Coarse Information Acquisition*

Youichiro Higashi, Kazuya Hyogo, Xiangyu Qu, and Norio Takeoka†

April 3, 2023

Abstract

If agents are not given sufficient information when making decisions, they will attempt to obtain more accurate information through information acquisition. The literature of rational inattention hypothesizes that a rational agent optimally chooses an experiment or information structure to obtain an additional piece of information. In reality, however, it is difficult to conduct accurate experiments due to lack of knowledge about the experiment itself and/or ambiguity about the payoff-relevant state space. This paper studies the choice behavior of decision makers who are aware that their information acquisition is not always accurate and that they can only choose coarse experiments. By adopting the choice theoretic model of information acquisition, provided in de Oliveira, Denti, Mihm, and Ozbek [12], we argue that one of their axioms, which is interpreted as preference for early resolution of risk and takes a form of quasi-convexity of preference, excludes the possibility of coarse experiments. By relaxing their quasi-convexity axiom, we axiomatically characterize models of information acquisition with coarse experiments.

Keywords: information acquisition, rational inattention, ambiguity, experiments, Bayes plausibility.

JEL classification: D81

*We would like to thank Yi-Hsuan Lin, Jawwad Noor, Gerelt Tserenjigmid, Chih-Chun Yang, the audiences at AMES 2022 (Keio University), Hitotsubashi Summer Institute 2021, and the seminar participants at Academia Sinica, Boston University, University of California Santa Cruz, and Yale University for their helpful comments. Part of this research was conducted while Norio Takeoka was visiting LEMMA, University of Paris 2 and Department of Economics, Boston University, whose hospitality is gratefully acknowledged. This study is supported by JSPS KAKENHI Grant Number JP21K01386(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; Qu is at CNRS, Centre d'Economie de la Sorbonne, France; 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 individuals are not given sufficient information when making decisions that directly affect their payoffs, they will attempt to obtain more accurate information through information acquisition. The literature of rational inattention (for example, Sims [24]) hypothesizes that a rational economic agent optimally chooses an additional piece of information while considering trade-offs between the benefits obtained from learning and the associated costs and constraints.

More formally, information acquisition is the choice of a function (called experiment) from the state space into a set of distributions on the signal space. By combining a prior on the state space and the Bayes' formula, a posterior can be obtained from the signal. Such information acquisition updates the prior to the posterior, allowing for more accurately informed decision making. Also, as is well known, choosing one experiment is equivalent to choosing one distribution (called an information structure) on the posterior probabilities that is consistent with the prior (Kamenica and Gentzkow [21]).

In reality, it is difficult to conduct accurate experiments for a variety of reasons. For example, consider a situation in which a decision maker seeks advice from an expert. In this case, choosing an expert corresponds to choosing an experiment. If the decision maker does not know precisely what information-gathering technology the expert has or what his/her interests and biases are, then choosing an expert is merely choosing a 'coarse' experiment, i.e., choosing multiple possibilities of experiments rather than single experiment. Also, if the decision maker does not have sufficient information about the state space, it may be difficult to have a probabilistic prior (Ellsberg [14]). If such ambiguity exists, even if the precise experiment is chosen, there may be more than one candidate for the prior distribution, and thus more than one corresponding information structure. In other words, prior ambiguity can be a source of coarse experiments.

This paper studies the choice behavior of decision makers who are aware that information acquisition is not always accurate and that they can only choose coarse experiments. By doing so, we establish a model of information acquisition that takes into account the practical aspect of being able to choose only coarse experiments, may due to ambiguity about information technology or payoff-relevant states, and examine how the perception of coarse experiments makes a difference compared to the standard model of information acquisition. In order to rigorously compare differences in choice behavior, this paper adopts an axiomatic approach. Using the existing literature on axiomatization of information acquisition (de Oliveira, Denti, Mihm, and Ozbek [12]) as a starting point, we first identify their axioms that rule out perception of coarse experiments. Then, we investigate what decision rules or utility representations can be obtained by weakening those axioms.

1.2 Background

To introduce the model, let Ω be a finite set of states and X be the set of lotteries over some prizes. A function $f : \Omega \to X$ is called an act. The set of all acts is denoted by $\mathcal{F}$. A non-empty finite subset $F \subset \mathcal{F}$ is interpreted as an opportunity set or menu. Imagine a situation where the agent chooses a menu and chooses an act from the menu subsequently. A primitive of the model is a preference $\succsim$ over menus.

Prior to the choice of an act from a menu, the agent may subjectively conduct information acquisition, which is interpreted as a subjective optimization among information structures over Ω. Formally, an information structure is $\pi \in \Delta(\Delta(\Omega))$, regarded as a probabilistic distribution over posteriors. In addition, on the subjectively feasible information structures, the Bayesian plausibility condition is required as $p^\pi = \overline{p}$, where p^π is the prior reduced from an information structure π and $\overline{p}$ is the agent's original prior over states. Given an expected utility function $u : X \to \mathbb{R}$ and a menu F, the value of information of π is computed as

$$b^u_F(\pi) = \int_{\Delta(\Omega)} \Big(\max_{f \in F} \sum_{\omega \in \Omega} u(f(\omega))p(\omega) \Big) \mathrm{d}\pi(p).$$

de Oliveira, Denti, Mihm, and Ozbek [12] axiomatically characterize the following representation of $\succsim$, called the rationally inattentive representation: There exist an expected utility $u : X \to \mathbb{R}$, a prior $\overline{p}$ over Ω, a cost function $c : \Pi(\overline{p}) \to \mathbb{R}_+ \cup \{\infty\}$, where $\Pi(\overline{p})$ is the set of information structures satisfying the Bayesian plausibility, such that $\succsim$ is represented by

$$U(F) = \max_{\pi \in \Pi(\overline{p})} \{ b^u_F(\pi) - c(\pi) \}. \tag{1}$$

An advantage of the axiomatic approach is that the set of axioms characterizing a particular representation clearly tells us what type of behavior is accommodated to or excluded from the representation model. One of the axioms characterizing (1), called Aversion to Contingent Planning, indeed excludes coarse experiments. We say that $\succsim$ satisfies Aversion to Contingent Planning (ACP) if for all F, G and $\alpha \in [0, 1]$,

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

where $\alpha F + (1-\alpha)G$ is defined as the menu $\{\alpha f + (1-\alpha)g \mid f \in F,\ g \in G\}$. Mathematically, this axiom requires quasi-convexity on $\succsim$. Since $\alpha F + (1 - \alpha)G$ 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.

As de Oliveira, Denti, Mihm, and Ozbek [12, p.628, footnote 5] correctly point out, ACP rules out a preference for hedging, which is reasonable when the agent is uncertain about the information acquisition technology available in the future. In particular, if $\succsim$ is restricted on the singleton sets, ACP implies that for all acts f, g and $\alpha \in [0, 1]$,

$$\{f\} \sim \{g\} \implies \{f\} \succsim \alpha\{f\} + (1 - \alpha)\{g\},$$

which reveals that the agent does not have preference for hedging uncertainty among states: a conflicting feature to ambiguity aversion of Gilboa and Schmeidler [19]. As stated above, ambiguity in the prior distribution can be a source of coarse experiment, so it can be seen that the ACP axiom, by eliminating the possibility of coarse experiments, also eliminates ambiguity in the prior belief.

1.3 Characterization Results

A choice object in experiments has been modeled by an information structure $\pi \in \Delta(\Delta(\Omega))$ satisfying the Bayes plausibility. For each menu F and each experiment $\pi \in \Pi$, the agent obtains a benefit of information, $b^u_F(\pi)$. In our model, a coarse experiment is captured by a set of information structures $\Pi \in \mathcal{K}(\Delta(\Delta(\Omega)))$, where $\mathcal{K}(Y)$ is the set of all non-empty compact subsets of a compact metric space Y. The agent faces subjective feasibility of coarse experiments. Formally, such a feasible set can be modeled as a set of multiple Π's, that is, any compact subset $\mathbf{\Pi} \subset \mathcal{K}(\Delta(\Delta(\Omega)))$.

For any coarse experiment $\Pi \in \mathbf{\Pi}$, $P^\Pi = \{p^\pi \mid \pi \in \Pi\} \subset \Delta(\Omega)$ is the set of priors induced from each information structure $\pi \in \Pi$. We say that $\mathbf{\Pi}$ satisfies the prior consistency if

$$P^\Pi = P^{\Pi'}$$

for all $\Pi, \Pi' \in \mathbf{\Pi}$. This common multiple priors across coarse experiments, denoted by $\overline{P}$, can be regarded as the agent's ambiguous prior belief over states.

If $\overline{P}$ is a singleton, that is, the agent has a single prior $\overline{p}$, the prior consistency is reduced to

$$p^\pi = \overline{p}$$

for all $\pi \in \Pi$ and $\Pi \in \mathbf{\Pi}$. This is the standard consistency condition for (precise) experiments, called the martingale property or the Bayesian plausibility (see Kamenica and Gentzkow [21]).

After choosing some $\Pi \in \mathbf{\Pi}$, the agent is completely ignorant of which experiments in Π will actually be conducted. Hence, given Π, the benefit of information is computed as

$$\min_{\pi \in \Pi} b^u_F(\pi).$$

The key assumption behind this formulation is that the set of signals obtained from experiments is sufficiently large that the experiment can be identified from the observed signals. Therefore, at the stage where the agent chooses an act from the menu, a probabilistic posterior can be obtained from the precise experiment that was identified.

The following representation, called the Coarse Information Choice (CIC) representation, captures the agent who optimally chooses a coarse experiment given the constraint: there exists a tuple $(u, \overline{P}, \mathbf{\Pi})$ such that $u : X \to \mathbb{R}$ is an expected utility function, $\overline{P}$ is a set of multiple priors, $\mathbf{\Pi} \subset \mathcal{K}(\Delta(\Delta(\Omega)))$ is compact and satisfies the prior consistency, and $\succsim$ is represented by

$$U(F) = \max_{\Pi \in \mathbf{\Pi}} \min_{\pi \in \Pi} b^u_F(\pi). \tag{2}$$

Since Π satisfies the prior consistency, (2) is simply written as a maxmin EU representaion of Gilboa and Schmeidler [19] on the set of singleton menus,

$$U(\{f\}) = \min_{p \in \overline{P}} \sum_{\omega} u(f(\omega))p(\omega).$$

As a special case, we characterize the case where $\overline{P}$ is a singleton, in which case the agent is a subjective utility maximizer on singleton menus. In general, ambiguity about states can be a source of coarse experiments, but, this special case highlights a coarseness of experiments that is not attributed to ambiguity. Even though he/she does not face any ambiguity about prior, the agent is still concerned about coarseness of experiments due to uncertain information technology.

As the representation (2) suggests, preference over menus reflects two effects, the incentive for optimal information choice and the hedging motive against coarseness of experiments, which may conflict to each other. Depending on which effect outweighs the other, attitudes toward contingent planning will change. The CIC representation provides an appropriate foundation for considering various attitudes toward contingent planning. Agents who expect to always be able to choose a precise experiment will exhibit aversion to contingent planning, as in the ACP axiom, because they only care about the incentive for optimal information choice. Conversely, agents who do not anticipate information choices will exhibit preference for contingent planning because they are only concerned about coarseness of experiments. For each case, we obtain an axiomatic characterization as a special form of the CIC representation.

To axiomatize the CIC representation, we first characterize a general representation, called the General CIC representation as an intermediate result. This class of representation has the same form of (2) but Π does not satisfy the prior consistency. To axiomatize the General CIC representation, we borrow techniques from the literature of choice under ambiguity. It is well-known that the representation (1) has a parallel relationship with the variational representation of Maccheroni, Marinacci, and Rustichini [22]. Their functional form is obtained by the conjugate theory, which relies on the concavity of the functional. Since the behavioral meaning of concavity of the functional is the ambiguity aversion, Chandrasekher, Frick, Iijima, and Le Yaouanq [6] and Xia [27] consider a generalization of the functional, called dual-self representations or Boolean representations, by dropping the quasi-concavity to accommodate various attitudes toward ambiguity. Our General CIC representation is a counterpart of their representations. In our models, quasi-concavity or convexity captures attitudes toward contingent planning. The quasi-convexity of preference follows from the optimal choice of information, while the quasi-concavity reflects aversion to coarse experiments. Since agents care about both aspects, their menu preferences generally satisfy neither quasi-convexity nor quasi-concavity.

1.4 Related Literature

Our model treats a set of information structures as coarse experiments. This is not an original idea of our model. In the existing literature, sets of experiments has been interpreted as ambiguous experiments. Çelen [4], Gensbittel, Renou, and Tomala [16], and Wang [26] generalize comparisons of informativeness in experiments, such as the Blackwell order, to admit ambiguous experiments. Çelen [4] considers ambiguity only in priors. Gensbittel, Renou, and Tomala [16] accommodate ambiguity both in priors and experiments by considering a set of joint distributions over states and signals. Wang [26] considers a set of experiments, functions from states into distributions over signals, and obtains a generalization of the Blackwell order regardless of the ambiguity of priors.

A set of experiments are also considered as an ambiguous communication device in the context of mechanism design and persuasion. Bose and Renou [3] consider a situation where a mechanism designer can choose a set of experiments in the communication stage of a mechanism. Ambiguous communication devices can induce ambiguous posteriors even when players have a single prior ex ante. Beauchene, Li, and Li [2] introduce a set of experiments in the model of Bayesian persuasion and show that the sender is strictly better off by choosing an ambiguous communication device.

In the literature on preference over menus, Dillenberger, Lleras, Sadowski, and Takeoka [13] extend the framework of Dekel, Lipman, and Rustichini [11] to preference of menus of acts and derive a subjective information structure from preference. Here, the agent uses a fixed experiment for all menus. To accommodate information acquisition, de Oliveira, Denti, Mihm, and Ozbek [12] generalize Dillenberger, Lleras, Sadowski, and Takeoka [13] and characterize a representation consistent with the rational inattention, where the agent behaves as if he/she optimally chooses an information structure by considering its benefits and costs. Epstein, Marinacci, and Seo [15] relax the Independence axiom and the Indifference to Randomization axiom in Dekel, Lipman, and Rustichini [11] and characterize representations with ambiguity over subjective states and/or coarse subjective states.

2 The Model

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 [12] and Higashi, Hyogo, and Takeoka [20].

- $\Omega = \{\omega_1, ..., \omega_n\}$: 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 $\mathbb{F}$

For any compact metric space Y, let $\mathcal{K}(Y)$ denote the set of all non-empty compact subsets of Y, which is endowed with the Hausdorff metric.

We assume that the agent has in mind the following timing of decisions:

(i) A menu F is chosen.

(ii) An experiment is conducted.

(iii) A signal arrives and the corresponding posterior over Ω is obtained.

(iv) An act $f \in F$ is chosen.

(v) A state ω is resolved and the lottery $f(\omega)$ is realized accordingly.

After choosing a menu, the agent conducts an experiment to acquire an additional piece of information about states. As a consequence of the experiment, the agent obtains a posterior. Afterward, the agent chooses an act from the menu.

Our primitive is a preference in the stage (i). If the agent anticipates the above timeline, the agent's preference over menus will reflect an experiment choice, including the associated feasible set of experiments and the associated costs, relevant in the stage (ii).

2.2 Functional Forms

Assume that the agent has an expected utility $u : X \to \mathbb{R}$ and an ambiguous belief about states, captured by a set of multiple priors $\overline{P} \subset \Delta(\Omega)$. Prior to the choice from a menu, the agent may have the opportunity to acquire information. A probability distribution $\pi \in \Delta(\Delta(\Omega))$ is interpreted as an experiment (information structure) about Ω. Given $u : X \to \mathbb{R}$ and a menu F, the benefit of information from an experiment $\pi \in \Delta(\Delta(\Omega))$ is defined by

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

The agent is not necessarily able to conduct a precise experiment. A coarse experiment is modeled as a subset of precise experiments, that is, $\Pi \in \mathcal{K}(\Delta(\Delta(\Omega)))$. We impose a consistency between the set of multiple priors $\overline{P}$ and a coarse experiment Π. For any π, $p^\pi \in \Delta(\Omega)$ denotes the initial prior associated with π, defined as

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

for each ω. We say that Π satisfies the prior consistency with $\overline{P}$ if

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

If $\overline{P}$ is a singleton, that is, the decision maker has a single prior $\overline{p}$, the prior consistency is reduced to

$$p^\pi = \overline{p}$$

for all $\pi \in \Pi$. This condition is known as the martingale property or the Bayesian plausibility (Kamenica and Gentzkow [21]).

The agent has in mind a subjectively feasible set of coarse experiments, $\mathsf{\Pi} \subset \mathcal{K}(\Delta(\Delta(\Omega)))$, which is modeled as a set of multiple Π's satisfying the prior consistency with multiple priors, that is, $P^\Pi = \overline{P}$ for all $\Pi \in \mathsf{\Pi}$. As described in stage (ii) of the time line, the agent conducts a coarse experiment after choosing a menu F, but the agent is completely ignorant of what experiment is actually conducted within Π. Thus, when facing a menu F, the (gross) benefit of coarse experiment Π is computed according its worst case scenario such as

$$\min_{\pi \in \Pi} b^u_F(\pi).$$

On the other hand, conducting an experiment is costly. To compute the net benefit of Π, the cost associated with the experiment should be subtracted from the gross benefit. We introduce a cost function $c : \mathsf{\Pi} \to \mathbb{R}_+ \cup \{\infty\}$ satisfying (i) $\min_{\Pi \in \mathsf{\Pi}} c(\Pi) = 0$ and (ii) Blackwell-monotonicity: if, for any $\pi \in \Pi$, there exists $\pi' \in \Pi'$ such that π' dominates π in the Blackwell order, then $c(\Pi') \geq c(\Pi)$. Condition (i) is a normalization of the cost function. Condition (ii) is a monotonic property of the cost function with respect to a generalization of the Blackwell order, consistent with Gensbittel, Renou, and Tomala [16] and Wang [26].

Now we are ready to introduce a counterpart of the rational inattentive representation.

Definition 1 $\succsim$ *admits a Costly Coarse Information Choice (Costly CIC) representation if there exists a tuple* $(u, \overline{P}, \mathsf{\Pi}, c)$ *such that* $u : X \to \mathbb{R}$ *is an expected utility function,* $\mathsf{\Pi} \subset \mathcal{K}(\Delta(\Delta(\Omega)))$ *is compact and prior consistent with* $\overline{P}$*,* $c : \mathsf{\Pi} \to \mathbb{R}_+ \cup \{\infty\}$ *is a cost function satisfying conditions (i) and (ii), and* $\succsim$ *is represented by*

$$U(F) = \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} b^u_F(\pi) - c(\Pi)\}. \tag{3}$$

On the singleton menus, by the prior consistency, together with condition (i) of c,

$$\begin{aligned} U(\{f\}) &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} b^u_{\{f\}}(\pi) - c(\Pi)\} \\ &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} \sum u(f(\omega)) p^\pi(\omega) - c(\Pi)\} \\ &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{p \in \overline{P}} \sum u(f(\omega)) p(\omega) - c(\Pi)\} \\ &= \min_{p \in \overline{P}} \sum u(f(\omega)) p(\omega). \end{aligned}$$

Thus, the agent follows a maxmin EU representation of Gilboa and Schmeidler [19] at the ex ante stage.

The following representation is a special case of the Costly CIC representation when $c(\Pi) = 0$ for all $\Pi \in \mathsf{\Pi}$.

Definition 2 $\succsim$ *admits a Coarse Information Choice (CIC) Representation if there exists a tuple $(u, \overline{P}, \mathsf{\Pi})$ such that $u : X \to \mathbb{R}$ is an expected utility function, $\overline{P}$ is a compact set of multiple priors, $\mathsf{\Pi} \subset \mathcal{K}(\Delta(\Delta(\Omega)))$ is a subjectively possible set of coarse experiments satisfying the prior consistency with $\overline{P}$, and $\succsim$ is represented by*

$$U(F) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b_F^u(\pi). \tag{4}$$

Again, the consistency condition between $\overline{P}$ and $\mathsf{\Pi}$ implies that on the singleton sets, the representation is reduced to the maxmin EU representation:

$$U(\{f\}) = \max_{\Pi \in \mathsf{\Pi}} \min_{p \in P^{\Pi}} \sum_{\omega \in \Omega} u(f(\omega))p(\omega) = \min_{p \in \overline{P}} \sum_{\omega \in \Omega} u(f(\omega))p(\omega).$$

2.3 Illustration

In this section, we will use a practical example to describe a possible application of the Costly CIC model. Suppose a policy maker needs to develop an environmental policy. Consider the simplest scenario in which the state Ω of the environment is binary, *good* (ω_1) or *bad* (ω_2). The policy maker is faced with two possible policy choices. (1) Maintain the existing policy, act f. Then if the good state is eventually realized, the payoff is 1. If the bad state is realized, the payoff is -1. (2) Adopt an active environmental policy, act g. Then, no matter what state is realized, the payoff is always 0. (Refer to the following table.)

Ω	ω_1	ω_2
f	1	-1
g	0	0

Due to the complexity of environmental issues, experts use different climate models based on their expertise. Therefore, the probabilities estimated by experts not only differ significantly, but also cannot manage to converge. Therefore, the policy maker's belief over the state is not a probability, but a set of probabilities.

For $a \in [0, 1]$, let $p_a \in \Delta(\Omega)$ be such that $p_a(\omega_1) = a$. Since the state space is binary, we know that $\{p_a : a \in [0, 1]\} = \Delta(\Omega)$. Therefore, assume that the multiple priors of policy maker is

$$\overline{P} = \big\{p_a \,|\, a \in [0.4, 0.6]\big\}.$$

Assume further that policy maker is risk neutral, i.e. $u(x) = x$. Without considering the experiment, the policy maker who obeys the maxmin expected utility will definitely choose policy g over f, since $U(\{f\}) = -0.2 < 0 = U(\{g\})$.

However, the Costly CIC model suggests that the policy maker can actually conduct further experiments on these two options in order to obtain more accurate information and thus make more effective policy choices. It is worth emphasizing that these experiments are coarse. Estimate of the likelihood that various signals might be realized is not a point, but an interval. As Manski, Sanstad, and DeCanio [23] argue, the coarseness frequently stems from partial identification in environmental issues. Therefore, coarse experiments are not singular, but are composed of many possibilities. Let $\delta_{p_a} \in \Delta(\Delta(\Omega))$ denote a second order belief that with probability 1 the first order belief is p_a. First, consider two possible experiments.

$$\Pi_0 = \left\{\alpha\delta_{p_{0.4}} + (1-\alpha)\delta_{p_{0.6}} \,|\, \alpha \in [0,1]\right\},$$
$$\Pi_1 = \left\{\beta\delta_{p_0} + (1-\beta)\delta_{p_1} \,|\, \beta \in [0.4, 0.6]\right\}.$$

Notice that Π_1 is more Blackwell-informative than Π_0. Moreover, it is clear that $P^{\Pi_0} = P^{\Pi_1} = \overline{P}$.

Take $\gamma \in [0,1]$, we define

$$\begin{aligned}\Pi_\gamma &= (1-\gamma)\Pi_0 + \gamma\Pi_1 \\ &= \left\{\beta\gamma\delta_{p_0} + \alpha(1-\gamma)\delta_{p_{0.4}} + (1-\alpha)(1-\gamma)\delta_{p_{0.6}} + (1-\beta)\gamma\delta_{p_1} \,\middle|\, \alpha \in [0,1], \beta \in [0.4, 0.6]\right\}.\end{aligned}$$

Since Π_0 and Π_1 satisfy prior consistency, so is Π_γ for each $\gamma \in [0,1]$. Therefore,

$$\mathsf{\Pi} = \left\{\Pi_\gamma \,|\, \gamma \in [0,1]\right\}$$

is a possible set of coarse experiments satisfying the prior consistency with $\overline{P}$. Since Π_0 is more informative than Π_0, consider a cost function $c(\Pi_\gamma) = \gamma^2$ for convenience.

Now, fix a Π_γ. For $\alpha \in [0,1]$ and $\beta \in [0.4, 0.6]$, let

$$\pi_\gamma(\alpha, \beta) = \beta\gamma\delta_{p_0} + \alpha(1-\gamma)\delta_{p_{0.4}} + (1-\alpha)(1-\gamma)\delta_{p_{0.6}} + (1-\beta)\gamma\delta_{p_1}.$$

Clearly, each $\pi(\alpha, \beta) \in \Pi_\gamma$. Therefore,

$$\begin{aligned}&b^u_{\{f,g\}}(\pi_\gamma(\alpha, \beta)) \\ &= \beta\gamma\max\{-1, 0\} + \alpha(1-\gamma)\max\{-0.2, 0\} + (1-\alpha)(1-\gamma)\max\{0.2, 0\} + (1-\beta)\gamma\max\{1, 0\} \\ &= (1-\alpha)(1-\gamma) \times 0.2 + (1-\beta)\gamma.\end{aligned}$$

Given this observation, it is evident to see

$$\begin{aligned}\min_{\pi \in \Pi_\gamma} b^u_{\{f,g\}}(\pi) = \min_{\alpha,\beta} b^u_{\{f,g\}}(\pi_\gamma(\alpha, \beta)) &= \min_{\alpha,\beta}(1-\alpha)(1-\gamma) \times 0.2 + (1-\beta)\gamma \\ &= 0.4\gamma.\end{aligned}$$

Hence, we have

$$\begin{aligned}U(\{f,g\}) &= \max_{\Pi \in \mathsf{\Pi}}\{\min_{\pi \in \Pi} b^u_{\{f,g\}}(\pi) - c(\Pi)\} \\ &= \max_\gamma 0.4\gamma - \gamma^2 \\ &= 0.04,\end{aligned}$$

where an optimum is achieved at $\gamma = 0.2$. Given Π, the policy maker prefers set of policy $\{f, g\}$ to single policy$f$ or g according to Costly CIC calculation. The insight from this example is that policy makers do not necessarily have to choose between two policies immediately when the possibility of information access exists. Rather, they can run experiment, even if the experiment itself is coarse, and wait for new information to emerge that will help make the choice.

3 Foundation of the CIC Model

3.1 Axioms

We provide a behavioral foundation of the CIC representation. We start with the basic axioms on $\succsim$ that are consistent with any type of 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 $F \sim F \cup \{g\}$.*

Since $F \subset F \cup \{g\}$, the latter menu is weakly preferred by preference for flexibility. If $\{f(\omega)\} \succsim \{g(\omega)\}$ for all $\omega \in \Omega$, however, for all states, f gives a preferred lottery than g does. In this sense, g is dominated by f. No matter what belief the agent has on states, g should not be chosen over f. Thus, adding g to F does not provide a strictly higher value of flexibility than F.

We assume the same unboundedness axiom as in Higashi, Hyogo, and Takeoka [20].

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*

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

Next, we discuss more substantial axioms that are specific to coarse experiment choice. Consider the agent who is aware of coarseness of information acquisition. In general, the evaluation of mixed menu $\alpha F+(1-\alpha)G$ reflects two conflicting effects, optimal information choice and preference for hedging from coarseness of information technology. As explained as an intuition behind the ACP axiom, if two menus are indifferent, the mixing of the two menus is not preferred in terms of optimal information choice, while the same mixing is preferred in terms of hedging because it may smooth out coarseness of information technology or ambiguity across states. Therefore, preference over menus does not necessarily satisfy the independence axiom because the mixing of two menus may alter the incentive for optimal information choice and the hedging motive.

On the other hand, if one of the two conflicting effects is shutdown, only the other effect prevails and the agent exhibits either preference for or avoidance of mixed menus. Furthermore, if both of the two effects are shutdown, then the agent will exhibit neutral attitude toward mixed menus, and hence, the independence axiom holds. We identify such instances below.

When a menu F is mixed with a menu of lotteries C, information acquisition at $\alpha F+(1-\alpha)C$ can be exclusively tailored for F because there is no role of information acquisition at C. Furthermore, the mixture with lotteries does not affect hedging motives from coarse experiments either because the benefit of information at the menu of lotteries C is constant among all experiments. This intuition suggests that the independence axiom should hold when menus are mixed with lotteries.[1]

Axiom 6 (Independence with Lotteries) *For all menus F, G, lottery x, and $\alpha \in (0,1)$*

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

Next, consider a mixture between a menu F and a singleton menu $\{f\}$. Note that there is no role of information acquisition at the singleton. Hence, the incentive for information choice is not altered at the mixed menu $\alpha F+(1-\alpha)\{f\}$. On the other hand, the mixture with the act f may affect hedging motives. For instance, the mixture may smooth out ambiguity across states. Thus, the agent only cares about the hedging motives at $\alpha F+(1-\alpha)\{f\}$. The above argument leads to the following axiom.

Axiom 7 (Weak Concavity) *For any menu F, act f, and $\alpha \in (0,1)$,*

$$F \sim \{f\} \Longrightarrow \alpha F+(1-\alpha)\{f\} \succsim \{f\}.$$

If F is a singleton menu such as $\{g\}$, Weak Concavity is identical with the ambiguity aversion axiom of Gilboa and Schmeidler [19].

[1] It is easy to see that a stronger axiom, where a singleton lottery $\{x\}$ is replaced with a finite set of lotteries C, is a necessary axiom for the CIC representation. Our Theorem 1, given below, shows that Axiom 6 is sufficient for the representation.

Finally, we consider mixing the same menu together. For any menu F and any $\alpha \in [0,1]$, consider the mixed menu $\alpha F + (1-\alpha)F$. In terms of optimal information choice, this mixture does not alter incentives for information acquisition compared with those at the menu F because this mixture is a contingent plan between the same menu. Moreover, this mixture does not affect the hedging motives against coarse experiments either because, for any experiment, its benefit of information is identical between the same menu and there is no additional benefit from their mixture. Hence, we assume

$$F \sim \alpha F + (1-\alpha)F. \tag{5}$$

Note also that (5) is a weak form of the ACP axiom. By definition of the mixture, $F \subset \alpha F + (1-\alpha)F$. By Preference for Flexibility, $\alpha F + (1-\alpha)F \succsim F$. Moreover, under ACP, $F \succsim \alpha F + (1-\alpha)F$. Thus, we have the ranking (5).

On the other hand, (5) has a strong implication on the choice from menus. For an illustration, consider three acts $f = (100, 0)$, $g = (0, 100)$, and their $\frac{1}{2}$-mixture $\frac{1}{2}f + \frac{1}{2}g$. Since this mixed act is a constant act giving 100 with probability 1/2 and 0 otherwise, an ambiguity averse agent seems to prefer this mixed act to both f and g. It is reasonable to assume that

$$\left\{\frac{1}{2}f + \frac{1}{2}g\right\} \succ \{f\}, \{g\}. \tag{6}$$

Note that rankings over singletons reflect the agent's prior belief. On the other hand, when $F = \{f, g\}$ and $\alpha = \frac{1}{2}$, (5) implies

$$\{f, g\} \sim \left\{f, g, \frac{1}{2}f + \frac{1}{2}g\right\}. \tag{7}$$

The ranking (7) suggests that in the ex post stage, $\frac{1}{2}f + \frac{1}{2}g$ is never chosen over f and g, which further suggests that after information acquisition, the agent does not have hedging motives for ambiguity across states presumably because the agent receives a precise signal, whereby an experiment is identified ex post and, consequently the agent can form a probabilistic belief.

The next axiom is a generalization of (5).

Axiom 8 (Indifference to Ex Post Randomization (IEPR)) *For all menus F, $n \in \mathbb{N} \setminus \{0\}$, and $\beta_1, ..., \beta_n \in \mathbb{R}_+$ such that $\sum_{i=1}^n \beta_i = 1$,*

$$F \sim \sum_{i=1}^{n} \beta_i F.$$

IEPR is reminiscent of the Indifferent to Randomization axiom of Dekel, Lipman, and Rustichini [11]. In their model, preference is defined over menus of lotteries, and Indifference to Randomization requires that any menu is indifferent to its convex hull. This axiom can be justified when the ex post choice from a menu is made by maximizing a linear or expected utility. Since we consider menus of finite elements as choice objects, convex hulls cannot be formally taken as choice objects. However, $\sum_i \beta_i F$ can be viewed as an analogue of the convex hull of F. Since $F \subset \sum_{i=1}^n \beta_i F$, the agent does not value flexibility provided by $(\sum_{i=1}^n \beta_i F) \setminus F$. As illustrated prior to the axiom, similar to the justification of Indifference to Randomization, IEPR is justified if the ex post choice from a menu is made by maximizing a linear utility function over acts, which means that the agent receives an accurate signal and forms a probabilistic belief ex post.

The main result of this section is as follows:[2]

Theorem 1 *$\succsim$ satisfies basic axioms, Independence with Lotteries, Weak Concavity, and IEPR if and only if it is represented by a CIC representation $(u, \overline{P}, \mathsf{\Pi})$.*

3.2 Proof Outline

To prove the sufficiency of the theorem, note first that on the singleton menus, Independence with Lotteries and Weak Concavity imply the certainty independence and the ambiguity aversion axiom of Gilboa and Schmeidler [19]. Thus, $\succsim$ admits a maxmin expected utility representation over $\mathcal{F}$ with an expected utility $u : X \to \mathbb{R}$ and multiple priors $\overline{P} \subset \Delta(\Omega)$. 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 $\{x_F\}$ with $F \sim \{x_F\}$.

We follow the construction of the support functions by de Oliveira, Denti, Mihm, and Ozbek [12]. For any $F \in \mathbb{F}$, a support function for F is defined as, for any posterior $p \in \Delta(\Omega)$,

$$\varphi_F(p) = \max_{f \in F} \sum_{\Omega} u(f(\omega))p(\omega). \tag{8}$$

By IEPR, the support function identifies the menu up to indifference: $\varphi_F = \varphi_G \implies F \sim G$. Let $\Phi_{\mathbb{F}} = \{\varphi_F \mid F \in \mathbb{F}\} \subset C(\Delta(\Omega))$ 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 an intermediate step, we first characterize a more general representation, which does not satisfy the prior consistency. We say that $\succsim$ admits a General CIC Representation if there exists a pair $(u, \mathsf{\Pi})$ such that $u : X \to \mathbb{R}$ is an expected utility function and $\mathsf{\Pi} \subset \mathcal{K}(\Delta(\Delta(\Omega)))$ is compact, and $\succsim$ is represented by

$$U(F) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b_F^u(\pi). \tag{9}$$

For each $\Pi \in \mathcal{K}(\Delta(\Delta(\Omega)))$, recall P^{Π} denotes the set of priors induced from Π. On the singletons, (9) is reduced to

$$U(\{f\}) = \max_{\Pi \in \mathsf{\Pi}} \min_{p \in P^{\Pi}} \sum_{\omega} u(f(\omega))p(\omega), \tag{10}$$

[2] Instead of Two-Sided Unboundedness, we can assume One-Sided Unboundedness axiom as in de Oliveira, Denti, Mihm, and Ozbek [12]: 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 $\{y\} \succ \{\alpha z + (1 - \alpha)x\}$.

which is called the dual-self representation or Boolean representation, studied in Chandrasekher, Frick, Iijima, and Le Yaouanq [6] and Xia [27]. Note that the above representation (10) involves an aspect of belief manipulation or a choice of optimistic beliefs unless the prior consistency with $\overline{P}$ is imposed.

To impose the desired consistency, the Weak Concavity axiom plays a central role.

3.3 Uniqueness

We describe the uniqueness properties of a CIC representation with $(u, \overline{P}, \mathsf{\Pi})$. We use the idea of half-space closure, introduced by Chandrasekher, Frick, Iijima, and Le Yaouanq [6] to establish the uniqueness of the dual-self representation. On top of their idea, we impose an addtional requirement of the prior consistency on the CIC representation.

For a CIC representation $(u, \overline{P}, \mathsf{\Pi})$, consider the set of experiments whose prior belongs to the set of multiple priors P, denoted by $\Pi(\overline{P}) = \{\pi \in \Delta(\Delta(\Omega)) \,|\, p^\pi \in \overline{P}\}$. Given $\mathsf{\Pi}$, define its half-space closure by

$$\overline{\mathsf{\Pi}} \equiv \mathrm{cl}\{H \cap \Pi(\overline{P}) \,|\, H \text{ is a closed half-space in } \Delta(\Delta((\Omega)) \text{ and } H \supset \Pi \text{ for some } \Pi \in \mathsf{\Pi}\},$$

where we call H a closed half-space in $\Delta(\Delta((\Omega))$ if $H = H_{\varphi,\lambda} \equiv \{\pi \in \Delta(\Delta(\Omega)) \,|\, \langle \varphi, \pi \rangle \geq \lambda\}$ for some $\lambda \in \mathbb{R}$ and some continuous function $\varphi : \Delta(\Omega) \to \mathbb{R}$. Chandrasekher, Frick, Iijima, and Le Yaouanq [6] uses these half-spaces H in $\Delta(\Delta(\Omega))$ to define the half-space closure. To take into account the consistency with priors, we additionally consider the intersection between H and $\Pi(\overline{P})$. Since $\mathsf{\Pi}$ satisfies the prior consistency with $\overline{P}$, we have $\Pi \subset H \cap \Pi(\overline{P})$. Thus, each $\Pi \in \mathsf{\Pi}$ is expanded in this manner to obtain its half-space closure.

Theorem 2 *Suppose $(u, \overline{P}, \mathsf{\Pi})$ is a CIC representation of $\succsim$. Then, following properties hold:*

(1) $(u, \overline{P}, \overline{\mathsf{\Pi}})$ is also a CIC representation of $\succsim$;

(2) For any expected utility function u', multiple priors $\overline{P}'$, and collection $\mathsf{\Pi}'$ of coarse experiments consistent with $\overline{P}'$, $(u', \overline{P}', \mathsf{\Pi}')$ is a CIC representation of $\succsim$ if and only if u' is a positive affine transformation of u, $\overline{P}' = \overline{P}$, and $\overline{\mathsf{\Pi}} = \overline{\mathsf{\Pi}'}$.

Theorem 2 ensures that the set of coarse experiments are pinned down up to the half-space closure.

3.4 Interpersonal Comparison

Given the uniqueness property of the CIC representation, we provide a comparative notion of attitude toward flexibility in terms of behavior and characterize its implication on the CIC representation. 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 by Dillenberger, Lleras, Sadowski, Takeoka [13], de Oliveira, Denti, Mihm, and Ozbek [12], and Higashi, Hyogo, and Takeoka [20].

Definition 3 $\succsim^1$ *is more averse to commitment than* $\succsim^2$ *if for all* $F \in \mathbb{F}$ *and* $f \in F$,

$$F \succsim^2 \{f\} \implies F \succsim^1 \{f\}.$$

Theorem 3 *Suppose that for* $i = 1, 2$, $\succsim^i$ *is represented by a CIC representation with* $(u^i, \overline{P}^i, \sqcap^i)$. *The following statements are equivalent:*

(a) $\succsim^1$ *is more averse to commitment than* $\succsim^2$*;*

(b) u^2 *is a positive affine transformation of* u^1, $\overline{P}^1 = \overline{P}^2$, *and* $\overline{\sqcap}^1 \supset \overline{\sqcap}^2$

Under the CIC representation, the condition $\overline{\sqcap}^1 \supset \overline{\sqcap}^2$ means that agent 1 has a larger constraint set of coarse experiments than agent 2. Hence, agent 1 values the flexibility provided by menu F more than agent 2.

4 Special Cases

4.1 Coarse Information Choice with a Single Prior

As a special case, suppose $\overline{P} = \{\overline{p}\}$. If $\sqcap$ is prior-consistent with $\{\overline{p}\}$, the consistency condition boils down to

$$p^{\pi} = \overline{p}$$

for all $\Pi \in \sqcap$ and all $\pi \in \Pi$.

Definition 4 $\succsim$ *admits a Coarse Experiment Choice (CEC) Representation if* $\succsim$ *admits a CIC representation* $(u, \{\overline{p}\}, \sqcap)$.

Note that on the singletons, U is reduced to an SEU representation,

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

The axiomatic foundation of this special case is obtained by strengthening both Independence with Lotteries and Weak Concavity. Recall the discussion prior to Weak Concavity and consider a mixture between a menu F and a singleton menu $\{f\}$. Since there is no role of information acquisition at the singleton, the incentive for information choice is not altered at $\alpha F + (1 - \alpha)\{f\}$. Moreover, if the agent has a single prior over states, the mixture with the act f does not affect hedging motives either. Thus, the independence axiom should hold when menus are mixed with singleton acts.

Axiom 9 (Independence with Singleton Acts (ISA)) *For all menus* F, G, *act* f, *and* $\alpha \in (0, 1)$

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

This axiom implies Independence with Lotteries and Weak Concavity. Moreover, it also implies the independence axiom on singletons, referred to as Singleton Independence: for all acts f, g, h, and $\alpha \in (0,1)$,

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

By Singleton Independence, the commitment ranking is represented by a subjective expected utility. This special case highlights the coarseness of experiments that is not attributable to prior ambiguity. Even though the agent is sure of prior over states, he/she still faces coarseness of experiments.

Theorem 4 *$\succsim$ satisfies basic axioms, ISA, and IEPR if and only if it is represented by a CEC representation $(u, \overline{p}, \Pi)$.*

For the uniqueness of the CEC representation, the same argument works if $\Pi(\overline{P}) = \{\pi \in \Delta(\Delta((\Omega)) \,|\, p^\pi \in \overline{P}\}$ defined in Section 3.3 is replaced with $\Pi(\overline{p}) = \{\pi \in \Delta(\Delta(\Omega)) \,|\, p^\pi = \overline{p}\}$. As shown in Theorem 2, u is unique up to positive affine transformation, $\overline{p}$ is unique, and the set of coarse experiments is identified up to the half-space closure.

4.2 Attitudes toward Contingent Planning

As discussed in Section 2, the evaluation of the mixed menus reflects two conflicting effect; the incentives for optimal information choice and the hedging motives against coarseness of experiments. The CIC representation accommodates both effects and sets a stage for studying various attitudes toward mixing menus or equivalently contingent planning. By assuming a particular attitude for contingent plannings, we characterize the corresponding special cases of the CIC representation.

In this subsection, we assume that $\succsim$ admits a CIC representation or equivalently satisfies all the axioms for the CIC representation. Take two indifferent menus F and G and any $\alpha \in (0,1)$. If the agent anticipates information acquisition after menu choice, a contingent planning $\alpha F + (1-\alpha)G$ is less preferred to F or G, while it is more preferred in terms of hedging coarse experiments. In general, the ranking between $\alpha F + (1-\alpha)G$ and F depends on which one of the two effects dominates the other and varies across different pairs of menus and the coefficient α. The following definition classifies attitudes toward contingent planning coherent across all pairs of menus and coefficients.

Definition 5 *(1) $\succsim$ exhibits aversion to contingent planning (ACP) if for all menus F and G and all $\alpha \in [0,1]$,*

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

(2) $\succsim$ exhibits preference for contingent planning (PCP) if for all menus F and G and all $\alpha \in [0,1]$,

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

(3) $\succsim$ *exhibits indifference to contingent planning (ICP) if for all menus* F *and* G *and all* $\alpha \in [0,1]$,

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

Part (1) states that the agent always avoids contingent planning, which is introduced as an axoim by de Oliveira, Denti, Mihm, and Ozbek [12]. As discussed in Section 1.2, ACP excludes the possibility of coarse experiments. Thus, the agent only cares about the incentives for optimal information choice.

Part (2) is the opposite behavior to part (1) and states that the agent always prefers contingent planning. This condition suggests that the hedging motive always dominates the incentive for optimal information acquisition presumably because the agent does not care about the latter aspect.

Finally, part (3) is obtained by assuming parts (1) and (2) simultaneously and states that the agent is always indifferent between a contingent planning and the original menu. The agent exhibits this neutral attitude toward contingent planning because the agent ignores both of the two effects.

We have the following characterizations.

Corollary 1 *Assume that* $\succsim$ *satisfies all the axioms of Theorem 1.*

(1) $\succsim$ *exhibits ACP if and only if it is represented by a CIC model with* $(u, \overline{p}, \{\{\pi\} | \pi \in \Pi\})$, *that is,*

$$U(F) = \max_{\pi \in \Pi} b_F^u(\pi). \tag{11}$$

(2) $\succsim$ *exhibits PCP if and only if it is represented by a CIC model with* $(u, \overline{P}, \{\Pi\})$, *that is,*

$$U(F) = \min_{\pi \in \Pi} b_F^u(\pi). \tag{12}$$

(3) $\succsim$ *exhibits ICP if and only if it is represented by a CIC model with* $(u, \overline{p}, \{\{\pi\}\})$, *that is,*

$$U(F) = b_F^u(\pi). \tag{13}$$

From part (1), an implication of ACP is that the agent only anticipates precise experiments $\{\pi\}$ as alternatives at the stage of information acquisition. Hence, we can interpret that the agent simply chooses a precise information structure π within a constraint. The CIC representation boils down to the formulation (11). This is a special case of the standard costly information acquisition, studied in de Oliveira, Denti, Mihm, and Ozbek [12], called the constrained information model. For the characterization, IEPR is redundant because it is implied from ACP. Note also that ACP and Weak Concavity jointly imply the neutrality to mixing two acts, and hence, the agent has a single prior $\overline{p}$ in this class.

Conversely, an implication of PCP is that the agent anticipates only one coarse experiment Π as a feasible option at the stage of information acquisition. Since information choice is always degenerate at Π, he/she does not respond to the incentive for optimal information choice. On the other hand, the agent faces coarseness of experiment Π and is completely ignorant about what experiment in Π is conducted. Hence, in this case, the CIC representation boils down to the formulation (12), regarded as the fixed coarse information model. Regarding the characterization, Weak Concavity is redundant because it is implied from PCP. Since PCP involves ambiguity aversion over states, the model accommodates multiple priors $\overline{P}$.[3]

Finally, part (3) is a joint implication of parts (1) and (2). If the agent exhibits ICP, he/she cares about neither the incentive for optimal information acquisition nor the hedging motives for coarse experiments. Consequently, the agent anticipates neither information choice nor coarse experiments. Thus, the CIC representation boils down to (13), where a single precise experiment π is adopted commonly across all menus. This class of representation is characterized by Dillenberger, Lleras, Sadowski, and Takeoka [13], called the subjective learning representation or the fixed information model.

5 Application: Optimal Sampling

The state space $\Omega = \Omega_1 \times \Omega_2$ is taken to be $\mathbb{R}^2$. The prior over Ω_i is given by a normal distribution $\omega_i \sim N(\mu_i, 1/\tau_i)$, where μ_i is the mean, and $\tau_i > 0$ is the precision.

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

$$u(y_1, y_2, \omega_1, \omega_2) = a(\omega_1 + \omega_2) - b|\omega_1 - y_1| - b|\omega_2 - y_2|, \ a > 0, \ b > 0.$$

A choice variable is y_i, which is interpreted as an investment decision depending on Ω_i. This payoff function takes its maximum at $y_i = \omega_i$, $i = 1, 2$, and the closer the investment decision is to the true state, the higher the payoff. Moreover, for all fixed y_i, higher states ω_i imply higher payoffs. As the payoffs change according to realization of ω_i, a choice of y_i is interpreted as a choice of act.

The agent can observe signals, whereby, the prior is updated to a posterior according to Bayes' rule. The signal s_i is correlated with ω_i according to a normal distribution $s_i \sim N(\omega_i, 1/\sigma)$, where $\sigma > 0$ is the precision of the signal. However, because of resource constraints, the agent can only acquire information about either one of Ω_1 and Ω_2. If signals about Ω_i are observed for n times, the value of information is given by

$$b^u(i, n; \tau_1, \tau_2) = \int \max_{y_1, y_2} \int u(y_1, y_2, \omega_1, \omega_2) \, \mathrm{d}p(\omega | s_1, \cdots, s_n) \, \mathrm{d}\pi_i^n(s_1, \cdots, s_n), \tag{14}$$

[3] If Singleton Independence is imposed in addition, P is forced to be a single prior $\{\overline{p}\}$. Even in this case, Π is not necessarily degenerate to a singleton; the agent still faces coarseness of experiments.

where $p(\omega|s_1,\cdots,s_n)$ is a posterior over Ω conditional upon the realization of signals $s_1,\cdots,s_n$, and $\pi_i^n(s_1,\cdots,s_n)$ is an ex ante probability of the signal realization up to n when the agent acquires information about Ω_i.

In this setting, an information structure is identified with a number of times for signal observations regarding Ω_i, denoted by (i,n). A more informative signal structure is obtained by greater sample size.

Following Cukierman [10], (21) is written as

$$b^u(i,n;\tau_1,\tau_2)=a(\mu_1+\mu_2)-b\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left\{\left(\frac{1}{\tau_i+n\sigma}\right)^{\frac{1}{2}}+\left(\frac{1}{\tau_{-i}}\right)^{\frac{1}{2}}\right\},$$

where π is the circular constant. Furthermore, to ensure that $b^u(i,n;\tau_1,\tau_2)>0$ for all n, throughout this section, we assume

$$\frac{a(\mu_1+\mu_2)}{b\left(\frac{2}{\pi}\right)^{\frac{1}{2}}}>\left(\frac{1}{\tau_1}\right)^{\frac{1}{2}}+\left(\frac{1}{\tau_2}\right)^{\frac{1}{2}}, \tag{15}$$

which requires that the sum of standard deviations of priors over Ω_1 and Ω_2 is smaller than some threshold.

Now assume that a prior over Ω is given by $(\tau_1,\tau_2)=(\overline{\tau},\overline{\tau})$ for some $\overline{\tau}>0$ satisfying (15).

5.1 Single Prior

Consider the situation where a prior over one of Ω_1 and Ω_2 becomes less precise and changes from $\overline{\tau}$ to $\tau<\overline{\tau}$. That is, a prior over Ω is now given as either $(\tau,\overline{\tau})$ or $(\overline{\tau},\tau)$. We still assume (15) for τ and $\overline{\tau}$. Since the argument is symmetric, we take $(\tau_1,\tau_2)=(\tau,\overline{\tau})$.

Below, we characterize an optimal sample size. The agent solves

$$\max_{(i,n)}\left\{b^u(i,n;\tau,\overline{\tau})-cn\right\},$$

where $c>0$ is a constant marginal cost of sampling.

Lemma 1 *If $(\tau_1,\tau_2)=(\tau,\overline{\tau})$ with $\tau<\overline{\tau}$, for all n, $b^u(1,n;\tau,\overline{\tau})>b^u(2,n;\tau,\overline{\tau})$.*

Proof. By definition of $b^u(i,n;\tau,\overline{\tau})$, it suffices to show that

$$\left(\frac{1}{\overline{\tau}+n\sigma}\right)^{\frac{1}{2}}+\left(\frac{1}{\tau}\right)^{\frac{1}{2}}>\left(\frac{1}{\tau+n\sigma}\right)^{\frac{1}{2}}+\left(\frac{1}{\overline{\tau}}\right)^{\frac{1}{2}}. \tag{16}$$

Since the function $f(z)=\left(\frac{1}{z}\right)^{\frac{1}{2}}$ is strictly decreasing and strictly convex, we have

$$\left(\frac{1}{\tau}\right)^{\frac{1}{2}}>\left(\frac{1}{\overline{\tau}}\right)^{\frac{1}{2}}\text{ and }\left(\frac{1}{\overline{\tau}+n\sigma}\right)^{\frac{1}{2}}<\left(\frac{1}{\tau+n\sigma}\right)^{\frac{1}{2}},$$

but the former dominates the latter. Hence we have (16). ■

Lemma 1 implies that the agent can obtain a higher benefit of information by acquiring information from states with a less precise prior. Thus, the information acquisition problem is reduced to

$$\begin{aligned} &\max_n \{b^u(1, n; \tau, \overline{\tau}) - cn\} \\ \iff &\max_n a(\mu_1 + \mu_2) - b\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left\{\left(\frac{1}{\tau + n\sigma}\right)^{\frac{1}{2}} + \left(\frac{1}{\overline{\tau}}\right)^{\frac{1}{2}}\right\} - cn. \end{aligned}$$

For simplicity, let us treat n as a continuous variable. Then, the FOC is given by

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

Clearly, if τ decreases, $\frac{\mathrm{d}b^u}{\mathrm{d}n}(n)$ shifts up, and hence, the optimal sample size increases. Hence, as the precision of a prior decreases, the agent acquires more information.

Proposition 1 *When an initial prior becomes less precise in the sense of mean-preserving spread, the agent acquires more information about less precise states.*

5.2 Multiple Priors

Next, consider the situation where a prior becomes more ambiguous and changes from the single prior $(\overline{\tau}, \overline{\tau})$ to multiple priors $\overline{P} = \{(\overline{\tau}, \overline{\tau}), (\tau, \overline{\tau}), (\overline{\tau}, \tau)\}$ with $\tau < \overline{\tau}$.

The agent decides from which state space and how many samples to observe. Choosing (i, n) corresponds to choosing a coarse experiment

$$\Pi(i, n) = \{\pi_i^n(\cdot|p) \,|\, p \in \overline{P}\}.$$

Let $\mathsf{\Pi} = \{\Pi(i, n) \,|\, i = 1, 2,\ n \geq 0\}$.

Below, we characterize an optimal sample size. The agent solves

$$\max_{\Pi(i,n) \in \mathsf{\Pi}} \left\{\min_{\pi_i^n(\cdot|p) \in \Pi(i,n)} b^u(\pi_i^n(\cdot|p)) - c(\Pi(i, n))\right\},$$

or equivalently,

$$\max_{(i,n)} \left\{\min_{p \in \overline{P}} b^u(i, n; p) - cn\right\},$$

where $c > 0$ is a constant marginal cost of sampling.

By Lemma 1, for all n,

$$\begin{aligned} &b^u(1, n; \overline{\tau}, \overline{\tau}) = b^u(2, n; \overline{\tau}, \overline{\tau}) > b^u(1, n; \tau, \overline{\tau}) \\ &= b^u(2, n; \overline{\tau}, \tau) > b^u(2, n; \tau, \overline{\tau}) = b^u(1, n; \overline{\tau}, \tau), \end{aligned}$$

which implies

$$\min_{p\in\overline{P}} b^u(1,n;p) = b^u(1,n,;\overline{\tau},\tau) = b^u(2,n,;\tau,\overline{\tau}) = \min_{p\in\overline{P}} b^u(2,n;p).$$

Therefore, the information acquisition problem is reduced to

$$\max_n a(\mu_1+\mu_2) - b\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\left\{\left(\frac{1}{\overline{\tau}+n\sigma}\right)^{\frac{1}{2}} + \left(\frac{1}{\tau}\right)^{\frac{1}{2}}\right\} - cn.$$

Then, the FOC is given by

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

Since $\tau < \overline{\tau}$, an optimal sample size under ambiguous priors is reduced compared to when the prior becomes less precise in the sense of risk.

Proposition 2 *When an initial prior becomes less precise in the sense of multiple priors, the agent does not acquire additional pieces of information compared to when the prior is not ambiguous.*

6 Discussion

As assumed in the rational inattention model, information acquisition is often costly. In this section, we discuss some necessary axioms for the Costly CIC representation.

The next axiom is provided by de Oliveira, Denti, Mihm, and Ozbek [12] for characterizing their rationally inattentive representation.

Axiom 10 (Independence of Degenerate Decisions (IDD)) *For all menus F, G, acts f, g, and $\alpha \in (0,1)$,*

$$\alpha F + (1-\alpha)\{f\} \succsim \alpha G + (1-\alpha)\{f\} \implies \alpha F + (1-\alpha)\{g\} \succsim \alpha G + (1-\alpha)\{g\}.$$

To understand an intuition of the axiom, note that when facing the mixed menu $\alpha F + (1-\alpha)\{f\}$, the agent only focuses on the menu αF in terms of information acquisition because there is no role of information acquisition at a commitment menu $\{f\}$. To rank the two mixed menus in the presumption of the axiom, the agent effectively compares αF and αG, and hence, the raking should be preserved when f is replaced with another act g.

In the functional form (3), IDD is not a necessary axiom. As explained above, when facing the mixed menu $\alpha F+(1-\alpha)\{f\}$, the agent only focuses on the menu αF in terms of information acquisition, while he/she also cares about the mixture between acts because of the hedging motives. Indeed, it may happen that if F is preferred to G when both menus are mixed with a constant act (or lottery), while the ranking is reversed when those are mixed with a non-constant act. However, the hedging motive disappears in the contingent planning between menus and lotteries because the prior does not matter for the evaluation of the latter. Indeed, (3) satisfies the following weakening of IDD.

Axiom 11 (Weak Independence with Lotteries) *For all menus F, G, lotteries x, y, and $\alpha \in (0,1)$,*

$$\alpha F + (1-\alpha)\{x\} \succsim \alpha G + (1-\alpha)\{x\} \implies \alpha F + (1-\alpha)\{y\} \succsim \alpha G + (1-\alpha)\{y\}.$$

Axiom 12 (Star-Shaped) *For all menus F and lotteries x, and $\alpha \in (0,1)$,*

$$\{\alpha x_F + (1-\alpha)x\} \succsim \alpha F + (1-\alpha)\{x\}.$$

Note that if $\succsim$ satisfies Independence with Lotteries,

$$\{\alpha x_F + (1-\alpha)x\} \sim \alpha F + (1-\alpha)\{x\}.$$

Axiom 13 (Strong Star-Shaped) *For all menus F and acts f, and $\alpha \in (0,1)$,*

$$\{\alpha x_F + (1-\alpha)f\} \succsim \alpha F + (1-\alpha)\{f\}.$$

This axiom implies Weak Convexity: $F \sim \{f\}$ implies $\alpha F + (1-\alpha)\{f\} \precsim \{f\}$. Note also that if $\succsim$ satisfies Independence with Acts,

$$\{\alpha x_F + (1-\alpha)f\} \sim \alpha F + (1-\alpha)\{f\}.$$

The following proposition gives some necessary axioms for Costly CIC representations.

Proposition 3 *(1) If $\succsim$ admits a Costly CIC representation (u, P, Π, c), then $\succsim$ satisfies the basic axioms, Weak Independence with Lotteries, IEPR, and Star-Shaped.*

(2) If $\succsim$ admits a Costly CIC representation $(u, \{\overline{p}\}, \Pi, c)$, then $\succsim$ additionally satisfies IDD and Strong Star-Shaped.

Appendix

A Preliminaries

A.1 Properties of Functionals

Following de Oliveira, Denti, Mihm, and Ozbek [12], 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 $\alpha \in (0,1)$; *translation invariant* if $V(\varphi + k\mathbf{1}) = V(\varphi) + k$ for all $\varphi \in \Psi$ and $k \in \mathbb{R}$ such that $\varphi + k\mathbf{1} \in \Psi$; *positively homogeneous* if $V(\alpha\varphi) = \alpha V(\varphi)$ for all $\varphi \in \Psi$ and $\alpha \geq 0$; *niveloid* if $V(\varphi) - V(\psi) \leq \sup\{\varphi(p) - \psi(p) \,|\, p \in \Delta(\Omega)\}$.

Cerreia-Vioglio, Maccheroni, Marinacci, Rustichini [5] analyzed the properties of niveloids including the following results. If V is a niveloid, it is monotone and translation invariant. The converse holds when Ψ is a tube, that is, $\Psi = \Psi + \mathbb{R}$. If V is a niveloid, V is Lipschitz continuous. If V is a niveloid, there is a niveloid that extends V to $C(\Delta(\Omega))$.

- Φ: the set of convex functions in $C(\Delta(\Omega))$
- Φ^*: the dual cone of Φ, that is,

$$\{\pi \in ca(\Delta(\Omega)) \,|\, \langle \varphi, \pi \rangle \geq 0 \text{ for all } \varphi \in \Phi\}.$$

  The set Φ^* is also a closed convex cone such that $0 \in \Phi^*$.
- 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 $\Phi_{\mathbb{F}}$:

1. $\Phi_{\mathbb{F}} + \mathbb{R} = \Phi_{\mathbb{F}}$
2. $\alpha\varphi_F \in \Phi_{\mathbb{F}}$ for every $\alpha \geq 0$
3. The set $\Phi_{\mathbb{F}}$ is dense in Φ.

A.2 Clarke Derivative and Differential

Consider a locally Lipschitz functional $V : C(\Delta(\Omega)) \to \mathbb{R}$. For every $\varphi, \xi \in C(\Delta(\Omega))$, *Clarke upper directional derivative* of V at φ in the direction of ξ is defined by

$$V^{\circ}(\varphi;\xi) \equiv \limsup_{\substack{\psi\to\varphi \\ t\searrow 0}} \frac{V(\psi + t\xi) - V(\psi)}{t}.$$

The *Clarke differential* of V at φ in the direction of ξ is defined by

$$\partial V(\varphi) = \{\chi \in ca(\Delta(\Omega)) \,|\, \langle \xi, \chi\rangle \leq V^{\circ}(\varphi;\xi), \forall \xi \in C(\Delta(\Omega))\}.$$

Following the literature, we use Clarke upper directional derivative to define Clarke differential, but an alternative notion is also useful. Consider a locally Lipschitz functional $V : C(\Delta(\Omega)) \to \mathbb{R}$. For every $\varphi, \xi \in C(\Delta(\Omega))$, *Clarke lower directional derivative* of V at φ in the direction of ξ is defined by

$$V_{\circ}(\varphi;\xi) \equiv \liminf_{\substack{\psi\to\varphi \\ t\searrow 0}} \frac{V(\psi + t\xi) - V(\psi)}{t}.$$

Since $V^{\circ}(\varphi;\xi) = -V_{\circ}(\varphi;-\xi)$ for any $\varphi, \xi \in C(\Delta(\Omega))$, the Clarke differential of V at φ in the direction of ξ is rewritten as

$$\partial V(\varphi) = \{\chi \in ca(\Delta(\Omega)) \,|\, \langle \xi, \chi\rangle \geq V_{\circ}(\varphi;\xi), \forall \xi \in C(\Delta(\Omega))\}.$$

Proposition A.3 in Ghirardato, Maccheroni, and Marinacci [17] shows the following results:[4]

Lemma 2 *If $V : C(\Delta(\Omega)) \to \mathbb{R}$ is locally Lipschitz and positively homogeneous, $\partial V(\varphi) \subset \partial V(\mathbf{0})$ for any $\varphi \in C(\Delta(\Omega))$.*

Lemma 3 *If $V : C(\Delta(\Omega)) \to \mathbb{R}$ is locally Lipschitz, monotone, and translation invariant, $\partial V(\varphi) \subset \Delta(\Delta(\Omega))$ for any $\varphi \in C(\Delta(\Omega))$.*

Lemma 4 *If $V : C(\Delta(\Omega)) \to \mathbb{R}$ is locally Lipschitz and positively homogeneous, $V^{\circ}(\mathbf{0};\xi) = \sup_{\varphi\in C(\Delta(\Omega))}\{V(\varphi+\xi) - V(\varphi)\}$ and $V_{\circ}(\mathbf{0};\xi) = \inf_{\varphi\in C(\Delta(\Omega))}\{V(\varphi+\xi) - V(\varphi)\}$ for any $\xi \in C(\Delta(\Omega))$.*

By Proposition 2.1.2 in Clarke [9] and $V^{\circ}(\varphi;\xi) = -V_{\circ}(\varphi;-\xi)$ for any $\varphi, \xi \in C(\Delta(\Omega))$, we have the following:

Lemma 5 *If $V : C(\Delta(\Omega)) \to \mathbb{R}$ is locally Lipschitz, $V^{\circ}(\eta;\xi) = \max_{\pi\in\partial V(\eta)}\langle \xi, \pi\rangle$ and $V_{\circ}(\eta;\xi) = \min_{\pi\in\partial V(\eta)}\langle \xi, \pi\rangle$ for any $\eta, \xi \in C(\Delta(\Omega))$.*

[4] See Proposition 47 and Lemma 48 in Ghirardato, Maccheroni, and Marinacci [18] for proofs.

We describe a characterization of Clarke differential. A functional V is *Gatèaux differentiable* at $\varphi \in C(\Delta(\Omega))$ if

1. $V'(\varphi;\xi) = \lim_{t\to 0} \frac{V(\varphi+t\xi)-V(\varphi)}{t}$ exists for any $\xi \in C(\Delta(\Omega))$;
2. the mapping $\xi \mapsto V'(\varphi;\xi)$ is a continuous linear functional from $C(\Delta(\Omega))$ to $\mathbb{R}$.

If V is Gatèaux differentiable at φ, its *Gatèaux derivative* at φ is denoted by $\nabla V(\varphi)$ and $V'(\varphi;\xi) = \langle \xi, \nabla V(\varphi)\rangle$ for any $\xi \in C(\Delta(\Omega))$.

A Borel subset N of $C(\Delta(\Omega))$ is *Haar-null* set if there exits a (not necessarily unique) probability measure μ on the Borel σ-algebra of $C(\Delta(\Omega))$ such that $\mu(\varphi + N) = 0$ for any $\varphi \in C(\Delta(\Omega))$; see Christensen [7] and Appendix A.3. in Ghirardato, Maccheroni, and Marinacci [17]. Haar-null sets are closed under translation and countable unions. In finite dimensions, Haar-null sets coincides with the sets of Lebesgue measure 0. If N is Haar-null in $C(\Delta(\Omega))$, $C(\Delta(\Omega)) \setminus N$ is dense in $C(\Delta(\Omega))$; see Thibault ([25]). An extension of Rademacher's theorem by Christensen [8] implies that if V is locally Lipschitz, there exists $D \subset C(\Delta(\Omega))$ such that $C(\Delta(\Omega)) \setminus D$ is Haar-null and V is Gatèaux differentiable on D. Given this result, Proposition 2.2 in Thibault [25] yields the following result on Clarke differentials.[5]

Lemma 6 (Thibault [25], Proposition 2.2) *Let V be a locally Lipschitz functional defined on $C(\Delta(\Omega))$. Then, there exists $D \subset C(\Delta(\Omega))$ such that $C(\Delta(\Omega)) \setminus D$ is Haar-null, V is Gatèaux differentiable on D, and for any $\varphi \in C(\Delta(\Omega))$, we have that*

$$\partial V(\varphi) = \overline{co}\{\lim_{n\to\infty} \nabla V(\varphi_n) \,|\, \varphi_n \in D, \varphi_n \to \varphi\}, \tag{17}$$

where $\overline{co}$ denotes weak closure of convex hull and* lim *denotes weak* limit.*

A.3 Boolean Representation

Chandrasekher, Frick, Iijima, and Le Yaouanq [6] provides a "Boolean" representation of a locally Lipschitz function defined on finite dimensional spaces. Based on their result, we derive a Boolean representation of a locally Lipschitz functional $V : C(\Delta(\Omega)) \to \mathbb{R}$. Let D be a subset of $C(\Delta(\Omega))$ in Lemma 6 and $\nabla V(\varphi)$ be Gatèaux derivative of V at $\varphi \in C(\Delta(\Omega))$.

Slightly modifying the proof of Lemma A.5 in Chandrasekher, Frick, Iijima, and Yaouanq [6], we obtain the following lemma.

Lemma 7 *For any $\varphi, \psi \in D$ and $\varepsilon > 0$, there exists $\xi \in D$ such that*

$$V(\xi) - V(\psi) + \langle \psi - \xi, \nabla V(\xi)\rangle \geq 0, V(\xi) - V(\varphi) + \langle \varphi - \xi, \nabla V(\xi)\rangle \leq \varepsilon.$$

[5] Christensen and Thibault obtain their results for separable Banach spaces. Since $\Delta(\Omega)$ is compact, Lemma 3.99 in Aliprantis and Border [1] implies that $C(\Delta(\Omega))$ is separable.

Proof. Pick any $\varphi, \psi \in D$ and $\varepsilon > 0$. Let $m \equiv V(\psi) - V(\varphi)$. If $\langle \psi - \varphi, \nabla V(\varphi)\rangle \geq m$, we can take $\xi = \varphi$. If $\langle \psi - \varphi, \nabla V(\psi)\rangle \geq m$, we can take $\xi = \psi$. Consider the case

$$\langle \psi - \varphi, \nabla V(\varphi)\rangle, \langle \psi - \varphi, \nabla V(\psi)\rangle < m. \tag{18}$$

We define a function $H : \mathbb{R} \to \mathbb{R}$ by

$$H(\lambda) \equiv V(\varphi + \lambda(\psi - \varphi)) - \lambda m - V(\varphi).$$

for any $\lambda \in \mathbb{R}$ with $\varphi + \lambda(\psi - \varphi) \in C(\Delta(\Omega))$. We show that H is differentiable at $\lambda = 0, 1$. Note that for any $\lambda \in \mathbb{R}$,

$$\begin{aligned} H'(\lambda) &= \lim_{t\to 0} \frac{V(\varphi + (\lambda + t)(\psi - \varphi)) - V(\varphi + \lambda(\psi - \varphi)}{t} - m \\ &= \lim_{t\to 0} \frac{V(\varphi + \lambda(\psi - \varphi) + t(\psi - \varphi)) - V(\varphi + \lambda(\psi - \varphi))}{t} - m. \end{aligned}$$

Since $\varphi, \psi \in D$, H is differentiable at $\lambda = 0, 1$. By (18), we have that $H(0) = H(1) = 0$, $H'(0) = \langle \psi - \varphi, \nabla V(\varphi)\rangle - m < 0$, and $H'(1) = \langle \psi - \varphi, \nabla V(\psi)\rangle - m < 0$. Hence, H is negative for small enough $\lambda > 0$ and positive for $\lambda < 1$ close enough to 1. Since H is locally Lipschitz continuous, the intermediate value theorem implies that the set $\{\lambda \in (0, 1) \,|\, H(\lambda) = 0\}$ is nonempty and closed. Let λ^* be its supremum.

Since H is locally Lipschitz continuous, we obtain $H(\lambda) = \int_{\lambda^*}^{\lambda} H'(\lambda')d\lambda'$ for all $\lambda > \lambda^*$. Since $H(\lambda) > 0$ for all $\lambda \in (\lambda^*, 1)$, we can choose $\lambda^{**} \in (\lambda^*, 1)$ close enough to λ^* such that H is differentiable at λ^{**} (by Rademacher's theorem) with $H'(\lambda^{**}) > 0$ and $H(\lambda^{**}) \in (0, \varepsilon)$. Note that

$$V^\circ(\varphi + \lambda^{**}(\psi - \varphi); \psi - \varphi) - m = \limsup_{\substack{\varphi \to \varphi + \lambda^{**}(\psi - \varphi) \\ t \searrow 0}} \frac{V(\varphi + t(\psi - \varphi)) - V(\varphi)}{t} - m \geq H'(\lambda^{**}) > 0.$$

By Lemma 5, $V^\circ(\eta; \xi) = \max_{\pi \in \partial V(\eta)} \langle \xi, \pi\rangle$ for any $\eta, \xi \in C(\Delta(\Omega))$. Hence, there exists $\pi \in \partial V(\varphi + \lambda^{**}(\psi - \varphi))$ such that

$$\langle \psi - \varphi, \pi\rangle - m \geq H'(\lambda^{**}) > 0. \tag{19}$$

By (17), there exists a sequence $\xi_n \to \varphi + \lambda^{**}(\psi - \varphi)$ such that $\xi_n \in D$ for any n and $\lim_{n\to\infty} \langle \phi, \nabla V(\xi_n)\rangle \in (\langle \phi, \pi\rangle - \varepsilon', \langle \phi, \pi\rangle + \varepsilon')$ for any $\phi \in C(\Delta(\Omega))$ and $\varepsilon' > 0$. By taking $\varepsilon' > 0$ small enough, we have that

$$\begin{aligned} \lim_{n\to\infty} (V(\xi_n) - V(\psi) + \langle \psi - \xi_n, \nabla V(\xi_n)\rangle) &> V(\varphi + \lambda^{**}(\psi - \varphi)) - V(\psi) + (1 - \lambda^{**})\langle \psi - \varphi, \pi\rangle - \varepsilon' \\ &= H(\lambda^{**}) + (1 - \lambda^{**})(V(\varphi) - V(\psi)) + (1 - \lambda^{**})\langle \psi - \varphi, \pi\rangle - \varepsilon' \\ &= H(\lambda^{**}) - (1 - \lambda^{**})m + (1 - \lambda^{**})\langle \psi - \varphi, \pi\rangle - \varepsilon' > 0, \end{aligned}$$

where the last inequality follows from $H(\lambda^{**}) > 0$ and $\langle \psi - \varphi, \pi \rangle - m > 0$ by (19). Similarly, by taking $\varepsilon' > 0$ small enough, we have that

$$\begin{aligned}\lim_{n\to\infty}(V(\xi_n) - V(\varphi) + \langle \varphi - \xi_n, \nabla V(\xi_n)\rangle) &< V(\varphi + \lambda^{**}(\psi - \varphi)) - V(\varphi) - \lambda^{**}\langle \psi - \varphi, \pi\rangle + \varepsilon' \\ &= H(\lambda^{**}) + \lambda^{**}m + V(\varphi) - V(\varphi) - \lambda^{**}\langle \psi - \varphi, \pi\rangle + \varepsilon' < \varepsilon,\end{aligned}$$

where the last inequality follows from $H(\lambda^{**}) < \varepsilon$ and $\langle \varphi - \psi, \pi\rangle + m < 0$ by (19). Thus, we can take $\xi = \xi_n \in D$ for large enough n. ■

By using Lemma 7, we can show that V has a Boolean representation in the same way as Lemma A.6 in Chandrasekher, Frick, Iijima, and Le Yaouanq [6]:

Lemma 8 *For any $\varphi \in D$, we have*

$$V(\varphi) = \max_{\psi \in D} \inf_{\xi \in K_\psi} V(\xi) + \langle \varphi - \xi, \nabla V(\xi)\rangle,$$

where $K_\psi \equiv \{\xi \in D \mid V(\xi) + \langle \psi - \xi, \nabla V(\xi)\rangle \geq V(\psi)\}$ for all $\psi \in D$.

B General CIC Representations

We prove the following theorem as an intermediate step for the CIC representation and special cases.

Theorem 5 *$\succsim$ satisfies basic axioms, Independence with Lotteries, and IEPR if and only if it is represented by a General CIC representation (u, Π).*

B.1 Sufficiency

First, we derive a utility representation $U : \mathbb{F} \to \mathbb{R}$ and define the functional $V : \Phi_{\mathbb{F}} \to \mathbb{R}$. Under Order, Preference for Flexibility, and Dominance, de Oliveira, Denti, Mihm, and Ozbek [12] proves the following:

Claim 1 (Claim 1 in de Oliveira, Denti, Mihm, and Ozbek [12]) *Consider menus F and G. Suppose that for each $g \in G$, there exists $f \in F$ such that $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$. Then, $F \succsim G$.*

By using Claim 1, Order, and Mixture Continuity, de Oliveira, Denti, Mihm, and Ozbek [12] proves the following property:

Claim 2 (Claim 2 in de Oliveira, Denti, Mihm, and Ozbek [12]) *Every menu F has a certainty equivalent $x_F \in X$ such that $\{x_F\} \sim F$.*

A preference $\succsim$ over $\mathcal{F}$ has a dual-self expected utility representation by Chandrasekher, Frick, Iijima, and Le Yaouanq [6].

Claim 3 *There exists an expected utility function $u : X \to \mathbb{R}$ with unbounded range and a nonempty compact collection $\mathbb{P} \subset \mathcal{K}(\Delta(\Omega))$ such that preference $\succsim$ over $\mathcal{F}$ is represented by function $U : \mathcal{F} \to \mathbb{R}$ defined by*

$$U(\{f\}) = \max_{P\in\mathbb{P}} \min_{p\in P} \sum_{\omega\in\Omega} u(f(\omega))p(\omega).$$

Proof. We show Monotonicity in Chandrasekher, Frick, Iijima, and Le Yaouanq [6], that is, $\{f(\omega)\} \succsim \{g(\omega)\}$ for any $\omega \in \Omega \implies \{f\} \succsim \{g\}$. Take two acts f and g such that $\{f(\omega)\} \succsim \{g(\omega)\}$ for any $\omega \in \Omega$. Dominance and Preference for Flexibility imply that $\{f\} \sim \{f, g\} \succsim \{g\}$. Moreover, Independence with lotteries implies Certainty Independence in Chandrasekher, Frick, Iijima, and Le Yaouanq [6]: for all $f, g \in \mathcal{F}$, $x \in X$, and $\alpha \in (0, 1)$, $\{f\} \succsim \{g\} \Longleftrightarrow \alpha\{f\} + (1-\alpha)\{x\} \succsim \alpha\{g\} + (1-\alpha)\{x\}$.

Under Order, Mixture Continuity, Two-Sided Unboundedness, Preference for Flexibility, Dominance, and Independence with lotteries, all the axioms except Archimedean in Chandrasekher, Frick, Iijima, and Le Yaouanq [6] are satisfied. By Claim 2, for each $f \in \mathcal{F}$, there exists $x_{\{f\}}$ such that $\{f\} \sim \{x_{\{f\}}\}$. Hence, we are able to replace Archimedean in Chandrasekher, Frick, Iijima, and Le Yaouanq [6] with Mixture Continuity. This implies that $\succsim$ on $\mathcal{F}$ has a dual-self expected utility representation by Theorem 1 in Chandrasekher, Frick, Iijima, and Le Yaouanq [6]. ■

We extend $U : \mathcal{F} \to \mathbb{R}$ to $\mathbb{F}$ by $U(F) = U(\{x_F\})$. By Claim 2, $U : \mathbb{F} \to \mathbb{R}$ represents $\succsim$. By Two-Sided Unboundedness, $U(\mathbb{F}) = \mathbb{R}$.

Define the functional $V : \Phi_{\mathbb{F}} \to \mathbb{R}$ by $V(\varphi_F) = U(F)$ as in de Oliveira, Denti, Mihm, and Ozbek [12]. To guarantee that V is well-defined, we prove that $\varphi_F = \varphi_G$ implies that $F \sim G$ for all menus F and G.

Let co$F$ denote the convex hull of menu F.

Claim 4 (Claim 4 in de Oliveira, Denti, Mihm, and Ozbek [12]) *Consider menus F and G. If $\varphi_F \geq \varphi_G$, for each $g \in G$, there exists $f \in coF$ such that $\{f(\omega)\} \succsim \{g(\omega)\}$.*

The following lemma is the same as Claim 5 in de Oliveira, Denti, Mihm, and Ozbek [12], but we use IEPR to prove it.

Lemma 9 *Consider menus F and G. If $G \subset coF$, $F \succsim G$.*

Proof. Let $G = \{g_1, ..., g_n\} \subset$ co$F$. For all $i = 1, ..., n$, we can write each $g_i = \sum_{j=1}^{m_i} \alpha^i_j f^i_j$ for $\alpha^i_1, ...\alpha^i_{m_i} \geq 0$ summing up to 1, and $f^i_1, ..., f^i_{m_i} \in F$. Hence,

$$G \subset \sum_{j=1}^{m_1} \cdots \sum_{j=1}^{m_n} \alpha^1_j \cdots \alpha^n_j F = \sum_{k=1}^{m} \beta_k F.$$

By Preference for Flexibility and IEPR, we obtain that $F \sim \sum_{k=1}^{m} \beta_k F \succsim G$. ■

By Claim 4, if $\varphi_F \geq \varphi_G$, there exists $H \subset \text{co}F$ such that for each $g \in G$, there exists $h \in H$ such that $\{h(\omega)\} \succsim \{g(\omega)\}$ for all $\omega \in \Omega$. By Lemma 9, we have that $F \succsim H$. Claim 1 implies that $H \succsim G$. Hence, if $\varphi_F \geq \varphi_G$, $F \succsim G$. This means that V is monotone and implies that V is well-defined.

Lemma 10 $V : \Phi_F \to \mathbb{R}$ *is normalized, monotone, translation invariant, and positively homogeneous.*

Proof. We have already shown that V is monotone. Claim 6 in de Oliveira, Denti, Mihm, and Ozbek [12] proves that V is normalized.

First, we show that V is translation invariant. Independence with lotteries implies that for any menus F, G, lotteries x, x', and $\alpha \in (0, 1)$,

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

Hence, Independence of Degenerated Decisions in de Oliveira, Denti, Mihm, and Ozbek [12] holds with lotteries x, x' instead of acts. Since a constant utility vector is identified with $\varphi_{\{x\}}$ for some $x \in X$, under Two-Sided Unboundedness, we can show that V is translation invariant in a similar way to Claim 6 in de Oliveira, Denti, Mihm, and Ozbek [12].

Next, we show that V is positively homogeneous. The proof is the same as in Corollary 1 in de Oliveira, Denti, Mihm, and Ozbek [12] except we use a lottery x instead of an act. Fix menu F arbitrary. First, we assume that $\alpha \in [0, 1]$ and take $x \in X$ such that $\{x\} \sim F$. Since Independence with lotteries implies that $F \sim \{x\} \iff F \sim \alpha F + (1-\alpha)\{x\}$ for all menus F, lotteries x, and $\alpha \in (0, 1)$, we have that $F \sim \alpha F + (1-\alpha)\{x\}$ for all $\alpha \in [0, 1]$. By translation invariance of V, we have that

$$V(\varphi_F) = V(\alpha\varphi_F + (1-\alpha)\varphi_{\{x\}}) = V(\alpha\varphi_F) + (1-\alpha)V(\varphi_{\{x\}}).$$

Since $V(\varphi_F) = V(\varphi_{\{x\}})$, We obtain $V(\alpha\varphi_F) = \alpha V(\varphi_F)$.

Second, we assume that $\alpha > 1$. Since V is positively homogeneous under $\alpha \in [0, 1]$, we have that

$$V(\alpha^{-1}(\alpha\varphi_F)) = \alpha^{-1}V(\alpha\varphi_F).$$

This implies that $V(\alpha\varphi_F) = \alpha V(\varphi_F)$. Hence, V is positively homogeneous. ■

Lemma 11 $V : C(\Delta(\Omega)) \to \mathbb{R}$ *is normalized, monotone, translation invariant, and positively homogeneous. Moreover, V is a niveloid, which implies that V is Lipschitz continuous.*

Proof. We extend V to $C(\Delta(\Omega))$. Since V is monotone and translation invariant on a tube $\Phi_{\mathbb{F}}$, it is a niveloid. This implies that there is an extension V on $C(\Delta(\Omega))$ that is a niveloid (hence, monotone and translation invariant), as mentioned in Appendix A.1. Moreover, V is normalized. Since there are several ways to extend V on $\Phi_{\mathbb{F}}$ to $C(\Delta(\Omega))$, we use the maximal one proved in Theorem 1 in Cerreia-Vioglio, Maccheroni, Marinacci, Rustichini [5]:

$$V(\varphi) = \inf\{V(\varphi') + k \mid \varphi' \in \Phi_{\mathbb{R}}, k \in \mathbb{R}, \varphi' + k\mathbf{1} \geq \varphi\}.$$

We show that V on $C(\Delta(\Omega))$ defined above is positively homogeneous. First of all, assume that $\alpha = 0$. Since V is normalized, we have that $V(\alpha\varphi) = V(0) = 0 = \alpha V(\varphi)$, as desired. For every $\varphi \in C(\Delta(\Omega))$ and $\alpha > 0$, note that

$$\begin{aligned}&\{(\varphi', k) \in \Phi_{\mathbb{F}} \times \mathbb{R} \mid \varphi' + k\mathbf{1} \geq \varphi\}\\ =&\{(\alpha\varphi', \alpha k) \in \Phi_{\mathbb{F}} \times \mathbb{R} \mid \alpha\varphi' + \alpha k\mathbf{1} \geq \varphi \text{ for some } (\varphi', k) \in \Phi_{\mathbb{F}} \times \mathbb{R}\}.\end{aligned}$$

Take any (φ', k) from the left-hand side. Since $\Phi_{\mathbb{F}}$ is a cone, $\frac{\varphi'}{\alpha} \in \Phi_{\mathbb{F}}$. Thus, $\alpha(\frac{\varphi'}{\alpha}) + \alpha(\frac{k}{\alpha})\mathbf{1} = \varphi' + k\mathbf{1} \geq \varphi$. By definition, $(\varphi', k) = (\alpha(\frac{\varphi'}{\alpha}), \alpha(\frac{k}{\alpha}))$ belongs to the right-hand side. Conversely, take any $(\alpha\varphi', \alpha k)$ from the right-hand side. Since $(\alpha\varphi', \alpha k) \in \Phi_{\mathbb{F}} \times \mathbb{R}$ and $\alpha\varphi' + \alpha k\mathbf{1} \geq \varphi$, by definition, $(\alpha\varphi', \alpha k)$ belongs to the left-hand side.

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

$$\begin{aligned}V(\alpha\varphi) &= \inf\{V(\varphi') + k \mid \varphi' \in \Phi_{\mathbb{R}}, k \in \mathbb{R}, \varphi' + k\mathbf{1} \geq \alpha\varphi\}\\ &= \inf\{V(\alpha\varphi') + \alpha k \mid \varphi' \in \Phi_{\mathbb{R}}, k \in \mathbb{R}, \alpha\varphi' + \alpha k\mathbf{1} \geq \alpha\varphi\}\\ &= \inf\{\alpha V(\varphi') + \alpha k \mid \varphi' \in \Phi_{\mathbb{R}}, k \in \mathbb{R}, \varphi' + k\mathbf{1} \geq \varphi\} = \alpha V(\varphi),\end{aligned}$$

as desired. ■

Following Theorem 1 in Chandrasekher, Frick, Iijima, and Le Yaouanq [6], we derive a Boolean representation of $\succsim$. Since V is a niveloid (equivalently, monotone and translation invariant), it is Lipschitz. Hence, Lemma 3 implies that $\partial V(\mathbf{0}) \subset \Delta(\Delta(\Omega))$. Let $\mathcal{K}(\Delta(\Delta(\Omega)))$ be the set of nonempty, closed, and convex subsets of $\Delta(\Delta(\Omega))$. Consider the collection $\mathsf{\Pi}$ by

$$\mathsf{\Pi} \equiv \mathrm{cl}\{\Pi_\varphi \mid \varphi \in C(\Delta(\Omega))\}, \tag{20}$$

where

$$\Pi_\varphi \equiv \{\pi \in \partial V(\mathbf{0}) \mid \langle \varphi, \pi \rangle \geq V(\varphi)\}, \tag{21}$$

and cl denotes the topological closure in $\mathcal{K}(\Delta(\Delta(\Omega)))$ under Hausdorff topology. Note that each Π_φ is a closed convex set of beliefs over beliefs. Since $\mathsf{\Pi}$ is a closed subset of the compact space $\mathcal{K}(\Delta(\Delta(\Omega)))$, $\mathsf{\Pi}$ is compact.

As an intermediate step, we show that for all $\varphi \in C(\Delta(\Omega))$,

$$V(\varphi) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} \langle \varphi, \pi \rangle. \tag{22}$$

By Lemma 6, there exists $D \subset C(\Delta(\Omega))$ such that $C(\Delta(\Omega)) \setminus D$ is Haar-null and V is Gatèaux differentiable on D. Since V is positively homogeneous, Lemma 2 implies that $\partial V(\varphi) \subset \partial V(\mathbf{0})$ for any $\varphi \in C(\Delta(\Omega))$. Hence, for any $\varphi \in D$, we have that $\pi_\varphi \equiv \partial V(\varphi) \in \partial V(\mathbf{0})$.

Lemma 12 *For any $\varphi \in D$, $V(\varphi) = \langle \varphi, \pi_\varphi \rangle$.*

Proof. Except notations and minor details, the proof is the same as that of Lemma B.2 in Chandrasekher, Frick, Iijima, and Le Yaouanq [6].

Fix any $\varphi \in D$. Since V is positively homogeneous, we have that $\alpha\varphi \in D$ and $\partial V(\varphi) = \partial V(\alpha\varphi)$ for any $\alpha \in (0, 1)$. Hence, the function $h : [0, 1] \to \mathbb{R}$ defined by $h(\alpha) = V(\alpha\varphi)$ is differentiable at any $\alpha \in (0, 1)$ and is Lipschitz continuous. This implies that $V(\varphi) = h(1) - h(0) = \int_0^1 h'(\alpha')d\alpha' = \int_0^1 \langle \varphi, \partial V(\alpha'\varphi)\rangle d\alpha' = \int_0^1 \langle \varphi, \partial V(\varphi)\rangle d\alpha' = \langle \varphi, \pi_\varphi \rangle$. ■

To show (22), first we pick any $\varphi, \psi \in D$ and as in Lemma 8, let $K_\psi \equiv \{\xi \in D \,|\, V(\xi) + \langle \psi - \xi, \pi_\xi \rangle \geq V(\psi)\}$. By Lemma 8 and 12, we have that

$$V(\varphi) = \max_{\psi \in D} \inf_{\xi \in K_\psi} V(\xi) + \langle \varphi - \xi, \pi_\xi \rangle = \max_{\psi \in D} \inf_{\xi \in K_\psi} \langle \varphi, \pi_\xi \rangle. \tag{23}$$

Lemma 12 implies that $\xi \in K_\psi$ if and only if $\pi_\xi \in \Pi^*_\psi \equiv \{\pi_\xi \,|\, \xi \in D, \langle \psi, \pi_\xi \rangle \geq V(\psi)\}$. By (17), we obtain that $\overline{co}\Pi^*_\psi = \Pi_\psi$, where $\overline{co}$ denotes weak* closure of convex hull. Hence, (23) implies that

$$V(\varphi) = \max_{\psi \in D} \inf_{\pi \in \Pi^*_\psi} \langle \varphi, \pi \rangle = \max_{\psi \in D} \min_{\pi \in \overline{co}\Pi^*_\psi} \langle \varphi, \pi \rangle = \max_{\psi \in D} \min_{\pi \in \Pi_\psi} \langle \varphi, \pi \rangle. \tag{24}$$

Next, fix any $\varphi, \psi \in C(\Delta(\Omega))$. There exists sequences $\varphi_n \to \varphi, \psi_n \to \psi$ such that $\varphi_n, \psi_n \in D$. For any n, take $\pi_n \in \Pi_{\psi_n}$ such that $\min_{\pi \in \Pi_{\psi_n}} \langle \varphi_n, \pi \rangle = \langle \varphi_n, \pi_n \rangle$ and consider a convergent subsequence π_{nk} with $\lim_{k\to\infty} \pi_{nk} = \pi^*$. Note that $\pi^* \in \Pi_\psi$. For any k, we have that $\langle \psi_{nk}, \pi_{nk} \rangle \geq V(\psi_{nk})$. By continuity of V, we obtain $\langle \psi, \pi^* \rangle \geq V(\psi)$.

By (24), for any k, we have that $\langle \varphi_{nk}, \pi_{nk} \rangle = \min_{\pi \in \Pi_{\psi_{nk}}} \langle \varphi_{nk}, \pi \rangle \leq V(\varphi_{nk})$. Hence, continuity of V implies that $\langle \varphi, \pi^* \rangle \leq V(\varphi)$. This implies that

$$V(\varphi) \geq \langle \varphi, \pi^* \rangle \geq \min_{\pi \in \Pi_\psi} \langle \varphi, \pi \rangle. \tag{25}$$

By applying (25) with $\psi = \varphi$, we have that $\min_{\pi \in \Pi_\varphi} \langle \varphi, \pi \rangle \geq V(\varphi) \geq \min_{\pi \in \Pi_\varphi} \langle \varphi, \pi \rangle$, where the first inequality follows from the definition of Π_φ. Hence, we have that for any $\varphi \in C(\Delta(\Omega))$,

$$V(\varphi) = \min_{\pi \in \Pi_\varphi} \langle \varphi, \pi \rangle.$$

Now, we derive a representation (22). Since (25) holds for any $\psi \in C(\Delta(\Omega))$, the definition of $\mathsf{\Pi}$ implies that for any $\Pi \in \mathsf{\Pi}$,

$$V(\varphi) \geq \min_{\pi \in \Pi} \langle \varphi, \pi \rangle$$

Take any $\psi \in C(\Delta(\Omega))$ such that $\psi \neq \varphi$. Fix any $\pi' \in \Pi_\psi$. If $\langle \varphi, \pi' \rangle \geq V(\varphi)$, we have that $\pi' \in \Pi_\varphi$. Thus, if $\pi' \notin \Pi_\varphi$, $\langle \varphi, \pi' \rangle < V(\varphi)$. Hence,

$$V(\varphi) = \min_{\pi \in \Pi_\varphi} \langle \varphi, \pi \rangle = \max_{\Pi_{\varphi'} \in \mathsf{\Pi}} \min_{\pi \in \Pi_{\varphi'}} \langle \varphi, \pi \rangle. \tag{26}$$

This implies that for any menu F,

$$U(F) = V(\varphi_F) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} \langle \varphi, \pi \rangle = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_F(\pi).$$

We have that for any act f,

$$U(\{f\}) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} \sum_{\Omega} u(f(\omega)) p^\pi(\omega) = \max_{P \in \mathbb{P}} \min_{p \in P} \sum_{\Omega} u(f(\omega)) p(\omega).$$

Hence, we can take $\mathbb{P} = \mathrm{cl}\{P^\Pi \,|\, \Pi \in \mathsf{\Pi}\}$ with $P^\Pi = \{p^\pi \in \Delta(\Omega) \,|\, \pi \in \Pi\}$.

B.2 Necessity

Assume that the preference is represented by the General CIC representation $(u, \mathsf{\Pi})$. We show that preference satisfies Independence with Lotteries. Take any menu F and any lottery x. For any $\alpha \in [0,1]$, it is enough to show that the representation U is mixture linear for $\alpha F + (1-\alpha)\{x\}$. Since $b^u_F(\pi)$ is mixture linear in F,

$$\begin{aligned}
U(\alpha F + (1-\alpha)\{x\}) &= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_{\alpha F + (1-\alpha)\{x\}}(\pi) \\
&= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} (\alpha b^u_F(\pi) + (1-\alpha) b^u_{\{x\}}(\pi)) \\
&= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} (\alpha b^u_F(\pi) + (1-\alpha) u(x)) \\
&= \alpha \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_F(\pi) + (1-\alpha) u(x) \\
&= \alpha U(F) + (1-\alpha) u(p),
\end{aligned}$$

as desired.

Next, we show that preference satisfies IEPR. Take any menu F and $\beta_i \geq 0$ with $\sum_i \beta_i = 1$. Since by linearity

$$b^u_{\sum_i \beta_i F}(\pi) = \sum_i \beta_i b^u_F(\pi) = b^u_F(\pi)$$

for any π,

$$U(F) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_F(\pi) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_{\sum_i \beta_i F}(\pi) = U(\sum_i \beta_i F),$$

as desired.

C Proof of Theorem 1

First we show sufficiency. By Theorem 5, $\succsim$ is represented by the General CIC representation $(u, \mathsf{\Pi})$. Since $\succsim$ satisfies Weak Concavity, it satisfies Singleton Concavity. Hence, $\succsim$

define on $\mathcal{F}$ is represented by a maxmin expected utility model by Gilboa and Schmeidler [19]:

$$V(\varphi_{\{f\}}) = U(\{f\}) = \min_{p \in \overline{P}} \sum_{\omega \in \Omega} u(f(\omega))p(\omega), \tag{27}$$

where $\overline{P} \subset \Delta(\Omega)$ is closed and convex and $u(X) = \mathbb{R}$ by Two-Sided Unboundedness. For all $f \in \mathcal{F}$, we have

$$\max_{\Pi \in \mathsf{\Pi}} \min_{p \in P^{\Pi}} u(f) \cdot p = U(\{f\}) = \min_{p \in \overline{P}} u(f) \cdot p.$$

We show that $\mathsf{\Pi}$ is prior-consistent with $\overline{P}$ under Weak Concavity. The next implication is obtained solely by Singleton Concavity.

Lemma 13 *For all* $\Pi \in \mathsf{\Pi}$, $\overline{P} \subset P^{\Pi}$.

Proof. Seeking a contradiction, suppose that there exists some $\Pi \in \mathsf{\Pi}$ with $\overline{P} \not\subset P^{\Pi}$. There exists $\overline{p} \in \overline{P}$ and $\overline{p} \notin P^{\Pi}$. By the separation hyperplane theorem, there exists $v \in \mathbb{R}^{\Omega} \setminus \{0\}$ such that $v \cdot \overline{p} < 0 \leq v \cdot p$ for all $p \in P^{\Pi}$. Let $f \in \mathcal{F}$ satisfy $v = u(f)$. The above inequalities imply

$$\min_{p \in \overline{P}} u(f) \cdot p \leq u(f) \cdot \overline{p} < 0 \leq \min_{p \in P^{\Pi}} u(f) \cdot p,$$

which in turn implies

$$\min_{p \in \overline{P}} u(f) \cdot p < 0 \leq \max_{\Pi \in \mathsf{\Pi}} \min_{p \in P^{\Pi}} u(f) \cdot p.$$

This contradicts to the fact that the restriction of U on the singletons coincides with the maxmin EU representation. ■

Next, we show the converse direction.

Lemma 14 *For any* $\varphi \in C(\Delta(\Omega))$ *and* $\varphi_{\{g\}} \in \Phi_{\mathcal{F}}$

$$V(\varphi + \varphi_{\{g\}}) \geq V(\varphi) + V(\varphi_{\{g\}}).$$

Proof. For any $\varphi_F \in \Phi_{\mathbb{F}}$ and $\varphi_{\{g\}} \in \Phi_{\mathcal{F}}$ such that $V(\varphi_F) = V(\varphi_{\{g\}})$, Weak Concavity implies that

$$V(\frac{1}{2}\varphi_F + \frac{1}{2}\varphi_{\{g\}}) = V(\varphi_{\frac{1}{2}F + \frac{1}{2}\{g\}}) \geq V(\varphi_F) = \frac{1}{2}V(\varphi_F) + \frac{1}{2}V(\varphi_{\{g\}}).$$

Since V is positively homogeneous, we have that

$$V(\varphi_F + \varphi_{\{g\}}) = 2V(\frac{1}{2}\varphi_F + \frac{1}{2}\varphi_{\{g\}}) \geq V(\varphi_F) + V(\varphi_{\{g\}}).$$

Next, consider the case $V(\varphi_F) \neq V(\varphi_{\{g\}})$. We may assume that $V(\varphi_F) > V(\varphi_{\{g\}})$. Let $k \equiv V(\varphi_F) - V(\varphi_{\{g\}}) > 0$. since $\varphi_{\{g\}} + k\mathbf{1} \in \Phi_{\mathbb{F}}$.

$$V(\varphi_{\{g\}} + k\mathbf{1}) = V(\varphi_{\{g\}}) + k = V(\varphi_F),$$

where the first equality follows from translation invariance. Hence, we have that

$$\begin{aligned}V(\varphi_F) &\leq V(\frac{1}{2}\varphi_F + \frac{1}{2}(\varphi_{\{g\}} + k\mathbf{1}) = V(\frac{1}{2}\varphi_F + \frac{1}{2}\varphi_{\{g\}}) + \frac{1}{2}k \\ &= V(\frac{1}{2}\varphi_F + \frac{1}{2}\varphi_{\{g\}}) + \frac{1}{2}(V(\varphi_F) - V(\varphi_{\{g\}})).\end{aligned}$$

Hence, we have that $V(\frac{1}{2}\varphi_F + \frac{1}{2}\varphi_{\{g\}}) \geq \frac{1}{2}V(\varphi_F) + \frac{1}{2}V(\varphi_{\{g\}})$. This implies that $V(\varphi_F + \varphi_{\{g\}}) \geq V(\varphi_F) + V(\varphi_{\{g\}})$ for any $\varphi_F \in \Phi_{\mathbb{F}}$ and $\varphi_{\{g\}} \in \Phi_{\mathcal{F}}$.

Consider an extension of V on $\Phi_{\mathbb{F}}$ to $\Phi_{\mathbb{F}} + \mathbb{R}$ in Lemma 11. Since $u(X) = \mathbb{R}$, we have that $\Phi_{\mathbb{F}} + \mathbb{R} = \Phi_{\mathbb{F}}$. Hence, a functional $V : C(\Delta(\Omega)) \to \mathbb{R}$ defined in Lemma 11 can be written as $V(\varphi) = \inf_{\varphi_F \in D(\varphi)} V(\varphi_F)$, where $D(\varphi) \equiv \{\varphi_F \in \Phi_{\mathbb{F}} \,|\, \varphi_F \geq \varphi\}$. We show that

$$\inf_{\varphi_F \in D(\varphi + \varphi_{\{g\}})} V(\varphi_F) \geq \inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{g\}}). \tag{28}$$

We show that for any $\varphi_F \geq \varphi + \varphi_{\{g\}}$, there exists $\varphi_G \in \Phi_{\mathbb{F}}$ such that $\varphi_G \in D(\varphi)$ and $\varphi_F \geq \varphi_G + \varphi_{\{g\}}$. Since $u(X) = \mathbb{R}$, for any $\varphi_{\{g\}} \in \Phi_{\mathbb{F}}$, there exists $\varphi_{\{h\}} \in \Phi_{\mathbb{R}}$ such that $\varphi_{\{h\}} = -\varphi_{\{g\}}$. Since $\Phi_{\mathbb{F}}$ is convex and $\varphi_F \in \Phi_{\mathbb{F}}$ implies that $\alpha\varphi_F \in \Phi_{\mathbb{F}}$ for $\alpha \in \mathbb{R}$, this implies that $\varphi_F - \varphi_{\{g\}} \in \Phi_{\mathbb{F}}$. By taking $\varphi_G = \varphi_F + \varphi_{\{h\}} = \varphi_F - \varphi_{\{g\}}$, we obtain φ_G with the desired properties.

Take any $\varepsilon > 0$. By the definition of infimum, there exists $\varphi_F^\varepsilon \in D(\varphi + \varphi_{\{g\}})$ such that $V(\varphi_F^\varepsilon) < \inf_{\varphi_F \in D(\varphi + \varphi_{\{g\}})} V(\varphi_F) + \varepsilon$. For such φ_F^ε, there exists $\varphi_G \in D(\varphi)$ such that $\varphi_F^\varepsilon \geq \varphi_G + \varphi_{\{g\}}$. By monotonicity of V, we have that $V(\varphi_G + \varphi_{\{g\}}) \leq V(\varphi_F^\varepsilon) < \inf_{\varphi_F \in D(\varphi + \varphi_{\{g\}})} V(\varphi_F) + \varepsilon$. By the definition of infimum,

$$\inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{g\}}) < \inf_{\varphi_F \in D(\varphi + \varphi_{\{g\}})} V(\varphi_F) + \varepsilon.$$

Therefore, we have that $\inf_{\varphi_F \in D(\varphi + \varphi_{\{g\}})} V(\varphi_F) \geq \inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{g\}})$ as $\varepsilon \to 0$. Hence, (28) holds.

Finally, we show that $V(\varphi + \varphi_{\{g\}}) \geq V(\varphi) + V(\varphi_{\{g\}})$ for any $\varphi \in C(\Delta(\Omega))$ and $\varphi_{\{g\}} \in \Phi_{\mathcal{F}}$. By (28),

$$\begin{aligned}V(\varphi + \varphi_{\{g\}}) = \inf_{\varphi_F \in D(\varphi + \varphi_{\{g\}})} V(\varphi_F) &\geq \inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{g\}}) \\ &\geq \inf_{\varphi_G \in D(\varphi)} V(\varphi_G) + V(\varphi_{\{g\}}) = V(\varphi) + V(\varphi_{\{g\}}).\end{aligned}$$

■

Lemma 15 *For any $\pi \in \partial V(\mathbf{0})$, $p^\pi \in \overline{P}$.*

Proof. By Lemma 5, we have that for any $\pi \in \partial V(\mathbf{0})$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,

$$V_\circ(\mathbf{0}; \varphi_{\{f\}}) = \min_{\pi \in \partial V(\mathbf{0})} \langle \varphi_{\{f\}}, \pi \rangle \leq \langle \varphi_{\{f\}}, \pi \rangle = \langle \varphi_{\{f\}}, p^\pi \rangle.$$

By Lemma 4 and 14, we have that for any $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,

$$\begin{aligned} V_\circ(\mathbf{0}; \varphi_{\{f\}}) &= \inf_{\varphi \in C(\Delta(\Omega))} \{V(\varphi + \varphi_{\{f\}}) - V(\varphi)\} \geq \inf_{\varphi \in C(\Delta(\Omega))} \{V(\varphi) + V(\varphi_{\{f\}} - V(\varphi))\} \\ &= V(\varphi_{\{f\}}) = \min_{p \in \overline{P}} \langle \varphi_{\{f\}}, p \rangle, \end{aligned}$$

where the last equality follows from (27). Hence, for any $\pi \in \partial V(\mathbf{0})$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,

$$\min_{p \in \overline{P}} \langle \varphi_{\{f\}}, p \rangle \leq \langle \varphi_{\{f\}}, \pi \rangle \Longleftrightarrow \min_{p \in \overline{P}} \sum_{\omega \in \Omega} u(f(\omega)) p(\omega) \leq \sum_{\omega \in \Omega} u(f(\omega)) p^\pi(\omega).$$

Hence, $p^\pi \in \overline{P}$ for any $\pi \in \partial V(\mathbf{0})$. Otherwise, a separating hyperplane theorem yields a contradiction. ■

Given Lemma 2 and equations (20) and (21), Lemma 15 implies $P^\Pi \subset \overline{P}$ for all $\Pi \in \mathsf{\Pi}$. Together with Lemma 13, we conclude that $\mathsf{\Pi}$ is prior consistent with $\overline{P}$.

The necessity of the axiom is proved as follows: First of all, by the prior consistency with $\overline{P}$, $P^\Pi = \overline{P}$ for all $\Pi \in \mathsf{\Pi}$. Thus,

$$U(\{f\}) = \max_{\Pi \in \mathsf{\Pi}} \min_{p \in P^\Pi} u(f) \cdot p = \min_{p \in \overline{P}} u(f) \cdot p.$$

Take any F and f with $F \sim \{f\}$. By the prior consistency with $\overline{P}$,

$$\begin{aligned} U(\alpha F + (1-\alpha)\{f\}) &= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_{\alpha F + (1-\alpha)\{f\}}(\pi) \\ &= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} (\alpha b^u_F(\pi) + (1-\alpha) u(f) \cdot p^\pi) \\ &\geq \max_{\Pi \in \mathsf{\Pi}} (\alpha \min_{\pi \in \Pi} b^u_F(\pi) + (1-\alpha) \min_{\pi \in \Pi} u(f) \cdot p^\pi) \\ &= \max_{\Pi \in \mathsf{\Pi}} (\alpha \min_{\pi \in \Pi} b^u_F(\pi) + (1-\alpha) \min_{p \in P^\Pi} u(f) \cdot p) \\ &= \max_{\Pi \in \mathsf{\Pi}} (\alpha \min_{\pi \in \Pi} b^u_F(\pi) + (1-\alpha) \min_{p \in \overline{P}} u(f) \cdot p) \\ &= \alpha \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_F(\pi) + (1-\alpha) \min_{p \in \overline{P}} u(f) \cdot p \\ &= \alpha U(F) + (1-\alpha) U(\{f\}). \end{aligned}$$

Now take any F and f with $F \sim \{f\}$. Since $U(\{F\}) = U(\{f\})$, the above claim implies

$$U(\alpha F + (1-\alpha)\{f\}) \geq \alpha U(F) + (1-\alpha) U(\{f\}) = U(F),$$

which ensures Weak Concavity, as desired.

D Proof of Theorem 2

We adapt the proof of uniqueness of the dual-self model in Chandrasekher, Frick, Iijima, and Le Yaouanq [6, Proposition 5] with appropriately taking into account the prior consistency. Let $(u, \overline{P}, \mathsf{\Pi})$ be a CIC representation for $\succsim$. We fix the unique functional $V : C(\Delta(\Omega)) \to \mathbb{R}$ representing $\succsim$ given in Lemma 11. By the construction of the representation in Theorem 1, $V(\varphi) = \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} \langle \varphi, \pi \rangle$.

Lemma 16 *The half-space closure $\overline{\mathsf{\Pi}}$ satisfies the prior consistency with $\overline{P}$.*

Proof. Take any (φ, λ) such that $\Pi \subset H_{\varphi,\lambda}$ for some $\Pi \in \mathsf{\Pi}$. Since Π satisfies the prior consistency with $\overline{P}$, $\Pi \subset \Pi(\overline{P})$. Hence, $\Pi \subset H_{\varphi,\lambda} \cap \Pi(\overline{P})$. We have $\overline{P} = P^{\Pi} \subset P^{H_{\varphi,\lambda} \cap \Pi(\overline{P})}$. Conversely, since $H_{\varphi,\lambda} \cap \Pi(\overline{P}) \subset \Pi(\overline{P})$, $P^{H_{\varphi,\lambda} \cap \Pi(\overline{P})} \subset P^{\Pi(\overline{P})} = \overline{P}$. Therefore, $P^{H_{\varphi,\lambda} \cap \Pi(\overline{P})} = \overline{P}$. Since any element of $\overline{\mathsf{\Pi}}$ is an accumulation point of the set of elements $H_{\varphi,\lambda} \cap \Pi(\overline{P})$, $\overline{\mathsf{\Pi}}$ satisfies the prior consistency. ■

Slight modification of the proof of Lemma D.1 in Chandrasekher, Frick, Iijima, and Le Yaouanq [6] yields the following:

Lemma 17 *Suppose that $(u, \overline{P}, \mathsf{\Pi})$ is a CIC representation of $\succsim$. Then, $\overline{\mathsf{\Pi}} = \mathrm{cl}\{H_{\varphi,\lambda} \cap \Pi(\overline{P}) \,|\, \varphi \in C(\Delta(\Omega)), \lambda \leq V(\varphi)\}$, where $H_{\varphi,\lambda} = \{\pi \in \Delta(\Delta(\Omega)) \,|\, \langle \varphi, \pi \rangle \geq \lambda\}$ for some $\lambda \in \mathbb{R}$ and $\varphi \in C(\Delta(\Omega))$.*

Proof. First, take any $\varphi \in C(\Delta(\Omega))$ and $\lambda \in \mathbb{R}$ such that $\lambda \leq V(\varphi)$. Since $(u, \overline{P}, \mathsf{\Pi})$ represents $\succsim$, there exists $\Pi \in \mathsf{\Pi}$ such that $\min_{\pi \in \Pi} \langle \varphi, \pi \rangle = V(\varphi)$. This implies $\Pi \subset H_{\varphi,V(\varphi)} \subset H_{\varphi,\lambda}$. Moreover, $\Pi \subset \Pi(\overline{P})$ by prior consistency of the representation. Thus, $H_{\varphi,\lambda} \cap \Pi(\overline{P}) \in \overline{\mathsf{\Pi}}$. Hence, we have that $\overline{\mathsf{\Pi}} \supset \mathrm{cl}\{H_{\varphi,\lambda} \cap \Pi(\overline{P}) \,|\, \varphi \in C(\Delta(\Omega)), \lambda \leq V(\varphi)\}$.

Next, consider any $\varphi \in C(\Delta(\Omega))$ and $\lambda \in \mathbb{R}$ such that there exists $\Pi' \in \mathsf{\Pi}$ with $\Pi' \subset H_{\varphi,\lambda}$. Since $(u, \overline{P}, \mathsf{\Pi})$ represents $\succsim$, $V(\varphi) \geq \min_{\pi \in \Pi'} \langle \varphi, \pi \rangle \geq \min_{\pi \in H_{\varphi,\lambda}} \langle \varphi, \pi \rangle$. This implies that $\lambda \leq V(\varphi)$. Hence, we have that $\overline{\mathsf{\Pi}} \subset \mathrm{cl}\{H_{\varphi,\lambda} \cap \Pi(\overline{P}) \,|\, \varphi \in C(\Delta(\Omega)), \lambda \leq V(\varphi)\}$.
■

Suppose that $(u', \overline{P}', \mathsf{\Pi}')$ is another CIC representation of $\succsim$. Since $\succsim$ on $\mathcal{F}$ is represented by a maxmin expected utility model, we have that u' is a positive affine transformation of u and $\overline{P} = \overline{P}'$. Since $\Pi(\overline{P}) = \Pi(\overline{P}')$, by Lemma 17 and the uniqueness of V, we have that $\overline{\mathsf{\Pi}} = \overline{\mathsf{\Pi}'}$.

Conversely, suppose that u' is a positive affine transformation of u, $\overline{P}' = \overline{P}$, and $\overline{\mathsf{\Pi}} = \overline{\mathsf{\Pi}'}$. We show that $(u', \overline{P}, \mathsf{\Pi}')$ represents $\succsim$. It suffices to show that $V(\varphi) = \max_{\Pi \in \mathsf{\Pi}'} \min_{\pi \in \Pi} \langle \varphi, \pi \rangle$ for any $\varphi \in C(\Delta(\Omega))$. Lemma 17 implies that $H_{\varphi,V(\varphi)} \cap \Pi(\overline{P}) \in \overline{\mathsf{\Pi}} = \overline{\mathsf{\Pi}'}$. Hence, there exists sequences of $\Pi'_n \in \mathsf{\Pi}'$ and half-spaces $H_n \supset \Pi'_n$ such that $H_n \to H_{\varphi,V(\varphi)}$. We have that for any $\varphi \in C(\Delta(\Omega))$, $\min_{\pi \in H_{\varphi,V(\varphi)}} \langle \varphi, \pi \rangle = V(\varphi) = \lim_n \min_{\pi \in H_n} \langle \varphi, \pi \rangle$ and $\min_{\pi \in H_n} \langle \varphi, \pi \rangle \leq \min_{\pi \in \Pi'_n} \langle \varphi, \pi \rangle$ for all n. Hence,

$$V(\varphi) \leq \max_{\Pi' \in \mathsf{\Pi}'} \min_{\pi \in \Pi'} \langle \varphi, \pi \rangle.$$

To show the above equation holds with equality, suppose that there exists $\Pi'' \in \sqcap'$ such that $\min_{\pi\in\Pi''}\langle\varphi,\pi\rangle - V(\varphi) \equiv \varepsilon > 0$. This implies that $H_{\varphi,V(\varphi)+\varepsilon} \supset \Pi''$. Hence, $H_{\varphi,V(\varphi)+\varepsilon} \cap \Pi(\overline{P}) \in \overline{\sqcap'}$ holds. Since $\overline{\sqcap'} = \overline{\sqcap}$, this is a contradiction to Lemma 17.

Note that the half-space closure of $\overline{\sqcap}$ is $\overline{\sqcap}$. The argument in the previous paragraph shows that $(u, \overline{P}, \overline{\sqcap})$ represents $\succsim$. Moreover, by Lemma 16, $\overline{\sqcap}$ is prior-consistent. Thus, $(u, \overline{P}, \overline{\sqcap})$ is a CIC representation of $\succsim$.

E Proof of Theorem 3

For each preference $\succsim^i$, take utility u^i, $\overline{P}^i$, and V^i as in Lemma 11 and Theorem 1. Since $\succsim^1$ and $\succsim^2$ coincide with each other on $\mathcal{F}$, u^2 is an affine transformation of u^1 and $\overline{P}^1 = \overline{P}^2 = \overline{P}$. Note that $\succsim^1$ is more averse to commitment than $\succsim^2$ if and only if u^2 is an affine transformation of u^1, $\overline{P}^1 = \overline{P}^2 = \overline{P}$, and $V^1(\varphi) \geq V^2(\varphi)$ for any $\varphi \in C(\Delta(\Omega))$. We only show that $V^1(\varphi) \geq V^2(\varphi)$ for any $\varphi \in C(\Delta(\Omega))$ if and only if $\overline{\sqcap}^1 \supset \overline{\sqcap}^2$.

Suppose that $V^1(\varphi) \geq V^2(\varphi)$ for any $\varphi \in C(\Delta(\Omega))$. This implies that $\{H_{\varphi,\lambda} \cap \Pi(\overline{P}) \mid \varphi \in C(\Delta(\Omega)), \lambda \leq V^1(\varphi)\} \supset \{H_{\varphi,\lambda} \cap \Pi(\overline{P}) \mid \varphi \in C(\Delta(\Omega)), \lambda \leq V^2(\varphi)\}$. By Lemma 17, we have that $\overline{\sqcap}^1 \supset \overline{\sqcap}^2$.

Conversely, if $\overline{\sqcap}^1 \supset \overline{\sqcap}^2$, we have that $\max_{\Pi\in\overline{\sqcap}^1} \min_{\pi\in\Pi}\langle\varphi,\pi\rangle \geq \max_{\Pi\in\overline{\sqcap}^2} \min_{\pi\in\Pi}\langle\varphi,\pi\rangle$. Since $(u^i, \overline{P}^i, \overline{\sqcap}^i)$ with $\overline{P}^2 = \overline{P}^1$ is a CIC representation of $\succsim^i$ for $i = 1, 2$, the inequality above means that $V^1(\varphi) \geq V^2(\varphi)$ for all $\varphi \in C(\Delta(\Omega))$.

F Proof of Theorem 4

First we show sufficiency. By Theorem 5, $\succsim$ is represented by the General CIC representation $(u, \sqcap)$. Since $\succsim$ satisfies Two-Sided Unboundedness, we have that $u(X) = \mathbb{R}$. Since $\succsim$ satisfies ISA, it satisfies Singleton Independence. Hence, an agent evaluate act $f \in \mathcal{F}$ by a subjective expected utility representation:

$$V(\varphi_{\{f\}}) = U(\{f\}) = \sum_{\omega\in\Omega} u(f(\omega))\overline{p}(\omega) = \langle\varphi_{\{f\}}, \delta_{\overline{p}}\rangle. \tag{29}$$

We want to show that $\sqcap$ is prior consistent with $\overline{p}$.

Lemma 18 *For any $\varphi \in C(\Delta(\Omega))$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,*

$$V(\varphi + \varphi_{\{f\}}) = V(\varphi) + V(\varphi_{\{f\}}). \tag{30}$$

Proof. First, we show that (30) holds for any $\varphi_F \in \Phi_{\mathbb{F}}$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$. For any menu F, we can take $x_F \in X$ such that $F \sim \{x_F\}$. ISA implies that $\alpha F + (1-\alpha)\{f\} \sim$

$\alpha\{x_F\} + (1-\alpha)\{f\}$ for any act f and $\alpha \in (0,1)$. Since $\succsim$ on $\mathcal{F}$ is represented by a subjective expected utility model, we have that for any $\alpha \in (0,1)$,

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

By taking $\alpha = \frac{1}{2}$, we have that $V(\frac{1}{2}\varphi_F + \frac{1}{2}\varphi_{\{f\}}) = \frac{1}{2}V(\varphi_F) + \frac{1}{2}V(\varphi_{\{f\}})$. Since V is positively homogeneous, we have that $V(\varphi_F + \varphi_{\{f\}}) = V(\varphi_F) + V(\varphi_{\{f\}})$, where $\frac{1}{2}\varphi_F + \frac{1}{2}\varphi_{\{f\}} = \varphi_{\frac{1}{2}F+\frac{1}{2}\{f\}} \in \Phi_{\mathbb{F}}$ implies that $\varphi_F + \varphi_{\{f\}} \in \Phi_{\mathbb{F}}$. This implies that (30) holds for any $\varphi_F \in \Phi_{\mathbb{F}}$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$.

Second, we show that (30) holds for any $\varphi \in C(\Delta(\Omega))$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$. Since $u(X) = \mathbb{R}$, we have that $\Phi_{\mathbb{F}} + \mathbb{R} = \Phi_{\mathbb{F}}$. Hence, a functional $V : C(\Delta(\Omega)) \to \mathbb{R}$ defined in Lemma 11 can be written as $V(\varphi) = \inf_{\varphi_F \in D(\varphi)} V(\varphi_F)$, where $D(\varphi) \equiv \{\varphi_F \in \Phi_{\mathbb{F}} \mid \varphi_F \geq \varphi\}$. We show that

$$\inf_{\varphi_F \in D(\varphi+\varphi_{\{f\}})} V(\varphi_F) = \inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{f\}}). \tag{31}$$

Since we can take $\varphi_F = \varphi_G + \varphi_{\{f\}}$, $\inf_{\varphi_F \in D(\varphi+\varphi_{\{f\}})} V(\varphi_F) \leq \inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{f\}})$ holds. As (28) in Lemma 14, we can show that $\inf_{\varphi_F \in D(\varphi+\varphi_{\{f\}})} V(\varphi_F) \geq \inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{f\}})$.

By (31), we have that for any $\varphi \in C(\Delta(\Omega))$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,

$$\begin{aligned} V(\varphi + \varphi_{\{f\}}) &= \inf_{\varphi_F \in D(\varphi+\varphi_{\{f\}})} V(\varphi_F) = \inf_{\varphi_G \in D(\varphi)} V(\varphi_G + \varphi_{\{f\}}) = \inf_{\varphi_G \in D(\varphi)} \{V(\varphi_G) + V(\varphi_{\{f\}})\} \\ &= \inf_{\varphi_G \in D(\varphi)} V(\varphi_G) + V(\varphi_{\{f\}}) = V(\varphi) + V(\varphi_{\{f\}}), \end{aligned}$$

where the third equality follows from (30) for $\varphi_F \in \Phi_{\mathbb{F}}$ and $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$. ■

Lemma 19 *For any $\pi \in \partial V(\mathbf{0})$, $p^\pi = \overline{p}$.*

Proof. Lemma 4 and 18 imply that for any $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,

$$\begin{aligned} V^\circ(\mathbf{0}; \varphi_{\{f\}}) &= \sup_{\varphi \in C(\Delta(\Omega))} \{V(\varphi + \varphi_{\{f\}}) - V(\varphi)\} = \sup_{\varphi \in C(\Delta(\Omega))} \{V(\varphi) + V(\varphi_{\{f\}}) - V(\varphi)\} \\ &= V(\varphi_{\{f\}}) = \left\langle \varphi_{\{f\}}, \delta_{\{\overline{p}\}} \right\rangle, \end{aligned}$$

where the last equality follows from (29). Similarly, Lemma 4 and 18 imply that for any $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,

$$\begin{aligned} V_\circ(\mathbf{0}; \varphi_{\{f\}}) &= \inf_{\varphi \in C(\Delta(\Omega))} \{V(\varphi + \varphi_{\{f\}}) - V(\varphi)\} = \inf_{\varphi \in C(\Delta(\Omega))} \{V(\varphi) + V(\varphi_{\{f\}}) - V(\varphi)\} \\ &= V(\varphi_{\{f\}}) = \left\langle \varphi_{\{f\}}, \delta_{\{\overline{p}\}} \right\rangle. \end{aligned}$$

Fix $\pi \in \partial V(\mathbf{0})$. By Lemma 5, we have that for any $\varphi_{\{f\}} \in \Phi_{\mathcal{F}}$,

$$\left\langle \varphi_{\{f\}}, \delta_{\{\overline{p}\}} \right\rangle = V_\circ(\mathbf{0}; \varphi_{\{f\}}) = \min_{\pi \in \partial V(\mathbf{0})} \langle \varphi_{\{f\}}, \pi \rangle \leq \left\langle \varphi_{\{f\}}, \pi \right\rangle \leq \max_{\pi \in \partial V(\mathbf{0})} \langle \varphi_{\{f\}}, \pi \rangle = V^\circ(\mathbf{0}; \xi) = \left\langle \varphi_{\{f\}}, \delta_{\{\overline{p}\}} \right\rangle.$$

Hence, we have that for any $f \in \mathcal{F}$,

$$\left\langle \varphi_{\{f\}}, \delta_{\{p^\pi\}} \right\rangle = \left\langle \varphi_{\{f\}}, \delta_{\{\overline{p}\}} \right\rangle \Longleftrightarrow \sum_{\omega \in \Omega} u(f(\omega)) p^\pi(\omega) = \sum_{\omega \in \Omega} u(f(\omega)) \overline{p}(\omega).$$

This implies that $p^\pi = \overline{p}$ for any $\pi \in \partial V(0)$. ■

By Lemma 19, we have $\mathsf{\Pi} \subset \mathcal{K}(\Delta(\Delta(\Omega)))$ that is prior-consistent with $\overline{p}$.

Turn to the necessity. We show that preference satisfies ISA. Take any menu F and any act f. For any $\alpha \in [0,1]$, it is enough to show that the representation U is mixture linear for $\alpha F + (1-\alpha)\{f\}$. Since $b^u_F(\pi)$ is mixture linear in F,

$$\begin{aligned} U(\alpha F + (1-\alpha)\{f\}) &= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_{\alpha F + (1-\alpha)\{f\}}(\pi) \\ &= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} (\alpha b^u_F(\pi) + (1-\alpha) b^u_{\{f\}}(\pi)) \\ &= \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} (\alpha b^u_F(\pi) + (1-\alpha) u(f) \cdot \overline{p}) \\ &= \alpha \max_{\Pi \in \mathsf{\Pi}} \min_{\pi \in \Pi} b^u_F(\pi) + (1-\alpha) u(f) \cdot \overline{p} \\ &= \alpha U(F) + (1-\alpha) U(\{f\}), \end{aligned}$$

as desired.

G Proof of Corollary 1

(1) If $\succsim$ satisfies ACP and Preference for Flexibility, by the proof of Claim 5 in de Oliveira, Denti, Mihm, and Ozbek [12], it satisfies IEPR. $\succsim$ is represented by a CIC representation $(u, \overline{P}, \mathsf{\Pi})$ by Theorem 5.

Under basic axioms and Independence with Lotteries, V is translation invariant. This implies that under ACP, V defined on $\Phi_{\mathbb{F}}$ is convex as shown in Claim 6 in de Oliveira, Denti, Mihm, and Ozbek [12]. As mentioned in Claim 6 and 7 in de Oliveira, Denti, Mihm, and Ozbek [12], an extension of V to $C(\Delta(\Omega))$ in Lemma 11 preserves convexity. If V defined on $C(\Delta(\Omega))$ is convex, Proposition 2.2.7 in Clarke [9] implies that $\partial V(\mathbf{0})$ coincides with the subdifferential of V at $\mathbf{0}$. Hence, we have that for any $\varphi \in C(\Delta(\Omega))$ and $\pi \in \partial V(\mathbf{0})$,

$$V(\varphi) \geq V(\mathbf{0}) + \langle \varphi - \mathbf{0}, \pi \rangle.$$

Since $V(\mathbf{0}) = 0$, $\langle \varphi, \pi \rangle \leq V(\varphi)$ for any $\varphi \in C(\Delta(\Omega))$ and $\pi \in \partial V(\mathbf{0})$. This implies that $\Pi_\varphi = \partial V(\mathbf{0})$ for any $\varphi \in C(\Delta(\Omega))$. Let $\Pi = \partial V(\mathbf{0})$. Then, for any $F \in \mathbb{F}$,

$$U(F) = \max_{\pi \in \Pi} \langle \varphi_F, \pi \rangle. \tag{32}$$

Take any acts f and g with $\{f\} \sim \{g\}$. ACP and Weak Concavity jointly imply $\{\alpha f + (1-\alpha)g\} \sim \{f\}$ for all $\alpha \in (0,1)$. By Maccheroni, Marinacci, and Rustichini [22, Corollary 20 and Lemma 29], this condition ensures that $\succsim$ on $\mathcal{F}$ is represented by a subjective expected utility model with prior $\overline{p}$. As in Claim 7 in de Oliveira, Denti, Mihm, and Ozbek [12], we will claim that $\Pi \subset \{\pi \in \Delta(\Delta(\Omega)) \,|\, p^\pi = \overline{p}\}$. Choose lotteries x and y such that $u(x) = 1$ and $u(y) = 0$. Fix $\omega \in \Omega$ and consider an act f yielding x on ω and y otherwise. By (32), we have that for any $\pi \in \Pi$,

$$\langle \varphi_{\{f\}}, \pi \rangle = \langle \varphi_{\{f\}}, \delta_{p^\pi} \rangle = p^\pi(\omega) \leq V(\varphi_{\{f\}}) = \overline{p}(\omega).$$

Hence, $p^\pi(\omega) \leq \overline{p}(\omega)$ for any $\omega \in \Omega$. Since $p^\pi \in \Delta(\Omega)$, $p^\pi(\omega) = \overline{p}(\omega)$ for any $\omega \in \Omega$. Thus, we have that $\Pi \subset \{\pi \in \Delta(\Delta(\Omega)) \,|\, p^\pi = \overline{p}\}$. Hence, (32) can be written as for any $F \in \mathbb{F}$,

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

(2) Since PCP implies Weak Concavity, the utility representation over singletons is taken to be a maxmin expected utility with multiple priors $\overline{P}$. As shown in part (1), V is translation invariant. Since $\succsim$ satisfies PCP, V defined on $\Phi_{\mathbb{F}}$ is concave. As mentioned in Claim 6 and 7 in de Oliveira, Denti, Mihm, and Ozbek [12], an extension of V to $C(\Delta(\Omega))$ in Lemma 11 preserves concavity. Proposition 2.2.7 in Clarke [9] implies that $\partial V(\mathbf{0})$ coincides with the superdifferential of V at $\mathbf{0}$. Hence, we have that for any $\varphi \in C(\Delta(\Omega))$ and $\pi \in \partial V(\mathbf{0})$,

$$V(\varphi) \leq V(\mathbf{0}) + \langle \varphi - \mathbf{0}, \pi \rangle.$$

Since $V(\mathbf{0}) = 0$, $\langle \varphi, \pi \rangle \geq V(\varphi)$ for any $\varphi \in C(\Delta(\Omega))$ and $\pi \in \partial V(\mathbf{0})$. This implies that $\Pi_\varphi = \partial V(\mathbf{0})$ for any $\varphi \in C(\Delta(\Omega))$. Let $\Pi = \partial V(\mathbf{0})$. Then, for any $F \in \mathbb{F}$,

$$U(F) = \min_{\pi \in \Pi} \langle \varphi_F, \pi \rangle. \tag{33}$$

We show the prior consistency, that is, $P(\Pi) = \overline{P}$. Since $\succsim$ satisfies Weak Concavity and the utility function over singletons is represented by a maxmin expected utility, by the same argument as in Lemmas 13 and 15, the prior consistency holds.

(3) Since ICP implies both ACP and PCP, V on $\Phi_{\mathbb{F}}$ satisfies linearity. It can be extended to $C(\Delta(\Omega))$ with preserving linearity. By combining the arguments as in parts (1) and (2), $V(\varphi) = \langle \varphi, \pi \rangle$ for any $\pi \in \partial V(\mathbf{0})$. Fix an arbitrary $\pi \in \partial V(\mathbf{0})$. As shown in part (1), the preference over singletons is represented by a subjective utility representation with a single prior $\overline{p}$. By the same argument as in part (1), $p^\pi = \overline{p}$, as desired.

H Proof of Proposition 3

(1) Assume that the preference is represented by the Costly CIC representation $(u, \overline{P}, \mathsf{\Pi}, c)$. We show that preference satisfies Weak Independence with Lotteries. Take any menu F,

any lottery x, and any $\alpha \in [0,1]$. Since $b^u_F(\pi)$ is mixture linear in F,

$$\begin{aligned}
U(\alpha F + (1-\alpha)\{x\}) &= \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} b^u_{\alpha F + (1-\alpha)\{x\}}(\pi) - c(\Pi)\} \\
&= \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} (\alpha b^u_F(\pi) + (1-\alpha) b^u_{\{x\}}(\pi)) - c(\Pi)\} \\
&= \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi) + (1-\alpha)u(x)\} \\
&= \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi)\} + (1-\alpha)u(x).
\end{aligned}$$

Therefore, for any other lottery y,

$$\begin{aligned}
& U(\alpha F + (1-\alpha)\{x\}) \geq U(\alpha G + (1-\alpha)\{x\}) \\
\iff & \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi)\} \geq \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_G(\pi) - c(\Pi)\} \\
\iff & U(\alpha F + (1-\alpha)\{y\}) \geq U(\alpha G + (1-\alpha)\{y\}),
\end{aligned}$$

as desired.

Next, we show that preference satisfies IEPR. Take any menu F and $\beta_i \geq 0$ with $\sum_i \beta_i = 1$. Since by linearity

$$b^u_{\sum_i \beta_i F}(\pi) = \sum_i \beta_i b^u_F(\pi) = b^u_F(\pi)$$

for any π,

$$U(F) = \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} b^u_F(\pi) - c(\Pi)\} = \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} b^u_{\sum_i \beta_i F}(\pi) - c(\Pi)\} = U(\sum_i \beta_i F),$$

as desired.

Next, we verify Star-Shaped. For all menus F and lotteries l,

$$\begin{aligned}
U(\alpha F + (1-\alpha)\{l\}) &= \max_{\Pi \in \mathbf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi)\} + (1-\alpha)u(l) \\
&= \alpha \max_{\Pi \in \mathbf{\Pi}} \left\{\min_{\pi \in \Pi} b^u_F(\pi) - \frac{c(\Pi)}{\alpha}\right\} + (1-\alpha)u(l). \qquad (34)
\end{aligned}$$

Since $c(\Pi) \leq \frac{c(\Pi)}{\alpha}$ for all $\alpha \in (0,1)$,

$$\max_{\Pi \in \mathbf{\Pi}} \left\{\min_{\pi \in \Pi} b^u_F(\pi) - \frac{c(\Pi)}{\alpha}\right\} \leq \max_{\Pi \in \mathbf{\Pi}} \left\{\min_{\pi \in \Pi} b^u_F(\pi) - c(\Pi)\right\} = U(F).$$

Thus, from(34),

$$\begin{aligned}
U(\alpha F + (1-\alpha)\{l\}) &\leq \alpha U(F) + (1-\alpha)U(\{l\}) \\
&= \alpha u(x_F) + (1-\alpha)u(l) \\
&= u(\alpha x_F + (1-\alpha)l).
\end{aligned}$$

(2) Assume that the preference is represented by the Costly CEC representation $(u, \{\overline{p}\}, \mathsf{\Pi}, c)$. Take any menu F, any act $\{f\}$, and any $\alpha \in [0,1]$. Since $b^u_F(\pi)$ is mixture linear in F,

$$\begin{aligned} U(\alpha F + (1-\alpha)\{f\}) &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} b^u_{\alpha F + (1-\alpha)\{f\}}(\pi) - c(\Pi)\} \\ &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} (\alpha b^u_F(\pi) + (1-\alpha) b^u_{\{f\}}(\pi)) - c(\Pi)\} \\ &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi) + (1-\alpha) u(f) \cdot \overline{p}\} \\ &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi)\} + (1-\alpha) u(f) \cdot \overline{p}. \end{aligned}$$

Therefore, for any other act g,

$$\begin{aligned} & U(\alpha F + (1-\alpha)\{f\}) \geq U(\alpha G + (1-\alpha)\{f\}) \\ \iff & \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi)\} \geq \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_G(\pi) - c(\Pi)\} \\ \iff & U(\alpha F + (1-\alpha)\{g\}) \geq U(\alpha G + (1-\alpha)\{g\}), \end{aligned}$$

as desired.

Next, we verify Strong Star-Shaped. For all menus F and acts f, as shown above,

$$\begin{aligned} U(\alpha F + (1-\alpha)\{f\}) &= \max_{\Pi \in \mathsf{\Pi}} \{\min_{\pi \in \Pi} \alpha b^u_F(\pi) - c(\Pi)\} + (1-\alpha) u(f) \cdot \overline{p} \\ &= \alpha \max_{\Pi \in \mathsf{\Pi}} \left\{ \min_{\pi \in \Pi} b^u_F(\pi) - \frac{c(\Pi)}{\alpha} \right\} + (1-\alpha) u(f) \cdot \overline{p}. \end{aligned} \tag{35}$$

Since $c(\Pi) \leq \frac{c(\Pi)}{\alpha}$ for all $\alpha \in (0,1)$,

$$\max_{\Pi \in \mathsf{\Pi}} \left\{ \min_{\pi \in \Pi} b^u_F(\pi) - \frac{c(\Pi)}{\alpha} \right\} \leq \max_{\Pi \in \mathsf{\Pi}} \left\{ \min_{\pi \in \Pi} b^u_F(\pi) - c(\Pi) \right\} = U(F).$$

Thus, from(35),

$$U(\alpha F + (1-\alpha)\{f\}) \leq \alpha U(F) + (1-\alpha) U(\{f\}).$$

Thus,

$$\begin{aligned} U(\alpha F + (1-\alpha)\{f\}) &\leq \alpha U(F) + (1-\alpha) U(\{f\}) \\ &= \alpha U(\{x_F\}) + (1-\alpha) U(\{f\}) \\ &= U(\{\alpha x_F + (1-\alpha) f\}). \end{aligned}$$

References

[1] Aliprantis, C. D. and K. C. Border (2006), *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Third Edition, Springer.

[2] Beauchene, D., J. Li, and M. Li (2019), “Ambiguous persuasion,” *Journal of Economic Theory*, 179, 312-365.

[3] Bose, S., and L. Renou (2014), “Mechanism design with ambiguous communication devices,” *Econometrica*, 82, 1853-1872.

[4] Çelen, B. (2012), “Informativeness of experiments for MEU, ” *Journal of Mathematical Economics*, 48, 404-406.

[5] Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and A. Rustichini (2004), “Niveloids and their extensions: risk measures on small domains, ” *Journal of Mathematical Analysis and Applications*, 413, 343-360.

[6] Chandrasekher, M., M. Frick, R. Iijima, and Y. L. Yaouanq (2022), “Dual-self representation of ambiguity preferences, ” *Econometrica*, 90, 1029-1061.

[7] Christensen, J. P. R. (1972), “On sets of Haar measure zero in Abelian Polish groups, ” *Israel Journal of Mathematics*, 13, 255-260.

[8] Christensen, J. P. R. (1974), *Topology and Borel Structure*, North-Holland.

[9] Clarke, F. H. (1990), *Optimization and Nonsmooth Analysis*, SIAM.

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

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

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

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

[14] Ellsberg, D. (1961), “Risk, ambiguity, and the Savage axioms, ” *Quarterly Journal of Economics*, 75, 643-669.

[15] Epstein, L. G., M. Marinacci, and K. Seo (2007), “Coarse contingencies and ambiguity, ” *Theoretical Economics*, 2, 355-394.

[16] Gensbittel, F., L. Renou, and T. Tomala (2015), “Comparisons of ambiguous experiments, ”working paper.

[17] Ghirardato, P., F. Maccheroni, M. Marinacci (2004), “Differentiating ambiguity and ambiguity attitude, ” *Journal of Economic Theory*, 118, 133-173.

[18] Ghirardato, P., F. Maccheroni, M. Marinacci (2004), “Ambiguity from a differential view point, ”working paper.

[19] Gilboa, I., and D. Schmeidler (1989), “Maxmin expected utility with non-unique prior, ”*Journal of Mathematical Economics*, 18, 141-153.

[20] Higashi, Y., K. Hyogo, and N. Takeoka (2022), “Costly subjective learning, ”working paper.

[21] Kamenica, E., and M. Gentzkow (2011), “Bayesian persuasion, ”*American Economic Review*, 101, 2590-2615.

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

[23] Manski, C. F., A. H. Sanstad, and S. J. DeCanio (2021), “Addressing partial identification in climate modeling and policy analysis, ”*Proceedings of the National Academy of Sciences*, 118, e2022886118.

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

[25] Thibault, L. (1982), “On generalized differentials and subdifferentials of Lipschitz functions, ”*Nonlinear Analysis*, 6, 1037-1053.

[26] Wang, Z. (2021), “Informativeness orders over ambiguous experiments, ”working paper.

[27] Xia, J. (2020), “Decision making under uncertainty: a game of two selves, ”working paper.