varphi(p)\, \mathrm{d}\pi(p)$. Moreover, when $u(X) = \mathbb{R}$, the martingale property is proved, that is, the above maximum is achieved within the subset $\Pi(\overline{p}) \subset \Delta(\Delta(\Omega))$.

Using an additional axiom, we show that the above representation can be transformed into the desired form. As $u : X \to \mathbb{R}$ is an expected utility, we can assume $u(x_0) = 0$, where x_0 is a neutral outcome. As $\succsim$ satisfies Neutral Outcome Independence, we can show that V is positively homogeneous, which in turn implies that $W(\pi, t)$ is homogeneous of degree one in t. By defining $\beta_+(\pi) = W(\pi, 1)$ and $\beta_-(\pi) = -W(\pi, -1)$, V can be rewritten as

$$V(\varphi) = \max_{\pi \in \Pi(\overline{p})} [\beta_+(\pi)(\langle \varphi, \pi \rangle)^+ - \beta_-(\pi)(\langle \varphi, \pi \rangle)^-],$$

where $(t)^+ = \max\{0, t\}$ and $(t)^- = \max\{0, -t\}$ for all $t \in \mathbb{R}$. This type of specification of W has a counterpart in decision making under ambiguity, which is called the confidence model (Chateauneuf and Faro [3]). Finally, we show that discounting and premium functions, β_+ and β_-, satisfy the desired properties.

4 Interpersonal comparison

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.

Definition 3 $\succsim_1$ *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 4 *Assume that* $\succsim_i$, $i = 1, 2$ *satisfy the axioms of Theorem 2.* $\succsim_1$ *is more averse to commitment than* $\succsim_2$ *if and only if there exists a homogeneous cost representation* $(u, \overline{p}, \beta_{+,i}, \beta_{-,i})$ *represents* $\succsim_i$, $i = 1, 2$, *and*

$$\beta_{+,1}(\pi) \geq \beta_{+,2}(\pi), \text{ and } \beta_{-.1}(\pi) \leq \beta_{-,2}(\pi)$$

for all π.

This theorem provides a characterization about aversion to commitment. Agent 1 has a discounting function and a premium function which are closer to one than agent 2 does, that is, agent 1 has lower costs for any information structure than agent 2.

Proof. 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. Hence, we can assume that $u_1 = u_2 = u$ and $\overline{p}_1 = \overline{p}_2 = \overline{p}$.

Take any $F \in \mathbb{F}$. Let x_F^i be agent i's lottery equivalent of F, that is, $\{x_F^i\} \sim_i F$. Since $\succsim_1$ is more averse to commitment than $\succsim_2$, $F \sim_2 \{x_F^2\}$ implies $F \succsim_1 \{x_F^2\}$. Therefore, we have $u(x_F^1) = U_1(F) \geq u(x_F^2) = U_2(F)$ for all F.

Take any π and F with $b_F^u(\pi) > 0$. By the representation, for $i = 1, 2$,

$$u(x_F^i) = U_i(F) = \max_{\pi' \in \Pi(\overline{p})} \beta_{+,i}(\pi) b_F^u(\pi') \geq \beta_{+,i}(\pi) b_F^u(\pi) > 0.$$

Therefore,

$$\beta_{+,1}(\pi) = \inf_{\{F \mid b_F^u(\pi) > 0\}} \frac{u(x_F^1)}{b_F^u(\pi)} \geq \inf_{\{F \mid b_F^u(\pi) > 0\}} \frac{u(x_F^2)}{b_F^u(\pi)} = \beta_{+,2}(\pi).$$

Take any F with $u(x_F^1) < 0$. By the above observation, $u(x_F^2) \leq u(x_F^1) < 0$, that is, $\{F \mid u(x_F^1) < 0\} \subset \{F \mid u(x_F^2) < 0\}$. Moreover, $u(x_F^1) < 0$ implies that $b_F^u(\pi) < 0$ for all $\pi \in \Pi(\overline{p})$. Thus, by Lemma 11,

$$\beta_{-,1}(\pi) = \sup_{\{F \mid u(x_F^1) < 0\}} \frac{u(x_F^1)}{b_F^u(\pi)} \leq \sup_{\{F \mid u(x_F^2) < 0\}} \frac{u(x_F^2)}{b_F^u(\pi)} = \beta_{-,2}(\pi),$$

as desired. ■

Appendix

A Preliminaries

Following de Oliveira, Denti, Mihm, and Ozbek [4], we introduce some notions and mathematical preliminaries needed for the subsequent analysis. The proofs are omitted.

- $C(\Delta(\Omega))$: the set of all real-valued continuous functions over $\Delta(\Omega)$ with the supnorm
- $ca(\Delta(\Omega))$: the set of all signed measures over $\Delta(\Omega)$ with the weak* topology
- $ca_+(\Delta(\Omega))$: the set of all positive measures over $\Delta(\Omega)$
- 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 $F \in \mathbb{F}$, let

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

- $\Phi_{\mathbb{F}}(\Phi_{\mathcal{F}}, \Phi_X)$: 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) $\Phi_{\mathbb{F}} \subset \Phi$

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

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

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

B Proof of Theorem 2

B.1 CSL representations

Theorem 1 of HHT shows that if $\succsim$ satisfies the basic axioms, Singleton Independence, and Aversion to Contingent Planning if and only if it admits a subjective learning representation. Below, we briefly explain the only if part as an intermediate result for our main theorem in the present paper.

First, we derive a utility representation $U : \mathbb{F} \to \mathbb{R}$ and define the functional $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ as in de Oliveira, Denti, Mihm, and Ozbek [4]. By Singleton Independence and the basic axioms, $\succsim$ over acts admits a subjective expected utility representation. There exists 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

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

Since every menu F has a certainty equivalent $x_F \in X$ such that $\{x_F\} \sim F$, we can extend $U : \mathcal{F} \to \mathbb{R}$ to $\mathbb{F}$ by $U(F) = U(x_F)$. This extension $U : \mathbb{F} \to \mathbb{R}$ represents $\succsim$. By Two-Sided Unboundedness, $U(\mathbb{F}) = \mathbb{R}$.

Define the functional $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ by $V(\varphi_F) = U(F)$. They show that V is well-defined because $\varphi_F = \varphi_G$ implies $F \sim G$. By Lemma 1 of HHT, $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is monotone, normalized, quasi-convex, and continuous.

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

$$V(\varphi) = \inf\{V(\varphi_F) | \, \varphi_F \in \Phi_{\mathbb{F}}, \varphi_F \geq \varphi\} \tag{12}$$

for all $\varphi \in C(\Delta(\Omega))$. Lemmas 2 and 3 of HHT shows that $V : C(\Delta(\Omega)) \to \mathbb{R}$ is a well-defined extension of $V : \Phi_{\mathbb{F}} \to \mathbb{R}$. Moreover, V is monotone, normalized, quasi-convex, and continuous.

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

$$\begin{aligned} B(\pi, t) &= \{\varphi \in C(\Delta(\Omega)) \,|\, \langle \varphi, \pi \rangle \geq t\}, \text{ and} \\ W(\pi, t) &= \inf_{\varphi \in B(\pi,t)} V(\varphi). \end{aligned} \tag{13}$$

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).

By Lemma 6 of HHT, V is written as

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

HHT shows that W satisfies the following properties:

Lemma 1 *(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 $(\pi, t) \in \Delta(\Delta(\Omega)) \times \mathbb{R}$.

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

(5) If $\pi \trianglerighteq \rho$, $W(\pi, t) \leq W(\rho, t)$ for all t.

Define

$$\Pi = \{\pi \in \Delta(\Delta(\Omega)) \,|\, W(\pi, t) > -\infty \text{ for some } t\}. \tag{14}$$

By Lemma 1 (4) implies $\delta_{\overline{p}} \in \Pi$. In particular, $\Pi \neq \emptyset$. Moreover, Lemma 12 of HHT shows that $\Pi \subset \Pi(\overline{p})$. Since any $\pi \notin \Pi$ never achieves the maximum of W, the representation U is rewritten as

$$U(F) = V(\varphi_F) = \max_{\pi \in \Pi(\overline{p})} W(\pi, \langle \varphi_F, \pi \rangle) = \max_{\pi \in \Pi} W(\pi, \langle \varphi_F, \pi \rangle), \tag{15}$$

which is a costly subjective learning representation.

B.2 Sufficiency

We show that if $\succsim$ satisfies Neutral Outcome Independence in addition, then (15) is written as a homogeneous cost representation.

Recall $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ defined as in Section B.1. Since u on X is mixture linear, without loss of generality, we can assume $u(x_0) = 0$ where x_0 is a neutral outcome.

Lemma 2 *$V : \Phi_{\mathbb{F}} \to \mathbb{R}$ is positively homogeneous.*

Proof. We show positive homogeneity of V. For $\alpha \in [0, 1]$,

$$\begin{aligned} V(\alpha\varphi_F) &= V(\alpha\varphi_F + (1-\alpha)0) = V(\varphi_{\alpha F+(1-\alpha)\{x_0\}}) \\ &= U(\alpha F + (1-\alpha)\{x_0\}) = U(\alpha\{x_F\} + (1-\alpha)\{x_0\}) \\ &= \alpha U(\{x_F\}) + (1-\alpha)0 = \alpha V(x_F). \end{aligned}$$

The second equality follows from linearity of φ. The forth equality follows from Neutral Outcome Independence.

For $\alpha \in (1, \infty)$, denote $\varphi_G = \alpha\varphi_F \in \Phi_{\mathbb{F}}$. By the above property,

$$V(\varphi_F) = V(\frac{1}{\alpha}\varphi_G) = \frac{1}{\alpha}V(\varphi_G) = \frac{1}{\alpha}V(\alpha\varphi_F),$$

as desired. ■

As in (12), V is extended to $C(\Delta(\Omega))$.

Lemma 3 *$V : C(\Delta(\Omega)) \to \mathbb{R}$ is positively homogeneous.*

Proof. We show that V satisfies positive homogeneity. First of all, as V is normalized, if $\alpha = 0$, $V(\alpha\varphi) = V(0) = 0 = \alpha V(\varphi)$ for all φ, as desired.

For every $\varphi \in C(\Delta(\Omega))$ and $\alpha > 0$, note that

$$\{\varphi_F \in \Phi_{\mathbb{F}} \,|\, \varphi_F \geq \varphi\} = \{\alpha\varphi_F \in \Phi_{\mathbb{F}} \,|\, \alpha\varphi_F \geq \varphi \text{ for some } \varphi_F \in \Phi_{\mathbb{F}}\}.$$

Indeed, take any φ_F from the left-hand side. Since $\Phi_{\mathbb{F}}$ is a cone, $\frac{\varphi_F}{\alpha} \in \Phi_{\mathbb{F}}$. Thus, $\alpha(\frac{\varphi_F}{\alpha}) = \varphi_F \geq \varphi$. By definition, $\varphi_F = \alpha(\frac{\varphi_F}{\alpha})$ belongs to the right-hand side. Conversely, take any $\alpha\varphi_F$ from the right-hand side. Since $\alpha\varphi_F \in \Phi_{\mathbb{F}}$ and $\alpha\varphi_F \geq \varphi$, by definition, $\alpha\varphi_F$ belongs to the left-hand side, as desired.

For all $\varphi \in C(\Delta(\Omega))$ and $\alpha > 0$, the above observation implies that

$$\begin{aligned} V(\alpha\varphi) &= \inf\{V(\varphi_F)|\,\varphi_F \in \Phi_{\mathbb{F}}, \varphi_F \geq \alpha\varphi\} \\ &= \inf\{V(\alpha\varphi_F)|\,\varphi_F \in \Phi_{\mathbb{F}}, \alpha\varphi_F \geq \alpha\varphi\} \\ &= \inf\{\alpha V(\varphi_F)|\,\varphi_F \in \Phi_{\mathbb{F}}, \varphi_F \geq \varphi\} = \alpha V(\varphi). \end{aligned}$$

■

Recall $W : \Delta(\Delta(\Omega)) \times \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ defined as in (13).

Lemma 4 *For all $\pi \in \Delta(\Delta(\Omega))$, $W(\pi, t)$ is homogeneous of degree one in t, that is, for all $\alpha > 0$, $W(\pi, \alpha t) = \alpha W(\pi, t)$.*

Proof. (i) By Lemma 4 of HHT and positive homogeneity of V,

$$\begin{aligned} W(\pi, \alpha t) &= \inf_{\varphi \in B(\pi, \alpha t)} V(\varphi) = \inf_{\varphi \in \alpha B(\pi, t)} V(\varphi) = \inf_{\varphi \in B(\pi, t)} V(\alpha\varphi) \\ &= \alpha \inf_{\varphi \in B(\pi, t)} V(\varphi) = \alpha W(\pi, t). \end{aligned}$$

■

For notational convenience, given $\pi \in \Delta(\Delta(\Omega))$, we define a scalar function defined for all $t \in \mathbb{R}$ by $W_\pi(t) = W(\pi, t)$. Since our concern is the infimum, let

$$\operatorname{dom} W_\pi = \{t \in \mathbb{R} \,|\, W_\pi(t) > -\infty\}.$$

Define Π as in (14).

Lemma 5 *For all $\pi \in \Pi$, either* $\operatorname{dom} W_\pi = \mathbb{R}$ *or* $\operatorname{dom} W_\pi = \mathbb{R}_+$.

Proof. Take any $\pi \in \Pi$. Assume that there exists $t^* < 0$ such that $W(\pi, t^*) > -\infty$. By Lemma 1 (1), $W(\pi, t) \geq W(\pi, t^*) > -\infty$ for all $t \geq t^*$. On the other hand, for all $t < t^* < 0$, by Lemma 4,

$$W(\pi, t) = W(\pi, \frac{t}{t^*} t^*) = \frac{t}{t^*} W(\pi, t^*) > -\infty.$$

Hence, $\operatorname{dom} W_\pi = \mathbb{R}$.

Next, assume that there exists no $t < 0$ such that $W(\pi, t) > -\infty$. That is, $W(\pi, t) = -\infty$ for all $t < 0$. Since $\pi \in \Pi$, there exists $t^* \geq 0$ such that $W(\pi, t^*) > -\infty$. If $t^* = 0$, by Lemma 1 (1), $W(\pi, t) \geq W(\pi, 0) > -\infty$ for all $t \geq 0$. Thus, $\operatorname{dom} W_\pi = \mathbb{R}_+$. If $t^* > 0$, by Lemma 4, for all $t > 0$,

$$W(\pi, t) = W(\pi, \frac{t}{t^*} t^*) = \frac{t}{t^*} W(\pi, t^*) > -\infty.$$

Moreover, since $W(\pi, t)$ is upper semi-continuous in t,

$$W(\pi, 0) = \lim_{t \searrow 0} W(\pi, t) = \lim_{t \searrow 0} \frac{t}{t^*} W(\pi, t^*) = 0 > -\infty.$$

Hence, $\operatorname{dom} W_\pi = \mathbb{R}_+$. ■

By Lemmas 4 and 5, $W(\pi, t)$ can be rewritten as follows: for all $\pi \in \Pi$ and $t > 0$,

$$W(\pi, t) = W(\pi, t \cdot 1) = tW(\pi, 1) = \beta_+(\pi) \cdot t,$$

where $\beta_+(\pi) := W(\pi, 1)$. Similarly, for all $\pi \in \Pi$ with $\mathrm{dom}W_\pi = \mathbb{R}$ and $t < 0$,

$$W(\pi, t) = W(\pi, (-t) \cdot -1) = -tW(\pi, -1) = \beta_-(\pi) \cdot t,$$

where $\beta_-(\pi) := -W(\pi, -1)$. If $\mathrm{dom}W_\pi = \mathbb{R}_+$, define $\beta_-(\pi) = \infty$. Thus, for $t \neq 0$, $W(\pi, t)$ is written as

$$W(\pi, t) = \beta_+(\pi)(t)^+ - \beta_-(\pi)(t)^-, \tag{16}$$

where for $t \in \mathbb{R}$, $(t)^+ = \max\{0, t\}$ and $(t)^- = \max\{0, -t\}$ and $-\infty \times 0 = 0$ with convention.

In particular, since $W(\delta_{\overline{p}}, t) = t$ by Lemma 1 (4), $\mathrm{dom}W_{\delta_{\overline{p}}} = \mathbb{R}$, which implies

$$\beta_+(\delta_{\overline{p}}) = \beta_-(\delta_{\overline{p}}) = 1. \tag{17}$$

Lemma 6 *For all $\pi \in \Pi$, $\beta_+(\pi)$ is real-valued, and $\beta_+(\pi) \geq 0$. If $\mathrm{dom}W_\pi = \mathbb{R}$, $\beta_-(\pi)$ is real valued, and $\beta_-(\pi) \geq 0$.*

Proof. Since $W(\pi, t)$ is real-valued for any $\pi \in \Pi$ and $t > 0$, $\beta_+(\pi)$ is real-valued. Next, we show that for all $\pi \in \Pi$, $\beta_+(\pi) \geq 0$. Suppose contrary that there exists $\pi \in \Pi$ such that $\beta_+(\pi) = W(\pi, 1) < 0$. For $t > 1$, $W(\pi, t) = W(\pi, 1) \cdot t < W(\pi, 1)$. This contradicts the fact that $W(\pi, t)$ is nondecreasing in t, shown in Lemma 1 (1).

Similarly, since $W(\pi, t)$ is real-valued for any $\pi \in \Pi$ with $\mathrm{dom}\, W_\pi = \mathbb{R}$, $\beta_-(\pi)$ is real-valued. Finally, we show that for all such π, $\beta_-(\pi) \geq 0$. Since $\langle \varphi_{\{x_0\}}, \pi \rangle = \langle \mathbf{0}, \pi \rangle \geq -1$, we have $W(\pi, -1) \leq 0$. Hence, $\beta_-(\pi) = -W(\pi, -1) \geq 0$. ■

The following lemma provides the case of $t = 0$.

Lemma 7 *For any $\pi \in \Pi$, $W(\pi, 0) = 0$.*

Proof. By Lemma 6, $W(\pi, t) = \beta_+(\pi)t$ for all $t > 0$ and $\pi \in \Pi$. Since $W(\pi, t)$ is upper semi-continuous in t,

$$W(\pi, 0) = \lim_{t \searrow 0} W(\pi, t) = \lim_{t \searrow 0} \beta_+(\pi)t = 0.$$

■

Lemma 8 Π *is closed and convex.*

Proof. To show that Π is closed, let $\pi^n \to \pi$ with $\pi^n \in \Pi$. By Lemma 7, $W(\pi^n, 0) \geq 0$. Since $W : \Delta(\Delta(\Omega)) \times \mathbb{R} \to \mathbb{R}$ is upper semi-continuous by Lemma 1 (3), $W(\pi, 0) \geq 0$, which implies $W(\pi, 0) > -\infty$ at π. Hence, $\pi \in \Pi$, as desired.

To show that Π is convex, take $\pi_1, \pi_2 \in \Pi$ and $\alpha \in [0, 1]$. There exist t_i, $i = 1, 2$, such that $W(\pi_i, t_i) > -\infty$. Since W is quasi-concave in (π, t) by Lemma 1 (2),

$$W(\alpha\pi_1 + (1-\alpha)\pi_2, \alpha t_1 + (1-\alpha)t_2) \geq \min[W(\pi_1, t_1), W(\pi_2, t_2)] > -\infty.$$

Thus, $\alpha\pi_1 + (1-\alpha)\pi_2 \in \Pi$. ■

By Lemma 7, (16) holds for all $t \in \mathbb{R}$. Now, we obtain that

$$V(\varphi) = \max_{\pi \in \Pi} [\beta_+(\pi)(\langle \varphi, \pi \rangle)^+ - \beta_-(\pi)(\langle \varphi, \pi \rangle)^-]. \tag{18}$$

Note that $V(\varphi) \geq 0$ is equivalent to $\langle \varphi, \pi \rangle \geq 0$ for some $\pi \in \Pi$, and $V(\varphi) < 0$ is equivalent to $\langle \varphi, \pi \rangle < 0$ for all $\pi \in \Pi$.

It follows from (18) that $\succsim$ is represented by

$$U(F) = V(\varphi_F) = \max_{\pi \in \Pi} [\beta_+(\pi)(\langle \varphi_F, \pi \rangle)^+ - \beta_-(\pi)(\langle \varphi_F, \pi \rangle)^-]. \tag{19}$$

To obtain a more explicit form of β_+ and β_-, we prepare the following lemma.

Lemma 9 *For all* $\pi \in \Pi$,

$$\beta_+(\pi) = \inf_{\{F \in \mathbb{F} \,|\, b^u_F(\pi) > 0\}} \frac{u(x_F)}{\langle \varphi_F, \pi \rangle}.$$

Proof. Since $W(\pi, \alpha)$ is homogeneous in α, for all $\pi \in \Pi$ and $\alpha > 0$, $W(\pi, \alpha) = \alpha W(\pi, 1) = \alpha \beta(\pi)$. By Lemma 9 of HHT, for any $\pi \in \Pi$ and $\alpha \in \mathbb{R}$

$$W(\pi, \alpha) = \inf_{\varphi \in B(\pi, \alpha)} V(\varphi) = \inf_{\varphi_F \in B(\pi, \alpha)} V(\varphi_F) = \inf_{\{F | \langle \varphi_F, \pi \rangle \geq \alpha\}} u(x_F).$$

We will claim that for any $\pi \in \Pi$ and $\alpha \in \mathbb{R}$,

$$\inf_{\{F | \langle \varphi_F, \pi \rangle \geq \alpha\}} u(x_F) = \inf_{\{F | \langle \varphi_F, \pi \rangle = \alpha\}} u(x_F). \tag{20}$$

Take any F with $\langle \varphi_F, \pi \rangle > \alpha$. For any $\lambda \in (0, 1)$, let $\lambda F + (1 - \lambda)\{x_0\}$ be denoted by λF. Since $\langle \varphi_{\lambda F}, \pi \rangle = \lambda \langle \varphi_F, \pi \rangle$, for any λ sufficiently close to one, $\langle \varphi_F, \pi \rangle > \langle \varphi_{\lambda F}, \pi \rangle > \alpha$. Moreover, since V is positively homogeneous,

$$u(x_F) = V(\varphi_F) > \lambda V(\varphi_F) = V(\lambda \varphi_F) = V(\varphi_{\lambda F}) = u(x_{\lambda F}).$$

That is, if F satisfies $\langle \varphi_F, \pi \rangle > \alpha$, $u(x_F)$ is not a lower bound of $\{u(x_F) \,|\, \langle \varphi_F, \pi \rangle \geq \alpha\}$. Thus, (20) holds.

By the above observations,

$$\begin{aligned}
\inf_{\{F \in \mathbb{F} \,|\, b^u_F(\pi) > 0\}} \frac{u(x_F)}{\langle \varphi_F, \pi \rangle} &= \inf_{\alpha > 0} \left(\inf_{\{F | \langle \varphi_F, \pi \rangle = \alpha\}} \frac{u(x_F)}{\alpha} \right) = \inf_{\alpha > 0} \frac{1}{\alpha} \left(\inf_{\{F | \langle \varphi_F, \pi \rangle = \alpha\}} u(x_F) \right) \\
&= \inf_{\alpha > 0} \frac{1}{\alpha} W(\pi, \alpha) = \beta_+(\pi),
\end{aligned}$$

as desired. ■

Lemma 10 *For all* $\pi \in \Pi$ *with* $\operatorname{dom} W_\pi = \mathbb{R}$,

$$\beta_-(\pi) = \sup_{\{F \in \mathbb{F} \,|\, b^u_F(\pi) < 0\}} \frac{u(x_F)}{\langle \varphi_F, \pi \rangle} = \sup_{\{F \in \mathbb{F} \,|\, u(x_F) < 0\}} \frac{u(x_F)}{\langle \varphi_F, \pi \rangle}.$$

Proof. Note that for all $\pi \in \Pi$ with $\mathrm{dom}W_\pi = \mathbb{R}$ and $\alpha < 0$, $W(\pi, \alpha) = -\alpha W(\pi, -1) = -\alpha\gamma(\pi)$. As in Lemma 9, we have

$$\sup_{\{F\in\mathbb{F}|\langle\varphi_F,\pi\rangle<0\}} \frac{u(x_F)}{\langle\varphi_F,\pi\rangle} = \sup_{\alpha>0}\left(\sup_{\{F\in\mathbb{F}\,|\,\langle\varphi_F,\pi\rangle=-\alpha\}} \frac{u(x_F)}{-\alpha}\right) = \sup_{\alpha>0} -\frac{1}{\alpha}\left(\inf_{\{F\in\mathbb{F}\,|\,\langle\varphi_F,\pi\rangle=-\alpha\}} u(x_F)\right)$$
$$= \sup_{\alpha>0}\left(-\frac{1}{\alpha}W(\pi,-\alpha)\right) = -W(\pi,-1) = \beta_-(\pi).$$

From the functional form, $\beta_-(\pi) \geq 0$ for any $\pi \in \Pi$ implies that $V(\varphi_F) < 0$ is equivalent to $\langle\varphi_F, \pi\rangle < 0$ for all $\pi \in \Pi$. Hence, $u(x_F) < 0$ is equivalent to $\langle\varphi_F, \pi\rangle < 0$ for all $\pi \in \Pi$. This implies that $\{F \in \mathbb{F}|\langle\varphi_F, \pi\rangle < 0, u(x_F) < 0\} = \{F \in \mathbb{F}|u(x_F) < 0\}$. Moreover, if $\langle\varphi_F, \pi\rangle < 0$, $u(x_F) < 0$, $\langle\varphi_G, \pi\rangle < 0$, and $u(x_G) \geq 0$, then

$$\frac{u(x_F)}{\langle\varphi_F,\pi\rangle} > 0 \geq \frac{u(x_G)}{\langle\varphi_G,\pi\rangle}.$$

Thus, we have that

$$\beta_-(\pi) = \sup_{\{F\in\mathbb{F}|\langle\varphi_F,\pi\rangle<0\}} \frac{u(x_F)}{\langle\varphi_F,\pi\rangle} = \sup_{\{F\in\mathbb{F}|\langle\varphi_F,\pi\rangle<0,u(x_F)<0\}} \frac{u(x_F)}{\langle\varphi_F,\pi\rangle} = \sup_{\{F\in\mathbb{F}|u(x_F)<0\}} \frac{u(x_F)}{\langle\varphi_F,\pi\rangle}.$$

■

As shown in Section B.1, $\Pi \subset \Pi(\overline{p})$. Let

$$\Pi^{\mathbb{R}} = \{\pi \in \Pi \,|\, \mathrm{dom}\, W_\pi = \mathbb{R}\}.$$

Note that β_+ is a real-valued function defined on Π, while β_- is a real-valued function defined on $\Pi^{\mathbb{R}}$. Note also that $\Pi^{\mathbb{R}} \neq \emptyset$. In fact, if $\Pi^{\mathbb{R}} = \emptyset$, $W(\pi, t) = -\infty$ for all $t < 0$ and $\pi \in \Pi$. But, since V is normalized, for all $t < 0$,

$$t = V(t\mathbf{1}) = \max_{\pi\in\Pi} W(\pi, \langle t\mathbf{1}, \pi\rangle) = \max_{\pi\in\Pi} W(\pi, t) = -\infty,$$

which is a contradiction. Moreover, $\Pi^{\mathbb{R}}$ is convex. Take any $\pi, \pi' \in \Pi^{\mathbb{R}}$ and $\alpha \in [0, 1]$. Since $W(\pi, t)$ is quasi-concave in (π, t),

$$W(\alpha\pi + (1-\alpha)\pi', -1) \geq \min[W(\pi, -1), W(\pi', -1)] > -\infty,$$

which implies $\mathrm{dom}W_{\alpha\pi+(1-\alpha)\pi'} = \mathbb{R}$. Thus, $\alpha\pi + (1-\alpha)\pi' \in \Pi^{\mathbb{R}}$.

Extend $\beta_+ : \Pi \to \mathbb{R}_+$ and $\beta_- : \Pi^{\mathbb{R}} \to \mathbb{R}_+$ to $\Pi(\overline{p})$ by

$$\beta_+^*(\pi) := \inf_{\{F\in\mathbb{F}\,|\,\langle\varphi_F,\pi\rangle>0\}} \frac{u(x_F)}{\langle\varphi_F,\pi\rangle} \tag{21}$$

and

$$\beta_-^*(\pi) := \sup_{\{F\in\mathbb{F}\,|\,\langle\varphi_F,\pi\rangle<0\}} \frac{u(x_F)}{\langle\varphi_F,\pi\rangle}. \tag{22}$$

By Lemmas 9 and 10, $\beta^*_+ = \beta_+$ on Π and $\beta^*_- = \beta_-$ on $\Pi^{\mathbb{R}}$. Moreover, for all $\pi \in \Pi(\overline{p}) \setminus \Pi$, by definition of β^*_+,$U(F) \geq \beta^*_+(\pi) b^u_F(\pi)$ for all F with $b^u_F(\pi) > 0$, and for all $\pi \in \Pi(\overline{p}) \setminus \Pi^{\mathbb{R}}$, $U(F) \geq \beta^*_-(\pi) b^u_F(\pi)$ for all F with $b^u_F(F) < 0$. Hence, β^*_+ on $\Pi(\overline{p}) \setminus \Pi$ and β^*_- on $\Pi(\overline{p}) \setminus \Pi^{\mathbb{R}}$ are in fact irrelevant for the representation. It follows from (19) that the representation $U(F)$ is rewritten as

$$U(F) = V(\varphi_F) = \max_{\pi \in \Pi(\overline{p})} [\beta^*_+(\pi)(\langle \varphi_F, \pi \rangle)^+ - \beta^*_-(\pi)(\langle \varphi_F, \pi \rangle)^-].$$

Finally, we show that β^*_+ and β^*_- have the desired properties.

Lemma 11

$$\beta^*_-(\pi) = \sup_{\{F \in \mathbb{F} \,|\, u(x_F) < 0\}} \frac{u(x_F)}{\langle \varphi_F, \pi \rangle}.$$

Proof. The proof is the same as in Lemma 10. Since $\beta^*_-(\pi) \geq 1 > 0$, the representation implies that $V(\varphi_F) < 0$ is equivalent to $\langle \varphi_F, \pi \rangle < 0$ for all $\pi \in \Pi(\overline{p})$. Hence, $u(x_F) < 0$ is equivalent to $\langle \varphi_F, \pi \rangle < 0$ for all $\pi \in \Pi$. This implies that $\{F \in \mathbb{F} | \langle \varphi_F, \pi \rangle < 0, u(x_F) < 0\} = \{F \in \mathbb{F} | u(x_F) < 0\}$. ■

Lemma 12 *β^*_+ is upper semi-continuous and β^*_- is lower semi-continuous.*

Proof. For each fixed F with $\langle \varphi_F, \pi \rangle > 0$, $\frac{u(x_F)}{\langle \varphi_F, \pi \rangle}$ is continuous in π. Since the infimum function among upper semi-continuous functions is also upper semi-continuous, β^*_+ is upper semi-continuous in π.

Similarly, since the supremum function among lower semi-continuous functions is also lower semi-continuous, β^*_- is lower semi-continuous in π. ■

Lemma 13 *β^*_+ is quasi-concave and β^*_- is quasi-convex.*

Proof. By (21),

$$\frac{1}{\beta^*_+(\pi)} = \frac{1}{\inf_F \frac{u(x_F)}{\langle \varphi_F, \pi \rangle}} = \sup_F \frac{\langle \varphi_F, \pi \rangle}{u(x_F)}.$$

Since $\langle \varphi_F, \pi \rangle$ is linear in π,

$$\sup_F \frac{\langle \varphi_F, \alpha\pi + (1-\alpha)\pi' \rangle}{u(x_F)} \leq \alpha \sup_F \frac{\langle \varphi_F, \pi \rangle}{u(x_F)} + (1-\alpha) \sup_F \frac{\langle \varphi_F, \pi' \rangle}{u(x_F)},$$

that is, $1/\beta^*_+$ is convex. Since the reciprocal of a quasi-convex function is quasi-concave, β^*_+ is quasi-concave.[3]

[3] If f is quasi-concave, $f(\pi') \geq f(\pi) \implies f(\alpha\pi + (1-\alpha)\pi') \geq f(\pi)$ for all $\alpha \in [0,1]$. This implies that

$$\frac{1}{f(\pi')} \leq \frac{1}{f(\pi)} \implies \frac{1}{f(\alpha\pi + (1-\alpha)\pi')} \leq \frac{1}{f(\pi')}$$

for all $\alpha \in [0,1]$. Thus, $1/f$ is quasi-convex. The converse also holds.

By (22),

$$\frac{1}{\beta^*_-(\pi)} = \frac{1}{\sup_F \frac{u(x_F)}{\langle \varphi_F, \pi\rangle}} = \inf_F \frac{\langle \varphi_F, \pi\rangle}{u(x_F)}.$$

Since $\langle \varphi_F, \pi\rangle$ is linear in π,

$$\inf_F \frac{\langle \varphi_F, \alpha\pi + (1-\alpha)\pi'\rangle}{u(x_F)} \geq \alpha \inf_F \frac{\langle \varphi_F, \pi\rangle}{u(x_F)} + (1-\alpha) \inf_F \frac{\langle \varphi_F, \pi'\rangle}{u(x_F)},$$

that is, $1/\beta^*_-$ is concave. Since the reciprocal of a quasi-concave function is quasi-convex, β^*_- is quasi-convex. ■

Lemma 14 *For all $\pi \in \Pi(\overline{p})$, $\beta^*_+(\pi) \leq 1$ and $\beta^*_-(\pi) \geq 1$.*

Proof. Take any lottery $x \in X$ with $u(x) > 0$. Then, $\varphi_{\{x\}}$ satisfies $\langle \varphi_{\{x\}}, \pi\rangle = u(x)$ for all π. By definition of β^*_+,

$$\beta^*_+(\pi) \leq \frac{u(x)}{\langle \varphi_{\{x\}}, \pi\rangle} = 1.$$

Take any lottery $x \in X$ with $u(x) < 0$. By definition of β^*_-,

$$\beta^*_-(\pi) \geq \frac{u(x)}{\langle \varphi_{\{x\}}, \pi\rangle} = 1.$$

■

Lemma 15 $\beta^*_+(\delta_{\bar{p}}) = \beta^*_-(\delta_{\bar{p}}) = 1$

Proof. Since $\beta^*_+(\delta_{\bar{p}}) = \beta_+(\delta_{\bar{p}})$ and $\beta^*_-(\delta_{\bar{p}}) = \beta_-(\delta_{\bar{p}})$, the result follows from (17). ■

Lemma 16 *For all $\pi, \rho \in \Pi(\overline{p})$, $\pi \trianglerighteq \rho \implies \beta^*_+(\pi) \leq \beta^*_+(\rho)$ and $\beta^*_-(\pi) \geq \beta^*_-(\rho)$.*

Proof. If $\pi \trianglerighteq \rho$, $\{F \in \mathbb{F} \,|\, b^u_F(\rho) > 0\} \subseteq \{F \in \mathbb{F} \,|\, b^u_F(\pi) > 0\}$ and $\langle \varphi_F, \pi\rangle \geq \langle \varphi_F, \rho\rangle$. By definition of β^*_+, we have $\beta^*_+(\pi) \leq \beta^*_+(\rho)$.

If $\pi \trianglerighteq \rho$, it follows from definition of β^*_- that only the denominator, which is negative, becomes greater. Hence, we obtain $\beta^*_-(\pi) \geq \beta^*_-(\rho)$. ■

Consequently,

$$\begin{aligned} U(F) = V(\varphi_F) &= \max_{\pi \in \Pi(\overline{p})} [\beta^*_+(\pi)(\langle \varphi_F, \pi\rangle)^+ - \beta^*_-(\pi)(\langle \varphi_F, \pi\rangle)^-] \\ &= \max_{\pi \in \Pi(\overline{p})} [\beta^*_+(\pi)(b^u_F(\pi))^+ - \beta^*_-(\pi)(b^u_F(\pi))^-] \end{aligned}$$

is a homogeneous cost representation.

B.3 Necessity

Since the homogeneous cost representation is a special case of the costly subjective learning representation, it is enough to check Neutral Outcome Independence.

Take any $F \in \mathbb{F}$ and $\alpha \in (0,1)$. Since $b^u_{\{x_0\}}(\pi) = u(x_0) = 0$,

$$\begin{aligned}
&U(\alpha F + (1-\alpha)\{x_0\}) \\
&= \max_{\pi\in\Pi}[\beta_+(\pi)(b^u_{\alpha F+(1-\alpha)\{x_0\}}(\pi))^+ - \beta_-(\pi)(b^u_{\alpha F+(1-\alpha)\{x_0\}}(\pi))^-] \\
&= \max_{\pi\in\Pi}[\beta_+(\pi)(\alpha b^u_F(\pi) + (1-\alpha)b_{\{x_0\}}(\pi))^+ - \beta_-(\pi)(\alpha b^u_F(\pi) + (1-\alpha)b_{\{x_0\}}(\pi))^-] \\
&= \max_{\pi\in\Pi}[\beta_+(\pi)(\alpha b^u_F(\pi))^+ - \beta_-(\pi)(\alpha b^u_F(\pi))^-] \\
&= \alpha \max_{\pi\in\Pi}[\beta_+(\pi)(b^u_F(\pi))^+ - \beta_-(\pi)(b^u_F(\pi))^-] = \alpha U(F).
\end{aligned}$$

Take any $F, G \in \mathbb{F}$ and $\alpha \in (0,1)$. By the above observation,

$$\begin{aligned}
&U(\alpha F + (1-\alpha)\{x_0\}) \geq U(\alpha G + (1-\alpha)\{x_0\}) \\
\Longleftrightarrow\ &\alpha U(F) \geq \alpha U(G) \\
\Longleftrightarrow\ &U(F) \geq U(G),
\end{aligned}$$

that is, the preference U represents satisfies Neutral Outcome Independence.

References

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

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

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

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

[5] Higashi, Y., K. Hyogo, and N. Takeoka (2022), "Costly subjective learning, "working paper.

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

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

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