Regret and Information Avoidance*

Zichang Wang†

This version: May 3rd, 2023
Job Market Paper

Abstract

Empirical evidence suggests that individuals selectively avoid information depending on past choices. We address these findings by studying an agent whose choice behavior can be modeled as if she trades off two conflicting effects of information. The first is a psychological cost from the regret about past choices that are revealed to be suboptimal by the information, whereas the second is the instrumental value of information for making better-informed choices in the future. Our main axioms reflect the agent's desire to have fewer options before the arrival of information and to have more options after the arrival of information. We also posit axioms that connect the agent's consumption choice with her information choice. We show that all parameters can be uniquely identified from the choice behavior. We also provide comparative statics on the agent's information aversion attitude.

Keywords: Regret, information avoidance, preference over menus of menus

*I am deeply indebted to Todd Sarver for his invaluable advice, guidance and encouragement 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, Jaden Chen, David Dillenberger, Alexander Jakobsen, Fei Li, Alfonso Maselli, Curtis Taylor, four anonymous reviewers for EC'23, numerous speakers at the Duke-UNC micro theory seminar, and seminar participants at the Duke economic theory lunch, the UNC student theory workshop, the University of Hong Kong, and the 6th Triangle Micro Conference. All errors are my own.

†Duke University, Department of Economics, 213 Social Sciences/Box 90097, Durham, NC 27708. Email: zichang.wang@duke.edu

1 Introduction

Information avoidance is the active avoidance of freely accessible information relevant to decision-making. It is puzzling because standard economic analyses suggest that information may help an individual make better decisions and can also be ignored at no cost. Thus, absent of strategic considerations, information should never be harmful. Nevertheless, there exists overwhelming evidence that information avoidance is widespread.[1]

Some existing theories can rationalize information avoidance in certain specific contexts.[2] However, few of them can directly account for empirical findings that information avoidance is often connected with choices made in the past. Such connections have been extensively documented in a longstanding literature in psychology.[3] The key observation from this literature is that after making a choice, people exhibit "selective exposure to information" by seeking supportive information and avoiding contradictory information to that choice.[4] In this paper, we directly address these findings by developing a model that formally links information avoidance to past choices.

How could an economic agent's preference for avoiding information be driven by a choice made in the past? One natural answer is through regret. Information could reveal an agent's past choice to be suboptimal and cause her to experience a sense of regret for not having chosen a different alternative. We refer to this (psychological) effect as the regret cost of information. In addition to the regret cost, our model will also incorporate the instrumental value of information. This instrumental value is derived from the agent's gain from making better-informed choices in the future. In reality, decisions are often made before all relevant information arrives, and it is common for information to reveal suboptimal past choices and to facilitate future decision-making 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 employment

[1] Lerman et al. (1999), Oster, Shoulson, and Dorsey (2013) and Persoskie et al. (2014) are among the large literature documenting individuals avoiding relevant medical information when they are at risk of certain diseases or health conditions. There is also a rapidly growing literature studying investors' aversion to financial information. Examples include Karlsson, Loewenstein, and Seppi (2009), Sicherman et al. (2016) and Hilbert et al. (2022). For a survey on information avoidance that draws from multiple disciplines, see Golman, Hagmann, and Loewenstein (2017).

[2] Among others, there are Caplin and Leahy (2001), Kőszegi (2003), Brunnermeier and Parker (2005), Dillenberger (2010) and Bénabou and Tirole (2011).

[3] This literature dates back to Festinger (1957), which was later followed by Frey and Wicklund (1978), Frey and Rosch (1984), Frey and Stahlberg (1986) and Jonas et al. (2001).

[4] Examples more directly related to consumer behavior include Ehrlich et al. (1957) and Brock and Balloun (1967), which observed that consumers may have the tendency to avoid information about products they have considered but did not buy or information about risks of products they have purchased.

opportunities from different industries for a new job might reveal to a worker that better early-career choices could have been made. This novel tradeoff between the regret cost of information and the instrumental value of information is at the core of our model. Our model can thus generate both information avoidance and information-seeking behavior, depending on which one of the two effects is dominant in specific decision situations.

To capture this tradeoff between the regret cost and instrumental value of information, we build a model of decision making under uncertainty that involves choices in three periods. Choices made before the arrival of information open up the possibility of regret, and choices that can be made after the arrival of information are sources for its instrumental value. The uncertainty is captured by a set of objective states of the world and information generally represents what the agent expects to learn about the state of the world during the course of her decision process.

Specifically, we consider an agent whose final choice can be modeled as an (Anscombe-Aumann) act, which is a function that specifies an outcome in each possible state of the world. In period 3 (the final period), the agent chooses an act from a menu (a set of acts) F after the arrival of information. The information is beneficial for this final choice since the agent's choice of act from F can be conditioned on what the information reveals. The menu F, however, was selected by the agent from a set of menus $\mathbb{F}$ in period 2 before the information arrives. Thus, the information could reveal that another menu G that was forgone from $\mathbb{F}$ is actually superior to the chosen menu F. More precisely, if the information reveals that G contains an act g that has higher value than all acts in F, then the agent would regret having chosen F over G. Finally, the agent in our model makes an initial choice in period 1 between different sets of menus, having in mind the subsequent choices of a menu and an act. Therefore, the first dimension of the choice domain in our model corresponds to sets of menus of acts. This modelling approach reflects a common decision procedure in reality in which the agent narrows down her set of options over time before making a final choice. Although it seems to be complicated, it is actually a simplification of standard dynamic decision problems.[5]

There is a second dimension of the choice domain in our model corresponding to information. We interpret it as that the agent has some control over the information that would appear in her three-period decision problem. We formalize this idea by modeling information as an *information structure*. Each information structure consists of a set of signal realizations and a collection of conditional probability distributions describing the

[5]For example, in Kreps and Porteus (1978), an agent's choice in each period could determine her consumption in that period and a set of options she can choose from in the next period. Our model simplifies this kind of problem by restricting the consumption to only take place in the final period.

likelihood for each signal realization to obtain in different states.[6] An agent anticipates that some information will arrive (i.e., some signal realization will be observed) during her decision process and different signal realizations might carry different values to her. Moreover, she can evaluate an information structure from an ex-ante perspective by averaging the values of all possible signal realizations. This second dimension of the choice domain allows us to interpret the agent's behavior as avoiding information, if she picks a less informative information structure, in the sense of the Blackwell order (Blackwell, 1951, 1953), when a more informative one is available.

In sum, we investigate preferences over pairs consisting of a set of menus and an information structure. We show that by simply observing the agent's preference over these pairs, we can determine whether her choices can be modeled *as if* she trades off the regret cost and the instrumental value of the information. Our main result is a representation theorem that features what we refer to as an *informational tradeoff (IT) representation*.

As a building block for the axiomatic characterization of the IT representation, we also consider a subjective informational tradeoff (SIT) representation in which the information structure is a parameter instead of a choice variable. We show that such a representation can be characterized from an agent's preference over sets of menus alone. Moreover, the information structure (as an unobservable parameter) can be elicited from this preference. This characterization is of interest on its own for two reasons. First, conceptually, the dual roles of information in our model do not depend on whether or not the agent has any control over its content. In other words, the agent might still recognize the regret cost and instrumental value of an information structure even if it is exogenously given. Second, from a more pragmatic point of view, the economic modeler may not always be able to observe the information structure anticipated by the agent (even if it is indeed chosen by the agent). It would thus be valuable to be able to elicit the information structure from the agent's consumption behavior.

We now describe the main axioms for our representations.

We have two main axioms for the SIT representation. The first axiom reflects the agent's desire to limit her options for her period-1 choice, since she might experience regret from comparing the menu she has chosen with the counterfactual outcomes represented by the other menus she didn't choose. Formally, suppose F and G are two menus and $\{F\} \succsim \{G\}$. That is, the agent prefers a singleton set containing only menu F to a singleton set containing only menu G. We assume this implies that the agent would (weakly) prefer to choose F over G whenever both F and G are contained in the same set of menus. Our justification for this assumption is from our interpretation that the agent's

[6] This is the (statistical) experiment studied in Blackwell (1951, 1953) and has become a common way to model information.

menu choice from a set of menus is made before any information could arrive. Therefore, the ex-ante comparison between F and G should not depend on whether other menus are present and should be in line with the ranking of the singleton sets $\{F\}$ and $\{G\}$. However, the agent also takes into account the value of each menu after the arrival of the information when evaluating sets of menus. More specifically, adding menu G to a set of menus $\mathbb{F}$ that already contains menu F will always make it (weakly) worse, since G will never be chosen over F from this set of menus but could contain acts that turns out to be better after some signal realization and contribute to the agent's regret. Summarizing the discussion above, the axiom states that if $\{F\} \succsim \{G\}$ and $F \in \mathbb{F}$, then $\mathbb{F} \succsim \mathbb{F} \cup \{G\}$. We refer to this axiom directly as Ex-Ante Regret.[7]

The second main axiom for the SIT representation reflects the agent's preference for a set of menus that allows her to decide later, all else equal. Formally, suppose $\mathbb{F}$ is a set of menus and F, G are two menus. The axiom states that $\mathbb{F} \cup \{F \cup G\} \succsim \mathbb{F} \cup \{F, G\}$. That is, if two sets of menus correspond to the same set of acts that can be ultimately chosen, then the agent would prefer the set that allows her to postpone her decision on which menu to commit to, because doing so would allow her more options to choose from after the arrival of information. We refer to this axiom as Interim Preference for Flexibility.[8]

Other axioms for the characterization of the SIT representation are Weak Order, Continuity, Independence, Finiteness and Domination. These axioms are more standard in the setting of preference over menus and are simply adapted to our setting of preference over sets of menus of acts.

Our axiomatic exercise to characterize the IT representation builds on the axiomatic characterization of the SIT representation. In other words, all axioms used for establishing a SIT representation will also be utilized for the IT representation. Note, however, that these axioms only discipline the agent's choices over sets of menus for some exogenously given information structure. We posit additional axioms that link the agent's consumption choice to her information choice.

[7]Readers familiar with the literature on regret might already notice the similarity of this axiom and the main axiom (dominance) of Sarver (2008). Despite the similarity, the axioms are imposed on different choice domains. Our axiom can be viewed as an adaption from the framework with menus of lotteries to the framework with sets of menus of acts.

[8]Takeoka (2006) posits an axiom similar to ours in a choice domain with menus of menus of lotteries. In the same domain as Takeoka (2006), Kopylov and Noor (2018) and Stovall (2018) each considers an axiom opposite to our axiom, where the agent either prefers to decide earlier or has interim preference for commitment.

1.1 Preview of Results

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

Let Ω be a finite set of states of the world. The agent's uncertainty about Ω is captured by a prior belief $\pi \in \Delta(\Omega)$.[9] An information structure is a mapping $\sigma : \Omega \to \Delta(S)$ where S is a finite set of signal realizations. Let X be a finite set of outcomes. An (Anscombe-Aumann) act is a mapping from Ω to $\Delta(X)$, specifying a lottery over X as the outcome of this act in each state of the world. Let f denote an act. Let F denote a menu of acts and let $\mathbb{F}$ denote a set of such 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 $\sigma(s)$ is the ex-ante probability that signal realization s is generated while μ_s^σ is the Bayesian posterior if realization s is observed.[10] The function $U(F, \mu_s^\sigma)$ captures the material utility of a menu F under posterior μ_s^σ with

$$U(F, \mu_s^\sigma) := \max_{f \in F} \sum_{\omega \in \Omega} \mu_s^\sigma(\omega) u\big(f(\omega)\big) \tag{2}$$

where $u : \Delta(X) \to \mathbb{R}$ is an affine function over lotteries that captures the agent's taste over outcomes. Equation (2) describes that the value of a menu F is the expected value of the best act in it and this contributes to the instrumental value of information.

Finally, $R(F, \mathbb{F}, \mu_s^\sigma)$ captures the agent's regret for having chosen menu F from a set of menus $\mathbb{F}$ at posterior μ_s^σ. Formally,

$$R(F, \mathbb{F}, \mu_s^\sigma) := K \left[\max_{G \in \mathbb{F}} U(G, \mu_s^\sigma) - U(F, \mu_s^\sigma) \right]. \tag{3}$$

That is, the agent's regret is proportional to the difference of the material utility for the menu she has chosen and the highest material utility she could have obtained if she had chosen another menu from $\mathbb{F}$. The intensity of regret is capture by the scalar $K \geq 0$.

In summary, equation (1) specifies that the agent evaluates the value of the pair $(\mathbb{F}, \sigma)$ by the highest expected net value that can be obtained by committing to some menu F in $\mathbb{F}$. The expected net value of a menu F is the weighted average of its net value after each possible signal realization from the information structure σ. This net value is obtained by subtracting the term capturing the regret as defined in equation (3) from the material value of the menu as defined in (2), reflecting the tradeoff between the regret cost and the instrumental value of information.

---

[9] For any finite set Y, we use $\Delta(Y)$ to denote the probability simplex over Y.

[10] Formally, $\sigma(s) = \sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega)$. For any signal realization s with $\sigma(s) > 0$, the posterior $\mu_s^\sigma \in \Delta(\Omega)$ is $\mu_s^\sigma(\omega) = \frac{\pi(\omega)\sigma(s|\omega)}{\sigma(s)}$, and we set the posterior μ_s^σ to be uniform if $\sigma(s) = 0$.

In a subjective information tradeoff (SIT) representation, a set of menus of acts $\mathbb{F}$ is evaluated by

$$V(\mathbb{F}) = \max_{F \in \mathbb{F}} \sum_{s \in S} \sigma(s) \Big[U(F, \mu_s^\sigma) - R(F, \mathbb{F}, \mu_s^\sigma) \Big] \tag{4}$$

where all terms $\sigma(s)$, μ_s^σ, $U(F, \mu_s^\sigma)$ and $R(F, \mathbb{F}, \mu_s^\sigma)$ are defined in the same way as in the IT representation. Therefore, equations (1) and (4) describe the same functional form. However, the important difference is that the information structure σ is a choice variable in equation (1) while it is a parameter in equation (4). We show that we can elicit all parameters in the SIT representation from the preferences over sets of menus, including the information structure.[11]

We now give a numerical example as an illustration for these representations and that our model can generate both information-avoiding and information-seeking behavior.

Example 1. Consider a student choosing between several colleges. College 1 offers an economics major (Econ) and a computer science major (CS) but does not allow any student to double major. College 2 offers only the economics major and College 3 offers both majors and allow students to double major. Suppose that upon graduation, students majoring in economics can choose between two jobs, banking (b) and consulting (c), while a computer science major has only one option to work as a software engineer (e). We thus interpret each major as a menu of jobs, that is, Econ $= \{b, c\}$ and CS $= \{e\}$. And we further interpret each college as a set of majors,[12] that is,

$$\begin{aligned}
\mathbb{F}_1 &= \{\text{Econ}, \ \text{CS}\} = \Big\{ \{b, c\}, \{e\} \Big\} \\
\mathbb{F}_2 &= \{\text{Econ}\} = \Big\{ \{b, c\} \Big\} \\
\mathbb{F}_3 &= \{\text{Econ}, \ \text{CS}, \ \text{Econ} \cup \text{CS}\} = \Big\{ \{b, c\}, \{e\}, \{b, c, e\} \Big\}
\end{aligned}$$

Suppose the student's choices are based on the career prospect associated with each job, but that there is some uncertainty about these careers. Further suppose that this uncertainty can be represented by a binary state space $\Omega = \{\omega_1, \omega_2\}$ that captures the relevant labor market conditions. Suppose the student's prior belief is such that the two states are equally likely, that is, $\pi(\omega_1) = \pi(\omega_2) = 0.5$. Suppose the state-dependent utility derived from the career prospect for each job can be summarized by

[11] Dillenberger, Lleras, Sadowski, and Takeoka (2014) are the first to study the identification of an agent's subjective information from preferences over menus. We use a technique similar to theirs as part of our identification strategy.

[12] Note that we interpret the option of double major to simply mean the student can choose among all three jobs upon graduation. On a different note, $\mathbb{F}_4 = \{\text{Econ} \cup \text{CS}\}$ represents a college that only offers a double major in economics and computer science, which is different from both $\mathbb{F}_1$ and $\mathbb{F}_3$.

u	b	c	e
ω_1	110	100	130
ω_2	60	90	50

Finally, suppose $K \geq 0$ and that the student learns the true state upon graduation. That is, after her major choice but before her job choice, she observes a signal realization generated from a fully revealing information structure described by

σ	s_1	s_2
ω_1	1	0
ω_2	0	1

That is, signal realization s_i will be generated with probability 1 contingent on the state of the world being ω_i. Therefore, observing s_i helps the student to be sure that ω_i is the state of the world.

We first illustrate the SIT representation by computing the value of each college to the student, taking this information structure as fixed. If the student chooses the economics major from $\mathbb{F}_1$, then banking (b) provides the best career prospect if s_1 is observed and consulting (c) is the better choice if s_2 is observed. However, she would feel a sense of regret for majoring in economics if she observes s_1, because she could have $u(e(\omega_1)) = 130$ if she had majored in CS instead of $\max_{f \in \text{Econ}} u(f(\omega_1)) = 110$. Indeed, the material utility (U) and corresponding disutility from regret (R) for choosing each major from $\mathbb{F}_1$ can be summarized by

U	Econ	CS
s_1	110	130
s_2	90	50

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

Since the prior is $\pi(\omega_1) = \pi(\omega_2) = 0.5$, each signal realization is generated with probability 0.5. Therefore, the expected net value for the Econ major in $\mathbb{F}_1$ is $0.5 \times (110 - 20K) + 0.5 \times (90 - 0) = 100 - 10K$. For CS, it is $0.5 \times (130 - 0) + 0.5 \times (50 - 40K) = 90 - 20K$. Since $K \geq 0$, the student would choose Econ over CS in $\mathbb{F}_1$ and the value of College 1 is thus $V(\mathbb{F}_1) = 100 - 10K$.

For College 2, there is only one major, so there is no regret associated with choosing the wrong major.[13] The value for $\mathbb{F}_2$ can be calculated as $V(\mathbb{F}_2) = 0.5 \times (110 - 0) + 0.5 \times (90 - 0) = 100$. Note that as long as $K > 0$, that is, as long as the agent is susceptible

[13] It is fair to ask about the possibility for the student to regret her college choice. We believe it is a natural first step for modeling regret by focusing on the regret generated from more recent decisions.

to feeling regret, then she would choose College 2 over College 1, since by offering the CS major that would never be chosen over the Econ major, College 1 only opens the student up for disutility from regret. This is exactly the motivation for our axiom on Ex-Ante Regret.

Lastly, through similar calculations, we can see that the student would choose to double major in College 3 by choosing "Econ $\cup$ CS", and the value for the college is $V(\mathbb{F}_3) = 0.5 \times (130 - 0) + 0.5 \times (90 - 0) = 110$. This value is higher than both values of College 1 and College 2, as doing a double major allows the student to best use the information about the state of the world, which also allows her to avoid feeling regret. This is exactly the motivation for our axiom on Interim Preference for Flexibility.

We next illustrate the IT representation and show that an agent represented by the IT representation may avoid information in some scenarios but seek information in other scenarios. In our example, the student might want to learn less about the true state of the world because of the anticipated regret. Formally, suppose the statistical experiment can be parameterized by its precision θ, that is, the information about the state of the world can be represented by

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

where $\theta \in [0.5, 1]$. That is, the information is noisy in the sense that even if s_i is observed, the student cannot be sure if ω_i is the true state. But s_i is still indicative about state ω_i since the Bayesian posterior μ_i after observing s_i is $\mu_i(\omega_i) = \theta \geq 0.5$. Suppose the student has chosen College 1 but can control the precision of the information by controlling θ. Our previous calculation shows that $W(\mathbb{F}_1, \sigma_1) = 100 - 10K$. Similar calculations establish that $W(\mathbb{F}_1, \sigma_{0.5}) = 95$. Therefore, as long as $K > 0.5$, the student would prefer a completely noisy information structure, $\theta = 0.5$, to a fully informative information structure, $\theta = 1$. As depicted in Figure 1, the optimal precision could depend on the regret intensity K.

More precisely, when $K = 0.4$ (the blue curve), the student strictly prefers the highest precision ($\theta = 1$) than any other precision. When $K = 0.8$ (the red curve), sticking with the completely noisy information structure ($\theta = 0$) is optimal.

Note both curves are first flat (for $\theta \in [1/2, 4/7]$), then decreasing (for $\theta \in [4/7, 3/4]$) and then increasing (for $\theta \in [3/4, 1]$). In the flat region, the consulting job (c) will have the best expected payoff after either signal realization s_1 or s_2. The agent thus anticipates no regret by choosing the Econ major. In this region, the value of the college does not depend on the precision of the information structure because it always equals to the ex-

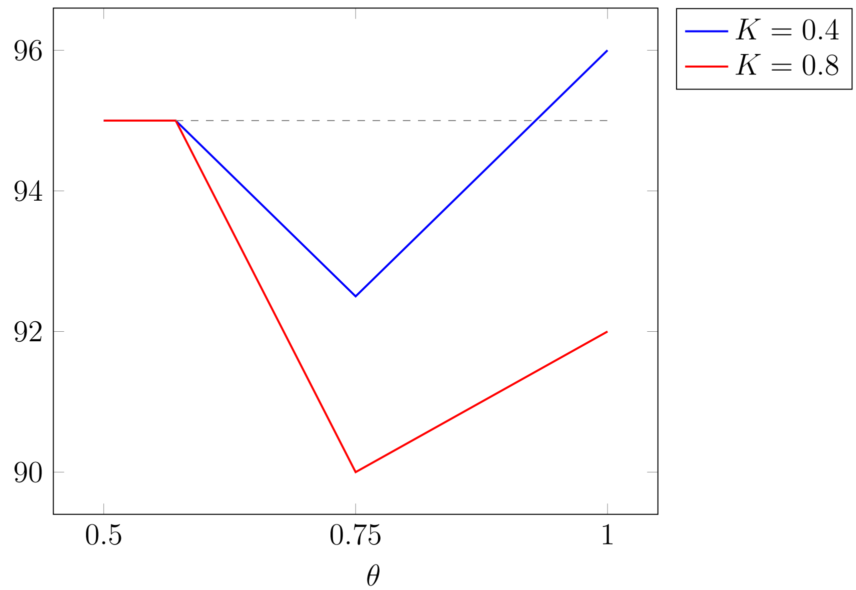

Figure 1: Value of College 1, $W(\mathbb{F}_1, \sigma_\theta)$, as a function of θ

ante expected payoff of the consulting job. In the intermediate region ($\theta \in [4/7, 3/4]$), the information is precise enough for the agent to regret choosing the Econ major but not precise enough for her to choose the banking job (b) after observing signal realization s_1. Therefore, the information carries a positive regret cost but zero instrumental value and the increase in its precision only reduces the value of the college. In the final region ($\theta \in [3/4, 1]$), the information is precise enough for the agent to choose banking (b) after observing s_1. The information thus carries a positive instrumental value and the net value of the information starts to increase with its precision.

The rest of the paper is organized as follows. In Section 2, we set up the model and formally describe our primitives. Section 3 contains our analysis and results regarding the subjective informational tradeoff representation, with the axioms in Section 3.1, the representation theorem in Section 3.2, and the uniqueness result in Section 3.3. In Section 4, we characterize the informational tradeoff representation by considering the larger choice domain where the agent's information choice is also observable. Additional axioms and the representation theorem are presented in Section 4.1. Section 5 concludes the paper by discussing some related literature and two extensions.

2 The Model

2.1 Information Structures

Let Ω be a finite set of states with $|\Omega| \geq 2$. An information structure is a Blackwell experiment with finitely many signal realizations. Formally, an *information structure* is a pair (S, σ) where S is a finite set of signal realizations and σ is a mapping from Ω to $\Delta(S)$. Write $\sigma(s \mid \omega)$ to denote the probability that signal realization s is generated contingent on the state being ω. Different information structures could have different sets of signal realizations. But for convenience, we simply write σ instead of (S, σ). Let $\mathcal{I}$ denote the set of all information structures.

2.2 Acts, Menus and Directions

Let X be a finite set of prizes with $|X| \geq 2$ and $\Delta(X)$ is the set of lotteries over X.[14] 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* is a nonempty compact subset of $\mathcal{F}$, typically denoted by F, G, H. Let $\mathcal{M}$ be the set of all menus. Endow $\mathcal{M}$ with the Hausdorff metric d_h.[15] $\mathcal{M}$ is compact.[16] A *direction* is a nonempty compact subset of $\mathcal{M}$, typically denoted by $\mathbb{F}, \mathbb{G}, \mathbb{H}$. A direction is effectively a set of menus. Let $\mathcal{D}$ be the set of all directions. We endow $\mathcal{D}$ with the Hausdorff metric d_H.[17] $\mathcal{D}$ is also compact.

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 $\alpha \in [0,1]$, define the convex combination of these two menus by $\alpha F + (1-\alpha) G := \{\alpha f + (1-\alpha) g \mid f \in F \text{ and } g \in G\}$. Similarly for any directions $\mathbb{F}, \mathbb{G} \in \mathcal{D}$ and $\alpha \in [0,1]$, their convex combination is defined as

$$\alpha \mathbb{F} + (1-\alpha)\mathbb{G} := \{\alpha F + (1-\alpha) G \mid F \in \mathbb{F} \text{ and } G \in \mathbb{G}\}.$$

[14] Equip $\Delta(X)$ with the usual mixture operation and endow $\Delta(X)$ with the Euclidean metric.

[15] This metric is defined by

$$d_h(F, G) := \max\left\{\max_{f \in F} \min_{g \in G} d(f, g), \max_{f \in G} \min_{g \in F} d(f, g)\right\}.$$

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

[17] This is the Hausdorff metric based on d_h, that is, d_H is defined by

$$d_H(\mathbb{F}, \mathbb{G}) := \max\left\{\max_{F \in \mathbb{F}} \min_{G \in \mathbb{G}} d_h(F, G), \max_{F \in \mathbb{G}} \min_{G \in \mathbb{F}} d_h(F, G)\right\}.$$

2.3 Primitive

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 direction and an information structure. We have in mind an agent facing a three-period decision problem. In period 1, the agent jointly chooses a direction $\mathbb{F}$ and an information structure σ.[18] The agent chooses a menu F from $\mathbb{F}$ in period 2, anticipating the information to arrive after this choice. A signal realization $s \in S$ 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. We do not explicitly model the agent's choices in periods 2 and 3, leaving them as part of the interpretation of the agent's period-1 preference. The timeline is summarized in Figure 2.[19]

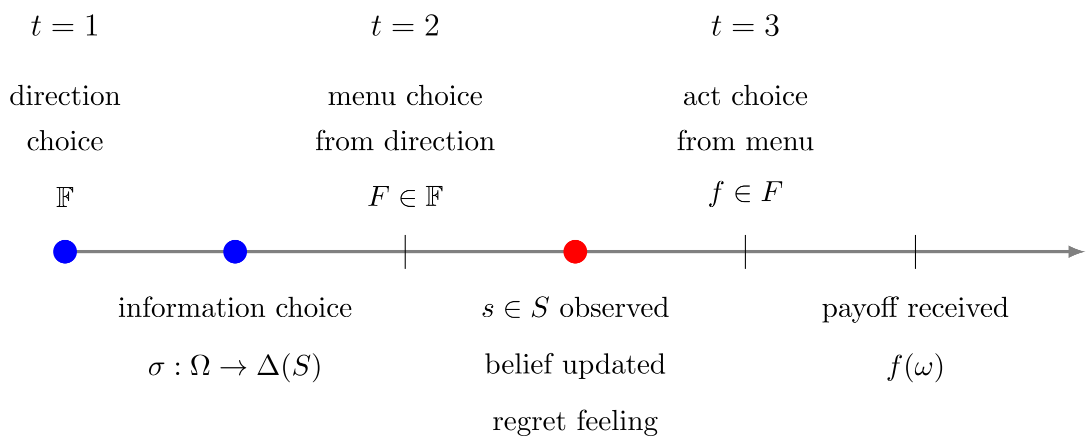

Figure 2: Timeline

We start our analysis by restricting attention to a subdomain. Specifically, we consider a binary relation $\succsim$ on $\mathcal{D}$ representing the agent's preference over directions. In this restricted domain, the modeler does not observe the agent's information choice and has to elicit the information structure anticipated by the agent through her preference over directions. The timeline for choices in this subdomain is identical to the timeline described in Figure 2. The only difference in interpretation is that the information structure is now an unobservable parameter instead of an observable choice variable. Therefore, we sometimes refer to this as the subjective version of our model. The analysis and results for the subjective version of the model are in Section 3. Further building on this analysis and results, the model with the larger choice domain is studied in Section 4.

[18] Our model makes no restriction on the order of the direction choice and the information choice. Either choice can be made first, depending on the decision scenario.

[19] In some applications, it might be more realistic for the information choice to be made after the menu choice. We show that this can be accommodated in a simple extension of our model in Section 5.2.

3 Subjective Informational Tradeoff Representation

We first give the formal definition of a subjective informational tradeoff representation based on the discussion in the Introduction.

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[20] $\sigma : \Omega \to \Delta(S)$ such that $\succsim$ can be represented by the function $V : \mathcal{D} \to \mathbb{R}$ defined by

$$V(\mathbb{F}) = \max_{F \in \mathbb{F}} \sum_{s \in S} \sigma(s) \Big[U(F, \mu_s^\sigma) - R(F, \mathbb{F}, \mu_s^\sigma) \Big] \tag{5}$$

where $\sigma(s) = \sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega)$ is the ex-ante probability of signal realization s, and

- μ_s^σ is the agent's posterior belief after observing s, with

$$\mu_s^\sigma(\omega) = \begin{cases} \frac{\pi(\omega)\sigma(s|\omega)}{\sigma(s)} & \text{if } \sigma(s) > 0 \\ \frac{1}{|\Omega|} & \text{if } \sigma(s) = 0 \end{cases}$$

- $U(F, \mu_s^\sigma)$ is the highest possible expected utility under belief μ_s^σ that can be obtained by choosing some act in menu F. That is,

$$U(F, \mu_s^\sigma) = \max_{f \in F} \sum_{\omega \in \Omega} \mu_s^\sigma(\omega) u\big(f(\omega)\big). \tag{6}$$

- $R(F, \mathbb{F}, \mu_s^\sigma)$ is the regret for having chosen F from $\mathbb{F}$ after observing s, that is,

$$R(F, \mathbb{F}, \mu_s^\sigma) = K \left[\max_{G \in \mathbb{F}} U(G, \mu_s^\sigma) - U(F, \mu_s^\sigma) \right], \tag{7}$$

  where $K \geq 0$ represents the agent's regret intensity.

The interpretation for the SIT representation is just as in the Introduction. When evaluating a direction, the agent anticipates that the information will not arrive until after she makes her menu choice. That is, she will select a menu before observing any signal realization generated from the information structure. On one hand, she can always (weakly) gain from the information by conditioning her choice of act on the signal realizations as captured by equation (6) and this constitutes the information value of information. On the other hand, she could experience regret after some signal realizations if her choice of menu

[20] We write σ instead of (S, σ) to denote information structures for ease of exposition. As we have mentioned before, we allow different information structures to have different sets of signal realizations.

is revealed to be inferior as captured by equation (7). The regret associated with each signal realization is proportional to the gap between the expected value of the best menu in $\mathbb{F}$ and the expected value of the menu she has chosen. The agent's ex-ante value of the direction under the anticipated information is therefore based on the difference between her expectation of the expected value of the chosen menu and her expectation of the regret.

An useful equivalent expression of the SIT representation can be obtained by combining equations (5)-(7):

$$\begin{aligned} V(\mathbb{F}) = \max_{F \in \mathbb{F}} & \left[(1+K) \sum_{s \in S} \max_{f \in F} \sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(f(\omega)) \right] \\ & - K \sum_{s \in S} \max_{G \in \mathbb{F}} \max_{g \in G} \sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(g(\omega)) \end{aligned} \tag{8}$$

Intuitively, the agent in our model chooses a menu from the direction to maximize her expectation of utility minus regret. However, given the information structure, the menu that maximizes her expected utility also minimizes the expectation of her regret. That is, the set of maximizers of equation (5) within a given direction $\mathbb{F}$ will not depend on the value of K. Therefore, although the regret may cause the agent in our model to prefer smaller directions (in terms of set inclusion), it does not distort her choice of menu from a direction since the same set of menus will be the maximizers for the positive term in equation (8) no matter how high the regret intensity level is.

3.1 Axioms

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

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

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

Axiom 3—Independence: *For any $\mathbb{F}, \mathbb{G}, \mathbb{H}$ 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 the reader 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 need to 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 $G \subseteq H \subseteq F$.

The name "critical" is intuitive: If $\mathbb{G}$ is critical for $\mathbb{F}$, then menus in $\mathbb{F}$ but outside $\mathbb{G}$ are all irrelevant for the agent's evaluation of $\mathbb{F}$. The intuition is similar for a menu G being critical for another menu F in a direction $\mathbb{F}$.

Axiom 4—Finiteness: *There exists a natural number N such that the followings hold:*

- *For every $\mathbb{F}$, there exists $\mathbb{G}$ with $|\mathbb{G}| < N$ such that $\mathbb{G}$ is critical for $\mathbb{F}$.*
- *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 only a finite number of menus within any direction matter for evaluating that direction and only a finite number of acts within any menu matter for the evaluation of the direction containing it. Effectively, this restriction reflects the fact the agent is only willing to entertain a finite number of beliefs over Ω as possible posteriors after the information arrival and this helps to guarantee that the anticipated information structure has finitely many signal realizations. 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 allows for the possibility of 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). He studies preferences over menus of lotteries to capture regret generated by an agent's subjective uncertainty about her taste over lotteries. We adapt his dominance axiom to our framework of preferences over directions of acts.

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 direction $\mathbb{F}$. This reflects the agent's desire to limit the size of the direction in some situations. In addition, Axiom 5 specifies that the exact situations in which the agent might exhibit this desire are when the added menu G will never be subsequently chosen from $\mathbb{F} \cup \{G\}$ because an ex-ante better menu F is already contained in $\mathbb{F}$. Intuitively, if G is not chosen in period 2, then it does not benefit the agent to add G to $\mathbb{F}$. But it can harm the agent if some act contained in G turns out to be better after some signal realizations from the anticipated information.

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: *For any direction $\mathbb{F}$ and menus F, G,*

$$\mathbb{F} \cup \{F \cup G\} \succsim \mathbb{F} \cup \{F, G\}.$$

Note that the two directions in the comparison share 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. Specifically, the agent can choose $F \cup G$ from the direction on the left hand side and choose acts from both F and G after observing any signal realization. But she has to decide between choosing F or G from the direction on the right hand side before the information could arrive. If she selects F, then she has to let go acts that are in G but not in F. Those acts could turn out to be better after certain signal realizations, potentially contributing to both lower payoff in the future and higher regret about the past. Same goes for choosing G. Other than the part about regret, this is the standard argument for the preference for flexibility as coined in Kreps (1979). This axiom differs from the standard argument for the preference for flexibility, however, since our agent not necessarily prefers a direction that contains larger sets.[21]

Stovall (2018) considers a related axiom in the setting of preferences over menus of menus of lotteries. His axiom, called "interim preference for commitment," states the opposite of our Axiom 6. That is, the agent prefers adding two menus separately to a set of menus comparing to adding their union. In his model, interim preference for commitment is used to capture the fact that the agent is subject to temptation in the intermediate stage.

Our next axiom is the standard nontriviality statement that the agent has strict preference over some pairs of outcomes.

Axiom 7—Nontriviality: *There exist lotteries $\ell, \ell' \in \Delta(X)$ with $\ell \succ \ell'$.*[22]

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

Definition 3:

- Say that an act f *dominates* another act g if $f(\omega) \succsim g(\omega)$ for all $\omega \in \Omega$.
- Say that a menu F dominates another menu G if for any $g \in G$, there exists $f \in F$ such that f dominates g.
- Say that a direction $\mathbb{F}$ dominates another direction $\mathbb{G}$ if for any $G \in \mathbb{G}$, there exists $F \in \mathbb{F}$ such that F dominates G.

[21] That is, we do not impose any axiom like " $\mathbb{F} \cup \{F\} \succsim \mathbb{F} \cup \{G\}$ whenever $G \subseteq F$."

[22] We follow a common practice of abusing notation by using ℓ to also denote the constant act that yields lottery ℓ in every state. For simplicity, we also write $\ell \succ \ell'$ for $\{\{\ell\}\} \succ \{\{\ell'\}\}$.

By definition, domination is built on state-by-state comparisons between acts. Therefore, if f dominates g, then the agent will never choose g over f under any belief over Ω. Similarly, if a menu F dominates another menu G, then F will have higher material value than G under any belief. Therefore, there will never be any regret about not choosing G from a direction also containing F. These observations motivate the next axiom.

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 reflects a sense of monotonicity of the agent's preference. If f is better than g in every state, then f itself should be preferred to g. The second half of part 1 states that adding a dominated act to a menu does not change her attitude toward the menu, because a dominated act contributes to neither the material value of the menu nor the regret. Similarly for part 2, if $\mathbb{F}$ dominates $\mathbb{G}$, then any menu $G \in \mathbb{G}$ is dominated by some menu $F \in \mathbb{F}$. Therefore, adding this collection of dominated menus to $\mathbb{F}$ will not change the agent's choice of menu from $\mathbb{F}$. It will not change the agent's anticipated regret, either. So the agent is indifferent between $\mathbb{F}$ and $\mathbb{F} \cup \mathbb{G}$.

3.2 Representation Theorem

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 for the SIT representation. We first prove a representation theorem (Theorem 6) for a preference over menus of menus of lotteries that is closely related. For convenience, we refer to a menu of menus of lotteries as a direction of lotteries. The setup for this related choice domain and the results are contained in Appendix A. This representation theorem features what we refer to as a "partial regret" representation. It can be viewed as an extended version of the regret representation in Sarver (2008) with three time periods. The agent chooses a direction of lotteries in period 1 and then selects a menu from this direction in period 2. Both choices are made before her subjective uncertainty about her taste over lotteries are resolved. She learns about her taste after the menu choice but before the lottery choice in period 3. This could make her regret her menu choice, but she can also choose the best lottery from this menu according to the revealed taste. Different from Sarver (2008), the agent still has some flexibility after the uncertainty about her taste for lotteries is resolved.

The second step is to establish a connection between the respective choice domains for the SIT representation and the partial regret representation. To do so, we build on a technique used in Dillenberger, Lleras, Sadowski, and Takeoka (2014, henceforth DLST) which involves a sequence of geometric arguments that connects a preference over menus of acts to a preference over menus of lotteries. Our proof involves connecting a preference over directions of acts to a preference over directions of lotteries.

We now apply the SIT representation to two specific types of directions.

Definition 4: For any menu F, let $D(F) := \{\{f\} \mid f \in F\}$. We say a direction $\mathbb{F}$ is an *early-commitment direction* if $\mathbb{F} = D(F)$ for some F.

The operation $D(F)$ is one natural way to make a menu F into a direction: We first collect each element $f \in F$ into a singleton menu and then collect all these singleton menus into a direction. In our model, this correspond to the case where the agent has to commit to a final choice (i.e., an act) before she observes any signal realization from the information structure. In other words, the agent is choosing between acts based on her prior. Facing early-commitment directions, information carries zero instrumental value and only contributes to an agent's regret. This is captured by our representation, since when restricting to early-commitment directions,

$$V\big(D(F)\big) = \max_{f\in F}\left[(1+K)\sum_{\omega\in\Omega}\pi(\omega)u(f(\omega))\right] - K\sum_{s\in S}\max_{g\in F}\sum_{\omega}\pi(\omega)\sigma(s\mid\omega)u(g(\omega)). \quad (9)$$

This corresponds to a finite version of the regret representation characterized in Sarver (2008) in the framework with acts. Sarver (2008) studies preferences over menus of lotteries and the agent's regret is from her uncertainty about future tastes for lotteries. Despite the similarity in the way to model regret, his paper does not emphasize the mechanism for the agent's taste change. On the contrary, we focus on the interpretation that information is the driving force for the agent's change in how she evaluates acts.

We say that a direction $\mathbb{F}$ is a *singleton direction* if $\mathbb{F} = \{F\}$ for some menu F. This corresponds to an alternative way of embedding the set of menus into the set of directions. When restricting to singleton directions, the SIT representation reduces to

$$V\big(\{F\}\big) = \sum_{s\in S}\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega)). \quad (10)$$

That is, a singleton direction corresponds to the case where the agent can make all her relevant decisions after observing a signal realization from the information structure, and this corresponds to a finite version of the subjective learning representation characterized in DLST.

3.3 Uniqueness of the SIT Representation

Recall that a subjective informational tradeoff representation for $\succsim$ has four parameters (π, u, K, σ), where π is the prior over Ω, u is an affine function over outcomes, K is a non-negative scalar and σ is an information structure.

Definition 5: 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 ν satisfies that for any $\mu \in \Delta(\Omega)$,

$$\nu(\mu) = \sum_{s \in S} \mathbf{1}[\mu = \mu_s^\sigma]\sigma(s),$$

where $\mathbf{1}(\cdot)$ is the indicator function, and $\sigma(s)$ and μ_s^σ are the ex-ante probability of observing s and Bayesian posterior after observing s, respectively.

Since the agent is Bayesian, the induced distribution over posteriors always averages back to the prior, that is, $\sum_{\mu \in \text{supp}(\nu)} \nu(\mu)\mu(\omega) = \pi(\omega)$ for any $\omega \in \Omega$. We say an information structure σ induces a *degenerate* distribution over posteriors if the induced distribution puts weight 1 on the prior belief.

Note from equation (5) 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 $\mathcal{D}$.

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 given the prior belief* π.
- $K_0 = K$ *if* σ *induces a nondegenerate distribution over posteriors. If* σ *induces a degenerate distribution over posteriors, then* $K_0, K \in [0, \infty)$.

Proof. See Appendix C.1. □

The identifications of the prior belief π and the taste u follow directly from the uniqueness result of the standard subjective expected utility model studied in Anscombe and Aumann (1963). One quick and intuitive way to understand our uniqueness result on the information structure σ and the regret intensity K is through the uniqueness results from Sarver (2008) and DLST.

When restricting attention to early-commitment directions (i.e., $\mathbb{F} = D(F)$ for some menu F), our representation reduces to a finite version of the regret representation in Sarver (2008) in the context of acts. This is the representation described in equation (9). Despite the difference in the primitive, we can think of applying Theorem 4 of Sarver (2008) to jointly identify σ and K as long as σ does not induce a degenerate distribution over posteriors.[23] However, we run into a similar issue as in Sarver (2008) when we try to separately identify σ and K using only the representation in equation (9). The issue is that we might not always be able to distinguish between the following two agents. The first is an agent 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 the second is an agent 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 over directions according to equation (9) if σ_2 is obtained by halving the probability of each signal realization in σ_1 in every state of the world and adding an extra signal realization that is generated with probability 0.5 in every state.

We can overcome this issue and achieve separate identification of the information structure σ and the regret intensity K by taking advantage of our larger choice domain. Specifically, we turn to the singleton directions. When restricting attention to singleton directions (i.e., $\mathbb{F} = \{F\}$ for some menu F), our representation reduces to a finite version of the subjective learning representation in DLST. This is the representation described in equation (10). DLST show that the information structure (modeled as a distribution over posteriors) can be uniquely identified. This helps us to separately identify σ and K, as long as σ does not induce a degenerate distribution over posteriors.

We can make a sharper statement on the uniqueness of the information structure in the SIT representation, provided that the identified prior belief π has full support. To make the statement, we formally introduce the notions of garbling and Blackwell equivalence.

Definition 6: Let $\sigma_1 : \Omega \to \Delta(S_1)$ and $\sigma_2 : \Omega \to \Delta(S_2)$ be two information structures. Say that σ_2 is a *garbling* of σ_1 if there exists $\phi : S_1 \to \Delta(S_2)$ such that

$$\sigma_2(s_2 \mid \omega) = \sum_{s_1 \in S_1} \phi(s_2 \mid s_1)\sigma_1(s_1 \mid \omega)$$

for all $s_2 \in S_2$ and $\omega \in \Omega$.

In words, σ_2 is a garbling of σ_1 if σ_2 can be obtained by adding some noise to the information structure σ_1. Blackwell's theorem (Blackwell, 1951, 1953) establishes that σ_2

[23] When it does, the agent anticipates zero regret from any direction because her evaluations for acts before and after the arrival of information are exactly the same. In such a case, all regret intensity levels will generate the same preference over directions.

is a garbling of σ_1 if and only if σ_1 is *more informative than* σ_2, where "σ_1 being more informative than σ_2" means that every Bayesian expected utility maximizer will weakly prefer σ_1 to σ_2 in every standard decision problem. And we say σ_1 is Blackwell equivalent to σ_2 if σ_1 is more informative than σ_2 and σ_2 is more informative than σ_1.

COROLLARY. *Suppose both* $(\pi_0, u_0, K_0, \sigma_0)$ *and* (π, u, K, σ) *represent* $\succsim$*. If* π *has full support, then* σ_0 *is Blackwell equivalent to* σ*.*

That is, we can identify the information structure up to its equivalence class specified by the Blackwell informativeness order when the identified prior has full support.

The notion of one information structure being more informative than another will also be used to capture the meaning of information avoidance in our model.

4 Informational Tradeoff Representation

We now move on to formally establish the informational tradeoff representation. First, we need to come back to the choice domain that involves both a choice of direction and a choice of information structure. Let $\succsim$ be a binary relation over $\mathcal{D} \times \mathcal{I}$,[24] where $\mathcal{D}$ is the set of all directions and $\mathcal{I}$ is the set of all information structures.[25] The interpretation for this binary relation is the same as discussed in Section 2.3.

DEFINITION 7: 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

$$\begin{aligned} W(\mathbb{F}, \sigma) = \max_{F \in \mathbb{F}} & \left[(1+K) \sum_{s \in S} \max_{f \in F} \sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(f(\omega)) \right] \\ & - K \sum_{s \in S} \max_{G \in \mathbb{F}} \max_{g \in G} \sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(g(\omega)) \end{aligned} \tag{11}$$

Equation (11) is expressed in the same fashion as equation (8): There is no direct reference to the material utility U of a menu or the regret cost R. But one can still see

[24] This is an abuse of notation as we also use $\succsim$ to denote a binary relation over $\mathcal{D}$. We will explicitly specify the domain whenever we use $\succsim$ from this point on.

[25] Jakobsen (2021) studies the axiomatic foundation for persuasion models using a related primitive. The modeler in his model observes a sender's preference over information structures indexed by menus of acts and a receiver's choice correspondence from menus indexed by signal realizations. Our papers are similar in that both can be used to study preference over information, but our primitive is different in that the modeler in our model observes the agent's preferences over pairs of directions and information structures instead of the conditional preferences along each single dimension of the choice domain.

that the first term captures the instrumental value of information since the agent can choose different acts after different signal realizations, and the second term corresponds to the regret cost.

The IT representation allows richer interpretations of our model comparing to the SIT representation. Fixing an information structure σ, the IT representation fully describes a preference over directions as in the SIT representation. Fixing a direction $\mathbb{F}$, we can now examine the agent's preference over information structures using the IT representation. This enables us to capture the information avoidance behavior of an agent. More precisely, we would say an agent *avoids information* if she exhibits a strict preference for a less informative information structure over an more informative one. The IT representation also allows the comparison of directions across different information structures.

To characterize the IT representation, we first connect a preference over $\mathcal{D} \times \mathcal{I}$ to a collection of preferences over $\mathcal{D}$ indexed by the information structures.

DEFINITION 8: Fix a binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ and an information structure $\sigma \in \mathcal{I}$, the *conditional preference* $\succsim^\sigma$ is the binary relation over $\mathcal{D}$ defined by

$$\mathbb{F} \succsim^\sigma \mathbb{G} \text{ if } (\mathbb{F}, \sigma) \succsim (\mathbb{G}, \sigma). \tag{12}$$

Through the conditional preferences, we can build on top of the axioms identified for the SIT representation to characterize the IT representation. All our previous axioms (Axioms 1-8) are imposed on a specific conditional preference and thus have no bite on the agent's choices between directions across different information structures. Our additional axioms will put restrictions that link the agent's direction choice and information choice.

4.1 Additional Axioms and Representation Theorem

Let $\succsim$ be a binary relation over $\mathcal{D} \times \mathcal{I}$, and let $\succsim^\sigma$ be the conditional binary relation over $\mathcal{D}$ induced by $\succsim$ and σ. On top of Axioms 1-8, we impose five additional axioms on $\succsim$ to characterize the IT representation. The first one is the standard rationality requirement for this preference over the larger choice domain.

AXIOM 9—Weak Order: $\succsim$ *is complete and transitive.*

Note that if a binary relation $\succsim$ over $\mathcal{D} \times \mathcal{I}$ is complete and transitive, then each conditional preference $\succsim^\sigma$ must also be complete and transitive. Therefore, Axiom 9 can be viewed as expanding Axiom 1.

The next axiom imposes an independence requirement on mixing directions 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* $\alpha\in(0,1)$*,*

$$(\mathbb{F},\sigma)\succsim(\mathbb{G},\sigma')\iff(\alpha\mathbb{F}+(1-\alpha)h,\sigma)\succsim(\alpha\mathbb{G}+(1-\alpha)h,\sigma') \tag{13}$$

Given the interpretation of our model, the agent's preference over acts (as reflected through the preference for directions like $\{\{f\}\}$) should not depend on the information structure. This is because when the agent has only one act that she can choose, the information can neither help her make better choices in the future nor cause her to regret her choice in the past. Therefore, even though we never formally defined the convex combination of two pairs, we can interpret the pair $(\alpha\mathbb{F}+(1-\alpha)h,\sigma)$ as $\alpha(\mathbb{F},\sigma)\oplus(1-\alpha)(h,\sigma)$. This could naturally be interpreted as that the agent has chosen σ as her information structure but could either face $\mathbb{F}$ or a single act h. Then the motivation for Axiom 10 is the same as the motivation for standard independence axioms.

To state the other axioms, some additional notions are needed. First, an information structure $\sigma\in\mathcal{I}$ is *null* (or *uninformative*) if it always induces a degenerate distribution over posteriors under any prior. It is uninformative in the sense that the agent's posterior belief always equals to the prior facing such an information structure. Let o denote a specific null information structure defined by $o:\Omega\to\Delta(\{s_o\})$. That is, this information structure has only one signal realization s_o and this signal realization is obtained with probability one no matter what the state of the world is. Our next axiom disciplines the agent's behavior when the observed information is exactly o.

Axiom 11—Strategic Rationality when No Information (SRNI): *For any menus* F,G*,*

$$\{F\}\succsim^o\{G\}\Longrightarrow\{F\cup G\}\sim^o\{F\}.$$

When the agent has chosen the information structure o, there will be no regret. She would evaluate any menu according to the expected value of its best alternative since her choices will be made based on the prior belief. Axiom 11 reflects this behavior with the standard strategic rationality statement: If F is preferred to G given o, then F must contain an act that is better than every act in G. Therefore, adding the acts in G to F should not change the agent's attitude toward F.

Suppose $\succsim^o$ has a SIT representation, we will show that Axiom 11 (SRNI) guarantees that the information structure identified from $\succsim^o$ agrees with o (i.e., they both induce a degenerate distribution over posteriors). This result will give us an anchor to pin down other information structures.

Any information structure that is Blackwell equivalent to o will always induce a degenerate distribution over posteriors, so o is not the only null information structure. However, we only need to impose the SRNI axiom on the preference over directions conditional on o

because the rest of our axioms will imply that Blackwell equivalent information structures induce same conditional preferences over directions.

Let $\sigma : \Omega \to \Delta(S)$ be an information structure and $F \in \mathcal{M}$ be a menu. A *plan* is a mapping $\gamma : S \to F$, with the interpretation that $\gamma(s)$ is the act chosen by the agent if she observes signal realization s. Intuitively, a plan describes the agent's commitment on how to react to the information. For example, if F is a singleton with $F = \{f\}$, then there is only one plan for any S. The plan involves choosing f after every possible signal realization, because it is the only option. For another example, if $F = \{f, g\}$ and $S = \{s_1, s_2\}$, then the agent has four different plans. The first plan is to choose f after s_1 and choose g after s_2 (we write fg in short). The other three are ff, gf and gg.

Let F^S denote the set of all plans (for a fixed information structure and a menu). 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

$$\gamma_\sigma(\omega) := \sum_{s \in S} \sigma(s \mid \omega)[\gamma(s)](\omega). \tag{14}$$

That is, γ_σ is obtained by reducing the "compound act" described by γ and σ to an act by averaging over different signal realizations. To illustrate, consider again the example where $F = \{f, g\}$ 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 realization 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 thus obtained by taking the convex combination $\sigma(s_1 \mid \omega)f(\omega) + \sigma(s_2 \mid \omega)g(\omega)$ state by state, as summarized by equation (14).

Definition 9: Fix a menu F and some information structure σ, *the menu induced by F and σ* is defined by

$$F_\sigma := \left\{\gamma_\sigma \mid \gamma \in F^S\right\}. \tag{15}$$

That is, F_σ is obtained by collecting all the acts that can be induced by the information structure σ and some plan $\gamma \in F^S$.

Axiom 12—Reduction: *For any menu $F \in \mathcal{M}$ and any information structure $\sigma \in \mathcal{I}$,*

$$(\{F\}, \sigma) \sim (\{F_\sigma\}, o).$$

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 realization). Therefore, the induced menu is $F_\sigma = \{f\}$ for any information structure σ. Axiom 12 thus implies that

$(f, \sigma) \sim (f, o)$ for any act f and information structure σ.[26] 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. This property of the agent's behavior gives us a common ground to align the conditional preferences indexed by different information structures.

Recall that a early-commitment direction is a direction $D(F) = \{\{f\} \mid f \in F\}$. Our last axiom concerns with mixtures of early-commitment directions $D(F)$ and their corresponding singleton directions $\{F\}$. As it turns out, their interactions have a clear implication on the regret intensity of an agent.

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

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

Axiom 13 describes the existence of a balance point between the two specific types of directions, the early-commitment directions and the singleton directions. The implicit assumption behind this axiom is as follows. If the agent strictly prefers σ_1 to o when facing a singleton direction $\{F\}$, then the agent will strictly prefer o to σ_1 when facing the corresponding early-commitment direction $D(F)$. 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 of this option and the option to face $\{F\}$ with no information at all. Moreover, Axiom 13 states that this balance point is the same across all menus F and information structures σ as long as the agent strictly benefits from σ compared to no information when facing $\{F\}$.

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.

[26]Similar to what we do in the previous section, we abuse notation a little by writing (f, σ) for the pair $(\{\{f\}\}, \sigma)$.

We now state our representation theorem for the IT representation.

THEOREM 3: *A binary relation $\succsim$ over $\mathcal{D}\times\mathcal{I}$ has an informational tradeoff representation if and only if the following conditions are both satisfied:*

- $\succsim$ *satisfies Weak Order, Act Independence, SRNI, Reduction and Balance.*
- $\succsim^\sigma$ *satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$.*

Proof. See Appendix C.5. □

4.2 Proof Sketch for Theorem 3

In this section, we formally state the intermediate results corresponding to the steps hinted in the previous section that help us to build the IT representation.

DEFINITION 10: A binary relation $\succsim$ over $\mathcal{D}\times\mathcal{I}$ has an *aligned informational tradeoff representation* if there exists a tuple $\big(\pi, u, (K^\sigma)_{\sigma\in\mathcal{I}}, (i^\sigma)_{\sigma\in\mathcal{I}}\big)$ that consists of a probability measure 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

$$\begin{aligned} W(\mathbb{F},\sigma) = \ &\max_{F\in\mathbb{F}} \left[(1+K^\sigma)\sum_{t\in S^\sigma}\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)i^\sigma(t\mid\omega)u(f(\omega))\right] \\ &- K^\sigma\sum_{t\in S^\sigma}\max_{G\in\mathbb{F}}\max_{g\in G}\sum_{\omega\in\Omega}\pi(\omega)i^\sigma(t\mid\omega)u(g(\omega)) \end{aligned} \tag{16}$$

This is called the "aligned" IT representation because the collection of IT representations for each σ is aligned together by sharing the same prior belief π and the same taste over outcomes u.

AXIOM 14—Stable Preference over Acts: *For any f, g and any σ, σ',*

$$(f,\sigma)\succsim(g,\sigma) \iff (f,\sigma')\succsim(g,\sigma')$$

Axiom 14 is motivated by the role of information in the decision process. When there is only one act that the agent can choose, any information is irrelevant for the decision since the agent can neither benefit from the information nor be hurt from any regret caused by the information.

LEMMA 1: *A binary relation $\succsim$ over $\mathcal{D}\times\mathcal{I}$ has an aligned informational tradeoff representation if and only if the following conditions are both satisfied:*

- $\succsim$ *satisfies Weak Order, Stable Preference over Acts and Act Independence.*
- $\succsim^\sigma$ *satisfies Axioms 2-8 for each* $\sigma \in \mathcal{I}$.

Proof. See Appendix C.2. $\square$

That is, by imposing Weak Order, Stable Preference over Acts and Act Independence on top of the axioms that ensure that each conditional preference $\succsim^\sigma$ has a SIT representation, we guarantee the existence of a utility representation for a preference over the two dimensional choice domain and at the same time make sure that the identified prior belief and taste over outcomes are information-independent.

LEMMA 2: *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. See Appendix C.3. $\square$

That is, imposing Axiom 11 on an aligned IT representation guarantees that the conditional preference $\succsim^o$ labeled with o indeed corresponds to a situation where the agent is anticipating a null information structure.

DEFINITION 11: A binary relation $\succsim$ over $\mathcal{D}\times\mathcal{I}$ has a *regret-varying informational tradeoff 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

$$\begin{aligned} W(\mathbb{F}, \sigma) = \max_{F\in\mathbb{F}} &\left[(1 + K^\sigma) \sum_{s\in S} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega)u(f(\omega))\right] \\ &- K^\sigma \sum_{s\in S} \max_{G\in\mathbb{F}} \max_{g\in G} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega)u(g(\omega)) \end{aligned} \tag{17}$$

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. Thus, if an agent's preference can be represented by a regret-varying IT representation, then we can model her behavior *as if* she is choosing between information structures taking into account the tradeoff of the benefit for future choices and the regret for past choices. It is not yet the IT representation because the regret intensity level could still vary across different information structures, thus the name "regret-varying."

LEMMA 3: *A binary relation* $\succsim$ *over* $\mathcal{D} \times \mathcal{I}$ *has a regret-varying informational tradeoff representation if and only the following conditions are both satisfied:*

- $\succsim$ *satisfies Weak Order, Act Independence, SRNI and Reduction;*
- $\succsim^\sigma$ *satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$.*

Proof. See Appendix C.4. □

Finally, adding Axiom 13 (Balance) on top of a regret-varying IT representation will make sure that the parameters for regret intensity identified from all conditional preferences over directions are the same.

4.3 Uniqueness of the IT Representation

As we have seen in Theorem 2, a threat to the identification of the regret intensity parameter is non-variation between the agent's prior and posterior beliefs. Since the modeler can observe the agent's information choice in the IT representation, such non-variation could only when the agent's prior belief is degenerate.

Definition 12: We say that a preference $\succsim$ over $\mathcal{D} \times \mathcal{I}$ *has a trivial preference over information* if $(\mathbb{F}, \sigma) \sim (\mathbb{F}, \sigma')$ for all $\mathbb{F} \in \mathcal{D}$ and $\sigma, \sigma' \in \mathcal{I}$. The preference $\succsim$ *has a non-trivial preference over information* if it does not have a trivial preference over information.

That is, a trivial preference for information means the agent is indifferent between any two information structures regardless of the direction she faces. The following lemma fully characterize the connection between a degenerate prior and the agent's preference for information.

Lemma 4: *Suppose $\succsim$ has an informational tradeoff representation with parameters (π, u, K). The prior belief π is degenerate if and only if $\succsim$ has a trivial preference for information.*

Proof of Lemma 4. Only 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.

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 $\{\{f, g\}\}$. □

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 4). Then the identification of the regret intensity K follows from Theorem 2. □

4.4 Comparing Information Aversion Attitudes

In this section, we focus our attention on preferences that have non-trivial preference 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 13: We say that $\succsim_1$ is *more information averse than* $\succsim_2$ if

$$(\mathbb{F}, \sigma) \succsim_2 (\mathbb{F}, \sigma') \Longrightarrow (\mathbb{F}, \sigma) \succsim_1 (\mathbb{F}, \sigma') \tag{18}$$

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, agent 1 is more information averse than agent 2 if whenever agent 2 prefers to avoid information, agent 1 will also prefer to avoid information.

We first focus our attention on comparing the information aversion attitude between agents who agree on their preference over the acts.

DEFINITION 14: We say that $\succsim_1$ and $\succsim_2$ *agree on their preference over acts* if $(f, \sigma) \succsim_1 (g, \sigma) \iff (f, \sigma) \succsim_2 (g, \sigma)$ for every $f, g \in \mathcal{F}$ and some $\sigma \in \mathcal{I}$.

Note that the agent's preference over acts is not affected by the information structure. We now can state our comparative statics result.

THEOREM 5: *Suppose $\succsim_1$ and $\succsim_2$ agree on their preference for acts, then the following are equivalent:*

1. *$\succsim_1$ is more information averse than $\succsim_2$.*
2. *$K_1 \geq K_2$.*

Proof. See Appendix C.6. □

5 Discussion and Extensions

We conclude by discussing some related literature and presenting two extensions of our model. In Section 5.1, we discuss related work on regret and information avoidance. In Section 5.2, we present an extension of our model to allow for the information choice to be made after the menu choice. In Section 5.3, we present another extension of our model to discuss the effects of allowing the agent to regret her choice of act.

5.1 Related Literature

Our paper contributes to the literature on regret and the literature on information avoidance.

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 (earlier examples include Bell (1982), Loomes and Sugden (1982, 1987) and Sugden (1993)). Sarver (2008) is the first to unveil the possibility of identifying regret through preferences over menus. Through the investigation of preferences over menus of lotteries, 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 a 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 utility representation similar to Sarver (2008) to explain choice overload where an agent tends to stick more with some default option when there are more options to choose from. Their key assumption is that regret might be asymmetric in the sense that sticking with the default option does not generate regret even if it later turns out to be inferior comparing to some other options, while deviating from the default option opens the agent up to regret. In both papers, the only relevant choice by the agent is made before their subjective uncertainty about their taste is resolved, so 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. Indeed, this tradeoff between the regret cost of information on past choices and the instrumental value of information through making better-informed future choices is at the core of our analysis. Krähmer and Stone (2013) apply a utility representation similar to the regret representation in Sarver (2008) to argue that ambiguity aversion as captured in Ellsberg's paradox could be explained as an aversion to anticipated regret, because drawing from an urn with an unknown composition of balls opens the agent up to regret while drawing from an unambiguous urn does not generate any regret. They touch upon the possibility of using regret to explain information avoidance by interpreting ambiguity aversion as a version of information aversion. With a direct focus on explicit information choice, our paper provides a systematic analysis of this issue. Moreover, our axiomatic approach uncovers the behavioral foundations for regret-driven information avoidance.

We are not the first to attempt to build a theoretical model that can account for information avoidance either. Information avoidance generally has many different appearances in different contexts, and some of these behaviors can be accounted for by existing theories.[27] 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) into the agent's utility function. This framework, even though non-axiomatic, is very general and allows the possibility for information avoidance driven by different anticipatory feelings, and one of them is anxiety. 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 the possibility of a 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 could break their unwarranted optimistic beliefs.[28] 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.

An important feature that distinguishes our paper from all these works is that we develop a model to formally link information avoidance with past choices. As we have argued in the Introduction, establishing this link is important because there is a great amount of evidence that people are more likely to avoid information after they make a relevant choice.

Two other relevant papers where past choices could influence preferences are Bénabou and Tirole (2011) and Eyster, Li, and Ridout (2021). Bénabou and Tirole (2011) considers a three-period model that generalizes Caplin and Leahy (2001). Their emphasis is that beliefs are treated as assets by some economic agents because those beliefs are valuable in building an agent's identity. Such agents might dislike information that could contradict

[27] Golman, Hagmann, and Loewenstein (2017) also contains a survey of these theoretical models.

[28] Oster, Shoulson, and Dorsey (2013) uses this theory to reconcile their empirical findings about individuals at risk of having Huntington disease avoiding simple and cheap genetic tests designed to reveal whether they have the disease until after symptoms start to show.

the belief they have invested in building. Eyster, Li, and Ridout (2021) builds an axiomatic model to study ex-post rationalization. They study agents who may distort their future choices in order to justify past choices. They thus focus on the specific type of information structure that fully reveals the state of the world after the agent's initial choice. Our paper is different from these two papers in that we build the preference for information directly into the choice domain, which allows us to formally study the connection between an agent's consumption choice and information choice.

5.2 Choosing Information after Choosing a Menu

An important feature of the interpretation for the IT representation is that the agent jointly chooses a pair of a direction and an information structure. In particular, as discussed in Section 2, the choice of information is made before the choice of menu from the chosen direction. This feature may give the impression that our agent has to commit to a specific information acquisition strategy long before she makes her subsequent choices from a direction, which is somewhat troubling in some potential applications.

To illustrate, consider our Example 1 about a student's choice regarding information about different jobs. In our current interpretation for the IT representation, the student needs to decide on the information structure (i.e., θ, the precision of the noisy signal) and commit to it before she could decide her major. The example would be more convincing if the student's choice about information is made after her major choice.

In this section, we discuss a simple extension of our model that addresses this concern. In this extension, we can interpret the representation as if the agent is choosing an information structure after choosing a menu from the direction. Consider the agent choosing a compact set of information structures, Σ, in period one. That is, Σ is a compact subset of $\mathcal{I}$. This set Σ should be interpreted as the agent's partial commitment about her future information acquisition strategies. Suppose the agent has a preference over pairs of directions and sets of information structures that can be represented by[29]

$$\max_{\sigma\in\Sigma} W(\mathbb{F},\sigma) = \max_{\sigma\in\Sigma}\max_{F\in\mathbb{F}} \sum_{s\in S_\sigma} \sigma(s)\left[\max_{f\in F}\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)u\big(f(\omega)\big) - R(F,\mathbb{F},\mu_s^\sigma)\right] \tag{19}$$

$$= \max_{F\in\mathbb{F}}\max_{\sigma\in\Sigma} \sum_{s\in S_\sigma} \sigma(s)\left[\max_{f\in F}\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)u\big(f(\omega)\big) - R(F,\mathbb{F},\mu_s^\sigma)\right] \tag{20}$$

[29]We will not present an complete axiomatic treatment for this representation since we are only using this for comparison purposes. But one relatively straightforward way to axiomatize this is to impose the standard strategic rationality axiom on the conditional preference for sets of information structures in addition to the other axioms posited in the main text.

where $\sigma(s)$ and μ_s^σ are defined as before, and

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

The fact that we can exchange the order of the two maximum operators follows from our agent being Bayesian and forward-looking. And this allows us to interpret this extended model as if the agent chooses her information structure after the menu choice.

5.3 Regretting the Choice of Act

Another important property of the informational tradeoff representation is that the agent only experiences regret once and the regret is only about her choice of menu. This property seems to indicate that the agent in our model cannot regret her choice of act even if she finds out about the true state of the world after she receives her actual payoff.[30] Even though this assumption seems to suggest an inconsistency in our modeling approach, it helps us cleanly isolate the tradeoff of the conflicting effects of information from other potential confounding factors.

In this section, we discuss another simple extension of our model to relax the assumption that the agent cannot 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.

We 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.[31] To model this, consider an agent who has a preference that can be represented by[32]

$$W_1(\mathbb{F},\sigma) := \max_{F\in\mathbb{F}}\sum_{s\in S}\sigma(s)\left[\widetilde{U}(F,\mathbb{F},\mu_s^\sigma) - R(F,\mathbb{F},\mu_s^\sigma)\right] \tag{21}$$

[30] Another implicit assumption for the interpretation for the IT representation is that the agent does not regret her direction choice. We believe this assumption may not be as problematic since it is a natural first step to model regret in a way such that more recent decisions are more salient.

[31] Generally, the agent can back out some information about the state of the world by observing the payoff she receives and the partition it induces on the state space Ω. Here we make the simplifying assumption that the agent observes the true state of the world when she receives her payoff.

[32] Again, we will not present an axiomatic treatment for this representation since it is only for comparison purposes.

where $\sigma(s)$ and μ_s^σ are defined as before. The regret term $R(F,\mathbb{F},\mu_s^\sigma)$ is also defined the same 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 direction $\mathbb{F}$ it is chosen from. Formally,

$$\widetilde{U}(F,\mathbb{F},\mu_s^\sigma) := \max_{f\in F}\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)\Big[u\big(f(\omega)\big) - \widetilde{R}(f,F,\mathbb{F},\omega)\Big] \tag{22}$$

where $\widetilde{R}(f,F,\mathbb{F},\omega)$ is the regret about the act choice f if the state is revealed to be ω. Formally,

$$\widetilde{R}(f,F,\mathbb{F},\omega) := K_1\left[\max_{G\in\mathbb{F}}\max_{g\in G}u\big(g(\omega)\big) - u\big(f(\omega)\big)\right] + K_2\left[\max_{h\in F}u\big(h(\omega)\big) - u\big(f(\omega)\big)\right] \tag{23}$$

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. Therefore, a representation W_1 in the form of equation (21) 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 direction.

LEMMA 5: *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 direction $\mathbb{F}$ and information structures $\sigma, \sigma' \in \mathcal{I}$.

Proof. See Appendix C.7. □

Intuitively, if $K_2 = 0$ and the agent's regret about her act choice solely comes from the counterfactual comparison with what she could have got from the entire direction $\mathbb{F}$, then this part of regret will only depend 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.

Another complication is about how anticipatory is the agent toward these regret feelings. For example, in the representation defined in equation (21), we assume that the regret term $R(F, \mathbb{F}, \mu_s^\sigma)$ is the same as in the baseline model. That is, we implicitly assume that the agent does not take future regret into account when considering her current regret. It might also be plausible to define the regret term differently by

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

where $\widetilde{U}(F, \mathbb{F}, \mu_s^\sigma)$ is defined as in equation (22). With this definition, the agent essentially "regret about her future regret" in the sense that she is so forward-looking that she includes her regret from the act choice in the future when evaluating the value of a menu chosen from the direction. Incorporating this consideration would make the model even more complicated. It would be an interesting avenue for future research to find a clean way to incorporate regret into an infinite-horizon discrete time model.

A The Framework with Lotteries

In this section of the Appendix, we present some preliminary results in the framework with lotteries. These results will be used to prove results in the main text.

A.1 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 direction is a nonempty compact subset of $\widehat{\mathcal{M}}$. Let $\widehat{\mathcal{D}}$ denote the set of all directions, with typical elements $\mathbb{A}, \mathbb{B}$. Endow $\widehat{\mathcal{D}}$ with the Hausdorff metric.

Let $\mathcal{V}$ denote the set of normalized expected utilities over $\Delta(Z)$, that is,

$$\mathcal{V} := \left\{ v \in \mathbb{R}^Z \mid \sum_{z \in Z} v_z = 0 \right\}.$$

Let $\mathcal{U}$ denote the set of doubly normalized expected-utility functions on $\Delta(Z)$, that is,

$$\mathcal{U} := \left\{ u \in \mathbb{R}^Z : \sum_{z \in Z} u_z = 0, \sum_{z \in Z} u_z^2 = 1 \right\}.$$

A.2 Redundancy of a Collection of Utilities

An important notion that will be repeatedly used in our proof is about the redundancy of a collection of linear functions. Formally,

Definition 15: 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 $\beta \in \mathbb{R}$.

By convention, an empty collection is not redundant.

Using the notion of redundancy, we can prove the following lemma.

Lemma 6: *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 6. The "if" part is straightforward.

To prove the "only if" part, first 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 $|J| = 1$, then let $J = \{u_j\}$, and by the definition of J, there exists $L > 0$ such that $U(A) = L \max_{p \in A} u_j(p)$. And $Lu_j(p) = U(\{p\}) = 0$ for all $p \in \Delta(Z)$ implies $u_j(p) = 0$, contradicting J being non-redundant. So $|J| \neq 1$.

Suppose by contradiction that $|J| \geq 2$, then $\{u_j\}_{j \in J}$ is a non-redundant collection of expected utilities, and by a standard result (e.g., Lemma A.1 of Kopylov (2009, JET)), there exists a menu of lotteries $A = \{p_j\}_{j \in J}$ such that p_j is the unique maximizer of u_j in A. Then, fix any $k \in J$,

$$U(A) = \sum_{j \in J} L_j \max_{p \in A} u_j(p) = \sum_{j \in J} L_j u_j(p_j) > \sum_{j \in J} L_j u_j(p_k) = U(\{p_k\}) = 0,$$

contradicting $U(A) = 0$. So $|J| < 2$. □

A.3 The Partial Regret Representation

Our primitive is a binary relation $\succsim$ over $\widehat{\mathcal{D}}$ (the set of all menus of menus of lotteries).

Definition 16: 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

$$\widehat{V}(\mathbb{A}) = \max_{A \in \mathbb{A}} \sum_{u \in \mathrm{supp}(\mu)} \mu(u) \left[\max_{p \in A} u(p) - R(A, \mathbb{A}, u) \right] \tag{24}$$

where

$$R(A, \mathbb{A}, u) := K \left[\max_{B \in \mathbb{A}} \max_{p \in B} u(p) - \max_{p \in A} u(p) \right]. \tag{25}$$

The interpretation for the PR representation is the same as discussed in the main text. The agent has subjective uncertainty about her future tastes. This uncertainty resolves after her committing to a menu from a direction but before her choosing a lottery to consume from the menu. Given a taste, she can choose the best lottery from her menu of choice but inevitably suffers from regret if her choice of menu is suboptimal.

A useful equivalent expression of the utility representation is

$$\widehat{V}(\mathbb{A}) = \max_{A \in \mathbb{A}} \left[(1 + K) \sum_{u \in \mathrm{supp}(\mu)} \mu(u) \max_{p \in A} u(p) \right] - K \sum_{u \in \mathrm{supp}(\mu)} \mu(u) \max_{B \in \mathbb{A}} \max_{p \in B} u(p). \tag{26}$$

We first present the axioms that characterize the PR representation. These should look familiar to the axioms presented in the main text. The proof for the representation theorem and some other intermediate results are in the next section.

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 for any $\mathbb{C}$ and $\alpha \in (0, 1]$,*

$$\alpha\mathbb{A} + (1 - \alpha)\mathbb{C} \succ \alpha\mathbb{B} + (1 - \alpha)\mathbb{C}.$$

For the next axiom, we need to define the notion of a critical subset in the framework with lotteries. The interpretation is very similar to that in the framework with acts.

Definition 17: Let $\mathbb{A}$ be a direction. 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 $\mathbb{A} \in \mathcal{D}$, there exists $\mathbb{B}$ with $|\mathbb{B}| < N$ such that $\mathbb{B}$ is critical for $\mathbb{A}$;

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 direction $\mathbb{A}$ and any 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 nontriviality 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.4 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) who establishes a finite version of Dekel, Lipman, and Rustichini (2001) for representations with subjective state spaces.

Definition 18: A binary relation $\succsim$ over $\widehat{\mathcal{D}}$ has a *nested DLR representation* if $\succsim$ can be represented by

$$V_{DLR}(\mathbb{A}) = \sum_{i \in I^+} \max_{A \in \mathbb{A}} U_i(A) - \sum_{i \in I^-} \max_{A \in \mathbb{A}} U_i(A) \tag{27}$$

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 collection of continuous linear functions from $\widehat{\mathcal{M}}$ to $\mathbb{R}$. (It is without loss to assume that $I^+ \cap I^- = \emptyset$. Let $I := I^+ \cup I^-$.) Moreover, for each $i \in I$, $U_i : \widehat{\mathcal{M}} \to \mathbb{R}$ has a DLR representation with

$$U_i(A) = \sum_{k \in P_i} \max_{p \in A} w_{ik}(p) - \sum_{j \in N_i} \max_{p \in A} v_{ij}(p) \tag{28}$$

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 that $P_i \cap N_i = \emptyset$) 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 7: *A binary relation $\succsim$ over $\widehat{\mathcal{D}}$ has a nested DLR representation if and only if it satisfies Axioms A.1-A.4.*

Proof. See the proof of Theorem 5 of Stovall (2018). □

Given the nested DLR representation, we can start to fine-tune the states and substates through the other axioms to get to the PR representation.

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 $\mathbb{A} \succsim \mathbb{A} \cup \mathbb{B}$.*

As suggested by its name, Axiom A.9 is a strengthening of Axiom A.5. However, together with Axioms A.1 and A.2, Axioms A.5 and A.9 are equivalent. This is summarized in the following lemma.

Lemma 8: *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. It is straightforward to see that if $\succsim$ satisfies Axiom A.9, then it satisfies Axiom A.5. We want to show that the other direction goes through, that is, Axiom A.5 implies

Axiom A.9. Suppose $\succsim$ satisfies Axiom A.5, and suppose two directions 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 the set of all menus of lotteries, $\widehat{\mathcal{M}}$, is compact when equipped with the Hausdorff metric. Therefore, $\widehat{\mathcal{M}}$ is separable. We can thus choose a countable dense subset of $\mathbb{B}$, say $\mathbb{B}^* = \{B_i \mid i = 1, 2, \dots\}$. Define a sequence of directions $(\mathbb{B}_n)_{n=1,2,\dots}$ by $\mathbb{B}_n = \{B_1, B_2, \dots, B_n\}$. 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, $(\mathbb{B}_n)_{n=1,2,\dots}$ 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 $\mathbb{A} \succsim \mathbb{A}_2 = \mathbb{A} \cup \mathbb{B}_2$. We can repeat this argument for each n and 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, $\mathbb{A} \succsim \lim_{n\to\infty} \mathbb{A} \cup \mathbb{B}_n = \mathbb{A} \cup \mathbb{B}$. $\square$

With this equivalence established, we show the effect of imposing Axiom A.9 on a nested DLR representation.

LEMMA 9: *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, that is, $|I^+| \leq 1$ in equation (27); 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 9.
If. Condition 1 implies that $\succsim$ has a nested DLR representation, so by Lemma 7, $\succsim$ satisfies Axioms A.1-A.4. We just need to check if Axiom A.9 is implied by conditions 1 and 2. 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 directions $\mathbb{A}, \mathbb{B} \in \widehat{\mathcal{D}}$. Axiom A.9 is satisfied.

If $|I^+| = 1$, then there is exactly one positive state, and

$$V_{DLR}(\mathbb{A}) = \max_{A\in\mathbb{A}} U_0(A) - \sum_{i\in I^-} \max_{A\in\mathbb{A}} U_i(A),$$

with $\{U_0\} \cup \{U_i\}_{i\in I^-}$ being a non-redundant collection of continuous linear functions from $\widehat{\mathcal{M}}$ to $\mathbb{R}$. Recall that $\widehat{v}(A) := V_{DLR}(\{A\}) = \sum_{i\in I^+} U_i(A) - \sum_{i\in I^-} U_i(A)$, and $\{A\} \succsim \{B\}$ if and only if $\widehat{v}(A) \geq \widehat{v}(B)$.

Suppose for any $B \in \mathbb{B}$, there exists $A \in \mathbb{A}$ such that $\widehat{v}(A) \geq \widehat{v}(B)$. Since U_0 is non-constant, condition 2 implies that $\alpha > 0$, and

$$U_0(A) = \alpha \underbrace{\left[U_0(A) - \sum_{i \in I^-} U_i(A) \right]}_{=\widehat{v}(A) = V_{DLR}(\{A\})} \text{ for all } A \in \widehat{\mathcal{M}}.$$

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 when $|I^+| = 1$.

This completes the proof of the "if" part.

Only if. By Lemma 7, Axioms A.1-A.4 imply the existence of a nested DLR representation for $\succsim$ as in Definition 18. We want to show that if Axiom A.9 is also satisfied, then $|I^+| \leq 1$ and condition 2 is also satisfied.

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 is 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 $\{U_k\}_{k \in I^+ \cup I^-}$.

  Thus we can apply a standard result (e.g., Lemma A.1 of Kopylov, 2009) to conclude that there exists 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 $\mathbb{A}_1 \neq \emptyset$ (since $|\mathbb{A}| \geq 2$) 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.

- 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, thus we can apply a standard result (e.g., Lemma A.1 of Kopylov, 2009) to conclude that there exists a collection of menus $\mathbb{A}' := \{B\} \cup \{A_k\}_{k \in I^+ \cup I^-}$ such that B is unique maximizer of $\widehat{v}$ in $\mathbb{A}$ and A_k is the unique maximizer of U_k in $\mathbb{A}$ for each $k \in I^+ \cup I^-$.

Construct two directions by $\mathbb{A}'_1 := \mathbb{A}' \setminus \{A_{i^*}\}$ and $\mathbb{B}'_1 := \{A_{i^*}\}$. Then $\mathbb{A}'_1 \neq \emptyset$ (since $B \in \mathbb{A}'_1$) 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.

- 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. Then the similar arguments as in Case 2 above will lead to a violation of Axiom A.9.

This completes the proof of the "only if" part. □

We move on to further fine-tune the states and substates of the nested DLR representation. A weapon for that is the following lemma.

LEMMA 10: *Let $\{U_i\}_{i \in I}$ be a finite collection of continuous linear functions from $\widehat{\mathcal{M}}$ to $\mathbb{R}$ such that the collection is non-redundant, and each $U_i : \widehat{\mathcal{M}} \to \mathbb{R}$ has a minimal finite DLR representation, that is,*

$$U_i(A) = \sum_{k \in P_i} \max_{p \in A} w_{ik}(p) - \sum_{j \in N_i} \max_{p \in A} v_{ij}(p)$$

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. *For any $i \in I$, any $k \in P_i$ and any $j \in N_i$,*

$$\left| \arg\max_{p \in A_i} w_{ik}(p) \right| = 1, \quad \left| \arg\max_{q \in A_i} v_{ij}(q) \right| = 1,$$

and for each fixed i, all these unique maximizers are distinct from each other.

Proof. See the proof of Lemma 3 of Stovall (2018). □

With Lemma 10 in hand, we can further fine-tune the substates in the nested DLR representation.

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 11: *Suppose $\succsim$ has a nested DLR representation as in Definition 18, 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. Suppose $\succsim$ has a nested DLR representation as in Definition 18, moreover, $|N_i| = 0$ for any $i \in I^+$ and $|P_i| \leq 1$ for any $i \in I^-$. We want to show that $\mathbb{A} \cup \{A \cup B\} \succsim \mathbb{A} \cup \{A, B\}$ for any direction $\mathbb{A}$ and menus A, B.

For any positive state $i \in I^+$, $U_i(A \cup B) \geq \max\{U_i(A), U_i(B)\}$ 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

$$U_i(C) = \max_{p \in C} w_{i0}(p) - \sum_{j \in N_i} \max_{p \in C} v_{ij}(p),$$

and 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 the negative terms in $V_{DLR}(\mathbb{A} \cup \{A, B\})$.

This completes the proof for the "if" part.

Only if. Suppose $\succsim$ has a nested DLR representation as in Definition 18, moreover, $\succsim$ satisfies Axiom A.6 (Interim Preference for Flexibility). We want to show that $|N_i| = 0$ for all $i \in I^+$ and $|P_i| \leq 1$ for all $i \in I^-$.

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 from $\widehat{\mathcal{M}}$ to $\mathbb{R}$. 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 10 are satisfied. Therefore, there exists a collection of menus $\{A_i\}_{i \in I}$ in the interior of $\Delta(Z)$ such that

1. $U_i(A_i) > U_i(A_j)$ for any $i, j \in I$ with $j \neq i$.

2. For any $i \in I$, any $k \in P_i$ and any $j \in N_i$,

$$\left|\arg\max_{p \in A_i} w_{ik}(p)\right| = 1, \ \left|\arg\max_{q \in A_i} v_{ij}(q)\right| = 1,$$

and for each fixed i, all these unique maximizers are distinct from each other.

Let $\mathbb{A} = \{A_i\}_{i \in I}$.

We will first show that $|P_i| \leq 1$ for all $i \in I^-$, that is, there is at most one positive sub-state in any negative state. We do this by proving its contrapositive.

Suppose by contradiction that $|P_{i^*}| \geq 2$ for some $i^* \in I^-$, that is, U_{i^*} has two or more positive sub-states. Then we want to construct A, B such that

$$V_{DLR}(\mathbb{A} \cup \{A \cup B\}) < V_{DLR}(\mathbb{A} \cup \{A, B\}).$$

We want A, B and $A \cup B$ to be all "closed" to A_{i^*} so that for any $i \in I$ with $i \neq i^*$,

$$U_i(A_i) > U_i(A_{i^*}) \approx \max\Big\{U_i(A \cup B), U_i(A), U_i(B)\Big\}$$

This will guarantee that the difference between $V_{DLR}(\mathbb{A} \cup \{A \cup B\})$ and $V_{DLR}(\mathbb{A} \cup \{A, B\})$ is generated solely by the difference in U_{i^*}. Since i^* is a negative state, to get the desired strict inequality, we want

$$\max_{C \in \mathbb{A} \cup \{A \cup B\}} U_{i^*}(C) > \max_{C \in \mathbb{A} \cup \{A, B\}} U_{i^*}(C).$$

For ease of exposition, we will write U_* for U_{i^*} and A_* for A_{i^*}. Similarly, we write P_* for P_{i^*} (the index set of positive sub-states 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 $\max\{a, b\}$.

By construction, A_* is the unique maximizer of U_* in $\mathbb{A}$, thus, it suffices to have

$$U_*(A_*) \vee U_*(A \cup B) > U_*(A_*) \vee U_*(A) \vee U_*(B)$$

which will happen if we have $U_*(A \cup B) > U_*(A_*), U_*(A), U_*(B)$. With this goal in mind, we start to construct A and B.

Define $p_k := \arg\max_{p \in A_*} w_{*k}(p)$. Since $|P_*| \geq 2$, we can pick k and k' with $k \neq k'$ from P_*. For any $\varepsilon > 0$, define

$$A^\varepsilon := A_* \cup \{p_k + \varepsilon w_{*k}\}, \ B^\varepsilon := A_* \cup \{p_{k'} + \varepsilon w_{*k'}\}.$$

Since w_{*k} is normalized and p_k is in the interior of $\Delta(Z)$ by construction, by a standard result, we can find some 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,

$$\begin{aligned}
\max_{p \in A^{\varepsilon}} w_{*k}(p) &= w_{*k}(p_k) \vee w_{*k}(p_k + \varepsilon w_{*k}) = w_{*k}(p_k) + \varepsilon \|w_{*k}\|^2 \\
\max_{p \in B^{\varepsilon}} w_{*k}(p) &= w_{*k}(p_k) \vee \Big(w_{*k}(p_{k'}) + \varepsilon w_{*k}(w_{*k'}) \Big)
\end{aligned}$$

By construction, $w_{*k}(p_{k'}) < w_{*k}(p_k)$, so there exists $\varepsilon_1 > 0$ such that

$$w_{*k}(p_{k'}) + \varepsilon_1 w_{*k}(w_{*k'}) < w_{*k}(p_k),$$

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 $\varepsilon_3 := \min\{\varepsilon_1, \varepsilon_2\}$, then

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

Note that

$$\begin{aligned}
U_*(A^{\varepsilon_3} \cup B^{\varepsilon_3}) &= \sum_{k \in P_*} \max_{p \in A^{\varepsilon_3} \cup B^{\varepsilon_3}} w_{*k}(p) - \sum_{j \in N_*} \max_{p \in A^{\varepsilon_3} \cup B^{\varepsilon_3}} v_{*j}(p) \\
U_*(A^{\varepsilon_3}) &= \sum_{k \in P_*} \max_{p \in A^{\varepsilon_3}} w_{*k}(p) - \sum_{j \in N_*} \max_{p \in A^{\varepsilon_3}} v_{*j}(p) \\
U_*(B^{\varepsilon_3}) &= \sum_{k \in P_*} \max_{p \in B^{\varepsilon_3}} w_{*k}(p) - \sum_{j \in N_*} \max_{p \in B^{\varepsilon_3}} v_{*j}(p)
\end{aligned}$$

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 $p_{k'}$ (by Lemma*), 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 A^ε and B^ε will satisfy our desired condition:

$$U_*(A^\varepsilon \cup B^\varepsilon) > U_*(A_*), U_*(A^\varepsilon), U_*(B^\varepsilon)$$

because: (i) they all have the same negative term, and (ii) $A^\varepsilon \cup B^\varepsilon \supset A^\varepsilon, B^\varepsilon \supseteq A_*$ so $A^\varepsilon \cup B^\varepsilon$ has a weakly larger positive term, but it must be strictly larger as well because A^ε contains a unique maximizer $p_k + \varepsilon w_{*k}$ in sub-state k and B^ε contains a unique maximizer $p_{k'} + \varepsilon w_{*k'}$ in sub-state k'.

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

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

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 for any $i \in I^+$, $|N_i| = 0$, that is, there is no negative sub-state in any positive state. We do this by proving its contrapositive.

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

$$A^\varepsilon := A_* \quad \text{and} \quad B^\varepsilon = (A_* \setminus \{q_j\}) \cup \{q_j - \varepsilon v_{*j}\}.$$

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

$$\begin{aligned}
v_{*j}(q_j - \varepsilon v_{*j}) &= v_{*j}(q_j) - \varepsilon \|v_{*j}\|^2 > \max_{p\in A_* \setminus \{q_j\}} v_{*j}(p) \\
v_{*j'}(q_j - \varepsilon v_{*j}) &= v_{*j'}(q_j) - \varepsilon v_{*j'}(v_{*j}) < \max_{p\in A_*} v_{*j'}(p),\ \forall j' \in N_* \text{ and } j' \neq j \\
w_{*k}(q_j - \varepsilon v_{*j}) &= w_{*k}(q_j) - \varepsilon w_{*k}(v_{*j}) < \max_{p\in A_*} w_{*k}(p),\ \forall k \in P_*
\end{aligned}$$

Fix such a ε, we will 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.

This completes the proof of Lemma 11. $\square$

Lemma 12 states that imposing the Inclusion axiom is equivalent to requiring there to be no negative substates in any negative state. Formally,

LEMMA 12: *Suppose $\succsim$ has a nested DLR representation as in Definition 18, then $\succsim$ satisfies Axiom A.7 (Inclusion) if and only if $|N_i| = 0$ for any $i \in I^-$.*

Proof. If. Suppose $\succsim$ has a nested DLR representation as in Definition 18, moreover, $|N_i| = 0$ for any $i \in I^-$. We want to show that $\mathbb{A} \cup \{B\} \succsim \mathbb{A}$ for any direction $\mathbb{A}$ and any menu B such that $B \subseteq A$ for some $A \in \mathbb{A}$.

Fix A, B with $B \subseteq A$. For any negative state $i \in I^-$, $U_i(A) \geq U_i(B)$ 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. Suppose $\succsim$ has a nested DLR representation as in Definition 18, moreover, $\succsim$ satisfies Axiom A.7 (Inclusion) is satisfied, we want to show that for any $i \in I^-$, $|N_i| = 0$. That is, there is no negative sub-state in any negative state. We do this by proving its contrapositive.

Let $\mathbb{A} = \{A_i\}_{i \in I}$ be constructed the same way as in the proof of Lemma 11. 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

$$A^\varepsilon := A_* \cup \{q_j - \varepsilon v_{*j}\} \quad \text{and} \quad B^\varepsilon = (A_* \setminus \{q_j\}) \cup \{q_j - \varepsilon v_{*j}\}.$$

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

$$\begin{aligned}
v_{*j}(q_j - \varepsilon v_{*j}) &= v_{*j}(q_j) - \varepsilon \|v_{*j}\|^2 > \max_{p \in A_* \setminus \{q_j\}} v_{*j}(p) \\
v_{*j'}(q_j - \varepsilon v_{*j}) &= v_{*j'}(q_j) - \varepsilon v_{*j'}(v_{*j}) < \max_{p \in A_*} v_{*j'}(p), \ \forall j' \in N_* \text{ and } j' \neq j \\
w_{*k}(q_j - \varepsilon v_{*j}) &= w_{*k}(q_j) - \varepsilon w_{*k}(v_{*j}) < \max_{p \in A_*} w_{*k}(p), \ \forall k \in P_*
\end{aligned}$$

Fix such a ε, we will 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^* is a negative state), which is a direct violation of Inclusion. $\square$

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

Proof of Theorem 6.
If. Suppose $\succsim$ has a FPR representation with parameters (μ, K), that is, $\succsim$ can be repre-

sented by

$$\widehat{V}(\mathbb{A}) = \max_{A\in\mathbb{A}} \left[(1+K)\sum_{u\in\text{supp}(\mu)} \mu(u)\max_{p\in A} u(p)\right] - K\sum_{u\in\text{supp}(\mu)} \mu(u)\max_{A\in\mathbb{A}}\max_{p\in A} u(p).$$

We want to show that Axioms A.1-A.8 are satisfied. It is straightforward to verify the necessity of the axioms using Lemmas 7, 9, 11, 12.

Only if. Axioms A.1-A.4 delivers a nested DLR representation (Lemma 7). Axiom A.5 guarantees that there is at most one positive state (part 1 of Lemma 9). 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 11). Axiom A.7 guarantees that there are no negetive substates in any negative state (Lemma 12). 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

$$V_S(\mathbb{A}) = \max_{A\in\mathbb{A}} \sum_{k\in P} \max_{p\in A} w_k(p) - \sum_{j\in N} \max_{A\in\mathbb{A}} \max_{p\in A} v_j(p). \tag{29}$$

For convenience, define

$$U_0(A) := \sum_{k\in P} \max_{p\in A} w_k(p) \qquad U_j(A) := \max_{p\in A} v_j(p) \text{ for each } j\in N$$

$$\widehat{v}(A) := V_S(\{A\}) = \sum_{k\in P} \max_{p\in A} w_k(p) - \sum_{j\in N} \max_{p\in A} v_j(p).$$

If $\succsim$ can be represented by V_S as defined in equation (29) 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.

Note that since $\{v_j\}_{j\in N}$ is non-redundant, $\{U_j\}_{j\in N}$ must also be non-redundant. We continue our proof by proving the following lemma.

Lemma 13: *If $\succsim$ can be represented by V_S defined in equation (29) with $|P| \geq 1$ and $\succsim$ satisfies Axiom A.5, then there exists $\alpha > 0$ such that $U_0 = \alpha v$.*

Proof of Lemma 13. 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 9, 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 for any menu A,

$$\sum_{k\in P} \max_{p\in A} w_k(p) = \beta \max_{p\in A} v_j(p).$$

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 9 again to conclude that there exists $\alpha > 0$ such that $U_0' = \alpha\widehat{v}$ 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,

$$U_0 = \frac{\beta}{\beta - 1} U_0' = \frac{\beta}{\beta - 1}(\alpha\widehat{v}).$$

With $\beta > 1$ and $\alpha > 0$, we can set $\alpha' := \frac{\alpha\beta}{\beta-1}$. Then $\alpha' > 0$ and $U_0 = \alpha'\widehat{v}$.

This completes the proof of Lemma 13. □

We have just shown that if Axioms A.1-A.8 are satisfied, then there exists $\alpha > 0$ 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

$$\sum_{j\in N} \max_{p\in A} v_j(p) = \frac{\alpha - 1}{\alpha} \sum_{k\in P} \max_{p\in A} w_k(p). \tag{30}$$

Claim: For equation (30) 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 (30) will represent a preference that exhibits the opposite of preference for flexibility, while the LHS of equation (30) represents a preference that exhibits preference for flexibility, contradiction. □

Now if $\alpha = 1$, then for any menu A,

$$\sum_{j\in N} \max_{p\in A} v_j(p) = 0.$$

By Lemma 6, 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 then get a FPR 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 FPR representation by setting $K = \alpha - 1 > 0$, (doubly) normalizing each v_j and scaling everything up by multiplicating K.

This completes the proof of Theorem 6. $\square$

A.5 Uniqueness of the PR representation

For the identification of the parameters μ and K, we build on the identification result of Dekel et al. (2001) and the uniqueness results in Sarver (2008).

Let $\widehat{V}$ be a PPR 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

$$\widehat{v}(A) := \widehat{V}(\{A\}) = \sum_{u\in \text{supp}(\mu)} \mu(u) \max_{p\in A} u(p). \tag{31}$$

That is, $\widehat{v}$ represents a preference over $\widehat{\mathcal{M}}$ generated by $\succsim$ restricting to singleton directions (directions containing only one menu). This is a special case of the representation captured in Dekel, Lipman, and Rustichini (2001) and Dekel, Lipman, Rustichini, and Sarver (2007) with preference for flexibility. Therefore, we can apply their identification result to conclude that the subjective belief over tastes, μ, is uniquely identified (since the expected utilities are doubly normalized). Formally,

Lemma 14: *Suppose both (μ, K) and (μ', K') represent $\succsim$, then $\mu' = \mu$.*

Now that μ is identified, we move on to the identification of K. We follow a similar chain of steps used to establishing the uniqueness results in Sarver (2008).

Define $\widehat{r} : \widehat{\mathcal{D}} \to \mathbb{R}$, for any direction $\mathbb{A} \in \widehat{\mathcal{D}}$, by

$$\begin{aligned}\widehat{r}(\mathbb{A}) &:= \min_{A\in\mathbb{A}} \sum_{u\in\mathrm{supp}(\mu)} \mu(u)R(A,\mathbb{A},u) \\ &= \min_{A\in\mathbb{A}} \sum_{u\in\mathrm{supp}(\mu)} \mu(u)K\left[\max_{B\in\mathbb{A}}\max_{q\in B} u(q) - \max_{p\in A} u(p)\right].\end{aligned} \tag{32}$$

The function $\widehat{r}(\mathbb{A})$ represents the minimal expected regret that the agent can experience when faced with direction $\mathbb{A}$. Note that given a direction $\mathbb{A}$, the menu A that maximizes v also minimizes expected regret. Therefore, for any direction $\mathbb{A} \in \widehat{\mathcal{D}}$, the agent will choose $A \in \mathbb{A}$ to maximize $v(A)$, and

$$\widehat{V}(\mathbb{A}) = \max_{A\in\mathbb{A}} \widehat{v}(A) - \widehat{r}(\mathbb{A}). \tag{33}$$

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, then we can first apply the mixture space theorem to guarantee that there exists $\alpha > 0$ such that $\widehat{v}' = \alpha\widehat{v}$ and $\widehat{r}' = \alpha\widehat{r}$.

But by Lemma 14, $\mu' = \mu$ implies that $\widehat{v}' = \widehat{v}$. Thus, $\alpha = 1$, and $\widehat{r}' = \widehat{r}$. □

To proceed with the identification of K, we first rule out a less interesting case where Axiom A.5 is trivially satisfied.

Lemma 15: *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 15. We show that $1 \iff 2$ and $1 \iff 3$.

$1 \implies 2$: 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 $\Longrightarrow$ 1: Suppose by contradiction that $\widehat{r}(\mathbb{A}) > 0$ for some $\mathbb{A} \in \widehat{\mathcal{D}}$, we want to show that $\succsim$ will not satisfy monotonicity. Fix a direction $\mathbb{A}$ such that $\widehat{r}(\mathbb{A}) > 0$, let $A \in \arg\max_{B\in\mathbb{A}} \widehat{v}(B)$, then $\{A\} \subseteq \mathbb{A}$ but $\{A\} \succ \mathbb{A}$ since

$$\widehat{V}(\{A\}) = \widehat{v}(A) > \widehat{v}(A) - \widehat{r}(\mathbb{A}) = \widehat{V}(\mathbb{A}).$$

Taking the contrapositive completes the proof.

3 $\Longrightarrow$ 1: Straightforward.

1 $\Longrightarrow$ 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 $\mathbb{A} := \{\{p_u\} : u \in \text{supp}(\mu)\}$. Then

$$\begin{aligned}\widehat{r}(\mathbb{A}) &= \min_{A\in\mathbb{A}} \sum_{u\in\text{supp}(\mu)} \mu(u)K\left[\max_{B\in\mathbb{A}}\max_{q\in B} u(q) - \max_{p\in A} u(p)\right]\\ &= \min_{p\in A_0} \sum_{u\in\text{supp}(\mu)} \mu(u)K\left[\max_{q\in A_0} u(q) - u(p)\right]\end{aligned}$$

But for any $p_u \in A_0$, $\max_{q\in A_0} u'(q) - u'(p) > 0$ for any $u' \neq u$. So $\widehat{r}(\mathbb{A}) > 0$. Taking the contrapositive completes the proof. $\square$

We say $\succsim$ has a nontrivial PR representation if there exist $\mathbb{A}$ and $\mathbb{B}$ 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 14, and by Theorem 7, $\widehat{v}' = \widehat{v}$, $\widehat{r}' = \widehat{r}$. Since $\succsim$ has a nontrivial FPR 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 $K' = K$. $\square$

B Proof of Theorem 1

B.1 Necessity of Axioms 1-8

Suppose $\succsim$ has a SIT representation with parameters (π, u, K, σ), we want to show that $\succsim$ satisfies Axioms 1-8.

By assumption, $\succsim$ can be represented by

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

It is without loss to assume that $0 \notin S$. For convenience, let

$$U_0(F) := \sum_{s\in S} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega)u(f(\omega)) \tag{34}$$

$$U_s(F) := \max_{g\in G} \sum_{\omega\in\Omega} \pi(\omega)\sigma(s \mid \omega)u(g(\omega)) \text{ for each } s \in S \tag{35}$$

It can be easily verified that U_0 and $(U_s)_{s\in S}$ are continuous linear functions from $\mathcal{M}$ to $\mathbb{R}$, where $\mathcal{M}$ is the set of all menus of acts equppied with the Hausdorff metric. Therefore, the binary relation $\succsim$ can be represented by

$$V(\mathbb{F}) = \max_{F\in\mathbb{F}} (1+K)U_0(F) - \sum_{s\in S} \max_{F\in\mathbb{F}} K \cdot U_s(F). \tag{36}$$

This is a finite DLR type representation over a general convex space, which is studied in Kopylov (2009).

Lemma 16: *If $\succsim$ has a SIT representation, then $\succsim$ satisfies Axioms 1-3 (Weak Order, Continuity and Independence).*

Proof of Lemma 16. By the arguments above, equation (36) implies that $\succsim$ has a finite DLR type representation over a general domain as characterized in Kopylov (2009). Therefore, $\succsim$ satisfies Axioms 1 and 2 by applying Theorem 2.1 of Kopylov (2009).

Axiom 3, our independence axiom, is slightly stronger than the independence axiom posited in Kopylov (2009). To be precise, we verify that $\succsim$ satisfies Axiom 3 directly. Let $\mathbb{F}, \mathbb{G}, \mathbb{H}$ be three directions and $\alpha \in (0,1)$.

$$\begin{aligned}
\mathbb{F} \succsim \mathbb{G} &\iff V(\mathbb{F}) \geq V(\mathbb{G}) \\
&\iff \alpha V(\mathbb{F}) + (1-\alpha)V(\mathbb{H}) \geq \alpha V(\mathbb{F}) + (1-\alpha)V(\mathbb{H}) \\
&\iff V\big(\alpha\mathbb{F} + (1-\alpha)\mathbb{H}\big) \geq V\big(\alpha\mathbb{G} + (1-\alpha)\mathbb{H}\big) \\
&\iff \alpha\mathbb{F} + (1-\alpha)\mathbb{H} \succsim \alpha\mathbb{G} + (1-\alpha)\mathbb{H}
\end{aligned}$$

where the third equivalence follows from the definition of the convex combination of two directions and the fact that $U_0, (U_s)_{s\in S}$ are all linear functions satisfying

$$U_s(\alpha F + (1-\alpha)G) = \alpha U_s(F) + (1-\alpha)U_s(G)$$

for any menus $F, G \in \mathcal{M}$ and any scalar $\alpha \in [0,1]$. □

Lemma 17: *If $\succsim$ has a SIT representation, then $\succsim$ satisfies Axioms 4 (Finiteness).*

Proof of Lemma 17. Recall that S is a finite set representing the possible signal realizations from the anticipated information structure σ. Let $N = |S| + 2$.

For any direction $\mathbb{F}$ such that $|\mathbb{F}| < N$, the direction itself is a critical subset whose cardinality is smaller than N. For any direction $\mathbb{F}$ such that $|\mathbb{F}| \geq N$ (including the countable and uncountable cases), let $F_0 \in \arg\max_{F \in \mathbb{F}} U_0(F)$, and $F_s \in \arg\max_{F \in \mathbb{F}} U_s(F)$ for each $s \in S$. Then $\mathbb{G} := \{F_0\} \cup \{F_s\}_{s \in S}$ is critical for $\mathbb{F}$ and $|\mathbb{G}| \leq |S| + 1 < N$. For the second part of Axiom 4, fix any direction $\mathbb{F}$ and menu $F \in \mathbb{F}$. If $|F| < N$, then F itself is critical for F in $\mathbb{F}$. If $F \notin \arg\max_{G \in \mathbb{F}} U_0(G)$ and $F \notin \arg\max_{G \in \mathbb{F}} U_s(G)$ for any $s \in S$, then this menu does not matter in the first place for the evaluation of $\mathbb{F}$, so the empty set $\emptyset$ is critical for F in $\mathbb{F}$. Lastly, if $|F| \geq N$ and F is a maximizer of some of U_0 and $(U_s)_{s \in S}$, we discuss two different cases:

- If F is a maximizer of U_0, then let f_s be a maximizer of $\sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega)u\big(f(\omega)\big)$ for each $s \in S$, then $G := \{f_s\}_{s \in S}$ is critical for F in $\mathbb{F}$, and $|G| \leq |S| < N$.
- If F is a maximizer of U_s for some $s \in S$, let g be a maximizer of $\sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega)u\big(f(\omega)\big)$, then $G := \{g\}$ is critical for F in $\mathbb{F}$, and $|G| = 1 < N$.

This completes the proof of Lemma 17. $\square$

LEMMA 18: *If $\succsim$ has a SIT representation, then $\succsim$ satisfies Axioms 5 (Ex-Ante Regret).*

Proof of Lemma 18. Recall that the SIT representation can be written as

$$V(\mathbb{F}) = \max_{F \in \mathbb{F}} (1 + K)U_0(F) - \sum_{s \in S} \max_{F \in \mathbb{F}} K \cdot U_s(F) \tag{37}$$

where

$$\begin{aligned} U_0(F) &:= \sum_{s \in S} \max_{f \in F} \sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega)u(f(\omega)) \\ U_s(F) &:= \max_{g \in G} \sum_{\omega \in \Omega} \pi(\omega)\sigma(s \mid \omega)u(g(\omega)) \text{ for each } s \in S \end{aligned}$$

Observe that $U_0(F) = V(\{F\})$ for any menu $F \in \mathcal{M}$. Suppose $\{F\} \succsim \{G\}$ and $F \in \mathbb{F}$, we want to show that $\mathbb{F} \succsim \mathbb{F} \cup \{G\}$.

$\{F\} \succsim \{G\}$ implies that $U_0(F) \geq U_0(G)$. Together with $F \in \mathbb{F}$, this implies

$$\max_{H \in \mathbb{F}} (1 + K)U_0(H) = \max_{H \in \mathbb{F} \cup \{G\}} (1 + K)U_0(H).$$

That is, $V(\mathbb{F})$ and $V(\mathbb{F}\cup\{G\})$ have the same positive term in equation (37). On the other hand, for any $s \in S$, we have

$$\max_{H\in\mathbb{F}} K\cdot U_s(F) \leq \max_{H\in\mathbb{F}\cup\{G\}} K\cdot U_s(F).$$

Therefore, $V(\mathbb{F}) \geq V(\mathbb{F}\cup\{G\})$, indicating $\mathbb{F} \succsim \mathbb{F}\cup\{G\}$. $\square$

LEMMA 19: *If $\succsim$ has a SIT representation, then $\succsim$ satisfies Axioms 6 (Interim Preference for Flexibility).*

Proof of Lemma 19. We want to show that for any $\mathbb{F}$ and any F, G,

$$\mathbb{F}\cup\{F\cup G\} \succsim \mathbb{F}\cup\{F, G\}.$$

This is easier to see using the alternative expression for the IT representation:

$$\begin{aligned} V(\mathbb{F}) = \; & \max_{F\in\mathbb{F}}\left[(1+K)\sum_{s\in S}\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega))\right] \\ & - K\sum_{s\in S}\max_{G\in\mathbb{F}}\max_{g\in G}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(g(\omega)) \end{aligned} \tag{38}$$

Note that the negative term can be equivalently written as

$$K\sum_{s\in S}\max_{g\in M(\mathbb{F})}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(g(\omega))$$

where $M(\mathbb{F}) = \{f \mid f \in F \text{ for some } F \in \mathbb{F}\}$ denotes the set of feasible acts as defined in the main text. It is clear that $M(\mathbb{F}\cup\{F\cup G\}) = M(\mathbb{F}\cup\{F, G\})$. Therefore, $V(\mathbb{F}\cup\{F\cup G\})$ and $V(\mathbb{F}\cup\{F, G\})$ have the same negative term in equation (38).

On the other hand, $V(\mathbb{F}\cup\{F\cup G\})$ has a weakly larger positive term because

$$\begin{aligned} & \sum_{s\in S}\max_{f\in F\cup G}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega)) \\ \geq \; & \max\left\{\sum_{s\in S}\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega)), \sum_{s\in S}\max_{f\in G}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega))\right\} \end{aligned}$$

Therefore, $V(\mathbb{F}\cup\{F\cup G\}) \geq V(\mathbb{F}\cup\{F, G\})$, indicating $\mathbb{F}\cup\{F\cup G\} \succsim \mathbb{F}\cup\{F, G\}$. $\square$

LEMMA 20: *If $\succsim$ has a SIT representation, then $\succsim$ satisfies Axioms 7 and 8 (Nontriviality and Domination).*

Proof of Lemma 20. For a lottery $\ell \in \Delta(X)$ and its corresponding constant act,

$$V(\{\{\ell\}\}) = \sum_{\omega \in \Omega} \pi(\omega) u(\ell) = u(\ell).$$

And Axiom 7 (Nontriviality) must be satisfied because we require u to be nonconstant.

For Axiom 8 (Domination), suppose f dominates g, then $f(\omega) \succsim g(\omega)$ for any $\omega \in \Omega$, which further implies that $u(f(\omega)) \geq u(g(\omega))$ for any $\omega \in \Omega$. Therefore,

$$V(\{\{f\}\}) = \sum_{\omega \in \Omega} \pi(\omega) u(f(\omega)) \geq \sum_{\omega \in \Omega} \pi(\omega) u(g(\omega)) = V(\{\{g\}\})$$

indicating $f \succsim g$. Moreover, $u(f(\omega)) \geq u(g(\omega))$ for any $\omega \in \Omega$ implies that

$$\sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(f(\omega)) \geq \sum_{\omega \in \Omega} \pi(\omega) \sigma(s \mid \omega) u(g(\omega)) \text{ for any } s \in S.$$

That is, g will never be chosen over f after any signal realization s. Therefore,

$$V(\{\{f, g\}\}) = \sum_{\omega \in \Omega} \pi(\omega) u(f(\omega)) = V(\{\{f\}\})$$

indicating $\{\{f, g\}\} \sim \{\{f\}\}$.

Similarly, if $\mathbb{F}$ dominates $\mathbb{G}$, then $\max_{F \in \mathbb{F}} U_0(F) = \max_{F \in \mathbb{F} \cup \mathbb{G}} U_0(F)$ and

$$\max_{F \in \mathbb{F}} U_s(F) = \max_{F \in \mathbb{F} \cup \mathbb{G}} U_s(F) \text{ for any } s \in S.$$

Therefore, $V(\mathbb{F}) = V(\mathbb{F} \cup \mathbb{G})$, indicating $\mathbb{F} \sim \mathbb{F} \cup \mathbb{G}$. $\square$

This completes the proof for the necessity of Axioms 1-8 for a SIT representation.

B.2 Sufficiency of Axioms 1-8

As we have mentioned in the main text, the proof of Axioms 1-8 being sufficient for a SIT representation is more involved. Here we provide a roadmap before we dive into the details. 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 we have 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. Similar translation techniques are used in Dillenberger, Lleras, Sadowski, and Takeoka (2014).

Step 1: Translate the preference over directions to a preference over "utility directions."

Note that Axioms 1-3 imply their corresponding axioms (weak order, continuity and independence) over acts (i.e., direction 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 result from the Anscombe-Aumann framework (for a detailed treatment, see 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

$$V(\{\{f\}\}) := \sum_{\omega \in \Omega} \bar{\pi}(\omega) u\big(f(\omega)\big).$$

Now for any act $f \in \mathcal{F}$, the composite function $u \circ f : \Omega \to [0, 1]$ specifies the utility associated with f in each state ω. We call this the *utility act* induced by f and also write $u(f)$ to denote the utility act induced by 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 directions. Endow $\mathcal{D}_u$ with the Hausdorff metric.

Naturally, we would believe that $\succsim$ as a binary relation over menus of menus of acts should induce a binary relation $\succsim_u$ over $\mathcal{D}_u$: If $\mathbb{F} \succsim \mathbb{G}$, then define $\mathbb{F}_u \succsim_u \mathbb{G}_u$ where $\mathbb{F}_u := u(\mathbb{F}) = \{u(F) \mid F \in \mathbb{F}\}$ and similar for $\mathbb{G}_u$. However, since u is generally not injective, for this definition to make sense, we need to guarantee that $u(\mathbb{F}) = u(\mathbb{G})$ implies $\mathbb{F} \sim \mathbb{G}$.

Lemma 21: *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 21. We say two AA acts f and g are *indistinguishable* if

$$\{\{f(\omega)\}\} \sim \{\{g(\omega)\}\} \text{ for all } \omega \in \Omega.$$

Note that f and g are indistinguishable if and only if $u(f) = u(g)$. If this looks weird at first sight, recall that $u(f)$ and $u(g)$ are not scalar values of a function but are utility acts, that is, they are 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. To see the if part, note that if $u(F) = u(G)$, then for any $f \in F$ there exists $g \in G$ such that $u(f) = u(g)$, which makes f and g indistinguishable. Similar arguments work for the other half of the arguments).

We say two directions $\mathbb{F}$ 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 direction $\mathbb{F} \in \mathcal{D}$, 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 (Domination),

$$\mathbb{F} \sim \mathbb{F} \cup \mathbb{G} \quad \text{and} \quad \mathbb{G} \sim \mathbb{F} \cup \mathbb{G}.$$

Therefore, $\mathbb{F} \sim \mathbb{G}$ by transitivty. □

We formally define a preference relation $\succsim_u$ over $\mathcal{D}_u$ by

$$\mathbb{F}_u \succsim_u \mathbb{G}_u \text{ if and only if } \mathbb{F} \succsim \mathbb{G} \text{ where } \mathbb{F} \in u^{-1}(\mathbb{F}_u) \text{ and } \mathbb{G} \in u^{-1}(\mathbb{G}_u). \tag{39}$$

This is a valid definition by the result of Lemma 21. We then show that if $\succsim$ satisfies Axioms 1-8, then $\succsim_u$ satisfies the suitably adapted versions of Axioms 1-8. The formal description of the axioms are as below.

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

Axiom B.2—Continuity: *For any $\mathbb{F}_u$, the sets $\{\mathbb{G}_u : \mathbb{G}_u \succsim_u \mathbb{F}_u\}$, $\{\mathbb{G}_u : \mathbb{F}_u \succsim_u \mathbb{G}_u\}$ are closed.*

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

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

We can define the notion of a critical direction and a critical menu within a direction with respect to $\succsim_u$ in a similar way that is defined in the main text.

Axiom B.4—Finiteness: *There exists a natural number N such that*

4.1. For every $\mathbb{F}_u \in \mathcal{D}_u$, there exists $\mathbb{G}_u$ with $|\mathbb{G}_u| < N$ such that $\mathbb{G}_u$ is critical for $\mathbb{F}_u$;

4.2. For every $\mathbb{F}_u \in \mathcal{D}_u$ and every $F_u \in \mathbb{F}_u$, there exists G_u with $|G_u| < N$ such that G_u is critical for F_u in $\mathbb{F}_u$.

Axiom B.5—Ex-Ante Regret: *If $\{F_u\} \succsim_u \{G_u\}$ and $F_u \in \mathbb{F}_u$, then $\mathbb{F}_u \succsim \mathbb{F}_u \cup \{G_u\}$.*

Axiom B.6—Interim Preference for Flexibility: *For any $\mathbb{F}_u$ and any F_u, G_u, $\mathbb{F}_u \cup \{F_u \cup G_u\} \succsim_u \mathbb{F}_u \cup \{F_u, G_u\}$.*

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

We can define the notion of domination with respect to utility acts, utility menus and utility directions in a similar way those are defined in the main text.

Axiom B.8—Domination:

- *If* f_u *dominates* g_u*, then* $f_u \succsim_u g_u$ *and* $\{\{f_u, g_u\} \sim_u \{\{f_u\}\}$*;*
- *If* $\mathbb{F}_u$ *dominates* $\mathbb{G}_u$*, then* $\mathbb{F}_u \sim_u \mathbb{F}_u \cup \mathbb{G}_u$.

Lemma 22: *If a binary relation* $\succsim$ *over directions satisfies Axioms 1-8, then the induced binary relation* $\succsim_u$ *over utility directions satisfies Axioms B.1-B.8.*

Proof. See Appendix D.1. □

This completes the first step in our translation.

Step 2: Translate the preference over utility directions to a preference over menus of menus of lotteries.

Recall that $\mathcal{F}_u$ is the collection of all utility acts and we can identify $\mathcal{F}_u$ with the set of all n-dimensional vectors where each entry is in $[0,1]$. That is, $\mathcal{F} = [0,1]^n$, where $n = |\Omega|$ is the cardinality of the state space Ω.

For convenience, let $\Omega = \{\omega_1, \dots, \omega_n\}$, and introduce an artificial state ω_0 and a new space $\mathcal{F}'$ defined by

$$\mathcal{F}' := \left\{ f' \in [0,n] \times [0,1]^n \;\middle|\; \sum_{i=0}^{n} f'(\omega_i) = n \right\}. \tag{40}$$

For notational simplicity, we suppress the subscript u but one should keep in mind that these are still interepreted as 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,

$$r(f') \in \mathcal{F}_u \text{ with } [r(f')](\omega_i) = f'(\omega_i),\ \forall i \in \{1, 2, \dots, n\}.$$

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

$$\mathbb{F}' \succsim_* \mathbb{G}' \text{ if } r(\mathbb{F}') \succsim_u r(\mathbb{G}').$$

To move on, we introduce yet another new space by

$$\mathcal{F}'' := \left\{ f'' \in [0,n]^{n+1} \;\middle|\; \sum_{i=0}^{n} f''(\omega_i) = n \right\}.$$

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

$$F^{n+1} := \left\{ \left(\frac{n}{n+1}, \dots, \frac{n}{n+1} \right) \right\}.$$

Then $F^{n+1} \in \mathcal{M}'$. Also observe that for any $F'' \in \mathcal{M}''$ and $\varepsilon \leq \frac{1}{n^2}$,

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

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

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

We then show that if $\succsim_u$ satisfies Axioms B.1-B.8, then $\succsim_*$ satisfies the suitably adapted versions of Axioms B.1-B.8. These adapted axioms are listed as Axioms B.1*-B.8*.

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

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

Axiom B.3*—Independence: *For any $\mathbb{F}', \mathbb{G}', \mathbb{H}'$ 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 can define the notion of a critical direction and a critical menu within a direction with respect to $\succsim_*$ in a similar way as before.

Axiom B.4*—Finiteness: *There exists a natural number N such that*

4.1. For every $\mathbb{F}' \in \mathcal{D}'$, there exists $\mathbb{G}'$ with $|\mathbb{G}'| < N$ such that $\mathbb{G}'$ is critical for $\mathbb{F}'$;

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

Axiom B.6*—Interim Preference for Flexibility: *For any $\mathbb{F}' \in \mathcal{D}'$ and any $F', G' \in \mathcal{M}'$, $\mathbb{F}' \cup \{F' \cup G'\} \succsim_* \mathbb{F}' \cup \{F', G'\}$.*

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

Note that Axiom B.8* (Inclusion) is not explicitly stated in Axioms B.1-B.8, it is implied from part 2 of Axiom B.8 (Domination) and closely related to Axiom A.7.

Lemma 23: *If a binary relation $\succsim_u$ over utility directions satisfies Axioms B.1-B.8, then the induced binary relation $\succsim_*$ over $\mathcal{D}'$ satisfies Axioms B.1*-B.8*.*

Proof. See Appendix D.2. □

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

The construction of $\succsim_*$ and $\succsim_{**}$ in Step 2 is the same as that in Dillenberger, Lleras, Sadowski, and Takeoka (2014).

We proceed by first listing the adapted versions of Axioms B.1*-B.8*.

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

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

Axiom B.3**—Independence: *For any $\mathbb{F}'', \mathbb{G}'', \mathbb{H}''$ 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 can define the notion of a critical direction and a critical menu within a direction with respect to $\succsim_{**}$ in a similar way as before.

Axiom B.4**—Finiteness: *There exists a natural number N such that*

4.1. For every $\mathbb{F}'' \in \mathcal{D}''$, there exists $\mathbb{G}''$ with $|\mathbb{G}''| < N$ such that $\mathbb{G}''$ is critical for $\mathbb{F}''$;

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

Axiom B.6**—Interim Preference for Flexibility: *For any* $\mathbb{F}'' \in \mathcal{D}''$ *and any* $F'', G'' \in \mathcal{M}''$, $\mathbb{F}'' \cup \{F'' \cup G''\} \succsim_{**} \mathbb{F}'' \cup \{F'', G''\}$.

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

Lemma 24: *If a binary relation* $\succsim_*$ *over* $\mathcal{D}'$ *satisfies Axioms B.1*-B.8*, then the induced binary relation* $\succsim_{**}$ *over* $\mathcal{D}''$ *satisfies Axioms B.1**-B.8**.*

Proof. See Appendix D.3. □

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

We are very close to a SIT representation. We proceed by first rescaling every element of $\mathcal{F}''$ with factor $\frac{1}{n}$. The rescaling will give us a unit simplex and make the corresponding domain $\mathcal{D}''$ 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 direction of acts satisfies Axioms 1-8, then the binary relation $\succsim_{**}$ defined over $\mathcal{D}''$ in step 2 satisfies Axioms B.1**-B.8**, and it is uniquely derived from the original binary relation $\succsim$ over directions of acts. Therefore, we can apply Theorem 6 to conclude that there exists 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

$$\widehat{W}(\mathbb{F}'') = \max_{F'' \in \mathbb{F}''} \left[(1+\widehat{K}) \sum_{\widehat{u} \in \text{supp}(\widehat{\mu})} \widehat{\mu}(\widehat{u}) \max_{f'' \in F''} \widehat{u}(f'') \right] - \widehat{K} \sum_{\widehat{u} \in \text{supp}(\widehat{\mu})} \widehat{\mu}(\widehat{u}) \max_{F'' \in \mathbb{F}''} \max_{f'' \in F''} \widehat{u}(f'') \tag{41}$$

where $\widehat{u}(f'')$ is the expected utility for lottery f'' under taste $\widehat{u}$. That is,

$$\widehat{u}(f'') = \sum_{\widehat{\omega} \in \widehat{\Omega}} \widehat{u}(\widehat{\omega}) f''(\widehat{\omega}) \text{ for any } \widehat{u} \in \widehat{\mathcal{U}} \text{ and } f'' \in \mathcal{F}''.$$

Moreover, $\widehat{\mu}$ is uniquely identified (Lemma 13), and $\widehat{K}$ is uniquely identified when $|\text{supp}(\widehat{\mu})| > 1$ (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

$$V(\mathbb{F}) = \max_{F \in \mathbb{F}} \left[(1+K) \sum_{\pi \in \text{supp}(\nu)} \nu(\pi) \max_{f \in F} u_\pi(f_u) \right] - K \sum_{\pi \in \text{supp}(\nu)} \nu(\pi) \max_{F \in \mathbb{F}} \max_{f \in F} u_\pi(f_u) \tag{42}$$

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

$$u_\pi(f_u) := \sum_{\omega \in \Omega} \pi(\omega) f_u(\omega) \text{ for all } f_u \in \mathcal{F}_u \text{ and } \pi \in \Delta(\Omega).$$

We now explore the additional constraint imposed on $\widehat{W}$ by Axiom B.8.

LEMMA 25: *Suppose $\succsim_{**}$ has a representation as defined in equation (41) and $\succsim_u$ satisfies Axiom B.8, then $\widehat{u}(\omega) \geq \widehat{u}(\omega_0)$ for any $\widehat{u} \in \operatorname{supp}(\widehat{\mu})$ and any $\omega \in \Omega$.*

Proof of Lemma 25. Suppose by contradiction that $\succsim_{**}$ can be represented as in equation (41), $\succsim_u$ satisfies Axiom B.8, but there exists $\widehat{u}_* \in \operatorname{supp}(\widehat{\mu})$ and $\omega_* \in \Omega$ such that $\widehat{u}_*(\omega_0) > \widehat{u}_*(\omega_*)$. We want to derive a contradiction. Let

$$f' := (n - \varepsilon, 0, \ldots, 0, \varepsilon, 0, \ldots, 0) \tag{43}$$

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

$$g' := (n, 0, \ldots, 0) \tag{44}$$

where n is assigned to state ω_0 and 0 is assigned to any state $\omega \in \Omega$. 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 by the definition of $\succsim_*$, we must have $\{\{f'\}\} \sim_* \{\{f', g'\}\}$.

Note that for any $\widehat{u} \in \operatorname{supp}(\widehat{\mu})$,

$$\begin{aligned}
\widehat{u}(f') &= (n - \varepsilon)\widehat{u}(\omega_0) + \varepsilon \widehat{u}(\omega_*) \\
\widehat{u}(g') &= n \widehat{u}(\omega_0)
\end{aligned}$$

In particular,

$$\widehat{u}_*(f') - \widehat{u}_*(g') = \varepsilon\Big(\widehat{u}_*(\omega_*) - \widehat{u}_*(\omega_0)\Big) < 0.$$

Therefore,

$$\begin{aligned}
\widehat{W}(\{\{f', g'\}\}) &= \sum_{\widehat{u} \in \operatorname{supp}(\widehat{\mu})} \widehat{\mu}(\widehat{u}) \Big[\widehat{u}(f') \vee \widehat{u}(g')\Big] \\
&= \sum_{\widehat{u} \neq \widehat{u}_*} \widehat{\mu}(\widehat{u}) \Big[\widehat{u}(f') \vee \widehat{u}(g')\Big] + \widehat{\mu}(\widehat{u}_*)\widehat{u}_*(g') \\
&> \sum_{\widehat{u} \neq \widehat{u}_*} \widehat{\mu}(\widehat{u}) \widehat{u}(f') + \widehat{\mu}(\widehat{u}_*)\widehat{u}_*(f') \\
&= \widehat{W}(\{\{f'\}\})
\end{aligned}$$

contradicting $\{\{f', g'\}\} \sim_* \{\{f'\}\}$. This completes the proof of Lemma 25. $\square$

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 simply 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 25, $\xi(\widehat{u}(\omega)) \geq 0$ 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 $\pi_{\widehat{u}}$. This can be done by letting $\pi_{\widehat{u}} \in \Delta(\Omega)$ be defined by

$$\pi_{\widehat{u}}(\omega) := \frac{\xi(\widehat{u}(\omega))}{\sum_{\omega' \in \Omega} \xi(\widehat{u}(\omega'))} \text{ for each } \omega \in \Omega.$$

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

$$\nu(\pi_{\widehat{u}}) := \frac{\sum_{\omega \in \Omega} \xi(\widehat{u}(\omega))}{\sum_{\widehat{u}' \in \text{supp}(\hat{\mu})} \sum_{\omega \in \Omega} \xi(\widehat{u}'(\omega))} \widehat{\mu}(\widehat{u}).$$

This will give us the representation we aim for.

Lastly, we can let the set of signal realizations 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. Note that 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 Lemma 1

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 $(\pi, u, i^\sigma, K^\sigma)$, then by Theorem 1, $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$. We now verify that $\succsim$ satisfies Stable Preference over Acts and Act Independence.

First note that for any act $f \in \mathcal{F}$ and any $\sigma \in \mathcal{I}$, the value of f under W is independent of σ since

$$W(f, \sigma) = \sum_{\omega \in \Omega} \pi(\omega) u(f(\omega)).$$

Therefore, Stable Preference over Acts must be satisfied.

To show that $\succsim$ satisfies Act Independence, note that W is affine in its first argument. That is, for any $\mathbb{F}, \mathbb{G} \in \mathcal{D}$ and any fixed information structure $\sigma \in \mathcal{I}$,

$$W(\alpha\mathbb{F} + (1-\alpha)\mathbb{G}, \sigma) = \alpha W(\mathbb{F}, \sigma) + (1-\alpha) W(\mathbb{G}, \sigma).$$

Therefore, for any $\mathbb{F}, \mathbb{G} \in \mathcal{D}$, $\sigma, \sigma' \in \mathcal{I}$ and $\alpha \in (0,1)$,

$$\begin{aligned}
& (\alpha\mathbb{F} + (1-\alpha)h, \sigma) \succsim (\alpha\mathbb{G} + (1-\alpha)h, \sigma') \\
\iff & W(\alpha\mathbb{F} + (1-\alpha)h, \sigma) \geq W(\alpha\mathbb{G} + (1-\alpha)h, \sigma') \\
\iff & \alpha W(\mathbb{F}, \sigma) + (1-\alpha) W(h, \sigma) \geq \alpha W(\mathbb{G}, \sigma') + (1-\alpha) W(h, \sigma') \\
\iff & W(\mathbb{F}, \sigma) \geq W(\mathbb{G}, \sigma') \\
\iff & (\mathbb{F}, \sigma) \succsim (\mathbb{G}, \sigma')
\end{aligned}$$

where the penultimate equivalence follows from the above result that $W(h, \sigma) = W(h, \sigma')$.

This completes the proof of the "only if" part.

If. Suppose $\succsim$ satisfies Weak Order, Stable Preference for Acts and Act Independence, and $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$, we want to show that $\succsim$ has an aligned informational tradeoff representation.

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

LEMMA 26: *If $\succsim$ satisfies Stable Preference over Acts, then there exists $\pi \in \Delta(\Omega)$ and $u : \Delta(X) \to \mathbb{R}$ 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 26. Note that when restricting to acts,

$$V^\sigma(f) = \sum_{\omega\in\Omega} \pi^\sigma(\omega) u^\sigma(f(\omega)) \text{ for any act } f \in \mathcal{F}.$$

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 directly from the uniqueness result of the Anscombe-Aumann framework (see Kreps (2018) for one treatment). $\square$

We are not quite done yet as we have not established that $\succsim$ actually has a utility representation. We can take care of this using the other axiom, Act Independence. Let $\widetilde{V}^\sigma$ be a SIT representation with parameters $(\pi, u, K^\sigma, i^\sigma)$ where π, u are as characterized in Lemma 26. 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. That is, we want to show that $(\mathbb{F}, \sigma) \succsim (\mathbb{G}, \sigma')$ if and only if $W(\mathbb{F}, \sigma) \geq W(\mathbb{G}, \sigma')$. Fix two pairs $(\mathbb{F}, \sigma)$ and $(\mathbb{G}, \sigma')$.

Lemma 27: *There exist acts f, g, h and $\alpha \in (0, 1]$ such that*

$$(\alpha\mathbb{F} + (1-\alpha)h, \sigma) \sim (f, \sigma)$$
$$(\alpha\mathbb{G} + (1-\alpha)h, \sigma') \sim (g, \sigma')$$

Proof of Lemma 27. It is without loss to normalize the taste function $u : \Delta(X) \to \mathbb{R}$ to have range $[0, 1]$ (one way to do this is to assign utility 0 to the worst outcome in X and utility 1 to the best outcome in X and this is always doable since X is finite). 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 $\sigma \in \mathcal{I}$. Using this, we can show that the value of any menu $F \in \mathcal{M}$ under any information structure is in $[0, 1]$, that is, $0 \leq \widetilde{V}^\sigma(\mathbb{F}) \leq 1$ 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 direction $\mathbb{F}$ and information structure $\sigma \in \mathcal{I}$. Therefore, for any act h and any $\sigma \in \mathcal{I}$,

$$\widetilde{V}^\sigma\Big(\alpha\mathbb{F} + (1-\alpha)h\Big) = \alpha\widetilde{V}^\sigma(\mathbb{F}) + (1-\alpha)\widetilde{V}^\sigma(h)$$
$$\widetilde{V}^{\sigma'}\Big(\alpha\mathbb{G} + (1-\alpha)h\Big) = \alpha\widetilde{V}^{\sigma'}(\mathbb{G}) + (1-\alpha)\widetilde{V}^{\sigma'}(h)$$

By our arguments above, there exists h such that $\widetilde{V}^{\sigma}(h) = \widetilde{V}^{\sigma'}(h) = 1$. Therefore 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

$$\widetilde{V}^{\sigma}\Big(\alpha^*\mathbb{F} + (1-\alpha^*)h\Big) = \alpha^*\widetilde{V}^{\sigma}(\mathbb{F}) + (1-\alpha^*)\widetilde{V}^{\sigma}(h) \geq 0$$
$$\widetilde{V}^{\sigma'}\Big(\alpha^*\mathbb{G} + (1-\alpha^*)h\Big) = \alpha^*\widetilde{V}^{\sigma'}(\mathbb{G}) + (1-\alpha^*)\widetilde{V}^{\sigma'}(h) \geq 0$$

By the surjectivity of u, this guarantees the existence of some acts $f, g \in \mathcal{F}$ such that

$$(\alpha^*\mathbb{F} + (1-\alpha^*)h, \sigma) \sim (f, \sigma)$$
$$(\alpha^*\mathbb{G} + (1-\alpha^*)h, \sigma') \sim (g, \sigma')$$

This completes the proof. $\square$

Fix two pairs $(\mathbb{F}, \sigma)$ and $(\mathbb{G}, \sigma')$. Let f, g, h and $\alpha \in (0,1]$ be as characterized from Lemma 27. 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 directions.

This completes the proof of the sufficiency of the axioms since W is by definition an aligned informational tradeoff representation. $\square$

C.3 Proof of Lemma 2

Proof. If. If i^o induces a degenerate distribution over posteriors, $\succsim^o$ is represented by

$$V^o(\mathbb{F}) = \max_{F\in\mathbb{F}} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)u(f(\omega)). \tag{45}$$

If $V^o(\{F\}) \geq V^o(\{G\})$, then $V^o(\{F \cup G\}) = V^o(\{F\})$, so Strategic Rationality when No Information (SRNI) is satisfied.

Only if. Suppose i^o induces a non-degenerate distribution over posteriors, that is, the support of the induced distribution over posteriors contains at least two different elements, μ_1 and μ_2. Then by a standard result, there exist two acts f and g such that f yields a strictly higher expected utility than g given posterior μ_1 and g yields a strictly higher expected utility than f given posterior μ_2. Therefore, $V^o(\{\{f, g\}\}) > \max\{V^o(\{\{f\}\}), V^o(\{\{g\}\})\}$, violating SRNI. Taking the contrapositive completes the proof. □

C.4 Proof of Lemma 3

Proof. Only if. Suppose $\succsim$ has a regret-varying IT representation, then $\succsim$ has an aligned informational tradeoff representation with $i^\sigma = \sigma$ for each $\sigma \in \mathcal{I}$. Therefore, we can apply Lemma 1 to conclude that $\succsim$ satisfies Weak Order and Act Independence, and $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$. Further more, $\succsim^o$ can be represented by (π, u, K^o, o), therefore, we can apply Lemma 2 to conclude that $\succsim$ satisfies SRNI.

We now check that $\succsim$ must satisfy Reduction. Note that the regret terms are zero for singleton directions $\{F\}$ and $\{F_\sigma\}$. Moreover,

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

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. Therefore, $\succsim$ must satisfy Reduction.

If. Suppose $\succsim$ satisfies Weak Order, Act Independence, SRNI and Reduction, and $\succsim^\sigma$

satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$, we want to show that $\succsim$ has a regret-varying IT representation.

As argued in the main text, if $\succsim$ satisfies Weak Order and Reduction, then $\succsim$ satisfies Stable Preference over Acts. Therefore, we can apply Lemma 1 to conclude that $\succsim$ has an aligned informational tradeoff 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 given π for any $\sigma \in \mathcal{I}$.

By Lemma 2, SRNI implies that i^o coincides with o, that is,

$$W(\mathbb{F}, o) = \max_{F\in\mathbb{F}} \max_{f\in F} \sum_{\omega\in\Omega} \pi(\omega)u(f(\omega)).$$

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 signal realizations 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 signal realizations 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 $W(\{F\}, \sigma) \neq W(\{F_\sigma\}, o)$). This completes the proof. $\square$

C.5 Proof of Theorem 3

Proof. Only if. Suppose $\succsim$ has an information tradeoff representation, then $\succsim$ has a regret-varying IT representation with $K^\sigma = K$ for all $\sigma \in \mathcal{I}$. Therefore, we can apply Lemma 3 to conclude that $\succsim$ satisfies Weak Order, Act Independence, SRNI and Reduction, and $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$.

We now check that $\succsim$ must satisfy Balance. Note that for any menu $F \in \mathcal{M}$ and any information structure $\sigma \in \mathcal{I}$,

$$W(D(F), \sigma) = (1 + K)W(\{F\}, o) - K \cdot W(\{F\}, \sigma). \tag{46}$$

Rearranging, we have

$$W(\{F\}, o) = \frac{1}{1+K} \cdot W(D(F), \sigma) + \frac{K}{1+K} \cdot W(\{F\}, \sigma). \tag{47}$$

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.

Now further suppose that for some $\alpha \in (0, 1]$,

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

Then

$$W(\{F\}, o) = \alpha \cdot \widetilde{W}(D(F), \sigma) + (1-\alpha) \cdot W(\{F\}, \sigma). \tag{48}$$

For equations (47) and (48) 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 (47) is an identity that holds for all pairs $(F, \sigma) \in \mathcal{M} \times \mathcal{I}$.

If. Suppose $\succsim$ satisfies Weak Order, Act Independence, SRNI, Reduction and Balance, and $\succsim^\sigma$ satisfies Axioms 2-8 for each $\sigma \in \mathcal{I}$. We want to show that $\succsim$ has an IT representation.

First, we can apply Lemma 3 to conclude that $\succsim$ has a regret-varying IT representation with parameters $\big(\pi, u, (K^\sigma)_{\sigma \in \mathcal{I}}\big)$. We want to show that if $\succsim$ also satisfies Balance, then there exists $K \geq 0$ such that setting $K^\sigma = K$ for all $\sigma \in \mathcal{I}$ will deliver the same utility representation as W.

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. So we only need to worry about information structures that induce non-degenerate distributions over posteriors.

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 menus F and 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. This completes the proof of the sufficiency of the axioms for the existence of an IT representation.

The uniquenuss of an IT representation is an immediate corollary of the uniqueness of the SIT represenation for each $\succsim^\sigma$. $\square$

C.6 Proof of Theorem 5

Proof of Theorem 5. Since $\succsim_1$ and $\succsim_2$ agree on their preference over acts, by a standard result, $\pi_1 = \pi_2$, $u_2 = au_1 + b$ for some $a > 0$ and $b \in \mathbb{R}$. Therefore, (π_1, u_1, K_2) also represents $\succsim_2$. Let W_1 be the functional representing $\succsim_1$ with (π_1, u_1, K_1) and W_2 be the functional representing $\succsim_2$ with (π_1, u_1, K_2).

We first show that 2 implies 1. 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 that

$$W_2(\mathbb{F}, \sigma) - W_2(\mathbb{F}, \sigma') \geq 0 \Longrightarrow W_1(\mathbb{F}, \sigma) - W_1(\mathbb{F}, \sigma') \geq 0 \tag{49}$$

It suffices to show that

$$\frac{W_1(\mathbb{F}, \sigma) - W_1(\mathbb{F}, \sigma')}{1 + K_1} - \frac{W_2(\mathbb{F}, \sigma) - W_2(\mathbb{F}, \sigma')}{1 + K_2} \geq 0 \tag{50}$$

The inequality in (50) holds with the more detailed algebra manipulation as follows.

$$\begin{aligned}
&\frac{W_1(\mathbb{F}, \sigma) - W_1(\mathbb{F}, \sigma')}{1 + K_1} - \frac{W_2(\mathbb{F}, \sigma) - W_2(\mathbb{F}, \sigma')}{1 + K_2} && (51)\\
=& \left(\frac{W_1(\mathbb{F}, \sigma)}{1 + K_1} - \frac{W_2(\mathbb{F}, \sigma)}{1 + K_2}\right) - \left(\frac{W_1(\mathbb{F}, \sigma')}{1 + K_1} - \frac{W_2(\mathbb{F}, \sigma')}{1 + K_2}\right) && (52)\\
=& -\frac{K_1}{1 + K_1} \sum_{s \in S} \sigma(s) \max_{f \in \mathbb{F}} \ U(f, \mu_s^\sigma) + \frac{K_2}{1 + K_2} \sum_{s \in S} \sigma(s) \max_{f \in \mathbb{F}} \ U(f, \mu_s^\sigma) && (53)\\
&+ \frac{K_1}{1 + K_1} \sum_{s' \in S'} \sigma'(s') \max_{f \in \mathbb{F}} U(f, \mu_{s'}^{\sigma'}) - \frac{K_2}{1 + K_2} \sum_{s' \in S'} \sigma'(s') \max_{f \in \mathbb{F}} \ U(f, \mu_{s'}^{\sigma'}) && (54)\\
=& \left(\frac{K_1}{1 + K_1} - \frac{K_2}{1 + K_2}\right) \left[\sum_{s' \in S'} \sigma'(s') \max_{f \in \mathbb{F}} \ U(f, \mu_{s'}^{\sigma'}) - \sum_{s \in S} \sigma(s) \max_{f \in \mathbb{F}} \ U(f, \mu_s^\sigma)\right] && (55)\\
\geq& \ 0 && (56)
\end{aligned}$$

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.

We then show that 1 implies 2. Suppose $\succsim_1$ is more information averse than $\succsim_2$.

By way of contradiction, suppose $K_2 > K_1$, we want to construct some $\mathbb{F}$ such that

$$(\mathbb{F}, o) \succsim_2 (\mathbb{F}, i) \text{ but } (\mathbb{F}, i) \succ_1 (\mathbb{F}, o) \tag{57}$$

where o is the fully uninformative 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 $f, g, h \in \mathcal{F}$:

$$\begin{array}{c|ccc} & u\circ f & u\circ g & u\circ h \\ \hline \omega & \frac{1}{\eta} & 0 & 1 \\ \omega' & 0 & y & 1 \\ \hat{\omega}\in\Omega\setminus\{\omega,\omega'\} & 0 & 0 & 0 \end{array} \tag{58}$$

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

$$W_1(\mathbb{F}, o) = W_2(\mathbb{F}, o) = \max\left\{\pi \cdot \frac{1}{\eta}, (1-\eta)y, 1\right\} = 1 \tag{59}$$

$$W_1(\mathbb{F}, i) = 1 + (1-\eta)y - (1-\eta)K_1(1-y) > 1 \tag{60}$$

$$W_2(\mathbb{F}, i) = 1 + (1-\eta)y - (1-\eta)K_2(1-y) \leq 1 \tag{61}$$

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

□

C.7 Proof of Lemma 5

Proof. Combining equations (22) and (23), we see that

$$\begin{aligned}\widetilde{U}(F, \mathbb{F}, \mu_s^\sigma) = {} & (1+K_1+K_2)\max_{f\in F}\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)u(f(\omega)) \\ & - K_1\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)\max_{G\in\mathbb{F}}\max_{g\in F}u(g(\omega)) - K_2\sum_{\omega\in\Omega}\mu_s^\sigma(\omega)\max_{h\in F}u(h(\omega))\end{aligned} \tag{62}$$

Plug this back to (21), we can write W_1 as

$$\begin{aligned}W_1(\mathbb{F}, \sigma) = {} & \max_{F\in\mathbb{F}}\Bigg[\sum_{s\in S}(1+K_0+K_1+K_2)\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega)) \\ & - K_2\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)\max_{h\in F}u(h(\omega))\Bigg] \\ & - K_0\sum_{s\in S}\max_{G\in\mathbb{F}}\max_{g\in G}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(g(\omega)) \\ & - K_1\sum_{\omega\in\Omega}\pi(\omega)\max_{G\in\mathbb{F}}\max_{g\in G}u(g(\omega))\end{aligned} \tag{63}$$

If $K_2 = 0$, this reduces to

$$\begin{aligned}W_1(\mathbb{F}, \sigma) = {} & (1+K_0+K_1)\max_{F\in\mathbb{F}}\sum_{s\in S}\max_{f\in F}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(f(\omega)) \\ & - K_0\sum_{s\in S}\max_{G\in\mathbb{F}}\max_{g\in G}\sum_{\omega\in\Omega}\pi(\omega)\sigma(s\mid\omega)u(g(\omega)) \\ & - K_1\sum_{\omega\in\Omega}\pi(\omega)\max_{G\in\mathbb{F}}\max_{g\in G}u(g(\omega))\end{aligned} \tag{64}$$

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 $K = K_0/(1+K_1)$. □

D Translation from Lotteries to Acts

D.1 Proof of Lemma 22

Proof. We can check that the operation $u(\cdot)$ that maps a primitive (acts, menus or directions) into its corresponding primitive as in utilities (utility acts, utility menus and utility directions) behave nicely with set operations and respects linearity. This guarantees that each Axiom from 1-6 and 8 implies their counterpart in Axiom B.1-B.6 and B.8.

Note that the statement of Axiom B.7 is slightly different from the statement of Axiom 7. We now check that Axioms 1-8 implies Axiom B.7.

First notice that by Axiom 7, there exists lotteries $\ell, \ell' \in \Delta(X)$ such that $\ell \succ \ell'$. Since these are constant acts, it must be that ℓ dominates ℓ'. Then by Axiom 8, $\{\{\ell\},\{\ell'\}\} \sim \{\{\ell\}\}$. Then by transitivity, $\{\{\ell\},\{\ell'\}\} \succ \{\{\ell'\}\}$. Let $\mathbb{F} = \{\{\ell\},\{\ell'\}\}$ and $\mathbb{G} = \{\{\ell'\}\}$. Then $\mathbb{F} \supseteq \mathbb{G}$ and $\mathbb{F} \succ \mathbb{G}$.

Then we can apply the definition of $\succsim_u$ to conclude that $u(\mathbb{F}), u(\mathbb{G}) \in \mathcal{D}_u$, and $u(\mathbb{F}) \supseteq u(\mathbb{G})$ with $u(\mathbb{F}) \succ_u u(\mathbb{G})$, Axiom B.7 is satisfied. □

D.2 Proof of Lemma 23

Proof that Axiom B.1 implies Axiom B.1.*

Fix any $\mathbb{F}', \mathbb{G}' \in \mathcal{D}'$. Then by the completeness of $\succsim_u$, either $r(\mathbb{F}') \succsim_u r(\mathbb{G}')$ or $r(\mathbb{G}') \succsim_u r(\mathbb{F}')$ or both, which further indicates that either $\mathbb{F}' \succsim_* \mathbb{G}'$ or $\mathbb{G}' \succsim_* \mathbb{F}'$ or both.

For transitivity, suppose $\mathbb{F}' \succsim_* \mathbb{G}'$ and $\mathbb{G}' \succsim_* \mathbb{H}'$, then by definition, $r(\mathbb{F}') \succsim_u r(\mathbb{G}')$ and $r(\mathbb{G}') \succsim_u r(\mathbb{H}')$, which implies $r(\mathbb{F}') \succsim_u r(\mathbb{H}')$ (by the transitivity of $\succsim_u$), which further implies that $\mathbb{F}' \succsim_* \mathbb{H}'$. $\square$

Proof that Axiom B.2 implies Axiom B.2.*
Fix any $\mathbb{F}' \in \mathcal{D}'$, consider the set $\{\mathbb{G}' \in \mathcal{D}' \mid \mathbb{F}' \succsim_* \mathbb{G}'\}$. By Axiom B.2, the set $\{\mathbb{G}_u \in \mathcal{D}_u \mid r(\mathbb{F}') \succsim_u \mathbb{G}\}$ is closed. Since r is a homeomorphism, it suffices to show that

$$\{\mathbb{G}' \in \mathcal{D}' \mid \mathbb{F}' \succsim_* \mathbb{G}'\} = r^{-1}\Big(\{\mathbb{G}_u \in \mathcal{D}_u \mid r(\mathbb{F}') \succsim_u \mathbb{G}_u\}\Big).$$

If $\mathbb{F}' \succsim_* \mathbb{G}'$, then $r(\mathbb{F}') \succsim_u r(\mathbb{G}')$, thus LHS $\subseteq$ RHS. If $r(\mathbb{F}') \succsim_u \mathbb{G}_u$, then $\mathbb{F}' \succsim_* r^{-1}(\mathbb{G}_u)$, thus LHS $\supseteq$ RHS. The arguments for $\{\mathbb{G}' \in \mathcal{D}' \mid \mathbb{G}' \succsim_* \mathbb{F}'\}$ are similar. $\square$

Proof that Axiom B.3 implies Axiom B.3.* First note that r is linear, that is, for any $\mathbb{F}', \mathbb{G}' \in \mathcal{D}'$ and $\alpha \in [0,1]$,

$$r(\alpha\mathbb{F}' + (1-\alpha)\mathbb{G}') = \alpha r(\mathbb{F}') + (1-\alpha)r(\mathbb{G}').$$

Then notice that r^{-1} is also linear. For any $\mathbb{F}, \mathbb{G} \in \mathcal{D}$ and $\alpha \in [0,1]$,

$$r^{-1}(\alpha\mathbb{F} + (1-\alpha)\mathbb{G}) = \alpha r^{-1}(\mathbb{F}) + (1-\alpha)r^{-1}(\mathbb{G}).$$

Therefore, for any $\mathbb{F}', \mathbb{G}', \mathbb{H}' \in \mathcal{D}'$ and any $\alpha \in (0,1]$,

$$\begin{aligned}
\mathbb{F}' \succsim_* \mathbb{G}' &\iff r(\mathbb{F}') \succsim_u r(\mathbb{G}') \\
&\iff \alpha r(\mathbb{F}') + (1-\alpha)r(\mathbb{H}') \succsim_u \alpha r(\mathbb{G}') + (1-\alpha)r(\mathbb{H}') \\
&\iff r(\alpha\mathbb{F}' + (1-\alpha)\mathbb{H}') \succsim_u r(\alpha\mathbb{G}' + (1-\alpha)\mathbb{H}') \\
&\iff \alpha\mathbb{F}' + (1-\alpha)\mathbb{H}' \succsim_* \alpha\mathbb{G}' + (1-\alpha)\mathbb{H}'
\end{aligned}$$

This is essentially the same proof as the proof of Claim 3 in DLST (2014). $\square$

Proof that Axiom B.4 implies Axiom B.4.*
Given $\succsim_u$, let N be a natural number satisfying Axiom B.4.

Fix any $\mathbb{F}' \in \mathcal{D}'$, by Axiom B.4, there exists $\mathbb{G}_u \in \mathcal{D}_u$ such that $\mathbb{G}_u$ is critical for $r(\mathbb{F}')$ with $|\mathbb{G}_u| < N$. Let $\mathbb{G}' = r^{-1}(\mathbb{G}_u)$. Then $|\mathbb{G}'| = |\mathbb{G}_u| < N$, and for any $\mathbb{H}'$ satisfying $\mathbb{G}' \subseteq \mathbb{H}' \subseteq \mathbb{F}'$, we have

$$r(\mathbb{G}') \subseteq r(\mathbb{H}') \subseteq r(\mathbb{F}') \Longrightarrow r(\mathbb{G}') \sim_u r(\mathbb{H}') \sim_u r(\mathbb{F}') \Longrightarrow G' \sim_* \mathbb{H}' \sim_* \mathbb{F}'.$$

Hence $\mathbb{G}'$ is critical for $\mathbb{F}'$ with $|\mathbb{G}'| < N$. Similar arguments for the second half. $\square$

Proof that Axiom B.5 implies Axiom B.5.*

First notice that for any $F' \in \mathcal{M}'$, $r(\{F'\}) = \{r(F')\}$ by construction. Therefore,

$$\begin{aligned}\{F'\} \succsim_* \{G'\}, F' \in \mathbb{F}' &\Longrightarrow \{r(F')\} \succsim_u \{r(G')\}, r(F') \in r(\mathbb{F}') \\ &\Longrightarrow r(\mathbb{F}') \succsim_u r(\mathbb{F}') \cup \{r(G')\} \\ &\Longrightarrow r(\mathbb{F}') \succsim_u r(\mathbb{F}' \cup \{G'\}) \Longrightarrow \mathbb{F}' \succsim_* \mathbb{F}' \cup \{G'\}\end{aligned}$$

This completes the proof. $\square$

Proof that Axiom B.6 implies Axiom B.6.*

First note that

$$r(\mathbb{F}' \cup \{F' \cup G'\}) = r(\mathbb{F}') \cup \{r(F') \cup r(G')\}$$

and that

$$r(\mathbb{F}' \cup \{F', G'\}) = r(\mathbb{F}') \cup \{r(F'), r(G')\}.$$

Since $r(\mathbb{F}') \in \mathcal{D}_u$, and $r(F'), r(G') \in \mathcal{M}_u$, then we can use Axiom B.6 to conclude that

$$r(\mathbb{F}') \cup \{r(F') \cup r(G')\} \succsim_u r(\mathbb{F}') \cup \{r(F'), r(G')\},$$

which further indicates that $\mathbb{F}' \cup \{F' \cup G'\} \succsim_* \mathbb{F}' \cup \{F', G'\}$. $\square$

Proof that Axiom B.7 implies Axiom B.7.*

By Axiom B.7, there exists $\mathbb{F}_u, \mathbb{G}_u \in \mathcal{D}_u$ such that $\mathbb{F}_u \supseteq \mathbb{G}_u$ and $\mathbb{F}_u \succ_u \mathbb{G}_u$. Therefore, $r^{-1}(\mathbb{F}_u), r^{-1}(\mathbb{G}_u) \in \mathcal{D}'$ satisfy $r^{-1}(\mathbb{F}_u) \supseteq r^{-1}(\mathbb{G}_u)$ and $r^{-1}(\mathbb{F}_u) \succ_* r^{-1}(\mathbb{G}_u)$. $\square$

Proof that Axiom B.8 implies Axiom B.8.*

Suppose $G' \subseteq F'$, then $r(G') \subseteq r(F')$, and by the definition of domination, $r(F')$ dominates $r(G')$. Suppose $F' \in \mathbb{F}'$, then $r(F') \in r(\mathbb{F}')$, and the second part of Axiom B.8 implies that $r(\mathbb{F}') \sim_u r(\mathbb{F}') \cup \{r(G')\}$, which further indicates that $\mathbb{F}' \sim_* \mathbb{F}' \cup \{G'\}$. $\square$

D.3 Proof of Lemma 24

Before we move on to study the properties of $\succsim_{**}$, we have the following simplifying result to help up use the definition of $\succsim_{**}$ more easily.

LEMMA 28: *Suppose $\succsim_u$ satisfies Axioms B.1-B.3 (Weak Order, Continuity and Independence), then $\succsim_{**}$ constructed as in Step 2 will satisfy*

$$\mathbb{F}'' \succsim_{**} \mathbb{G}'' \iff \frac{1}{n^2}\mathbb{F}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\} \succsim_* \frac{1}{n^2}\mathbb{G}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\}.$$

Proof of Lemma 28. Since $\succsim_u$ satisfies Axioms B.1-B.3, then by Lemma 23, $\succsim_*$ satisfies Axioms B.1*-B.3*. Recall that

$$F^{n+1} = \left\{ \left(\frac{n}{n+1}, \dots, \frac{n}{n+1} \right) \right\}.$$

That is, F^{n+1} is the singleton menu containing only a "uniform lottery."

"$\Longleftarrow$:" Suppose $\frac{1}{n^2}\mathbb{F}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\} \succsim_* \frac{1}{n^2}\mathbb{G}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\}$. We know that $\{F^{n+1}\} \in \mathcal{D}'$ and $\succsim_*$ satisfies B.3*. Therefore, for any $\varepsilon \in [0, 1/n^2)$, $\alpha := n^2\varepsilon \in [0, 1)$ and

$$\begin{aligned}
\varepsilon\mathbb{F}'' + (1-\varepsilon)\{F^{n+1}\} &= \alpha\left(\frac{1}{n^2}\mathbb{F}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\}\right) + (1-\alpha)\{F^{n+1}\} \\
&\succsim_* \alpha\left(\frac{1}{n^2}\mathbb{G}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\}\right) + (1-\alpha)\{F^{n+1}\} \\
&= \varepsilon\mathbb{G}'' + (1-\varepsilon)\{F^{n+1}\}
\end{aligned}$$

"$\Longrightarrow$:" Suppose $\mathbb{F}'' \succsim_{**} \mathbb{G}''$, then we can find a sequence of ε_n converging from below to $1/n^2$, then the result follows from the continuity of $\succsim_*$. $\square$

An immediate but important implication of Lemma 25 is that the constructed binary relation $\succsim_{**}$ is uniquely determined by $\succsim_*$.

Moving forward, we will do the operation "mix $\mathbb{F}''$ with $\{F^{n+1}\}$ on weight $1/n^2$" a lot in the following proofs. For an easier exposition, we give this operation a name by defining $t : \mathcal{D}'' \to \mathcal{D}'$ given by

$$t(\mathbb{F}'') := \frac{1}{n^2}\mathbb{F}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\} \tag{65}$$

Proof that Axioms B.1-B.3* imply Axiom B.1**.*
Completeness: Fix any $\mathbb{F}'', \mathbb{G}'' \in \mathcal{D}''$, $t(\mathbb{F}''), t(\mathbb{G}'') \in \mathcal{D}'$. By the completeness of $\succsim_*$, $t(\mathbb{F}') \succsim_* t(\mathbb{G}')$ or $t(\mathbb{G}') \succsim_* t(\mathbb{F}')$ or both. By Lemma 25, Axioms B.1*-B.3* guarantee that $\mathbb{F}'' \succsim_{**} \mathbb{G}''$ if and only if $t(\mathbb{F}'') \succsim_* t(\mathbb{G}'')$. Thus, $\succsim_{**}$ is complete.

Transitivity:

$$\begin{aligned}
\mathbb{F}'' \succsim_{**} \mathbb{G}'',\ \mathbb{G}'' \succsim_{**} \mathbb{H}'' &\Longrightarrow t(\mathbb{F}'') \succsim_* t(\mathbb{G}''),\ t(\mathbb{G}'') \succsim_* t(\mathbb{H}'') \\
&\Longrightarrow t(\mathbb{F}'') \succsim_* t(\mathbb{H}'') \\
&\Longrightarrow \mathbb{F}'' \succsim_{**} \mathbb{H}''
\end{aligned}$$

where the second implication follows from $\succsim_*$ being transitive. $\square$

Proof that Axioms B.1-B.3* imply Axiom B.2**.*

We consider the lower contour set. Same arguments work for the upper contour set.

Fix any sequence $\{\mathbb{G}''_m\}_{m=1}^{\infty} \subset \{\mathbb{G}'' \mid \mathbb{F}'' \succsim_{**} \mathbb{G}''\}$, that is, $\mathbb{F}'' \succsim_{**} \mathbb{G}''_m$ for all $m \in \mathbb{N}$, which further indicates that $t(\mathbb{F}'') \succsim_{*} t(\mathbb{G}''_m)$ for all $m \in \mathbb{N}$. If $\mathbb{G}''_m \to \mathbb{G}''_*$ as $m \to \infty$ for some $\mathbb{G}''_* \in \mathcal{D}''$, then by construction, $t(\mathbb{G}''_m) \to t(\mathbb{G}''_*)$. Since $\succsim_*$ satisfies Continuity (Axiom B.2*), this implies that $t(\mathbb{F}'') \succsim_* t(\mathbb{G}''_*)$. Thus, $\mathbb{F}'' \succsim_{**} \mathbb{G}''_*$. □

Proof that Axioms B.1-B.3* imply Axiom B.3**.*

First note that the mapping $t : \mathcal{D}'' \to \mathcal{D}'$ is linear, that is, for any $\alpha \in [0,1]$,

$$\begin{aligned} t\Big(\alpha\mathbb{F}'' + (1-\alpha)\mathbb{G}''\Big) &= \frac{1}{n^2}\Big(\alpha\mathbb{F}'' + (1-\alpha)\mathbb{G}''\Big) + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\} \\ &= \frac{1}{n^2}\Big(\alpha\mathbb{F}'' + (1-\alpha)\mathbb{G}''\Big) + \left(1 - \frac{1}{n^2}\right)\Big(\alpha\{F^{n+1}\} + (1-\alpha)\{F^{n+1}\}\Big) \\ &= \alpha t(\mathbb{F}'') + (1-\alpha)t(\mathbb{G}'') \end{aligned}$$

Note that the second equality only goes through because F^{n+1} is a singleton menu. The decomposition is not generally doable for non-singleton menus.

Using the linearity of t, we prove the following stronger version of Independence. For any $\mathbb{F}'', \mathbb{G}'', \mathbb{H}''$ and any $\alpha \in (0,1]$,

$$\begin{aligned} \mathbb{F}'' \succsim_{**} \mathbb{G}'' &\iff t(\mathbb{F}'') \succsim_* t(\mathbb{G}'') \\ &\iff \alpha t(\mathbb{F}'') + (1-\alpha)t(\mathbb{H}'') \succsim_* \alpha t(\mathbb{G}'') + (1-\alpha)t(\mathbb{H}'') \\ &\iff t\Big(\alpha\mathbb{F}'' + (1-\alpha)\mathbb{H}''\Big) \succsim_* t\Big(\alpha\mathbb{G}'' + (1-\alpha)\mathbb{H}''\Big) \\ &\iff \alpha\mathbb{F}'' + (1-\alpha)\mathbb{H}'' \succsim_{**} \alpha\mathbb{G}'' + (1-\alpha)\mathbb{H}'' \end{aligned}$$

where the first and the last equivalence follow from Lemma 25, the second equivalence follows from $\succsim_*$ satisfying Axiom B.3* and the third equivalence follows from the linearity of t we showed above. □

Proof that Axioms B.1-B.4* imply Axiom B.4**.*

First note that the mapping t respects set inclusion, that is, if $\mathbb{G}'' \subseteq \mathbb{F}''$, then

$$t(\mathbb{G}'') = \frac{1}{n^2}\mathbb{G}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\} \subseteq \frac{1}{n^2}\mathbb{F}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\} = t(\mathbb{F}'').$$

Moreover, $|t(\mathbb{F}'')| = |\mathbb{F}''|$ for any finite $\mathbb{F}''$.

Now Axiom B.4 implies Axiom B.4*, which guarantees the existence of a natural number N satisfying the conditions for $\succsim_*$ over $\mathcal{D}'$ corresponding to the two conditions above. We argue the same N works for $\succsim_{**}$.

Fix any $\mathbb{F}'' \in \mathcal{D}''$, $t(\mathbb{F}'') \in \mathcal{D}'$, and thus by Axiom B.4*, there exists $\mathbb{G}'$ such that $\mathbb{G}'$ is critical for $t(\mathbb{F}'')$ and $|\mathbb{G}'| < N$. Since $\mathbb{G}' \subseteq t(\mathbb{F}'')$, there exists $\mathbb{G}'' \subseteq \mathbb{F}''$ such that $\mathbb{G}' = t(\mathbb{G}'')$. So $|\mathbb{G}''| = |\mathbb{G}'| < N$, and we argue that $\mathbb{G}''$ is critical for $\mathbb{F}''$. Fix any $\mathbb{H}''$ such that $\mathbb{G}'' \subseteq \mathbb{H}'' \subseteq \mathbb{F}''$, then $\mathbb{G}' = t(\mathbb{G}'') \subseteq t(\mathbb{H}'') \subseteq t(\mathbb{F}'')$. This implies that $G' \sim_* t(\mathbb{H}'') \sim_* t(\mathbb{F}'')$ since $\mathbb{G}'$ is critical for $t(\mathbb{F}'')$, which further implies that $\mathbb{H}'' \sim_{**} \mathbb{F}''$.

Similar arguments work for the second condition. □

Proof that Axioms B.1-B.3* and B.5* imply Axiom B.5**.*
First note that the mapping $t : \mathcal{D}'' \to \mathcal{D}'$ respects set unions: For any $\mathbb{F}'', \mathbb{G}'' \in \mathcal{D}''$,

$$\begin{aligned} t\Big(\mathbb{F}'' \cup \mathbb{G}''\Big) &= \frac{1}{n^2}\Big(\mathbb{F}'' \cup \mathbb{G}''\Big) + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\} \\ &= \left(\frac{1}{n^2}\mathbb{F}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\}\right) \cup \left(\frac{1}{n^2}\mathbb{G}'' + \left(1 - \frac{1}{n^2}\right)\{F^{n+1}\}\right) \\ &= t(\mathbb{F}'') \cup t(\mathbb{G}'') \end{aligned}$$

Now suppose $\mathbb{F}''$ and $\mathbb{G}''$ satisfy that for any $G'' \in \mathbb{G}''$, there exists $F'' \in \mathbb{F}''$ such that $\{F''\} \succsim_{**} \{G''\}$, we want to show that $\mathbb{F}'' \succsim_{**} \mathbb{F}'' \cup \mathbb{G}''$.

For any $G' \in t(\mathbb{G}'')$, there exists $G'' \in \mathbb{G}''$ such that $\{G'\} = t(\{G''\})$. By assumption, there exists $F'' \in \mathbb{F}''$ such that $\{F''\} \succsim_{**} \{G''\}$. Then $t(\{F''\}) \succsim_* t(\{G''\}) = \{G'\}$ and $t(F'') \in t(\mathbb{F}'')$. (This is an abuse of notation, by $t(F'')$ we mean the only menu contained in $t(\{F''\})$.) Then the assumption in Axiom B.5* is satisfied and we can conclude that $t(\mathbb{F}'') \succsim_* t(\mathbb{F}'') \cup t(\mathbb{G}'') = t(\mathbb{F}'' \cup \mathbb{G}'')$, which further indicates that $\mathbb{F}'' \succsim_{**} \mathbb{F}'' \cup \mathbb{G}''$. □

Proof that Axioms B.1-B.3* and B.6* imply Axiom B.6**.*
We first formalize the notation (abuse) appeared above. Let $F'' \in \mathcal{M}''$ be a menu, then

$$t(F'') := \frac{1}{n^2}F'' + \left(1 - \frac{1}{n^2}\right)F^{n+1} \in \mathcal{M}'.$$

And $t(F'' \cup G'') = t(F'') \cup t(G'')$ for all $F'', G'' \in \mathcal{M}''$. Now for any $\mathbb{F}'' \in \mathcal{D}''$ and any $F'', G'' \in \mathcal{M}''$, $t(\mathbb{F}'') \in \mathcal{D}'$, $t(F''), t(G'') \in \mathcal{M}'$. Thus, by Axiom 6*,

$$t(\mathbb{F}'') \cup \{t(F'') \cup t(G'')\} \succsim_* t(\mathbb{F}'') \cup \{t(F''), t(G'')\}.$$

But $t\Big(\mathbb{F}'' \cup \{F'' \cup G''\}\Big) = t(\mathbb{F}'') \cup \{t(F'' \cup G'')\} = t(\mathbb{F}'') \cup \{t(F'') \cup t(G'')\}$, and similar for the RHS. Therefore, $\mathbb{F}'' \cup \{F'' \cup G''\} \succsim_{**} \mathbb{F}'' \cup \{F'', G''\}$. □

Proof that Axioms B.1-B.3* and B.7* imply Axiom B.7**.* If $F'' \supseteq G''$, then $t(F'') \supseteq t(G'')$, and by Axiom 7*, $t(\mathbb{F}'') \cup \{t(F''), t(G'')\} \succsim_* t(\mathbb{F}'') \cup \{t(F'')\}$. □

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.

BÉNABOU, R. AND J. TIROLE (2011): "Identity, Morals, and Taboos: Beliefs as Assets," *The Quarterly Journal of Economics*, 126, 805–855.

BLACKWELL, D. (1951): "Comparison of Experiments," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93–102.

——— (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 265–272.

BROCK, T. C. AND J. L. BALLOUN (1967): "Behavioral Receptivity to Dissonant Information," *Journal of Personality and Social Psychology*, 6, 413.

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 M. Rosch (1984): "Information Seeking After Decisions: The Roles of Novelty of Information and Decision Reversibility," *Personality and Social Psychology Bulletin*, 10, 91–98.

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.

Frey, D. and R. A. Wicklund (1978): "A Clarification of Selective Exposure: The Impact of Choice," *Journal of Experimental Social Psychology*, 14, 132–139.

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.

Kopylov, I. and J. Noor (2018): "Commitments and Weak Resolve," *Economic Theory*, 66, 1–19.

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

Krähmer, D. and R. Stone (2013): "Anticipated Regret as an Explanation of Uncertainty Aversion," *Economic Theory*, 52, 709–728.

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

Kreps, D. M. (1979): "A Representation Theorem for "Preference for Flexibility"," *Econometrica: Journal of the Econometric Society*, 565–577.

Kreps, D. M. and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica: Journal of the Econometric Society*, 185–200.

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.

Takeoka, N. (2006): "Representation with a Unique Subjective Decision Tree," *working paper*.