HIAS-E-101

Response time and revealed information structure

Tomohito Aoyama

Hitotsubashi Institute for Advanced Study, Hitotsubashi University.

December, 2020

Response time and revealed information structure*

Tomohito Aoyama[†]

November 30, 2020

Abstract

Consider a decision-maker who has an opportunity to wait for information before making a choice. He can obtain more information by waiting more, but this is costly. As a result, he endogenously determines the length of time to choose an alternative, which is called the *response time*. The present study models such a decision-maker as if he solves an optimal stopping problem. The model incorporates a dynamic information structure formalized as an evolving information partition, which is called filtration. I axiomatically characterize the model using behavioral data consisting of choices and response times that depend on choice situations and states. That is, from the data, we can identify filtration that governs the decision-maker's learning process as well as other model parameters. This result implies that using response time helps us understand the human cognitive process.

KEYWORDS: Response time, Subjective learning, Information acquisition

JEL CLASSIFICATION: D01, D81, D83

*I am grateful to Koji Abe, Yasuyuki Miyahara, Hisao Hisamoto, and Hideo Suehiro for their continuous support and guidance. I also thank Norio Takeoka and Masaki Miyashita for very useful comments and suggestions. I also thank audiences at SWET 2019, the 2019 annual meeting of Association of Behavioral Economics and Finance, and the 2020 spring meeting of The Japanese Economic Association, and seminar participants at Kobe University and Hitotsubashi University. This work was supported by Grant-in-Aid for JSPS Fellows Grant Number JP18J12252.

†Hitotsubashi Institute for Advanced Study, Hitotsubashi University. Email: t.aoyama@r.hit-u.ac.jp

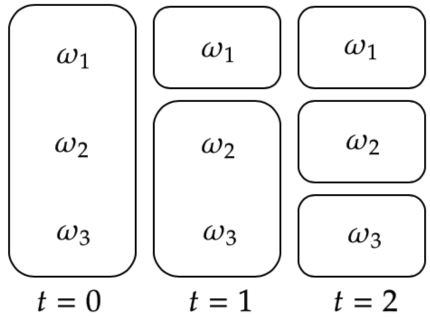

Figure 1: Filtration

1 Introduction

In many economic environments, choice timing is not fixed exogenously; instead, it is a choice variable of the decision-maker (DM). As a result, the amount of time consumed to choose, or *response time* (RT), reflects the decision process of the DM. In this paper, I theoretically show that response time data, together with choice data, help us identify the decision process.

Assuming that the available behavioral data are choice and response time data conditional on the state of the world, I axiomatically characterize a model that incorporates endogenously determined learning processes. The model I consider, *optimal stopping representation* (OSR), describes a DM who decides when and what to choose according to his private learning process. The learning process is described by information filtration, that is, a sequence of information partitions corresponding to knowledge evolution over time. Filtration is not directly observable from an analyst's view, while it is fixed from the perspective of the DM. One of the contributions of this study is the identification of the DM's filtration and other model parameters such as expected utility function, subjective probability, and waiting cost.

My approach to identify filtration is as follows. If DM would learn a realization of some event at a point in time, he uses this information if doing so is profitable. Information he has would thus be reflected in choice and response time. Therefore, filtration is defined as the smallest one that is necessary to describe his behavior. Once I define the filtration in this way, I can elicit other parameters for an OSR with a way of Ellis (2018) where I modify suitably to fit the setting of this paper.

I explain the identification strategy for filtration and waiting cost through an example. Consider a state space $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and a choice situation $B_x =$

$\{f_x, g_x\}$, where $f_x = (x, x, 0)$ and $g_x = (0, 0, x)$. The bet f_x pays \$$x$ in ω_1 and ω_2, and g_x pays \$$x$ in ω_3. Suppose the subjective filtration of the DM $\mathcal{F}$ is as depicted in Figure 1. At $t = 0$, the DM has no additional information beyond his prior belief. However, waiting until $t = 1$, he learns whether the true state is ω_1. Further, waiting until $t = 2$, he learns the true state completely. Assume waiting longer is costly. How can the analyst identify the filtration? First, suppose x is large enough that DM has a strong incentive to learn the state. Then, DM first waits until $t = 1$ and learns whether the state is ω_1. If it is ω_1, he stops and chooses f_x. Otherwise, he continues to wait until $t = 2$ to learn the true state. Then, he chooses f_x if it is ω_2, and g_x if it is ω_3. In this case, the analyst concludes that the DM distinguishes ω_1 and other states at $t = 1$ by observing the difference of RTs in ω_1 and in other states. Moreover, observing the difference of choices in ω_2 and ω_3, the analyst can conclude that the DM can distinguish ω_2 and ω_3 at $t = 2$. Integrating these observations, the analyst elicits the subjective filtration. Next, consider decreasing x. Then, the incentive to acquire additional information gets weaker than that given the original. At some point, the DM may decide to stop at $t = 1$ in all states. The amount of x decreased by that time bounds the difference in the costs of two RTs. This procedure identifies the waiting cost.

Optimal stopping representation is a dynamic extension of the *optimal inattention representation* (OIR) studied by Ellis (2018), who characterizes a model of rational inattention, assuming that the observable data are choices conditional on the states. In an OIR, the DM chooses an alternative after a single information acquisition. The learning process is described by an information partition that is selected by the DM for each choice situation. However, in OSR, the DM selects a stopping time for each choice situation, holding the information filtration fixed. The two representations cannot be distinguished solely in terms of choice. However, they explain the DM's choice through different informations. Consider the example in the previous paragraph. From the original menu B_x, the DM chooses f_x in ω_1 and ω_2, but chooses g_x in ω_3. In this choice situation, Ellis's strategy identifies the information the DM acquires as $\{\{\omega_1, \omega_2\}, \{\omega_3\}\}$.[1] However, the DM stops waiting at different timings in ω_1 and ω_2. As I illustrated, the analyst can use this RT data to conclude that the DM learns the true state completely. That is, the identified information partition is $\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$.

Finally, I present a generalization of OSR. People often make an effort to acquire information quickly, but also do so slowly on occasion. Because OSR has a fixed filtration, it does not allow such a flexible choice of information structure. A natural way to model this class of information acquisition is to incorporate a choice of

[1] This is what he calls canonical attention rule.

filtration, not only response time. In Section 5, I introduce a model of this kind and show that it has no testable implication beyond the optimal inattention model. This means that, as long as we are concerned with the data this paper treats, it is hard to extend OIR further than OSR, maintaining an explanatory power on response times.

Literature

Here I overview related decision-theoretic studies.

Decision theoretic approaches to response time

Some recent decision-theoretic papers study what can be learned about the cognitive process using RT. Among others, Duraj and Lin (2019) is the closest to the present paper. They behaviorally characterize a DM who solves the optimal stopping problem with constant waiting cost or geometric discounting, given a filtration over state space. The most critical difference between Duraj and Lin (2019) and this paper is that they assume that the analyst can directly observe a DM's filtration while I treat filtration as a subjective one. In this paper, I assume that the analyst can observe the true state, which allows the identification of filtration.

I briefly review other studies. Echenique and Saito (2017) characterized a model in which the differences of choice values determine response time. Koida (2017) studies a sequence of incomplete preference relations that become more comparable over time; in his model, RT is determined when two alternatives become comparable. Baldassi et al. (2018) characterize drift-diffusion model and its multi-alternative extension.

Dynamic information acquisition

The aim of this paper is to reveal how a DM' s uncertainty resolves over time. Several existing studies analyzed problems in this regard. Takeoka (2007) adopted a menu of menus of acts as an alternative and characterized a model with two-stage costless information acquisition. de Oliveira and Lamba (2019) ask a problem on judging whether a DM's action sequence can be rationalized by information flow. Dillenberger et al. (2018) study an infinite horizon decision model in which the state evolves following a Markov process and the DM acquires information by choosing an information partition. While all of these studies assume exogenous choice timing, I assume it as endogenous.

Rational Inattention

Economic agents often feel that information acquisition is costly, perhaps for limitation of cognitive ability, and thus, they may avoid acquiring all information even if that is materially costless. This insight, called *rational inattention*, was introduced to economics by Sims (2003). Several decision-theoretic studies provide behavioral foundations for rational inattention theory using different primitives.

An early paper of this kind is Hyogo (2007). In his model, the DM chooses a pair of an experiment and a menu, where the anticipated information value of the experiment is subjective. The following recent studies assume that the analyst does not observe choices of experiments and rather infers them. Caplin and Dean (2015) use choice probabilities conditional on states as data. 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 and characterize rational inattention model with additive information cost. Higashi et al. (2020) characterize a generalization and a discount counterpart of de Oliveira et al. (2017). As I explained, Ellis (2018) uses a state-conditional choice correspondence.

The rest of this paper is organized as follows. In section 2, the analytical framework and optimal stopping representation are introduced. In section 3, I present the axioms and show the representation results. In section 4, the identification result and comparative statics are presented. In section 5, I introduce a more general model and compare the predictive power of my models and Ellis's OIR. Section 6 concludes. Proofs are collected in section 7.

2 Setup and model

2.1 Setup

This subsection introduces the framework. Let Ω be a finite state space. Let X be a convex subset of a metrizable vector space and let d be its compatible metric. Let $\mathcal{A}$ be the set of functions that take Ω to X. Each element of $\mathcal{A}$is called an *act*. With a natural isomorphism, X is regarded as 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. For typical elements of the sets above, I write $x, y, z \in X$, $f, g, h \in \mathcal{A}$, and $A, B, C \in \mathcal{K}$.

For any algebra[2] Q over Ω and $\omega \in \Omega$, $Q(\omega)$ denotes the smallest element of Q

[2] An algebra over Ω is a subset of 2^Ω, which is closed under complementation and union, and

that contains ω. A collection of algebras $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ is a filtration if $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \leq t$. For filtrations $\mathcal{F}, \mathcal{G}$, write $\mathcal{F} \ll \mathcal{G}$ if $\mathcal{F}_t \subset \mathcal{G}_t$ for all $t \geq 0$. For any collection $\mathcal{E}$ of events, $\sigma(\mathcal{E})$ denotes the algebra generated by $\mathcal{E}$.

I call a function $\lambda : \Omega \to \mathbb{R}_+$ a stopping time. A stopping time is interpreted as planning on how long to wait for information. Typical stopping times are denoted as λ, μ. For a stopping time λ and $t \in \mathbb{R}_+$, let $\{\lambda \leq t\} = \{\omega \in \Omega | \lambda(\omega) \leq t\}$. This is the event that he stops before t following λ. A stopping time λ is adapted to a filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ if $\{\lambda \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. When this is the case, the DM can follow λ using the information represented by $\mathcal{F}$. For a filtration $\mathcal{F}$, let $\mathcal{T}_{\mathcal{F}}$ be the set of stopping times adapted to $\mathcal{F}$. I sometimes denote $\mathcal{T}_{\mathcal{F}}$ just as $\mathcal{T}$ when what $\mathcal{F}$ denotes for is clear. For a filtration $\mathcal{F}$ and $\lambda \in \mathcal{T}_{\mathcal{F}}$, let $\mathcal{F}_\lambda = \{\mathcal{F}_{\lambda(\omega)}(\omega) | \omega \in \Omega\}$, which represents the information obtained by using λ given $\mathcal{F}$.[3]

A data (c, τ) is a pair of choice correspondence and response time, both conditional on states. A choice correspondence is a function $c : \mathcal{K} \times \Omega \to \mathcal{K}$ that satisfies $c(B, \omega) \subset B$ for any $(B, \omega) \in \mathcal{K} \times \Omega$. A value $c(B, \omega)$ is what DM chooses given B, and ω realizes. The response time is a function $\tau : \mathcal{K} \times \Omega \to \mathbb{R}_+$. A value $\tau(B, \omega)$ is the amount of time the DM waited before the decision when B is given, and ω is the true state. Both c and τ are state-dependent because the DM acquires partial information on the state and is reflected in choice and response time. Given B, the function $\tau(B, \cdot)$ is a stopping time, which is abbreviated as τ_B.

2.2 Model

Here I explain the model to be studied. The story of the model is as follows: Given a menu, the DM first chooses a stopping time that bears waiting cost. While waiting for information, he successively acquires finer information on the true state. Stopping at the time designated by τ, he uses the obtained information to update his prior belief following the Bayes rule. Then, he chooses an alternative $f \in B$ that maximizes the expected utility. Anticipating this, the choice of stopping time is done to maximize the ex-ante net utility. Now I introduce the model formally.

Definition 1. *An optimal stopping representation (OSR) $(u, \pi, \mathcal{F}, \gamma)$ is a quadruple of*

- *a continuous affine function $u : X \to \mathbb{R}$ with $u(X) = \mathbb{R}$,*
- *a full-support probability π over Ω,*

contains Ω. The reason I use algebras, instead of parititons, to formalize information is to simplify Definition 3.

[3]This definition of $\mathcal{F}_\lambda$ is equivalent to the usual one in probability theory in this framework.

- *a filtration $\mathcal{F}$ over Ω,*
- *a cost function $\gamma : \mathcal{T}_\mathcal{F} \to \mathbb{R}_+ \cup \{\infty\}$ such that $\gamma(\lambda) \leq \gamma(\mu)$ whenever $\lambda \leq \mu$ and $\gamma(0) = 0$.*

It represents a data (c, τ) if it satisfies the following conditions:

$$\tau_B \in \arg\max_{\lambda \in \mathcal{T}_\mathcal{F}} \mathrm{E}_\pi \left[\max_{f \in B} \mathrm{E}_\pi[u(f)|\mathcal{F}_\lambda] \right] - \gamma(\lambda), \tag{1}$$

$$c(B, \omega) = \arg\max_{f \in B} \mathrm{E}_\pi[u(f)|\mathcal{F}_{\tau_B}](\omega) \text{ for all } \omega \in \Omega. \tag{2}$$

The first line of the representation requires that given a menu B, the observed response time τ_B maximizes the DM's ex-ante expected net utility. The second line requires that he chooses alternatives that maximize conditional expected utility. If the data satisfy these conditions, they are interpreted as generated by the DM, who solves the optimal stopping problem.

3 Axiomatic characterization

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

3.1 Optimal stopping axioms

Before introducing optimal stopping axioms, I explain how to elicit the filtration $\mathcal{F}$ from the data. First, I define a collection of binary relations $\{\bowtie_t\}_{t \in \mathbb{R}_+}$ over Ω that reflects the information DM has at each time point. For $r, s \in \mathbb{R}$, let $r \wedge s = \min\{r, s\}$.

Definition 2. *Two states ω and ω' are not distinguished until $t \in \mathbb{R}_+$ if there exists no $B \in \mathcal{K}$ that satisfies the following two conditions:*

- *$\tau(B, \omega) \wedge \tau(B, \omega') \leq t$,*
- *$\tau(B, \omega) \neq \tau(B, \omega')$ or $c(B, \omega) \neq c(B, \omega')$.*

In this case, I write $\omega \bowtie_t \omega'$.

This definition is based on a two-step procedure to judge whether two states are distinguished by DM at some time point. Suppose there is a menu B such that $\tau(B,\omega) \leq t$. First, if $\tau(B,\omega) \neq \tau(B,\omega')$, the DM can distinguish ω and ω' at $\tau(B,\omega)$, and hence, at t. Otherwise, he should stop at the same time in ω and ω'. Second, if $c(B,\omega) \neq c(B,\omega')$, the DM could distinguish two states also in this case at t. If neither observation is obtained given any menu, there is no reason to consider the two states as distinguished until t. The obtained relation $\bowtie_t$ generates an algebra over Ω that represents the information the DM has at t.

Definition 3. *(Revealed) filtration is the indexed collection $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ of algebras over Ω, given by*

$$\mathcal{F}_t = \sigma(\{\{\omega' \,|\omega' \bowtie_t \omega\} \,|\omega \in \Omega\}). \tag{3}$$

Therefore, filtration is defined as the minimal one that is consistent with both choice data and response time. By definition, the stopping time τ_B is $\mathcal{F}$-adapted for any $B \in \mathcal{K}$.

The next axiom *Dynamic Subjective Consequentialism* requires that the DM respects the revealed filtration that is elicited in Definition 3.

Axiom 1 (*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*

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

then

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

Given B, the DM acquires information $\mathcal{F}_{\tau_B}$ using τ_B and knows the realization of an event $\Delta \in \mathcal{F}_{\tau_B}$ that contains the true state ω. If $f, g \in B$ agree on Δ, he treats them as if they are the same act. This is what *DSC* states.

The next axiom *Information Monotonicity* requires that response time reflects the amount of information necessary for the choices given each menu.

Axiom 2 (*IM: Information Monotonicity*)**.** *If $c(A,\omega) \neq c(A,\omega')$ implies $c(B,\omega) \neq c(B,\omega')$ for any $\omega, \omega' \in \Omega$, then $\tau_A \leq \tau_B$.*

This axiom states that if choice behavior given B requires more information than when given A in the sense that whenever the DM's choice distinguishes ω and ω' at A, it is the case at B, then the DM waits longer given B. In particular, if A is singleton, the DM has no incentive for information acquisition. Then, in this case, he stops faster than in any other situation. As long as the information structure is independent of choice situations and evolves over time, stopping time should become longer as choice requires finer information.

3.2 Optimal inattention axioms

I turn to optimal inattention axioms originally appeared in Ellis that axiomatizes an optimal inattention representation. Here, I briefly introduce them. The axiom *Independence of Nonrelevant Alternative* is a variant of the Weak Axiom of Revealed Preference that is imposed to the conditional choice correspondence.

Axiom 3 (*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 the next axiom, I introduce a mixing operation over acts. For $f, g \in \mathcal{A}$ and $\alpha \in [0,1]$, let $\alpha f + (1-\alpha)g$ be an act defined by $[\alpha f + (1-\alpha)g](\omega) = \alpha f(\omega) + (1-\alpha)g(\omega)$. This operation is naturally extended as follows. For $f \in \mathcal{A}$ and $B \in \mathcal{K}$, let $\alpha f + (1-\alpha)B = \{\alpha f + (1-\alpha)g | g \in B\}$. The axiom Attention Constrained Independence is a form of independence axiom for choice correspondence, which is an implication of additive information cost.

Axiom 4 (*ACI: Attention Constrained Independence*)**.** *If* $\alpha g + (1-\alpha)f \in c(\alpha g + (1-\alpha)B, \omega)$*, then* $\alpha h + (1-\alpha)f \in c(\alpha h + (1-\alpha)B, \omega)$*.*

For the next axiom, I define a preference relation over outcomes as follows. 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 and symmetric parts of $\succeq^R$, respectively.

Monotonicity states that if an act is chosen from a menu and it is state-wise dominated by another act in it, then the latter must also be chosen.[4]

[4]In the original form of Monotonicity in Ellis includes an assumption of a no null-state. I replace it here for Axiom $C(2)$ in the next subsection.

Axiom 5 (*M: Monotonicity*)**.** *For $f, g \in B$, if*

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

then

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

3.3 Technical axioms

Here, I pose technical axioms. Introducing preference relations for menus is necessary to state the axioms. Let $\succeq^D$ be a relation over menus such that $A \succeq^D B$ if and only if $A \cap c(B, \omega) \neq \emptyset$ for all $\omega \in \Omega$. The relation $A \succeq^D B$ means that, given A, the DM can follow the optimal policy he uses if B is given. Therefore, if he can choose one menu from A and B, he weakly prefers A. In this sense, this is a directly revealed preference relation for menus. Next, let $\succeq^I$ denote the transitive closure of $\succeq^D$.[5] That is, assuming his hypothetical preferences for menus are transitive, $\succeq^D$ is extended to $\succeq^I$. The relation $\succeq^I$ is an indirectly revealed ranking obtained in this way.

The next axiom is *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. I write $g \succeq^I f$ if $\{g\} \succeq^I \{f\}$ for simplicity.

Axiom 6 (*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, which is equivalent to Ellis's original condition under *Monotonicity.* The second part is an implication of the assumption of a no null state. Suppose a state ω is a non-null state and x is strictly preferred over y. Then, if f_n and g_n are sufficiently similar to $x\omega y$ and y, respectively, g_n is not preferred over f_n because the DM does not neglect

[5] That is, $A \succeq^I B$ if and only if there exists $B_1, \ldots, B_n$ such that $A \succeq^D B_1 \succeq^D \cdots \succeq^D B_n \succeq^D B$.

the outcome in ω. Together with other axioms, it guarantees the existence of a full-support subjective probability.[6]

The next axiom is *Unboundedness*, which is used to calibrate the cost function.

Axiom 7 (*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*

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

3.4 Representation theorem

Now I present a sufficiency result. The axioms I offer are sufficient for the data to have an OSR.

Theorem 1. *If c and τ satisfy INRA, ACI, M, DSC, IM, C, and U, then there exist u, π, and γ such that $(u, \pi, \mathcal{F}, \gamma)$ is an OSR of (c, τ).*

Next, I turn to the necessity of the axioms.

Theorem 2. *For any OSR $(u, \pi, \mathcal{F}, \gamma)$, there exists a data (c, τ) that satisfies the following:*

1. *(c, τ) satisfies ACI, IM, M, DSC, C, and U.*
2. *There is a dense open subset $\mathcal{D}$ of $\mathcal{K}$ and a conditional choice correspondence c' that satisfies INRA such that the following holds:*

$$\text{for any } B \in \mathcal{D} \text{ and } \omega \in \Omega,\; c(B, \omega) = c'(B, \omega).$$

In general, an OSR has multiple data consistent with it. While such data may not satisfy all the postulated axioms in general, there is at least one set of data that does. This is why the above theorem is written as an existence result. Theorem 1 and Theorem 2 shows that the axioms I posed almost characterize OSR.

3.5 Sketch of the proof of Theorem 1

To show the sufficiency result, I consider a preference relation $\succeq$ for *plans*, which is a pair (F, λ) of a function $F : \Omega \to \mathcal{A}$ and a stopping time λ such that F is $\mathcal{F}_\lambda$-measureble. Given a menu B, the DM chooses a stopping time λ and finds the best

[6] 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 because the non-triviality of $\succeq^I$ is not guaranteed.

act given the event $\mathcal{F}_\lambda(\omega)$ for each ω. These choice of stopping time and anticipated choices are summarized by the plan (F,λ). The relation $\succeq$ is derived from $\succeq^I$. Suppose $F(\omega)\in c(A,\omega)$ for each ω and $\lambda=\tau_A$. Similarly, suppose $G(\omega)\in c(B,\omega)$ and $\mu=\tau_B$. Then, the DM follows (F,λ) given A, and (G,μ) given B. From these observations, I define $(F,\lambda)\succeq(G,\mu)$ if $A\succeq^I B$. In this approach, the key axiom is *INRA*, which guarantees that the preferences for menus are determined by that for plans.

Then, I construct a utility representation

$$V(F,\lambda)=\mathrm{E}_\pi[\mathrm{E}_\pi[u(F(\omega)(\omega))|\mathcal{F}_\lambda]]-\gamma(\lambda).$$

Axiom *ACI* implies that $\succeq^R$ and the restriction of $\succeq^I$ to singleton menus satisfy independence. This observation is used to elicit u and π, together with other axioms. Also, *ACI* and *IM* jointly imply the translation invariance of $\succeq$: if $(\alpha F+(1-\alpha)f,\lambda)\succeq(\alpha G+(1-\alpha)f,\mu)$, then $(\alpha F+(1-\alpha)g,\lambda)\succeq(\alpha G+(1-\alpha)g,\mu)$. This is the implication of additive waiting cost and allows the calibration of γ.

Once the utility representation is obtained, it is used to explain the data (c,τ). Consider a menu B, and plans (F,τ_B) and (G,λ) such that $F(\omega)\in c(B,\omega)$ and $\mathrm{Im}G\subset B$. That is, (F,τ_B) is the plan actually chosen given B, and (G,λ) is another candidate. It is shown that $(F,\tau_B)\succeq(G,\lambda)$ or $V(F,\tau_B)\geq V(G,\lambda)$ holds. Therefore observed the choices of a stopping time and acts is interpred as the maximization of the net utility.

4 Identification and comparative statics

In this section, I present identification results and comparative statics on filtration. The identification of filtration is partial. That is, it cannot be uniquely identified but can be bounded to some extent. The next proposition shows that filtration defined in Definition 3 is the coarsest one that is consistent with the observed behavior.

Proposition 1. *Suppose $\mathcal{F}=\{\mathcal{F}\}_{B\in\mathcal{K}}$ is the revealed filteration of (c,τ). If (c,τ) has an OSR $(u,\pi,\mathcal{G},\gamma)$, then $\mathcal{F}\ll\mathcal{G}$.*

Next, I present the identification result for other parameters. Let $\mathcal{T}^*=\{\tau_B|B\in\mathcal{K}\}$ be the set of stopping times the DM uses given some menu.

Proposition 2. *Suppose that $(u,\pi,\mathcal{F},\gamma)$ and $(u',\pi',\mathcal{F}',\gamma')$ represent (c,τ) and $\gamma(\lambda)>0$ for any $\lambda\in\mathcal{T}_\mathcal{F}$ with $\mathcal{F}_\lambda\neq\{\Omega\}$. Then,*

1. *$\pi=\pi'$,*

2. *There exist* $\alpha > 0$ *and* $\beta \in \mathbb{R}$ *such that*

$$u' = \alpha u + \beta$$

and

$$\gamma'(\tau) = \alpha\gamma(\tau)$$

for any $\tau \in \mathcal{T}^*$.

That is, under the assumption that any stopping time is costly whenever it conveys some information, subjective probability is uniquely identified. Expected utility is unique up to affine transformation. Moreover, cost function is also identified over $\mathcal{T}^*$.

Next, I present the comparative statics of filtration. The relation $\omega \bowtie_t \omega'$ is interpreted as a DM not distinguishing ω and ω' at t. Based on this interpretation, I define the following comparative statics notion.

Definition 4. *For data* (c, τ) *and* $(\tilde{c}, \tilde{\tau})$, *let* $\bowtie_t$ *and* $\tilde{\bowtie}_t$ *be their relation defined as Definition 2, respectively. Individual* (c, τ) *learns faster than another* $(\tilde{c}, \tilde{\tau})$ *if, for any* t *and* $\omega, \omega' \in \Omega$,

$$\omega \bowtie_t \omega' \Rightarrow \omega \tilde{\bowtie}_t \omega'$$

for any $t \geq 0$.

This notion corresponds to the fineness of the filtration.

Corollary 1. *An individual* (c, τ) *learns faster than* $(\tilde{c}, \tilde{\tau})$ *if and only if their filtrations* $\mathcal{F}$ *and* $\tilde{\mathcal{F}}$ *satisfy* $\mathcal{F} \gg \tilde{\mathcal{F}}$.

This result follows from Definition 3 as a corollary. The proof is obvious and so omitted.

5 Generalization and model comparisons

An OSR describes a DM whose information structure is fixed. On the contrary, people sometimes collect information quickly and sometimes slowly. Therefore, a generalization of OSR that models joint choices of filtration, stopping time, and acts is also worth considering. In this section, I formalize such a model. As a result, it is shown that this model has no testable implications beyond the OIR of Ellis.

Let Π be the set of pairs $(\mathcal{F}, \lambda)$ of filtration and stopping time adapted to the former. The following model is a generalization of OSR.

Definition 5. *An generalized optimal stopping representation (GOSR)* $(u, \pi, \{\mathcal{F}^B\}_{B\in\mathcal{K}}, \gamma)$ *is a quadruple of*

- *a continuous affine function* $u : X \to \mathbb{R}$ *with* $u(X) = \mathbb{R}$,
- *a full-support probability* π *over* Ω,
- *a collection* $\{\mathcal{F}^B\}_{B\in\mathcal{K}}$ *of filtrations indexed by menus,*
- *a cost function* $\gamma : \Pi \to \mathbb{R}_+ \cup \{\infty\}$ *such that* $\gamma(\mathcal{F}, \lambda) \leq \gamma(\mathcal{G}, \mu)$ *whenever* $\mathcal{F} \ll \mathcal{G}$ *and* $\lambda \leq \mu$.

It represents a data (c, τ) *if the following conditions hold:*

$$(\mathcal{F}^B, \tau_B) \in \arg\max_{(\mathcal{F},\lambda)\in\Pi} \mathrm{E}_\pi[\max_{f\in B} \mathrm{E}_\pi[u(f)|\mathcal{F}^B_\lambda]] - \gamma(\lambda), \tag{4}$$

$$c(B, \omega) = \arg\max_{f\in B} \mathrm{E}_\pi[u(f)|\mathcal{F}_{\tau_B}](\omega) \text{ for all } \omega \in \Omega. \tag{5}$$

In contrast to OSR, GOSR has a collection of filtrations (not a fixed filtration) as a parameter. It models a DM who jointly chooses a filtration and response time that maximizes the net utility. Optimal stopping representation is a special case of GOSR where γ is finite only for some fixed filtration.

For comparison with GOSR, the optimal attention representation (OIR) of Ellis is introduced. Let $\mathbb{P}$ denote the set of all partitions over Ω.

Definition 6. *An optimal inattention representation (OIR) is quadruple* $(u, \pi, \hat{\gamma}, \hat{\mathrm{P}})$ *of*

- *a continuous affine function* $u : X \to \mathbb{R}$ *with* $u(X) = \mathbb{R}$,
- *a full-support probability* π *over* Ω,
- *a cost function* $\hat{\gamma} : \mathbb{P} \to \mathbb{R}_+ \cup \{\infty\}$ *such that* $\hat{\gamma}(\{\Omega\}) = 0$ *and* $\hat{\gamma}(Q) \geq \hat{\gamma}(R)$ *if* $Q \gg R$,
- *a function* $\hat{\mathrm{P}} : \mathcal{K} \to \mathbb{P}$.

It represents c *if the following conditions hold:*

$$\hat{\mathrm{P}}(B) \in \arg\max_{Q\in\mathbb{P}} \mathrm{E}_\pi[\max_{f\in B} \mathrm{E}_\pi[u(f)|Q]] - \hat{\gamma}(Q), \tag{6}$$

$$c(B, \omega) = \arg\max_{f\in B} \mathrm{E}_\pi[u(f)|\hat{\mathrm{P}}(B)](\omega). \tag{7}$$

The OIR is a model of optimal inattention in which information acquisition is modeled as choosing costly partitions. In the model, given a menu, the DM chooses a partition optimally to maximize the net utility.

In this paper, I have studied additional implications to observables, adding temporal structure to information flow. Such a model is not meaningful as long as it has some implications for RT. The next result shows the (non-)existence of implication of the models I considered.

Proposition 3.

1. *If (c, τ) has a GOSR, then c has an OIR.*
2. *If c has an OIR, for any response time τ, (c, τ) has a GOSR.*
3. *There exists a data (c, τ) that does not have OSR, while c has an OIR.*

First, any choice consistent with a GOSR can be explained by an OIR. Therefore, as long as choice data is the only concern, GOSR, and in particular OSR, is not distinguished from OIR. Second, GOSR has no testable implication beyond OIR. Third, on the contrary, OSR has additional implications. Therefore, complementing choice data with RT gives additional information.

6 Concluding Remarks

In this paper, I discuss the behavioral implications of a model of DM who dynamically acquires information before choice and the extent to which its parameters can be identified. In the identification strategy, response time data plays an important role. It reflects the time point when additional information becomes available, making it possible to identify the filtration. Compared with the OIR of Ellis, the model with a fixed filtration has an additional implication on RT, while a general model that allows a menu-dependent choice of filtration does not.

In Section 5, I show that GOSR has no testable implication beyond OIR. In order to treat GOSR meaningfully, one needs another primitives. One possible alternative is state-dependent joint distribution over choice and RT. But this extension is yet to be studied.

The proof of sufficiency result largely follows that of Ellis. I note that the proof of Ellis has two correctable gaps. In his proof, he considers plans that are measurable with respect to some revealed partition, and then constructs a utility representation over them. The first gap is that even if a plan F requires a partition coarser than some one actually used, the existence of a menu B with $P(B) = \sigma(F)$ is not immediately guaranteed. Then, the plan may not be comparable with other plans in terms of $\succeq^I$. Therefore, such plans must be excluded through the elicitation of parameters. To avoid the same problem I define $\succeq$ only on plans that are implemented with observed stopping times, and then extend it to all plans. The second gap is that non-degeneracy of $\succeq^I$, which is needed to elicit the subjective probability, is not shown. In my proof, C (2) guarantees non-degeneracy.

7 Proofs

7.1 Proof of Theorem 1

7.1.1 Basic properties of choice correspondence

In this subsection, I investigate the basic properties of the choice correspondence. First, I construct an expected utility function. Let $\mathcal{K}(X)$ be the set of all non-empty compact subsets of X.[7]

Lemma 1. *Then, 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. I 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 $y \succeq^R x$. This implies that $\{y \in X | y \succeq^R x\}$ is closed.

Next, I 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. Therefore, $x \succeq^R z$.

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. I 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 $y \notin \max_{\succeq^R} B$

[7] Because *Monotonicity* axiom in this paper is weaker than that of Ellis, I modified the proof of Lemma 1 is appropriately.

and take $x \in \max_{\succeq^R} B$. Then, $x \succ^R y$ and $x \in c(B,\omega)$ by the first inclusion. By *INRA*, $c(\{x,y\},\omega) = \{x,y\} \cap c(B,\omega)$ holds for all ω. Then, if $y \in c(B,\omega)$, $y \in c(\{x,y\},\omega)$ and so $y \succeq^R x$. But this is a contradiction. Thus $y \notin c(B,\omega)$. That is, $c(B,\omega) \subset \max_{\succeq^R} B$.

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

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 from A; secondly, the choice behavior respects the acquired information.

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 *M* implies $f \in c(A \cup B,\omega)$. Therefore $A \cap c(A \cup B,\omega) \neq \emptyset$ for any $\omega \in \Omega$. Applying *INRA* shows the first part.

Turn to the second part. Define an act h by

$$h(\omega') = \begin{cases} f(\omega') & \text{if } \omega' \in \Delta, \\ g(\omega') & \text{otherwise} \end{cases}$$

and let $B' = B \cup \{h\}$. Then by $u \circ g = u \circ h$ and the first part, $B \cap c(B',\omega') = c(B,\omega')$ for all ω'. Since $f(\omega') = h(\omega')$ for $\omega' \in \Delta$, *DSC* implies $h \in c(B',\omega)$. Finally, $u \circ g = u \circ h$ and *M* imply $g \in c(B',\omega)$, and so $g \in c(B,\omega)$. The second part is proved. □

7.1.2 Transformation of acts into utility acts

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

First, I construct a set $Y \subset X$ such that $u(Y) = \mathbb{R}$ and show the restriction $u|_Y : Y \to \mathbb{R}$ 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 by 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 $\overline{v}_n : \mathbb{R} \to X$ as

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

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.

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 $v(\hat{x}_k) = v_n(\hat{x}_k) \to v_n(\hat{x}) = v(\hat{x})$ 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 $\mathcal{A}_Y = \{f \in \mathcal{A} |\ \forall \omega \in \Omega\ f(\omega) \in Y\}$. Define a function $\Phi^* : \mathcal{A} \to \mathbb{R}^\Omega$ by $\Phi^*(f) = u \circ f$, and denote its restriction to $\mathcal{A}_Y$ by $\Phi : \mathcal{A}_Y \to \mathbb{R}^\Omega$. Lastly, define $\Psi : \mathcal{A} \to \mathcal{A}_Y$ by $\Psi = \Phi^{-1} \circ \Phi^*$.

Lemma 4.

1. *Φ^* is continuous.*
2. *Φ is a homeomorphism.*
3. *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 $|u(x) - u(y)| < \epsilon$. Therefore, 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, $\|u \circ f_k - u \circ f\| \to 0$.

Consider the second part. The continuity of Φ follows from that of Φ^*. I shall prove that Φ is a bijection. For $\hat{f} \in \mathbb{R}^\Omega$, $\Phi(v \circ \hat{f}) = u \circ v \circ \hat{f} = \hat{f}$. 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, I show that the inverse function $\Phi^{-1} : \mathbb{R}^\Omega \to \mathcal{A}_Y$ is continuous. Note that $\Phi^{-1}(\hat{f}) = v \circ \hat{f}$. Take a sequence $\{\hat{f}_k\}_{k=1}^\infty \in \mathbb{R}^\Omega$ such that $\hat{f}_k \to \hat{f}$ in $\mathbb{R}^\Omega$. 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 $d_\infty(\Phi^{-1}(\hat{f}_k), \Phi^{-1}(\hat{f})) < \epsilon$. The third part is trivial. □

Next, I show that the choice behavior depends only on the state-dependent utilities of the acts in the choice sets. For a moment, I 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 the followings hold:*

1. *$[c(A, \omega)]^u = [c(B, \omega)]^u$ for each $\omega \in \Omega$.*
2. *$\tau_A = \tau_B$.*

Proof. Lemma 2 (1) implies

$$c(A, \omega) = A \cap c(A \cup B, \omega),\ c(B, \omega) = B \cap c(A \cup B, \omega) \tag{8}$$

for each $\omega \in \Omega$. Thus for $f \in c(A, \omega)$, $f \in c(A \cup B, \omega)$. Take $g \in B$ such that $f^u = g^u$. *Monotonicity* implies $g \in c(A \cup B, \omega)$. Then $g \in c(B, \omega)$. The first part is proved. Next, I show that, for any $f \in A$ and $g \in B$ with $f^u = g^u$, $f \in c(A, \omega)$ iff $g \in c(B, \omega)$. Suppose $f \in A$, $g \in B$ and $f^u = g^u$. If $f \in c(A, \omega)$, $f \in c(A \cup B, \omega)$ by (8). Then, by axiom *M*, $g \in c(A \cup B, \omega)$. Again by (8), $g \in c(B, \omega)$.

Turn to the second part. Suppose $c(A, \omega) \neq c(A, \omega')$ and take $f \in c(A, \omega) \backslash c(A, \omega')$. If $c(B, \omega) = c(B, \omega')$, then there is some $g \in B$ such that $g^u = f^u$ and $g \in c(B, \omega')$. But this implies $f \in c(A, \omega')$, which is a contradiction. Thus $c(A, \omega) \neq c(A, \omega')$ iff $c(B, \omega) \neq c(B, \omega')$. *IM* implies $\tau_A = \tau_B$. □

The next lemma states that the choice correspondence is translation invariant. Let $\tilde{\mathcal{A}} = \{f^u | f \in \mathcal{A}\}$ and $\tilde{\mathcal{K}} = \{A^u | A \in \mathcal{K}\}$. For $\tilde{f} \in \tilde{\mathcal{A}}$ and $\alpha \in \mathbb{R}$, let $\alpha \tilde{f} \in \tilde{A}$ be a function defined by $\alpha \tilde{f}(\omega) = \alpha(\tilde{f}(\omega))$. Similarly, for $\tilde{A} \in \tilde{\mathcal{K}}$ and $\alpha \in \mathbb{R}$, let $\alpha \tilde{A} \in \mathcal{K}$ be the set defined by $\alpha \tilde{A} = \{\alpha \tilde{f} | \tilde{f} \in \tilde{A}\}$. For $\tilde{A}, \tilde{B} \in \tilde{\mathcal{K}}$, let $\tilde{A} + \tilde{B} = \{\tilde{f} + \tilde{g} | \tilde{f} \in \tilde{A}, \tilde{g} \in \tilde{B}\}$.

Lemma 6. *If $A^u = B^u + g^u$, the followings hold:*

1. $[c(A,\omega)]^u = [c(B,\omega)]^u + g^u$.
2. $\tau_A = \tau_B$

Proof. Take $C \in \mathcal{K}$, $h \in \mathcal{A}$, and $x \in X$ such that $C^u = 2B^u$, $h = 2g^u$, and $x^u = 0$. Then, $B^u = \frac{1}{2}C^u + \frac{1}{2}x^u = (\frac{1}{2}C + \frac{1}{2}x)^u$ and $A^u = B^u + g^u = \frac{1}{2}C^u + \frac{1}{2}h^u = (\frac{1}{2}C + \frac{1}{2}h)^u$. By Lemma 5, $[c(A,\omega)]^u = [c(\frac{1}{2}C + \frac{1}{2}h, \omega)]^u$ and $[c(B,\omega)]^u = [c(\frac{1}{2}C + \frac{1}{2}x,\omega)]^u$ hold.

By *ACI*, it holds that

$$\frac{1}{2}f + \frac{1}{2}h \in c\left(\frac{1}{2}C + \frac{1}{2}h, \omega\right) \Leftrightarrow \frac{1}{2}f + \frac{1}{2}x \in c\left(\frac{1}{2}C + \frac{1}{2}x, \omega\right), \tag{9}$$

which implies $[c(\frac{1}{2}C + \frac{1}{2}h,\omega)]^u = [c(\frac{1}{2}C + \frac{1}{2}x,\omega)]^u + \frac{1}{2}h^u$. Then, $[c(A,\omega)]^u = [c(\frac{1}{2}C + \frac{1}{2}x,\omega)]^u + \frac{1}{2}h^u = [c(B,\omega)]^u + g^u$.

Turn to the second part. By equation (9), $c\left(\frac{1}{2}C + \frac{1}{2}h, \omega\right) \neq c\left(\frac{1}{2}C + \frac{1}{2}h, \omega'\right)$ iff $c\left(\frac{1}{2}C + \frac{1}{2}x, \omega\right) \neq c\left(\frac{1}{2}C + \frac{1}{2}x, \omega'\right)$. Then, *IM* implies $\tau_{\frac{1}{2}C+\frac{1}{2}x} = \tau_{\frac{1}{2}C+\frac{1}{2}h}$. Note that $B^u = (\frac{1}{2}C + \frac{1}{2}x)^u$ and Lemma 5 (2) imply $\tau_{\frac{1}{2}C+\frac{1}{2}x} = \tau_B$. Similarly, $\tau_{\frac{1}{2}C+\frac{1}{2}h} = \tau_A$ holds. Therefore $\tau_A = \tau_B$. □

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

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

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

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

by Lemma 6. Note

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

for any compact subset B of $\mathcal{A}_Y$. Once I found a representation of $\tilde{c}$ and $\tilde{\tau}$, the obtained parameters work for c and τ. In the proof of Theorem 1, I write $\mathbb{R}^\Omega$ as $\mathcal{A}$, $\tilde{c}$ as c, and $\tilde{\tau}$ as τ for simplicity.

7.1.3 Preliminary

A *plan* is a pair (F, λ) of function $F : \Omega \to \mathcal{A}$ and a $\mathcal{F}$-adapted stopping time λ, such that F is $\mathcal{F}_\lambda$-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 uses given some menu. Let $\mathbb{H}^* = \{(F, \lambda) \in \mathbb{H} | \lambda \in \mathcal{T}^*\}$, which consists of plans that can be implemented through a response time that is actually used when DM faces some menu. I especially pay attention to $\mathbb{H}^*$ because this property facilitates the calibration of γ. For $B \in \mathcal{K}$, let $\hat{c}(B) = \{F \in B^\Omega \,|\, F(\omega) \in c(B, \omega) \text{ for all } \omega\}$. Henceforth, the notation $\{f_\omega\}_{\omega\in\Omega}$ sometimes denotes for the function $\omega \mapsto f_\omega$ and sometimes for the set $\{f_\omega \,|\, \omega \in \Omega\}$.

In the next section, a preference relation over plans is defined from preference over menus. For that purpose, for each plan $(F, \lambda) \in \mathbb{H}^*$, the existence of a menu given which DM implements the plan is shown. But in general, there may not be a menu B that satisfies $F \in \hat{c}(B)$. So, as a substitute, I construct a menu with which the specified response time is implemented, and the same utility level is given at each state. Lemma 7 and Lemma 8 serves this purpose. Lemma 7 says that for any $\lambda \in \mathcal{T}^*$, there is a menu so that λ is used and utilities obtained at any state is zero. A plan $\{g_\omega\}_{\omega\in\Omega}$ is a selector of $c(B, \cdot)$ if $g_\omega \in c(B, \omega)$ for all ω.

Lemma 7. *For any $\lambda \in \mathcal{T}^*$, there exists $B_\lambda \in \mathcal{K}$ and $\{f_\omega\}_{\omega\in\Omega}$ that satisfy the followings*

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

Proof. Take any $B \in \mathcal{K}$ such that $\tau_B = \lambda$. 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 $B_\lambda = B - h$. Note that, by Lemma 6, $c(B_\lambda, \omega) = c(B, \omega) - h$ and $\tau_{B_\lambda} = \tau_B = \lambda$. Define a plan $\{f_\omega\}_{\omega\in\Omega}$ as $f_\omega = g_\omega - h$, and then the following hold: $f_\omega \in B_\lambda$, $f_\omega(\omega) = 0$, and $f_\omega \in c(B_\lambda, \omega)$. □

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

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

then write such a menu B as B^F_λ.

Lemma 8. *For any $(F, \lambda) \in \mathbb{H}^*$, there exists $B^F_\lambda \in \mathcal{K}$ and a function $\overline{F} : \Omega \to \mathcal{A}$ such that*

$$\bar{F} \in \hat{c}(B^F_\lambda),\ (\bar{F})^*(\omega) = F^*(\omega),\ \tau(B^F_\lambda, \cdot) = \lambda.$$

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

Any choice behavior requires enough information to do so. The next lemma states this fact. For each B, let $\mathrm{P}(B) = \sigma(\{\{\omega' \in \Omega | c(B,\omega'), c(B,\Omega)\} | \omega \in \Omega\})$.

Lemma 9. *For any $B \in \mathcal{K}$, $\mathrm{P}(B) \subset \mathcal{F}_{\tau_B}$.*

Proof. Suppose $c(B,\omega) \neq c(B,\omega')$. Whether $\tau(B,\omega) \geq \tau(B,\omega')$ or $\tau(B,\omega) \leq \tau(B,\omega')$, ω and ω' are distinguished at $\tau(B,\omega)$. □

7.1.4 Preference relation between plans

Here a preference relation between plans in $\mathbb{H}^*$ is defined. I start by considering preference relations between menus. 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 $A + f = \{g + f \,|\, g \in A\}$. For $F \in \mathcal{A}^\Omega$ and $f \in \mathcal{A}$, let $F + f \in \mathcal{A}^\Omega$ be defined 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 $\hat{c}(B+f) = \hat{c}(B) + f$, $F + f \in \hat{c}(B + f)$. In addition, $\mathrm{Im}\,(F + f) = (\mathrm{Im}\, F) + f \subset A + f$. The first part is proved. Suppose $A \succeq^D C_1 \succeq^D \cdots \succeq^D C_n \succeq^D B$. Then, by the first part, $(A + f) \succeq^D (C_1 + f) \succeq^D \cdots \succeq^D (C_n + f) \succeq^D (B + f)$. □

If a use of a plan (F, λ) is observed given B, it is a best plan. The second part of the next lemma states this observation. For $f, g \in \mathcal{A}$, define an act $f \wedge g \in \mathcal{A}$ by $(f \wedge g)(\omega) = \min\{f(\omega), g(\omega)\}$.

Lemma 11.

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

Proof. Let $h_\omega = F(\omega) \wedge G(\omega)$ and consider a menu $C = \{h_\omega\}_{\omega\in\Omega}$. For any $h \in C$, there exists $g \in B$ such that $g \geq h$. So *INRA* and *M* implies $B \cap c(B \cup C, \omega) = c(B, \omega)$ for any $\omega \in \Omega$. So $c(B, \omega) \neq c(B, \omega')$ implies $c(B \cup C, \omega) \neq c(B \cup C, \omega')$. Then, *IM* implies $\tau_B \leq \tau_{B\cup C}$, and $\mathcal{F}_{\tau_B} \subset \mathcal{F}_{\tau_{B\cup C}}$.

Let

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

Note that G is $\mathcal{F}_{\tau_B}$-measurable since $\mathrm{P}(B) \subset \mathcal{F}_{\tau_B}$. 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. For any $\omega' \in \Delta_\omega$,

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

Note that, for all ω, $G(\omega) \in c(B \cup C, \omega)$. Then, by $\Delta_\omega \in \mathcal{F}_{\tau_{B\cup C}}$, the equation above, and Lemma 2 (2), $h_\omega \in c(B \cup C, \omega)$ holds. Conclude $C \succeq^D B \cup C \succeq^D B$. It is easy to show that $A \succeq^I C$. Combining these shows the first part. Applying the first part to the menus $B^F_{\tau_B}$, B, and B^G_μ shows the second part. □

Elicitation of the subjective probability requires the continuity of the preference. For this reason, I 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 $A_n + f \succeq^I B_n + f$. Taking $n \to \infty$ proves 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 $C_n \to C$ with $A_n \succeq^I B_n$ and $B'_n \succeq^I C_n$. Passing to a subsequence, wlog assume $d_h(B_n, B), d_h(B'_n, B) < n^{-1}$. Then, $B_n + n^{-1} \succeq^I B'_n - n^{-1}$. Moreover, $A_n + n^{-1} \succeq^I B_n + n^{-1}$ 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. Taking $n \to \infty$ proves the second part. □

Finally, I define the preference relation over plans. For $(F, \lambda), (G, \mu) \in \mathbb{H}^*$, let

$$(F, \lambda) \succeq (G, \mu) \Leftrightarrow B^F_\lambda \succeq^* B^G_\mu.$$

And this relation is translation invariant in the following sense:

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

$$(F,\lambda) \succeq (G,\mu) \Rightarrow (F+f,\lambda) \succeq (G+f,\mu).$$

Proof. From Lemma 11 (1), $B^F_\lambda + f \sim^I B^{F+f}_\lambda$ follows. By definition, $(F,\lambda) \succeq (G,\mu)$ means $B^F_\lambda \succeq^* B^G_\mu$. This and the linearity of $\succeq^*$ imply $B^F_\lambda + f \succeq^* B^G_\mu + f$. Combine this with $B^{F+f}_\lambda \succeq^I B^F_\lambda + f$ and $B^G_\mu + f \succeq^I B^{G+f}_\mu$. Then, by the transitivity of $\succeq^*$, $B^{F+f}_\lambda \succeq^* B^{G+f}_\mu$ or $(F+f,\lambda) \succeq (G+f,\mu)$. □

7.1.5 Representation

Now I turn to the elicitation of subjective probability and cost function. The first is the 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 $\int f d\pi \geq (>) \int g d\pi$. For any ω, because $1\omega 0 \succ^* 0$, $\pi(\omega) > 0$ holds. That is, π is full-support. □

Next, I elicit the cost function on $\mathcal{T}^*$ and one-way utility representation on $\mathbb{H}^*$. The idea of calibration is as follows. If $(F,\lambda) \succeq (G,\mu)$ and λ is more costly than μ, the utility from F compared to G is large enough to compensate the cost increase. The difference of utility between F and G is evaluated using π under expected utility criterion. For a moment, denote f_λ for a plan $(F,\lambda) \in \mathbb{H}^*$ where $F \in \hat{c}(B^F_\lambda)$ and $F^* = f$. In the proof of the next lemma, I denote $\pi(f)$ for the integration $\int f d\pi$.

Lemma 15.

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

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

where

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

2. For $\lambda, \mu \in \mathcal{T}^*$, if $\lambda \le \mu$, then $\gamma^*(\lambda) \le \gamma^*(\mu)$.

Proof. By *IM*, stopping time for singleton menus $\{f\}$ are the same. Take some f and let $\underline{\lambda} = \tau_{\{f\}}$. Let $M_{\lambda,\mu} = \{f \in \mathcal{A} | f_\lambda \succeq 0_\mu\}$ and let

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

Claim. $\inf_{f \in M_{\lambda,\mu}} \int f d\pi \ge \gamma^*(\lambda) - \gamma^*(\mu)$

$\vdash$ Take $g_n \in M_{\mu,\underline{\lambda}}$ with $\pi(g_n) \to \gamma^*(\mu)$ and $h_n \in M_{\lambda,\mu}$ with $\pi(h_n) \to \inf_{h \in M_{\lambda,\mu}} \int h d\pi$. Since $[g_n]_\mu \succeq 0_{\underline{\lambda}}$, $0_\mu \succeq [-g_n]_{\underline{\lambda}}$. Combining with $[h_n]_\lambda \succeq 0_\mu$, I have $[h_n]_\lambda \succeq [-g_n]_{\underline{\lambda}}$ or $[g_n + h_n]_\lambda \succeq 0_{\underline{\lambda}}$. Thus,

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

That is, $\inf_{h \in M_{\lambda,\mu}} \int h d\pi \ge \gamma^*(\lambda) - \gamma^*(\mu)$. $\dashv$

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

$$\int F^* - G^* d\pi \ge \gamma^*(\lambda) - \gamma^*(\mu),$$

or $V(F, \lambda) \ge V(G, \mu)$. This shows the first part.

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

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

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

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

Here, $\mathcal{T}^*(\lambda) = \{\tilde{\lambda} \in \mathcal{T}^* | \tilde{\lambda} \geq \lambda\}$. Moreover, γ is increasing; $\gamma(\lambda) \geq \gamma(\mu)$ whenever $\lambda \geq \mu$.

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

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

Take the supremum of the right hand side. □

The next lemma states that benefit from any implementable choice is bounded by an optimal choice.

Lemma 17. *For $B \in \mathcal{K}$ and an algebra Q of Ω, the following hold:*

1. *If $F \in B^\Omega$ and F is Q-measurable,*

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

2. *There is some $G \in B^\Omega$ that satisfies*

$$\mathrm{E}[G^*|Q](\omega) = \max_{f \in B} \mathrm{E}[f|Q](\omega).$$

Proof. Let $\{Q(\omega) | \omega \in \Omega\} = \{\Delta_1, \ldots, \Delta_I\}$. For the first part, observe

$$\mathrm{E}[F^*] = \sum_{i=1}^{I} \pi(\Delta_i) \mathrm{E}[F^*|\Delta_i] \leq \sum_i \pi(\Delta_i) \max_{f \in B} \mathrm{E}[f|\Delta_i] = \mathrm{E}[\max_{f \in B} \mathrm{E}[f|Q]].$$

For the second part, take any $g_i \in \arg\max_{f \in B} \mathrm{E}[f|\Delta_i]$ for each i and define $G \in B^\Omega$ by $G(\omega) = g_i$ for $\omega \in \Delta_i$. Then, for $\omega \in \Delta_i$,

$$\mathrm{E}[G^*|Q](\omega) = \mathrm{E}[G^*|\Delta_i] = \mathrm{E}[g_i|\Delta_i] = \max_{f \in B} \mathrm{E}[f|Q](\omega).$$

□

Now, I show the optimality of implemented response times.

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

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

Proof. Let $F \in \hat{c}(B)$. Take any $(G, \mu) \in \mathbb{H}$ with $\mathrm{Im}G \subset B$ and $\mathrm{E}[G^*|\mathcal{F}_\mu] = \max_{f \in B} \mathrm{E}[f|\mathcal{F}_\mu]$. Then, by Lemma 16 and Lemma 17,

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

Note μ is arbitrary. □

Finally, I 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, I show $c(B, \omega) \subset \arg\max_{f \in B} \mathrm{E}[f|\mathcal{F}_{\tau_B}](\omega)$ for all $\omega \in \Omega$. Take any ω and $f \in c(B, \omega)$. 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 $\mathcal{F}_{\tau_B}$-measurable selector F of $c(B, \cdot)$ that satisfies $F(\omega) = f$ and define a plan G as

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

By construction, $\mathrm{E}[G^*] > \mathrm{E}[F^*]$. 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, I 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 $\{\mathrm{P}(B)(\omega)|\omega \in \Omega\} = \{\Delta_0, \ldots, \Delta_I\}$ where $\Delta_0 = \mathrm{P}(B)(\omega^*)$. Take a $\mathrm{P}(B)$-measurable $F \in \hat{c}(B)$ 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 $\omega \in \Delta_i$. I show that $\hat{c}(B^n) = \{G^n\}$.

Note that $g_i \in \arg\max_{h \in B} \mathrm{E}[h|\Delta_i]$ by the inclusion proved before. Since $g_i^n(\omega) = g_i(\omega)$ for $\omega \in \Delta_i$, $g_i^n \in \arg\max_{h \in B^n} \mathrm{E}[h|\Delta_i]$. To show that g_i^n is the only maximizer, take any $g \in B^n \backslash \{g_i^n\}$. Then, there is some $g' \in B$ that satisfies $g'(\omega) - n^{-1} = g(\omega)$ for $\omega \in \Delta_i$. So

$$\mathrm{E}[g|\Delta_i] = \mathrm{E}[g'|\Delta_i] - n^{-1} < \mathrm{E}[g_i|\Delta_i] = \mathrm{E}[g_i^n|\Delta_i],$$

which shows $\arg\max_{h \in B^n} \mathrm{E}[h|\Delta_i] = \{g_i^n\}$. Next, I show $g_i^n \neq g_j^n$ for $i \neq j$. If $g_i^n = g_j^n$, $g_i(\omega) = g_j(\omega) + n^{-1}$ for $\omega \in \Delta_j$. So $\mathrm{E}[g_i|\Delta_j] = \mathrm{E}[g_j|\Delta_j] + n^{-1}$, which is a contradiction. Thus, $\mathrm{P}(B^n) = \mathrm{P}(B)$. Then, by *IM*, $\tau_{B_n} = \tau_B$ for all n. Note $g_i^n \to g_i$ and $B^n \to B$, and apply *C* (1) to obtain $g_i \in c(B, \omega)$. □

Lemma 18 and 19 prove Theorem 1.

7.2 Proof of Theorem 2

Fix an OSR $(u, \pi, \mathcal{F}, \gamma)$. First I construct a data (c, τ) that satisfies M. To this end, two lemmas are necessary.

Lemma 20. *For any algebra Q, if there is some $\lambda_0 \in \mathcal{T}$ such that $Q \subset \mathcal{F}_{\lambda_0}$, the function*

$$\lambda(\omega) = \min\{t \in \mathbb{R}_+ \,|\, \mathcal{F}_t(\omega) \subset Q(\omega)\}$$

is a well-defined $\mathcal{F}$-adapted stoping time. Moreover, for any $\mathcal{F}$-adapted stopping time μ, $\lambda \leq \mu$ if $Q \subset \mathcal{F}_\mu$.

Proof. Take any ω. By the assumption, there is some $t \geq 0$ with $\mathcal{F}_t(\omega) \subset Q(\omega)$. Let $t^* = \inf\{t \in \mathbb{R}_+ | \mathcal{F}_t(\omega) \subset Q(\omega)\}$. Let $t_n \downarrow t^* \geq 0$ and $\mathcal{F}_{t_n}(\omega) \subset Q(\omega)$. Since Ω is finite, $\mathcal{F}_{t_n}(\omega) = \mathcal{F}_{t^*}$ for sufficiently large n. Then, $\mathcal{F}_{t^*}(\omega) \subset Q(\omega)$. That is, $t^* = \min\{t \in \mathbb{R}_+ | \mathcal{F}_t(\omega) \subset Q(\omega)\}$ and λ is well-defined.

Next, I show that λ is $\mathcal{F}$-adapted. Note

$$\lambda(\omega) \leq t \Leftrightarrow \exists s \leq t,\ \mathcal{F}_s(\omega) \subset Q(\omega) \Leftrightarrow \omega \in \bigcup_{s \in [0,t]} \{\omega \,|\, \mathcal{F}_s(\omega) \subset Q(\omega)\}$$

and

$$\{\omega \,|\, \mathcal{F}_s(\omega) \subset Q(\omega)\} = \bigcup \{\mathcal{F}_s(\omega) \,|\, \omega \in \Omega,\ \mathcal{F}_s(\omega) \subset Q(\omega)\}.$$

So $\{\lambda \leq t\}$ is written as a union of elements of $\mathcal{F}_t$. That is, λ is $\mathcal{F}$-adapted. The inclusion $Q \subset \mathcal{F}_\lambda$ is obvious. For the last part, suppose $\mu(\omega) < \lambda(\omega)$. By definition of λ, $\mathcal{F}_{\mu(\omega)}(\omega) \not\subset Q(\omega)$ and thus $Q \not\subset \mathcal{F}_\mu$. □

For each $B \in \mathcal{K}$, $\lambda \in \mathcal{T}$, and $\omega \in \Omega$, let $\tilde{c}(B, \lambda, \omega) = \arg\max_{f \in B} \mathrm{E}[u(f)|\mathcal{F}_\lambda](\omega)$ and let $\tilde{\mathrm{P}}(B, \lambda)$ be the algebra generated by $\omega \mapsto \tilde{c}(B, \lambda, \omega)$.

Lemma 21. *For any $B \in \mathcal{K}$ and $\lambda, \mu \in \mathcal{T}$, if $\tilde{P}(B, \mu) \subset \mathcal{F}_\lambda$ and $\lambda \leq \mu$, then $\tilde{c}(B, \lambda, \omega) = \tilde{c}(B, \mu, \omega)$ for any ω.*

Proof. Fix any $\omega \in \Omega$. First, I show $\tilde{c}(B, \mu, \omega) \subset \tilde{c}(B, \lambda, \omega)$. Take any $f \in \tilde{c}(B, \mu, \omega)$. Note that $f \in \tilde{c}(B, \mu, \omega')$ for any $\omega' \in \mathcal{F}_\lambda(\omega) \subset \tilde{P}(B, \mu)(\omega)$ by assumption. Then, for any $g \in B$,

$$\mathrm{E}[u(f)|\mathcal{F}_\lambda](\omega) = \mathrm{E}[\mathrm{E}[u(f)|\mathcal{F}_\mu]|\mathcal{F}_\lambda](\omega) \geq \mathrm{E}[\mathrm{E}[u(g)|\mathcal{F}_\mu]|\mathcal{F}_\lambda](\omega) = \mathrm{E}[u(g)|\mathcal{F}_\lambda](\omega).$$

That is, $f \in \tilde{c}(B, \lambda, \omega)$.

Next, I show $\tilde{c}(B, \mu, \omega) \supset \tilde{c}(B, \lambda, \omega)$. Suppose $f \notin \tilde{c}(B, \mu, \omega)$ and take $g \in \tilde{c}(B, \mu, \omega)$. By assumption, for any $\omega' \in \mathcal{F}_\lambda(\omega) \subset \tilde{\mathrm{P}}(B, \mu)(\omega)$, $g \in \tilde{c}(B, \mu, \omega')$ hold. Then,

$$\mathrm{E}[u(g)|\mathcal{F}_\lambda](\omega) = \mathrm{E}[\mathrm{E}[u(g)|\mathcal{F}_\mu]|\mathcal{F}_\lambda](\omega) > \mathrm{E}[\mathrm{E}[u(f)|\mathcal{F}_\mu]|\mathcal{F}_\lambda](\omega) = \mathrm{E}[u(f)|\mathcal{F}_\lambda](\omega)$$

and so $f \notin \tilde{c}(B, \lambda, \omega)$. □

I construct a data that satisfies *IM*. For each $B \in \mathcal{K}$, let $\tilde{\mathbb{P}}(B) = \{\tilde{\mathrm{P}}(B, \lambda) \,|\, \lambda \in \arg\max_{\lambda \in \mathcal{T}} \mathrm{E}[\max_{f \in B} \mathrm{E}[u(f)|\mathcal{F}_\lambda]] - \gamma(\lambda)\}$. Let $\tilde{\mathbb{P}}$ be the set of algebras Q such that there exists some $\mu \in \mathcal{T}$ with $Q \subset \mathcal{F}_\mu$ and enuemrate it as $\tilde{\mathbb{P}} = \{Q_1, \ldots, Q_n\}$. Let λ_i be the minimal stopping time that satisfies $Q_i \subset \mathcal{F}_{\lambda_i}$, whose existence is guaranteed by Lemma 20.

Let $\mathcal{K}_1 = \{B \in \mathcal{K} \,|\, Q_1 \in \tilde{\mathbb{P}}(B)\}$. For $i = 2, \ldots, n$, let $\mathcal{K}_i = \{B \in \mathcal{K} \,|\, Q_i \in \tilde{\mathbb{P}}(B)\} \setminus \bigcup_{j=1}^{i-1} \mathcal{K}_j$. Then, $\mathcal{K}_1, \ldots, \mathcal{K}_n$ partition $\mathcal{K}$.

I show that

$$\lambda_i \in \arg\max_{\lambda \in \mathcal{T}_\mathcal{F}} \mathrm{E}[\max_{f \in B} \mathrm{E}[u(f)|\mathcal{F}_\lambda]] - \gamma(\lambda). \tag{10}$$

for $B \in \mathcal{K}_i$. Take any $\mu \in \arg\max_{\lambda \in \mathcal{T}_\mathcal{F}} \mathrm{E}[\max_{f \in B} \mathrm{E}[u(f)|\mathcal{F}_\lambda]] - \gamma(\lambda)$ with $\tilde{\mathrm{P}}(B, \mu) = Q_i$. Then, because $Q_i \subset \mathcal{F}_{\lambda_i}$ and $\lambda_i \leq \mu$, Lemma 21 shows $\tilde{c}(B, \lambda_i, \omega) = \tilde{c}(B, \mu, \omega)$ for any ω. Then, since $\lambda_i \leq \mu$, (10) holds.

For $B \in \mathcal{K}_i$, let $\tau_B = \lambda_i$ and $c(B, \omega) = \tilde{c}(B, \lambda_i, \omega)$. Then, the OSR represents(c, τ) satisfies the requirement. I already showed that (c, τ) is represented by the given OSR. Suppose $\mathrm{P}(A) \subset \mathrm{P}(B)$. Then, by construction, τ_A is the minimal stopping time such that $\mathrm{P}(A) \subset \mathcal{F}_{\tau_A}$ and thus $\tau_A \leq \tau_B$, which is *IM*.

Next, I show (c, τ) satisfies the rest of axioms in the first part.

ACI: By the construction of τ, $\mathcal{F}_{\tau_{\alpha g + (1-\alpha)B}} = \mathcal{F}_{\tau_{\alpha h + (1-\alpha)B}}$.

$$\begin{aligned} \alpha g + (1-\alpha) f \in c(\alpha g + (1-\alpha)B, \omega) &\Leftrightarrow \alpha g + (1-\alpha) f \in \arg\max_{\tilde{f} \in \alpha g + (1-\alpha)B} \mathrm{E}[u(\tilde{f})|\mathcal{F}_{\tau_{\alpha g + (1-\alpha)B}}]](\omega) \\ &\Leftrightarrow \alpha h + (1-\alpha) f \in \arg\max_{\tilde{f} \in \alpha h + (1-\alpha)B} \mathrm{E}[u(\tilde{f})|\mathcal{F}_{\tau_{\alpha h + (1-\alpha)B}}]](\omega) \\ &\Leftrightarrow \alpha h + (1-\alpha) f \in c(\alpha h + (1-\alpha)B, \omega). \end{aligned}$$

DSC: Suppose $f, g \in B$, $\omega \in \Delta \in \mathcal{F}_{\tau_B}$, and $f(\omega') = g(\omega')$ for any $\omega' \in \Delta$. Assume $f \in c(B, \omega)$. Then, by the definition of OSR, $f \in \arg\max_{h \in B} \mathrm{E}[u(h)|\mathcal{F}_{\tau_B}](\omega)$. But

$$\mathrm{E}[u(f)|\mathcal{F}_{\tau_B}](\omega) = \int_\Delta u(f) \mathrm{d}\pi(\cdot|\Delta) = \int_\Delta u(g) \mathrm{d}\pi(\cdot|\Delta) = \mathrm{E}[u(g)|\mathcal{F}_{\tau_B}](\omega).$$

So, $g \in \arg\max_{h\in B} \mathrm{E}[u(h)|\mathcal{F}_{\tau_B}](\omega) = c(B,\omega)$.

C(1): Suppose $B_n \to B$, $f_n \to f$, $f_n \in c(B_n, \omega)$ and

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

Then, $\mathcal{F}_{\tau_{B_n}} = \mathcal{F}_{\tau_B}$ for all n.

Since $f_n \in c(B_n, \omega)$, $f_n \in \arg\max_{h\in B_n} \mathrm{E}[u(h)|\mathcal{F}_{\tau_{B_n}}](\omega)$. Letting $\Delta = \mathcal{F}_{\tau_B}(\omega)$, this can be written as

$$\int_\Delta u(f_n)\mathrm{d}\pi(\cdot|\Delta) = \max_{h\in B_n} \int_\Delta u(h)\mathrm{d}\pi(\cdot|\Delta).$$

Let $n \to \infty$. Then, the LHS converges to $\int_\Delta u(f)\mathrm{d}\pi(\cdot|\Delta)$ By Berge maximum theorem (Aliprantis and Border (2006), p. 570), the RHS converges to $\max_{h\in B} \int_\Delta u(h)\mathrm{d}\pi(\cdot|\Delta)$. Thus $f \in \arg\max_{h\in B} \mathrm{E}[u(h)|\Delta](\omega) = c(B,\omega)$.

C (2): Let $U(B) = \max_{\lambda\in\mathcal{T}} \mathrm{E}[\max_{f\in B} \mathrm{E}[u(f)|\mathcal{F}_\lambda]] - \gamma(\lambda)$. I show $A \succeq^I B$ implies $U(A) \geq U(B)$. For this, it is enough to show $A \succeq^D B$ implies $U(A) \geq U(B)$. Suppose $F \in \hat{c}(B)$ and $\mathrm{Im}F \subset A$. Then, $U(A) = \max_\lambda \mathrm{E}[\max_{f\in A} \mathrm{E}[u(f)|\mathcal{F}_\lambda]] - \gamma(\lambda) \geq \mathrm{E}[F^*] - \gamma(\tau_B) = U(B)$.

Now turn to the proof of the necessity of C (2). For any singleton menu $\{f\}$, $U(\{f\}) = \mathrm{E}[u(f)]$. Moreover, $U(\{f_n\}) \to U(\{f\})$ when $f_n \to f$ by the continuity of u. Suppose $x > y$, $f_n \to x\omega y$, and $g_n \to y$. En route to a contradiction, assume $\{g_n\} \succeq^I \{f_n\}$ for all n. Then $U(\{g_n\}) \geq U(\{f_n\})$ holds for all n. Let $n \to \infty$ and obtain $U(\{y\}) \geq U(\{x\omega y\})$ or $\mathrm{E}[y] \geq \mathrm{E}[x\omega y]$. However, since π is full-support, $E[y] < E[x\omega y]$. Contradiction.

M and U follows from Theorem 2 in Ellis since c has an OIR. The second part is also a direct implication of Theorem 2 in Ellis. □

7.3 Proof of Proposition 1

Lemma 22. *Suppose (c,τ) has an OSR $(u,\pi,\mathcal{G},\gamma)$. If $\tau(B,\omega), \tau(B,\omega') \leq t$ and $\omega' \in \mathcal{G}_t(\omega)$, then $c(B,\omega) = c(B,\omega')$ and $\tau(B,\omega) = \tau(B,\omega')$.*

Proof. Observe $\omega' \in \mathcal{G}_t(\omega) \subset \mathcal{G}_{\tau_B(\omega)}(\omega) = \mathcal{G}_{\tau_B}(\omega)$. By the symmetric argument, $\mathcal{G}_{\tau_B}(\omega) = \mathcal{G}_{\tau_B}(\omega')$. Because τ_B is $\mathcal{G}$-adapted, $\tau_B(\omega) = \tau_B(\omega')$. For any $f \in B$,

$$c(B,\omega) = \mathrm{E}[u(f)|\mathcal{G}_{\tau_B}](\omega) = \mathrm{E}[u(f)|\mathcal{G}_{\tau_B}](\omega') = c(B,\omega').$$

□

Lemma 23. *Suppose (c,τ) has an OSR $(u,\pi,\mathcal{G},\gamma)$. If $\omega' \in \mathcal{G}_t(\omega)$, then either $\tau(B,\omega), \tau(B,\omega') \leq t$ or $t < \tau(B,\omega), \tau(B,\omega')$ holds.*

Proof. Note that τ_B is $\mathcal{G}$-adapted for any $B \in \mathcal{K}$, or $\{\tau_B \le t\} \in \sigma(\mathcal{G}_t)$ for any $t \ge 0$. If $\omega \in \{\tau_B \le t\}$, then $\mathcal{G}_t(\omega) \subset \{\tau_B \le t\}$. So $\omega' \in \mathcal{G}_t(\omega)$ implies $\omega, \omega' \in \{\tau_B \le t\}$. If $\omega \notin \{\tau_B \le t\}$, then $\mathcal{G}_t(\omega) \cap \{\tau_B \le t\} = \emptyset$. In this case, $\omega, \omega' \notin \{\tau_B \le t\}$. In both cases, the conclusion of the statement holds. □

Proof of Proposition 1

En route to a contradiction, assume the proposition does not hold. Then, there is some $\Delta \in \mathcal{F}_t \backslash \mathcal{G}_t$ for some t. Because $\mathcal{F}_t = \sigma(\{\{\omega' | \omega' \bowtie_t \omega\} | \omega \in \Omega\})$, any $\omega \in \Delta$ and $\omega' \in \Delta^c$ are distinguished until t. Since $\Delta \notin \mathcal{G}_t$, for some $\omega \in \Delta$ and $\omega' \in \Delta^c$, $\omega' \in \mathcal{G}_t(\omega)$ holds. Otherwise, for any $\omega \in \Delta$, $\mathcal{G}_t(\omega) \subset \Delta$ and so $\Delta = \bigcup_{\omega \in \Delta} \mathcal{G}_t(\omega)$ and contradicts to the assumption $\Delta \notin \sigma(\mathcal{G}_t)$. Fix such ω and ω'. Then, there exists a menu B that satisfies the followings:

1. $\tau(B, \omega) \wedge \tau(B, \omega') \le t$,
2. $\tau(B, \omega) \ne \tau(B, \omega')$ or $c(B, \omega) \ne c(B, \omega')$.

By the assumption and Lemma 23, $\tau(B, \omega), \tau(B, \omega') \le t$. Then, by Lemma 22, $c(B, \omega) = c(B, \omega')$ and $\tau(B, \omega) = \tau(B, \omega')$. A contradiction. □

7.4 Proof of Proposition 2

For the OSR $(u, \pi, \mathcal{F}, \gamma)$, define the corresponding OIR $(u, \pi, \hat{\gamma}, \hat{\mathrm{P}})$ by

$$\hat{\mathrm{P}}(B) = \mathcal{F}_{\tau_B},$$
$$\hat{\gamma}(Q) = \begin{cases} \inf\{\gamma(\lambda) | \lambda \in \mathcal{T},\ Q \subset \mathcal{F}_\lambda\} & \text{if } \exists \lambda \in \mathcal{T} \text{ s.t. } Q \subset \mathcal{F}_\lambda \\ \infty & \text{otherwise.} \end{cases}$$

Similary, for the OSR $(u', \pi', \mathcal{F}', \gamma')$ define an OIR $(u', \pi', \hat{\gamma}', \hat{\mathrm{P}}')$.

By Theorem 3 in Ellis, the following hold:

1. $\pi = \pi'$,
2. there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $u' = \alpha u + \beta$ and $\hat{\gamma}'(Q) = \alpha \hat{\gamma}(Q)$.

I show the uniqueness of cost function. By the optimality of τ_B in OSR, τ_B takes the smallest cost to acquire $\mathcal{F}_{\tau_B}$ and so $\hat{\gamma}(\mathcal{F}_{\tau_B}) = \gamma(\tau_B)$. Then,

$$\gamma'(\tau_B) = \hat{\gamma}'(\mathcal{F}_{\tau_B}) = \alpha \hat{\gamma}(\mathcal{F}_{\tau_B}) = \alpha \gamma(\tau_B).$$

7.5 Proof of Proposition 3

Consider the first part. Suppose (c, τ) has an GOSR $(u, \pi, \{\mathcal{F}^B\}_{B\in\mathcal{K}}, \gamma)$. Define

$$\hat{\mathrm{P}}(B) = \mathcal{F}^B_{\tau_B},$$
$$\hat{\gamma}(Q) = \begin{cases} \inf\{\gamma(\mathcal{F},\lambda) | (\mathcal{F},\lambda) \in \Pi,\ Q \subset \mathcal{F}_\lambda\} & \text{if } \exists(\mathcal{F},\lambda) \in \Pi \text{ s.t. } Q \subset \mathcal{F}_\lambda, \\ \infty & \text{otherwise.} \end{cases}$$

for each $B \in \mathcal{K}$ and $Q \in \mathbb{P}$. I show that $(u, \pi, \hat{\gamma}, \hat{\mathrm{P}})$ is an OIR of c. The second requirement of the OIR directly follows from that of OSR and the definition of $\hat{\mathrm{P}}$. For notational simplicity, let $\hat{U}(B, Q) = \mathrm{E}[\max_{f\in B} \mathrm{E}[u(f)|Q]] - \hat{\gamma}(Q)$ and $U(B, \mathcal{F}, \lambda) = \mathrm{E}[\max_{f\in B} \mathrm{E}[u(f)|\mathcal{F}_\lambda]] - \gamma(\lambda)$. Consider the first requirement of OIR. Take any $Q \in \mathbb{P}$. If there is no $(\mathcal{G}, \mu) \in \Pi$ such that $Q \subset \mathcal{G}_\mu$, then $\hat{\gamma}(Q) = \infty$ and so $\hat{U}(B, \hat{\mathrm{P}}(B)) \geq \hat{U}(B, Q)$. Otherwise, note that $U(B, \mathcal{F}^B, \tau_B) \geq U(B, \mathcal{G}, \mu)$ for any $(\mathcal{G}, \mu) \in \Pi$. So,

$$\hat{U}(B, \hat{\mathrm{P}}(B)) = U(B, \mathcal{F}^B, \tau_B) \geq \max_{\substack{(\mathcal{G},\mu)\in\Pi;\\ Q\subset\mathcal{G}_\mu}} U(B, \mathcal{G}, \mu) = \hat{U}(B, Q),$$

where the equalities follow from the definition of $\hat{\mathrm{P}}(\cdot)$ and $\hat{\gamma}(\cdot)$. The first part is proved.

Turn to the second part. Let (c, τ) be any data such that c has an OIR. Suppose $(u, \pi, \hat{\gamma}, \hat{\mathrm{P}})$ is an OIR of c. Let $\mathcal{F}^B_t = \hat{\mathrm{P}}(B)$ for all $B \in \mathcal{K}$ and $t \geq 0$. Let $\gamma(\mathcal{F}, \lambda) = \hat{\gamma}(\mathcal{F}_0)$ for any $\mathcal{F}, \lambda$. Take any $(\mathcal{G}, \lambda) \in \Pi$, and let $Q = \mathcal{G}_0$. Then,

$$\gamma(\mathcal{F}^B, \tau_B) = \hat{\gamma}(\mathcal{F}^B_0) = \hat{\gamma}(\hat{\mathrm{P}}(B)) \leq \hat{\gamma}(Q) = \hat{\gamma}(\mathcal{G}_0) = \gamma(\mathcal{G}, \lambda)$$

and therefore,

$$U(B, \mathcal{F}^B, \tau_B) = \hat{U}(B, \hat{\mathrm{P}}(B)) \geq \hat{U}(B, Q) = U(B, \mathcal{G}, \lambda).$$

This is condition (4) of GOSR. Condition (5) is trivial.

Turn to the third part. Let $\Omega = \{\omega_1, \omega_2\}$ and $X = \mathbb{R}$. Let c be the choice correspondence defined by the OIR defined by the following parameters: $u(x) = x$, $\pi = (1/2, 1/2)$, $\gamma(\{\Omega\}) = 0$, and $\gamma(\{\{\omega_1\}, \{\omega_2\}\}) = 1$. Let $\tau_B = 0$ for all $B \in \mathcal{K}$. Then, (c, τ) does not have an OSR. Since $\gamma(\{\{\omega_1\}, \{\omega_2\}\}) < \infty$, there is some $B \in \mathcal{K}$ such that $c(B, \omega_1) \neq c(B, \omega_2)$. Since $\tau_B = 0$, $\mathcal{F}_0 = \{\{\omega_1\}, \{\omega_2\}\}$. Next, let $f_i^\epsilon = \epsilon_{\{\omega_i\}}0$ for $\epsilon > 0$, and $i = 1, 2$. I consider a menu $B^\epsilon = \{f_1^\epsilon, f_2^\epsilon\}$. Then, for ϵ small enough, $c(B^\epsilon, \omega_1) = c(B^\epsilon, \omega_2)$ by the condition (6) of OIR. However, if (c, τ) has an OSR, $c(B^\epsilon, \omega_i) = \{f_i^\epsilon\}$ for $i = 1, 2$. Contradiction. □

References

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

Baldassi, C., Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M. 2018. A behavioral characterization of the Drift Diffusion Model. Working paper.

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

Chambers, C. P., Liu, C., Rehbeck, J. 2020. Costly information acquisition. Journal of Economic Theory, 186, p. 104979.

de Oliveira, H., Denti, T., Mihm, M., Ozbek, K. 2017. Rationally inattentive preferences and hidden information costs: Rationally inattentive preferences. Theoretical Economics, 12, 621–654.

de Oliveira, H., Lamba, R. 2019. Rationalizing Dynamic Choices. Working Paper.

Dillenberger, D., Krishna, V. R., Sadowski, P. 2018. Subjective Information Choice Processes. Working Paper.

Duraj, J., Lin, Y.-H. 2019. Identification and Welfare Analysis in Sequential Sampling Models. Working paper.

Echenique, F., Saito, K. 2017. Response time and utility. Journal of Economic Behavior & Organization, 139, 49–59.

Ellis, A. 2018. Foundations for optimal inattention. Journal of Economic Theory, 173, 56–94.

Higashi, Y., Hyogo, K., Takeoka, N. 2020. Costly Subjective Learning. Working Paper, p. 54.

Hyogo, K. 2007. A subjective model of experimentation. Journal of Economic Theory, 133, 316–330.

Koida, N. 2017. A multiattribute decision time theory. Theory and Decision, 83, 407–430.

Sims, C. A. 2003. Implications of rational inattention. Journal of Monetary Economics, 50, 665–690.

Takeoka, N. 2007. Subjective probability over a subjective decision tree. Journal of Economic Theory, 136, 536–571.