Essays on Decision under Uncertainty

Aoyama, Tomohito

(Degree)
博士（経営学）
(Date of Degree)
2020-03-25
(Date of Publication)
2021-03-01
(Resource Type)
doctoral thesis
(Report Number)
甲第7686号
(URL)
https://hdl.handle.net/20.500.14094/D1007686

博士論文

Essays on Decision under Uncertainty

2020 年 1 月 17 日

神戸大学経営学研究科

宮原泰之研究室

経営学専攻

学籍番号 152B001B

氏名 青山知仁

Essays on Decision under Uncertainty

Tomohito Aoyama

January 17, 2020

Abstract

This essay consists of three papers, which study decision under uncertainty. The first paper uses reponse time and choice as the data. It first elicit subjective filtration that describes a rought sketch of human coginitive process. Then, imposing intuitive axioms, I identify the rest of parameters: subjective probability, expected utility, and cost function. This result implies that using response time somewhat helps us understand cognition.

The second paper is on rational inattention, which studies bounded rationality. This generalizes an existing study by allowing infinite state space. It serves as a pre-analysis for the first paper.

The third paper is on ambiguity, or a type of uncertainty that can not be described with single probability. Building on a variant of smooth ambiguity model, we define unambiguous events in terms of model. Then, I characterize a model of decision under ambiguity under which exogenously given events are unambiguous in this sense.

Acknowlegement

I am grateful to Koji Abe for his advice and warm encouragement. Not only him, I also thank Hideo Suehiro, Hisao Hisamoto, and Yasuyuki Miyahara for leading me to economics. I also thank the participants of Decision Theory Workshop for helpful comments, and instructors of Osaka University for letting me attend their amazing classes.

Contents

Chapter 1

Introduction

This essay contains three papers on axiomatic decision theory, a method in economics to analyze what we can learn from observable data, concering with choices that individual makes. In this chapter, I review the decision theoretic literature and highlight my contribution in this field.

1.1 Literature review

Modeling of belief under uncertainty and providing its foundation is a central goal of decision theory. The standard model of belief is *subjective probability*. So-called subjective expected utility (SEU) model is a class of utility function that calculates expected utility of each alternative using a subjective probability. Since the analytical method and fundamental concept of decision theory were prepared through the studies of subjective probability, I first review its history.

While subjective probability theory is normatively appealing, it is criticized from various viewpoints. An important one is that observed behavior of economic agents deviates from a prediction of subjective probability when the assessment of the probability of events is difficult. Such a class of uncertainty is called *ambiguity*. Chapter 4 of this thesis studies a model of ambiguity. Thus I review ambiguity literature.

Studies of subjective probability and ambiguity consider rational agents, while it is said that economic agents in the real world are only boundedly rational. That is, they do not follow the usual principle of expected utility maximization. Acknowledging the limits of prediction power of conventional models, decision theoretic analysis of bounded rationality emerged. Chapter 3 in this thesis conducts an axiomatic study of bounded rationality. So here, I review this literature.

Behavioral economics is becoming closer to psychology, both methodologically and conceptually. In this trend, economic researchers are accepting non-choice data such as response time, fMRI data, and gaze data. Chapter 2 studies how we can use response time data, together with choice data, to understand the decision process. So here, I review existing decision theoretic literature on non-choice data.

1.1.1 Subjective probability

Modern decision theory starts from De Finetti (1931), which is the first formal treatment of subjective probability. Probability theory in mathematics treats state space, of which subsets are interpreted as events, and measures on it are interpreted as a probability. But it is unclear whether uncertainty should be modeled as a probability. de Finetti asked the appropriateness of this formalization. To answer this problem, he analyzed binary relation over events, which is called *qualitative probabilty*. Qualitative probability represents judgments on which event is more probable given two events. de Finetti imposed several axioms on binary relation and showed they are equivalent to the existence of a probability measure that is consistent with the qualitative probability: more probable events have higher probabilities. His axioms can be interpreted normatively. In a nutshell, de Finetti showed that modeling uncertainty is appropriate if you want to abide by his axioms. Moreover, he initiated axiomatic decision theory by giving its methodology.

Meanwhile, Ramsey (1926) gave another methodological foundation of decision theory: identification of mental concept. In modern language, he axiomatized subjective expected utility model in terms of preference relation, while his analysis is rough. His motivation is to criticize Keynes, who considered probabilistic logical relation: if P holds, then Q with some probability. Keynes says there are probabilistic logical relations and claims it is objective; that is, the probabilities are the same for everyone, which was what Ramsey criticized. He assumed that analyst can observe preferences for objective lotteries with two outcomes, where each outcome obtain with the probability of one half. Assuming expected utility representation, an analyst can calibrate expected utility function. Next, he considers choice for contingent bets in the following form: you get x yen if a proposition is true, while you get y yen otherwise. Given the calibrated expected utility function, the analyst can calibrate the subjective probability. In this way, Ramsey showed that one can elicit subjective probability from choice behavior. This is the first work that connected subjective probability as a mental concept to observable choice behavior.

von Neumann and Morgenstern (1947) is the book that initiated game theory and expected utility (EU) theory as its foundation. In the economic literature before vNM, utility function was an ordinal concept. In contrast to Ramsey, expected utility theory utilizes general objective probability as alternatives. The expected utility theory vNM invented was a key to refine the concept of subjective expected utility theory.

Savage (1954) is the work that completed the concept of subjective probability by integrating the theories of de Finetti, Ramsey, and von Neumann and Morgenstein. He assumes there is an objective state space, and an alternative is a bet on which state will realize, which is called an act. Assuming preferences for acts are observable, he axiomatically characterized subjective expected utility representation of the preference relation. The elicitation method is as follows. The first step is constructing qualitative probability from the observed choice. Suppose some outcome is better than another. Then, using these, the analyst can reveal which event decision-maker (DM) believes more probable. Next, imposing some regularity assumptions on observed preferences, he guarantees the revealed qualitative probability satisfies the axioms of de Finetti. Then, using the elicited subjective probability, he transforms each act into objective probability over the outcomes. Finally, he applies the representation theorem of vNM and completes the identification. In this way, Savage answered the question: what is subjective uncertainty? If DM follows Savage's axioms, then his choice behavior is consistent with subjective expected utility representation, and we call the elicited probability as subjective probability, whether or not DM actually believes this probability is the correct law. This is the standard concept of subjective probability.

Now, there are many extensions and variants of Savage's theory. Here I cite two important studies. One is Machina and Schmeidler (1992) (MS). After the work of vNM, experimental studies showed that subjects do not exactly follow the expected utility theory. For example, Allais (1953) directly falsified independence axiom. The falsification of EU was followed by many works that aim to replace EU with more general choice behavior. See Machina for this literature. Stimulated by non-expected utility theories, MS pointed out that even when DM has a subjective probability, alternatives need not be evaluated by calculating expected values of utility. The model MS consider incorporates a subjective probability. DM first translates a Savagean act into a probability over outcomes with it. Then, he values the probability in possibly non-EU way.

The second is Anscombe and Aumann (1963) (AA). Their contribution is rather technical. The elicitation of Savage was difficult because he assumed no structural assumption, topological or algebraic, on state space and outcome space. AA added linear structure to the space of alternatives. Specifically, they changed what each alternative give in each state from abstract outcomes into lotteries. As vNM did, we can mix lotteries. Extending this mixture operation to Anscombe-Aumann acts makes it possible to elicit subjective probability more easily. Since the domain includes constant act, or objective lotteries, applying EU theorem, one can identify an expected utility u. Then, transform each AA-act f into the utility act $u \circ f$. The mixture operation of acts naturally translates into the usual definition of convex mixture. Upper countor sets of preference in this domain are half-space, and the linear function corresponding to the half-space is the subjective probability. By this geometric argument, it is relatively easy to identify the subjective probability in this domain. Decision theory of uncertainty after AA would usually adopt the AA framework except when the use of Savage's domain is inevitable for conceptual reasons.

1.1.2 Ambiguity

Up to now, I explained the history of how the concept of subjective probability has grown. Meanwhile, it is pointed out that uncertainty need not to be represented by probability. This point is made by Knight (2012) and Ellsberg (1961). Especially, Ellsberg constructs a thought experiment that sharply shows one's choice behavior violates Savage's sure-thing principle. Suppose there is an urn that contains ninety balls, which are colored red, black, and yellow. You know there are thirty red balls in the urn, but you do not know the number of black or yellow. In this setting, Ellsberg constructs four Savagean acts that give money if the picked ball is some specific color. This type of uncertainty, which can not be described with additive probability, is called *ambiguity*.

This insight stimulated decision theoretic literature, though it took long to model ambiguity. Schmeidler (1989) is the first to model ambiguity. His idea is that even if the choice is inconsistent with *additive* probability, it can be consistent with *non-additive* probability. Non-additive probability, or *capacity* was invented by Choquet, who defined integration with such an object. This type of integration is called Choquet integral. Schmeidler first characterized Choquet integral in terms of property as a functional (Schmeidler (1986)). Then, he weakened independence axiom of Anscomb-Aumann model and represented the preference relation with a utility function of Choquet expected utility (CEU) form.

Similarly, weakening the axioms of Anscombe and Aumann, Gilboa and Schmeidler (1989) characterized *maximin expected utility model* (MEU). MEU is a generalized form of SEU, which incorporates ambiguity aversion by modeling DM's belief as a set of probability, called *multiple prior*. DM evaluates an act by calculating SEU with each probability in multiple prior and then take the minimum value as the value of the act. MEU is a special case of CEU in which capacity is convex. This maxmin expected utility model is the standard ambiguity model. It was applied in many economic fields, and many extensions and generalizations were studied in decision theoretic literature. For example, *variational preference* of Maccheroni et al. (2006) is a generalization of MEU. While variational preference includes many important classes of preference, for example, *multiplier preference* of Strzalecki (2011), their technical contribution is also large. The variational technique they brought to decision theory is repeatedly used since then.

There is another direction of modeling ambiguity, called *smooth ambiguity model* (SAM), which describes possible differentiable indifference curves, in contrast, variational preference whose indifference curves are kinked. SAM describes ambiguity with second-order belief, that is, probability over probabilities. For example, if second-order belief assigns a probability of $1/2$ to probabilities π and π', then we interpret that as believing π and π' are the true law of uncertainty with equal confidence. SAM evaluates act with a two-step procedure. First, it calculates SEU of each act; second, average the utilities after transforming with a function that represents ambiguity attitude. SAM has two axiomatizations up to now. Klibanoff et al. (2005) is the first to axiomatize SAM. But their primitive is not preferences observable from the view of analyst. They assume preferences for second-order acts are also a part of data. See Epstein (2010) for more on this point. Seo (2009) is another study that axiomatizes SAM. He utilized the original primitive of Anscombe-Aumann that has objective randomization before the resolution of subjective uncertainty, which we call here random act. This makes the space of primitive convex, which makes the analysis tractable. Using this domain, he could use only preferences for random acts. In a nutshell, Seo's contribution is the treatment of SAM as a positive theory.

Various ambiguity model are different conceptualization of ambiguity. But what is ambiguity at all? Epstein and Zhang (2001) tackled this problem and answered by providing the concept of unambiguous events. An unambiguous event is an event that DM is assigning probability in terms of models. They defined ambiguity in terms of behavior and axiomatized SEU representation *restricted on ambiguous events*. Under their axioms, the behavior is consistent with SEU when comparing acts measurable with respect to unambiguous events, but the model says nothing on other comparisons. Asking what unambiguous events are corresponds to asking what are events not the sources of ambiguity.

1.1.3 Bounded rationality

While rational choice models were developed, psychology literature acutely criticized these theories. I noted the experiments by Allais and Ellsberg that strongly deny the validity of EU and SEU. Not only are these experiments the challenge to rational decision theory. The most influential and stark criticism was made by Kahneman and Tversky (1979). They, through laboratory experiments, provided many pieces of evidence that contradict EU. Moreover, they build a model that captures deviations from EU, even if the modification is somewhat ad hoc. The behavioral economics literature, which follows them, aims to provide better models of boundedly rational choice theory (Camerer et al. (2004)).

Here, I cite Eliaz and Ok (2006) for an illustration of behavioral decision theory. First, one problem of rational choice theory is the assumption that anything can be comparable. Intuitively, sometimes one can not say which alternative is better in shopping centers. This line is persued by Eliaz and Ok (2006). They considered a choice correspondence and elicited two new binary relations to capture the concept of comparability. While both relations are what we usually call *indifference* relation, one represents *same value*, while another represents *incomparability*. Thus behaviorally, these relations are distinguishable. Like this study, decision theoretic analysis enables us to understand abstract ideas of *bounded rationality*.

Recently, a natural formalization of bounded rationality was suggested: *rational inattention* (RIA), which captures the idea that DM does not always think with full force since it is mentally costly. Rational inattention was incorporated by Sims (2003) as a new assumption on a macroeconomic model. Because this is the model with fundamental importance for behavioral economics, many decision theoretic analyses were conducted. This line is first done by Caplin and Dean (2015), who gave revealed preference theory of RIA, which is applicable to finite data. An elegant menu choice foundation is given by de Oliveira et al. (2017). The most important study of RIA for this thesis is Ellis (2018).

1.1.4 Use of non-choice data

As I have written, decision theory asks what we can learn from choice behavior. But now, the decision theoretic community seems to be asking the necessity to restrict our attention to choices. Some researchers already pointed out that the use of non-choice data, including response time, gaze data, and fMRI, possibly facilitate behavioral science (Caplin and Schotter (2008)).

An early study that adopted fMRI data is Caplin and Schotter (2008). They used both fMRI data and choice behavior and conducted revealed preference analysis of the reward prediction error hypothesis. Moreover, they used an fMRI experiment to test it using their theory, resulting in an affirmative answer on whether the theory is valid or not.

More recently, many decision theoretic works study *response time* (RT), that is, time consumed to choose an alternative. Recent trend of RT studies starts from Fudenberg et al. (2018). They constructed a variant of *drift-diffusion model* (DDM), in which DM does not know how much the values of two choices differ and estimated it. After they point on how non-choice data works to predict laboratory behavior, axiomatic studies also started. For example, Fudenberg et al. (2019), using joint distributions of choice and RT, axiomatized DDM, and provided a theory of its estimation and hypothetical testing. See Chapter 2 for more literature review.

1.2 Summary of later chapters

Here I summarize the contents in later chapters.

1.2.1 Response time and revealed information structure

My study in chapter 2 aims to ask whether RT serves to understand the human cognitive process. In the paper, I use a domain that extends that of Ellis, by adding state-conditional response time. From this primitive, I elicit subjective filtration, subjective probability, expected utility, and waiting cost. Especially, subjective filtration is important. The human cognitive process can be understood from the view of information processing, which is believed in today's computational neuroscience community. A simple and tractable representation is filtration over state space. I partially identify subjective filtration using data of RT and choice.

1.2.2 Axiomatization of optimal inattention model with infinite state space

Ellis (2018) used state-conditional choice correspondence and rationalized it with RIA model. He assumed finite state space, while infinite state space is potentially important. For example, describing uncertainty in infinite horizon, normal distribution, and infinite economy requires infinite state space. In order to give an axiomatic foundation of RIA in such an environment, I generalized Ellis's work in Chapter 3. While my work is not conceptually new, more abstraction was required to complete the proof because of a measurability issue. And this served as a pre-analysis for the study in Chapter 2.

1.2.3 Second-order beliefs and unambiguous events

The theory of Epstein and Zhang (2001) is silent on the choices of non-measurable acts. While this makes the theory general analysis of ambiguity, sometimes it is useful to model unambiguous events in terms of model, since models are intuitive sketches of mental concept. With this motivation, in Chapter 4, I characterized a special case of SOSEU model, which is the name of SAM model of Seo (2009), that it assigns some fixed probability to exogenously given events. This result makes what the seemingly intuitive modeling of unambiguous events means clear and let us examine its behavioral implication. A problem yet to be solved is an endogenous elicitation of unambiguous events. The domain of Seo does not span the full space of utility acts. This makes it impossible to uniquely identify the belief and the problem difficult.

Chapter 2

Response time and revealed information structure

2.1 Introduction

In many economic environments, choice timing is not exogenously fixed, but rather itself is a choice variable of the decision-maker (DM). The amount of time consumed to choose, or *response time* (RT), reflects the decision process of DM. This is why behavioral scientists, including recent experimental economists, measure and analyze response time. In this chapter, we theoretically show that response time data, together with choice data, help us understand the decision process.

Assuming the available behavioral data are choice and response time conditional on the state of the world, we axiomatically characterize a model that incorporates endogenously determined learning process. The model we consider describes a DM who decides when and what to choose according to his private learning process. The learning process is described by an information filtration that is a collection of information partitions that evolves over time. We call this as *subjective filtration*. Subjective filtration is not directly observable from the analyst's view while it is fixed from the point of DM. Nevertheless, our axioms enable us identify the DM's subjective filtration. And then, we also elicit utility, subjective probability, and waiting cost.

Our approach to identifying subjective filtration is as follows. If DM would learn a realization of some event at a point in time, he uses this information if that is profitable. Thus any information he has would be reflected in choices. Therefore we define subjective filtration as the smallest one that is necessary to describe his behavior.

The model this chapter considers is a dynamic extension of *optimal inattention representation* studied by Ellis (2018) (henceforth, Ellis), to whom we owe many proof techniques. Ellis axiomatically characterizes a model of rational inattention, assuming the observable data are choices conditional on the states. In his model, DM chooses an alternative after one-shot information acquisition. In contrast to ours, the learning process is described by an information partition that is selected by the DM given a choice situation. On the other hand, in our model, DM selects stopping time given a choice situation, holding the information filtration fixed. The models of Ellis and me can not be distinguished solely in terms of choice. Note, however, Ellis's model does not predict response time.

Literature

This study lies at the intersection of some study areas, which we overview here. First, I review the non-axiomatic literature of RT and indicate how today's behavioral scientists exploit RT data. Second, I turn to decision theoretic literature on RT, and I highlight differences between existing studies and this one. Third, I review decision theoretic literature on dynamic information acquisition. The main contribution of this chapter is the revelation of subjective filtration. Some existing decision theoretic studies also considered elicitation of subjective filtration. I review them to highlight what is added by the use of RT. Fourth, I review rational inattention literature. This study is a dynamic extension of Ellis, one of rational inattention studies. So I compare Ellis and others to explain why I use his primitive.

Response time

Choice has been the most important data economic researchers collect and use. Based on the revealed preference approach, they have studied many mental concepts by providing their operational definition. However, recently they are rapidly acknowledging RT, being affected by cognitive psychology literature.

In this literature, the most influential paradigm of joint analyses of RT and choice is *drift-diffusion model* (DDM). It assumes that DM sequentially acquires information on which choice is better, and he maximizes expected utility given the benefit and cost of waiting for information. Its standard form is given by Edwards (1965) and Ratcliff (1978). Their model was extended into various directions. Among them, Fudenberg et al. (2018) consider a new and natural assumption that DM does not know how much the values of alternatives differ, characterize the optimal policy, and estimate it. Estimation and hypothetical testing of DDM are inevitable to analyze laboratory data. Fudenberg et al. (2019) is a study of this direction. See Spiliopoulos and Ortmann (2018) for more information on RT literature.

Response time study is done with a consensus that "[response time]'s usefulness lies primarily in revealing additional information about a decision maker's underlying cognitive processes or preferences [...] and the effects of deliberation costs on behavior " (Spiliopoulos and Ortmann (2018), p.2). The purpose of this chapter is to theoretically ask how much this research program is possible.

Decision theoretic approaches to RT

As the importance of RT study grows, some recent decision-theoretic papers study what we can learn about the cognitive process using RT. Among others, Duraj and Lin (2019) is the closest to this study.[1] Their data is joint distributions of chosen alternative and RT given various menus in discrete time horizon. They behaviorally characterize DM who solves optimal stopping problem with constant waiting cost or geometric discounting, given a filtration over state space. That is, they and I adopted different formalizations of waiting cost.[2] While the idea resembles this study, there are differences in the parameter to be identified. While I partially identify subjective filtration, they assume that filtration is objective, or analyst can directly observe DM's filtration. Even in laboratories, it is difficult to directly observe filtration, which is interpreted as cognitive process.[3] This difference is caused by the richness of the primitives. While they assume analysts can not observe true state, I assume she can. In a nutshell, their and my studies are not nested, and are complementary.

[1] I briefly review other decision theoretic studies. Echenique and Saito (2017) characterized a reduced-form model in which response time is determined by the difference of choice values. Koida (2017) studies a sequence of incomplete preference relations that become more comparable over time, which is caused by the contraction of Bewley type beliefs. In his model, RT is determined by when two alternatives become comparable. Baldassi et al. (2018) characterize DDM and its multi-alternative extension, and present an algorithm that describes the formation of the consideration set.

[2] Conceptually, their constant waiting cost model is a special case of the discrete-time version of my model. Meanwhile, their discounting model differs not only conceptually, but also in terms of behavior, as they show.

[3] One may think fMRI technology lets us observe the cognitive process. However, the analyses of fMRI data are based on experimental design, observed behavior, and a priori assumption on the data generating process. Thus we can not take fMRI as direct observation of the cognitive process.

Decision theoretic approach to dynamic information acquisition

The aim of this chapter is to reveal how DM's uncertainty resolves over time. Several existing studies analyzed this problem.

Takeoka (2007) adopted menu of menu of acts as alternative and characterized a model with two-stage uncertainty resolution. Because the data he uses is a preference relation over menus, the revealed filtration is interpreted as DM's expectation on how he learns after his facing menus *before* they are given.[4] In contrast, my primitive is behavior after facing menu, and thus my revealed filtration is interpreted as DM's expectation *after* his facing menus. Moreover, the length of information acquisition is an endogenous variable in my model. Note that, however, his primitive can describe decision on whether or not to stop learning by using an alternative that contains a singleton menu, while there are menus not of this type.

de Oliveira and Lamba (2019) ask a problem on judging whether DM's action sequence can be rationalized by some information flow. Their and my study are similar, while the length of action sequences is exogenously fixed in their model, and thus their model does not predict choice timing.

Dillenberger et al. (2018) study an infinite horizon decision model in which the state evolves following a Markov process, and DM acquires information by choosing an information partition. In their model, choices of information partition are constrained, where constraint evolves depending on history. Their information constraint is so general that it unifies many such objects considered in the literature. They show an identification result and representation theorem.

Note that all of the three studies above assume discrete time horizon, whereas my model assumes continuous time horizon.

Rational Inattention

Economic agents often feels information acquisition is costly, perhaps for limitation of cognitive ability, and so he may avoid acquiring all information even if that is materially costless. This insight is called *rational inattention*, which is introduced to economics by Sims (2003). The model I propose here is a model of dynamic information acquisition, and it is strongly related to rational inattention literature.

Rational inattention, which is a tractable formalization of bounded rationality, is now studied through the lens of decision theory. While such studies use differ-

[4] Menu choice studies share implicit assumptions of this type.

ent primitives, the intuition of models are almost the same: DM chooses costly information, updates his belief, and chooses an alternative. Most studies assume DM has an additive information cost function, and some assume that information acquisition discounts the gain.

Caplin and Dean (2015) use choice probability conditional on the state as data. Their study differs from other ones in that it is applicable to finite data and so, to experimental data, in principle. Chambers et al. (2020) studies a generalization and a discount counterpart of Caplin and Dean (2015). de Oliveira et al. (2017) use preference relation over menus of acts. They applied variational technique to elicit parameters. Their study is one of the earliest axiomatic studies of inattention.

I already explained the model of Ellis (2018). Seemingly, his primitive resembles Caplin and Dean (2015) and proof technique is similar to de Oliveira et al. (2017). However, his contribution is large. He invents a new concept named *plan*, which describes what he chooses in each state. As explained, his primitive itself is a state-conditional choice correspondence. He constructs preference relation over plans from the primitive and applied variational technique to preferences for plans. Potentially, this proof method can be used for future research, as revealed preference analysis of standard choice correspondence bore plentiful field.

In section 2.2, I introduce the analytical framework and optimal stopping representation. In section 2.3.4, I present the representation theorem and comment on its interpretation and future research.

2.2 Framework and model

2.2.1 Framework

This subsection introduces the framework. Each metric space S introduced below is endowed with its Borel σ-algebra $\mathcal{B}(S)$. Let Ω be a finite set, which is interpreted as the set of states that describe uncertainty. Let X be a convex subset of a metrizable topological vector space and let d be its compatible metric. Let $\mathcal{A}$ be the set of functions from Ω to X. Each element of $\mathcal{A}$, interpreted as an alternative, is called an *act*. With a natural isomorphism, we regard X is the set of constant acts. The set $\mathcal{A}$ is endowed with the uniform metric $d_\infty(f, g) = \max_{\omega\in\Omega} d(f(\omega), g(\omega))$. Let $\mathcal{K}$ be the set of all non-empty compact sets of $\mathcal{A}$ that is endowed with the Hausdorff metric d_h. Each element of $\mathcal{K}$ is interpreted as a set of alternatives. For typical elements of the sets above, we write $x, y, z \in X$, $f, g, h \in \mathcal{A}$, and $A, B, C \in \mathcal{K}$. For any σ-algebra $\mathcal{F}$ over Ω and $\omega \in \Omega$, $\mathcal{F}(\omega)$ denotes for the smallest element of $\mathcal{F}$ that contains ω.

Our data set is a conditional choice correspondence and a conditional response time. Conditional choice correspondence is a function $c : \mathcal{K} \times \Omega \to \mathcal{K}$ that satisfies $c(B, \omega) \subset B$ for any $(B, \omega) \in \mathcal{K} \times \Omega$. Conditional response time is a function $\tau : \mathcal{K} \times \Omega \to \mathbb{R}_+$. For each $B \in \mathcal{K}$, define $c_B : \Omega \to \mathcal{K}$ and τ_B by

$$c_B(\omega) = c(B, \omega) \text{ and } \tau_B(\omega) = \tau(B, \omega).$$

2.2.2 Model

Here we explain the model we analyze. Given a menu, DM first chooses response time τ. Stopping at time $\tau(\omega)$, DM learns a realization of an event. Then, he updates his prior belief following Bayes rule and chooses an alternative $f \in B$. Finally, an outcome $f(\omega)$ is given, and the waiting cost realizes. We introduce notation and a term to formalize the ideas above. For a stopping time $\tilde{\tau} : \Omega \to \mathbb{R}_+$ and $t \in \mathbb{R}_+$, let $\{\tilde{\tau} \leq t\} = \{\omega \in \Omega | \tilde{\tau}(\omega) \leq t\}$. This is the event that he stops before t according to τ. We say a stopping time $\tilde{\tau}$ is adapted to a filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ if $\{\tilde{\tau} \leq t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}_+$. In this case, we say τ is $\mathbb{F}$-adapted. When this is the case, DM who follows τ can decide whether to stop using the information represented by $\{\mathcal{F}_t\}$. Stopping times are ordered in point-wise manner: $\tau \geq \sigma$ if $\tau(\omega) \geq \sigma(\omega)$ for all ω. Now we introduce the model.

Definition 1. *We say a quadruple $(u, \pi, \mathbb{F}, \gamma)$ is an optimal stopping representation of (c, τ) if these parameters satisfy the following conditions:*

$$\tau_B \in \arg\max_{\tau \in \mathcal{T}} \mathrm{E}[\max_{f \in B} \mathrm{E}[u(f)|\mathcal{F}_\tau]] - \gamma(\tau)$$
$$c_B(\omega) = \arg\max_{f \in B} \mathrm{E}[u(f)|\mathcal{F}_{\tau_B}](\omega) \textit{ for all } \omega \in \Omega.$$

Here,

- $u : X \to \mathbb{R}$ *is an affine function with* $u(X) = \mathbb{R}$,
- π *is a full-support probability over* Ω *with which expectations are taken,*
- $\mathbb{F} = \{\mathcal{F}_t\}_{t \in \mathbb{R}}$ *is a filtration over* Ω,
- $\mathcal{T}$ *is the set of* $\mathbb{F}$*-adapted stopping times,*

- $\mathcal{F}_\tau = \{\Delta \in 2^\Omega | \forall t \in \mathbb{R}_+ \ \Delta \cap \{\tau \le t\} \in \mathcal{F}_t\}$,
- $\gamma : \mathcal{T} \to \overline{\mathbb{R}}$ *is a cost function such that* $\gamma(\tau) \le \gamma(\sigma)$ *whenever* $\tau \le \sigma$.

The first line of the representation requires that given a menu B, the observed response time τ_B maximizes DM's ex-ante expected net utility assuming his ex-post choices are optimal. The second line requires that he chooses alternatives that maximize conditional expected utility.

2.3 Main result

The axioms we impose to (c, τ) can be classified into three groups. The first group consists of axioms of optimal inattention. These axioms first appeared in Ellis (2018). They guarantee the existence of fundamental preference relation behind the choice correspondence and impose structural assumption on it. The second group consists of axioms of optimal stopping. They are new axioms that describe consistent relationships between choice and response time. The third group consists of technical axioms. In all axioms, variables with no quantifier are understood as bounded by a universal quantifier.

2.3.1 Axioms of optimal inattention

The first axiom Independence of Nonrerevant Alternative is a variant of the Weak Axiom of Revealed Preference that is adapted to the conditional choice correspondence.

Axiom 1 (*INRA: Independence of Never Relevant Acts*)**.** *If* $A \subset B$ *and* $A \cap c(B, \omega) \neq \emptyset$ *for any* $\omega \in \Omega$*, then* $c(A, \omega) = A \cap c(B, \omega)$ *for all* $\omega \in \Omega$.

For $f, g \in \mathcal{A}$ and $\alpha \in [0, 1]$, let $\alpha f + (1 - \alpha)g$ be an act such that $[\alpha f + (1 - \alpha)g](\omega) = \alpha f(\omega) + (1 - \alpha)g(\omega)$. For $f \in \mathcal{A}$ and $B \in \mathcal{K}$, let $\alpha f + (1 - \alpha)B = \{\alpha f + (1 - \alpha)g | g \in B\}$.

Attention Constrained Independence is a weaker version of indepedence axiom, which is an implication of additive information cost.

Axiom 2 (*ACI: Attention Constraiend Independence*)**.** *If* $\alpha g + (1-\alpha)f \in c(\alpha g + (1 - \alpha)B, \omega)$*, then* $\alpha h + (1 - \alpha)f \in c(\alpha h + (1 - \alpha)B, \omega)$

We define a preference relation over outcomes. For $x, y \in X$, define

$$x \succeq^R y \Leftrightarrow \text{there exists an } \omega \in \Omega \text{ such that } x \in c(\{x, y\}, \omega).$$

Let $\succ^R$ and $\sim^R$ be the asymmetric part and symmetric part of $\succeq^R$, respectively.

Monotonicity states that if an act chosen from a menu at some state and it is state-wise dominated by another act, then the latter one must also be chosen at the state.

Axiom 3 (*M: Monotonicity*)**.** *For $f, g \in B$, if*

$$f(\omega) \succeq^R g(\omega) \textit{ for all } \omega \in \Omega,$$

then

$$g \in c(B, \omega) \Rightarrow f \in c(B, \omega).$$

2.3.2 Axiom of optimal stopping

From the data (c, τ), we elicit the filtration $\{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$. Introduce a symbol $\diamond$, which is not an element of $\mathcal{K}$. Then, endow $\mathcal{K} \cup \{\diamond\}$ with the σ-algebra $\mathcal{B}(\mathcal{K}) \cup \{\{\diamond\}\}$, which contains all the Borel sets of $\mathcal{K}$ and set $\{\diamond\}$. Next, for each menu B and $t \in \mathbb{R}_+$, define a function $c_B^t : \Omega \to \mathcal{K} \cup \{\diamond\}$ by

$$c_B^t(\omega) = \begin{cases} c(B, \omega) & \text{if } \tau(B, \omega) \leq t \\ \diamond & \text{otherwise.} \end{cases} \tag{2.1}$$

This function captures the information necessary to implement the choices c_B. Now we dfine subjective filtration over Ω.

Definition 2. *Subjective filtration is the indexed collection $\mathbb{F} = \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ of σ-algebras over Ω given by*

$$\mathcal{F}_t = \sigma(c_B^t; B \in \mathcal{K}). \tag{2.2}$$

That is, $\mathcal{F}_t$ the smallest σ-algebra that makes all of the functions $\{c_B^t\}_{B \in \mathcal{K}}$ measurable. So $\mathcal{F}_t$ is the minimal information to explain the choice behavior by time point t. It is easily shown that τ_B is $\mathbb{F}$-adapted for any $B \in \mathcal{K}$.

The next axiom *Dynamic Subjective Consequentialism* requires that DM respects the revealed filtration.

Axiom 4 (*DSC: Dynamic Subjective Consequentialism*)**.** *For $f, g \in B$, $\omega \in \Omega$, and $\Delta \in \mathcal{F}_{\tau_B}$ such that $\omega \in \Delta$, if*

$$f(\omega') = g(\omega') \textit{ for all } \omega' \in \Delta,$$

then

$$f \in c(B, \omega) \Leftrightarrow g \in c(B, \omega).$$

Given B, DM acquires information $\mathcal{F}_{\tau_B}$ using τ_B and thus he knows the realization of $\Delta \in \mathcal{F}_{\tau_B}$ that containes the true state ω. If $f, g \in B$ agrees on Δ, he treats them as if they are the same act.

The next axiom, *Information Monotinicity*, states that lenghs of response times are consistent with the fineness of choices.

Axiom 5 (*IM: Information Monotonicity*)**.** *If*

$$c(A, \omega) \neq c(A, \omega') \Rightarrow c(B, \omega) \neq c(B, \omega') \textit{ for any } \omega, \omega' \in \Omega,$$

then,

$$\tau_A \leq \tau_B.$$

We interpret the inequality $c(A, \omega) \neq c(A, \omega')$ as an implication of acquiring inoformation enough to distinguish ω and ω' when facing A. Thus the relationship between the choice behavior when facing different menus assumed in this axiom means that DM acquires more information when facing B. The axiom requires that DM waits for more when facing B than when facing A. In most realistic situations, waiting more gives better information and costs more. Thus more complex choice behavior is only possible by waiting more.

The next axiom, *Time Invariance*, states that response times given a mixture of an act and a menu is the same even if the mixed act is changed, if the mixing weight and the menu is left the same.

Axiom 6 (*TI: Time Invariance*)**.**

$$\tau_{\alpha f + (1-\alpha)B} = \tau_{\alpha g + (1-\alpha)B}.$$

2.3.3 Technical axioms

Define a relation $\succeq^D$ over menus by

$$A \succeq^D B \Leftrightarrow \text{For all } \omega \in \Omega,\ A \cap c(B, \omega) \neq \emptyset.$$

The relation $A \succeq^D B$ means that, given A, DM can emulate the optimal policy he uses if B is given. Thus if he can choose one from A or B, he prefers A. In this sense, this is a directly revealed preference ranking for menus. Let $\succeq^I$ denotes the transitive closure of $\succeq^D$. That is,

$$A \succeq^I B \Leftrightarrow \text{There exists } n \in \mathbb{N} \text{ and } B_1, \ldots B_n \text{ such that } A \succeq^D B_1 \succeq^D \cdots \succeq^D B_n \succeq^D B.$$

To interpret the relation, remind that we interpret $\succeq^D$ as the preference ranking for menus. Assuming he is rational in the sense of having transitive raking, we can extend the ranking. The relation $\succeq^I$ is the indirectly revealed ranking obtained in this way.

The next is axiom *Continuity.* For $x, y \in X$ and $\omega \in \Omega$, let $x\omega y$ denote the act that gives x in ω and y in other states. We write $g \succeq^I f$ if $\{g\} \succeq^I \{f\}$ for simplicity.

Axiom 7 (*C: Continuity*)**.**

1. *For any $\omega \in \Omega$, $\{B_n\}_{n=1}^{\infty}$ and $\{f_n\}_{n=1}^{\infty}$ such that $B_n \to B$ and $f_n \to f$ with $f_n \in c(B_n, \omega)$, if*

$$\tau_{B_n} = \tau_B \text{ for any } n \in \mathbb{N},$$

then

$$f \in c(B, \omega).$$

2. *For any $x, y \in X$ such that $x \succ^R y$, $\omega \in \Omega$, and sequences $f_n \to x\omega y$ and $g_n \to y$, there exists $n \in \mathbb{N}$ such that $g_n \not\succeq^I f_n$.*

The first part is a weak form of upper hemicontinuity of c. The second part states is an implication of the assumption that there is no null state. Suppose a state ω can be the true state. Then, $x\omega y$ is strictly preferred over y since it gives a better outcome at ω. Thus, if f_n and g_n are sufficiently close to $x\omega y$ and y respectively, g_n is not prefered over f_n. Together with other axioms, it guarantees the existence of a full-support subjective probability.[5]

[5] The axioms of Ellis admit a conditional choice correspondence that generates the trivial $\succeq^I$ relation. His argument on eliciting subjective probability has a gap since it requires the non-triviality of $\succeq^I$, which is not guaranteed.

The next axiom is *Unboundedness*. This is necessary to calibrate the cost function.

Axiom 8 (*U: Unboundedness*)**.** *There exist* $x, y \in X$ *such that* $x \succ^R y$ *and, for any* $\beta \in (0, 1)$*, there exist* $z^*, z_* \in X$ *such that*

$$\beta z^* + (1 - \beta)y \succ^R x, \; y \succ^R \beta z_* + (1 - \beta)y.$$

2.3.4 Representation theorem

We present our representation result.

Theorem 1. *If* c *and* τ *satisfies INRA, ACI, M, DSC, IM, TI, C, and U, then there exist* $u, \pi,$ *and* γ *such that* $(u, \pi, \mathbb{F}, \gamma)$ *is an optimal stopping representation of* (c, τ)*.*

This theorem shows the intuitive axioms we considered above are sufficient for the data to be explained by optimal stopping model. I asked what we can learn using both RT and choice. To answer this question, I analyzed a model that jointly predict RT and choice. It incorporates dynamic information flow, represented as subjective filtration. Subjective filtration is a rough representation of the cognitive process that describes how uncertainty resolves from the view of DM. I showed that RT and choice together partially identify subjective filtration. This result implies that using RT somewhat helps us understand the human cognitive process.

2.4 Concluding remarks

Point identification of filtration is a problem yet to be solved.[6] It is difficult to identify filtration using the current primitive because of the incompleteness of $\succeq^I$. Using a triplet $(\succeq, c, \tau)$ of preference relation over menus, conditional correspondence, and conditional response time may serve to point identification.

The followings are my current conjecture. Using the menu preference, one can calibrate any menu with constant acts. Moreover, this may serve to add structural assumptions on γ. The cost function of this study is too general to identify filtration. An important special case of cost function is expectated cost form:

$$\gamma(\tau) = \int_\Omega \tilde{\gamma}(\tau(\omega)) \, \mathrm{d}\pi,$$

[6]As far as I know, there is no such identification result with the assumption of endogenous choice timing.

where $\tilde{\gamma}(\cdot)$ is state-independenct waiting cost, and π is a subjective probability. Assuming this cost function form may let us identify the filtration uniquely.

2.5 Proof

2.5.1 Roadmap

Here I give a roadmap of the proof of Theorem 1. First, from Lemma 6 to Lemma 5, I replace the acts with utility acts to simplify the rest of the proof. Lemma 6 shows that choice from a menu is invariant under its translation in the new domain. Lemma 7 shows that the chosen stopping time given any menu is sufficient to implement the choice behavioir then.

Next, we construct a preference relation for *plan* is derived. Here, a plan is DM's action plan that describes when he stops and what he chooses in each state. This is done via a preference relation over menus. A menu corresponds to a plan if given the menu, DM implements the plan. Lemma 9, which is proved using Lemma 8, states that, for each pair of choice plan and stopping time, there is a menu given which both of them are implemented. This lemma guarantees that preferences for plans is well-defined. Then, to elicit the parameters, we show regularities the preferences for menus and that for plans. Lemma 6 shows that the relation $\succeq^I$ is translation invariant. Lemma 11 shows that if a plan is implemented given a menu, then it is the best one that DM can choose.

Then, we turn to the elicitation of parameters. Lemma 14 and Lemma 15 elicits a subjective probability and cost function, respectively. Lemma 18 and Lemma 19 completes the proof.

2.5.2 Basic properties of choice correspondence

In this subsection, we investigate the basic properties of the choice correspondence. First, we construct an expected utility function. Let $\mathcal{K}(X)$ be the set of all non-empty compact subsets of X.

Lemma 1. *There exists a continuous affine function $u : X \to \mathbb{R}$ such that, for any $B \in \mathcal{K}(X)$,*

$$x \in c(B, \omega) \Leftrightarrow u(x) \geq u(y) \; \forall y \in B.$$

and $u(X) = \mathbb{R}$.

Proof. We first show that $\succeq^R$ is continuous, that is,

$$\{y \in X | y \succeq^R x\} \text{ and } \{y \in X | x \succeq^R y\},$$

are closed. First, note that Axiom M implies that, for any $x, y \in X$ and $\omega, \omega' \in \Omega$, $c(\{x,y\},\omega) = c(\{x,y\},\omega')$. Suppose $y_n \to y$, $y_n \succeq^R x$ and take any $\omega \in \Omega$. Since $c(\{y,x\},\cdot)$ and each $c(\{y_n,x\},\cdot)$ is constant, *IM* implies $\tau_{\{y_n,x\}} = \tau_{\{y,x\}}$ for all n. In addition, $\{y_n, x\} \to \{y, x\}$ in $\mathcal{K}$. Then, Axiom C implies $y \in c(\{y,x\},\omega)$, or $y \succeq^R x$.

Next we show that $\succeq^R$ is transitive. Suppose $x \succeq^R y$ and $y \succeq^R z$, and take any $\omega \in \Omega$. Note that $c(\{x,y,z\},\omega)$ is nonempty. If $z \in c(\{x,y,z\},\omega)$, then $y \succeq^R z$ and axiom M imply $y \in c(\{x,y,z\},\omega)$. Likewise, if $y \in c(\{x,y,z\},\omega)$, then $x \in c(\{x,y,z\},\omega)$. In conclusion, $x \in c(\{x,y,z\},\omega)$ holds for any ω. Then, by *INRA*, $x \in c(\{x,z\},\omega)$ holds for any ω. Thus, $x \succeq^R z$.

Note that, since $\succeq^R$ is complete, transitive, and continuous relation, $\max_{\succeq^R} B = \{x \in B | \, x \succeq^R y \text{ for all } y \in B\}$ is nonempty. We show that $\max_{\succeq^R} B = c(B,\omega)$ holds for any ω. First, suppose $y \in \max_{\succeq^R} B$ and take $x \in c(B,\omega)$. Then, $y \succeq^R x$ and axiom M imply $y \in c(B,\omega)$. Thus $\max_{\succeq^R} B \subset c(B,\omega)$ Next, suppose $y \notin \max_{\succeq^R} B$ and take $x \in \max_{\succeq^R} B$. Then, $x \succ^R y$ and $x \in \max_{\succeq^R} B$ by the first inclusion. By *INRA*, $c(\{x,y\},\omega) = \{x,y\} \cap c(B,\omega)$ holds for all ω. Then, if $y \in c(B,\omega)$, then $y \in c(\{x,y\},\omega)$ and so $y \succeq^R x$. But this contradicts the way x and y are taken. Thus $y \notin c(B,\omega)$. We showed $c(B,\omega) \subset \max_{\succeq^R} B$.

As Ellis showed in the proof of his Lemma 1, *ACI* implies Independence of $\succeq^R$. Then, applying expected utility theorem, construct a utility representation u of $\succeq^R$ that is affine and continuous. By axiom U, u is unbonded. $\square$

The next lemma states that choice behavior follows two regularities: First, adding acts that are dominated by existing ones does not change what to be chosen; secondly, the choice behavior respects the acquired information. For $B \in \mathcal{K}$ and $\omega \in \Omega$, let $\mathrm{P}(B)(\omega) = \{\omega' \in \mathcal{K} | \; c(B,\omega') = c(B,\omega)\}$, the partition generated by c_B. Equivalently, we can see $\mathrm{P}(B)$ as the σ-algebra $\{c_B^{-1}(V) | V \in \mathcal{O}(\mathcal{K})\}$ over Ω generated by $c(B,\cdot)$. For $f \in \mathcal{A}$ and u, let $u \circ f$ denote for their composite function.

Lemma 2.

1. *Assume that, for any $g \in B$, there exists $f \in A$ such that $u \circ f \geq u \circ g$. Then*

$$c(A,\omega) = A \cap c(A \cup B, \omega).$$

2. *For $f, g \in B$, $\omega \in \Omega$, and $\Delta \in \mathcal{F}_{\tau_B}$ with $\omega \in \Delta$, if $f \in c(B,\omega)$ and $u(g(\omega')) = u(f(\omega'))$ for all $\omega' \in \Delta$, then $g \in c(B,\omega)$.*

Proof. Suppose that $g \in B \cap c(A \cup B, \omega)$. By the assumption, there is some $f \in A$ such that $u \circ f \geq u \circ g$. This and Axiom M implies $f \in c(A \cup B, \omega)$. We have shown $A \cap c(A \cup B, \omega) \neq \emptyset$ for any $\omega \in \Omega$. Applying *INRA* completes the first part. The second part is proved as Lemma 2 in Ellis using *DSC*. □

2.5.3 Transforming acts into utility acts

We collect preliminary results to work on real-valued functions, instead of acts. Let $\mathcal{B}_b = \mathbb{R}^\Omega$ which is endowed with the uniform norm $\|\cdot\|$. Let $\mathcal{K}(\mathcal{B}_b)$ be the set of compact sets of $\mathcal{B}_b$.

First, we construct a set $Y \subset X$ such that $u(Y) = \mathbb{R}$ and the restriction of u to Y is a homeomorphism. For each $n \in \mathbb{Z}$, take $x_n \in X$ such that $u(x_n) = n$. Let $Y_n = \{(1-\alpha)x_n + \alpha x_{n+1} | \alpha \in [0,1]\}$ and $Y = \bigcup_{n\in\mathbb{Z}} Y_n$. Let v be the inverse function of $u|_Y$, which exists because of the definition of Y. For each $n \in \mathbb{Z}$, let $v_n : [n, n+1] \to Y_n$ be the inverse function of $u|_{Y_n}$.

Lemma 3.

1. *For each $n \in \mathbb{Z}$, the function v_n is uniformly continuous.*

2. *The function $u|_Y$ is a homeomorphism.*

Proof. Define a function $\overline{v}_n : \mathbb{R} \to X$ by

$$\overline{v}_n(\beta) = (1 - (\beta - n))x_n + (\beta - n)x_{n+1}.$$

For $\beta \in [n, n+1]$, $\overline{v}_n(\beta) \in Y_n$ holds, and besides the affinity of u implies $u(\overline{v}_n(\beta)) = \beta$. So the restriction of $\overline{v}_n$ to $[n, n+1]$ is v_n. Because the addition and the scalar multiplication is continuous in any topological vector space, so is $\overline{v}_n$. Since it is affine, it is uniformly continuous, and so is v_n. The first part is completed.

Turn to the second part. It is sufficient to show that v is continuous. To this end, take a sequence $\{\hat{x}_k\}_{k=1}^\infty \subset \mathbb{R}$ and suppose $\hat{x}_k \to \hat{x}$ in $\mathbb{R}$. If $\hat{x} \in (n, n+1)$ for some $n \in \mathbb{Z}$, for sufficently large k, $\hat{x}_k \in (n, n+1)$. The first part implies $v(\hat{x}_k) = v_n(\hat{x}_k) \to v_n(\hat{x}) = v(\hat{x})$ as $k \to \infty$. Turn to the case that $\hat{x} = n$ for some n. Then, for sufficiently large k, $\hat{x}_k \in (n-1, n+1)$. Take any $\epsilon > 0$. From the coninuity of v_{n-1} and v_n, there exist some $\delta > 0$ such that if $\hat{x}_k \in [n-1, n]$ and $|\hat{x}_k - \hat{x}| < \delta$, then $d(v_{n-1}(\hat{x}_k), v_{n-1}(x_k)) < \epsilon$; and if $\hat{x}_k \in [n, n+1]$ and $|\hat{x}_k - \hat{x}| < \delta$, then $d(v_n(\hat{x}_k), v_n(\hat{x})) < \epsilon$. So $v(\hat{x}_k) \to v(\hat{x})$ in Y. □

Let $\mathcal{A}_Y = \{f \in \mathcal{A} | \ \forall \omega \in \Omega \ f(\omega) \in Y\}$. In the next lemma, we denote $u \circ f$ for the composite function of u and f. Define $\Phi^* : \mathcal{A} \to \mathcal{B}_b$, $\Phi : \mathcal{A}_Y \to \mathcal{B}_b$, and $\Psi : \mathcal{A} \to \mathcal{A}_Y$ by

$$\Phi^*(f) = u \circ f, \ \Phi(f) = u \circ f, \text{ and } \Psi = \Phi^{-1} \circ \Phi^*.$$

Lemma 4.

1. *Φ^* is continuous.*

2. *Φ is a homeomorphism.*

3. *Ψ continuous and $u \circ \Psi(f) = u \circ f$ for any $f \in \mathcal{A}$.*

Proof. Consider the first part. Let $f_k \to f$ in $\mathcal{A}$. Because u is a continuous affine function, it is uniformly continuous. So, for any $\epsilon > 0$, there is some $\delta > 0$ such that $d(x, y) < \delta$ implies $|u(x) - u(y)| < \epsilon$. Note that, for sufficiently large k, $d_\infty(f_k, f) < \delta$. So, for all $\omega \in \Omega$, $|u(f_k(\omega)) - u(f(\omega))| < \epsilon$. That is, $\|u \circ f_k - u \circ f\| \to 0$. The first part is complete.

Consider the second part. The continuity of Φ follows from that of Φ^*. We shall prove that Φ is a bijection. For $\hat{f} \in \mathcal{B}_b$, $\Phi(v \circ \hat{f}) = u \circ v \circ \hat{f} = \hat{f}$. So it is onto. Take $f, g \in \mathcal{A}_Y$ such that $u \circ f = u \circ g$. Then, $f = v \circ u \circ f = v \circ u \circ g = g$. So it is one-to-one.

Finally, we show that the inverse function $\Phi^{-1} : \mathcal{B}_b \to \mathcal{A}_Y$ is continuous. Note that $\Phi^{-1}(\hat{f}) = v \circ \hat{f}$. Take a sequence $\{\hat{f}_k\}_{k=1}^\infty \in \mathcal{B}_b$ such that $\hat{f}_k \to \hat{f}$ in $\mathcal{B}_b$. Take $n \in \mathbb{N}$ such that $-n < \hat{f} < n$. Fix any $\epsilon > 0$. The uniform continuity of v_j implies that for each $j = -n-1, \dots, n$, there exist $\delta_j > 0$ such that, if $x, y \in [j, j+1]$ and $|x - y| < \delta_j$, then $d(v_j(x), v_j(y)) < \epsilon$. Let $\delta = \min_j \delta_j$. Fix a sufficently large k so that $\|\hat{f}_k - \hat{f}\| < \min\{\delta, 1\}$. Then, for any $\omega \in \Omega$, there exists j such that $\hat{f}_k(\omega), \hat{f}(\omega) \in [j, j+1]$. Hence, for all $\omega \in \Omega$, $d(v(\hat{f}_k(\omega)), v(\hat{f}(\omega))) < \epsilon$. So $d_\infty(\Phi^{-1}(\hat{f}_k), \Phi^{-1}(\hat{f})) \le \epsilon$. □

Next, we show that the choice behavior depends only on the state-dependent utilities of the acts in the choice sets. For a moment, we denote $u \circ f$ as f^u and let $A^u = \{f^u | f \in A\}$ for $A \in \mathcal{K}$.

Lemma 5. *If $A^u = B^u$, then $[c(A, \omega)]^u = [c(B, \omega)]^u$ for each $\omega \in \Omega$.*

Proof. Note that Lemma 2 (1) implies

$$c(A, \omega) = A \cap c(A \cup B, \omega), \ c(B, \omega) = B \cap c(A \cup B, \omega)$$

for each $\omega \in \Omega$. Thus for $f \in c(A, \omega)$, we have $f \in c(A \cup B, \omega)$. Take $g \in B$ such that $f^u = g^u$. Axiom M implies $g \in c(A \cup B, \omega)$. Then $g \in c(B, \omega)$. □

The next lemma states that the choice correspondence is translation invariant. For $f \in \mathcal{A}$ and $\alpha \in \mathbb{R}$, define $\alpha f \in \mathcal{A}$ by $(\alpha f)(\omega) = \alpha f(\omega)$. For each $B \in \mathcal{K}$ and $\alpha \in \mathbb{R}$, let $\alpha B = \{\alpha f | f \in B\}$.

Lemma 6. *The equation* $[c(B+g,\omega)]^u = [c(B,\omega)]^u + g^u$ *holds.*

Proof. Suppose $A^u = 2B^u$, $f^u = 2g^u$, and $x^u = 0$. Then, $B^u = (\frac{1}{2}A + \frac{1}{2}x)^u$ and $B^u + g^u = (\frac{1}{2}A + \frac{1}{2}f)^u$. Then, Lemma 5 implies

$$[c(B+g,\omega)]^u = \left[c\left(\frac{1}{2}A + \frac{1}{2}f, \omega\right)\right]^u. \tag{2.3}$$

On the other hand, ACI implies

$$\frac{1}{2}g + \frac{1}{2}f \in c\left(\frac{1}{2}A + \frac{1}{2}f, \omega\right) \Leftrightarrow \frac{1}{2}g + \frac{1}{2}x \in c\left(\frac{1}{2}A + \frac{1}{2}x, \omega\right).$$

Thus

$$\begin{aligned}\left[c\left(\frac{1}{2}A + \frac{1}{2}f, \omega\right)\right]^u &= \left[c\left(\frac{1}{2}A + \frac{1}{2}x, \omega\right)\right]^u + \frac{1}{2}f^u \\ &= \left[c\left(\frac{1}{2}A + \frac{1}{2}x, \omega\right)\right]^u + g^u = [c(B,\omega)]^u + g^u \end{aligned} \tag{2.4}$$

Combining (2.3) and (2.4) completes the proof. □

Now we define a choice correspondence $\tilde{c} : \mathcal{K}(\mathcal{B}_b) \times \Omega \to \mathcal{K}(\mathcal{B}_b)$ and $\tilde{\tau} : \mathcal{K}(\mathcal{B}_b) \times \Omega \to \mathbb{R}_+$ by

$$\tilde{c}(B,\omega) = \Phi[c(\Phi^{-1}(B),\omega)] \text{ and } \tilde{\tau}(B,\omega) = \tau(\Phi^{-1}(B),\omega).$$

Then, $\tilde{c}$ and $\tilde{\tau}$ inherit all the properties of c and τ. Besides, $\tilde{c}$ is translation invariant:

$$\tilde{c}(B+f,\omega) = \tilde{c}(B,\omega) + f$$

by the virtue of Lemma 6. Note

$$c(B,\omega) = \Phi^{-1}(\tilde{c}(\Phi(B),\omega)) \text{ and } \tau(B,\omega) = \tilde{\tau}(\Phi(B),\omega).$$

for $B \in \mathcal{K}(\mathcal{A}_Y)$. Once we found a representation of $\tilde{c}$ and $\tilde{\tau}$, the obtained parameters work for c and τ. From now on, we write $\mathcal{B}_b$ as $\mathcal{A}$, $\tilde{c}$ as c, and $\tilde{\tau}$ as τ for simplicity.

2.5.4 Preliminary

Any choice behavior is done with enough information to do so. The next lemma states this fact.

Lemma 7. *For any $B \in \mathcal{K}$,* $\mathrm{P}(B) \subset \mathcal{F}_{\tau_B}$.

Proof. For any $V \in \mathcal{O}(\mathcal{K})$ and $t \geq 0$, we have

$$c_B^{-1}(V) \cap \{\tau_B \leq t\} = \{\omega | c_B^t(\omega) \in V\} \in \mathcal{F}_t.$$

Thus, by definition of $\mathcal{F}_{\tau_B}$, $c_B^{-1}(V) \in \mathcal{F}_{\tau_B}$. Remind that $\mathrm{P}(B)$ is the partition generated by c_B and thus $P(B) = \{c_B^{-1}(V) | V \in \mathcal{O}(\mathcal{K})\}$. Therefore, $\mathrm{P}(B) \subset \mathcal{F}_{\tau_B}$. □

A *plan* is a pair (F, τ) of function $F : \Omega \to \mathcal{A}$ and a $\{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$-adapted stopping time τ, where F is $\mathcal{F}_\tau$-mesurable. Let $\mathbb{H}$ denote the set of all plans. Let $\mathcal{T}^* = \{\tau_B | B \in \mathcal{K}\}$ be the set of stopping times DM sometimes uses. Finally, let $\mathbb{H}^* = \{(F, \tau) \in \mathbb{H} | \tau \in \mathcal{T}^*\}$. The set $\mathbb{H}^*$ consists of plans that is implemented via a response time that is actually used when DM faces some menu. We especially pay attention $\mathbb{H}^*$ because this property facilitates the calibration of γ. For $B \in \mathcal{K}$, let $\hat{c}(B)$ be the set of functions $F : \Omega \to B$ that satisfy $F(\omega) \in c(B, \omega)$. Henceforth, for any $f \in \mathcal{A}$, the notation $\{f_\omega\}_{\omega \in \Omega}$ sometimes denotes for the function $\omega \mapsto f_\omega$ and sometimes for the set $\{f_\omega \,|\, \omega \in \Omega\}$. Let $\mathcal{O}(S)$ be the set of all open sets in S.

In the next section, we will construct a preference relation over plans from preference over menus. So, for each plan $(F, \tau) \in \mathbb{H}^*$, we need a menu given which DM implements it. But in general, there may not be a menu B with $F \in \hat{c}(B)$. So, as a substitute, we construct a menu with which the specified response time is implemented, and the same utility level is given at each state.

Lemma 8 and 9 serves this purpose. Lemma 8 says that for any $\tau \in \mathcal{T}^*$, we can construct a menu so that τ is used and utilities obtained at any state is zero.

Lemma 8. *For any $\tau \in \mathcal{T}^*$, there exists $B_\tau \in \mathcal{K}$ and $\{f_\omega\}_{\omega \in \Omega}$ that satisfy the followings*

1. $\tau(B_\tau, \cdot) = \tau$
2. $f_\omega(\omega) = 0$, $f_\omega \in c(B_\tau, \omega)$.
3. *For any* $g \in \mathcal{A}$, $f_\omega + g \in c(B_\tau + g, \omega)$.

Proof. Take any $B \in \mathcal{K}$ such that $\tau_B = \tau$. Take a selector $\{g_\omega\}_{\omega\in\Omega}$ of the correspondence $c(B,\cdot)$. Define a function h as $h(\omega) = g_\omega(\omega)$. Define a menu $B_\tau = B - h$. Note that $c(B_\tau,\omega) = c(B,\omega) - h$ and $\tau_{B_\tau} = \tau_B = \tau$. Define a plan $\{f_\omega\}_{\omega\in\Omega}$ by $f_\omega = g_\omega - h$, and then the following hold: $f_\omega \in B_\tau$, $f_\omega(\omega) = 0$, and $f_\omega \in c(B_\tau,\omega)$. □

Next, for $F \in \mathbb{H}^*$, let $F^* \in \mathcal{A}$ be an act defined by $F^*(\omega) = F(\omega)(\omega)$. For any $(F,\tau) \in \mathbb{H}^*$, if there exists some menu $B \in \mathcal{K}$ and $\overline{F} : \Omega \to \mathcal{A}$ such that

$$\overline{F} \in \hat{c}(B),\ (\overline{F})^*(\omega) = F^*(\omega),\ \text{and}\ \tau(B,\cdot) = \tau,$$

then write such a menu B as B^F_τ.

Lemma 9. *For any $(F,\tau) \in \mathbb{H}^*$, $B^F_\tau \in \mathcal{K}$ is well-defined.*

Proof. Define $B^F_\tau := B_\tau + F^*$, where B_τ is the menu constructed applying Lemma 8 to τ. There is a plan $\{f_\omega\}_{\omega\in\Omega} \in \hat{c}(B_\tau)$ such that $f_\omega(\omega) = 0$. Then, let $\bar{F}(\omega) = f_\omega + F^*$ and observe $\overline{F}(\omega)(\omega) = F(\omega)(\omega)$ and $\overline{F} \in \hat{c}(B^F_\tau)$. □

For $F \in \mathbb{H}^*$, there may well be multiple plans $\overline{F}$ with the properties above. We denote $\overline{F}$ for one of them. The non-uniqueness does not cause a problem.

2.5.5 Preference relation between plans

We shall construct a preference relation between plans in $\mathbb{H}^*$. We start with constructing preference relations between menus. Recall that

$A \succeq^D B \Leftrightarrow$ For all $\omega \in \Omega$, $A \cap c(B,\omega) \neq \emptyset$,
$A \succeq^I B \Leftrightarrow$ There exists $n \in \mathbb{N}$ and $B_1, \dots B_n$ such that $A \succeq^D B_1 \succeq^D \cdots \succeq^D B_n \succeq^D B$.

The next lemma states that the translation invariance of c inherits to the relations $\succeq^D$, $\succeq^I$. For $A \in \mathcal{K}$ and $f \in \mathcal{A}$, let $A + f = \{g + f \mid g \in A\}$. For $F \in \mathcal{A}^\Omega$ and $f \in \mathcal{A}$, let $F + f \in \mathcal{A}^\Omega$ as $(F + f)(\omega) = F(\omega) + f$.

Lemma 10.

1. *If $A \succeq^D B$, then $(A + f) \succeq^D (B + f)$.*
2. *If $A \succeq^I B$, then $(A + f) \succeq^I (B + f)$.*

Proof. Suppose $A \succeq^D B$ and take $F \in \hat{c}(B)$ such that $\operatorname{Im} F \subset A$. Because $\hat{c}(B+f) = \hat{c}(B) + f$, $F + f \in \hat{c}(B+f)$ holds. In addition, $\operatorname{Im}(F+f) = (\operatorname{Im} F) + f \subset A + f$. Completed the first part.

Suppose $A \succeq^D C_1 \succeq^D \cdots \succeq^D C_n \succeq^D B$. Then, using the first part, $(A+f) \succeq^D (C_1 + f) \succeq^D \cdots \succeq^D (C_n + f) \succeq^D (B+f)$. Completed the second part. $\square$

If an implementation of a plan (F, τ) is observed when facing B, it should be the best plan DM can choose. The second part of the next lemma states this fact. For $f, g \in \mathcal{A}$, let $f \wedge g \in \mathcal{A}$ be the act defined by $(f \wedge g)(\omega) = \min\{f(\omega), g(\omega)\}$. We denote the closure of any set $S \subset \mathcal{A}$ as $\mathrm{cl}S$.

Lemma 11.

1. *Consider menus $A, B \in \mathcal{K}$ and plans $F, G \in \mathcal{A}^\Omega$ such that* $\mathrm{Im}F \subset A$, *$G \in \hat{c}(B)$, $F^* \geq G^*$, and F is $\mathcal{F}_{\tau_B}$-measurable. Then, $A \succeq^I B$ holds.*
2. *Suppose $F \in \hat{c}(B)$,* $\mathrm{Im}G \subset B$, *and $(G, \sigma) \in \mathbb{H}^*$. Then, $B^F_{\tau_B} \succeq^I B^G_\sigma$ holds.*

Proof. The set $C = \{f_\omega \wedge g_\omega\}_{\omega \in \Omega}$ is compact since it is a bounded closed set of $\mathbb{R}^\Omega$ and thus it is a menu. Write $h_\omega = F(\omega) \wedge G(\omega)$. For any $h \in C$, there exists $g \in B$ such that $g \geq h$, and so *INRA* implies

$$\forall \omega \in \Omega \ B \cap c(B \cup C, \omega) = c(B, \omega).$$

Next, we show that DM acquires more information when $B \cup C$ is given than when B is given.
Claim. $\mathrm{P}(B) \subset \mathrm{P}(B \cup C)$.

$\vdash$ The function $\varphi : \Omega \to \mathcal{K}$ defined by $\varphi(\omega) = c(B \cup C, \omega)$ is $(\mathrm{P}(B \cup C), \mathcal{B}(\mathcal{K}))$-measurable. And $\psi : \mathcal{K} \to \mathcal{K}$, $\psi(D) = B \cap D$ is continuous. Then, the composition $\psi \circ \varphi$ is $(\mathrm{P}(B \cup C), \mathcal{B}(\mathcal{K}))$-measurable. For any $V \in \mathcal{O}(\mathcal{K})$,

$$\{\omega | c(B, \omega) \in V\} = \{\omega | B \cap c(B \cup C, \omega) \in V\} = (\psi \circ \varphi)^{-1}(V) \in \mathrm{P}(B \cup C).$$

This means $\mathrm{P}(B) \subset \mathrm{P}(B \cup C)$. $\dashv$

The claim above and *IM* imply $\tau_B \leq \tau_{B \cup C}$ and thus $\mathcal{F}_{\tau_B} \subset \mathcal{F}_{\tau_{B \cup C}}$. Let

$$\Delta_\omega = \{\omega' \in \Omega | F(\omega') = F(\omega) \text{ and } G(\omega') = G(\omega)\}.$$

Then, because F and G are $\mathcal{F}_{\tau_B}$-measurable and $\mathcal{F}_{\tau_B} \subset \mathcal{F}_{\tau_{B\cup C}}$, $\Delta_\omega \in \mathcal{F}_{\tau_{B\cup C}}$ holds. Note that for any $\omega' \in \Delta_\omega$,

$$h_\omega(\omega') = F(\omega)(\omega') \wedge G(\omega)(\omega') = F(\omega')(\omega') \wedge G(\omega')(\omega') = F(\omega')(\omega') = F(\omega)(\omega').$$

Thus, because $\Delta_\omega \in \mathcal{F}_{\tau_{B\cup C}}$, *IM* and Lemma 2 (2), $h_\omega \in c(B \cup C, \omega)$. Conclude $C \succeq^D B \cup C \succeq^D B$. On the other hand, it is straightforward to show that $A \succeq^I C$. Combining these and complete the first part.

Applying the first part to the menus $B^F_{\tau_B}$, B, and B^F_σ shows the second part. □

We will need continuity of preference relation when eliciting subjective probability. For this reason, we use the topological closure $\succeq^*$ of $\succeq^I$:

$$A \succeq^* B \Leftrightarrow \text{There exist sequences } A_n \to A \text{ and } B_n \to B \text{ such that } A_n \succeq^I B_n.$$

Naturally, $\succeq^*$ is also translation invariant and transitive.

Lemma 12.

1. *If $A \succeq^* B$, then $A + f \succeq^* B + f$.*
2. *If $A \succeq^* B$ and $B \succeq^* C$, then $A \succeq^* C$.*

Proof. Assume $A \succeq^* B$ and take sequences $A_n \to A$, $B_n \to B$ with $A_n \succeq^I B$. Then, by Lemma 6, we have $A_n + f \succeq^I B_n + f$. Take $n \to \infty$ and complete the first part.

Assume $A \succeq^* B \succeq^* C$ and take sequences $A_n \to A$, $B_n \to B$, $B'_n \to B$, and $C_n \to C$ such that $A_n \succeq^I B_n$ and $B'_n \succeq^I C_n$. Wlog assume $d_h(B_n, B), d_h(B'_n, B) < n^{-1}$. Then, $A_n + n^{-1} \succeq^I B_n + n^{-1}$ and $B'_n - n^{-1} \succeq^I C_n - n^{-1}$. From the transitivity of $\succeq^I$, $A_n + n^{-1} \succeq^I C_n - n^{-1}$ follows. Take $n \to \infty$ and complete the second part. □

Finally, we define the preference relation over plans. For $F, G \in \mathbb{H}^*$, let

$$(F, \tau) \succeq (G, \sigma) \Leftrightarrow B^F_\tau \succeq^* B^G_\sigma.$$

And this relation is translation invariance in the following sense:

Lemma 13. *For $(F, \tau), (G, \sigma) \in \mathbb{H}^*$ and $f \in \mathcal{A}$,*

$$(F, \tau) \succeq (G, \sigma) \Rightarrow (F + f, \tau) \succeq (G + f, \tau).$$

Proof. First note that $B^F_\tau + f \sim^I B^{F+f}_\tau$ follows from Lemma 11 (1). By definition, $(F, \tau) \succeq (G, \sigma)$ means $B^F_\tau \succeq^* B^G_\sigma$. This and the linearity of $\succeq^*$ imply $B^F_\tau + f \succeq^* B^G_\sigma + f$. Combine this with $B^{F+f}_\tau \succeq^I B^F_\tau + f$ and $B^G_\sigma + f \succeq^I B^{G+f}_\sigma$. □

2.5.6 Representation

Now we turn to the elicitation of subjective probability and cost function. The first is that of subjective probability. For $f, g \in \mathcal{A}$, write $f \succeq^* g$ if $\{f\} \succeq^* \{g\}$ for notational simplicity.

Lemma 14. *There is a full-support probability π over Ω such that, for any $f, g \in \mathcal{A}$,*

$$f \succeq^* g \text{ implies } \int f d\pi \geq \int g d\pi.$$

Proof. By C (2) and the definition of $\succeq^*$, $1\omega 0 \succ^* 0$ holds for any ω and thus $\succeq^*$ is non-degenerate. The relation $\succeq^*$ is reflexive, transitive, monotonic, linear, continuous, and non-degenerate. Follow the argument in Lemma 9 of Ellis and obtain a probability π such that $f \succeq^* (\succ^*) g$ implies $\int f d\pi \geq (>) \int g d\pi$. Note that, for any ω, $1\omega 0 \succ^* 0$ and thus $\pi(\omega) > 0$. That is, π is full-support. $\square$

Next, we construct cost function on $\mathcal{T}^*$ and one-way utility representation on $\mathbb{H}^*$. The constructed cost function respects the point-wise order of response times. The idea of calibration is as follows. If $(F, \tau) \succeq (G, \sigma)$ and τ is more costly than σ, the benefit from F compared to G is large enough to compensate the cost increase. But benefit from each choice is calculated using the elicited probability and so we can evaluate the difference of cost. For a moment, denote f_τ for a plan $(F, \tau) \in \mathbb{H}^*$ where $F \in \hat{c}(B^F_\tau)$ and $F^* = f$. In the proof of the next lemma, for any act f, I sometimes denote $\pi(f)$ for the integration $\int f \mathrm{d}\pi$.

Lemma 15.

1. *There exists $\gamma^* : \mathcal{T}^* \to \overline{\mathbb{R}}$ and $V^* : \mathbb{H}^* \to \overline{\mathbb{R}}$ such that*

$$(F, \tau) \succeq (G, \sigma) \Rightarrow V^*(F, \tau) \geq V^*(G, \sigma),$$

where

$$V^*(F, \tau) = \int F^* d\pi - \gamma^*(\tau).$$

2. *For $\tau, \sigma \in \mathcal{T}^*$, if $\tau \leq \sigma$, then $\gamma^*(\tau) \leq \gamma^*(\sigma)$.*

Proof. By *IM*, given any singleton menu $\{f\}$, response time is always the same: $\underline{\tau} = \tau_{\{f\}}$ for some $\underline{\tau}$. Let $M_{\tau,\sigma} = \{f \in \mathcal{A} | f_\tau \succeq 0_\sigma\}$ and let

$$\gamma^*(\tau) = \inf_{f \in M_{\tau,\underline{\tau}}} \int f d\pi,$$
$$V^*(F,\tau) = \int F^* d\pi - \gamma^*(\tau).$$

Claim. $\inf_{f \in M_{\tau,\sigma}} \int f d\pi \geq \gamma^*(\tau) - \gamma^*(\sigma)$

$\vdash$ Take $g_n \in M_{\sigma,\underline{\tau}}$ with $\pi(g_n) \to \gamma^*(\sigma)$ and $h_n \in M_{\tau,\sigma}$ with $\pi(h_n) \to \inf_{h \in M_{\tau,\sigma}} \int h d\pi$.

Since $[g_n]_\sigma \succeq 0_{\underline{\tau}}$, $0_\sigma \succeq [-g_n]_{\underline{\tau}}$. Combining with $[h_n]_\tau \succeq 0_\sigma$, we have $[h_n]_\tau \succeq [-g_n]_{\underline{\tau}}$ or $[g_n + h_n]_\tau \succeq 0_{\underline{\tau}}$. Thus,

$$\gamma^*(\tau) = \inf_{f \in M_{\tau,\underline{\tau}}} \int f d\pi \leq \int g_n + h_n d\pi \to \gamma^*(\sigma) + \inf_{h \in M_{\tau,\underline{\tau}}} \int h d\pi.$$

That is, $\inf_{h \in M\tau,\sigma} \int h d\pi \geq \gamma^*(\tau) - \gamma^*(\sigma)$. $\dashv$

If $(F,\tau) \succeq (G,\sigma)$, then $[F^*]_\tau \succeq [G^*]_\sigma$ or $[F^* - G^*]_\tau \succeq 0_\sigma$. Then, by the claim above,

$$\int (F^* - G^*) d\pi \geq \gamma^*(\tau) - \gamma^*(\sigma),$$

or $V(F,\tau) \geq V(G,\sigma)$. This shows the first part.

Consider $\tau, \sigma \in \mathcal{T}$ such that $\tau \leq \sigma$. Then, take $F \in \hat{c}(B_\tau)$ and $G \in \hat{c}(B_\sigma)$ such that $F^* = G^* = 0$. Applying Lemma 11 (1), we obtain $(F,\tau) \succeq (G,\sigma)$, or $-\gamma^*(\tau) \geq -\gamma^*(\sigma)$ in terms of the representation. This shows the second part. □

Next, we extend the domain of V^* to $\mathbb{H}$ and show that the extension V is maximized by implemented plans.

Lemma 16. *If $F \in \hat{c}(B)$, $(G,\sigma) \in \mathbb{H}$, and $\mathrm{Im}G \subset B$, then $V(F,\tau_B) \geq V(G,\sigma)$, where $V : \mathbb{H} \to \overline{\mathbb{R}}$ and $\gamma : \mathcal{T} \to \overline{\mathbb{R}}$ is defined by*

$$V(F,\tau) = \int F^* d\pi - \gamma(\tau),$$
$$\gamma(\tau) = \inf_{\tilde{\tau} \in \mathcal{T}^*(\tau)} \gamma^*(\tilde{\tau}),$$

where $\mathcal{T}^*(\tau) = \{\tilde{\tau} \in \mathcal{T}^* | \tilde{\tau} \geq \tau\}$. Moreover, γ is increasing, that is, $\gamma(\tau) \geq \gamma(\sigma)$ whenever $\tau \geq \sigma$.

Proof. Take F and (G, σ) as the hypothesis of the statement. Then, by Lemma 11 (2), $(F, \tau_B) \succeq (G, \tilde{\sigma})$ holds for any $\tilde{\sigma} \in \mathcal{T}^*(\sigma)$. Since $\gamma(\tau_B) = \gamma^*(\tau_B)$, this implies

$$\int F^* d\pi - \gamma(\tau_B) \geq \int G^* - \gamma^*(\tilde{\sigma}).$$

Taking the supremum of the right hand side complete the proof. $\square$

The next lemma states that benefit from any implementable choice is bounded by an optimal choice. In the proof of the next lemma, let $N_\epsilon(f) = \{g \in \mathcal{A} | \, |f - g| < \epsilon\}$ for any $f \in \mathcal{A}$.

Lemma 17. *If* $\operatorname{Im} F \subset B$ *and* $\sigma(F) \subset \mathcal{F}$ *for some* σ*-algebra* $\mathcal{F}$ *over* Ω*, then*

$$\mathrm{E}[F^*] \leq E\left[\sup_{f \in B} \mathrm{E}[f|\mathcal{F}]\right].$$

Proof. It is sufficient to show that, for any $\Delta \in \mathcal{F}$,

$$\int_\Delta F^* d\pi \leq \int_\Delta \sup_{f \in B} \mathrm{E}[f|\mathcal{F}] d\pi.$$

Take any $\Delta \in \mathcal{F}$ and choose $\epsilon > 0$. Then, by compactness of B, there exist $f_1, \ldots, f_n \in \mathrm{cl}(\operatorname{Im} F)$ such that $\mathrm{cl}(\operatorname{Im} F) \subset \bigcup_{i=1}^n N_\epsilon(f_i)$. Define $\Delta_1, \ldots, \Delta_n$ by

$$\Delta_1 = \{\omega \in \Delta | F(\omega) \in N_\epsilon(f_1)\},$$
$$\Delta_{k+1} = \{\omega \in \Delta | F(\omega) \in N_\epsilon(f_{k+1})\} \setminus \bigcup_{i=1}^k \Delta_i.$$

Then, $\Delta = \sum_{i=1}^n \Delta_i$ and $\Delta_i \in \mathcal{F}$.

$$\begin{aligned}
\int_\Delta F^* d\pi &= \sum_{i=1}^n \int_{\Delta_i} F^* d\pi \leq \sum_{i=1}^n \int_{\Delta_i} \mathrm{E}[f_i + \epsilon | \mathcal{F}] d\pi \\
&\leq \epsilon + \sum_{i=1}^n \int_{\Delta_i} \sup_{f \in B} \mathrm{E}[f|\mathcal{F}] d\pi = \epsilon + \int_\Delta \sup_{f \in B} \mathrm{E}[f|\mathcal{F}] d\pi.
\end{aligned}$$

Because ϵ is arbitrary, $\int_\Delta F^* d\pi \leq \int_\Delta \sup_{f \in B} \mathrm{E}[f|\mathcal{F}] d\pi$. $\square$

Now, we show the optimality of implemented response times.

Lemma 18. *For any* $B \in \mathcal{K}$*,*

$$\tau_B \in \arg\max_{\tau \in \mathcal{T}} \mathrm{E}[\sup_{f \in B} \mathrm{E}[f|\mathcal{F}_\tau]] - \gamma(\tau).$$

Proof. Fix any $\sigma \in \mathcal{T}$. Take any $B \in \mathcal{K}$ and approximate it with finite nested menus: $B_n \to B$ and $B_n \subset B_{n+1}$. For each $B_n = \{f_1^n, \dots f_{N(n)}^n\}$, define

$$\Delta_1^n = \{\omega | \mathrm{E}[f_1^n|\mathcal{F}_\sigma](\omega) \geq \mathrm{E}[f_j^n|\mathcal{F}_\sigma](\omega)\ \forall j \neq 1\},$$
$$\Delta_{i+1}^n = \{\omega | \mathrm{E}[f_{i+1}^n|\mathcal{F}_\sigma](\omega) \geq \mathrm{E}[f_j^n|\mathcal{F}_\sigma](\omega)\ \forall j \neq i+1\} \backslash \bigcup_{j=1}^{i} \Delta_j^n.$$

Then define a plan G_n by $G_n(\omega) = f_i^n$ if $\omega \in \Delta_i^n$, which is $\mathcal{F}_\sigma$-measurable. The constructed plan G_n may depend on the way the conditional expectations are taken, but it is well-defined once we fix conditional expectations.

The plan G_n is an element of B^Ω, which is compact in product topology. Thus it is wlog to assume $\{G_n\}$ converges to some $G \in B^\Omega$, which is also $\mathcal{F}_\sigma$-measurable. Since

$$\mathrm{E}[G_n^*|\mathcal{F}_\sigma](\omega) = \sup_{f \in B_n} \mathrm{E}[f|\mathcal{F}_\sigma](\omega) \text{ a.s.}$$

holds for each n,

$$\mathrm{E}[G^*|\mathcal{F}_\sigma](\omega) = \sup_{f \in B} \mathrm{E}[f|\mathcal{F}_\sigma](\omega) \text{ a.s..}$$

Let $F \in \hat{c}(B)$ and $\tau = \tau_B$. Then, by Lemma 16 and Lemma 17,

$$\mathrm{E}[\sup_{f \in B} \mathrm{E}[f|\mathcal{F}_\tau]] - \gamma(\tau) \geq \mathrm{E}[F^*] - \gamma(\tau) \geq \mathrm{E}[G^*] - \gamma(\sigma) = \mathrm{E}[\sup_{f \in B} \mathrm{E}[f|\mathcal{F}_\sigma]] - \gamma(\sigma).$$

This completes the proof. $\square$

Finally, we show the optimality of the implemented choices.

Lemma 19. *$c(B, \omega) = \arg\max_{f \in B} \mathrm{E}[f|\mathcal{F}_{\tau_B}](\omega)$ for all ω*

Proof. First, we show $c(B, \omega) \subset \arg\max_{f \in B} \mathrm{E}[f|\mathcal{F}_{\tau_B}](\omega)$ for all $\omega \in \Omega$. Take $f \in c(B, \omega)$ and any ω. En route to a contradiction, suppose there exists $g \in B$ with $\mathrm{E}[g|\mathcal{F}_{\tau_B}](\omega) > \mathrm{E}[f|\mathcal{F}_{\tau_B}](\omega)$. Take any selctor F of $c(B, \cdot)$ that satisfies $F(\omega) = f$ and define a plan G by

$$G(\omega') = \begin{cases} g & \text{if } \omega' \in \mathcal{F}_{\tau_B}(\omega) \\ F(\omega') & \text{otherwise.} \end{cases}$$

Since $(G, \tau_B) \in \mathbb{H}$, Lemma 16 implies $V(F, \tau_B) \geq V(G, \tau_B)$, or $\mathrm{E}[F^*] \geq \mathrm{E}[G^*]$, which is a contradiction. Thus $c(B, \omega) \subset \arg\max_{f \in B} \mathrm{E}[f|\mathcal{F}_{\tau_B}](\omega)$.

Next, we consider the converse inclusion. Fix any ω^* and take $f \in \arg\max_{h \in B} \mathrm{E}[h|\mathcal{F}_{\tau_B}](\omega^*)$. Let $\mathrm{P}(B) = \{\Delta_0, \ldots, \Delta_I\}$ where $\Delta_0 = \mathrm{P}(B)(\omega^*)$. Take $F \in \hat{c}(B)$ and set $g_0 = f$, $g_i = F(\omega)$ for $\omega \in \Delta_i$. Define new acts $g_i^n = g_i - n^{-1}\mathbb{1}_{\Delta_i^c}$ for each i and consider a menu

$$B^n = (B - n^{-1}) \cup \{g_i^n\}_{i=0}^I.$$

Let G^n be a plan defined by $G^n(\omega) = g_i^n$ for $\omega \in \Delta_i$. We show that $\hat{c}(B^n) = \{G^n\}$. To this end, take any $(H, \sigma) \in \mathbb{H}$ with $\mathrm{Im}\, H \subset B^n$ and $H(\omega) \neq G(\omega)$ for some ω. Note that, by the construction of B^n, there is a plan $(\tilde{H}, \sigma)$ with the following properties: $\mathrm{Im}\tilde{H} \subset B$, $\tilde{H}^* \geq H^*$, and $\tilde{H}^*(\omega) > H^*(\omega)$ with positive probability. Then, by the latter two properties $V(\tilde{H}, \sigma) > V(H, \sigma)$ holds. In addition, by Lemma 16, $V(F, \tau_B) \geq V(\tilde{H}, \sigma)$. Combining these inequalities shows $V(F, \tau_B) > V(H, \sigma)$. On the otherhand, $(G^n, \tau_B) \in \mathbb{H}$ and $V(G^n, \tau_B) = V(F, \tau_B)$. Then, Lemma 16 implies $\hat{c}(B^n) = \{G^n\}$ and we conclude $\mathrm{P}(B^n) = \mathrm{P}(B)$. This and *IM* imply $\tau_{B_n} = \tau_B$. Note that $B^n \to B$ and $g_0^n \to f$, and $g_0^n \in c(B^n, \omega^*)$. Apply *C* (1) and conclude $f \in c(B, \omega^*)$. □

Lemma 18 and 19 completes the proof of Theorem 1.

Chapter 3

Axiomatization of optimal inattention model with infinite state space

3.1 Introduction

In typical economic environments, agents may not be fully attentive to all available information, possibly because paying attention requires cognitive stress. This phenomenon is called *rational inattention*. Building on this intuition, Sims (2003) modeled a decision-maker (DM) with limited information-processing capacity. Inspired by his works, many alternative models were provided. Recent experiments confirmed limited attention to information in laboratories.

In decision theoretic literature, de Oliveira et al. (2017) and Ellis (2018) (henceforth, Ellis) axiomatizd diffrerent models of rational inattention. Their models share a story of rational inattention: DM first acquires costly information, and then chooses an alternative according to it. But they used different primitives. While de Oliveira et al. (2017) adopted a menu choice framework, Ellis introduced a new primitive, state-conditional choice correspondence.

This chapter generalizes the analysis of Ellis and axiomatizes a model in which information is modeled as a σ-algebra over a possibly infinite state-space. Since various economic environments are described with infinite state spaces, this generalization is important to extend the applicability. For example, infinite state space is often required to introduce continuous probability distribution, such as normal distribution. The generalization I conduct is important to provide an axiomatic foundation to optimal inattention model in such cases.

The identification of parameters is roughly the same as that of Ellis. From the choice correspondence, I construct an incomplete preference relation over menus. From this relation, preferences for *plans*, functions that describe what DM chooses in each state, are defined. I apply the variational technique to elicit the parameters from this relation. There are four differences in the identification processes in this paper and Ellis'. First, I transform each act into a utility act to simplify the analysis. Second, I first construct a utility representation only for plans that generates σ-algebras actually chosen given some menu, and then extend it to all the plans. Ellis do not follow this step and my argument fills this gap. Third, I posit a new axiom *Cyclical Consistency* that guaranees that the preference relation for plans is not degenerate. In the proof of Ellis, which do not assume this axiom, this relation may well be degenerate and so the elicitation of subjective probability is not justified. Fourth, since the state-space is infinite and the representation includes integration, some care on the measurability issue is taken.

The rest of the chapter is organized as follows. Section 2 describes the setup and the model. Section 3 introduces the axioms and states the representation theorem. Section 4 provides the proofs.

3.2 Model

3.2.1 Setup

This subsection introduces the setup. Each metric space S introduced below is endowed with its Borel σ-algebra $\mathcal{B}(S)$[1]. Let (Ω, Σ) be a measurable space, where the set Σ is a σ-algebra. The set Ω is possibly infinite and interpreted as the set of states that describes uncertainty. Let $\mathbb{P}$ be the set of all sub σ-algebra of Σ. Let X be a convex subset of a metrizable topological vector space and let d be its compatible metric. Let $\mathcal{A}$ be a set of $(\Sigma, \mathcal{B}(X))$-measurable functions from Ω to X. With a natural isomorphism, we regard X as the set of constant function and assume $X \subset \mathcal{A}$. Each element of $\mathcal{A}$, interpreted as a choice, is called an *act*. A concrete definition of $\mathcal{A}$ will be provided in the next paragraph. The set $\mathcal{A}$ is endowed with the uniform metric $d_\infty(f, g) = \sup_{\omega \in \Omega} d(f(\omega), g(\omega))$. Let $\mathcal{K}$ be the set of all non-empty compact sets of $\mathcal{A}$ that is endowed with the Hausdorff metric d_h. Each element of $\mathcal{K}$ is interpreted as a choice sets consisting of acts. For typical elements of the sets above, we write $\mathcal{F}, \mathcal{G}, \mathcal{H} \in \mathbb{P}$, $x, y, z \in X$, $f, g, h \in \mathcal{A}$, and $A, B, C \in \mathcal{K}$.

[1]Since I admit infinite state-space, the introduction of σ-algebra is necessary.

Our choice data is a conditional choice correspondence $c : \mathcal{K} \times \Omega \to \mathcal{K}$ such that $c(B, \omega) \subset B$ for any $(B, \omega) \in \mathcal{K} \times \Omega$. We assume $c(B, \cdot) : \Omega \to \mathcal{K}$, interpreted as a function of ω, is $(\Sigma, \mathcal{B}(\mathcal{K}))$-measurable. For each $B \in \mathcal{K}$, let $P(B)$ be the σ-algebra over Ω generated by $c(B, \cdot)$.

We can elicit the preference relation over outcomes from c. For $x, y \in X$, define

$$x \succeq^R y \Leftrightarrow \text{there exists an } \omega \in \Omega \text{ such that } x \in c(\{x, y\}, \omega).$$

For technical reason, we focus on acts that is bounded in terms of $\succeq^R$. That is,

$$f \in \mathcal{A} \Leftrightarrow \text{there exist } x, y \in X \text{ such that, for all } \omega \in \Omega,\ x \succeq^R f(\omega) \succeq^R y.$$

3.2.2 Model

We provide an axiomatic foundation of the following model, called the Optimal Inattention Model (OIM).

$$P(B) \in \arg\max_{\mathcal{F} \in \mathbb{P}} \mathrm{E}[\max_{f \in B} \mathrm{E}[u(f)|\mathcal{F}]] - \gamma(\mathcal{F}),$$
$$c(B, \omega) = \arg\max_{f \in B} E[u(f)|P(B)](\omega) \ \pi\text{-a.s.}$$

In the above expression, $u : X \to \mathbb{R}$ denotes an expected utility function, which is affine and unbouned above and below. Let $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$. The function $\gamma : \mathbb{P} \to \overline{\mathbb{R}}$ is a cost function, which is monotone for inclusion relation of σ algebras. Finally, the expectation is taken using some subjective probability π, which is countably additive.

This model is a snapshot of a rationally inattentive decision-maker. Before choosing an act, DM decides how much information he acquires. This information amount is coded as the σ-algebra $\mathcal{F}$, with which he updates the prior belief. The above model requires that the information is optimally chosen in consideration of the benefit of using it and the cost of its acquisition.

3.3 Main result

In this section, we introduce the axioms to characterize OIM. Most of them are the same as those adopted in Ellis. Thus, we follow Ellis to label each of the axioms. For a technical reason, the continuity axiom is strengthened, and a new axiom *Cyclical Consistency* is added.

The first axiom Independence of Nonrerevant Alternative is a variant of the Weak Axiom of Revealed Preference that is adapted to the conditional choice correspondence.

Axiom 1 (*INRA*)**.** *If $A \subset B$ and $A \cap c(B, \omega) \neq \emptyset$ for any $\omega \in \Omega$, then $c(A, \omega) = A \cap c(B, \omega)$.*

For $f, g \in \mathcal{A}$ and $\alpha \in [0, 1]$, let $\alpha f + (1-\alpha)g$ be an act such that $[\alpha f + (1-\alpha)g](\omega) = \alpha f(\omega) + (1-\alpha)g(\omega)$. For $f \in \mathcal{A}$ and $B \in \mathcal{K}$, let $\alpha f + (1-\alpha)B = \{\alpha f + (1-\alpha)g | g \in B\}$.

Attention constrained Independence is a variant of indepedence axiom that holds even with information cost.

Axiom 2 (*ACI*)**.** *If $\alpha g + (1-\alpha)f \in c(\alpha g + (1-\alpha)B, \omega)$, then $\alpha h + (1-\alpha)f \in c(\alpha h + (1-\alpha)B, \omega)$*

Monotonicity states that if an act that is dominated by another one in all the state and the dominated one is chosen, then DM must also choose the dominated one.

Axiom 3 (*M*)**.** *For $f, g \in B$, if $f(\omega) \succeq^R g(\omega)$ for all $\omega \in \Omega$, then*

$$g \in c(B, \omega) \Rightarrow f \in c(B, \omega).$$

The next assumption *Subjective Consequentialism* requires that DM respects the obtained information $P(B)$.

Axiom 4 (*SC*)**.** *For $f, g \in B$ and $\omega \in \Omega$, if*

$$f(\omega') = g(\omega') \text{ for all } \omega' \in P(B)(\omega),$$

then

$$f \in c(B, \omega) \Leftrightarrow g \in c(B, \omega).$$

Define a relation $\succeq^D$ by

$$A \succeq^D B \Leftrightarrow \text{For all } \omega \in \Omega,\ A \cap c(B, \omega) \neq \emptyset.$$

The relation $A \succeq^D B$ means that, facing A, DM can emurate the optimal policy he uses if B is given. And thus, A is directly selected over B. Let $\succeq^I$ denotes the transitive closure of $\succeq^D$. That is,

$$A \succeq^I B \Leftrightarrow \text{There exists } n \in \mathbb{N} \text{ and } B_1, \dots B_n \text{ such that } A \succeq^D B_1 \succeq^D \cdots \succeq^D B_n \succeq^D B.$$

The relation $A \succeq^I B$ means that A is indirectly selected over B. For a sequnce of events $\{E_n\}_{n=1}^{\infty} \subset \Sigma$, write $E_n \downarrow \emptyset$ if $E_1 \supset E_2 \supset \dots$ and $\bigcap_{n=1}^{\infty} E_n = \emptyset$. The next axiom *Cyclical Consistency*, *CC* for short, is a new technical axiom.

Axiom 5 (*CC*)**.** *For $x, y \in X$, if $x \succeq^I y$ and $y \succeq^I x$, then $x \sim^R y$.*

The next is *Continuity.* Its first part requires countable additivity of the subjective probability. Without this axiom, elicited subjective probability may be only finitely additive. In that case, conditional expectations with the subjective probability, which is necessary for OIM, may not exist. On contrast, conditional expectations for countably additive probabilities are always well-defined. The second part is a weak form of upper hemicontinuity and is the same as in Ellis. For $x, z \in X$ and an event E, let

$$xEy(\omega) = \begin{cases} x & \text{if } \omega \in E \\ z & \text{otherwise.} \end{cases}$$

Let $\succ^R$ be the asymmetric part of $\succeq^R$.

Axiom 6 (*C*)**.**
1. *For $x, y, z \in X$ and $\{E_n\}_{n=1}^{\infty}$, if*

$$x \succ^R y \succ^R z \text{ and } E_n \downarrow \emptyset,$$

   then, there exists some $n \in \mathbb{N}$ such that

$$y \succeq^I xE_nz.$$

2. *For any $\omega \in \Omega$ and sequences $\{B_n\}_{n=1}^{\infty}$ and $\{f_n\}_{n=1}^{\infty}$ such that $B_n \to B$ and $f_n \to f$, and $f_n \in c(B_n, \omega)$, if*

$$P(B_n)(\omega) = P(B)(\omega) \text{ for any } n \in \mathbb{N},$$

   then

$$f \in c(B, \omega).$$

The last axiom is *Unboundedness*, which is necessary to calibrate the cost function. It states that no matter how an outcome is preferred to another one, a mixture of sufficiently better or worse one to an appropriate outcome reverses the preference.

Axiom 7 *(U). There exist $x, y \in X$ such that $x \succ^R y$ and, for any $\beta \in (0,1)$, there exist $z^*, z_* \in X$ such that*

$$\beta z^* + (1-\beta)y \succ^R x,\ y \succ^R \beta z_* + (1-\beta)y.$$

Now I present the suffuciency result.

Theorem 1. *If c satisfies INRA, ACI, M, SC, C, U, and CC, then there exists (u, π, γ) such that*

$$P(B) \in \arg\max_{\mathcal{F}\in\mathbb{P}} \mathrm{E}[\max_{f\in B} \mathrm{E}[u(f)|\mathcal{F}]] - \gamma(\mathcal{F})$$
$$c(B, \omega) \subset \arg\max_{f\in B} \mathrm{E}[u(f)|P(B)](\omega)\ \pi\text{-}a.s.$$

The theorem states that the axioms I posited above is a sufficient condition for the choice correpondence to have an OIM. While Ellis assumed finite state-space, his result is generalized to the infinite-state case. Note that, in the representation, the chosen acts are not all of the utility-maximizing ones. This subtlety comes because a single state may have zero probability in infinite state-space.

3.4 Proof

3.4.1 Roadmap

Here I provide a roadmap of the proof of Theorem 1. We construct an expected utility function in Lemma 1 and show, in Lemma 2, that choice behavior only depends on utilities given in each state by acts in menus. Through Lemma 3 to Lemma 6, I translate the set of menus $\mathcal{K}$ to the set of compact subsets of measurable real functions $\tilde{f} : \Omega \to \mathbb{R}$.

Then, I introduce a notion *plan*. A plan F is a function $\mathcal{F} : \Omega \to \mathcal{A}$ that describes what DM would choose in each state. Later, I will construct a preference relation for plans to calibrate parameters from it. But I first define preferences for menus from the primitive, and then define those for plans from this. In Lemma 7 and Lemma 8, I show that, for any plan, there exists a menu given which DM implements the plan. This is done to show that preferences for plans are well-defined later.

The preference relation for menus is the topological closure of $\succeq^I$. Through Lemma 10 to Lemma 13, I show structural properties of the preference relation $\succeq$ for plans. Its two important properties are translation invariance and that it is consistent with the choice correspondence: if a plan is implemented, that is the best one DM can choose.

Then, I elicit a subjective probability in Lemma 14 and a cost function in Lemma 15, which define a utility representation over a restricted domain of plans. Lemma 16 shows the monotonicity of the cost function with respect to the inclusion relation of σ-algebras. Through Lemma 17 to Lemma 19, I extend the representation to all of the plans. Lemma 20, 21, and 22 completes the proof of the sufficiency result.

3.4.2 Basic properties of choice correspondence

In this subsection, we investigate the basic properties of the choice correspondence.

First, we construct an expected utility function. Let $\mathcal{K}(X)$ be the set of all non-empty compact subsets of X.

Lemma 1. *There exists an continuous affine function $u : X \to \mathbb{R}$ such that, for any $B \in \mathcal{K}(X)$,*

$$x \in c(B, \omega) \Leftrightarrow u(x) \geq u(y) \ \forall y \in B.$$

and $u(X) = \mathbb{R}$.

Proof. Apply Lemma 1 in Ellis and obtain a function u with the properties we want except continuity. We show that the sets

$$\{y \in X | y \succeq^R x\} \text{ and } \{y \in X | x \succeq^R y\},$$

are closed. First, note that M implies that, for any $x, y \in X$ and $\omega, \omega' \in \Omega$, $c(\{x, y\}, \omega) = c(\{x, y\}, \omega')$. Suppose $y_n \to y$, $y_n \succeq^R x$ and take any $\omega \in \Omega$. Then, it follows that $\{y_n, x\} \to \{y, x\}$ in $\mathcal{K}$ and Axiom C implies $y \in c(\{y, x\}, \omega)$, or $y \succeq^R x$.

So for all $x, y \in X$ such that $x \succ^R y$, the set $\{z | x \succ^R z \succ^R y\}$ is open. Because u is an affine function bounded on an open set, it is continuous. □

The next lemma states that choice behavior follows two regularities: first, adding acts that are dominated by existing ones do not change what to be chosen; secondly, the choice behavior respects the acquired information. For $B \in \mathcal{K}$ and $\omega \in \Omega$, let $P(B)(\omega) = \{\omega' \in \mathcal{K} | \ c(B, \omega') = c(B, \omega)\}$. Because $\mathcal{K}$ is a metric space, its singleton is a closed set. Since $P(B)(\omega)$ is an inverse image of such a set, it is a measurable set.

Lemma 2.

1. Assume that, for any $g \in B$, there exists $f \in A$ such that $u \circ f \geq u \circ g$. Then

$$c(A, \omega) = A \cap c(A \cup B, \omega).$$

2. For $f, g \in B$, if $f \in c(B, \omega)$ and $u(g(\omega')) = u(f(\omega'))$ for all $\omega' \in P(B)(\omega)$, then $g \in c(B, \omega)$.

Proof. Suppose that $g \in B \cap c(A \cup B, \omega)$. By the assumption, there is some $f \in A$ such that $u \circ f \geq u \circ g$. This and M implies $f \in c(A \cup B, \omega)$. We have shown $A \cap c(A \cup B, \omega) \neq \emptyset$ for any $\omega \in \Omega$. Apply *INRA*, and the first part is completed.

The second part is proved as Lemma 2 in Ellis. □

3.4.3 Transforming acts into utility acts

We collect preliminary results to work on a real function space instead of acts. While this preparation is not done in Ellis, it is useful to consider infinite state space cases.

Let $\mathcal{B}_b$ be the set of all $(\Sigma, \mathcal{B}(\mathbb{R}))$-measurable bounded functions $f : \Omega \to \mathbb{R}$, which is endowed with the uniform norm $\|\cdot\|$. In this section, we translate each act $f \in \mathcal{A}$ into its utility act $u \circ f \in \mathcal{B}_b$ to work on $\mathcal{B}_b$ from the next section.

First, we construct a set $Y \subset X$ such that $u(Y) = \mathbb{R}$ and the restriction of u to Y is a homeomorphism. For each $n \in \mathbb{Z}$, take $x_n \in X$ such that $u(x_n) = n$. Let $Y_n = \{(1-\alpha)x_n + \alpha x_{n+1} | \alpha \in [0, 1]\}$ and $Y = \bigcup_{n \in \mathbb{Z}} Y_n$. Let v be the inverse function of $u|_Y$, which exists because of the definition of Y. For each $n \in \mathbb{Z}$, let $v_n : [n, n+1] \to Y_n$ be the inverse function of $u|_{Y_n}$.

Lemma 3.

1. For each $n \in \mathbb{Z}$, the function v_n is uniformly continuous.

2. The function u is a homeomorphism.

Proof. Define a function $\overline{v}_n : \mathbb{R} \to X$ by

$$\overline{v}_n(\beta) = (1 - (\beta - n))x_n + (\beta - n)x_{n+1}.$$

For $\beta \in [n, n+1]$, $\overline{v}_n(\beta) \in Y_n$, and besides the affinity of u implies $u(\overline{v}_n(\beta)) = \beta$. So the restriction of $\overline{v}_n$ to $[n, n+1]$ is v_n. Because the addition and the scalar multiplication is continuous in any topological vector space, $\overline{v}_n$ is continuous. Since it is affine, it is uniformly continuous, and so is v_n. The first part is complete.

Turn to the second part. It is sufficient to show that v is continuous. To this end, take an sequence $\{\hat{x}_k\}_{k=1}^{\infty} \in \mathbb{R}$ and suppose $\hat{x}_k \to \hat{x}$ in $\mathbb{R}$. If $\hat{x} \in (n, n+1)$

for some $n \in \mathbb{Z}$, for sufficently large k, $\hat{x}_k \in (n, n+1)$. The first part implies $v(\hat{x}_k) = v_n(\hat{x}_k) \to v_n(\hat{x}) = v(\hat{x})$. Turn to the case that $\hat{x} = n$ for some n. Then, for sufficiently large k, $\hat{x}_k \in (n-1, n+1)$. Take any $\epsilon > 0$. From the coninuity of v_{n-1} and v_n, there exist some $\delta > 0$ such that if $\hat{x}_k \in [n-1, n]$ and $|\hat{x}_k - \hat{x}| < \delta$, then $d(v_{n-1}(\hat{x}_k), v_{n-1}(x_k)) < \epsilon$; and if $\hat{x}_k \in [n, n+1]$ and $|\hat{x}_k - \hat{x}| < \delta$, then $d(v_n(\hat{x}_k), v_n(\hat{x})) < \epsilon$. So $v(\hat{x}_k) \to v(\hat{x})$ in Y. □

Let $\mathcal{A}_Y = \{f \in \mathcal{A} | \ \forall \omega \in \Omega \ f(\omega) \in Y\}$. In the next lemma, we denote $u \circ f$ for the composite function of u and f. Define functions $\Phi^* : \mathcal{A} \to \mathcal{B}_b$, $\Phi : \mathcal{A}_Y \to \mathcal{B}_b$, $\Psi : \mathcal{A} \to \mathcal{A}_Y$ by

$$\Phi^*(f) = u \circ f, \ \Phi(f) = u \circ f, \ \text{and} \ \Psi = \Phi^{-1} \circ \Phi^*.$$

Lemma 4.

1. *The function $\Phi^* : \mathcal{A}$ is continuous.*
2. *The function $\Phi : \mathcal{A}_Y$ is a homeomorphism.*
3. *The function Ψ is continuous and $u \circ \Psi(f) = u \circ f$ for any $f \in \mathcal{A}$.*

Proof. Consider the first part. Let $f_k \to f$ in $\mathcal{A}$. Because u is a continuous affine function, it is uniformly continuous. So, for any $\epsilon > 0$, there is some $\delta > 0$ such that $d(x, y) < \delta$ implies $|u(x) - u(y)| < \epsilon$. Note that, for sufficiently large k, $d_\infty(f_k, f) < \delta$. So, for all $\omega \in \Omega$, $|u(f_k(\omega)) - u(f(\omega))| < \epsilon$. That is, $\|u \circ f_k - u \circ f\| \to 0$. The first part is complete.

Consider the second part. The continuity of Φ follows from that of Φ^*. We shall prove that Φ is a bijection. For $\hat{f} \in \mathcal{B}_b$, $\Phi(v \circ \hat{f}) = u \circ v \circ \hat{f} = \hat{f}$. So it is onto. Take $f, g \in \mathcal{A}_Y$ such that $u \circ f = u \circ g$. Then, $f = v \circ u \circ f = v \circ u \circ g = g$. So it is one-to-one.

Finally, we show that the inverse function $\Phi^{-1} : \mathcal{B}_b \to \mathcal{A}_Y$ is continuous. Note that $\Phi^{-1}(\hat{f}) = v \circ \hat{f}$. Take a sequence $\{\hat{f}_k\}_{k=1}^\infty \in \mathcal{B}_b$ such that $\hat{f}_k \to \hat{f}$ in $\mathcal{B}_b$. Take $n \in \mathbb{N}$ such that $-n < \hat{f} < n$. Fix any $\epsilon > 0$. The uniform continuity of v_j implies that for each $j = -n-1, \ldots, n$, there exist $\delta_j > 0$ such that, if $x, y \in [j, j+1]$ and $|x - y| < \delta_j$, then $d(v_j(x), v_j(y)) < \epsilon$. Let $\delta = \min_j \delta_j$. Fix a sufficently large k so that $\|\hat{f}_k - \hat{f}\| < \min\{\delta, 1\}$. Then, for any $\omega \in \Omega$, there exists j such that $\hat{f}_k(\omega), \hat{f}(\omega) \in [j, j+1]$. Hence, for all $\omega \in \Omega$, $d(v(\hat{f}_k(\omega)), v(\hat{f}(\omega))) < \epsilon$. So $d_\infty(\Phi^{-1}(\hat{f}_k), \Phi^{-1}(\hat{f})) \leq \epsilon$. □

Next, we show that the choice behavior depends only on the state-dependent utilities of the acts in the choice sets. For a moment, we denote $u \circ f$ as f^u and let $A^u = \{f^u | f \in A\}$ for the sets of acts.

Lemma 5. *If $A^u = B^u$, then $[c(A,\omega)]^u = [c(B,\omega)]^u$ for any $\omega \in \Omega$.*

Proof. Note that Lemma 2 (1) imply

$$c(A,\omega) = A \cap c(A \cup B, \omega),\ c(B,\omega) = B \cap c(A \cup B, \omega).$$

for any $\omega \in \Omega$. Take any $f \in c(A,\omega)$. Then $f \in c(A \cup B, \omega)$ follows from the above-mentioned remark. Take $g \in B$ such that $f^u = g^u$. Lemma 2 (2) implies $g \in c(A \cup B, \omega)$. Then $g \in c(B,\omega)$. □

The next lemma states that the choice correspondence is translation invariant in terms of utility. For each $\tilde{B} \in \mathcal{K}(\mathcal{B}_b)$ and $\alpha \in \mathbb{R}$, Let $\alpha\tilde{f}$ is defined by $(\alpha\tilde{f})(\omega) = \alpha\tilde{f}(\omega)$ for $\alpha \in \mathbb{R}$ and $\tilde{f} \in \mathcal{B}_b$ and let $\alpha\tilde{B} = \{\alpha\tilde{f} | \tilde{f} \in \tilde{B}\}$

Lemma 6. *The equation $[c(B+g,\omega)]^u = [c(B,\omega)]^u + g^u$ holds.*

Proof. Suppose $A^u = 2B^u$, $f^u = 2g^u$, and $x^u = 0$. Then, $B^u = (\frac{1}{2}A + \frac{1}{2}x)^u$ and $B^u + g^u = (\frac{1}{2}A + \frac{1}{2}f)^u$. Then, Lemma 5 implies

$$[c(B+g,\omega)]^u = \left[c\left(\frac{1}{2}A + \frac{1}{2}f, \omega\right)\right]^u. \tag{3.1}$$

On the other hand, ACI implies

$$\frac{1}{2}g + \frac{1}{2}f \in c\left(\frac{1}{2}A + \frac{1}{2}f, \omega\right) \Leftrightarrow \frac{1}{2}g + \frac{1}{2}x \in c\left(\frac{1}{2}A + \frac{1}{2}x, \omega\right).$$

Thus

$$\begin{aligned}
\left[c\left(\frac{1}{2}A + \frac{1}{2}f, \omega\right)\right]^u &= \left[c\left(\frac{1}{2}A + \frac{1}{2}x, \omega\right)\right]^u + \frac{1}{2}f^u \\
&= \left[c\left(\frac{1}{2}A + \frac{1}{2}x, \omega\right)\right]^u + g^u \\
&= [c(B,\omega)]^u + g^u
\end{aligned} \tag{3.2}$$

Combining (3.1) and (3.2) completes the proof. □

Now we define a choice correspondence $\tilde{c} : \mathcal{K}(\mathcal{B}_b) \times \Omega \to \mathcal{K}(\mathcal{B}_b)$ by

$$\tilde{c}(B,\omega) = \Phi[c(\Phi^{-1}(B),\omega)].$$

Then, $\tilde{c}$ inherits all the properties of c. Besides, $\tilde{c}$ is translation invariant:

$$\tilde{c}(B+f,\omega) = \tilde{c}(B,\omega) + f$$

by the virtue of Lemma 6. Note

$$c(B,\omega)=\Phi^{-1}(\tilde{c}(\Phi(B),\omega)),$$

for $B\in\mathcal{K}(\mathcal{A}_Y)$. Once we found an OIM of $\tilde{c}$, the obtained subjective probability and cost function work for c. From now on, we write $\mathcal{B}_b$ as $\mathcal{A}$, and $\tilde{c}$ as c for simplicity.

3.4.4 Some auxiliary results

When DM chooses information, he preplans what he will choose conditional on future signals. Here, we shall elicit a preference relation between such plans. Using it, we can calibrate the information cost γ. Formally, we call a $(\Sigma,\mathcal{B}(\mathcal{A}))$-measurable bounded function $F:\Omega\to\mathcal{A}$ a *plan*. Let $\mathbb{H}$ denote the set of all plans.

Let $\sigma(F)$ be the σ-algebra over Ω generated by F. This is the minimal information that enables DM to implement F. Let

$$\mathbb{H}^*=\{F\in\mathbb{H}\,|\,\text{there is some } B\in\mathcal{K} \text{ such that } \sigma(F)=P(B)\}.$$

The set $\mathbb{H}^*$ consists of plans that can be implemented by fully using some information DM aquires given some menu. We especially pay attention to plans in $\mathbb{H}^*$ because this property facilitates the calibration of γ.

For $B\in\mathcal{K}$, let $\hat{c}(B)$ be the set of $(\Sigma,\mathcal{B}(B))$-measurable functions $F:\Omega\to B$ that satisfy $F(\omega)\in c(B,\omega)$. The set $\hat{c}(B)$ is nonempty by the virtue of Aliprantis and Border (2006) (AB, henceforth) Theorem 18.13.

Let $\mathbb{P}^*=\{P(B)\,|\,B\in\mathcal{K}\}$. This is the set of information DM uses. Henceforth, the notation $\{f_\omega\}_{\omega\in\Omega}$ sometimes denotes for the plan $\omega\mapsto f_\omega$ and sometimes for the set $\{f_\omega\,|\,\omega\in\Omega\}$. Let $\mathcal{O}(S)$ be the set of all open sets in S.

Lemma 7. *For any $\mathcal{F}\in\mathbb{P}^*$, there exists some $B_{\mathcal{F}}\in\mathcal{K}$ and a plan $F=\{f_\omega\}_{\omega\in\Omega}$, and satisfy the following.*
1. $P(B_{\mathcal{F}})=\mathcal{F}$,
2. $f_\omega(\omega)=0,\ f_\omega\in c(B_{\mathcal{F}},\omega)$,
3. For any $g\in\mathcal{A}$, $f_\omega+g\in c(B_{\mathcal{F}}+g,\omega)$.

Proof. Take $B\in\mathcal{K}$ such that $P(B)=\mathcal{F}$. Apply AB Theorem 18.13 and take a measurable selector $\{g_\omega\}_{\omega\in\Omega}$ of the correspondence $c(B,\cdot)$.

Next, I construct a new act h as $h(\omega)=g_\omega(\omega)$. Its measurability is not trivial and this is what I show next.

Claim. The function $h : \Omega \to \mathbb{R}$, defined by $h(\omega) = g_\omega(\omega)$ is in $\mathcal{A}$.

$\vdash$ What we show is the boundedness and the $(\Sigma, \mathcal{B}(\mathbb{R}))$-measurability of h. The boundedness follows from the compactness of B.

Turn to the measurability. Define functions $\varphi : \Omega \to \Omega \times B$, $\varphi(\omega) = (\omega, g_\omega)$, and $\psi : \Omega \times B \to \mathbb{R}$, $\psi(\omega, g) = g(\omega)$. Note $h = \psi \circ \varphi$ and turn to the measurability of φ and ψ.

First, φ is a vector of two measurable functions and so is $(\Sigma, \Sigma \otimes \mathcal{B}(B))$-measurable. To prove the measurability of ψ, we check that it satisfies the assumptions of AB p.153 Lem 4.51. Since B is a compact space of metrizable space, it is separable. Fix $g \in B$ and let $\psi^g(\omega) = \psi(\omega, g)$. Then, for any $V \in \mathcal{O}(\mathbb{R})$, $(\psi^g)^{-1}(V) = g^{-1}(V) \in \Sigma$. Thus, the function ψ^g is measurable. Fix $\omega \in \Omega$ and let $\psi_\omega : B \to \mathbb{R}$, $\psi_\omega(g) = \psi(\omega, g)$. For any $V \in \mathcal{O}(\mathbb{R})$, $(\psi_\omega)^{-1}(V) = \{g \in B \,|\, g(\omega) \in V\}$ and this is relative open set in B, and so ψ_ω is continuous. Apply the lemma mentioned above and conclude that ψ is $(\Sigma \otimes \mathcal{B}(B), \mathcal{B}(B))$-measurable. $\dashv$

Define a menu $B_\mathcal{F} = B - h$. Note that $c(B_\mathcal{F}, \omega) = c(B, \omega) - h$ and thus $P(B_\mathcal{F}) = \mathcal{F}$. Define a plan $\{f_\omega\}_{\omega \in \Omega}$ by $f_\omega = g_\omega - h$, and then the following hold: $f_\omega \in B_\mathcal{F}$, $f_\omega(\omega) = 0$, and $f_\omega \in c(B_\mathcal{F}, \omega)$. $\square$

For $F \in \mathbb{H}^*$, let $F^* \in \mathcal{A}$ be an act defined by $F^*(\omega) = F(\omega)(\omega)$. Recall that $\sigma(F)$ is the σ-algebra of Ω generated by F.

Lemma 8. *For any $F \in \mathbb{H}^*$, there exists $B^F \in \mathcal{K}$ and a plan $\overline{F}$ such that*

$$\overline{F} \in \hat{c}(B^F),\ \overline{F}(\omega)(\omega) = F(\omega)(\omega),\ \text{and } P(B^F) = \sigma(F).$$

Proof. Define $B^F := B_{\sigma(F)} + F^*$, where $B_{\sigma(F)}$ is the menu constructed applying Lemma 7 to $\sigma(F)$. There is a plan $\{f_\omega\}_{\omega \in \Omega} \in \hat{c}(B^F)$ such that $f_\omega(\omega) = 0$. Then, $\overline{F}(\omega)(\omega) = F(\omega)(\omega)$ and $\overline{F} \in \hat{c}(B^F)$. $\square$

For $F \in \mathbb{H}^*$, there may well be multiple plans $\overline{F}$ with the properties above. We denote $\overline{F}$ for one of them. The non-uniqueness does not cause any problem.

For $f, g \in \mathcal{A}$, let $f \wedge g \in \mathcal{A}$ be the act defined by $(f \wedge g)(\omega) = \min\{f(\omega), g(\omega)\}$. We denote the closure of any set $S \subset \mathcal{A}$ as $\mathrm{cl}S$.

Lemma 9. *For $A, B \in \mathcal{K}$, $f_\omega \in c(A, \omega)$, and $g_\omega \in c(B, \omega)$, the set*

$$C = \mathrm{cl}\{f_\omega \wedge g_\omega\}_{\omega \in \Omega}$$

is compact.

Proof. Take a sequence $\{h_n\}_{n=1}^{\infty}$ of C. Then, there exist $f_n \in \{f_\omega\}_{\omega\in\Omega}$, $g_n \in \{g_\omega\}_{\omega\in\Omega}$, and $h'_n \in \mathcal{A}$ such that $\|h'_n\| < n^{-1}$ and $h_n = (f_n \wedge g_n) + h'_n$.

Note the compactness of A and B and, passing to a subsequence, wlog assume

$$f_n \to f \in \mathrm{cl}\{f_\omega\}_{\omega\in\Omega} \text{ and } g_n \to g \in \mathrm{cl}\{g_\omega\}_{\omega\in\Omega}.$$

Note that the function $\wedge$ is jointly continuous. Thus $f \wedge g \in C$ and $h_n \to f \wedge g$. □

3.4.5 Preference relation between plans

Let us reveal the preference relation between plans in $\mathbb{H}^*$. We start with constructing preference relations between menus. Recall that

$$A \succeq^D B \Leftrightarrow \text{For all } \omega \in \Omega,\ A \cap c(B,\omega) \neq \emptyset.$$

and

$$A \succeq^I B \Leftrightarrow \text{There exists } n \in \mathbb{N} \text{ and } B_1, \ldots B_n \text{ such that } A \succeq^D B_1 \succeq^D \cdots \succeq^D B_n \succeq^D B.$$

The next lemma states that the translation invariance of c inherits to the relations $\succeq^D$, $\succeq^I$. For $A \in \mathcal{K}$, $f \in \mathcal{A}$, let $A + f = \{g + f \mid g \in A\}$. For $F \in \mathbb{H}^*$ and $f \in \mathcal{A}$, define a plan $F + f$ by $(F + f)(\omega) = F(\omega) + f$.

Lemma 10.
1. If $A \succeq^D B$, then $(A + f) \succeq^D (B + f)$.
2. If $A \succeq^I B$, then $(A + f) \succeq^I (B + f)$.

Proof. Suppose $A \succeq^D B$ and take $F \in \hat{c}(B)$ such that $\mathrm{Im}\, F \subset A$. Because $\hat{c}(B+f) = \hat{c}(B)+f$, $F+f \in \hat{c}(B+f)$. Then, $\mathrm{Im}\,(F+f) = (\mathrm{Im}\, F)+f \subset A+f$. Completed the first part.

Suppose $A \succeq^D C_1 \succeq^D \cdots \succeq^D C_n \succeq^D B$. Then, using the first part, $(A+f) \succeq^D (C_1 + f) \succeq^D \cdots \succeq^D (C_n + f) \succeq^D (B + f)$. Completed the second part. □

Lemma 11. *Consider $A, B \in \mathcal{K}$ and plans $\{f_\omega\}_{\omega\in\Omega} \in \hat{c}(A)$, $\{g_\omega\}_{\omega\in\Omega} \in \hat{c}(B)$. If $f_\omega(\omega) = g_\omega(\omega)$ for all $\omega \in \Omega$ and $P(A) = P(B)$, then $A \succeq^I B \succeq^I A$.*

Proof. The set $C = \mathrm{cl}\{f_\omega \wedge g_\omega\}_{\omega\in\Omega}$ is compact by Lemma 9. Write $h_\omega = f_\omega \wedge g_\omega$. For any $h \in C$, there exists $f \in A$ such that $f \geq h$, and so *INRA* implies

$$\forall\omega \in \Omega,\ A \cap c(A \cup C, \omega) = c(A, \omega).$$

Claim. $P(A) \subset P(A \cup C)$.

$\vdash$ The function $\varphi : \Omega \to \mathcal{K}, \varphi(\omega) = c(A \cup C, \omega)$ is $(P(A \cup C), \mathcal{B}(\mathcal{K}))$-measurable. And $\psi : \mathcal{K} \to \mathcal{K},\ \psi(D) = A \cap D$ is continuous. Then, the composition $\psi \circ \varphi$ is $(P(A \cup C), \mathcal{B}(\mathcal{K}))$-measurable. For any $V \in \mathcal{O}(\mathcal{K})$,

$$\{\omega | c(A, \omega) \in V\} = \{\omega | A \cap c(A \cup C, \omega) \in V\} = (\psi \circ \varphi)^{-1}(V) \in P(A \cup C).$$

This means $P(A) \subset P(A \cup C)$. $\dashv$

The inclusion $P(A) \subset P(A \cup C)$ implies $P(A \cup C)(\omega) \subset P(A)(\omega)$ for any ω. Fix some $\omega \in \Omega$. Then, $\forall \omega' \in P(A \cup C)(\omega),\ f_\omega(\omega') = f_{\omega'}(\omega') = h_{\omega'}(\omega') = h_\omega(\omega')$. This implies $h_\omega \in c(A \cup C, \omega)$ and $\{h_\omega\}_{\omega \in \Omega} \in \hat{c}(A \cup C)$. This and *INRA* implies $\{h_\omega\}_{\omega \in \Omega} \in \hat{c}(C)$. As a result

$$C \succeq^D (A \cup C) \succeq^D A \succeq^D (A \cup C) \succeq^D C$$

, that is, $A \succeq^I C \succeq^I A$.

Apply the same argument to B and C and obtain $C \succeq^I B \succeq^I C$. Hence $A \succeq^I B \succeq^I A$. □

Lemma 12. *For $F, G \in \mathbb{H}^*$ and $B \in \mathcal{K}$, if $F \in \hat{c}(B)$ and $\operatorname{Im} G \subset B$, then $B^F \succeq^I B^G$.*

Proof. First, we shall show that, $\mathrm{cl}(\operatorname{Im} G) \succeq^I B^G$. Take $\overline{G} \in \hat{c}(B^G)$ such that $\overline{G}(\omega)(\omega) = G(\omega)(\omega)$. Let $C = \mathrm{cl}\{G(\omega) \wedge \overline{G}(\omega)\}_{\omega \in \Omega}$. Then, $\overline{G}(\omega), G(\omega) \wedge \overline{G}(\omega) \in c(B^G \cup C, \omega)$. Hence,

$$C \succeq^D (B^G \cup C) \succeq^D B^G.$$

On the other hand, $\mathrm{cl}(\operatorname{Im} G) \succeq^D \mathrm{cl}(\operatorname{Im} G) \cup C \succeq^D C$. Combine these and obtain $\mathrm{cl}(\operatorname{Im} G) \succeq^I B^G$.

Next, we show $B^F \succeq^I B$. Note $P(B^F) = \sigma(F) \subset P(B)$. Take $\overline{F} \in \hat{c}(B^F)$ that $F(\omega)(\omega) = \overline{F}(\omega)(\omega)$. Let $C = \mathrm{cl}\{F(\omega) \wedge \overline{F}(\omega)\}_{\omega \in \Omega}$ and note

$$B^F \succeq^D (B^F \cup C) \succeq^D C \succeq^D (B \cup C) \succeq^D B,$$

or $B^F \succeq^I B$.

The assumption $\mathrm{Im}G \subset B$ implies $B \succeq^D \mathrm{cl}(\mathrm{Im}G)$. Since

$$B^F \succeq^I B \succeq^D \mathrm{cl}(\mathrm{Im}G) \succeq^I B^G,$$

$B^F \succeq^I B^G$ holds. □

For $F, G \in \mathbb{H}^*$, define

$$F \succeq G \Leftrightarrow B^F \succeq^I B^G.$$

Lemma 13. *For $F, G \in \mathbb{H}^*$ and $f \in \mathcal{A}$,*

$$F \succeq G \text{ implies } F + f \succeq G + f.$$

Proof. First, we show $B^{(F+f)} \succeq^I (B^F + f)$. To this end, take $\overline{F} \in \hat{c}(B^F)$ such that $\overline{F}(\omega)(\omega) = F(\omega)(\omega)$. Let $H \in \hat{c}(B^{F+f})$ such that $H(\omega)(\omega) = (F+f)(\omega)(\omega) = \overline{F}(\omega)(\omega) + f(\omega)$.

Note that the translation invariance $c(B^F + f, \omega) = c(B^F, \omega) + f$ implies $P(B^F + f) = P(B^F) = \sigma(F)$. But on the other hand $P(B^{(F+f)}) = \sigma(F + f) = \sigma(F)$.

Then, $\overline{F} + f \in \hat{c}(B^F + f)$, $H \in \hat{c}(B^{(F+f)})$, $(\overline{F} + f)(\omega)(\omega) = H(\omega)(\omega)$, and $P(B^F + f) = P(B^{(F+f)})$. Apply Lemma 11 and obtain

$$B^{(F+f)} \succeq^I (B^F + f) \succeq^I B^{(F+f)}.$$

In the same way obtain

$$B^{(G+f)} \succeq^I (B^G + f) \succeq^I B^{(G+f)}.$$

From the assumption $F \succeq G$, $B^F \succeq^I B^G$ holds. The translation invariance of $\succeq^I$ implies $(B^F + f) \succeq^I (B^G + f)$. Combining this with the conclusion of the last paragraph completes the proof. $\square$

3.4.6 Representation

For $f, g \in \mathcal{A}$, write $f \succeq g$ if $\{f\} \succeq \{g\}$. For a function $f : \Omega \to \mathbb{R}$ and $\pi \in \Delta\Omega$, $\pi(f)$ denotes for the integration $\int_\Omega f \, \mathrm{d}\pi$ if it exists.

Lemma 14. *There is a countably additive probability π defined on Σ such that, for any $f, g \in \mathcal{A}$,*

$$f \succeq g \text{ implies } \int f d\pi \geq \int g d\pi.$$

Proof. Let $K = \mathrm{co}\{f \in \mathcal{A} | f \succeq 0\}$. The set K is a convex set with a non-empty interior. Besides, $0 \notin \mathrm{int} K$ holds. If not, for any $f \in \mathcal{A}$, there exists $n \in \mathbb{N}$ so that $\frac{1}{n} f \in K$ and so $\frac{1}{n} f \succeq 0$. Then, the translation invariance and the transitivity of

$\succeq$ implies $\frac{2}{n}f \succeq 0$. Repeat this operation and obtain $f \succeq 0$. In the same manner, $-f \succeq 0$ or $0 \succeq f$, and thus $f \sim 0$. Hence the relation $\succeq$ is degenerate. By virtue of Axiom U, there exists $x, y \in X$ such that $x \succ^R y$. Axiom CC implies that $x \succ y$ and obtain a contradiction.

Apply the Separation Theorem and take a continuous linear functional π defined on $\mathcal{A}$ such that $\forall f \in K^* \; \pi(f) \geq 0$ and $\forall f \in \mathrm{int} K^* \; \pi(f) > 0$. Note that the topological dual of $\mathcal{B}_b(\Sigma)$ is the space of finitely additive charge, i.e., $\pi(f)$ is the integration $\int f d\pi$ of f by some finitely additive charge π.

Because $f \in \mathcal{K}$ for any $f \geq 0$, the functional π is positive. Normalize it as $\pi(1) = 1$, then π is a probability.

Suppose $1 > \epsilon > 0$ and $E_n \downarrow \emptyset$. Then, Axiom C (1) implies that there exists $n \in \mathbb{N}$ such that $\epsilon \succeq 1E_n0$, or $\epsilon \geq \pi(E_n)$. So $\pi(E_n) \to 0$. That is, π is countably additive. $\square$

Let $f_{\mathcal{F}}$ denotes for the plan $\{f_\omega\}_{\omega \in \Omega}$ such that $f_\omega \in c(B^f_{\mathcal{F}}, \omega)$ and $f_\omega(\omega) = f(\omega)$.

Lemma 15. *There exists a function $\gamma^* : \mathbb{P}^* \to \overline{\mathbb{R}}$ such that, for any $F, G \in \mathbb{H}^*$,*

$$F \succeq G \Rightarrow V^*(F) \geq V^*(G),$$

where $V^* : \mathbb{H}^* \to \overline{\mathbb{R}}$ is defined by

$$V^*(F) = \int F^* d\pi - \gamma^*(\sigma(F)).$$

Proof. Let $\Sigma_0 = \{\emptyset, \Omega\}$ be the trivial σ-algebra. For $\mathcal{F}, \mathcal{G} \in \mathbb{P}^*$, let

$$M_{\mathcal{F},\mathcal{G}} = \{f \in \mathcal{A} | f_{\mathcal{F}} \succeq 0_{\mathcal{G}}\},$$

and define $\gamma^* : \mathbb{P} \to \overline{\mathbb{R}}$ and $V^* : \mathbb{H}^* \to \overline{\mathbb{R}}$ by

$$\gamma^*(\mathcal{F}) = \inf_{f \in M_{\mathcal{F},\Sigma_0}} \int f d\pi,$$

$$V^*(F) = \int F^* d\pi - \gamma^*(\sigma(F)).$$

Claim. $\inf_{f \in M_{\mathcal{F},\mathcal{G}}} \int f d\pi \geq \gamma^*(\mathcal{F}) - \gamma^*(\mathcal{G})$.

$\vdash$ Let $g_n \in M_{\mathcal{G},\Sigma_0}$ such that $\pi(g_n) \to \gamma(\mathcal{G})$ and $h_n \in M_{\mathcal{F},\mathcal{G}}$ such that $\pi(h_n) \to \inf_{h \in M_{\mathcal{F},\mathcal{G}}} \pi(h)$. Because $[g_n]_{\mathcal{G}} \succeq 0_{\Sigma_0}$, $0_{\mathcal{G}} \succeq [-g_n]_{\Sigma_0}$. Combining with $[h_n]_{\mathcal{F}} \succeq 0_{\mathcal{G}}$, and because of invariance of $\succeq$, $[g_n + h_n]_{\mathcal{F}} \succeq 0_{\Sigma_0}$. Then,

$$\gamma^*(\mathcal{F}) = \inf_{f \in M_{\mathcal{F},\mathcal{G}}} \int f d\pi \leq \int g_n + h_n d\pi \to \gamma(\mathcal{G}) + \inf_{h \in M_{\mathcal{F},\mathcal{G}}} \int h d\pi.$$

Then, $\inf_{h\in M_{\mathcal{F},\mathcal{G}}} \int h d\pi \geq \gamma^*(\mathcal{F}) - \gamma^*(\mathcal{G})$. $\dashv$

Let $F, G \in \mathbb{H}^*$ $F \succeq G, \mathcal{F} = \sigma(\mathcal{F}),\ \mathcal{G} = \sigma(G)$. Then,

$$[F^*]_{\mathcal{F}} \sim F \succeq G \sim [G^*]_{\mathcal{G}}.$$

Substracting G^* from both hands, obtain $[F^* - G^*]_{\mathcal{F}} \succeq 0_{\mathcal{G}}$. From $F^* - G^* \in M_{\mathcal{F},\mathcal{G}}$ and the claim,

$$\int F^* - G^* d\pi \geq \gamma^*(\mathcal{F}) - \gamma^*(\mathcal{G}).$$

That is, $V^*(F) \geq V^*(G)$. □

Lemma 16. *For $\mathcal{F}, \mathcal{G} \in \mathbb{P}^*$, if $\mathcal{F} \subset \mathcal{G}$, then $\gamma^*(\mathcal{F}) \leq \gamma^*(\mathcal{G})$.*

Proof. Consider $B_{\mathcal{F}}$ and $B_{\mathcal{G}}$ and take $F \in \hat{c}(B_{\mathcal{F}})$ and $G \in \hat{c}(B_{\mathcal{G}})$, such that $F(\omega)(\omega) = G(\omega)(\omega) = 0$. Let $C = \mathrm{cl}\{F(\omega) \wedge G(\omega)\}_{\omega\in\Omega}$. Then, $F \in \hat{c}(B_{\mathcal{F}} \cup C)$ and so $B_{\mathcal{F}} \succeq^I C$. On the other hand, $G \in \hat{c}(B_{\mathcal{G}} \cup C)$.

For the use in the proof of a claim, let $\mathcal{B}(\mathcal{A}) \otimes \mathcal{B}(\mathcal{A}) = \sigma(\{B \times B' \mid B, B' \in \mathcal{B}(\mathcal{A})\})$.

Claim. The plan $H(\omega) = F(\omega) \wedge G(\omega)$ is $\mathcal{G}$-measurable.

$\vdash$ Let $\varphi : \Omega \to \mathcal{A} \times \mathcal{A}$, $\varphi(\omega) = (F(\omega), G(\omega))$. This is $(\mathcal{G}, \mathcal{B}(\mathcal{A}) \otimes \mathcal{B}(\mathcal{A}))$-measurable because of measurability of F and G. Next let $\psi : \mathcal{A} \times \mathcal{A} \to \mathcal{A},\ \psi(f, g) = f \wedge g$. Because ψ is continuous, it is $(\mathcal{B}(\mathcal{A}) \otimes \mathcal{B}(\mathcal{A}), \mathcal{B}(\mathcal{A}))$-measurable. Note $H = \psi \circ \varphi$ and complete the proof. $\dashv$

Note that

$$c(B_{\mathcal{G}} \cup C, \omega) \cap B_{\mathcal{G}} = c(B_{\mathcal{G}}, \omega)$$

implies $P(B_{\mathcal{G}} \cup C) \supset P(B_{\mathcal{G}})$. Then, because

$$P(B_{\mathcal{G}} \cup C)(\omega) \subset \mathcal{G}(\omega) \subset \mathcal{F}(\omega),$$

for any $\omega' \in P(B_{\mathcal{G}} \cup C)(\omega)$,

$$F(\omega)(\omega') \wedge G(\omega)(\omega') = F(\omega')(\omega') \wedge G(\omega')(\omega') = 0.$$

Then, $F(\omega) \wedge G(\omega) \in c(B_{\mathcal{G}} \cup C, \omega)$, that is, $C \succeq^D B_{\mathcal{G}} \cup C$. Combining this with $B_{\mathcal{F}} \succeq^I C$, conclude $B_{\mathcal{F}} \succeq^I B_{\mathcal{G}}$. This implies $F \succeq G$ and, in terms of the representation, $-\gamma^*(\mathcal{F}) \geq -\gamma^*(\mathcal{G})$. □

Next, we extend the representation V^* to $\mathbb{H}$. To this end, we need a preliminary lemma. For $\mathcal{F} \in \mathbb{P}$, let $\mathbb{P}^*(\mathcal{F}) = \{\mathcal{G} \in \mathbb{P}^* | \ \mathcal{F} \subset \mathcal{G} \subset \Sigma\}$. For $\mathcal{A}$ and $\mathcal{F} \in \mathbb{P}^*$, define $B^f_{\mathcal{F}} = B_{\mathcal{F}} + f$.

Lemma 17. *For any $F, G \in \mathbb{H}$ such that $\mathrm{Im} G \subset B$, if $F \in \hat{c}(B)$ for some $B \in K$ and $P(B) = \mathcal{F}$, then*

$$B^{F^*}_{\mathcal{F}} \succeq^I B^{G^*}_{\mathcal{G}} \textit{ for any } \mathcal{G} \in \mathbb{P}^*(\sigma(G)).$$

Proof. Take any $\tilde{F}, \tilde{G} \in \mathbb{H}^*$

$$\begin{aligned} \tilde{F}(\omega)(\omega) = F(\omega)(\omega), \ \sigma(\tilde{F}) = \mathcal{F}, \\ \tilde{G}(\omega)(\omega) = G(\omega)(\omega), \ \sigma(\tilde{G}) = \mathcal{G}. \end{aligned}$$

We show $B^{\tilde{F}} \succeq^I B^{\tilde{G}}$.

It is easy to see $B^{\tilde{F}} \succeq^I B$. To prove $B \succeq^I B^{\tilde{G}}$, let $g_\omega = G(\omega) \wedge \tilde{G}(\omega)$ and $C = \mathrm{cl}\{g_\omega\}_{\omega \in \Omega}$. Because of the definition of $\tilde{G}$, $\{g_\omega\}_{\omega \in \Omega}$ is a measurable plan. For any $\omega' \in \mathcal{G}(\omega)$.

$$g_\omega(\omega') = G(\omega)(\omega') \wedge \tilde{G}(\omega)(\omega') = G(\omega')(\omega') \wedge \tilde{G}(\omega')(\omega') = \tilde{G}(\omega')(\omega'). \tag{3.3}$$

So $C \succeq^D (B^{\tilde{G}} \cup C) \succeq^D B^{\tilde{G}}$.

On the otherhand, $\tilde{G} \in \hat{c}(B^{\tilde{G}})$, *INRA*, and (3.3) implies $g_\omega \in c(C, \omega)$. Hence $B \succeq^D (B \cup C) \succeq^D C$. Combining everything, complete the proof. □

Lemma 18. *If $F \in \hat{c}(B)$ and $\mathrm{Im}\, G \subset B$, $V(F) \geq V(G)$, where $V : \mathbb{H} \to \mathbb{R}$ is defined by*

$$\begin{aligned} V(F) &= \int F^* d\pi - \gamma(\sigma(F)), \\ \gamma(\sigma(F)) &= \inf_{\tilde{\mathcal{F}} \in \mathbb{P}^*(\sigma(F))} \gamma^*(\tilde{\mathcal{F}}). \end{aligned}$$

Proof. Let $\mathcal{F} = P(B)$ and take any $\mathcal{G} \in \mathbb{P}^*(\sigma(G))$. Then, $B^{F^*}_{\mathcal{F}} \succeq^I B^{G^*}_{\mathcal{G}}$.

Let $\tilde{F} \in \hat{c}(B^{F^*}_{\mathcal{F}})$, $\tilde{F}^* = F^*$, $\sigma(\tilde{F}) = \mathcal{F}$, and $\tilde{G} \in \hat{c}(B^{G^*}_{\mathcal{G}})$, $\tilde{G}^* = G^*$, $\sigma(\tilde{G}) = \mathcal{G}$. Then, $V(F) \geq V^*(\tilde{F}) \geq V^*(\tilde{G})$. Take the supremum of the most right-hand side, conclude $V(F) \geq V(G)$. □

Lemma 19. *If $F \in \hat{c}(B)$, then $\gamma(\sigma(F)) = \gamma(P(B))$.*

Proof. Because $\sigma(F) \subset P(B)$, it is sufficient to show $\gamma(\sigma(F)) \geq \gamma(P(B))$. To this end, take any $\mathcal{G} \in \mathbb{P}^*(\sigma(F))$ and consider $B_{\mathcal{G}}^{F^*}$. Write $\mathcal{F} = P(B)$. Take $G \in \hat{c}(B_{\mathcal{G}}^{F^*})$ such that $G^*(\omega) = F^*(\omega)$, and let $C = \mathrm{cl}\{F(\omega) \wedge G(\omega)\}_{\omega \in \Omega}$.

Then, $G \in \hat{c}(B_{\mathcal{G}}^{F^*} \cup C)$. Let $H(\omega) = F(\omega) \wedge G(\omega)$ and note that $H \in \hat{c}(B_{\mathcal{G}}^{F^*} \cup C)$. That is, $C \succeq^I B_{\mathcal{G}}^{F^*}$. Because $B_{\mathcal{F}}^{F^*} \succeq^D B \succeq^D B \cup C \succeq^D C \succeq^D B_{\mathcal{G}}^{F^*}$, $B_{\mathcal{F}}^{F^*} \succeq^I B_{\mathcal{G}}^{F^*}$, or $[F^*]_{\mathcal{F}} \succeq [F^*]_{\mathcal{G}}$. Then

$$E[F^*] - \gamma(\mathcal{F}) = V([F^*]_{\mathcal{F}}) \geq V([F^*]_{\mathcal{G}}) = E[F^*] - \gamma(\mathcal{G}),$$

that is, $\gamma(\mathcal{F}) \leq \gamma(\mathcal{G})$. Remind that σ-algebra $\mathcal{G} \in \mathbb{P}^*(\sigma(F))$ was taken arbitrary and conclude $\gamma(\mathcal{F}) \leq \gamma(\sigma(F))$. □

Because π is countably additive, the conditional expectation of any act is well-defined.

Lemma 20. *If* $\mathrm{Im}\, F \subset B$ *and* $\sigma(F) \subset \mathcal{F}$*, then*

$$E[F^*] \leq E\left[\max_{f \in B} E[f|\mathcal{F}]\right].$$

Proof. It is sufficient to show that, for any $\Delta \in \mathcal{F}$,

$$\int_\Delta F^* d\pi \leq \int_\Delta \max_{f \in B} E[f|\mathcal{F}] d\pi.$$

Take any $\Delta \in \mathcal{F}$ and choose $\epsilon > 0$. Then, by virtue of the compactness of B, there exist $f_1, \ldots, f_n \in \mathrm{cl}(\mathrm{Im}\, F)$ such that $\mathrm{cl}(\mathrm{Im}\, F) \subset \bigcup_{i=1}^n B_\epsilon(f_i)$. Define $\Delta_1, \ldots, \Delta_n$ by

$$\Delta_1 = \{\omega | F(\omega) \in B_\epsilon(f_1)\},$$
$$\Delta_{k+1} = \{\omega | F(\omega) \in B_\epsilon(f_{k+1})\} \setminus \bigcup_{i=1}^k \Delta_i.$$

Then, $\Delta = \sum_{i=1}^n \Delta_i$ and $\Delta_i \in \mathcal{F}$.

$$\begin{aligned}\int_\Delta F^* d\pi &= \sum_{i=1}^n \int_{\Delta_i} F^* d\pi \leq \sum_{i=1}^n \int_{\Delta_i} E[f_i + \epsilon|\mathcal{F}] d\pi \\ &= \epsilon + \sum_{i=1}^n \int_{\Delta_i} \max_{f \in B} E[f|\mathcal{F}] d\pi = \epsilon + \int_\Delta \max_{f \in B} E[f|\mathcal{F}] d\pi.\end{aligned}$$

Because ϵ is arbitrary, $\int_\Delta F^* d\pi \leq \int_\Delta \max_{f \in B} E[f|\mathcal{F}] d\pi$. □

Lemma 21. *$P(B) \in \arg\max_{\mathcal{F}\in\mathbb{P}} E[\max_{f\in B} E[f|\mathcal{F}]] - \gamma(\mathcal{F})$.*

Proof. Take any $\mathcal{G} \in \mathbb{P}$. Define a function $\nu : \Omega \times B \to \mathbb{R}$ by $\nu(\omega, f) = E[f|\mathcal{G}](\omega)$. This is a $(\mathcal{G} \otimes \mathcal{B}(B), \mathcal{B}(\mathbb{R}))$-measurable Caratheodory function. Then, Apply AB p.605 Theorem 18.19 (3) and take a $\mathcal{G}$-measurable plan G such that

$$E[G^*] = E\left[E[G^*|\mathcal{G}]\right] = E\left[\max_{f\in B} E[f|\mathcal{G}]\right].$$

Apply Lemma 18 and obtain

$$V(G) = E[G^*] - \gamma(\sigma(G)) \geq E\left[\max_{f\in B} E[f|\mathcal{G}]\right] - \gamma(\mathcal{G}).$$

Take any $F \in \hat{c}(B)$. Then, Lemma 20 and Lemma 19 imply

$$E\left[\max_{f\in B} E[f|P(B)]\right] - \gamma(P(B)) \geq V(F).$$

Then, combining with two equalities above, Lemma 18 implies

$$E\left[\max_{f\in B} E[f|P(B)]\right] - \gamma(P(B)) \geq V(F) \geq V(G) \geq E\left[\max_{f\in B} E[f|\mathcal{G}]\right] - \gamma(\mathcal{G}).$$

$\square$

Lemma 22. *$c(B, \omega) \subset \arg\max_{f\in B} E[f|P(B)](\omega)$ π-a.s.*

Proof. En route to a contradiction, suppose there exists an event $S_0 \in \Sigma$ such that $\pi(S_0) > 0$ and for any $\omega \in S_0$,

$$c(B, \omega) \not\subset \arg\max_{f\in B} E[f|P(B)](\omega).$$

Write $c^*(\omega) = \arg\max_{f\in B} E[f|P(B)](\omega)$.

We shall show that there is an $\epsilon > 0$ and $S \in \Sigma$ such that $\pi(S) > 0$ and

$$c(\omega)\backslash(c^*(\omega) + N_\epsilon(0)) \neq \emptyset$$

for any $\omega \in S$. If not, for any $n \in \mathbb{N}$,

$$c(\omega) \subset c^*(\omega) + N_{\frac{1}{n}}(0) \ \pi\text{-a.s.} \tag{3.4}$$

By the assumption, there are $\omega \in S_0$ and $f \in B$ such that $f \in c(\omega)\backslash c^*(\omega)$. Because c^* is compact valued, for sufficiently large n, $f \notin c^*(\omega) + N_{\frac{1}{n}}(0)$. This is a contradiction to (3.4).

Take an $\epsilon > 0$ and $S \in \Sigma$ such that $\pi(S) > 0$ and

$$c(\omega)\backslash(c^*(\omega) + N_\epsilon(0)) \neq \emptyset.$$

Let $\varphi'(\omega) = c(\omega)\backslash(c^*(\omega) + N_\epsilon(0))$. For $V \in \mathcal{O}(\mathcal{K})$,

$$c^*(\omega) + N_\epsilon(0) \in V \Leftrightarrow c^*(\omega) \in V - N_\epsilon(0) = \bigcup_{f \in N_\epsilon(0)} (V - f)$$

and note the last set is open. Thus $c^*(\omega) + N_\epsilon(0)$ is an open valued measurable correspondence. So φ' is also measurable by virtue of AB Lemma 18.4 (3). Next, let

$$\varphi(\omega) = \begin{cases} \varphi'(\omega) & \omega \in S \\ c(\omega) & \omega \notin S. \end{cases}$$

For any $V \in \mathcal{O}(\mathcal{K})$,

$$\varphi^{-1}(V) = (S \cap (\varphi')^{-1}(V)) \cup (S^c \cap c^{-1}(V)) \in \Sigma$$

and so φ is measurable.

Take a measurable selector F of φ. Note that $F \in \hat{c}(B)$. On the otherhand, there exists a plan G such that $\mathrm{Im} G \subset B$ and

$$V(G) \geq E[\max_{f \in B} E[f|P(B)]] - \gamma(P(B)).$$

By the construction of F,

$$V(F) = E[F^*] - \gamma(P(B)) < E[\max_{f \in B} E[f|P(B)]] - \gamma(P(B)) \leq V(G).$$

This contradicts Lemma 18. $\square$

The proof of the main result is completed.

Chapter 4

Second-order beliefs and unambiguous events

4.1 Introduction

Since Ellsberg (1961) found a class of uncertainty, named ambiguity, it has been one of the central subjects of decision theory. One of Ellsberg's thought experiments is as follows: There is an urn containing 90 balls, each of which is colored red, green, or blue. Thirty balls are known to be red, but there is no information about the ratio of green and blue balls. One ball will be taken out from the urn. A decision-maker (henceforth DM) is confronted with four bets on its color. Bet 1 gives her 100 dollars if the ball is red and nothing otherwise. Bet 2 gives 100 dollars if the ball is green. Bet 3 gives 100 dollars if the ball is red or blue. Bet 4 gives 100 dollars if the ball is green or blue. DM is asked to choose between bet 1 and 2, and then between 3 and 4. It is well known that typical subjects prefer to bet 1 over 2 and bet 4 over 3. This behavior contradicts any models based on a probabilistic belief. The class of uncertainty that cannot be represented by probabilities is called ambiguity.

An approach to analyzing ambiguity is to provide models that capture the above-mentioned Ellsberg type behavior. Since the pioneering works of Schmeidler (1989) and Gilboa and Schmeidler (1989), many alternative models have been provided. Another stream is to define ambiguity by distinguishing events to which DM assigns probabilistic likelihood. Early works in this line are Zhang (2002) and Epstein and Zhang (2001). The present chapter contributes to the latter literature by providing a set of axioms that characterize a special case of second-order subjective expected utility (SOSEU) model of Seo (2009). Under the obtained model, exogenously given events can be interpreted as unambiguous events. A similar result holds for a more general model, second-order maxmin expected utility (SOMEU) of Nascimento and Riella (2013).

SOSEU consists of three parameters: an expected utility function u, a strictly increasing real function v, and a probability measure on the set of probability measures on the state space S, $m \in \Delta(\Delta S)$. Second-order probability m represents the ignorance of DM about probabilistic law. Write an ambiguous act that represents a bet on subjective uncertainty as $f : S \to \Delta X$, where ΔX is the set of lotteries over outcomes. A utility function of SOSEU is written as

$$V(f) = \int_{\Delta S} v\left(\int_S u(f(s))d\mu(s)\right) dm(\mu),$$

and it captures the Ellsberg type behavior. The idea of SOSEU appears in Savage (1954) and the model is axiomatized by Klibanoff et al. (2005) and Seo (2009) in different settings.

Under SOSEU, it seems natural to define unambiguous events as events such that the distributions of the probability assigned to them, induced by second-order belief, degenerate.[1] In this chapter, I assume that we know which events are such unambiguous events. That is, unambiguous events are exogenously given. The main theorem in this chapter shows that the axioms I introduce later characterize the preference relations that are represented by SOSEU representation under which exogenous events are unambiguous in terms of the model.

Many authors have written on probabilistic beliefs on endogenous and exogenous events. Zhang (2002) and Epstein and Zhang (2001) provide different definitions of unambiguous events based on the intuition of the Sure-Thing Principle of Savage (1954) and P4* of Machina and Schmeidler (1992), respectively. Kopylov (2007) refines their arguments. Sarin and Wakker (1992) axiomatizes Choquet expected utility whose capacity is additive on exogenously given events. Qu (2013) axiomatizes maximin expected utility (MEU) whose multiple priors degenerate on exogenously given events and defines unambiguous events under MEU. The papers most related to the present one are Klibanoff et al. (2005) and

[1] This is a model-based definition and some authors argue that the concept of ambiguity should be formalized without referring to any particular model (Epstein and Zhang (2001)). However, it should not be dismissed immediately. Some economists believe that intuitive stories of models make us more comfortable to rely on predictions of them, and they enhance the modeler's reasoning process (Dekel and Lipman (2010)). In this point of view, it is worthwhile to first provide a definition of unambiguous events based on well-established models such as SOSEU and then investigate its behavioral implications.

Klibanoff et al. (2011). Klibanoff et al. (2005) provided an alternative foundation of SOSEU (smooth ambiguity model in their terminology) and proposed a definition of unambiguous events, which works under their model. Klibanoff et al. (2011) compares their definition with earlier ones.

In section 2, the setup is presented, and the main results are described. In section 3, I compare the Theorem 1 in this chapter with the characterization of unambiguous events in Klibanoff et al. (2005), and discuss the possibility of future research. All proofs are collected in the appendix.

4.2 Model and results

4.2.1 Setup

I adopt the original framework of Anscombe and Aumann (1963) as in Seo (2009). Let X denote the set of prizes and assume it is a separable metric space. For any topological space Y, denote ΔY the set of all Borel probability measures on Y. Any set of probability measures is endowed with the weak topology. Let S denote the set of states and assume it is finite. A subset of S is called an event. A function from S to ΔX is called an act. Let $\mathcal{F}$ denote the set of acts. $\mathcal{F}$ is endowed with product topology. The choice set of DM is the set of probability measures on $\mathcal{F}$, $\Delta\mathcal{F}$. Prizes are denoted by x, y, z. Elements of ΔX are denoted by p, q, r. Acts are denoted by f, g, h. Elements of $\Delta\mathcal{F}$ are denoted by P, Q, R.

The sets X and $\mathcal{F}$ can be seen as subsets of ΔX and $\Delta\mathcal{F}$ respectively in an obvious manner. An element of ΔX can be identified with a constant act. So $X \subset \Delta X \subset \mathcal{F} \subset \Delta\mathcal{F}$. For any family of events $\mathcal{A}$, let $\mathcal{F}_{\mathcal{A}}$ denote the set of acts that are measurable with respect to $\mathcal{A}$.

DM's choice behavior is modeled as a binary relation $\succeq$ on $\Delta\mathcal{F}$. The relations $\succ$ and $\sim$ denote the asymmetric and symmetric part of $\succeq$.

Let $\mathbb{U} = \{\mathcal{U} | \mathcal{U} \subset 2^S, S \in \mathcal{U}, A \in \mathcal{U} \Rightarrow A^c \in \mathcal{U}\}$. As shown below, under suitable axioms, $\mathcal{U} \in \mathbb{U}$ can be interpreted as a set of unambiguous events. For example, in Ellsberg's experiment above, DM knows that the probability a red ball is taken out is one third. So R, the event that a red ball is taken out, is an unambiguous event. One can take $\mathcal{U} = \{S, \emptyset, R, R^c\}$ as the set of her unambiguous events.

4.2.2 Unambiguous events under SOSEU

I list the axioms of Seo (2009) that characterize SOSEU representations. The first two are common in the literature.

Axiom 8 (Order). *$\succeq$ is complete, transitive.*

Axiom 9 (Continuity). *$\succeq$ is continuous.*

A mixing operation on $\mathcal{F}$ is defined as $(\alpha f \oplus (1-\alpha)g)(s) = \alpha f(s) + (1-\alpha)g(s)$ for any $f, g \in \mathcal{F},\ s \in S$. Then, $\mathcal{F}$ is a mixture space under this operation. The next axiom requires preferences to satisfy independence on the set of one-stage lotteries.

Axiom 10 (Second-Stage Independence). *For any $\alpha \in (0,1]$ and one stage lotteries $p, q, r \in \Delta X$, $\alpha p \oplus (1-\alpha)r \succeq \alpha q \oplus (1-\alpha)r \Leftrightarrow p \succeq q$.*

For $P, Q \in \Delta\mathcal{F}$ and $\alpha \in [0,1]$, $\alpha P + (1-\alpha)Q \in \Delta\mathcal{F}$ is a lottery such that $(\alpha P + (1-\alpha)Q)(B) = \alpha P(B) + (1-\alpha)Q(B)$ for any Borel set $B \subset \mathcal{F}$. The next axiom requires preferences to satisfy independence on the set of lotteries of acts.

Axiom 11 (First-Stage Independence). *For any $\alpha \in (0,1]$ and probabilities $P, Q, R \in \Delta\mathcal{F}$, $\alpha P(1-\alpha)R \succeq \alpha Q + (1-\alpha)R \Leftrightarrow P \succeq Q$.*

For $f \in \mathcal{F}$ and $\mu \in \Delta S$, define $\Psi(f,\mu) = \mu(s_1)f(s_1) \oplus \cdots \oplus \mu(s_{|S|})f(s_{|S|}) \in \Delta X$. If DM has a probabilistic belief μ, she identifies an act f with the one-stage lottery $\Psi(f,\mu)$. For $P \in \Delta\mathcal{F}$ and $\mu \in \Delta S$, define $\Psi(P,\mu) \in \Delta(\Delta X)$ by $\Psi(P,\mu)(B) = P(\{f \in \mathcal{F} | \Psi(f,\mu) \in B\})$. This two-stage lottery gives some first-stage lottery p with the probability that the original lottery P assigns to the acts that are identified with p under μ. The next axiom requires that if a second-order lottery P is preferred to Q no matter which first-order subjective probability $\mu \in \Delta S$ is true, then P is indeed preferred to Q.

Axiom 12 (Dominance). *For any $P, Q \in \Delta\mathcal{F}$, if $\Psi(P,\mu) \succeq \Psi(Q,\mu)$ for all $\mu \in \Delta S$, then $P \succeq Q$.*

Seo (2009) shows that these behavioral regularities characterize the following utility representation.

Definition 1. *A tuple (u, v, m) is a second-order subjective expected utility (SOSEU) representation of $\succeq$ if $u : \Delta X \to \mathbb{R}$ is bounded continuous and mixture linear,*

$v : u(\Delta X) \to \mathbb{R}$ is bounded continuous and strictly increasing, $m \in \Delta(\Delta S)$ and V represents $\succeq$, where

$$V(P) = \int_{\mathcal{F}} U(f) dP(f),$$
$$U(f) = \int_{\Delta S} v\left(\int_S u(f) d\mu\right) dm(\mu).$$

An SOSEU representation (u, v, m) is nondegenerate if u is not constant.

Under SOSEU, it is natural to define unambiguous events as those which the distributions of probability DM assigns to them degenerate. The formal definition is as follows.

Definition 2. *An event is unambiguous under $m \in \Delta(\Delta S)$ if there exists $\alpha \in [0, 1]$ such that $\mu(E) = \alpha$ m-a.s.. A family $\mathcal{U} \in \mathbb{U}$ is unambiguous under m if each $E \in \mathcal{U}$ is unambiguous under m.*

To guarantee the unambguousness of $\mathcal{U}$, I introduce a new axiom. For this purpose, define $\Delta_{\mathcal{U}} = \{\mu \in \Delta S | \forall f \in \mathcal{F}_{\mathcal{U}}\ f \sim \Psi(f, \mu)\}$. This is the set of first-order probabilities such that DM is indifferent between any acts measurable with respect to $\mathcal{U}$ and the corresponding one-stage lottery constructed with it.

Axiom 13 ($\mathcal{U}$-Dominance)**.** *For $P, Q \in \Delta\mathcal{F}$, if $\Psi(P, \mu) \succeq \Psi(Q, \mu)$ for all $\mu \in \Delta_{\mathcal{U}}$ then $P \succeq Q$.*

$\mathcal{U}$-dominance is stronger than Dominance in that only first-order probabilities in $\Delta_{\mathcal{U}}$, a set smaller than ΔS, are relevant to DM. This can be interpreted as having stronger confidence in probabilistic law.

Finally, we put an auxiliary axiom to elicit probabilities.

Axiom 14 (Nondegenerate)**.** *$P \succ Q$ for some $P, Q \in \Delta\mathcal{F}$.*

This axiom states that DM is not indifferent between all alternatives.

Now the main result can be stated.

Theorem 2. *The followings are equivalent. The axioms above except Axiom 5 characterize preference relations that have SOSEU representation under which $\mathcal{U}$ is unambiguous in terms of the model.*

1. *$\succeq$ satisfies Axioms 8–11, 13, and 14.*
2. *$\succeq$ has a nondegenerate SOSEU representation (u, v, m) under which $\mathcal{U}$ is unambiguous.*

This theorem demonstrates that a preference satisfies the required axioms if and only if it has an SOSEU representation, and each event in $\mathcal{U}$ is unambiguous under the second-order belief. Hence DM behaves as if she is an SEU maximizer when choosing among unambiguous acts. If one imposes $\mathcal{U}$-dominance for larger $\mathcal{U}$, the model becomes closer to SEU. Hence the theorem reveals what lies between SEU and SOSEU.

The model includes nondegenerate SOSEU as a special case.

Corollary 1. *Suppose* $\mathcal{U} = \{\emptyset, S\}$. *Followings are equivalent.*

1. $\succeq$ *satisfies Order, Continuity, First-stage independence, Second-stage independence,* $\mathcal{U}$*-dominance, Nondegenerate.*
2. $\succeq$ *has a nondegenerate SOSEU representation.*

So far, we have treated the set of unambiguous events as exogenous ones. The next goal is to obtain the largest set of unambiguous events endogenously. We give a definition that describes a behavioral characterization of unambiguous events under SOSEU.

Definition 3. *An event* E *is said to be unambiguous if* $\succeq$ *satisfies* $\{E, E^c, S, \emptyset\}$*-Dominance.*

Then, the natural question to be investigated is to consider whether $\mathcal{U}_{\succeq}$ is the maximum set of unambiguous events in the sense that (1) there is an SOSEU under which any event in $\mathcal{U}_{\succeq}$ is unambiguous and (2) $\mathcal{U}_{\succeq}$ is the set of all the events such that there is an SOSEU representation under which the event is unambiguous. But this is still an open question.

4.2.3 Unambiguous events under SOMEU

In this subsection, I show a result similar to Theorem 2 for a more general model that appears Nascimento and Riella (2013), called second-order maxmin expected utility .

Definition 4. *A tuple* (u, v, M) *is a second-order maxim expected utility (SOMEU) representation of* $\succeq$ *if* $u : \Delta X \to \mathbb{R}$ *is bounded continuous and mixture linear,* $v : u(\Delta X) \to \mathbb{R}$ *is bounded continuous and strictly increasing,* $M \subset \Delta(\Delta S)$ *is a nonempty, closed, and convex set, and* V *represents* $\succeq$*, where*

$$V(P) = \min_{m \in M} \int_{\Delta S} \left[\int_{\mathcal{F}} v \left(\int_S u(f) d\mu \right) dP(f) \right] dm(\mu).$$

An SOMEU representation (u, v, M) *is nondegenerate if* u *is not constant.*

I list axioms that characterize SOMEU preferences. I adopt axioms different from Nascimento and Riella (2013) for simplicity. The following two axioms are counterparts of Uncertainty Aversion and Certainty Independence proposed in Gilboa and Schmeidler (1989), respectively. Because both are weaker than First-stage independence, SOMEU is a generalization of SOSEU.

Axiom 15 (Convexity)**.** *For all $P, Q \in \Delta\mathcal{F}$, and $\lambda \in (0,1)$, if $P \sim Q$ then $\lambda P + (1-\lambda)Q \succeq Q$.*

Axiom 16 (First-Stage Certainty Independence)**.** *For any $P, Q \in \Delta\mathcal{F}$, $\overline{P} \in \Delta(\Delta X)$, and $\alpha \in (0,1]$, $\alpha P + (1-\alpha)\overline{P} \succeq \alpha Q + (1-\alpha)\overline{P} \Leftrightarrow P \succeq Q$.*

The model-based definition of unambiguous events is extended to the case in which a belief is modeled as a set M of second-order probabilities. An event is unambiguous under a second-order belief M if all the probabilities it assign a common degenerate probability to the event.

Definition 5. *An event E is unambiguous under $M \subset \Delta(\Delta S)$ if there exists an $\alpha \in [0,1]$ such that $\mu(E) = \alpha$ m-a.s. for all $m \in M$.*

The following theorem is the counterpart of Theorem 1 adapted for SOMEU.

Theorem 3. *For any $\mathcal{U} \in \mathbb{U}$, followings are equivalent.*

1. *$\succeq$ satisfies Axioms 8–10 and 13–16.*
2. *$\succeq$ has a nondegenerate SOMEU representation (u, v, m) under which $\mathcal{U}$ is unambiguous.*

This theorem demonstrates that the same definition of unambiguous events works as well under SOMEU.

4.2.4 Unambiguous events under second-order Bewley representation

In the decision-theoretic literature, preferences under ambiguity are sometimes described as an incomplete preference relation. Bewley representation is a utility representation for such incomplete relation for acts. Under the model, DM prefers an act to another one if the former is better than the latter in terms of all probability he considers. Its second-order counterpart, Nascimento and Riella (2013) consider is second-order Bewley representation, which incorporates a set of second-order probabilities as a representation of belief. Here, I consider a special case of their model under which some events are unambiguous.

Definition 6. *A tuple (u, v, M) is a second-order Bewley representation of $\succeq$ if $u : \Delta X \to \mathbb{R}$ is bounded continuous and mixture linear, $v : u(\Delta X) \to \mathbb{R}$ is bounded continuous and strictly increasing, $M \subset \Delta(\Delta S)$ and $P \succeq Q$ if and only if*

$$\int_{\Delta S} \left[\int_{\mathcal{F}} v \left(\int_S u(f) d\mu \right) dP(f) \right] dm(\mu) \geq \int_{\Delta S} \left[\int_{\mathcal{F}} v \left(\int_S u(f) d\mu \right) dQ(f) \right] dm(\mu)$$

for all $m \in M$. A second-order Bewley representation (u, v, M) is nondegenerate if u is not constant.

Nascimento and Riella (2013) axiomatized second-order Bewley representation postulating the following axioms.

Axiom 17 (Preference Relation)**.** *The binary relation $\succeq$ is reflexive and transitive.*

Let $\succeq^\bullet = \succeq|_{\Delta(\Delta X)}$ be a binary relation over $\Delta(\Delta X)$.

Axiom 18 (Partial Completeness)**.** *The binary relation $\succeq^\bullet$ is complete.*

This axiom states that DM can always decide which one of two second-order probabilities is better.

The definition of unambiguous events for second-order Bewley representation is the same as Definition 5. Now I characterize second-order Bewley representation under which $\mathcal{U}$ is unambiguous.

Theorem 4. *For any $\mathcal{U} \in \mathbb{U}$, followings are equivalent.*
1. *$\succeq$ satisfies Axioms 9–12, 14,17, and 18.*
2. *$\succeq$ has a nondegenerate second-order Bewley representation (u, v, M) under which $\mathcal{U}$ is unambiguous.*

4.3 Proof

4.3.1 Preliminaries

First, we characterize the set $\Delta_{\mathcal{U}}$.

Lemma 23. *Suppose $\succeq$ satisfies Preference Relation, Partial Completeness, Continuity, Second-Stage Independence, $\mathcal{U}$-dominance, Nondegenerate. Then, there exists an additive set function $\mu^* : \mathcal{U} \to [0, 1]$ such that $\Delta_{\mathcal{U}} = \{\mu \in \Delta S | \forall E \in \mathcal{U},\ \mu(E) = \mu^*(E)\}$.*

Proof. *If $\Delta_{\mathcal{U}} = \emptyset$, then $\forall P, Q \in \Delta\mathcal{F},\ P \sim Q$ by $\mathcal{U}$-dominance and this contradicts Nondegenerate. Hence $\Delta_{\mathcal{U}}$ is nonempty.*

Take any expected utility function $u : \Delta X \to \mathbb{R}$ that represents $\succeq|_{\Delta X}$. For any $\mu \in \Delta_{\mathcal{U}}$, partiton $\pi \subset \mathcal{U}$, and $f \in \mathcal{F}_\pi$, define $V^\mu_\pi(f) = \sum_{E\in\pi} u(f(E))\mu(E)$. Because $f \sim \Psi(f,\mu), V^\mu_\pi(\cdot) = u \circ \Psi(\cdot,\mu)$ represents $\succeq|_{\mathcal{F}_\pi}$.

In light of uniqueness result of SEU representation, all $\mu \in \Delta_{\mathcal{U}}$ agree on $\mathcal{U}$. Therefore one can define $\mu^* : \mathcal{U} \to [0,1]$ by $\mu^*(E) = \mu_0(E)$ for all $E \in \mathcal{U}$ using any fixed μ_0. From the definition of $\Delta_{\mathcal{U}}$, $\Delta_{\mathcal{U}} \subset \{\mu \in \Delta S |$for all$E \in \mathcal{U},\ \mu(E) = \mu^*(E)\}$.

In order to prove converse inclusion, take any $\mu \in \Delta S$ such that for all $E \in \mathcal{U},\ \mu(E) = \mu^*(E)$. Take any $f \in \mathcal{F}_{\mathcal{U}}$. Then $f \in \mathcal{F}_\pi$ for some partition $\pi \subset \mathcal{U}$ and

$$V^{\mu_0}_\pi(f) = \sum_{E\in\pi} u(f(E))\mu_0(E) = \sum_{E\in\pi} u(f(E))\mu(E) = u(\Psi(f,\mu)) = V^{\mu_0}_\pi(\Psi(f,\mu)).$$

Because $V^{\mu_0}_\pi$ represents $\succeq|_{\mathcal{F}_\pi}$, $f \sim \Psi(f,\mu)$. This means $\mu \in \Delta_{\mathcal{U}}$. □

4.3.2 Proof of Theorem 1

Now we turn to the proof of Theorem 1. We only prove the sufficiency of axioms. Assume Axioms 8–11. Because $\mathcal{U}$-Dominance implies Dominance, Theorem 4.2 in Seo (2009) can be applied, and there is an SOSEU representation (u, v, m). Under this representation, the utility function of $\succeq$ is written as $V(P) = \int_{\mathcal{F}} \left[\int_{\Delta S} v\left(\int_S u(f)d\mu\right) dm(\mu)\right] dP(f)$.

For each $P \in \Delta\mathcal{F}$, define $\xi_P : \Delta S \to \mathbb{R}$ by $\xi_P(\mu) = V(\Psi(P,\mu)) = \int_{\mathcal{F}} v(\int_S u(f)d\mu)dP$ and set $\Phi = \{\xi_P | P \in \Delta\mathcal{F}\}$. Define $I : \Phi \to \mathbb{R}$ as $I_0(\xi) = \int_{\Delta S} \xi dm$, then

$$\begin{aligned} I(\xi_P) &= \int_{\Delta S} \left[\int_{\mathcal{F}} v\left(\int_S u(f)d\mu\right) dP(f)\right] dm(\mu) \\ &= \int_{\mathcal{F}} \left[\int_{\Delta S} v\left(\int_S u(f)d\mu\right) dm(\mu)\right] dP(f) \\ &= V(P). \end{aligned}$$

Note that I is normalized and linear, that is, $I(\alpha) = \alpha$ and $I(\beta\xi+\zeta) = \beta I(\xi)+I(\zeta)$ when $\alpha, \beta\xi + \zeta, \xi, \zeta \in \Phi$.

Take any $\xi_P, \xi_Q \in \Phi$ such that $\xi_P(\mu) \geq \xi_Q(\mu)$ for any $\mu \in \Delta_{\mathcal{U}}$. This is equivalent to $V(\Psi(P,\mu)) \geq V(\Psi(Q,\mu))$ for any $\mu \in \Delta_{\mathcal{U}}$. Because V represents $\succeq$ and $\succeq$ satisfies $\mathcal{U}$-Dominance, $V(P) \geq V(Q)$, that is, $I(\xi_P) \geq I(\xi_Q)$. Hence, for any $\xi, \zeta \in \Phi$, $\xi|_{\Delta_{\mathcal{U}}} = \zeta|_{\Delta_{\mathcal{U}}}$ implies $I(\xi) = I(\zeta)$.

Set $\hat{\Phi} = \{\xi|_{\Delta_{\mathcal{U}}} | \xi \in \Phi\}$ and define $\hat{I} : \hat{\Phi} \to \mathbb{R}$ by $\hat{I}(\hat{\xi}) = I(\xi)$ $(\hat{\xi} = \xi|_{\Delta_{\mathcal{U}}})$. The functional $\hat{I}$ inherits normalizedness and linearity from I. Using linearity, I can be extended to $span(\hat{\Phi})$. Then it can be extended to $C(\Delta_{\mathcal{U}})$, using Ok(2007) p.594 Proposition 12. We write the obtained functional also as $\hat{I}$. Apply the Riesz Representation Theorem and take $\hat{m} \in \Delta(\Delta_{\mathcal{U}})$ such that $\hat{I}(\xi) = \int_{\Delta_{\mathcal{U}}} \xi d\hat{m}$. Set $\overline{m}(B) = \hat{m}(B \cap \Delta_{\mathcal{U}})$ for each Borel set $B \subset \Delta S$. The tuple $(u, v, \overline{m})$ is an SOSEU representation of $\succeq$. Lemma 1 completes the proof of Theorem 1.

4.3.3 Proof of Theorem 2

Assume that $\succeq$ satisfies Axioms 8–10 and 14–16. These axioms imply the axioms imposed Theorem 2 in Nascimento and Riella (2013). Hence $\succeq$ has a SOMEU representation. Repeat the same argument as the proof of Theorem 1 and obtain a SOMEU representation (u, v, M) under which $\mathcal{U}$ is unambiguous. $\square$

4.3.4 Proof of Theorem 3

Assume $\succeq$ satisfies Axioms 9–12, 14, 17, and 18. These axioms imply the axioms imposed Theorem 5 in Nascimento and Riella (2013). Hence $\succeq$ has a Second-Order Bewley representation. Repeat the same argument as the proof of Theorem 1 and obtain a Second-Order Bewley representation (u, v, M) under which $\mathcal{U}$ is unambiguous. $\square$

Bibliography

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

Allais, M. (1953): "Le Comportement de l'Homme Rationnel Devant Le Risque: Critique Des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21, 503–546.

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

Baldassi, C., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2018): "A Behavioral Characterization of the Drift Diffusion Model," *Working paper*.

Camerer, C. F., G. Loewenstein, and M. Rabin, eds. (2004): *Advances in Behavioral Economics*, Princeton University Press.

Caplin, A. and M. Dean (2015): "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105, 2183–2203.

Caplin, A. and A. Schotter (2008): *The Foundations of Positive and Normative Economics: A Handbook*, Oxford University Press.

Chambers, C. P., C. Liu, and J. Rehbeck (2020): "Costly Information Acquisition," *Journal of Economic Theory*, 186, 104979.

De Finetti, B. (1931): "Sul Significato Soggettivo Della Probabilita," *Fundamenta mathematicae*, 17, 298–329.

de Oliveira, H., T. Denti, M. Mihm, and K. Ozbek (2017): "Rationally Inattentive Preferences and Hidden Information Costs: Rationally Inattentive Preferences," *Theoretical Economics*, 12, 621–654.

de Oliveira, H. and R. Lamba (2019): "Rationalizing Dynamic Choices," *Working Paper.*

Dekel, E. and B. L. Lipman (2010): "How (Not) to Do Decision Theory," *Annual Review of Economics*, 2, 257–282.

Dillenberger, D., V. R. Krishna, and P. Sadowski (2018): "Subjective Information Choice Processes," *Working Paper.*

Duraj, J. and Y.-H. Lin (2019): "Identification and Welfare Analysis in Sequential Sampling Models," *Working paper.*

Echenique, F. and K. Saito (2017): "Response Time and Utility," *Journal of Economic Behavior & Organization*, 139, 49–59.

Edwards, W. (1965): "Optimal Strategies for Seeking Information: Models for Statistics, Choice Reaction Times, and Human Information Processing," *Journal of Mathematical Psychology*, 2, 312–329.

Eliaz, K. and E. A. Ok (2006): "Indifference or Indecisiveness? Choice-Theoretic Foundations of Incomplete Preferences," *Games and economic behavior*, 56, 61–86.

Ellis, A. (2018): "Foundations for Optimal Inattention," *Journal of Economic Theory*, 173, 56–94.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 643–669.

Epstein, L. G. (2010): "A Paradox for the "Smooth Ambiguity" Model of Preference," *Econometrica*, 78, 2085–2099.

Epstein, L. G. and J. Zhang (2001): "Subjective Probabilities on Subjectively Unambiguous Events," *Econometrica*, 69, 265–306.

Fudenberg, D., W. K. Newey, P. Strack, and T. Strzalecki (2019): "Testing the Drift-Diffusion Model," *Working Paper.*

Fudenberg, D., P. Strack, and T. Strzalecki (2018): "Speed, Accuracy, and the Optimal Timing of Choices," *American Economic Review*, 108, 3651–3684.

Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–151.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–292.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, 73, 1849–1892.

——— (2011): "Definitions of Ambiguous Events and the Smooth Ambiguity Model," *Economic Theory*, 48, 399–424.

Knight, F. H. (2012): *Risk, Uncertainty and Profit*, Courier Corporation.

Koida, N. (2017): "A Multiattribute Decision Time Theory," *Theory and Decision*, 83, 407–430.

Kopylov, I. (2007): "Subjective Probabilities on "Small" Domains," *Journal of Economic Theory*, 133, 236–265.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498.

Machina, M. J. and D. Schmeidler (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 745–780.

Nascimento, L. and G. Riella (2013): "Second-Order Ambiguous Beliefs," *Economic Theory*, 52, 1005–1037.

Qu, X. (2013): "Maxmin Expected Utility with Additivity on Unambiguous Events," *Journal of Mathematical Economics*, 49, 245–249.

Ramsey, F. P. (1926): "Truth and Probability," in *Philosophy of Probability: Contemporary Readings*, ed. by A. Eagle, Routledge.

Ratcliff, R. (1978): "A Theory of Memory Retrieval," *Psychological Review*, 85, 59–108.

Sarin, R. and P. Wakker (1992): "A Simple Axiomatization of Nonadditive Expected Utility," *Econometrica*, 60, 1255–1272.

SAVAGE, L. (1954): *The Foundations of Statistics*, New York: John Wiley and Sons.

SCHMEIDLER, D. (1986): "Integral Representation without Additivity," *Proceedings of the American mathematical society*, 97, 255–261.

——— (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57, 571–587.

SEO, K. (2009): "Ambiguity and Second-Order Belief," *Econometrica*, 77, 1575–1605.

SIMS, C. A. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665–690.

SPILIOPOULOS, L. AND A. ORTMANN (2018): "The BCD of Response Time Analysis in Experimental Economics," *Experimental Economics*, 21, 383–433.

STRZALECKI, T. (2011): "Axiomatic Foundations of Multiplier Preferences," *Econometrica*, 79, 47–73.

TAKEOKA, N. (2007): "Subjective Probability over a Subjective Decision Tree," *Journal of Economic Theory*, 136, 536–571.

VON NEUMANN, J. AND O. MORGENSTERN (1947): *Theory of Games and Economic Behavior, 2nd Rev.*, Princeton university press.

ZHANG, J. (2002): "Subjective Ambiguity, Expected Utility and Choquet Expected Utility," *Economic Theory*, 20, 159–181.