Regret and Information Avoidance*

Zichang Wang†

This version: August 7, 2025

Abstract

Empirical evidence suggests that individuals selectively avoid information based on past decisions. We address these findings by studying an agent who behaves as if she trades off two conflicting effects of information. The first is a psychological cost from the regret about past decisions that are revealed to be suboptimal, whereas the second is the instrumental value in guiding future decisions. We take as primitive preferences over consumption-information pairs and provide axiomatic representation theorems incorporating both instrumental value and regret in the value of information. New axioms are posited to connect the agent's consumption choice with her information choice. We show that all parameters can be uniquely identified from the preference. Defining information avoidance as preferring a Blackwell less informative experiment, we provide comparative statics on the agent's information aversion attitude.

KEYWORDS: Regret, information avoidance, value of information, preference for information, menu preference

JEL Code: D81, D83

*This paper is adapted from the first chapter of my dissertation at Duke University. I am deeply indebted to Todd Sarver for his invaluable advice and guidance throughout the development of this paper. I am grateful to Arjada Bardhi and Philipp Sadowski for many helpful discussions and detailed comments on an early draft of the paper. For feedback and suggestions, I want to thank Attila Ambrus, Andrew Caplin, Jaden Chen, David Dillenberger, Faruk Gul, Alexander Jakobsen, Fei Li, Yucheng Liang, Alfonso Maselli, Collin Raymond, Uzi Segal, Wing Suen, Curtis Taylor, Chen Zhao, four anonymous reviewers for EC'23, numerous speakers at the Duke-UNC micro theory seminar, and seminar participants at the University of Hong Kong, the 6th Triangle Micro Conference, and the 2023 Bounded Rationality in Choice (BRIC) conference. All errors are my own.

†Xiamen University, School of Economics and Wang Yanan Institute for Studies in Economics (WISE), 422 Siming South Road, Xiamen, Fujian, 361005, China. Email: zichang.wang@xmu.edu.cn

1 Introduction

Information avoidance describes an agent's unwillingness to obtain information that is both free and useful. Even though standard economic theories leave no room for information to be harmful in individual decision-making problems, substantial empirical evidence indicates that information avoidance is a pervasive phenomenon in real-world settings.[1]

What makes it more puzzling are the findings suggesting that information avoidance is often selective in decisions made in the past. More specifically, individuals might only avoid an information source *after* a decision has been made, while the same information source will be consulted if that decision were nonexistent. This is referred to as "selective exposure to information" or "selective information avoidance."

- A large empirical literature in psychology documents selective information avoidance about consumption goods: Consumers only avoid information about products they did not choose or about risks associated with products they have consumed (e.g., Ehrlich et al. 1957, Frey and Stahlberg 1986, and Jonas et al. 2001).
- In the context of financial information, Sicherman et al. (2016) find that stock investors log into their account less frequently to avoid information on their individual stocks when the stock market (as a whole) is not performing well. Moreover, investors who hold more bonds than stocks show a weaker tendency to avoid information and investors holding only bonds exhibit no information avoidance at all.

Why would an agent refuse to be better informed for future decisions due to a choice made in the past? One natural reason is regret. Information could reveal an agent's past decision to be suboptimal and cause her to experience a sense of loss for not having chosen a different alternative. We refer to this (psychological) effect as the *regret cost* of information.

Information usually arrives in the middle of a sequence of decisions, thus revealing suboptimal past decisions (causing regrets) and guiding future decisions (generating instrumental value) at the same time. For example, a physical examination that helps to screen for potential health problems might also lead to the discovery that an old lifestyle was unhealthy. Inquiries about positions from different industries for a new job might reveal to a worker that better early-career choices could have been made. Stock analyses for a new investment project might unveil that a better portfolio could have been constructed in the past.

[1] Lerman et al. (1999), Oster, Shoulson, and Dorsey (2013) and Persoskie et al. (2014) are among the large literature documenting individuals avoiding medical information when at risk of certain diseases or health conditions. A rapidly growing literature studies investors' aversion to financial information. Examples include Karlsson, Loewenstein, and Seppi (2009), Sicherman et al. (2016) and Hilbert et al. (2022).

In this paper, we address the findings on selective information avoidance by proposing and axiomatizing a model based on the tradeoff between the regret cost and the instrumental value of any information source. In our model, an agent avoids information when (and only when) its regret cost outweighs its instrumental value. A decision made in the past (or the absence thereof) exposes the agent to (or protects the agent from) the possibility of regret, causing selectiveness of information avoidance on past decisions.

We build a model of decision making under uncertainty that involves choices in three periods. The uncertainty is captured by a set of payoff-relevant states and information is represented by Blackwell experiments over these states. In period 3 (the final period), the agent chooses an (Anscombe-Aumann) act from a menu of acts F after the arrival of information. That is, the act choice can be customized according to what the information reveals. In contrast, the menu of acts F itself has been selected from a menu of menus $\mathbb{F}$ in period 2 *before* the information arrives. Thus, the information could cause regret if it reveals that some unchosen menu G contains an act g that is better than all acts in F. In period 1, the agent chooses a *consumption-information pair* consisting of a menu of menus of acts and a Blackwell experiment, having in mind the subsequent choices of a menu and an act.

Using a menu of menus approach to model the consumption component reflects the common decision procedure in which an agent narrows down her set of options over time before settling on a final alternative. Using a Blackwell experiment to model the information component allows a notion of information avoidance that is more broadly applicable.

In our model, information avoidance is defined as an agent preferring a (Blackwell) less informative experiment to a more informative one. This general notion nests the common notion of information avoidance where an agent prefers no information to full information (fully revealing experiments) or partial information (general experiments).

Taking preferences over consumption-information pairs as primitive, our key result is a representation theorem that features an *informational tradeoff (IT) representation*, where the agent behaves *as if* she takes into account both the instrumental value and the regret cost when evaluating an information source. As a result, our model can generate both information-avoiding and information-seeking behavior depending on whether the regret cost or the instrumental value is dominant.

The information component of the choice domain opens the possibility for comparative static analysis on the degree of information aversion. We say agent 1 is *more information averse than* agent 2 if agent 1 avoids information whenever agent 2 avoids information. Another main result of the paper is a clean characterization of the comparative statics on agents' information aversion attitudes. It is impossible to establish such a result without data on preference for information.

In the axiomatization process of the IT representation, we obtain a subjective informational tradeoff (SIT) representation in which the agent's information is an unobserved preference parameter instead of a choice variable. We show that such a representation can be characterized from axioms on the agent's preference over menus of menus alone. Moreover, the unobserved information can be elicited from this preference.

The SIT representation is of interest on its own. Conceptually, the agent might still recognize both the regret cost and the instrumental value of an information structure even if it is exogenously given. Pragmatically, it is valuable to be able to elicit an agent's information because a modeler may not be able to observe this parameter of interest.

We have two key axioms for the SIT representation.

- The first key axiom is called "Ex-Ante Regret". It reflects the agent's desire to limit her options before the arrival of information. Formally, let F and G be two menus, and $\mathbb{F}$ a set of menus. The axiom states that if $\{F\} \succsim \{G\}$ and $F \in \mathbb{F}$, then $\mathbb{F} \succsim \mathbb{F} \cup \{G\}$.[2] Due to the existence of a better menu F, having the extra option of G actually makes $\mathbb{F}$ less preferred. The intuition is that, adding a menu that will not be chosen subsequently does not increase material value but could cause more regret.
- The second key axiom is called "Interim Preference for Flexibility". It reflects the agent's preference to have more options after the arrival of information. Formally, the axiom states that $\mathbb{F} \cup \{F \cup G\} \succsim \mathbb{F} \cup \{F, G\}$ for any menu of menus $\mathbb{F}$ and any menus F and G. That is, when the set of acts that can be ultimately chosen are the same, the agent prefers to keep more options open so she has more flexibility in period 3.

Our axiomatization of the IT representation builds on the axiomatic characterization of the SIT representation. In addition to axioms for the SIT representation, we impose axioms to discipline the interaction between the agent's consumption choice and information choice.

A novel axiom termed "Balance" is utilized to characterize the menu of menus where the regret cost and instrumental value for an information source are balanced out. More specifically, given a menu of acts $F = \{f_1, f_2, \dots\}$. An early-commitment scenario is $D(F) = \big\{\{f_1\}, \{f_2\}, \dots\big\}$ and a full-flexibility scenario is $\big\{F\big\} = \big\{\{f_1, f_2, \dots\}\big\}$. Information only causes regret in an early-commitment scenario whereas it only adds to the instrumental value in a full-flexibility scenario. The Balance axiom guarantees that the regret generated by any information for $D(F)$ is proportional to the instrumental value generated by the same piece of information for $\{F\}$.

We now describe the functional forms identified from our representation theorems.

[2] Despite its similarity with the main axiom (dominance) of Sarver (2008), our axiom is adapted to be imposed on the preference over menus of menus of acts instead of menus of lotteries.

Let Ω and X be finite sets of states and outcomes, respectively.

An information structure is a mapping $\sigma : \Omega \to \Delta(S)$ where S is a finite set of signals. An act is a mapping $f : \Omega \to \Delta(X)$. Let F be a menu of acts and $\mathbb{F}$ a menu of menus. The informational tradeoff (IT) representation for a pair $(\mathbb{F}, \sigma)$ is

$$W(\mathbb{F}, \sigma) := \max_{F\in\mathbb{F}} \sum_{s\in S} \sigma_s \Big[U(F, \mu_s^\sigma) - R(F, \mathbb{F}, \mu_s^\sigma) \Big] \tag{1}$$

where σ_s is the ex-ante probability that signal s is generated while μ_s^σ is the Bayesian posterior if s is observed,[3] and

$$U(F, \mu_s^\sigma) := \max_{f\in F} \sum_{\omega\in\Omega} \mu_s^\sigma(\omega) u\big(f(\omega)\big)$$

captures the material value of a menu F under posterior μ_s^σ, where $u : \Delta(X) \to \mathbb{R}$ is an affine function over lotteries that captures the agent's taste over outcomes, and

$$R(F, \mathbb{F}, \mu_s^\sigma) := K \left[\max_{G\in\mathbb{F}} U(G, \mu_s^\sigma) - U(F, \mu_s^\sigma) \right]$$

captures the agent's regret for having chosen menu F from a set of menus $\mathbb{F}$ at posterior μ_s^σ. That is, the agent's regret is proportional to the difference of the material value for the menu she has chosen and the highest material value she could have obtained from another menu in $\mathbb{F}$. The intensity of regret is captured by the scalar $K \geq 0$.

In a SIT representation, a menu of menus of acts $\mathbb{F}$ is evaluated by

$$V(\mathbb{F}) = \max_{F\in\mathbb{F}} \sum_{s\in S} \sigma_s \Big[U(F, \mu_s^\sigma) - R(F, \mathbb{F}, \mu_s^\sigma) \Big] \tag{2}$$

where σ_s, μ_s^σ, $U(F, \mu_s^\sigma)$ and $R(F, \mathbb{F}, \mu_s^\sigma)$ are the same as in the IT representation. Although $W(\cdot, \cdot)$ and $V(\cdot)$ correspond to the same functional form, they describe functions over different domains. In particular, the information structure σ is a choice variable in an IT representation while it is an unobserved parameter in a SIT representation.

The SIT representation generalizes two well-established representations in the literature: The regret representation in Sarver (2008) and the subjective learning representation in Dillenberger, Lleras, Sadowski, and Takeoka (2014, henceforth DLST).

We establish uniqueness of all parameters in both the IT representation and the SIT representation. In particular, we solve the non-identification problem in Sarver (2008) where the regret intensity parameter K cannot be uniquely identified.

We proceed by illustrating our representations with the following numerical example.

[3] Formally, $\sigma_s = \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega)$ where $\pi \in \Delta(\Omega)$ is a prior and $\sigma(s \mid \omega)$ is the probability for observing signal s when the state is ω. For any signal s with $\sigma_s > 0$, the Bayesian posterior $\mu_s^\sigma \in \Delta(\Omega)$ is $\mu_s^\sigma(\omega) = \frac{\pi(\omega)\sigma(s|\omega)}{\sigma_s}$. Without loss of generality, we set the posterior μ_s^σ to be uniform if $\sigma_s = 0$.

A numerical example. A college student aims to get a job with the best career prospect upon graduation. Each job is modeled as an act depending on the general labor market condition. Each major is a menu of jobs: the economics major could be Econ $= \{b, c\}$ whereas the computer science major could be CS $= \{c, e\}$ where b, c, e stands for banking (b), consulting (c) and engineer (e), respectively. Colleges are modeled as menus of majors with $\mathbb{F}_1 = \{\text{Econ, CS}\} = \{\{b, c\}, \{e\}\}$, and $\mathbb{F}_2 = \{\text{Econ}\} = \{\{b, c\}\}$.

Uncertainty about career prospects can be captured by a binary state space $\Omega = \{\omega_1, \omega_2\}$. The student's prior belief is $\pi(\omega_1) = \pi(\omega_2) = 0.5$. The state-dependent utility is in table (a).

u	b	c	e
ω_1	110	100	130
ω_2	60	90	50

(a) State-dependent utilities.

σ	s_1	s_2
ω_1	1	0
ω_2	0	1

(b) A revealing information structure.

Suppose $K \geq 0$ and that the student learns the true state upon graduation, which corresponds to a fully revealing Blackwell experiment as in table (b).

We first illustrate the SIT representation. Suppose the student has chosen the economics major from $\mathbb{F}_1$, her ex-post regret will be $R(\text{Econ}, \mathbb{F}_1, s_1) = 20K$ if signal s_1 is observed and 0 if signal s_2 is observed.[4] The material utility (U) and ex-post regret (R) for choosing each major from $\mathbb{F}_1$ can be summarized by the following tables.

U	Econ	CS
s_1	110	130
s_2	90	50

(a) Ex-post material utility.

R	Econ	CS
s_1	$20K$	0
s_2	0	$40K$

(b) Ex-post regret.

The ex-ante probability for each signal is $\sigma_{s_1} = \sigma_{s_2} = 0.5$. Therefore, the expected net value is $100 - 10K$ for the economics major in $\mathbb{F}_1$, and $90 - 20K$ for the computer science major. Since $K \geq 0$, the student would choose Econ from $\mathbb{F}_1$. Therefore, $V(\mathbb{F}_1) = 100 - 10K$.

The value for $\mathbb{F}_2$ is $V(\mathbb{F}_2) = 100$. With $K > 0$, the student would strictly prefer College 2 over College 1 because College 1 only exposes her to regret by offering the CS major.

To illustrate the IT representation, consider the parametrized information structures

σ_θ	s_1	s_2
ω_1	θ	$1-\theta$
ω_2	$1-\theta$	θ

[4] The signal s_1 reveals the state to be ω_1. The student's best option is b with $u(b(\omega_1)) = 110$ while career e with $u(e(\omega_1)) = 130$ is no longer available. Calculation for $R(\text{Econ}, \mathbb{F}_1, s_2)$ is similar.

where $\theta \in [0.5, 1]$ is a parameter for the precision of the signals.

Suppose a student from College 1 can control the precision θ. Applying the IT representation, we have $W(\mathbb{F}_1, \sigma_1) = 100 - 10K$ and $W(\mathbb{F}_1, \sigma_{0.5}) = 95$. Therefore, as long as $K > 0.5$, the student would prefer no information ($\theta = 0.5$) to full information ($\theta = 1$).

In Figure 1, we vary θ from 0.5 to 1 and draw the value of information as a function of θ. The upper curve corresponds to a regret intensity $K = 0.4$ while the lower curve is drawn under $K = 0.8$. Both curves are non-monotonic, thus allowing room for both information-avoiding and information-seeking behavior.

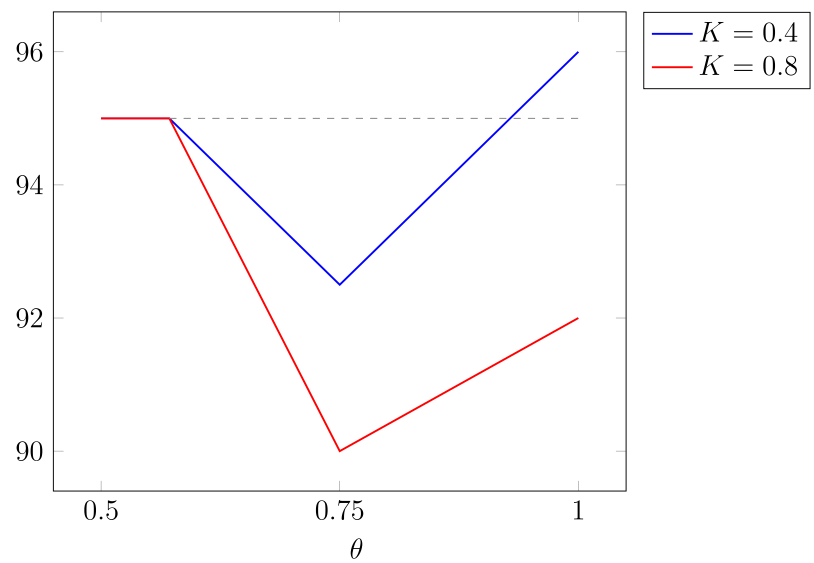

Figure 1: Value of Information σ_θ, $W(\mathbb{F}_1, \sigma_\theta)$, as a function of θ

The remainder of the paper is organized as follows. We set up the model in Section 2. We provide an axiomatization for the SIT representation in Section 3 to build towards the axiomatic theorem for the IT representation in Section 4. Section 5 concludes with a discussion of the related literature and an extension.

2 The Model

Let Ω be a finite set of states with $|\Omega| \geq 2$. An *information structure* is a Blackwell experiment $\sigma : \Omega \to \Delta(S)$ where $S \subseteq [0, 1]$ is a finite set of signals.[5] We write $\sigma(s \mid \omega)$ to denote the probability of observing signal s when the state is ω. Let $\mathcal{I}$ denote the set of all information structures.

Let X be a finite set of prizes with $|X| \geq 2$. An (Anscombe-Aumann) *act* is a mapping $f : \Omega \to \Delta(X)$. Let $\mathcal{F}$ be the set of all acts, endowed with the Euclidean metric d. A *menu*

[5]For finite Y, $\Delta(Y)$ is the probability simplex over Y, equipped with the usual mixture operation.

is a nonempty compact subset of $\mathcal{F}$, typically denoted by F, G. Let $\mathcal{M}$ be the set of all menus. Endow $\mathcal{M}$ with the Hausdorff metric d_h.[6] A *menu of menus* is a nonempty compact subset of $\mathcal{M}$, typically denoted by $\mathbb{F}, \mathbb{G}$. Let $\mathcal{D}$ be the set of all menus of menus. We endow $\mathcal{D}$ with the Hausdorff metric d_H.[7] Both $\mathcal{M}$ and $\mathcal{D}$ are compact.[8]

The set $\mathcal{F}$ is equipped with the standard mixture operation. If $f, g \in \mathcal{F}$ and $\alpha \in [0,1]$, then $\alpha f + (1-\alpha)g$ is an act defined by $\big(\alpha f + (1-\alpha)g\big)(\omega) := \alpha f(\omega) + (1-\alpha)g(\omega)$. For any $F, G \in \mathcal{M}$ and $\alpha \in [0,1]$, $\alpha F + (1-\alpha)G := \{\alpha f + (1-\alpha)g \mid f \in F \text{ and } g \in G\}$. For any $\mathbb{F}, \mathbb{G} \in \mathcal{D}$ and $\alpha \in [0,1]$, $\alpha \mathbb{F} + (1-\alpha)\mathbb{G} := \{\alpha F + (1-\alpha)G \mid F \in \mathbb{F} \text{ and } G \in \mathbb{G}\}$.

The primitive of our model is a binary relation on $\mathcal{D} \times \mathcal{I}$, representing the agent's preference over pairs consisting of a menu of menus and an information structure. We have in mind an agent facing a three-period decision problem. In period 1, the agent jointly chooses a menu of menus $\mathbb{F}$ and an information structure σ. In period 2, the agent chooses a menu F from $\mathbb{F}$, anticipating the information to arrive after this choice. A signal $s \in S$ is then generated according to σ and observed by the agent. In period 3, the agent updates her belief and chooses an act f from F.[9] We do not explicitly model the agent's choices in periods 2 and 3, leaving them as part of the interpretations of the model. The timeline for events is summarized in Figure 2.

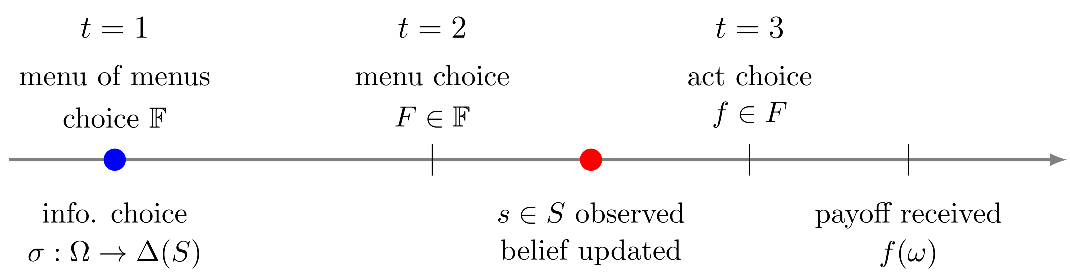

Figure 2: Timeline of Events

3 Subjective IT Representation

We first consider a binary relation $\succsim$ on $\mathcal{D}$, that is, the agent's preference over menus of menus of acts. The timeline of events in this subdomain is almost identical to that in Figure 2, except that the agent only chooses a menu of menus $\mathbb{F}$ in period 1 while the information structure becomes a preference parameter that is not directly observed by the modeler. We show in Section 3.1 that this parameter can be elicited from $\succsim$.

[6]This is the metric defined by $d_h(F,G) := \max\{\max_{f\in F}\min_{g\in G} d(f,g), \max_{f\in G}\min_{g\in F} d(f,g)\}$.

[7]d_H is defined by $d_H(\mathbb{F},\mathbb{G}) := \max\{\max_{F\in\mathbb{F}}\min_{G\in\mathbb{G}} d_h(F,G), \max_{F\in\mathbb{G}}\min_{G\in\mathbb{F}} d_h(F,G)\}$.

[8]See, for example, Aliprantis and Border (2006, Theorem 3.85).

[9]In some applications, it is more realistic for the information choice to be made after the menu choice. Our model can easily accommodate this change since our agent is both forward-looking and Bayesian.

Definition 1: A *subjective informational tradeoff (SIT) representation* is a tuple (π, u, K, σ) that consists of a probability measure π on Ω, a non-constant affine function $u : \Delta(X) \to \mathbb{R}$, a constant $K \geq 0$, and an information structure $\sigma : \Omega \to \Delta(S)$ such that $\succsim$ can be represented by the function $V : \mathcal{D} \to \mathbb{R}$ defined by

$$V(\mathbb{F}) = \max_{F \in \mathbb{F}} \sum_{s \in S} \sigma_s \Big[U(F, \mu_s^\sigma) - R(F, \mathbb{F}, \mu_s^\sigma) \Big] \tag{3}$$

where $\sigma_s = \sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega)$ is the ex-ante probability of signal s, μ_s^σ is the agent's Bayesian posterior after observing s,[10] and

$$U(F, \mu_s^\sigma) = \max_{f \in F} \sum_{\omega \in \Omega} \mu_s^\sigma(\omega) u\big(f(\omega)\big) \tag{4}$$

is the highest expected value that can be obtained by an act in F. Moreover,

$$R(F, \mathbb{F}, \mu_s^\sigma) = K \left[\max_{G \in \mathbb{F}} U(G, \mu_s^\sigma) - U(F, \mu_s^\sigma) \right] \tag{5}$$

is the regret for having chosen menu F from the menu of menus $\mathbb{F}$ after observing signal s. Lastly, $K \geq 0$ represents the agent's regret intensity.

We impose eight axioms on $\succsim$. The first three axioms are standard in the setting of preferences over menus of lotteries and are simply adapted to our setting.

Axiom 1—Weak Order: *$\succsim$ is complete and transitive.*

Axiom 2—Continuity: *For any $\mathbb{F} \in \mathcal{D}$, the sets $\{\mathbb{G} : \mathbb{F} \succsim \mathbb{G}\}$, $\{\mathbb{G} : \mathbb{G} \succsim \mathbb{F}\}$ are closed.*

Axiom 3—Independence: *For any $\mathbb{F}, \mathbb{G}, \mathbb{H} \in \mathcal{D}$ and any $\alpha \in (0, 1)$,*

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

We refer readers to Dekel, Lipman, and Rustichini (2001), Dekel, Lipman, Rustichini, and Sarver (2007) and Kopylov (2009) for a discussion of these axioms.

To state the next axiom, we introduce the notion of a "critical" subset.

Definition 2: We say $\mathbb{G}$ is *critical* for $\mathbb{F}$ if $\mathbb{G} \subseteq \mathbb{F}$ and $\mathbb{H} \sim \mathbb{F}$ for any $\mathbb{H}$ satisfying $\mathbb{G} \subseteq \mathbb{H} \subseteq \mathbb{F}$. We say G is *critical for F in* $\mathbb{F}$ if $G \subseteq F \in \mathbb{F}$ and $(\mathbb{F} \setminus \{F\}) \cup \{H\} \sim \mathbb{F}$ for any menu H satisfying $G \subseteq H \subseteq F$.

In words, if $\mathbb{G}$ is critical for $\mathbb{F}$, then menus in $\mathbb{F}$ but outside $\mathbb{G}$ are irrelevant for the evaluation of $\mathbb{F}$. The intuition is similar for a menu G being critical for F in $\mathbb{F}$.

[10] If $\sigma_s > 0$, $\mu_s^\sigma(\omega) = \pi(\omega)\sigma(s \mid \omega)/\sigma_s$. Without loss, we set μ_s^σ to be uniform whenever $\sigma_s = 0$.

Axiom 4—Finiteness: *There exists a natural number N such that: (4.1) For every $\mathbb{F}$, there exists $\mathbb{G}$ with $|\mathbb{G}| < N$ such that $\mathbb{G}$ is critical for $\mathbb{F}$; and (4.2) For every $\mathbb{F}$ and every $F \in \mathbb{F}$, there exists G with $|G| < N$ such that G is critical for F in $\mathbb{F}$.*

In essence, Axiom 4 states that: Given a menu of menus, only a finite number of menus within in it and only a finite number of acts within any menu matter for its evaluation. This restricts the agent to entertain only a finite number of beliefs over Ω as possible posteriors.

Axiom 4 is adapted from the Finiteness axiom from Stovall (2018), who studies preferences over menus of menus of lotteries. We refer the reader to Stovall (2018) for a more detailed discussion of this axiom and its connection with the finiteness axioms stated in Dekel, Lipman, and Rustichini (2009) and Kopylov (2009).

The following axiom is the key to capture regret:

Axiom 5—Ex-Ante Regret: *If $\{F\} \succsim \{G\}$ and $F \in \mathbb{F}$, then $\mathbb{F} \succsim \mathbb{F} \cup \{G\}$.*

Axiom 5 is adapted from the Dominance axiom from Sarver (2008). Contrary to standard models, Axiom 5 allows for the possibility that $\mathbb{F} \succ \mathbb{F} \cup \{G\}$, that is, the agent may strictly prefer not to add a menu G to a menu of menus $\mathbb{F}$. The exact situations in which the agent might want to limit the size of a menu of menus are when the added menu G will never be subsequently chosen from $\mathbb{F} \cup \{G\}$. Intuitively, if G is not chosen in period 2, then it carries no instrumental value. However, its presence can cause regret if the information reveals G to have higher material value than F.

The next axiom reflects the fact that information can still benefit the agent because she values flexibility for her intermediate choice.

Axiom 6—Interim Preference for Flexibility: $\mathbb{F} \cup \{F \cup G\} \succsim \mathbb{F} \cup \{F, G\}$ *for any menu of menus $\mathbb{F}$ and menus F, G.*

Note that the two menus of menus in the comparison correspond to the same set of acts that can be ultimately chosen. They only differ in how much flexibility can be retained after the arrival of information. The agent can choose $F \cup G$ from the menu of menus on the left hand side and choose acts from both F and G after observing any signal. But she has to decide between F or G from the right-hand-side menu of menus before the information arrives. Selecting F means she has to let go acts that are in G but not in F.

We follow a common practice that uses ℓ to also denote the constant act that yields lottery ℓ in every state. For simplicity, we also write $\ell \succ \ell'$ for $\{\{\ell\}\} \succ \{\{\ell'\}\}$.

Axiom 7—Nontriviality: *There exist lotteries $\ell, \ell' \in \Delta(X)$ with $\ell \succ \ell'$.*

To state the last axiom, we need to introduce the notion of domination.

DEFINITION 3: An act f *dominates* act g if $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$. A menu F *dominates* menu G if for any $g \in G$, there exists $f \in F$ such that f dominates g. A menu of menus $\mathbb{F}$ *dominates* $\mathbb{G}$ if for any $G \in \mathbb{G}$, there exists $F \in \mathbb{F}$ such that F dominates G.

This definition is built on state-by-state comparisons between acts.

AXIOM 8—Domination: *(1) If f dominates g, then $f \succsim g$ and $\{\{f, g\}\} \sim \{\{f\}\}$; (2) If $\mathbb{F}$ dominates $\mathbb{G}$, then $\mathbb{F} \sim \mathbb{F} \cup \mathbb{G}$.*

The first half of part 1 imposes monotonicity to the agent's preference over acts. The second half of part 1 states that adding a dominated act to a menu does not change an agent's attitude toward the menu. The intuition is that a dominated act contributes to neither the material value nor the regret. The intuition for part 2 is similar.

THEOREM 1: *A binary relation $\succsim$ over $\mathcal{D}$ has a subjective informational tradeoff representation if and only if it satisfies Axioms 1-8.*

The proof of Theorem 1 is contained in Appendix B.

The necessity of the axioms is relatively easy to check.

There are two steps to prove the sufficiency of Axioms 1-8. First, we axiomatize a representation (through Theorem 6) tightly connected with the SIT representation. This so-called *partial regret* (PR) representation is for preferences over menus of menus of lotteries. The PR representation is the counterpart of the SIT representation within a framework using lotteries. The setup for this framework and preliminary results are contained in Appendix A. In the second step, we use a sequence of geometric arguments to establish a connection between preferences over menus of menus of acts (for the SIT representation) and preferences over menus of menus of lotteries (for the PR representation).

3.1 Uniqueness of the SIT Representation

A *distribution over posteriors,* denoted by ν, is a finitely-supported probability measure on $\Delta(\Omega)$. We say that ν is *induced by* a prior π and an information structure $\sigma : \Omega \to \Delta(S)$ if $\nu(\mu) = \sum_{s \in S} \mathbf{1}[\mu = \mu_s^\sigma]\sigma_s$ for any $\mu \in \Delta(\Omega)$, where $\mathbf{1}[\cdot]$ is the indicator function, and σ_s and μ_s^σ are defined the same as before.

Note from equation (3) that if two information structures σ and σ' induce the same distribution over posteriors given a prior π, then the SIT representations with parameters (π, u, K, σ) and (π, u, K, σ') represent the same preference over $\mathcal{D}$. Therefore, we can only identify the information up to the distribution over posteriors it induces (in other words, the equivalence class of experiments that are Blackwell equivalent).

We say an information structure σ induces a *degenerate* distribution over posteriors if the induced distribution puts weight 1 on the prior belief.

THEOREM 2: *Suppose both* $(\pi_0, u_0, K_0, \sigma_0)$ *and* (π, u, K, σ) *represent* $\succsim$, *then* $\pi_0 = \pi$. $u_0 = au + b$ *for some* $a > 0$ *and* $b \in \mathbb{R}$. σ_0 *and* σ *induce the same distribution over posteriors.* $K_0 = K$ *if* σ *induces a non-degenerate distribution over posteriors. If* σ *induces a degenerate distribution over posteriors, then* $K_0, K \in [0, \infty)$.

The proof of Theorem 2 is relegated to Appendix C.1.

The identifications of the prior π and the taste u are standard.

Separate identification of the regret intensity K and the information σ is achieved through applying the SIT representation to two special types of menus of menus.

For any menu F, let $D(F) := \{\{f\} \mid f \in F\}$. We say a menu of menus $\mathbb{F}$ corresponds to an *early-commitment* if $\mathbb{F} = D(F)$ for some F.

If $\mathbb{F}$ corresponds to an early-commitment, $\mathbb{F}$ contain only singleton menus. That is, the agent has to commit to an act before she receives any information. Thus, information carries zero instrumental value and only contributes to regret. Formally,

$$V\big(D(F)\big) = \max_{f \in F} \left[(1+K) \sum_{\omega \in \Omega} \pi(\omega) u(f(\omega)) \right] - K \sum_{s \in S} \max_{g \in F} \sum_{\omega} \pi(\omega)\sigma(s \mid \omega) u(g(\omega)). \tag{6}$$

This corresponds to a finite version of the regret representation characterized in Sarver (2008) in the framework with acts. As in Sarver (2008), we can jointly identify σ and K as long as σ does not induce a degenerate distribution over posteriors. However, we run into a similar problem as in Sarver (2008) when we try to separately identify σ and K using only the representation in equation (6). The problem is that we might not always be able to distinguish an agent A who has a large intensity of regret but anticipates an information structure that is not likely to update her belief away from the prior belief, and another agent B who has a low level of regret intensity but anticipates an information structure that is more likely to update her belief away from her prior. More precisely, (σ_1, K) and $(\sigma_2, 2K)$ will generate the same preference according to equation (6) if σ_2 is obtained by mixing σ_1 with an uninformative Blackwell experiment with equal weights.

Luckily, we can solve this problem by taking advantage of our larger choice domain. We say that a menu of menus $\mathbb{F}$ corresponds to *full-flexibility* if $\mathbb{F} = \{F\}$ for some menu F. That is, the agent can make all her relevant decisions after the information arrives. Therefore, information causes zero regret and only carries instrumental value. With full-flexibility, the SIT representation reduces to

$$V\big(\{F\}\big) = \sum_{s \in S} \max_{f \in F} \sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega) u(f(\omega)). \tag{7}$$

Equation (7) corresponds to a finite version of the subjective learning representation characterized in DLST. DLST show that the information structure can be uniquely identified up to the induced distribution over posteriors. The identification of K follows once σ is identified.

4 Informational Tradeoff Representation

With the axioms for the SIT representation in hand, we switch back to the larger choice domain $\mathcal{D} \times \mathcal{I}$. To minimize notation, we use $\succsim$ to denote a binary relation over $\mathcal{D} \times \mathcal{I}$ from this point on.[11]

DEFINITION 4: A binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ has an *informational tradeoff (IT) representation* if there exists a triple (π, u, K) that consists of a probability measure π on Ω, a non-constant affine function $u : \Delta(X) \to \mathbb{R}$, and a constant $K \geq 0$ such that $\succsim$ can be represented by the function $W : \mathcal{D} \times \mathcal{I} \to \mathbb{R}$ defined by

$$W(\mathbb{F}, \sigma) := \max_{F \in \mathbb{F}} \sum_{s \in S} \sigma_s \Big[U(F, \mu_s^\sigma) - R(F, \mathbb{F}, \mu_s^\sigma)\Big] \tag{8}$$

where σ_s, μ_s^σ, material utility for a menu $U(F, \mu_s^\sigma)$, and the regret $R(F, \mathbb{F}, \mu_s^\sigma)$ are all defined in the same way as before.

As illustrated in the numerical example in the Introduction, given a menu of menus of acts $\mathbb{F}$, it is possible that $W(\mathbb{F}, \sigma) > W(\mathbb{F}, \sigma')$ even when σ' is Blackwell more informative than σ. Formally, we say an agent *avoids information* facing $\mathbb{F}$ if $(\mathbb{F}, \sigma) \succ (\mathbb{F}, \sigma')$ with σ' being Blackwell more informative than σ. Blackwell's informativeness theorem dictates that being Blackwell more informative is equivalent to generating a higher expected utility in any standard decision problem. Therefore, using such preference reversals is an intuitive (and in a sense canonical) way to define information avoidance.

Given a preference $\succsim$ over $\mathcal{D} \times \mathcal{I}$, each information structure $\sigma \in \mathcal{I}$ induces a preference over menus of menus of acts. Formally, a *conditional preference* $\succsim^\sigma$ is a binary relation over $\mathcal{D}$ where we say $\mathbb{F} \succsim^\sigma \mathbb{G}$ if $(\mathbb{F}, \sigma) \succsim (\mathbb{G}, \sigma)$.

In addition to previous axioms for the SIT representation (which will be imposed within each $\succsim^\sigma$), we impose new axioms to discipline the agent's choices between menus of menus across *different* information structures.

The first one is the standard rationality requirement.

AXIOM 9—Weak Order: $\succsim$ *(as a preference over $\mathcal{D} \times \mathcal{I}$) is complete and transitive.*

[11] Jakobsen (2021) uses a similar primitive to axiomatize Bayeisan persuasion models. In his paper, the modeler observes a sender's preference over information indexed by menus of acts and a receiver's choice correspondence from menus indexed by signals.

Note that if a binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ is complete and transitive, then the induced conditional preferences $\succsim^\sigma$ must also be complete and transitive.

The next axiom imposes an independence requirement on mixing menus of menus with acts. For convenience, we write $\alpha\mathbb{F} + (1-\alpha)h$ for $\alpha\mathbb{F} + (1-\alpha)\{\{h\}\}$.

Axiom 10—Act Independence: *For any* $(\mathbb{F}, \sigma)$, $(\mathbb{G}, \sigma')$, *any act* h *and any* $\alpha \in (0,1)$,

$$(\mathbb{F}, \sigma) \succsim (\mathbb{G}, \sigma') \iff (\alpha\mathbb{F} + (1-\alpha)h, \sigma) \succsim (\alpha\mathbb{G} + (1-\alpha)h, \sigma').$$

Axiom 10 reflects the intuitive assumption that information should have no impact on the agent's preference over acts. This is because the information causes no regret and generates no instrumental value when the agent is constrained to some $\{\{h\}\}$. Therefore, the pair $(\alpha\mathbb{F} + (1-\alpha)h, \sigma)$ can be interpreted as if the agent has chosen σ as her information structure but faces uncertainty about whether she will get $\mathbb{F}$ or $\{\{h\}\}$. In this sense, the motivation for Axiom 10 is the same as any standard independence axiom.

An information structure $\sigma \in \mathcal{I}$ is *null* (or *uninformative*) if it always induces a degenerate distribution over posteriors under any prior.

Let $o : \Omega \to \Delta(\{0\})$ denote a specific null information structure. It has only one signal 0 which is obtained with probability one in every state. Our next axiom disciplines the agent's behavior when the observed information is exactly o.

Axiom 11—Strategic Rationality under Null Information (SRNI): *For any menus* F, G, *if* $\{F\} \succsim^o \{G\}$, *then* $\{F \cup G\} \sim^o \{F\}$.

Axiom 11 is motivated by the fact that null information should cause no regret since there is no change in how the agent evaluate each option. Anticipating o, the agent evaluates any menu according to its material value under her prior belief. Axiom 11 is the standard strategic rationality requirement in this situation.

Suppose $\succsim^o$ has a SIT representation, Theorem 2 guarantees the identification of some information structure σ^o. We show that if $\succsim^o$ also satisfies Axiom 11 (SRNI), then the identified σ^o and o both induce a degenerate distribution over posteriors.

Lemma 1: *If* $\succsim^o$ *has a SIT representation* (π^o, u^o, K^o, i^o), *then* $\succsim^o$ *satisfies Axiom 11 if and only if* i^o *induces a degenerate distribution over posteriors.*

Proof. *"If."* If i^o induces a degenerate distribution over posteriors, then $\succsim^o$ is represented by $V^o(\mathbb{F}) = \max_{F \in \mathbb{F}} \max_{f \in F} \sum_{\omega \in \Omega} \pi(\omega) u(f(\omega))$. If $V^o(\{F\}) \geq V^o(\{G\})$, then $V^o(\{F \cup G\}) = V^o(\{F\})$. SRNI is thus satisfied.

"Only if." Suppose i^o induces a distribution over posteriors that contains at least two different elements, μ_1 and μ_2, in its support. Then by a standard result, there exist two

acts f and g such that f is strictly better than g in μ_1 and g is strictly better than f in μ_2. Therefore, $V^o(\{\{f,g\}\}) > \max\{V^o(\{\{f\}\}), V^o(\{\{g\}\})\}$, violating SRNI. $\square$

The experiment o is not the only null information structure. However, we only need to impose the SRNI axiom on $\succsim^o$ because the rest of our axioms will imply that Blackwell equivalent information structures induce the same conditional preference.

Let $\sigma : \Omega \to \Delta(S)$ be an information structure and F be a menu. A *plan* is a mapping $\gamma : S \to F$, where $\gamma(s)$ is the act chosen by the agent if she observes signal s. Intuitively, a plan describes the agent's reaction to the information. For example, if F is a singleton with $F = \{f\}$, then the unique plan involves choosing f after every possible signal. For another example, if $F = \{f, g\}$ and $S = \{s_1, s_2\}$, then the agent has four different plans. The first is to choose f after s_1 and choose g after s_2 (fg in short). The other three are ff, gf and gg.

Given an information structure $\sigma : \Omega \to \Delta(S)$ and a plan $\gamma \in F^S$, the *act induced by* σ *and* γ, denoted by γ_σ, is an act defined by $\gamma_\sigma(\omega) := \sum_{s\in S} \sigma(s \mid \omega)[\gamma(s)](\omega)$. That is, γ_σ is obtained by reducing the "compound act" described by γ and σ by averaging over different signals. To illustrate, if $F = \{f, g\}$ and $S = \{s_1, s_2\}$, the information structure $\sigma : \Omega \to \Delta(S)$ together with the plan fg describe the following compound act: Contingent on ω being the state of the world, the agent will obtain $f(\omega)$ (when the signal is s_1 and she chooses f according to the plan fg) with probability $\sigma(s_1 \mid \omega)$ and obtain $g(\omega)$ with probability $\sigma(s_2 \mid \omega)$. The induced act $(fg)_\sigma$ is obtained by the convex combination $\sigma(s_1 \mid \omega)f(\omega) + \sigma(s_2 \mid \omega)g(\omega)$.

Let F^S denote the set of all plans fixing F and S.

Definition 5: Fix a menu F and some information structure σ, *the menu induced by* F *and* σ is defined by $F_\sigma := \left\{\gamma_\sigma \mid \gamma \in F^S\right\}$.

That is, F_σ is obtained by collecting all the acts that can be induced by the information structure σ and some plan $\gamma \in F^S$. In particular, for singleton menus like $F = \{f\}$, the set of plans F^S is a singleton containing only one plan (choosing act f after any signal). Therefore, the induced menu is $F_\sigma = \{f\}$ for any information structure σ.

Axiom 12—Reduction: $(\{F\}, \sigma) \sim (\{F_\sigma\}, o)$ *for any* F *and any* σ.

Axiom 12 is motivated by the common notion that information carries only instrumental value (but no intrinsic value) when there is no regret. It states that the agent only cares about the state-contingent distributions over outcomes generated by the information structure.

An important implication of Axiom 12 is that $(f, \sigma) \sim (f, o)$ for any act f and information structure σ.[12] Together with the completeness and transitivity of $\succsim$, Axiom 12 implies that the agent has a stable preference over acts that is independent of the information structures.

[12]Similar as before, we abuse notation a little by writing (f, σ) for the pair $(\{\{f\}\}, \sigma)$.

Our last axiom concerns with mixtures of early-commitments $D(F)$ and their corresponding full-flexibilities $\{F\}$. As it turns out, their interactions have a clear implication on the regret intensity of an agent.

Axiom 13—Balance: *If* $(\{F\}, \sigma_1) \succ (\{F\}, o)$, $(\{G\}, \sigma_2) \succ (\{G\}, o)$, $\alpha \in (0, 1]$,

$$\Big(\alpha D(F) + (1-\alpha)\{F\}, \sigma_1\Big) \sim (\{F\}, o) \iff \Big(\alpha D(G) + (1-\alpha)\{G\}, \sigma_2\Big) \sim (\{G\}, o).$$

If the agent strictly prefers σ_1 to o when facing $\{F\}$, then the agent must strictly prefer o to σ_1 when facing $D(F)$ because the regret for $D(F)$ is proportionl to the instrumental value for $\{F\}$. The mixture $\alpha D(F) + (1-\alpha)\{F\}$ effectively describes a scenario where the agent needs to first choose an act from F, but has uncertainty about whether she has to stick to this choice (corresponding to $D(F)$) or she has the chance to re-optimize within F (corresponding to $\{F\}$). Axiom 13 implies that when facing such an uncertainty, there is a weight α that precisely balances out the expected regret generated from facing $D(F)$ and the expected benefit from facing $\{F\}$. At this balance point, the agent is indifferent between facing $\alpha D(F) + (1-\alpha)\{F\}$ with information σ and facing $\{F\}$ with no information at all.

Axiom 13 implies the existence of a unique scalar $\alpha^* \in (0, 1]$ such that

$$\Big(\alpha^* D(F) + (1-\alpha^*)\{F\}, \sigma\Big) \sim (\{F\}, o)$$

for all menu $F \in \mathcal{M}$ and information structure $\sigma \in \mathcal{I}$. Such a unique balance point α^* exists independent of the consumption choice F and the information choice σ because our interpretation for the IT representation involves a single parameter that represents the agent's regret intensity in every decision situation.

Theorem 3: *A binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ has an informational tradeoff representation if and only if: (1) $\succsim$ satisfies Weak Order, Act Independence, SRNI, Reduction and Balance; and (2) $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$.*

The proof of Theorem 3 is relegated to Appendix C.2.

4.1 Uniqueness of the IT Representation

As we have seen, non-variation between the agent's prior and posterior beliefs is a threat to identification. Since the modeler can observe the agent's information choice in the IT representation, such non-variation could only happen when the agent's prior belief is degenerate.

We say an agent *has a trivial preference over information* if $(\mathbb{F}, \sigma) \sim (\mathbb{F}, \sigma')$ for all $\mathbb{F} \in \mathcal{D}$ and $\sigma, \sigma' \in \mathcal{I}$. That is, the agent is indifferent between any two information structures regardless of the menu of menus. The agent *has a non-trivial preference over information* if she does not have a trivial preference over information.

LEMMA 2: *Suppose $\succsim$ has an IT representation with parameters (π, u, K). This preference $\succsim$ has a trivial preference for information if and only if the prior belief π is degenerate.*

Proof. "If." Suppose π is degenerate, that is, there exists exactly one $\omega \in \Omega$ such that $\pi(\omega) = 1$. Then the only possible posterior is $\mu_s^\sigma = \pi$ for any information structure σ and any signal realization $s \in S$. Hence $W(\mathbb{F}, \sigma) = W(\mathbb{F}, \sigma')$ for all $\mathbb{F} \in \mathcal{D}$ and $\sigma, \sigma' \in \mathcal{I}$, and $\succsim$ has a trivial preference for information.

"Only if." Suppose $\succsim$ has a trivial preference for information, and suppose by contradiction that π is non-degenerate. That is, there exists $\omega, \omega' \in \Omega$ such that $\pi(\omega), \pi(\omega') > 0$. Since the taste function u is non-constant. It follows from a standard result that we can construct two acts f, g such that the agent strictly prefers a fully informative experiment to a fully uninformative experiment when facing $\{\{f, g\}\}$. □

As long as the agent has a non-trivial preference for information, all parameters for the IT representation can be uniquely identified.

THEOREM 4: *Suppose $\succsim$ has an informational tradeoff representation (π, u, K) and $\succsim$ has a non-trivial preference for information. Then (π', u', K') also represents $\succsim$ if and only if $\pi' = \pi$, $u' = au + b$ for some $a > 0$ and $b \in \mathbb{R}$, and $K' = K$.*

Proof of Theorem 4. The "if" direction is straightforward.

"Only if." Suppose (π', u', K') also represents $\succsim$. Then (π, u) and (π', u') agree on their preference for acts. By the standard result from the Anscombe-Aumann framework, $\pi' = \pi$ and $u' = au + b$ for some $a > 0$ and $b \in \mathbb{R}$. Since $\succsim$ has a non-trivial preference for information, the prior π is non-degenerate (Lemma 2). Then the identification of the regret intensity K follows from Theorem 2. □

4.2 Comparing Information Aversion Attitudes

In this section, we focus our attention on agents with non-trivial preferences over information. Suppose $\succsim_1$, $\succsim_2$ each has an informational tradeoff representation with parameters (π_1, u_1, K_1) and (π_2, u_2, K_2), respectively.

DEFINITION 6: We say that $\succsim_1$ is *more information averse than* $\succsim_2$ if

$$(\mathbb{F}, \sigma) \succsim_2 (\mathbb{F}, \sigma') \Longrightarrow (\mathbb{F}, \sigma) \succsim_1 (\mathbb{F}, \sigma') \tag{9}$$

for any $\mathbb{F} \in \mathcal{D}$ and any $\sigma, \sigma' \in \mathcal{I}$ such that σ' is Blackwell more informative than σ.

That is, our notion of being "more information averse" requires that whenever agent 2 prefers a less informative experiment to a more informative one given a certain menu of menus, agent 1 will have the same avoidance given the same menu of menus.

We focus our attention on comparing the information aversion attitudes between agents who share the same preference over acts. Formally, we say that $\succsim_1$ and $\succsim_2$ *have the same preference over acts* if $(f,\sigma) \succsim_1 (g,\sigma) \iff (f,\sigma) \succsim_2 (g,\sigma)$ for every $f, g \in \mathcal{F}$.

THEOREM 5: *Suppose $\succsim_1$ and $\succsim_2$ have the same preference over acts, then $\succsim_1$ is more information averse than $\succsim_2$ if and only if $K_1 \geq K_2$.*

Proof of Theorem 5. Since $\succsim_1$ and $\succsim_2$ have the same preference over acts, $\pi_1 = \pi_2$, $u_2 = au_1 + b$ for some $a > 0$ and $b \in \mathbb{R}$. Therefore, (π_1, u_1, K_2) also represents $\succsim_2$. Let W_1, W_2 be the functionals representing $\succsim_1$ and $\succsim_2$, respectively.

"*If.*" Suppose $K_1 \geq K_2$. Fix any $\mathbb{F} \in \mathcal{D}$ and $\sigma, \sigma' \in \mathcal{I}$ such that $\sigma' \trianglerighteq \sigma$, we want to show

$$W_2(\mathbb{F},\sigma) - W_2(\mathbb{F},\sigma') \geq 0 \Longrightarrow W_1(\mathbb{F},\sigma) - W_1(\mathbb{F},\sigma') \geq 0 \tag{10}$$

It suffices to show that $\frac{W_1(\mathbb{F},\sigma)-W_1(\mathbb{F},\sigma')}{1+K_1} - \frac{W_2(\mathbb{F},\sigma)-W_2(\mathbb{F},\sigma')}{1+K_2} \geq 0$, which holds because

$$\begin{aligned}
&\frac{W_1(\mathbb{F},\sigma) - W_1(\mathbb{F},\sigma')}{1+K_1} - \frac{W_2(\mathbb{F},\sigma) - W_2(\mathbb{F},\sigma')}{1+K_2} \\
=& \left(\frac{W_1(\mathbb{F},\sigma)}{1+K_1} - \frac{W_2(\mathbb{F},\sigma)}{1+K_2}\right) - \left(\frac{W_1(\mathbb{F},\sigma')}{1+K_1} - \frac{W_2(\mathbb{F},\sigma')}{1+K_2}\right) \\
=& -\frac{K_1}{1+K_1}\sum_{s\in S}\sigma(s)\max_{f\in\mathbb{F}}\ U(f,\mu_s^\sigma) + \frac{K_2}{1+K_2}\sum_{s\in S}\sigma(s)\max_{f\in\mathbb{F}}\ U(f,\mu_s^\sigma) \\
&+ \frac{K_1}{1+K_1}\sum_{s'\in S'}\sigma'(s')\max_{f\in\mathbb{F}} U(f,\mu_{s'}^{\sigma'}) - \frac{K_2}{1+K_2}\sum_{s'\in S'}\sigma'(s')\max_{f\in\mathbb{F}}\ U(f,\mu_{s'}^{\sigma'}) \\
=& \left(\frac{K_1}{1+K_1} - \frac{K_2}{1+K_2}\right)\left[\sum_{s'\in S'}\sigma'(s')\max_{f\in\mathbb{F}}\ U(f,\mu_{s'}^{\sigma'}) - \sum_{s\in S}\sigma(s)\max_{f\in\mathbb{F}}\ U(f,\mu_s^\sigma)\right] \geq 0,
\end{aligned}$$

where the second equality follows from plugging in the equivalent expression of W and replace $\max_{B\in\mathbb{F}}\max_{f\in B}$ with $\max_{f\in\mathbb{F}}$, and the last inequality follows from $K_1 \geq K_2$ and that σ' is Blackwell more informative than σ.

"*Only if.*" Suppose $\succsim_1$ is more information averse than $\succsim_2$. Suppose by contradiction that $K_2 > K_1$, we want to construct some $\mathbb{F}$ such that $(\mathbb{F}, o) \succsim_2 (\mathbb{F}, i)$ but $(\mathbb{F}, i) \succ_1 (\mathbb{F}, o)$, where o is the null experiment and i is the fully informative experiment.

Since $\succsim_1$ and $\succsim_2$ both have non-trivial preference for information, $\pi_1 = \pi_2$ is not degenerate. Thus, we can find two states ω and ω' such that $\pi_1(\omega), \pi_1(\omega') > 0$. Let $\eta = \frac{\pi_1(\omega)}{\pi_1(\omega)+\pi_1(\omega')}$. Consider the following three acts $f, g, h \in \mathcal{F}$:

$$\begin{array}{c|ccc}
 & u\circ f & u\circ g & u\circ h \\
\hline
\omega & \frac{1}{\eta} & 0 & 1 \\
\omega' & 0 & y & 1 \\
\hat{\omega} \in \Omega\setminus\{\omega,\omega'\} & 0 & 0 & 0
\end{array} \tag{11}$$

where $y \in (\frac{K_1}{1+K_1}, \frac{K_2}{1+K_2})$. Let $\mathbb{F} = \{\{f, g\}, \{h\}\}$. Then

$$W_1(\mathbb{F}, o) = W_2(\mathbb{F}, o) = \max\left\{\pi \cdot \frac{1}{\eta}, (1-\eta)y, 1\right\} = 1 \tag{12}$$

$$W_1(\mathbb{F}, i) = 1 + (1-\eta)y - (1-\eta)K_1(1-y) > 1 \tag{13}$$

$$W_2(\mathbb{F}, i) = 1 + (1-\eta)y - (1-\eta)K_2(1-y) \leq 1 \tag{14}$$

Such an $\mathbb{F}$ can always be constructed when $K_2 > K_1$. This completes the proof. □

5 Discussion and Extension

We conclude by discussing some related literature and an extension.

5.1 Related Literature

We are not the first to model regret as the difference between what the agent actually gets from a certain choice and the counterfactual best outcome she could have got if she made a different choice.[13] Sarver (2008) is the first to unveil the possibility of identifying regret through preferences over menus. Sarver (2008) develops an axiomatic model in which an agent anticipates regret from the resolution of her subjective uncertainty regarding her taste over lotteries. His dominance axiom helps to differentiate regret from other motives for desiring a smaller menu, like temptation. Our paper is similar in spirit to capture regret, but we study preferences over a larger choice domain with an emphasis on the role of information. This larger choice domain also helps us to overcome the identification issue for the parameter of regret intensity in Sarver (2008). Buturak and Evren (2017) axiomatize a similar utility representation to explain choice overload where an agent tends to stick more to the default option when there are more options to choose from.

In both papers, the only relevant choice by the agent is made before their subjective uncertainty about their taste is resolved. Thus, this resolution is not beneficial for their future choices. Our model explicitly allows the co-existence of regret for past choices and benefit for future choices.

We are not the first to attempt to build a theoretical model that can account for information avoidance either. Golman, Hagmann, and Loewenstein (2017) provides a survey on information avoidance from multiple disciplines. Caplin and Leahy (2001) introduce a psychological state into the standard expected utility framework and build anticipatory feelings about the future (represented by a belief) directly into the agent's utility function. This

[13]Earlier examples include Bell (1982), Loomes and Sugden (1982, 1987) and Sugden (1993).

framework, even though non-axiomatic, is very general and allows the possibility for information avoidance driven by different anticipatory feelings. Kőszegi (2003) utilizes a similar framework with a focus on patient behavior and builds a model where patients avoid relevant medical information in order to avoid the anxiety from the anticipation of a potential bad outcome. Brunnermeier and Parker (2005) model economic agents who can optimally set their own beliefs facing uncertainty, and such an agent might want to avoid information that breaks their unwarranted optimistic beliefs. Dillenberger (2010) builds an axiomatic model that features a preference for one-shot over gradual resolution of uncertainty. As a result, the agent in his model might behave as if she is averse to some information if the information represents a partial resolution of uncertainty.

Our paper differs from all these works due to our focus on rationalizing the linkage between information avoidance and past decisions.

Finally, our work is broadly related to theories on cognitive dissonance (Festinger 1957) where past decisions could affect subsequent decisions. Yang (2024) axiomatically characterizes a multi-self model where an agent distorts her belief to reconcile with past decisions. Eyster et al. (2021) studies an agent who distort future decisions to justify past decisions. Our explanation for information avoidance (as a specific form of cognitive dissonance) does not involve any belief or choice distortion.

5.2 An Extension

One important property of the IT representation is that the agent only experiences regret once and the regret is only about her choice of menu. This assumption mainly helps us to cleanly isolate the tradeoff of the conflicting effects of information from other potential confounding factors. That said, our model can be made more realistic by allowing the agent to have regrets over her choice of act (after the state of the world is revealed), her choice of menu of menus, and/or her choice of information.

In this section, we first discuss a simple extension of our model by allowing the agent to regret her choice of act. We'll see that such a relaxation may not result in a change of the agent's preference for information. Therefore, the extended model does not necessarily deliver a different prediction on an agent's information avoidance behavior even though it causes the model to be much more complicated and less intuitive.

In addition, we will briefly discuss the reason for not entertaining the regrets for the choice of menu of menus and the choice of information.

Regretting the choice of act. Consider an agent who might also regret her choice of act in period 3. This regret could arise because the information arrived before the act choice only partially resolves the uncertainty about the state of the world. Recall that the agent's choice of act is based on a posterior belief, and this posterior belief is not degenerate in general. Therefore, it might be that the agent observes the true state of the world after she receives her payoff, and regret her choice of act if another act might be better given the realized state. To model this, consider the following variation of the IT representation

$$W_1(\mathbb{F},\sigma) := \max_{F\in\mathbb{F}} \sum_{s\in S} \sigma(s)\left[\widetilde{U}(F,\mathbb{F},\mu_s^\sigma) - R(F,\mathbb{F},\mu_s^\sigma)\right] \tag{15}$$

where $\sigma(s)$, μ_s^σ, and the regret term $R(F,\mathbb{F},\mu_s^\sigma)$ are defined as before. That is,

$$R(F,\mathbb{F},\mu_s^\sigma) = K_0\left[\max_{G\in\mathbb{F}}\max_{g\in G}\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)u\big(g(\omega)\big) - \max_{f\in F}\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)u\big(f(\omega)\big)\right]$$

where $K_0 \geq 0$ represents the regret intensity. However, the "material utility" of a menu F under a posterior belief μ_s^σ, denoted by $\widetilde{U}$, is now also dependent on the menu of menus $\mathbb{F}$ from which it is chosen. Formally,

$$\widetilde{U}(F,\mathbb{F},\mu_s^\sigma) := \max_{f\in F}\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)\left[u\big(f(\omega)\big) - \widetilde{R}(f,F,\mathbb{F},\omega)\right] \tag{16}$$

where $\widetilde{R}(f,F,\mathbb{F},\omega)$ is the regret for choosing f if the state turns out to be ω. Formally,

$$\widetilde{R}(f,F,\mathbb{F},\omega) := K_1\left[\max_{G\in\mathbb{F}}\max_{g\in G}u\big(g(\omega)\big) - u\big(f(\omega)\big)\right] + K_2\left[\max_{h\in F}u\big(h(\omega)\big) - u\big(f(\omega)\big)\right] \tag{17}$$

where the first term reflects the agent's regret toward the counterfactual outcome she could get if she can re-optimize her choice of menu, while the second term reflects the agent's regret toward the counterfactual outcome she could get if she can only re-optimize her choice of act from the chosen menu F. Parameters $K_1, K_2 \geq 0$ represent the respective regret intensities. Thus, utility W_1 in the form of equation (15) has five parameters, (π, u, K_0, K_1, K_2).

We are interested in comparing two agents' attitudes toward information. The first one is an agent who can be represented by an IT representation W with parameters (π, u, K), and the second one is an agent who can be represented by W_1 with parameters (π, u, K_0, K_1, K_2). The following lemma states that under some regularity conditions about the regret intensities, the possibility of experiencing about her act choice will not change an agent's preference for information given any menu of menus.

Lemma 3: *If representations W and W_1 share the same parameters (π, u) and the regret intensity parameters satisfy $K = \frac{K_0}{1+K_1}$ and $K_2 = 0$, then*

$$W(\mathbb{F},\sigma) \geq W(\mathbb{F},\sigma') \iff W_1(\mathbb{F},\sigma) \geq W_1(\mathbb{F},\sigma')$$

for any menu of menus $\mathbb{F}$ and information structures $\sigma, \sigma' \in \mathcal{I}$.

Proof. See Appendix C.5. □

When $K_2 = 0$, the agent's regret about her choice of act comes solely from the counterfactual comparison with what she could have got from the entire menu of menus $\mathbb{F}$. This part will depend only on her prior belief but not on her choice of information. And the relative regret intensity levels need to satisfy $K = K_0/(1 + K_1)$. The result hinges on these parametric assumptions. That is, if the relative strength of the regret is mismatched or if the agent's regret about her act choice also comes from the counterfactual comparison with what she could have got from the menu she has chosen, then her preference over information could change. The assumption that $K_2 = 0$ also reflects the subtlety we must face when considering regret in a multiple stage setup: We must be precise about what is the reference point (i.e., the counterfactual outcome) the agent is considering for her comparison.

Regretting the choice of menu of menus. It is also natural to think about a variation of the IT representation where the agent also regrets her period-1 choice (the choice of menu of menus). This would correspond to a student regretting her college choice (in addition to the major choice). Sarver (2008) provides a nice interpretation that could also apply to our IT representation. That is, our representations do not rule out the possibility for the agent to regret her period-1 choice. Instead, our representations can be interpreted as the *additional regret* from the menu choice *on top of* the regret from the choice over menu of menus. We refer interested readers to section 4.3 of Sarver (2008) for more details.

Regretting the choice of information. In an IT representation, there is no regret associated with the information choice. We believe this is a reasonable first step to take when modeling explicit preference over information. It is conceivable that a myopic agent might feel a sense of regret in the future for make "bad" decisions due to the lack of information. However, our information preference is over Blackwell experiments, not their signal realizations. Thus, there seems to be no obvious answer on if or how one should model such a regret in information choice. We believe this could be an intriguing avenue for future research.

A The Framework with Lotteries

Let Z be a finite set of outcomes. Let $\Delta(Z)$ denote the set of lotteries over Z, with typical elements p, q. Endow $\Delta(Z)$ with the standard Euclidean metric. A menu is a nonempty compact subset of $\Delta(Z)$. Let $\widehat{\mathcal{M}}$ denote the set of all menus, with typical elements A, B. Endow $\widehat{\mathcal{M}}$ with the Hausdorff metric. A menu of menus is a nonempty compact subset of $\widehat{\mathcal{M}}$. Let $\widehat{\mathcal{D}}$ denote the set of all menus of menus, with typical elements $\mathbb{A}, \mathbb{B}$. Endow $\widehat{\mathcal{D}}$ with the Hausdorff metric.

Definition 7: Let $\{U_1, \ldots, U_m\}$ be a collection of continuous linear functions from $\widehat{\mathcal{M}}$ to $\mathbb{R}$. We say this collection is *redundant* if there exists U_i that is constant or if there exists $i \neq j$ such that $U_j = \alpha U_i + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$. We say a collection is *non-redundant* if it is not redundant.

Similarly, if $\{u_1, \ldots, u_n\}$ is a collection of expected utility functions over $\Delta(Z)$, then we say this collection is *redundant* if there exists u_i that is constant or if there exists $i \neq j$ such that $u_j = \alpha u_i + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$. We take the convention that an empty collection is not redundant. Let $\mathcal{V} := \left\{v \in \mathbb{R}^Z \mid \sum_{z \in Z} v_z = 0\right\}$ denote the set of normalized expected utilities over $\Delta(Z)$. Let $\mathcal{U} := \left\{u \in \mathbb{R}^Z : \sum_{z \in Z} u_z = 0, \sum_{z \in Z} u_z^2 = 1\right\}$ denote the set of doubly normalized expected-utility functions on $\Delta(Z)$.

Lemma 4: *Let $\{u_i\}_{i \in I}$ be a collection of normalized expected utility functions over $\Delta(Z)$, and let $U : \widehat{\mathcal{M}} \to \mathbb{R}$ be defined by $U(A) := \sum_{i \in I} \max_{p \in A} \ u_i(p)$. Then, $U(A) = 0$ for all $A \in \widehat{\mathcal{M}}$ if and only if $u_i(p) = 0$ for all $i \in I$ and all $p \in \Delta(Z)$.*

Proof of Lemma 4. The "if" part is straightforward. For the "only if" part, let $J \subseteq I$ be a maximal collection of non-redundant subset of I. That is, for any $i \in I \setminus J$, u_i is constant or there exists some $j \in J$ such that $u_i = \alpha u_j$ for some $\alpha > 0$. It suffices to show that $U(A) \equiv 0$ implies $J = \emptyset$. Suppose by contradiction that $|J| = 1$ with $J = \{u_j\}$. Then $U(A) = L \max_{p \in A} u_j(p)$ with some $L > 0$. Thus, $u_j(p) = \frac{U(\{p\})}{L} = 0$ for all $p \in \Delta(Z)$, contradiction. Suppose by contradiction that $|J| \geq 2$, then by a standard result (Lemma A.1 of Kopylov 2009), there exists $A = \{p_j\}_{j \in J}$ such that p_j is the unique maximizer of u_j in A. Fix any $k \in J$, $U(A) = \sum_{j \in J} L_j u_j(p_j) > \sum_{j \in J} L_j u_j(p_k) = U(\{p_k\}) = 0$, contradiction. $\square$

A.1 The Partial Regret Representation

Definition 8: A binary relation $\succsim$ over $\widehat{\mathcal{D}}$ has a *partial regret (PR) representation* if there exists a finitely-supported probability measure μ over $\mathcal{U}$ and a scalar $K \geq 0$ such that $\succsim$ is represented by the function $\widehat{V} : \widehat{\mathcal{D}} \to \mathbb{R}$ defined by

$$\widehat{V}(\mathbb{A}) = \max_{A \in \mathbb{A}} \sum_{u \in \operatorname{supp}(\mu)} \mu(u) \left[\max_{p \in A} u(p) - R(A, \mathbb{A}, u)\right] \tag{18}$$

where $R(A, \mathbb{A}, u) := K \left[\max_{B \in \mathbb{A}} \max_{p \in B} u(p) - \max_{p \in A} u(p)\right].$

A useful equivalent expression for the PR representation is

$$\widehat{V}(\mathbb{A}) = (1 + K) \max_{A \in \mathbb{A}} \sum_{u \in \operatorname{supp}(\mu)} \mu(u) \max_{p \in A} u(p) - K \sum_{u \in \operatorname{supp}(\mu)} \mu(u) \max_{B \in \mathbb{A}} \max_{q \in B} u(p)$$

Following are the axioms for the PR representation.

Axiom A.1—Weak Order: *$\succsim$ is complete and transitive.*

Axiom A.2—Continuity: *For any $\mathbb{A}$, the sets $\{\mathbb{B} : \mathbb{B} \succsim \mathbb{A}\}$, $\{\mathbb{B} : \mathbb{A} \succsim \mathbb{B}\}$ are closed.*

Axiom A.3—Independence: *If $\mathbb{A} \succ \mathbb{B}$, then $\alpha\mathbb{A} + (1-\alpha)\mathbb{C} \succ \alpha\mathbb{B} + (1-\alpha)\mathbb{C}$ for any $\mathbb{C}$ and $\alpha \in (0, 1]$.*

Definition 9: Let $\mathbb{A}$ be a menu of menus. Say that $\mathbb{B}$ is *critical for* $\mathbb{A}$ if $\mathbb{B} \subseteq \mathbb{A}$ and $\mathbb{B}' \sim A$ for all B' satisfying $\mathbb{B} \subseteq \mathbb{B}' \subseteq \mathbb{A}$. Say that B *is critical for A in* $\mathbb{A}$ if $B \subseteq A \in \mathbb{A}$ and $(\mathbb{A} \setminus \{A\}) \cup \{B'\} \sim \mathbb{A}$ for all B' satisfying $B \subseteq B' \subseteq A$.

Axiom A.4—Finiteness: *There exists a natural number N such that: (4.1) For every $\mathbb{A}$, there exists $\mathbb{B}$ with $|\mathbb{B}| < N$ such that $\mathbb{B}$ is critical for $\mathbb{A}$; and (4.2) For every $\mathbb{A} \in \mathcal{D}$ and every $A \in \mathbb{A}$, there exists B with $|B| < N$ such that B is critical for A in $\mathbb{A}$.*

Axiom A.5—Ex-Ante Regret: *If $\{A\} \succsim \{B\}$ and $A \in \mathbb{A}$, then $\mathbb{A} \succsim \mathbb{A} \cup \{B\}$.*

Axiom A.6—Interim Preference for Flexibility: *For any menu of menus $\mathbb{A}$ and menus A, B, $\mathbb{A} \cup \{A \cup B\} \succsim \mathbb{A} \cup \{A, B\}$.*

Axiom A.7—Inclusion: *If $B \subseteq A$ and $A \in \mathbb{A}$, then $\mathbb{A} \cup \{B\} \succsim \mathbb{A}$.*

Axiom A.8—Nontriviality: *There exists $\mathbb{A}$ and $\mathbb{B}$ such that $\mathbb{B} \subseteq \mathbb{A}$ with $\mathbb{A} \succ \mathbb{B}$.*

Axioms A.1-A.5 corresponds to Axioms 1-5, respectively. The inclusion axiom (Axiom A.7) corresponds to a weakened version of the domination axiom (Axiom 8), and the non-triviality axiom (Axiom A.8) is stated slightly different than Axiom 7.

Theorem 6: *A binary relation $\succsim$ over $\widehat{\mathcal{D}}$ has a partial regret representation if and only if it satisfies Axioms A.1-A.8.*

A.2 Proof of Theorem 6

We start by establishing a nested DLR representation from Axioms A.1-A.4. DLR refers to Dekel, Lipman, and Rustichini (2009).

Definition 10: A binary relation $\succsim$ over $\widehat{\mathcal{D}}$ has a *nested DLR representation* if $\succsim$ can be represented by $V_{DLR}(\mathbb{A}) = \sum_{i \in I^+} \max_{A \in \mathbb{A}} U_i(A) - \sum_{i \in I^-} \max_{A \in \mathbb{A}} U_i(A)$ where I^+ is the index set for the positive states and I^- is the index set for the negative states and $\{U_i\}_{i \in I^+ \cup I^-}$ is a non-redundant (it is without loss to assume $I^+ \cap I^- = \emptyset$). Let $I := I^+ \cup I^-$. For each $i \in I$, $U_i : \widehat{\mathcal{M}} \to \mathbb{R}$ has a DLR representation $U_i(A) = \sum_{k \in P_i} \max_{p \in A} w_{ik}(p) - \sum_{j \in N_i} \max_{p \in A} v_{ij}(p)$ where P_i is the index set for the positive substates for U_i and N_i is the index set of negative substates for U_i (it is without loss to assume $P_i \cap N_i = \emptyset$) and $\{w_k\}_{k \in P_i} \cup \{v_j\}_{j \in N_i}$ is a non-redundant collection of normalized expected utilities over $\Delta(Z)$.

LEMMA 5: *A binary relation* $\succsim$ *over* $\widehat{\mathcal{D}}$ *has a nested DLR representation if and only if it satisfies Axioms A.1-A.4.*

Lemma 5 follows from Theorem 5 of Stovall (2018).

The rest of the proof mainly focuses on fine-tuning the states and substates in the nested DLR representation through the other axioms.

AXIOM A.9—Strong Ex-Ante Regret: *If for any* $B \in \mathbb{B}$*, there exists* $A \in \mathbb{A}$ *such that* $\{A\} \succsim \{B\}$*, then* $\mathbb{A} \succsim \mathbb{A} \cup \mathbb{B}$*.*

Axioms A.5 and A.9 are equivalent when Axioms A.1 and A.2 are satisfied.

LEMMA 6: *Suppose* $\succsim$ *satisfies Axioms A.1 (Weak Order) and A.2 (Continuity), then it satisfies Axiom A.5 if and only if it satisfies Axiom A.9.*

Proof. The "*if*" part is straightforward.

"*Only if.*" Suppose $\succsim$ satisfies Axiom A.5, and suppose two menus of menus of lotteries $\mathbb{A}$ and $\mathbb{B}$ are such that for all $B \in \mathbb{B}$, there exists $A \in \mathbb{A}$ such that $\{A\} \succsim \{B\}$.

First note that $\widehat{\mathcal{M}}$ is compact with the Hausdorff metric. Therefore, $\widehat{\mathcal{M}}$ is separable. We can thus choose a countable dense subset of $\mathbb{B}$, say $\mathbb{B}^* = \{B_i \mid i \in \mathbb{N}\}$. Define a sequence of menus of menus $(\mathbb{B}_n)_{n\in\mathbb{N}}$ by $\mathbb{B}_n = \{B_1, B_2, \ldots, B_n\}$. Then $\mathbb{B}_n \subseteq \mathbb{B}_{n+1}$ for all n and $\mathbb{B}^* = \bigcup_{n=1}^{\infty} \mathbb{B}_n$. By a standard result, $(\mathbb{B}_n)_{n\in\mathbb{N}}$ converges to the closure of $\mathbb{B}^*$, that is, $\mathbb{B}_n \to \overline{\mathbb{B}^*} = \mathbb{B}$ as n goes to infinity. Note that $B_1 \in \mathbb{B}^* \subseteq \mathbb{B}$, so there exists $A_1 \in \mathbb{A}$ such that $\{A_1\} \succsim \{B_1\}$ by assumption, and Axiom A.5 implies that $\mathbb{A} \succsim \mathbb{A} \cup \{B_1\}$. Let $\mathbb{A}_1 = \mathbb{A} \cup \{B_1\} = \mathbb{A} \cup \mathbb{B}_1$. Similarly, $B_2 \in \mathbb{B}^* \subseteq \mathbb{B}$, so there exists $A_2 \in \mathbb{A}$ such that $\{A_2\} \succsim \{B_2\}$, and Axiom A.5 implies that $\mathbb{A}_1 \succsim \mathbb{A}_1 \cup \{B_2\}$. Let $\mathbb{A}_2 = \mathbb{A}_1 \cup \{B_2\} = \mathbb{A} \cup \mathbb{B}_2$. By transitivity, $\mathbb{A} \succsim \mathbb{A}_1$ and $\mathbb{A}_1 \succsim \mathbb{A}_2$ imply that $\mathbb{A} \succsim \mathbb{A}_2 = \mathbb{A} \cup \mathbb{B}_2$. By an induction argument, $\mathbb{A} \succsim \mathbb{A} \cup \mathbb{B}_n$ for all n. Therefore, $\mathbb{A} \cup \mathbb{B}_n$ is in the lower contour set of $\mathbb{A}$ for all n, that is, $\mathbb{A} \cup \mathbb{B}_n \in \{\mathbb{C} \in \widehat{\mathcal{D}} \mid \mathbb{A} \succsim \mathbb{C}\}$. Since $\succsim$ satisfies Axiom A.2 (Continuity), this lower contour set is closed, $\mathbb{A} \succsim \lim_{n\to\infty} \mathbb{A} \cup \mathbb{B}_n = \mathbb{A} \cup \mathbb{B}$. □

LEMMA 7: *A binary relation* $\succsim$ *over* $\widehat{\mathcal{D}}$ *satisfies Axioms A.1-A.4 and Axiom A.9 if and only if the following two conditions hold:*

1. $\succsim$ *has a nested DLR representation with at most one positive state; and*

2. *Let* $\widehat{v}(A) := V_{DLR}(\{A\}) = \sum_{i\in I^+} U_i(A) - \sum_{i\in I^-} U_i(A)$*. There exists a non-negative scalar* $\alpha \geq 0$ *such that* $\sum_{i\in I^+} U_i(A) = \alpha\widehat{v}(A)$ *for all* $A \in \widehat{\mathcal{M}}$*.*

Proof of Lemma 7. "If." Since $\succsim$ has a nested DLR representation, it satisfies Axioms A.1-A.4 (by Lemma 5). If $|I^+| = 0$, that is, $I^+ = \emptyset$, then there is no positive state. Thus, $\mathbb{A} \succsim \mathbb{A} \cup \mathbb{B}$ for any two menus of menus $\mathbb{A}, \mathbb{B} \in \widehat{\mathcal{D}}$. Axiom A.9 is satisfied.

If $|I^+| = 1$, then $V_{DLR}(\mathbb{A}) = \max_{A\in\mathbb{A}} U_0(A) - \sum_{i\in I^-} \max_{A\in\mathbb{A}} U_i(A)$, where $\{U_0\}\cup\{U_i\}_{i\in I^-}$ is non-redundant. Suppose for any $B \in \mathbb{B}$, there exists $A \in \mathbb{A}$ such that $\widehat{v}(A) \geq \widehat{v}(B)$. Condition 2 implies that $\alpha > 0$, and $U_0(A) = \alpha[U_0(A) - \sum_{i\in I^-} U_i(A)]$ for all $A \in \widehat{M}$ where $U_0(A) - \sum_{i\in I^-} U_i(A) = \widehat{v}(A) = V_{DLR}(\{A\})$. Therefore, condition 2 implies that for any $B \in \mathbb{B}$, there exists $A \in \mathbb{A}$ such that $U_0(A) \geq U_0(B)$, which further indicates that $\max_{A\in\mathbb{A}\cup\mathbb{B}} U_0(A) = \max_{A\in\mathbb{A}} U_0(A)$. Since $\mathbb{A}\cup\mathbb{B} \supseteq \mathbb{A}$, $\max_{A\in\mathbb{A}\cup\mathbb{B}} U_i(A) \geq \max_{A\in\mathbb{A}} U_i(A)$ for any $i \in I^-$. Thus, $V_{DLR}(\mathbb{A}) \geq V_{DLR}(\mathbb{A}\cup\mathbb{B})$, and $\mathbb{A} \succsim \mathbb{A}\cup\mathbb{B}$. Axiom A.9 is satisfied.

"Only if." By Lemma 5, Axioms A.1-A.4 imply the existence of a nested DLR representation for $\succsim$ as in Definition 10. If $|I^+| = 0$, then $I^+ = \emptyset$ and there is no positive state, condition 1 is satisfied. And we can set $\alpha = 0$ so that condition 2 is satisfied.

If $|I^+| > 0$, then $I^+ \neq \emptyset$ and there is at least one positive state. It suffices to show that: If Axiom A.9 is satisfied, then for any positive state $i \in I^+$, there exists $\alpha > 0$ such that $U_i(A) = \alpha\widehat{v}(A)$. Suppose by contradiction that for some $i^* \in I^+$, U_{i^*} does not represent the same preference over $\widehat{\mathcal{M}}$ as $\widehat{v}$. Consider three cases:

Case 1: $\widehat{v}$ represents a trivial preference over $\widehat{\mathcal{M}}$. That is, $\widehat{v}(A) = 0$ for all $A \in \widehat{\mathcal{M}}$. Then there at least another state in $I^+\cup I^-$ that is different from i^* (otherwise $\widehat{v}(A) = U_{i^*}(A)$, contradiction). So there are at least two elements in $\{U_k\}_{k\in I^+\cup I^-}$. A standard result (e.g., Lemma A.1 of Kopylov, 2009) guarantees the existence of a collection of menus $\mathbb{A} := \{A_k\}_{k\in I^+\cup I^-}$ such that A_k is the unique maximizer of U_k in $\mathbb{A}$ for each $k \in I^+\cup I^-$, and $|\mathbb{A}| \geq 2$. Let $\mathbb{A}_1 := \mathbb{A}\setminus\{A_{i^*}\}$ and $\mathbb{B}_1 := \{A_{i^*}\}$. Then $\mathbb{A}_1 \neq \emptyset$ and for any $B \in \mathbb{B}_1$, there exists $A \in \mathbb{A}_1$ such that $\{A\} \succsim \{B\}$ (since $\widehat{v}(A) = 0 = \widehat{v}(B)$ for any $A, B \in \widehat{\mathcal{M}}$). Moreover, $\max_{A\in\mathbb{A}_1\cup\mathbb{B}_1} U_{i^*}(A) > \max_{A\in\mathbb{A}_1} U_{i^*}(A)$ and $\max_{A\in\mathbb{A}_1\cup\mathbb{B}_1} U_k(A) = \max_{A\in\mathbb{A}_1} U_k(A)$ for any $k \neq i^*$. Therefore, $V_{DLR}(\mathbb{A}_1\cup\mathbb{B}_1) > V_{DLR}(\mathbb{A}_1)$, indicating $\mathbb{A}_1\cup\mathbb{B}_1 \succ \mathbb{A}_1$, violating Axiom A.9. Contradiction.

Case 2: $\{\widehat{v}\}\cup\{U_k\}_{k\in I^+\cup I^-}$ is non-redundant. There are at least two elements ($\widehat{v}$ and U_{i^*}) in this collection, a standard result (e.g., Lemma A.1 of Kopylov, 2009) guarantees the existence of a collection of menus $\mathbb{A}' := \{B\}\cup\{A_k\}_{k\in I^+\cup I^-}$ such that B is the unique maximizer of $\widehat{v}$ in $\mathbb{A}$ and A_k is the unique maximizer of U_k in $\mathbb{A}$ for each $k \in I^+\cup I^-$. Consider $\mathbb{A}'_1 := \mathbb{A}'\setminus\{A_{i^*}\}$ and $\mathbb{B}'_1 := \{A_{i^*}\}$. Then $\mathbb{A}'_1 \neq \emptyset$ and $\mathbb{A}'_1 \geq \mathbb{B}'_1$ (since $B \in \mathbb{A}'_1$ and $\widehat{v}(B) > \widehat{v}(A_{i^*})$). Moreover, $\max_{A\in\mathbb{A}'_1\cup\mathbb{B}'_1} U_{i^*}(A) > \max_{A\in\mathbb{A}'_1} U_{i^*}(A)$ and $\max_{A\in\mathbb{A}'_1\cup\mathbb{B}'_1} U_k(A) = \max_{A\in\mathbb{A}'_1} U_k(A)$ for any $k \neq i^*$. Therefore, $V_{DLR}(\mathbb{A}'_1\cup\mathbb{B}'_1) > V_{DLR}(\mathbb{A}'_1)$, indicating $\mathbb{A}'_1\cup\mathbb{B}'_1 \succ \mathbb{A}'_1$, violating Axiom A.9. Contradiction.

Case 3: There exists some $j \in I^+\cup I^-$ such that $j \neq i^*$ but $\widehat{v}$ represents the same preference over $\widehat{\mathcal{M}}$ as U_j. Then $\{\widehat{v}\}\cup\{U_k\}_{k\in I^+\cup I^-\setminus\{j\}}$ is non-redundant. Similar arguments as in Case 2 above will lead to a violation of Axiom A.9. □

LEMMA 8: *Let $\{U_i\}_{i\in I}$ be a finite and non-redundant collection of continuous linear functions. If each $U_i : \widehat{\mathcal{M}} \to \mathbb{R}$ has a minimal finite DLR representation $U_i(A) = \sum_{k\in P_i} \max_{p\in A} w_{ik}(p) - \sum_{j\in N_i} \max_{p\in A} v_{ij}(p)$ where for each $i \in I$, $\{w_{ik}\}_{k\in P_i} \cup \{v_{ij}\}_{j\in N_i}$ is a non-redundant collection of normalized expected utilities on $\Delta(Z)$, then there exists a collection of menus $\{A_i\}_{i\in I}$ (all in the interior of $\Delta(Z)$) such that:*

1. *$U_i(A_i) > U_i(A_j)$ for any $i \in I$ and any $j \neq i$.*
2. *$\left|\arg\max_{p\in A_i} w_{ik}(p)\right| = 1$, $\left|\arg\max_{q\in A_i} v_{ij}(q)\right| = 1$ for any $i \in I$, any $k \in P_i$ and any $j \in N_i$. For each i, these maximizers are all distinct from each other.*

Proof. See the proof of Lemma 3 of Stovall (2018). $\square$

The next lemma states that imposing the Interim Preference for Flexibility axiom is equivalent to requiring there to be no negative substates in any positive state and at most one positive substate in any negative state. Formally,

LEMMA 9: *Suppose $\succsim$ has a nested DLR representation as in Definition 10, then $\succsim$ satisfies Axiom A.6 (Interim Preference for Flexibility) if and only if $|N_i| = 0$ for all $i \in I^+$ and $|P_i| \leq 1$ for all $i \in I^-$.*

Proof of Lemma 9. "If." For any $i \in I^+$, $U_i(A \cup B) \geq \max\{U_i(A), U_i(B)\}$ since $|N_i| = 0$. Therefore, the positive terms in $V_{DLR}(\mathbb{A} \cup \{A \cup B\})$ is weakly larger than the positive terms in $V_{DLR}(\mathbb{A} \cup \{A, B\})$.

For any negative state $i \in I^-$: If $|P_i| = 0$, then $U_i(C) = -\sum_{j\in N_i} \max_{p\in C} v_{ij}(p)$, and $-U_i(A \cup B) \geq \max\{-U_i(A), -U_i(B)\}$. If $|P_i| = 1$, then $U_i(C) = \max_{p\in C} w_{i0}(p) - \sum_{j\in N_i} \max_{p\in C} v_{ij}(p)$. Since $\max_{p\in A\cup B} w_{i0}(p) = \max\{\max_{p\in A} w_{i0}(p), \max_{p\in B} w_{i0}(p)\}$, we can again conclude that $-U_i(A \cup B) \geq \max\{-U_i(A), -U_i(B)\}$. Therefore, the negative terms in $V_{DLR}(\mathbb{A} \cup \{A \cup B\})$ is weakly larger than that in $V_{DLR}(\mathbb{A} \cup \{A, B\})$.

"Only if." By assumption, $\succsim$ can be represented by $V_{DLR}(\mathbb{A}) = \sum_{i\in I^+} \max_{A\in\mathbb{A}} U_i(A) - \sum_{i\in I^-} \max_{A\in\mathbb{A}} U_i(A)$ where $\{U_i\}_{i\in I^+\cup I^-}$ is a non-redundant collection of continuous linear functions. Let $I := I^+ \cup I^-$. For each $i \in I$, $U_i(A) = \sum_{k\in P_i} \max_{p\in A} w_{ik}(p) - \sum_{j\in N_i} \max_{p\in A} v_{ij}(p)$, where $\{w_{ik}\}_{k\in P_i} \cup \{v_{ij}\}_{j\in N_i}$ is a non-redundant collection of normalized expected-utilities on $\Delta(Z)$. Thus, all assumptions of Lemma 8 are satisfied. Therefore, there exists a collection of menus $\{A_i\}_{i\in I}$ in the interior of $\Delta(Z)$ such that both conditions 1 and 2 in Lemma 8 are satisfied.

Let $\mathbb{A} = \{A_i\}_{i\in I}$. We first show that $|P_i| \leq 1$ for all $i \in I^-$. Suppose by contradiction that $|P_{i^*}| \geq 2$ for some $i^* \in I^-$ (i.e., U_{i^*} has two or more positive substates). We want to find A, B such that $V_{DLR}(\mathbb{A} \cup \{A \cup B\}) < V_{DLR}(\mathbb{A} \cup \{A, B\})$.

For ease of exposition, we write U_* for U_{i^*} and A_* for A_{i^*}. Similarly, we write P_* for P_{i^*} (the index set of positive substates in i^*), N_* for N_{i^*}, w_{*k} for w_{i^*k} for any $k \in P_*$ and v_{*j} for v_{i^*j} for any $j \in N_*$. For $a, b \in \mathbb{R}$, write $a \vee b$ to denote $\max\{a, b\}$.

Let $p_k := \arg\max_{p \in A_*} w_{*k}(p)$. Since $|P_*| \geq 2$, we can pick k and k' with $k \neq k'$ from P_*. For any $\varepsilon > 0$, define $A^\varepsilon := A_* \cup \{p_k + \varepsilon w_{*k}\}$ and $B^\varepsilon := A_* \cup \{p_{k'} + \varepsilon w_{*k'}\}$. Since w_{*k} is normalized and p_k is in the interior of $\Delta(Z)$, a standard result guarantees the existence of an ε small enough so that $p_k + \varepsilon w_{*k}$ is also in the interior of $\Delta(Z)$.

Note that for any $\varepsilon > 0$, $w_{*k}(p_k + \varepsilon w_{*k}) = w_{*k}(p_k) + \varepsilon \|w_{*k}\|^2 > w_{*k}(p_k)$. Therefore, $\max_{p \in A^\varepsilon} w_{*k}(p) = w_{*k}(p_k) \vee w_{*k}(p_k + \varepsilon w_{*k}) = w_{*k}(p_k) + \varepsilon \|w_{*k}\|^2$, and

$$\max_{p \in B^\varepsilon} w_{*k}(p) = w_{*k}(p_k) \vee \Big(w_{*k}(p_{k'}) + \varepsilon w_{*k}(w_{*k'})\Big)$$

Since $w_{*k}(p_{k'}) < w_{*k}(p_k)$, there exists $\varepsilon_1 > 0$ with $w_{*k}(p_{k'}) + \varepsilon_1 w_{*k}(w_{*k'}) < w_{*k}(p_k)$, which further indicates that $\max_{p \in B^{\varepsilon_1}} w_{*k}(p) = w_{*k}(p_k)$. With similar arguments, we could find some $\varepsilon_2 > 0$ such that $\max_{p \in B^{\varepsilon_2}} w_{*k'}(p) > w_{*k'}(p_{k'}) = \max_{p \in A^{\varepsilon_2}} w_{*k'}(p)$. Let $\varepsilon_3 := \min\{\varepsilon_1, \varepsilon_2\}$,

$$\max_{p \in A^{\varepsilon_3}} w_{*k}(p) > w_{*k}(p_k) = \max_{p \in B^{\varepsilon_3}} w_{*k}(p)$$
$$\max_{p \in A^{\varepsilon_3}} w_{*k'}(p) = w_{*k'}(p_{k'}) < \max_{p \in B^{\varepsilon_3}} w_{*k'}(p).$$

Note that $U_*(A^{\varepsilon_3} \cup B^{\varepsilon_3}) = \sum_{k \in P_*} \max_{p \in A^{\varepsilon_3} \cup B^{\varepsilon_3}} w_{*k}(p) - \sum_{j \in N_*} \max_{p \in A^{\varepsilon_3} \cup B^{\varepsilon_3}} v_{*j}(p)$. Moreover, we have $U_*(A^{\varepsilon_3}) = \sum_{k \in P_*} \max_{p \in A^{\varepsilon_3}} w_{*k}(p) - \sum_{j \in N_*} \max_{p \in A^{\varepsilon_3}} v_{*j}(p)$ and $U_*(B^{\varepsilon_3}) = \sum_{k \in P_*} \max_{p \in B^{\varepsilon_3}} w_{*k}(p) - \sum_{j \in N_*} \max_{p \in B^{\varepsilon_3}} v_{*j}(p)$. For each $j \in N_*$,

$$\max_{p \in A^{\varepsilon_3} \cup B^{\varepsilon_3}} v_{*j}(p) = \max_{p \in A_1} v_{*j}(p) \vee v_{*j}(p_k + \varepsilon_3 w_{*k}) \vee v_{*j}(p_{k'} + \varepsilon_3 w_{*k'}).$$

Since A_* is finite and v_{*j} has a unique maximizer in A_* that is not p_k or $p_{k'}$, we have a small number $\varepsilon_j > 0$ with $\varepsilon_j < \varepsilon_3$ such that

$$\max_{p \in A^{\varepsilon_j} \cup B^{\varepsilon_j}} v_{*j}(p) = \max_{p \in A^{\varepsilon_j}} v_{*j}(p) = \max_{p \in B^{\varepsilon_j}} v_{*j}(p) = \max_{p \in A_*} v_{*j}(p).$$

Let $\varepsilon := \min_{j \in N_*} \varepsilon_j$, then $U_*(A^\varepsilon \cup B^\varepsilon) > \max\{U_*(A_*), U_*(A^\varepsilon), U_*(B^\varepsilon)\}$ because: (i) they all have the same negative term, and (ii) $A^\varepsilon \cup B^\varepsilon$ must have a strictly larger positive term because A^ε contains a unique maximizer $p_k + \varepsilon w_{*k}$ in substate k and B^ε contains a unique maximizer $p_{k'} + \varepsilon w_{*k'}$ in substate k'.

Finally, we can further scale down ε to guarantee that for any $i \in I$ with $i \neq i^*$,

$$U_i(A_i) > \max\{U_i(A^\varepsilon \cup B^\varepsilon), U_i(A^\varepsilon), U_i(B^\varepsilon)\}$$

because: (i) A_i is the unique maximizer of U_i in $\mathbb{A}$ and (ii) $A^\varepsilon \cup B^\varepsilon$, A^ε and B^ε can all be made arbitrarily close to A_*.

We will then show that $|N_i| = 0$ for any $i \in I^+$. Suppose by contradiction that $|N_{i^*}| \geq 1$ for some $i^* \in I^+$. Since $|N_*| \geq 1$, we can fix some $j \in N_*$. Let $q_j := \arg\max_{p \in A_*} v_{*j}(p)$. For any $\varepsilon > 0$, define $A^\varepsilon := A_*$ and $B^\varepsilon = (A_* \setminus \{q_j\}) \cup \{q_j - \varepsilon v_{*j}\}$. We can find ε small enough such that $q_j - \varepsilon v_{*j}$ is in the interior of $\Delta(Z)$. For any such ε, $v_{*j}(q_j - \varepsilon v_{*j}) = v_{*j}(q_j) - \varepsilon \|v_{*j}\|^2 < v_{*j}(q_j) = \max_{p \in A_*} v_{*j}(p)$. We can make ε small enough so that $v_{*j}(q_j - \varepsilon v_{*j}) = v_{*j}(q_j) - \varepsilon \|v_{*j}\|^2 > \max_{p \in A_* \setminus \{q_j\}} v_{*j}(p)$, and

$$v_{*j'}(q_j - \varepsilon v_{*j}) = v_{*j'}(q_j) - \varepsilon v_{*j'}(v_{*j}) < \max_{p \in A_*} v_{*j'}(p), \ \forall j' \in N_* \text{ and } j' \neq j$$

$$w_{*k}(q_j - \varepsilon v_{*j}) = w_{*k}(q_j) - \varepsilon w_{*k}(v_{*j}) < \max_{p \in A_*} w_{*k}(p), \ \forall k \in P_*$$

For such an ε, we have $U_*(B^\varepsilon) > U_*(A^\varepsilon) = U_*(A_*) = U_*(A^\varepsilon \cup B^\varepsilon)$. We can further scale down ε so that $U_i(A_i) > U_i(A^\varepsilon), U_i(B^\varepsilon)$ for any $i \in I$ and $i \neq i^*$. Then $V_{DLR}(\mathbb{A} \cup \{A^\varepsilon, B^\varepsilon\}) > V_{DLR}(\mathbb{A} \cup \{A^\varepsilon \cup B^\varepsilon\})$, violating Axiom A.6. Contradiction. $\square$

Lemma 10 states that imposing the Inclusion axiom is equivalent to requiring there to be no negative substates in any negative state. Formally,

LEMMA 10: *Suppose $\succsim$ has a nested DLR representation as in Definition 10, then $\succsim$ satisfies Axiom A.7 (Inclusion) if and only if $|N_i| = 0$ for any $i \in I^-$.*

Proof of Lemma 10. "If." Fix A, B with $B \subseteq A$. For any $i \in I^-$, $U_i(A) \geq U_i(B)$ since $|N_i| = 0$. Since $A \in \mathbb{A}$, we must have $-\max_{C \in \mathbb{A}} U_i(C) = -\max_{C \in \mathbb{A} \cup \{B\}} U_i(C)$. For any positive state $i \in I^+$, $\max_{C \in \mathbb{A} \cup \{B\}} U_i(C) \geq \max_{C \in \mathbb{A}} U_i(C)$. Thus, $V_{DLR}(\mathbb{A} \cup \{B\}) \geq V_{DLR}(\mathbb{A})$.

"Only if." Let $\mathbb{A} = \{A_i\}_{i \in I}$ be constructed the same way as in the proof of Lemma 9. Suppose by contradiction that $|N_{i^*}| \geq 1$ for some $i^* \in I^-$. For ease of exposition, write U_* for U_{i^*}, A_* for A_{i^*}, P_* for P_{i^*}, N_* for N_{i^*}, w_{*k} for w_{i^*k} for any $k \in P_*$, and v_{*j} for v_{i^*j} for any $j \in N_*$.

Since $|N_*| \geq 1$, we can fix some $j \in N_*$. Let $q_j := \arg\max_{p \in A_*} v_{*j}(p)$. For any $\varepsilon > 0$, define $A^\varepsilon := A_* \cup \{q_j - \varepsilon v_{*j}\}$ and $B^\varepsilon = (A_* \setminus \{q_j\}) \cup \{q_j - \varepsilon v_{*j}\}$. We can find ε small enough such that $q_j - \varepsilon v_{*j}$ is in the interior of $\Delta(Z)$. For any such ε, $v_{*j}(q_j - \varepsilon v_{*j}) = v_{*j}(q_j) - \varepsilon \|v_{*j}\|^2 < v_{*j}(q_j) = \max_{p \in A_*} v_{*j}(p)$. Moreover, we can make ε small enough so that $v_{*j}(q_j - \varepsilon v_{*j}) = v_{*j}(q_j) - \varepsilon \|v_{*j}\|^2 > \max_{p \in A_* \setminus \{q_j\}} v_{*j}(p)$, and

$$v_{*j'}(q_j - \varepsilon v_{*j}) = v_{*j'}(q_j) - \varepsilon v_{*j'}(v_{*j}) < \max_{p \in A_*} v_{*j'}(p), \ \forall j' \in N_* \text{ and } j' \neq j$$

$$w_{*k}(q_j - \varepsilon v_{*j}) = w_{*k}(q_j) - \varepsilon w_{*k}(v_{*j}) < \max_{p \in A_*} w_{*k}(p), \ \forall k \in P_*$$

For such an ε, we have $U_*(B^\varepsilon) > U_*(A^\varepsilon) = U_*(A_*)$. We can further scale down ε so that $U_i(A_i) > U_i(A^\varepsilon), U_i(B^\varepsilon)$ for any $i \in I$ and $i \neq i^*$. Then $V_{DLR}(\mathbb{A} \cup \{A^\varepsilon, B^\varepsilon\}) < V_{DLR}(\mathbb{A} \cup \{A^\varepsilon\})$ (since $i^* \in I^-$), which violates Axiom A.7. Contradiction. $\square$

Now we are ready to present the proof for Theorem 6.

Proof of Theorem 6. "If." Lemmas 5, 7, 9, 10 imply the neccessity of the axioms.

"Only if." Axioms A.1-A.4 deliver a nested DLR representation (Lemma 5). Axiom A.5 guarantees that there is at most one positive state (part 1 of Lemma 7). Axiom A.6 guarantees that there are no negative substates in any positive state and at most one positive substate in any negative state (Lemma 9). Axiom A.7 guarantees that there are no negetive substates in any negative state (Lemma 10). Thus, there can be at most one positive state (which contains only positive substates) and there might be multiple negative states (each of which contains at most one positive substate and no negative substates).

Therefore, Axioms A.1-A.7 imply that there exists two non-redundant collections of normalized expected utilities over $\Delta(Z)$, $\{w_k\}_{k\in P}$ and $\{v_j\}_{j\in N}$, such that $\succsim$ can be represented by the function $V_S : \widehat{\mathcal{D}} \to \mathbb{R}$ defined by

$$V_S(\mathbb{A}) = \max_{A\in\mathbb{A}} \sum_{k\in P} \max_{p\in A} w_k(p) - \sum_{j\in N} \max_{A\in\mathbb{A}} \max_{p\in A} v_j(p). \tag{19}$$

For convenience, define $U_0(A) := \sum_{k\in P} \max_{p\in A} w_k(p)$, $U_j(A) := \max_{p\in A} v_j(p)$ for each $j \in \mathbb{N}$ and $\widehat{v}(A) := V_S(\{A\}) = \sum_{k\in P} \max_{p\in A} w_k(p) - \sum_{j\in N} \max_{p\in A} v_j(p)$. If $\succsim$ can be represented by V_S as defined in equation (19) and $\succsim$ satisfies Axiom A.8, then $|P| \geq 1$ and U_0 is non-constant. Otherwise there will be no positive state in the representation and $\mathbb{A} \succsim \mathbb{B}$ for any $\mathbb{A} \subseteq \mathbb{B}$, violating Axiom A.8.

Since $\{v_j\}_{j\in N}$ is non-redundant, $\{U_j\}_{j\in N}$ must also be non-redundant.

LEMMA 11: *If $\succsim$ can be represented by V_S defined in equation (19) with $|P| \geq 1$ and $\succsim$ satisfies Axiom A.5, then there exists $\alpha > 0$ such that $U_0 = \alpha v$.*

Proof of Lemma 11. We discuss two possible cases.

Case 1: Suppose $\{U_0\} \cup \{U_j\}_{j\in N}$ is non-redundant. Then $V_S(\mathbb{A}) := \max_{A\in\mathbb{A}} U_0(A) - \sum_{j\in N} \max_{A\in\mathbb{A}} U_j(A)$ is a nested DLR representation with exactly one positive state. By part 2 of Lemma 7, this implies that there exists $\alpha > 0$ such that $U_0 = \alpha v$.

Case 2: Suppose $\{U_0\} \cup \{U_j\}_{j\in N}$ is redundant. By our previous results, $\{U_j\}_{j\in N}$ is not redundant and U_0 is non-constant. Therefore, there exists some $\beta > 0$ and exactly one $j \in N$ such that $U_0 = \beta U_j$. (N cannot be empty, otherwise U_0 is the only function in the collection and $\{U_0\}$ is not redundant.) It must be that $\beta > 1$, otherwise there will be no positive state (U_0 is absorbed by U_j) and Axiom A.8 will be violated. $U_0 = \beta U_j$ means that $\sum_{k\in P} \max_{p\in A} w_k(p) = \beta \max_{p\in A} v_j(p)$ for any menu A. Now this can only happen when $|P| = 1$, otherwise the LHS will exhibit strict preference for flexibility in some cases while the RHS will always exhibit strategic rationality. Let $P = \{k\}$, then $w_k = \beta v_j$, and $U_0 = \beta \max_{p \in A} v_j(p)$. Let $U_0' := (\beta - 1)U_j$, then $\{U_0'\} \cup \{U_{j'}\}_{j' \in N,\ j' \neq j}$ is non-redundant. Then we can apply Lemma 7 again so that there exists $\alpha > 0$ with $U_0' = \alpha \widehat{v}$ where

$$\widehat{v}(A) := V_S(\{A\}) = U_0(A) - \sum_{j \in N} U_j(A) = U_0'(A) - \sum_{j' \in N,\ j' \neq j} U_{j'}(A).$$

Therefore, $U_0 = \frac{\beta}{\beta-1} U_0' = \frac{\beta}{\beta-1}(\alpha \widehat{v})$, and $U_0 = \alpha' \widehat{v}$ where $\alpha' = \frac{\alpha\beta}{\beta-1} > 0$. □

Thus, Axioms A.1-A.8 guarantee the existence of some $\alpha > 0$ such that $U_0 = \alpha \widehat{v}$. That is, $\sum_{k \in P} \max_{p \in A} w_k(p) = \alpha \left[\sum_{k \in P} \max_{p \in A} w_k(p) - \sum_{j \in N} \max_{p \in A} v_j(p) \right]$, which further indicates that

$$\sum_{j \in N} \max_{p \in A} v_j(p) = \frac{\alpha - 1}{\alpha} \sum_{k \in P} \max_{p \in A} w_k(p). \tag{20}$$

Claim: For equation (20) to hold, it must be that $\alpha \geq 1$.

Proof of the Claim. Suppose by contradiction that $\alpha < 1$. Then $(\alpha - 1)/\alpha < 0$, and the RHS of equation (20) will represent a preference that exhibits the opposite of preference for flexibility, while the LHS of equation (20) represents a preference that exhibits preference for flexibility, contradiction. □

If $\alpha = 1$, then for any menu A, $\sum_{j \in N} \max_{p \in A} v_j(p) = 0$. By Lemma 4, this indicates that $N = \emptyset$. And $V_S(\mathbb{A}) = \max_{A \in \mathbb{A}} \sum_{k \in P} \max_{p \in A} w_k(p)$ with $|P| \geq 1$. We can get a PR representation by setting $K = 0$ and (doubly) normalizing each w_k.

If $\alpha > 1$, then $|N| \geq 1$, and for any menu A,

$$U_0(A) = \sum_{k \in P} \max_{p \in A} w_k(p) = \frac{\alpha}{\alpha - 1} \sum_{j \in N} \max_{p \in A} v_j(p),$$

which further indicates that

$$V_S(\mathbb{A}) = \max_{A \in \mathbb{A}} \left[\frac{\alpha}{\alpha - 1} \sum_{j \in N} \max_{p \in A} v_j(p) \right] - \sum_{j \in N} \max_{A \in \mathbb{A}} \max_{p \in A} v_j(p).$$

Since $\alpha > 1$, we can get a PR representation by setting $K = \alpha - 1 > 0$, (doubly) normalizing each v_j and scaling everything up by multiplicating K. □

A.3 Uniqueness of the PR representation

Let $\widehat{V}$ be a PR representation for $\succsim$ with parameters (μ, K). Define $\widehat{v} : \widehat{\mathcal{M}} \to \mathbb{R}$, for any menu $A \in \widehat{\mathcal{M}}$, by $\widehat{v}(A) := \widehat{V}(\{A\}) = \sum_{u \in \text{supp}(\mu)} \mu(u) \max_{p \in A} u(p)$. That is, $\widehat{v}$ represents a preference over $\widehat{\mathcal{M}}$ generated by $\succsim$ restricting to singleton menus of menus (menus of menus containing only one menu). This is a special case of the representation captured in Dekel et al. (2001) and Dekel et al. (2007) with preference for flexibility. Their identification results indicate that μ is uniquely identified. Formally,

LEMMA 12: *Suppose both (μ, K) and (μ', K') represent $\succsim$, then $\mu' = \mu$.*

To identify K, we follow steps similar to that in Sarver (2008). Define $\widehat{r} : \widehat{\mathcal{D}} \to \mathbb{R}$, for any menu of menus $\mathbb{A} \in \widehat{\mathcal{D}}$, by $\widehat{r}(\mathbb{A}) := \min_{A\in\mathbb{A}} \sum_{u\in\text{supp}(\mu)} \mu(u)R(A, \mathbb{A}, u)$ Thus, $\widehat{r}(\mathbb{A})$ is the agent's minimal expected regret facing $\mathbb{A}$. Given a menu of menus $\mathbb{A}$, the menu A that maximizes v also minimizes expected regret. Therefore, for any $\mathbb{A} \in \widehat{\mathcal{D}}$, the agent will choose $A \in \mathbb{A}$ to maximize $v(A)$, and $\widehat{V}(\mathbb{A}) = \max_{A\in\mathbb{A}} \widehat{v}(A) - \widehat{r}(\mathbb{A})$.

THEOREM 7: *Two PR representations (μ, K) and (μ', K') represent the same preference $\succsim$ if and only if $\widehat{v}' = \widehat{v}$ and $\widehat{r}' = \widehat{r}$.*

Proof of Theorem 7. The "if" part is straightforward. For the "only if" part, suppose (μ, K) and (μ', K') represent the same preference. The mixture space theorem guarantees that there exists $\alpha > 0$ such that $\widehat{v}' = \alpha\widehat{v}$ and $\widehat{r}' = \alpha\widehat{r}$. Then by Lemma 12, $\mu' = \mu$ implies that $\widehat{v}' = \widehat{v}$. Thus, $\alpha = 1$, and $\widehat{r}' = \widehat{r}$. □

We then rule out a less interesting case where Axiom A.5 is trivially satisfied.

LEMMA 13: *Suppose $\succsim$ has a PR representation with parameters (μ, K), then the following are equivalent: (1) $\widehat{r}(\mathbb{A}) = 0$ for all $\mathbb{A} \in \widehat{\mathcal{D}}$; (2) $\succsim$ satisfies monotonicity. That is, if $\mathbb{B} \subseteq \mathbb{A}$, then $\mathbb{A} \succsim \mathbb{B}$; (3) $K = 0$ or $\mu = \delta_u$ for some $u \in \mathcal{U}$ (or both).*

Proof of Lemma 13. $1 \implies 2$: Suppose $\widehat{r}(\mathbb{A}) = 0$ for all $\mathbb{A} \in \widehat{\mathcal{D}}$. If $\mathbb{A} \supseteq \mathbb{B}$, then $\widehat{V}(\mathbb{A}) = \max_{A\in\mathbb{A}} \widehat{v}(A) \geq \max_{B\in\mathbb{B}} \widehat{v}(B) = \widehat{V}(\mathbb{B})$. Thus, $\mathbb{A} \succsim \mathbb{B}$.

$2 \implies 1$: Suppose $\widehat{r}(\mathbb{A}) > 0$ for some $\mathbb{A} \in \widehat{\mathcal{D}}$. Let $A \in \arg\max_{B\in\mathbb{A}} \widehat{v}(B)$, then $\{A\} \subseteq \mathbb{A}$ but $\{A\} \succ \mathbb{A}$ since $\widehat{V}(\{A\}) = \widehat{v}(A)$ while $\widehat{V}(\mathbb{A}) = \widehat{v}(A) - \widehat{r}(\mathbb{A})$.

$1 \implies 3$: Suppose by contradiction that $K > 0$ and $|\text{supp}(\mu)| \geq 2$, then there exists a menu of lotteries $A_0 = \{p_u\}_{u\in\text{supp}(\mu)}$ such that p_u is the unique maximizer of u in A. Let $\mathbb{A} := \{A_0\}$. Then, $\widehat{r}(\mathbb{A}) = \min_{p\in A_0} \sum_{u\in\text{supp}(\mu)} \mu(u)K\left[\max_{q\in A_0} u(q) - u(p)\right]$. But for any $p_u \in A_0$, $\max_{q\in A_0} u'(q) - u'(p) > 0$ for any $u' \neq u$. So $\widehat{r}(\mathbb{A}) > 0$. □

A PR representation is *nontrivial* if there exist $\mathbb{A}, \mathbb{B}$ such that $\mathbb{B} \subseteq \mathbb{A}$ but $\mathbb{B} \succ \mathbb{A}$.

THEOREM 8: *Suppose $\succsim$ has a nontrivial PR representation and $\succsim$ has two PR representations (μ, K) and (μ', K'), then $\mu' = \mu$ and $K' = K$.*

Proof. $\mu' = \mu$ by Lemma 12, and by Theorem 7, $\widehat{v}' = \widehat{v}$, $\widehat{r}' = \widehat{r}$. Since $\succsim$ has a nontrivial PR representation, $K, K' > 0$. Since $\mu' = \mu$, it must be that $\widehat{r}(\mathbb{A})/K = \widehat{r}'(\mathbb{A})/K'$ for all $\mathbb{A} \in \widehat{\mathcal{D}}$. Thus, $\widehat{r}(\mathbb{A}) = \widehat{r}'(\mathbb{A})$ for all $\mathbb{A}$ implies that $K' = K$. □

B Proof of Theorem 1

To save space, we omit the proof for the necessity of these axioms from the main text. Interested readers can find the complete proof in the online appendix. Here we focus on the proof for the sufficiency of Axioms 1-8.

We have characterized the partial regret (PR) representation in Appendix A. The PR representation looks very similar to the SIT representation, only with the caveat that it is in the framework of menus of menus of lotteries. In this proof, we will connect Axioms 1-8 in the main text with Axioms A.1-A.8 in the lottery framework in Appendix A through two translations. The first translation will be from the preference over menus of menus of acts to menus of menus of "utility acts," and the second translation will be from the menus of menus of "utility acts" to menus of menus of lottery.

A useful equivalent expression of the SIT representation is

$$\begin{aligned} V(\mathbb{F}) = \max_{F\in\mathbb{F}} &\left[(1+K)\sum_{s\in S}\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega))\right] \\ &- K\sum_{s\in S}\max_{G\in\mathbb{F}}\max_{g\in G}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(g(\omega)) \end{aligned} \tag{21}$$

Step 1: From $\succsim$ to a preference over "menus of menus of utility acts."

Note that Axioms 1-3 imply their corresponding axioms (weak order, continuity and independence) over acts (i.e., menu of menus containing only one singleton menu, like $\{\{f\}\}$). Axioms 7 (Nontriviality) and part 1 of Axiom 8 (Domination) imply the nontriviality axiom and the monotonicity axiom in the Anscombe-Aumann framework. Therefore, by the standard result from the Anscombe-Aumann framework (see, for example, Kreps 2018), Axioms 1-3, 7 and 8 imply that there exists a unique probability measure $\bar{\pi}\in\Delta(\Omega)$ and a surjective affine utility index $u:\Delta(X)\to[0,1]$ such that the preference $\succsim$ restricting to acts can be represented by $V(f):=\sum_{\omega\in\Omega}\bar{\pi}(\omega)u\big(f(\omega)\big)$. For any act $f\in\mathcal{F}$, the composition $u\circ f:\Omega\to[0,1]$ specifies the utility associated with f in each ω. We call this the *utility act* induced by f, denoted by $u(f)$.

Let $\mathcal{F}_u := u(\mathcal{F}) = \{u(f)\mid f\in\mathcal{F}\}$, that is, $\mathcal{F}_u$ is the set of utility acts induced by AA acts from $\mathcal{F}$. Since u is surjective, $\mathcal{F}_u = [0,1]^{|\Omega|}$. Endow $\mathcal{F}_u$ with the Euclidean metric. Let $\mathcal{M}_u$ be the set of all non-empty compact subsets of $\mathcal{F}_u$, with typical elements F_u, G_u. We call these utility menus. Endow $\mathcal{M}_u$ with the Hausdorff metric. Let $\mathcal{D}_u$ be the set of all non-empty compact subsets of $\mathcal{M}_u$, with typical elements $\mathbb{F}_u, \mathbb{G}_u$. We call these utility menus of menus. Endow $\mathcal{D}_u$ with the Hausdorff metric.

LEMMA 14: *If $\succsim$ satisfies Axioms 1-3, 7 and 8, then $u(\mathbb{F}) = u(\mathbb{G})$ implies $\mathbb{F}\sim\mathbb{G}$.*

Proof of Lemma 14. We say two AA acts f and g are *indistinguishable* if $\{\{f(\omega)\}\} \sim \{\{g(\omega)\}\}$ for all $\omega \in \Omega$. Note that f and g are indistinguishable if and only if $u(f) = u(g)$ (since $u(f)$ and $u(g)$ are not scalars, but functions from Ω to $[0,1]$).

We say two menus F and G are indistinguishable if for any $f \in F$ there exists $g \in G$ that is indistinguishable to f and for any $g' \in G$ there exists $f' \in F$ that is indistinguishable to g'. For a menu $F \in \mathcal{M}$, let $u(F) := \{u(f) \mid f \in F\}$. Note that F and G are indistinguishable if and only if $u(F) = u(G)$. (The "only if" part is easy. For the "if" part: If $u(F) = u(G)$, then for any $f \in F$ there exists $g \in G$ such that $u(f) = u(g)$, so f and g are indistinguishable.)

We say two menus of menus $\mathbb{F}$ and $\mathbb{G}$ are indistinguishable if for any $F \in \mathbb{F}$ there exists $G \in \mathbb{G}$ that is indistinguishable to F and for any $G' \in \mathbb{G}$ there exists $F' \in \mathbb{F}$ that is indistinguishable to G'. For a menu of menus $\mathbb{F} \in \mathcal{D}$, let $u(\mathbb{F}) := \{u(F) \mid F \in \mathbb{F}\}$. Then $\mathbb{F}$ and $\mathbb{G}$ are indistinguishable if and only if $u(\mathbb{F}) = u(\mathbb{G})$. If $\mathbb{F}$ and $\mathbb{G}$ are indistinguishable, then $\mathbb{F}$ dominates $\mathbb{G}$ and $\mathbb{G}$ dominates $\mathbb{F}$, then by the second part of Axiom 8, $\mathbb{F} \sim \mathbb{F} \cup \mathbb{G}$ and $\mathbb{G} \sim \mathbb{F} \cup \mathbb{G}$. Therefore, $\mathbb{F} \sim \mathbb{G}$ by transitivty. $\square$

Let $\succsim_u$ over $\mathcal{D}_u$ be defined by $\mathbb{F}_u \succsim_u \mathbb{G}_u$ if and only if $\mathbb{F} \succsim \mathbb{G}$ where $\mathbb{F} \in u^{-1}(\mathbb{F}_u)$ and $\mathbb{G} \in u^{-1}(\mathbb{G}_u)$. Lemma 14 guarantees that $\succsim_u$ is a preference.

Step 2: From $\succsim_u$ to a preference over menus of menus of lotteries.

Recall that $\mathcal{F}_u$ is the collection of all utility acts. Thus, we can identify $\mathcal{F}_u$ with $[0,1]^{|\Omega|}$. For convenience, let $\Omega = \{\omega_1, \ldots, \omega_n\}$, and introduce an artificial state ω_0 and a new space $\mathcal{F}'$ defined by $\mathcal{F}' := \left\{f' \in [0,n] \times [0,1]^n \;\middle|\; \sum_{i=0}^n f'(\omega_i) = n\right\}$. We suppress the subscript u for ease of notation, but these are still utility acts.

Endow $\mathcal{F}'$ with the standard Euclidean metric. Consider $r : \mathcal{F}' \to \mathcal{F}_u$ where $r(f')$ is the vector in $\mathcal{F}$ that agrees with the last n components of f'. That is, $r(f') \in \mathcal{F}_u$ and for any $i = 1, 2, \ldots, n$, $[r(f')](\omega_i) = f'(\omega_i)$. It is easy to verify that r is a homeomorphism between $\mathcal{F}'$ and $\mathcal{F}_u$. Let r^{-1} denote its inverse. Let $\mathcal{M}'$ denote the set of nonempty compact subsets of $\mathcal{F}'$, with typical elements F', G'. Endow $\mathcal{M}'$ with the Hausdorff metric. Let $\mathcal{D}'$ denote the set of nonempty compact subsets of $\mathcal{M}'$, with typical elements $\mathbb{F}', \mathbb{G}'$. Endow $\mathcal{D}'$ with the Hausdorff metric. We slightly abuse notation and let r also denote the homeomorphism from $\mathcal{M}'$ to $\mathcal{M}_u$ and $\mathcal{D}'$ to $\mathcal{D}_u$, with $r(F') = \{r(f') \mid f' \in F'\} \in \mathcal{M}_u$ and $r(\mathbb{F}') = \{r(F') \mid F' \in \mathbb{F}'\} \in \mathcal{D}_u$. With this construction, we can define a preference $\succsim_*$ over $\mathcal{D}'$ by $\mathbb{F}' \succsim_* \mathbb{G}'$ if $r(\mathbb{F}') \succsim_u r(\mathbb{G}')$.

To move on, consider $\mathcal{F}'' := \left\{f'' \in [0,n]^{n+1} \;\middle|\; \sum_{i=0}^n f''(\omega_i) = n\right\}$. Endow $\mathcal{F}''$ with the standard Euclidean metric. Let $\mathcal{M}''$ denote the set of nonempty compact subsets of $\mathcal{F}''$, with typical elements F'', G''. Endow $\mathcal{M}''$ with the Hausdorff metric. Let $\mathcal{D}''$ denote the set of nonempty compact subsets of $\mathcal{M}''$, with typical elements $\mathbb{F}'', \mathbb{G}''$. Endow $\mathcal{D}''$ with the Hausdorff metric.

Fix a menu F^{n+1} defined by $F^{n+1} := \left\{\left(\frac{n}{n+1}, \ldots, \frac{n}{n+1}\right)\right\}$. Then $F^{n+1} \in \mathcal{M}'$. Also observe that for any $F'' \in \mathcal{M}''$ and $\varepsilon \leq \frac{1}{n^2}$,

$$\varepsilon F'' + (1-\varepsilon)F^{n+1} = \left\{ \varepsilon f'' + (1-\varepsilon)\left(\frac{n}{n+1}, \ldots, \frac{n}{n+1}\right) \;\middle|\; f'' \in F'' \right\} \in \mathcal{M}'.$$

Finally, define a relation $\succsim_{**}$ on $\mathcal{D}''$ by: $\mathbb{F}'' \succsim_{**} \mathbb{G}''$ if

$$\varepsilon\mathbb{F}'' + (1-\varepsilon)\{F^{n+1}\} \succsim_{*} \varepsilon\mathbb{G}'' + (1-\varepsilon)\{F^{n+1}\} \text{ for all } \varepsilon < \frac{1}{n^2}.$$

Step 3: Verify that $\succsim_{**}$ is the unique extension of $\succsim_{*}$ to $\mathcal{D}''$ satisfying adapted versions of Axioms A.1-A.8.

The construction of $\succsim_{*}$ and $\succsim_{**}$ in Step 2 is the same as that in DLST.

Axiom 1*—Weak Order: *$\succsim_{**}$ is complete and transitive.*

Axiom 2*—Continuity: *For any $\mathbb{F}''$, $\{\mathbb{G}'' : \mathbb{G}'' \succsim_{**} \mathbb{F}''\}$, $\{\mathbb{G}'' : \mathbb{F}'' \succsim_{**} \mathbb{G}''\}$ are closed.*

Axiom 3*—Independence: *For any $\mathbb{F}'', \mathbb{G}'', \mathbb{H}''$ and any $\alpha \in (0,1)$, $\mathbb{F}'' \succsim_{**} \mathbb{G}''$ if and only if $\alpha\mathbb{F}'' + (1-\alpha)\mathbb{H}'' \succsim_{**} \alpha\mathbb{G}'' + (1-\alpha)\mathbb{H}''$.*

Axiom 4*—Finiteness: *There exists a natural number N such that: (4.1) For every $\mathbb{F}''$, there exists $\mathbb{G}''$ with $|\mathbb{G}''| < N$ such that $\mathbb{G}''$ is critical for $\mathbb{F}''$; and (4.2) For every $\mathbb{F}'' \in \mathcal{D}''$ and every $F'' \in \mathbb{F}''$, there exists G'' with $|G''| < N$ such that G'' is critical for F'' in $\mathbb{F}''$.*

Axiom 5*—Ex-Ante Regret: *If $\{F''\} \succsim_{**} \{G''\}$ and $F'' \in \mathbb{F}''$, then $\mathbb{F}'' \succsim_{**} \mathbb{F}'' \cup \{G''\}$.*

Axiom 6*—Interim Preference for Flexibility: $\mathbb{F}'' \cup \{F'' \cup G''\} \succsim_{**} \mathbb{F}'' \cup \{F'', G''\}$.

Axiom 7*—Nontriviality: *There exist $\mathbb{F}'', \mathbb{G}'' \in \mathcal{D}''$ such that $\mathbb{F}'' \supseteq \mathbb{G}''$ and $\mathbb{F}'' \succ_{**} \mathbb{G}''$.*

Axiom 8*—Inclusion: *If $G'' \subseteq F''$ and $F'' \in \mathbb{F}''$, then $\mathbb{F}'' \cup \{G''\} \succsim_{**} \mathbb{F}''$.*

Lemma 15: *If a binary relation $\succsim$ over $\mathcal{D}$ satisfies Axioms A.1-A.8, then the induced binary relation $\succsim_{**}$ over $\mathcal{D}''$ satisfies Axioms 1*-8*.*

To save space, the proof of Lemma 15 is relegated to the online appendix.

Step 4: Apply Theorem 6 in Appendix A and translate the resulting PR representation to a SIT representation.

To proceed, we rescale every element of $\mathcal{F}''$ with factor $\frac{1}{n}$. This gives us a unit simplex and makes $\mathcal{D}''$ formally equivalent to the choice domain in Appendix A.

Through steps 2 and 3, we have verified that if a binary relation $\succsim$ over the menu of menus of acts satisfies Axioms 1-8, then the induced binary relation $\succsim_{**}$ defined over $\mathcal{D}''$ is unique and satisfies Axioms B.1**-B.8**. Therefore, Theorem 6 guarantees the existence of a finitely-supported probability measure $\widehat{\mu}$ over $\widehat{\mathcal{U}}$ (the set of doubly normalized expected utilities over $\Delta(\widehat{\Omega})$ where $\widehat{\Omega} = \Omega \cup \{\omega_0\})$ and a scalar $\widehat{K} \geq 0$ such that $\succsim_{**}$ can be represented by

$$\widehat{W}(\mathbb{F}'') = \max_{F'' \in \mathbb{F}''} \left[(1+\widehat{K}) \sum_{\widehat{u} \in \mathrm{supp}(\widehat{\mu})} \widehat{\mu}(\widehat{u}) \max_{f'' \in F''} \widehat{u}(f'') \right] - \widehat{K} \sum_{\widehat{u} \in \mathrm{supp}(\widehat{\mu})} \widehat{\mu}(\widehat{u}) \max_{F'' \in \mathbb{F}''} \max_{f'' \in F''} \widehat{u}(f'') \quad (22)$$

where $\widehat{u}(f'') = \sum_{\widehat{\omega} \in \widehat{\Omega}} \widehat{u}(\widehat{\omega}) f''(\widehat{\omega})$ is the expected utility for lottery f'' under taste $\widehat{u}$. Moreover, $\widehat{\mu}$ is uniquely identified (Lemma 13), and $\widehat{K}$ is uniquely identified when $|\mathrm{supp}(\widehat{\mu})| > 1$ (Theorem 8). Note that $\widehat{W}$ also represents $\succsim_*$ when restricting to $\mathcal{D}'$.

To get to the SIT representation, we aim for a representation for $\succsim$ of the form

$$V(\mathbb{F}) = \max_{F \in \mathbb{F}} \left[(1+K) \sum_{\pi \in \mathrm{supp}(\nu)} \nu(\pi) \max_{f \in F} u_\pi(f_u) \right] - K \sum_{\pi \in \mathrm{supp}(\nu)} \nu(\pi) \max_{F \in \mathbb{F}} \max_{f \in F} u_\pi(f_u) \quad (23)$$

where ν is a finitely-supported probability measure over $\Delta(\Omega)$ representing a distribution over posteriors induced by a prior and an information structure, and $u_\pi(f_u) := \sum_{\omega \in \Omega} \pi(\omega) f_u(\omega)$ for all $f_u \in \mathcal{F}_u$ and $\pi \in \Delta(\Omega)$. We now explore the additional constraint imposed on $\widehat{W}$ by Axiom B.8.

Lemma 16: *Suppose $\succsim_{**}$ has a representation as defined in equation (22) and $\succsim_u$ satisfies Axiom B.8, then $\widehat{u}(\omega) \geq \widehat{u}(\omega_0)$ for any $\widehat{u} \in \mathrm{supp}(\widehat{\mu})$ and any $\omega \in \Omega$.*

Proof of Lemma 16. Suppose by contradiction that $\succsim_{**}$ can be represented as in equation (22), $\succsim_u$ satisfies Axiom B.8, but there exists $\widehat{u}_* \in \mathrm{supp}(\widehat{\mu})$ and $\omega_* \in \Omega$ such that $\widehat{u}_*(\omega_0) > \widehat{u}_*(\omega_*)$. Consider $f' := (n-\varepsilon, 0, \ldots, 0, \varepsilon, 0, \ldots, 0)$ where $n - \varepsilon$ is assigned to state ω_0, ε is assigned to state ω_* and 0 is assigned to any other state. We can find ε small enough so that $f' \in \mathcal{F}'$. Let $g' := (n, 0, \ldots, 0)$ where n is assigned to state ω_0. Therefore, $r(f')$ dominates $r(g')$ (recall that the notion of domination is defined over $\mathcal{F}_u$ and thus only concerns the last n coordinates). Thus, part 1 of Axiom B.8 dictates that $\{\{r(f')\}\} \sim_u \{\{r(f'), r(g')\}\}$, and $\{\{f'\}\} \sim_* \{\{f', g'\}\}$ by construction.

Note that for any $\widehat{u} \in \mathrm{supp}(\widehat{\mu})$, $\widehat{u}(f') = (n-\varepsilon)\widehat{u}(\omega_0) + \varepsilon \widehat{u}(\omega_*)$ and $\widehat{u}(g') = n\widehat{u}(\omega_0)$. In particular, $\widehat{u}_*(f') - \widehat{u}_*(g') = \varepsilon \Big(\widehat{u}_*(\omega_*) - \widehat{u}_*(\omega_0) \Big) < 0$. Therefore,

$$\widehat{W}(\{\{f', g'\}\}) = \sum_{\widehat{u} \in \mathrm{supp}(\widehat{\mu})} \widehat{\mu}(\widehat{u}) \Big[\widehat{u}(f') \vee \widehat{u}(g') \Big] = \sum_{\widehat{u} \neq \widehat{u}_*} \widehat{\mu}(\widehat{u}) \Big[\widehat{u}(f') \vee \widehat{u}(g') \Big] + \widehat{\mu}(\widehat{u}_*) \widehat{u}_*(g')$$

while $\widehat{W}(\{\{f'\}\}) = \sum_{\widehat{u} \neq \widehat{u}_*} \widehat{\mu}(\widehat{u}) \widehat{u}(f') + \widehat{\mu}(\widehat{u}_*) \widehat{u}_*(f') < \widehat{W}(\{\{f', g'\}\})$, contradiction. $\square$

Given our construction of $\widehat{W}$, there are two normalizations we can do that will allow us to replace the finitely-supported probability measure $\widehat{\mu}$ on $\widehat{\mathcal{U}}$ with a unique finitely-supported probability measure ν on $\Delta(\Omega)$. For all $\omega \in \Omega$ and for all $\widehat{u}$, define $\xi(\widehat{u}(\omega)) :=$

$\widehat{u}(\omega) - \widehat{u}(\omega_0)$. Since $\sum_{i=0}^{n} f'(\omega_i) = n$ for all $f' \in \mathcal{F}'$ and ξ adds a constant to every $\widehat{u}$, $\arg\max_{f''\in F''} \left(\sum_{i=0}^{n} f''(\omega_i)\xi(\widehat{u}(\omega_i))\right) = \arg\max_{f''\in F''} \left(\sum_{i=1}^{n} f''(\omega_i)\widehat{u}(\omega_i)\right)$ for all $F'' \in \mathcal{M}''$ and all $\widehat{u} \in \text{supp}(\widehat{\mu})$. Furthermore, by Lemma 16, $\xi(\widehat{u}(\omega)) \geq 0$ for all $\omega \in \Omega$ and all $\widehat{u} \in \text{supp}(\widehat{\mu})$. Therefore, we could transform $\xi \circ \widehat{u} : \Omega \to \mathbb{R}$ into a probability measure $\pi_{\widehat{u}}$ defined by $\pi_{\widehat{u}}(\omega) := \frac{\xi(\widehat{u}(\omega))}{\sum_{\omega'\in\Omega} \xi(\widehat{u}(\omega'))}$ for each $\omega \in \Omega$. Finally, we can get the SIT representation by adjusting the weights on each $\pi_{\widehat{u}}$ by defining $\nu \in \Delta_0(\Delta(\Omega))$ by $\nu(\pi_{\widehat{u}}) := \frac{\sum_{\omega\in\Omega} \xi(\widehat{u}(\omega))}{\sum_{\widehat{u}'\in\text{supp}(\widehat{\mu})} \sum_{\omega\in\Omega} \xi(\widehat{u}'(\omega))} \widehat{\mu}(\widehat{u})$. Lastly, we can set S equal to $\{1, 2, \ldots, m\}$ where $m = |\text{supp}(\nu)|$ is the size of the support for ν, and the conditional probability distributions $\sigma(s \mid \omega)$ can be uniquely determined.

C Other Omitted Proofs

C.1 Proof of Theorem 2

Proof. The uniqueness of the SIT representation follows from the uniqueness of the PR representation. Through the proof of Theorem 1, we have established that if $\succsim$ has a SIT representation with parameters (π, u, K, σ), then the translated preference $\succsim_{**}$ has a PR representation with parameters (ν, K) where ν is a distribution over posteriors induced by π and σ and K is a non-negative regret intensity level.

Therefore, if $(\pi_0, u_0, K_0, \sigma_0)$ and (π, u, K, σ) both represent $\succsim$ through a SIT representation, then (ν_0, K_0) and (ν, K) both represent the induced preference $\succsim_{**}$ through a PR representation. Then we can apply Lemma 13 and Theorem 8 to conclude that $\nu_0 = \nu$ and $K_0 = K$ whenever $|\text{supp}(\nu)| > 1$. This completes the proof. $\square$

C.2 Proof of Theorem 3

Definition 11: A binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ has an *aligned informational tradeoff representation* if there exists a tuple $\left(\pi, u, (K^\sigma)_{\sigma\in\mathcal{I}}, (i^\sigma)_{\sigma\in\mathcal{I}}\right)$ that consists of a prior π on Ω, a non-constant affine function $u : \Delta(X) \to \mathbb{R}$, a collection of information structures $i^\sigma : \Omega \to \Delta(S^\sigma)$ and a collection of non-negative scalars K^σ such that $\succsim$ can be represented by the function $W : \mathcal{D} \times \mathcal{I} \to \mathbb{R}$ defined by

$$\begin{aligned} W(\mathbb{F}, \sigma) = \max_{F\in\mathbb{F}} &\left[(1 + K^\sigma) \sum_{t\in S^\sigma} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega) i^\sigma(t \mid \omega) u(f(\omega))\right] \\ &- K^\sigma \sum_{t\in S^\sigma} \max_{G\in\mathbb{F}} \max_{g\in G} \sum_{\omega\in\Omega} \pi(\omega) i^\sigma(t \mid \omega) u(g(\omega)) \end{aligned} \tag{24}$$

This is called the "aligned" IT representation because the collection of IT representations for each σ is aligned together by the same π and u.

Axiom 14—Stable Preference over Acts: *For any act f, g and any information σ, σ', $(f, \sigma) \succsim (g, \sigma)$ if and only if $(f, \sigma') \succsim (g, \sigma')$.*

Axiom 14 is motivated by the role of information in the decision process. When there is only one act to be chosen, any information is irrelevant.

Lemma 17: *A binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ has an aligned informational tradeoff representation if and only if: (1) $\succsim$ satisfies Weak Order, Stable Preference over Acts and Act Independence; and (2) $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$.*

The proof of Lemma 17 is relegated to Appendix C.3.

In sum, Stable Preference over Acts and Act Independence guarantee the existence of a utility representation for a preference over the two dimensional choice domain such that the identified prior belief and taste over outcomes are information-independent.

Definition 12: A binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ has a *regret-varying IT representation* if there exists a tuple $\left(\pi, u, (K^\sigma)_{\sigma \in \mathcal{I}}\right)$ that consists of a probability measure π on Ω, a non-constant affine function $u : \Delta(X) \to \mathbb{R}$ and a collection of non-negative scalars K^σ such that $\succsim$ can be represented by the function $W : \mathcal{D} \times \mathcal{I} \to \mathbb{R}$ defined by

$$\begin{aligned} W(\mathbb{F}, \sigma) = \max_{F \in \mathbb{F}} & \left[(1 + K^\sigma) \sum_{s \in S} \max_{f \in F} \sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(f(\omega)) \right] \\ & - K^\sigma \sum_{s \in S} \max_{G \in \mathbb{F}} \max_{g \in G} \sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(g(\omega)) \end{aligned} \tag{25}$$

Comparing to an aligned IT representation, a regret-varying IT representation has the feature that σ is no longer just a label, it is indeed the information structure anticipated by the agent. It is not yet the IT representation because the regret intensity level could still vary across different information structures.

Lemma 18: *A binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ has a regret-varying informational tradeoff representation if and only if: (1) $\succsim$ satisfies Weak Order, Act Independence, SRNI and Reduction; and (2) $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$.*

The proof of Lemma 18 is relegated to Appendix C.4.

Finally, the Balance axiom guarantees that the parameters for regret intensity identified from all conditional preferences over menus of menus are the same.

Proof of Theorem 3. "Only if." Suppose $\succsim$ has an IT representation, then $\succsim$ has a regret-varying IT representation with $K^\sigma = K$ for all $\sigma \in \mathcal{I}$. By Lemma 18, $\succsim$ satisfies Weak

Order, Act Independence, SRNI and Reduction, and $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$. As for the Balance axiom, note that $W(D(F),\sigma) = (1+K)W(\{F\},o) - K \cdot W(\{F\},\sigma)$ for any F and any σ. Rearranging, we have

$$W(\{F\},o) = \frac{1}{1+K} \cdot W(D(F),\sigma) + \frac{K}{1+K} \cdot W(\{F\},\sigma). \tag{26}$$

Suppose $(\{F\},\sigma) \succ (\{F\},o)$, then $W(\{F\},\sigma) > W(\{F\},o)$, which further indicates that $W(\{F\},\sigma) > W(D(F),\sigma)$, otherwise the above equation cannot hold. Further suppose that $\Big(\alpha D(F) + (1-\alpha)\{F\},\sigma\Big) \sim (\{F\},o)$ for some $\alpha \in (0,1]$, then

$$W(\{F\},o) = \alpha \cdot \widetilde{W}(D(F),\sigma) + (1-\alpha) \cdot W(\{F\},\sigma). \tag{27}$$

For equations (26) and (27) to hold at the same time, it must be that $\alpha = \frac{1}{1+K}$. Then $\succsim$ must satisfy Balance by the fact that equation (26) is an identity.

"If." First, we can apply Lemma 18 to conclude that $\succsim$ has a regret-varying IT representation with parameters $\big(\pi, u, (K^\sigma)_{\sigma\in\mathcal{I}}\big)$.

If some σ induces a degenerate distribution over posteriors, then we can set K^σ to be any non-negative scalar without affecting the value of W.

Suppose by contradiction that $K^\sigma \neq K^{\sigma'}$ for some $\sigma,\sigma' \in \mathcal{I}$ such that σ and σ' each induces a non-degenerate distribution over posteriors. We want to show that Balance must be violated. By assumption, we can find F, G such that $(\{F\},\sigma) \succ (\{F\},o)$ and $(\{G\},\sigma') \succ (\{G\},o)$. Let $\alpha = 1/(1+K^\sigma)$ and $\alpha' = 1/(1+K^{\sigma'})$, then $\alpha \neq \alpha'$. Without loss of generality, suppose $\alpha > \alpha'$. Follow a similar computation in the proof for the necessity of Balance,

$$\begin{aligned} \Big(\alpha' D(F) + (1-\alpha')\{F\},\sigma\Big) \succ \Big(\alpha D(F) + (1-\alpha)\{F\},\sigma\Big) &\sim (\{F\},o) \\ \Big(\alpha' D(G) + (1-\alpha)\{G\},\sigma'\Big) &\sim (\{G\},o) \end{aligned}$$

which is a direct violation of Balance. The uniquenuss of an IT representation is an immediate corollary of the uniqueness of the SIT represenation for each $\succsim^\sigma$. □

C.3 Proof of Lemma 17

Proof. "Only if." Suppose $\succsim$ has an aligned IT representation, then $\succsim$ must be complete and transitive. Moreover, each $\succsim^\sigma$ has a SIT representation with parameters $(\pi, u, i^\sigma, K^\sigma)$. By Theorem 1, $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$.

For any act f and any $\sigma \in \mathcal{I}$, $W(f,\sigma) = \sum_{\omega\in\Omega} \pi(\omega)u(f(\omega))$. That is, the value of f is independent of σ. Stable Preference over Acts is thus satisfied.

To show that $\succsim$ satisfies Act Independence, note that W is affine in its first argument. That is, $W(\alpha\mathbb{F} + (1-\alpha)\mathbb{G},\sigma) = \alpha W(\mathbb{F},\sigma) + (1-\alpha)W(\mathbb{G},\sigma)$ for any $\mathbb{F},\mathbb{G} \in \mathcal{D}$ and any

$\sigma \in \mathcal{I}$. The rest follows from applying this linearity and the fact that $W(h,\sigma) = W(h,\sigma')$ for any act h and any pairs of σ and σ'.

"If." As mentioned in the main text, $\succsim$ being complete and transitive implies that $\succsim^\sigma$ is complete and transitive for each $\sigma \in \mathcal{I}$. By Theorem 1, each $\succsim^\sigma$ has a SIT representation V^σ with parameters $(\pi^\sigma, u^\sigma, K^\sigma, i^\sigma)$.

LEMMA 19: *If $\succsim$ satisfies Stable Preference over Acts, then there exists π and u such that $\pi^\sigma = \pi$ and $u^\sigma = \alpha^\sigma u + b^\sigma$ for some $\alpha^\sigma > 0$ and $b^\sigma \in \mathbb{R}$ for all $\sigma \in \mathcal{I}$.*

Proof of Lemma 19. For any act $f \in \mathcal{F}$, $V^\sigma(f) = \sum_{\omega\in\Omega} \pi^\sigma(\omega)u^\sigma(f(\omega))$. That is, each V^σ reduces to a subjective expected utility representation with parameters (π^σ, u^σ) when restricting to acts. Suppose $\succsim$ satisfies Stable Preference over Acts, then $f \succsim^\sigma g$ if and only if $f \succsim^{\sigma'} g$ for any acts f, g and any information structures σ, σ'. Therefore, $\succsim^\sigma$ and $\succsim^{\sigma'}$ induce the same preference over acts. Then the result follows from the uniqueness result of Anscombe and Aumann (1963). $\square$

The Act Independence axiom helps to establish that $\succsim$ actually has a utility representation. Let $\widetilde{V}^\sigma$ be a SIT representation with parameters $(\pi, u, K^\sigma, i^\sigma)$ where π, u are as characterized in Lemma 19. Then $\widetilde{V}^\sigma$ represents $\succsim^\sigma$ for each $\sigma \in \mathcal{I}$. Consider a function $W : \mathcal{D} \times \mathcal{I} \to \mathbb{R}$ defined by $W(\mathbb{F}, \sigma) := \widetilde{V}^\sigma(\mathbb{F})$. Suppose $\succsim$ also satisfies Act Independence, we will now show that $\succsim$ can be represented by W.

LEMMA 20: *For any $(\mathbb{F}, \sigma)$ and $(\mathbb{G}, \sigma')$, there exist acts f, g, h and $\alpha \in (0, 1]$ such that $(\alpha\mathbb{F} + (1-\alpha)h, \sigma) \sim (f, \sigma)$ and $(\alpha\mathbb{G} + (1-\alpha)h, \sigma') \sim (g, \sigma')$*

Proof of Lemma 20. It is without loss to normalize the taste function $u : \Delta(X) \to \mathbb{R}$ to have range $[0, 1]$. Moreover, $u : \Delta(X) \to [0, 1]$ is surjective, so for any $a \in [0, 1]$, there exists an act $f \in \mathcal{F}$ such that $\widetilde{V}^\sigma(f) = a$ for all information $\sigma \in \mathcal{I}$. Therefore, $0 \leq \widetilde{V}^\sigma(\{F\}) \leq 1$ for any $F \in \mathcal{M}$ and $\sigma \in \mathcal{I}$. Finally, this indicates that $\widetilde{V}^\sigma(\mathbb{F}) \leq \widetilde{V}^\sigma(M(\mathbb{F})) \leq 1$ for any menu of menus $\mathbb{F}$ and information structure $\sigma \in \mathcal{I}$. Therefore, for any act h and any $\sigma \in \mathcal{I}$, $\widetilde{V}^\sigma\Big(\alpha\mathbb{F} + (1-\alpha)h\Big) = \alpha\widetilde{V}^\sigma(\mathbb{F}) + (1-\alpha)\widetilde{V}^\sigma(h)$, similar for $\widetilde{V}^{\sigma'}\Big(\alpha\mathbb{G} + (1-\alpha)h\Big)$. By our arguments above, there exists h such that $\widetilde{V}^\sigma(h) = \widetilde{V}^{\sigma'}(h) = 1$. For such an h, $\widetilde{V}^\sigma(h) \geq \widetilde{V}^\sigma(\mathbb{F})$ and $\widetilde{V}^{\sigma'}(h) \geq \widetilde{V}^{\sigma'}(\mathbb{G})$, which further implies that we can find $\alpha^* \in (0, 1]$ such that $\widetilde{V}^\sigma\Big(\alpha^*\mathbb{F} + (1-\alpha^*)h\Big) = \alpha^*\widetilde{V}^\sigma(\mathbb{F}) + (1-\alpha^*)\widetilde{V}^\sigma(h) \geq 0$ and $\widetilde{V}^{\sigma'}\Big(\alpha^*\mathbb{G} + (1-\alpha^*)h\Big) = \alpha^*\widetilde{V}^{\sigma'}(\mathbb{G}) + (1-\alpha^*)\widetilde{V}^{\sigma'}(h) \geq 0$. Since u is surjective, there exists some acts $f, g \in \mathcal{F}$ such that $(\alpha^*\mathbb{F} + (1-\alpha^*)h, \sigma) \sim (f, \sigma)$ and at the same time, $(\alpha^*\mathbb{G} + (1-\alpha^*)h, \sigma') \sim (g, \sigma')$. $\square$

Fix $(\mathbb{F},\sigma)$ and $(\mathbb{G},\sigma')$. Let f,g,h and $\alpha\in(0,1]$ satisfy Lemma 20. Then

$$\begin{aligned}
(\mathbb{F},\sigma)\succsim(\mathbb{G},\sigma') &\iff (\alpha\mathbb{F}+(1-\alpha)h,\sigma)\succsim(\alpha\mathbb{G}+(1-\alpha)h,\sigma')\\
&\iff (f,\sigma)\succsim(g,\sigma')\\
&\iff (f,\sigma)\succsim(g,\sigma)\\
&\iff \sum_{\omega\in\Omega}\pi(\omega)u(f(\omega))\geq\sum_{\omega\in\Omega}\pi(\omega)u(g(\omega))\\
&\iff W(\alpha\mathbb{F}+(1-\alpha)h,\sigma)\geq W(\alpha\mathbb{G}+(1-\alpha)h,\sigma')\\
&\iff W(\mathbb{F},\sigma)\geq W(\mathbb{G},\sigma')
\end{aligned}$$

The first equivalence holds since $\succsim$ satisfies Act Independence. The second equivalence holds by the construction of f,g,h and α. The third equivalence follows from $\succsim$ satisfying Stable Preference for Acts. The fourth equivalence holds since $\widetilde{V}^\sigma$ represents $\succsim^\sigma$. The fifth equivalence holds since $W(\mathbb{F},\sigma)=\widetilde{V}^\sigma(\mathbb{F})$ so $W(\cdot,\sigma)$ can represent $\succsim^\sigma$ and $(\alpha\mathbb{F}+(1-\alpha)h,\sigma)\sim(f,\sigma)$ by construction. The last equivalence holds since W is affine in the menus of menus. This completes the proof of the sufficiency of the axioms since W is by definition an aligned informational tradeoff representation. □

C.4 Proof of Lemma 18

Proof. "Only if." Suppose $\succsim$ has a regret-varying IT representation, then $\succsim$ has an aligned IT representation with $i^\sigma=\sigma$ for each $\sigma\in\mathcal{I}$. $\succsim$ must satisfy Reduction since

$$\begin{aligned}
W(\{F_\sigma\},o)=\max_{f\in F_\sigma}\sum_{\omega\in\Omega}\pi(\omega)u(f(\omega))&=\max_{\gamma\in F^S}\sum_{\omega\in\Omega}\pi(\omega)u\Big(\gamma_\sigma(\omega)\Big)\\
&=\max_{\gamma\in F^S}\sum_{\omega\in\Omega}\pi(\omega)u\left(\sum_{s\in S}\sigma(s\mid\omega)[\gamma(s)](\omega)\right)\\
&=\max_{\gamma\in F^S}\sum_{\omega\in\Omega}\pi(\omega)\sum_{s\in S}\sigma(s\mid\omega)u\Big([\gamma(s)](\omega)\Big)\\
&=\sum_{s\in S}\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega))=W(\{F\},\sigma)
\end{aligned}$$

where the second equality follows from the construction of F_σ, the third equality follows from the definition of an induced act γ_σ, the fourth equality follows from u being an affine function, and the fifth equality follows from the fact that maximizing over the set of all plans is equivalent to maximizing over the menu of acts contingent on each signal realization.

By Lemma 17, $\succsim$ satisfies Weak Order and Act Independence, and $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma\in\mathcal{I}$. In addition, $\succsim^o$ can be represented by (π,u,K^o,o). By Lemma 1, $\succsim$ satisfies SRNI.

"If." As argued in the main text, if $\succsim$ satisfies Weak Order and Reduction, then $\succsim$ satisfies Stable Preference over Acts. By Lemma 17, $\succsim$ has an aligned IT representation with parameters $\left(\pi, u, \left(K^{\sigma}\right)_{\sigma\in\mathcal{I}}, \left(i^{\sigma}\right)_{\sigma\in\mathcal{I}}\right)$, we want to show that if $\succsim$ satisfies SRNI and Reduction, then i^{σ} and σ induce the same distribution over posteriors for any σ. By Lemma 1, $W(\mathbb{F}, o) = \max_{F\in\mathbb{F}} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)u(f(\omega))$ since SRNI implies that i^{o} coincides with o. By Reduction, $W(\{F\}, \sigma) = W(\{F_{\sigma}\}, o)$ for any menu $F \in \mathcal{M}$ and information structure $\sigma \in \mathcal{I}$. Note that

$$W(\{F\}, \sigma) = \sum_{t\in S^{\sigma}} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega) i^{\sigma}(t \mid \omega) u(f(\omega))$$
$$W(\{F_{\sigma}\}, o) = \sum_{s\in S} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega) u(f(\omega))$$

where the first equation follows from applying W to the pair $(\{F\}, \sigma)$ and S^{σ} is the set of signals for the identified information structure i^{σ}, that is, $i^{\sigma} : \Omega \to \Delta(S^{\sigma})$. The second equation follows from the previous derivation in the proof for the necessity of Reduction (S is the set of signals for σ and generally different from S^{σ}). Therefore, σ and i^{σ} must induce the same distribution over posteriors (otherwise, by a standard result, we can find a menu F such that $W(\{F\}, \sigma) \neq W(\{F_{\sigma}\}, o)$). □

C.5 Proof of Lemma 3

Proof. Combining equations (15), (16) and (17), we can write W_1 as

$$\begin{aligned} W_1(\mathbb{F}, \sigma) = \ & \max_{F\in\mathbb{F}} \Bigg[\sum_{s\in S} (1 + K_0 + K_1 + K_2) \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega) u(f(\omega)) \\ & \qquad - K_2 \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega) \max_{h\in F} u(h(\omega)) \Bigg] \\ & - K_0 \sum_{s\in S} \max_{G\in\mathbb{F}} \max_{g\in G} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega) u(g(\omega)) \\ & - K_1 \sum_{\omega\in\Omega} \pi(\omega) \max_{G\in\mathbb{F}} \max_{g\in G} u(g(\omega)) \end{aligned} \tag{28}$$

If $K_2 = 0$, this reduces to

$$\begin{aligned} W_1(\mathbb{F}, \sigma) = \ & (1 + K_0 + K_1) \max_{F\in\mathbb{F}} \sum_{s\in S} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega) u(f(\omega)) \\ & - K_0 \sum_{s\in S} \max_{G\in\mathbb{F}} \max_{g\in G} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega) u(g(\omega)) \\ & - K_1 \sum_{\omega\in\Omega} \pi(\omega) \max_{G\in\mathbb{F}} \max_{g\in G} u(g(\omega)) \end{aligned} \tag{29}$$

Note that the last term does not depend on σ. For convenience, let

$$U_1(\mathbb{F},\sigma) := \max_{F\in\mathbb{F}} \sum_{s\in S} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega)u(f(\omega))$$
$$U_2(\mathbb{F},\sigma) := \sum_{s\in S} \max_{G\in\mathbb{F}} \max_{g\in G} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega)u(g(\omega))$$

Therefore,

$$\begin{aligned}
& W_1(\mathbb{F},\sigma) \geq W_1(\mathbb{F},\sigma') \\
\iff & (1+K_0+K_1)\cdot U_1(\mathbb{F},\sigma) - K_0 \cdot U_2(\mathbb{F},\sigma) \geq (1+K_0+K_1)\cdot U_1(\mathbb{F},\sigma') - K_0 \cdot U_2(\mathbb{F},\sigma') \\
\iff & \frac{1+K_0+K_1}{1+K_1}\cdot U_1(\mathbb{F},\sigma) - \frac{K_0}{1+K_1}\cdot U_2(\mathbb{F},\sigma) \geq \frac{1+K_0+K_1}{1+K_1}\cdot U_1(\mathbb{F},\sigma') - \frac{K_0}{1+K_1}\cdot U_2(\mathbb{F},\sigma') \\
\iff & (1+K)\cdot U_1(\mathbb{F},\sigma) - K\cdot U_2(\mathbb{F},\sigma) \geq (1+K)\cdot U_1(\mathbb{F},\sigma') - K\cdot U_2(\mathbb{F},\sigma') \\
\iff & W(\mathbb{F},\sigma) \geq W(\mathbb{F},\sigma')
\end{aligned}$$

where the third equivalence follows from the assumption that $K = K_0/(1+K_1)$. □

References

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

Anscombe, F. J. and R. J. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199–205.

Bell, D. E. (1982): "Regret in Decision Making under Uncertainty," *Operations Research*, 30, 961–981.

Brunnermeier, M. K. and J. A. Parker (2005): "Optimal Expectations," *American Economic Review*, 95, 1092–1118.

Buturak, G. and Ö. Evren (2017): "Choice Overload and Asymmetric Regret," *Theoretical Economics*, 12, 1029–1056.

Caplin, A. and J. Leahy (2001): "Psychological Expected Utility Theory and Anticipatory Feelings," *The Quarterly Journal of Economics*, 116, 55–79.

Dekel, E., B. L. Lipman, and A. Rustichini (2001): "Representing Preferences with a Unique Subjective State Space," *Econometrica*, 69, 891–934.

——— (2009): "Temptation-Driven Preferences," *The Review of Economic Studies*, 76, 937–971.

Dekel, E., B. L. Lipman, A. Rustichini, and T. Sarver (2007): "Representing Preferences with a Unique Subjective State Space: A Corrigendum," *Econometrica*, 75, 591–600.

Dillenberger, D. (2010): "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior," *Econometrica*, 78, 1973–2004.

Dillenberger, D., J. S. Lleras, P. Sadowski, and N. Takeoka (2014): "A Theory of Subjective Learning," *Journal of Economic Theory*, 153, 287–312.

Ehrlich, D., I. Guttman, P. Schönbach, and J. Mills (1957): "Postdecision Exposure to Relevant Information," *The Journal of Abnormal and Social Psychology*, 54, 98.

Eyster, E., S. Li, and S. Ridout (2021): "A Theory of Ex Post Rationalization," *arXiv preprint arXiv:2107.07491*.

Festinger, L. (1957): *A Theory of Cognitive Dissonance*, Stanford University Press.

Frey, D. and D. Stahlberg (1986): "Selection of Information after Receiving More or Less Reliable Self-Threatening Information," *Personality and Social Psychology Bulletin*, 12, 434–441.

Golman, R., D. Hagmann, and G. Loewenstein (2017): "Information Avoidance," *Journal of Economic Literature*, 55, 96–135.

Hilbert, L. P., M. K. Noordewier, and W. W. van Dijk (2022): "The Prospective Associations Between Financial Scarcity and Financial Avoidance," *Journal of Economic Psychology*, 88, 102459.

Jakobsen, A. M. (2021): "An Axiomatic Model of Persuasion," *Econometrica*, 89, 2081–2116.

Jonas, E., S. Schulz-Hardt, D. Frey, and N. Thelen (2001): "Confirmation Bias in Sequential Information Search after Preliminary Decisions: An Expansion of Dissonance Theoretical Research on Selective Exposure to Information." *Journal of Personality and Social Psychology*, 80, 557.

Karlsson, N., G. Loewenstein, and D. Seppi (2009): "The Ostrich Effect: Selective Attention to Information," *Journal of Risk and Uncertainty*, 38, 95–115.

Kopylov, I. (2009): "Finite Additive Utility Representations for Preferences over Menus," *Journal of Economic Theory*, 144, 354–374.

Kőszegi, B. (2003): "Health Anxiety and Patient Behavior," *Journal of Health Economics*, 22, 1073–1084.

Kreps, D. (2018): *Notes on the Theory of Choice*, Routledge.

Lerman, C., C. Hughes, B. J. Trock, R. E. Myers, D. Main, A. Bonney, M. R. Abbaszadegan, A. E. Harty, B. A. Franklin, J. F. Lynch, et al. (1999): "Genetic Testing in Families with Hereditary Nonpolyposis Colon Cancer," *JAMA*, 281, 1618–1622.

LOOMES, G. AND R. SUGDEN (1982): "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," *The Economic Journal*, 92, 805–824.

——— (1987): "Some Implications of a More General Form of Regret Theory," *Journal of Economic Theory*, 41, 270–287.

OSTER, E., I. SHOULSON, AND E. DORSEY (2013): "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease," *American Economic Review*, 103, 804–30.

PERSOSKIE, A., R. A. FERRER, AND W. M. KLEIN (2014): "Association of Cancer Worry and Perceived Risk with Doctor Avoidance: An Analysis of Information Avoidance in a Nationally Representative US Sample," *Journal of Behavioral Medicine*, 37, 977–987.

SARVER, T. (2008): "Anticipating Regret: Why Fewer Options May Be Better," *Econometrica*, 76, 263–305.

SICHERMAN, N., G. LOEWENSTEIN, D. J. SEPPI, AND S. P. UTKUS (2016): "Financial Attention," *The Review of Financial Studies*, 29, 863–897.

STOVALL, J. E. (2018): "Temptation with Uncertain Normative Preference," *Theoretical Economics*, 13, 145–174.

SUGDEN, R. (1993): "An Axiomatic Foundation for Regret Theory," *Journal of Economic Theory*, 60, 159–180.

YANG, Y. (2024): "An Axiomatic Model of Cognitive Dissonance," *working paper*.