A Theory-Based Decision Model

Edi Karni[*,†]
Johns Hopkins University

December 21, 2020

Abstract

The paper proposes a theory-based model of decision making under uncertainty. According to this model predictions of the outcomes of acts are derived from theories. Realized act-outcome pairs are used to update the decision-makers beliefs regarding the validity of the relevant theories. Consequently, acts are, simultaneously, initiatives that have material consequences and information generating experiments. Pure experiments (that is, information generating acts devoid of direct material consequences), are characterized and the value of information they generate is defined in terms of their informational structures. An incentive-compatible mechanism is introduced, by which the subjective probabilities the decision-makers holds regarding the validity of the theories are elicited.

Keywords: Theory-based decisions, experimentation, value of information, subjective probabilities, probability elicitation.

JEL classification: D8, D81, D83

[*]Johns Hopkins University, Department of Economics E-mail: karni@jhu.edu.
[†]I am grateful to Itzhak Gilboa for his useful comments.

"On principle, it is quite wrong to try founding a theory on observable magnitudes alone. In reality the very opposite happens. It is the theory which decides what we can observe." – Albert Einstein in a conversation with Werner Heisenberg.

1 Introduction

The objective of this paper is twofold: First, to develop a theory-based decision model and to analyze its implications for individual choice under uncertainty. Second, develop a theory of choice of experiments. Two traits of human nature are pertinent for this purpose. The first is the inclination to distill, in the form of general laws, regularities in empirical observations and to invoke these laws when making predictions. The second is the curiosity-driven inclination to the explore, by means of observation and experimentation, the physical and social surroundings.

In the context of decision making under uncertainty, theories, a term that I use interchangeably with hypotheses or propositions, are such general laws that decision makers use, explicitly or implicitly, to predict the outcomes of their actions. The main premise of this paper is that, in many situations, decision makers choose among acts on the basis of the predictions of what they consider to be relevant theories weighted by their beliefs in the validity of these theories or propositions.[1] Observations that, in this context, consist of act-outcome pairs are used to update the decision makers' aforementioned beliefs. Consequently, a theory is modified or discarded when it is falsified by observations, which happens when an act results in an outcome that is outside the set of outcomes that, according to the theory, are possible given the act. Observations that are consistent with the existing theories induce update the weights that the decision makers assign to the predictions of the theories.

It is common to distinguish between acts and experiments. Acts are initiatives that have material consequences and are devoid of informational content, and experiments are information-generating initiatives devoid of direct material implications. This dichotomy is an idealized simplification. In reality, acts result in outcomes that inform decision makers about the validity of the theories and are, therefore, also experiments of a kind. Experiments may have material implications (e.g., cost) and are, therefore, acts of a kind. In

[1] In this sense, this paper extends Ramsey's (1926) pioneering essay "Truth and Probabilities," which explored the choice-based foundations of the degrees of belief in the truth of propositions.

this paper, I develop a general theory of choice that amalgamates the material and informational aspects of acts and consider the choice of experiments as a special case of this theory.

Unlike Savage's vision of a grand world "theory of personal probability and utility is, as I see it, a sort of framework into which I hope to fit a large class of decision problems." (Drèze 1987, p. 78) the theory that I propose is context-dependent. According to this approach, it is the context that lends the different components their specific interpretations. For instance, in the context of medical decisions, theories predict the probable states of health (i.e., outcomes) that would result from treatments (i.e., acts). In the context of financial decision, theories are models of the financial markets that predict the probable values (i.e., outcomes) of different portfolios positions (i.e., acts).

To grasp the ideas to be explored, consider the following simple example. An urn is known to contain 100 ball 50 of which are red and 50 are black. Two balls are drawn sequentially, at random, and their colors observed. A decision maker can place a bet on the event A, that the two balls are of the same color, or on the complementary event A^c, the two balls are of different colors. The decision maker entertains two hypotheses regarding the random process generating the outcomes. Hypothesis I is that the draws are with replacement; hypothesis II is that the draws are without replacement. The two hypotheses imply distinct distributions on the two events. According to hypothesis I, the two events are equally probable; according to hypothesis II, the probability of event A is $49/99$ and that of A^c is $50/99$. Suppose that the decision maker's prior degrees of belief regarding the validity of the two hypotheses are p and $1-p$, respectively. If the underlying stochastic process is decided by nature, than the two hypotheses may be regarded as "states of nature." According to this interpretation, states of nature are abstract general laws, or theories, depicting of nature (e.g., Einstein's general relativity and Newton's mechanics). In this scenario, the stochastic process takes place behind a "veil of ignorance" and decision makers gets to observe only the colors of the balls. Acts are bets on the color combinations of the balls and their payoffs are their material consequences. In addition, because they require that the color combinations that obtain be verifiable, betting generate information that enable the decision maker to update his belief about the validity of the two hypotheses.

Much of the theory of decision making under uncertainty focuses on identifying betting patterns that would allow an observer to infer a decision maker's beliefs regarding the likely realization of the events, such as A and A^c, and quantify those beliefs by probabilities. The question I seek to answer here is whether an observer can elicit, using incentive-compatible procedures, the decision makers subjective probabilities of these events an using these probabilities to recover the probabilities the decision maker assigns the underlying hypotheses. For example, suppose that implementing one of the available elicitation schemes (e.g., Karni [2009]), the observer concludes that the decision maker's subjective probability of the event A is 0.496. The observer can recover the probabilities p by solving the equation $[99p + (1-p)\,98]/198 = 0.496$.[2]

The following examples serve to set the stage and buttress the argument in favor of the proposed model.

Education signals: An employer who regularly hires employees is interested in their productivity. Employees' productivity is idiosyncratic cannot be ascertained except by actually employing them. Suppose, for simplicity, that the employer entertains alternative hypotheses regarding the relationship between an employee's productivity and her education level. Distinct hypotheses maintain that the employees' education levels and productivity are positively correlated but to different degrees. The employer holds a prior belief about the validity of the alternative hypotheses, based on which she decides to required a certain level of education to fill a job vacancy. After having observed the productivity of the employee, the employer updates her belief about the validity of the underlying hypotheses, which she than relies upon the next time she looks to fill a job vacancy.

Medical decisions: A patient showing certain symptoms seeks medical treatment. The attending physician may entertain different hypotheses regarding the possible underlying affliction, which she holds with different degrees of confidence. Each hypothesis predicts distinct probable outcomes of the available treatments. Once a treatment is administered, its outcome provide information regarding the underlying affliction and may be used to update the physician's belief about the likely affliction that is the cause of the symptoms.

Monetary policy: To reduce the unemployment rate, the central bank considers quantitative easing to inflate the prices. There are competing theories regarding the effect

[2] Appling Bayes' rule to obtain the posterior beliefs regarding the validity of hypoteses I and II, it is posible to predict the posterior probabilities that a Bayesian decision maker assings to the events A and A^c and the odds he will accept to bet on these events. Eliciting the posterior probabilities of these events and comparing them to the predictions is a test of Bayesianism. A more general treatment of the elicitation issue is provided in Section 5, below.

of such policies on unemployment. The Phillips curve model predicts persistent negative correlation between the rate of inflation and the rate of unemployment; the long-run Phillips curve model predicts that a higher rate of inflation may reduce the rate of unemployment temporarily but that in the long run, once inflationary expectations are formed, the rate of unemployment attains its natural level at the higher inflation rate. The monetary authority may entertain probabilistic beliefs about the validity of these theories. Based on these beliefs, the central bank implements a policy. Once the effects of the policy are observed and analyzed, it updates its beliefs regarding the validity of the alternative theories and invokes its posterior beliefs the next time it is called upon to intervene.

In these examples theories, or hypotheses, do not necessarily assign acts unique outcomes. Rather, they predict a distribution on a set of possible outcomes, where the randomness may be due to factors that are either unobserved or have not been properly accounted for by the theories.

The rest of the paper is organized as follows. The next section describes the analytical framework. Section 3 depicts the properties of the preference relations and their representations. Section 4 models experiments and discusses the value of information. Section 5 introduces a novel, incentive compatible, mechanism designed to elicit the decision maker's subjective degrees of beliefs in the truth of theories. Section 6 includes further discussion of the model and a reviews the related literature. Section 7 provides the proofs.

2 The Analytical Framework

2.1 Theories and decision models

Theories formalize the causal relationships between acts and outcomes. To formalize this idea denote by Ω an abstract topological state space and let $\mathcal{B}$ be the Borel sigma algebra on Ω. Denote by X a finite set of *outcomes* and let F be the set of random variables (i.e., measurable functions) on Ω taking values in X. Elements of F are *acts.*

A *theory* is a measure space $(\Omega, \mathcal{B}, \mu_t)$, where μ_t is a probability measure. Let T denote the index set whose generic element t indexes the theory $(\Omega, \mathcal{B}, \mu_t)$. Thus, each theory maps the set of acts to probability distributions on X as follows: $t\left(f\right)\left(x\right) = \mu_t\left(f^{-1}\left(x\right)\right)$, for all $x \in X$. The support of $t\left(f\right)$ is the set $X\left(t, f\right) := \{x \in X \mid \mu_t\left(f^{-1}\left(x\right)\right) > 0\}$. Let ΔX denote the simplex in $\mathbb{R}^{|X|}$ then theories are functions from F to ΔX. I assume that the set of theories T is finite.

An *observation* is an act-outcome pair, $(f, x) \in F \times X$. An observation, (f, x) is *consistent* with a theory, t, if $x \in X(t, f)$. If (f, x) is such that $x \notin X(t, f)$ then t is said to be *falsified*.

A *decision model* is a set $\{F, X, (\Omega, \mathcal{B}, \mu_t)_{t \in T}\}$. It is possible to reformulate the decision model so that acts are functions from the set T to ΔX. Specifically, $f : T \to \Delta X$ is defined by $f(t) = t(f)$.[3] Thus, acts are identified with the vectors in $\mathbb{R}^{|T \times X|}$ as follows: $f \in F$ is identified with $\left(\mu_t \left(f^{-1}(x)\right)\right)_{(t,x) \in T \times X}$.

The analytical framework encompasses two layers of randomness. The first layer is modeled by the state space, Ω, representing the presence of variables, not accounted for by the theories, that introduce randomness to theoretical predictions of the outcomes of acts. The second layer concerns the uncertainty regarding truth of the theories themselves.

2.2 The choice set and preference relations

Decisions are choices of sequences of acts. To simplify the exposition I model decisions as two-stage sequential processes. In the first stage, a choice of an act and the resulting outcome produce an observation, $(f, x) \in F \times X$, on the basis of which the decision maker updates his beliefs about the validity the underlying theories and revises his preferences before choosing the subsequent act. Formally, let $\mathcal{Z} := \{\zeta : F \times X \to F\}$, the set of mappings representing choice of acts in the second stage contingent on every conceivable observations. Then, each $\zeta \in \mathcal{Z}$ is identified with the vector $\left(\mu_t \left(\zeta^{-1}(f, x)(x')\right)\right)_{(t,x,x') \in T \times X \times X}$ $\mathbb{R}^{|T \times X| \times |X|}$.

The *choice set* is $\mathbb{C} := F \times \mathcal{Z}$, whose generic element, (f, ζ), consists of act and a contingent plan for choosing a second-stage act. Because $F \subset \mathbb{R}^{|T \times X|}$ and $\mathcal{Z} \subset \mathbb{R}^{|T \times X| \times |X|}$, they are connected separable topological spaces. Thus, the choice set endowed with the product topology is a connected separable topological space.

A *preference relation* $\succcurlyeq$ on $\mathbb{C}$ is a binary relation that has the following interpretation. For all $(f, \zeta), (f', \zeta') \in \mathbb{C}$, $(f, \zeta) \succcurlyeq (f', \zeta')$ means that choosing the act f in the first stage and pursuing the strategy ζ in the second is at least as preferred as choosing the act f' in the first stage and pursuing the strategy ζ' in the second. The strict preference relation, $\succ$,

[3] According to this formulation the decision model is analoguos to the analytical framework of Anscombe and Aumann (1963) with theories replacing the states.

and the indifference relation, $\sim$, are the asymmetric and symmetric parts of $\succcurlyeq$, respectively.

The preference relations, $\succcurlyeq$, is *nontrivial* if the corresponding strict preference relations is non-empty. I assume throughout that the preference relations are nontrivial. The two components of $\mathbb{C}$ are *essential* if $\neg\left((f,\zeta) \sim \left(f,\zeta'\right), \forall \zeta,\zeta' \in \mathcal{Z}\right)$ and $\neg\left((f,\zeta) \sim (f',\zeta), \forall f,f' \in F\right)$.

3 Preference Relations: Structures and Representations

3.1 The axiomatic structure

The following axioms are standard and require no elaboration.

(A.1) (**Weak Order**) $\succcurlyeq$ on $\mathbb{C}$ is complete and transitive.

(A.2) (**Continuity**) For each $(f,\zeta) \in \mathbb{C}$ the sets $\{\left(f',\zeta'\right) \in \mathbb{C} \mid \left(f',\zeta'\right) \succcurlyeq (f,\zeta)\}$ and $\{\left(f',\zeta'\right) \in \mathbb{C} \mid (f,\zeta) \succcurlyeq \left(f',\zeta'\right)\}$ are closed in the product topology.

The third axiom asserts that the two components of the elements of the choice set exert independent influences on the decision maker's well-being. Formally,

(A.3) (**Component independence)** For all $f,f' \in F$ and $\zeta,\zeta' \in \mathcal{Z}$, $(f',\zeta) \succcurlyeq (f,\zeta)$ if and only if $\left(f',\zeta'\right) \succcurlyeq \left(f,\zeta'\right)$ and $\left(f,\zeta'\right) \succcurlyeq (f,\zeta)$ if and only if $\left(f',\zeta'\right) \succcurlyeq (f',\zeta)$.

The next axiom maintains that, if (f',ζ) and $\left(f,\zeta'\right)$ represent compensating variation (i.e., $(f',\zeta) \sim \left(f,\zeta'\right)$), meaning that the intensity of preferences as measure by the difference between the decision maker's evaluations of ζ and ζ' is the same as that between f' and f, and if, in addition, (f'',ζ) and $\left(f,\zeta''\right)$ represent another compensating variation indicating that the intensity of preference of ζ'' over ζ is the same as that between f'' over f, then the "value added" of ζ'' over ζ' is the same as that of f'' over f'.

(A.4) (**Thomsen condition**) For all $(f,\zeta), \left(f',\zeta'\right), \left(f'',\zeta''\right) \in \mathbb{C}$, $(f',\zeta) \sim \left(f,\zeta'\right)$, and $(f'',\zeta) \sim \left(f,\zeta''\right)$ imply that $\left(f'',\zeta'\right) \sim (f',\zeta'')$.

To state the next axiom I introduce the following additional notations and definitions: For every $f,f' \in F$ and $\alpha \in [0,1]$ define $\alpha f + (1-\alpha) f' \in F$ by $(\alpha f + (1-\alpha) f')(t) = \alpha f(t) + (1-\alpha) f'(t)$. Thus, F is a convex set. Similarly, for every ζ,$\zeta' \in \mathcal{Z}$, and $\alpha \in [0,1]$ define $\alpha\zeta + (1-\alpha)\zeta' \in \mathcal{Z}$ by $\alpha\left(\zeta + (1-\alpha)\zeta'\right)(f,x) = \alpha\zeta(f,x) + (1-\alpha)\zeta'(f,x)$, for all $(f,x) \in F \times X$.[4] Hence, $\mathcal{Z}$ is a convex set. For all $(f,\zeta), \left(f',\zeta'\right) \in \mathbb{C}$ and $\alpha \in (0,1]$, define $\alpha(f,\zeta) + (1-\alpha)\left(f',\zeta'\right) \in \mathbb{C}$ by: $\alpha(f,\zeta) + (1-\alpha)\left(f',\zeta'\right) = \left(\alpha f + (1-\alpha f'), \alpha\zeta + (1-\alpha)\zeta'\right)$.

[4] Recall that $\zeta(f,x) \in F$. Hence, the mixture operation is on F..

The next axiom is the well-known independence axiom of expected utility theory applied to $\mathbb{C}$, and have the usual separability justification, namely, that the preference between two probability mixtures of act-contingent plan pairs, is independent of the pair that are common to the two mixtures.

(A.5) (**Independence**) For all $(f,\zeta), (f',\zeta'), (f'',\zeta'') \in \mathbb{C}$ and $\alpha \in (0,1]$, $(f,\zeta) \succcurlyeq (f',\zeta')$ if and only if $\alpha(f,\zeta) + (1-\alpha)(f'',\zeta'') \succcurlyeq \alpha(f',\zeta') + (1-\alpha)(f'',\zeta'')$.

Given $\succcurlyeq$ on $\mathbb{C}$, define the sub-relation $\succcurlyeq_1$on F and the sub-relation $\succcurlyeq_2$on $\mathcal{Z}$, respectively, by: $f \succcurlyeq_1 f'$ if $(f,\zeta) \succcurlyeq (f',\zeta)$, for all $\zeta \in \mathcal{Z}$, and $\zeta \succcurlyeq_2 \zeta'$ if $(f,\zeta) \succcurlyeq (f,\zeta')$, for all $f \in F$. By component independence, these preference relations are well defined. The sub preference relations, $\succcurlyeq_i$, $i = 1,2$ inherit the separability inherent in the independence axiom. In particular, for all $f, f', f'' \in F$ and $\alpha \in (0,1]$, $f \succcurlyeq_1 f'$, if and only if $\alpha f + (1-\alpha) f'' \succcurlyeq_1 \alpha f' + (1-\alpha) f''$ and, for all $\zeta, \zeta', \zeta'' \in \mathcal{Z}$ and $\alpha \in (0,1]$, $\zeta \succcurlyeq_2 \zeta'$ if and only if $\alpha\zeta + (1-\alpha)\zeta'' \succcurlyeq_2 \alpha\zeta' + (1-\alpha)\zeta''$.

The next axiom asserts that the theories, being abstract ideas, guide the decision making process but do not impact the decision maker's well-being directly. To state this assertion formally, I introduce the following notations: Given $f \in F$ let $f_{-t}\,p$ be the act that is obtained by replacing the $t-th$ coordinate of f , (i.e., $\mu_t(f^{-1}(\cdot)) \in \Delta X$) with $p \in \Delta X$. A theory t is *irrelevant* if $(f_{-t}p) \sim_1 (f_{-t}p')$, for all $p, p' \in \Delta X$ and $f \in F$, otherwise it is *relevant*.[5] To grasp the meaning of a theory that is irrelevant, consider the act of jumping of a chair and let the outcomes be landing on the floor and winning $\$z$ or staying suspended in the air and winning $\$z'$. Suppose that the competing theories are gravity, t , predicting landing on the floor, ℓ, and suspended gravity, t', predicting remaining suspended in the air, s. On Earth the latter theory is manifestly invalid, and thus irrelevant, if the decision maker is indifferent among all acts $f := (f(t) = (\ell, z), f(t') = (s, z'))$ and $f' := (f'(t) = (\ell, z), f'(t') = (s, z''))$, for all $\$z'$ and $\$z''$.

(A.6) (**Monotonicity**) For all $f \in F$, $p, p' \in \Delta X$ and all relevant $t, t' \in T$, $(f_{-t}p) \succcurlyeq_1 (f_{-t}p')$ if and only if $(f_{-t'}p) \succcurlyeq_1 (f_{-t'}p')$.

[5] In the theory of decision making under uncertainty the same condition with theories replaced by events indicates that the decision maker believes that the event to be null. In the present context the same condition indicats that the decision maker believes that the theory is invalid and, consequetly, irrelevant insofar as the evaluation of the acts is concerned.

3.2 Representation

A preference relation $\succcurlyeq$ on $\mathbb{C}$ is said to have a *utility representation* if, for all $(f,\zeta), (f',\zeta') \in \mathbb{C}$, $(f,\zeta) \succcurlyeq (f',\zeta')$ if and only if $U(f,\zeta) \geq U(f',\zeta')$. The representation is *additive* if there are real-valued functions U_1 on F and U_2 on $\mathcal{Z}$ such that $U(f,\zeta) = U_1(f) + U_2(\zeta)$, and is an *expected utility representation* if, in addition, the functions U_1 and U_2 are affine.

Theorem 1: *Let* $\succcurlyeq$ on $\mathbb{C}$ *be a preference relation and assume the two components of* $\mathbb{C}$ *are essential. Then the following two conditions are equivalent:*

(i) $\succcurlyeq$ *is nontrivial, continuous, weak order satisfying component independence, the Thomsen condition, independence, and monotonicity.*

(ii) There are non-constant real-valued functions u_1 *and, for all* $(f,x) \in F \times X$, $u_2((f,x),\cdot)$ *on* X *and probability distributions* π *and, for all* $(f,x) \in F \times X$, $\pi(\cdot \mid f,x)$ *on* T *such that* $\succcurlyeq$ *has expected utility representation:*

$$(f,\zeta) \mapsto \Sigma_{x\in X} u_1(x)\, \Sigma_{t\in T}\mu_t\left(f_t^{-1}(x)\right)\pi(t) + \Sigma_{x'\in X} u_2\left((f,x),x'\right) \Sigma_{t\in T}\mu_t\left(\zeta(f,x)^{-1}(x')\right)\pi(t \mid (f,x)).$$

Moreover, the utility functions are unique up to cardinal, unit-comparable, transformation and the probability distributions are unique.[6]

The decision maker is *Bayesian* if, for all $t \in T$, and $(f,x) \in F \times X$, $\pi(t \mid (f,x)) = \mu_t\left(f^{-1}(x)\right)\pi(t) / \Sigma_{t'\in T}\mu_{t'}\left(f^{-1}(x)\right)\pi(t')$. In general, the first stage decision may involves a material-information trade-off. It allows for the possibility that an act be chosen that entails material loss if its informational content permits better second-stage choice.

4 Experiments

4.1 Preferences and representation

Experiments, are acts whose outcomes are referred to as *signals*. Experiments are *free* if they are devoid of material implications (that is, they have no implications for the decision maker's well-being) and whose sole significance is their information content. Formally, experiments are random variables on a measurable space $(\Omega, \mathcal{B})$ taking values in a set of signals, Y. Let E denote the set of experiments, then $\mathcal{E} \subset F$. Let $\mathcal{Y} := \{\zeta : \mathcal{E} \times Y \to F\}$ be

[6] Cardinal unit comaprable transformation means that if $\hat{u}_1$, $\hat{u}_2((f,x),\cdot)$, $(f,x) \in F \times X$, is another set of functions that constitute an additive representation of $\succcurlyeq$ then $\hat{u}_1 = cu_1 + a$ and $u_2((f,x),\cdot) = cu_2((f,x),\cdot) + a(f,x)$, $c > 0$. This property is sometimes referred to as *jointly cardinal.*

sets of mappings representing strategies of choosing acts contingent on the observations.[7] The relevant choice set is $\mathcal{C} := \mathcal{E} \times \mathcal{Y}$, whose generic element, $(\widetilde{y}, \zeta)$, is an experiment-strategy pair.

A central tenet of the subjective expected utility theory is that information affects the decision makers beliefs while leaving their tastes intact. To formalize this principle I propose a variation of the model of the preceding section in which the first-stage decision is a choice of an experiment, $\widetilde{y} \in \mathcal{E}$, to be followed, in the second stage, by a choice of an act contingent on the observations $(\widetilde{y}, y) \in \mathcal{E} \times Y$. The idea is that, based on the observation obtained in the first stage, the decision maker updates his beliefs about the validity the underlying theories and, consequently, his preferences over the second-stage acts. The main concern and challenge is to formalized the idea that neither the experiment itself nor the signal produce affect the decision-maker well being except through the update of his beliefs about the validity of the underlying theories. This is done by the next two axioms.

The first axiom asserts that free experiments have no direct material implications. Specifically, it requires that if the second stage strategy is fixed so that decision maker is unable to use the information, then the decision maker is indifferent among the experiments. To state the axiom I introduce the following additional notations: Denote by ζ_f the *constant strategy* $\zeta_f(\widetilde{y}, y) = f$, for all $(\widetilde{y}, y) \in \mathcal{E} \times Y$. I identify the set $\{\zeta_f \in \mathcal{Y} \mid f \in F\}$ of constant strategies with F.

(A.7) (**Experimental neutrality**) For all $(\widetilde{y}, f), (\widetilde{y}', f) \in \mathcal{E} \times F$, $(\widetilde{y}, f) \sim (\widetilde{y}', f)$.

The next axiom asserts that experimental observations are carriers of information and, as such, have no direct effect on the decision makers material well-being in the second stage. Formalizing this assertion requires that the effect of information on the second-stage choice be isolated. To do so I introduce the following additional notations: For each $f \in F$, $(\widetilde{y}, y) \in \mathcal{E} \times Y$, and $x \in X$, define $p_f(x \mid (\widetilde{y}, y)) = \Sigma_{t \in T} \mu_t \left(f^{-1}(x)\right) \pi(t \mid (\widetilde{y}, y))$. That is the $p_f(\cdot \mid (\widetilde{y}, y)) \in \Delta X$ is the posterior distribution on the outcomes of the act f conditional on the experimental observation $(\widetilde{y}, y)$. Denote by $\zeta_{-(\widetilde{y}, y)} f \in \mathcal{Y}$ the strategy that is obtained by replacing the $(\widetilde{y}, y)$ coordinate of ζ with the act f.

(A.8) (**Observation independence)** For all $f, g, f', g' \in F$ and $(\widetilde{y}, y), (\widetilde{y}', y') \in \mathcal{E} \times Y$, if $p_f(\cdot \mid \widetilde{y}, y) = p_{f'}(\cdot \mid (\widetilde{y}', y'))$ and $p_g(\cdot \mid \widetilde{y}, y) = p_{g'}(\cdot \mid \widetilde{y}', y')$ then $\zeta_{-(\widetilde{y}, y)} f \succcurlyeq \zeta_{-(\widetilde{y}, y)} g$ if and only if $\zeta_{-(\widetilde{y}', y')} f' \succcurlyeq \zeta_{-(\widetilde{y}', y')} g'$.

[7] Since $\mathcal{E} \subset F$ and $\mathcal{Y} \subset X$, we have that $\mathcal{Y} \subset \mathcal{Z}$.

Given $\{f_1, ..., f_n\} \subseteq F$ and $\{\alpha \in [0,1]^n \mid \Sigma_{s\in 1}^n \alpha_s = 1\}$, define the mixture $\Sigma_{s=1}^n \alpha_s f_s$ by $(\Sigma_{s=1}^n \alpha_s f_s)(t) = \Sigma_{s=1}^k \alpha_s f_s(t)$, for all $t \in T$, where $f'_s(t) = \mu_t\left(f^{-1}(\cdot)\right) \in \Delta X$. Extend the definition of the indifference relation so that $\zeta_{-(\widetilde{y},y)}\Sigma_{s=1}^n \alpha_s f_s \sim \zeta_{-(\widetilde{y}',y')}\Sigma_{s=1}^n \alpha_s f'_s$ if $p_f(\cdot \mid (\widetilde{y}, y)) = p_{f'}(\cdot \mid (\widetilde{y}', y'))$. Hence, we may assume without loss of generality that F is a convex set (i.e., identify F with its convex hull), which implies the existence of the acts in F hypothesized in (A.8).

Theorem 2 characterizes the representation of preference ranking of experiments.

Theorem 2: *Let $\succcurlyeq$ on $\mathcal{C}$ be preference relation and assume the two components of $\mathcal{C}$ are essential. Then the following conditions are equivalent:*

(i) $\succcurlyeq$ is nontrivial, continuous, weak order satisfying component independence, the Thomsen condition, independence, monotonicity, experimental neutrality and observation independence.

(ii) There is a non-constant function $u : X \to \mathbb{R}$, and probability distributions, $\pi(\cdot \mid (\widetilde{y}, y))$, $\forall (\widetilde{y}, y) \in \mathcal{E} \times Y$, on T, and $\eta(\cdot \mid \widetilde{y})$, $\forall \widetilde{y} \in \mathcal{E}$, on Y such that $\succcurlyeq$ has a utility representation

$$(\widetilde{y}, \zeta) \mapsto \Sigma_{x\in X} u(x) \Sigma_{y\in Y} \Sigma_{t\in T} \mu_t\left(\zeta^{-1}(\widetilde{y}, y)(x)\right) \pi(t \mid (\widetilde{y}, y)) \eta(y \mid \widetilde{y}).$$

Moreover, u is unique up to positive affine transformation and sets of the probability distributions are unique.

Because free experiments are devoid of material implications, the first stage utility does not figure in the representation and, consequently, unlike in Theorem 1, the prior probabilities on T are not defined. One way to obtain prior probabilities on T is to endow the experiments with material consequences by allowing the decision maker to place bets on the signals. Formally, define *bets* to be functions $\beta : Y \to \Delta M$, where ΔM denotes the lotteries on the real interval, M, whose elements represent monetary payoffs. Let G denote the set of bets, or gambles, and define a new kind of acts that are made of experiment-bet pairs $(\widetilde{y}, \beta) \in \mathcal{E} \times G$. The outcomes of an act $(\widetilde{y}, \beta)$ are $\{(y, \beta(y)) \mid y \in Y\}$. Let $\mathcal{C}^* = (\mathcal{E} \times G) \times \mathcal{Y}$. Then, by Theorems 1 and 2, we get:

Corollary: *Let $\succcurlyeq$ on $\mathcal{C}^*$ be preference relation and assume the two components of $\mathcal{C}^*$ are essential. Then the following conditions are equivalent:*

(i) $\succcurlyeq$ is nontrivial, continuous, weak order satisfying component independence, the Thomsen condition, independence, monotonicity, experimental neutrality and observation independence.

(ii) There is a non-constant functions $u_1 : M \to \mathbb{R}$ and $u_2 : X \to \mathbb{R}$ and sets of probability distributions, π and $\pi(\cdot \mid (\widetilde{y}, y))$, $(\widetilde{y}, y) \in \mathcal{E} \times Y$, on T such that $\succcurlyeq$ has a representation

$$((\widetilde{y}, \beta), \zeta) \mapsto \Sigma_{y \in Y} \left[\Sigma_{z \in M} u_1(z) \beta(y)(z) + \Sigma_{x \in X} u_2(x) \Sigma_{t \in T} \mu_t \left(\zeta^{-1}(\widetilde{y}, y)(x)\right) \pi(t \mid (\widetilde{y}, y))\right] \eta(y \mid \widetilde{y}),$$

where $\eta(y \mid \widetilde{y}) = \Sigma_{t \in T} \mu_t \left(\widetilde{y}^{-1}(y)\right) \pi(t)$, for all $y \in Y$.

Moreover, the utility functions are unique up to cardinal unit-comparable transformation and the probability distributions π, and $\pi(\cdot \mid (f, x))$, $(f, x) \in F \times X$, are unique.

An alternative approach to obtaining a prior applies to Bayesian decision makers. Let $\widetilde{y}_\varnothing$ denote a non-informative experiment. Formally, $\mu_t \left(\widetilde{y}_\varnothing^{-1}(y)\right) = \mu_{t'} \left(\widetilde{y}_\varnothing^{-1}(y)\right)$, for all $y \in Y$ and $t, t' \in T$.[8] Suppose that π is the decision maker prior on T then, by Bayes rule, $\pi(t \mid (\widetilde{y}_\varnothing, y)) = \pi(t)$, for all $t \in T$.

4.2 Comparison of experiments

Experiments are valuable because the information they produce allow decision makers to improve their choices of acts. Distinct experiments may be more or less valuable depending on the information they generate. If they are not free, the choice of an experiment is itself a decision problem. In what follows I discuss the choice of experiments and the value of information from the Bayesian perspective.

Corresponding to every $\widetilde{y} \in \mathcal{E}$ there is a $\mid Y \mid \times \mid T \mid$ left-stochastic matrix, $I(\widetilde{y})$, dubbed the *information structure,* whose (y, t) entry, $\mu_t \left(\widetilde{y}_t^{-1}(y)\right)$, is the probability that the signal y is generated by the experiment $\widetilde{y}$ conditional on theory t being true. Let $B \subset F$ be a budget set of *feasible* acts. Having chosen the experiment $\widetilde{y}$ and observing the signal $y \in Y$ in the first stage, the decision maker chooses the second stage act from the feasible set. Given a utility function $u : X \to \mathbb{R}$, and a subjective prior distribution π on T, define a function $J_{(u,\pi)} : \mathcal{E} \times 2^F \to \mathbb{R}$ by:

$$J_{(u,\pi)}(\widetilde{y}, B) = \Sigma_{y \in Y} \eta(y \mid \widetilde{y}) \left[\max_{f \in B} \Sigma_{x \in X} u(x) \Sigma_{t \in T} \mu_t \left(f^{-1}(x)\right) \pi(t \mid (\widetilde{y}, y))\right],$$

where $\pi(t \mid (\widetilde{y}, y)) = \mu_t \left(\widetilde{y}^{-1}(y)\right) \pi(t) / \Sigma_{t' \in T} \mu_{t'} \left(\widetilde{y}^{-1}(y)\right) \pi(t')$ and $\eta(y \mid \widetilde{y}) = \Sigma_{t \in T} \mu_t \left(\widetilde{y}^{-1}(y)\right) \pi(t)$ denotes, respectively, the (subjective) posteriors probability of $t \in T$ conditional on $(\widetilde{y}, y)$

[8] Choosing $\widetilde{y}_\varnothing$ is equivalent to choosing an act without experimentation.

and probability of $y \in Y$ conditional on $\widetilde{y}$. Then $J_{(u,\pi)}$ represents the decision maker's expected utility if the experiment $\widetilde{y}$ is chosen in the first stage and followed, in the second stage, by a choice of an act that is optimal given the set, B, of feasible acts and the information acquired.

The decision maker's problem is: Given $B \in 2^F$ choose $\widetilde{y} \in \mathcal{E}$ so as to maximize $J_{(u,\pi)}(\widetilde{y}, B)$. Assuming that it exists, let $f^*_{(u,\pi)}(\widetilde{y} \mid B)$ denote the solution to the decision maker's problem. Define the value of $\widetilde{y}$ conditional on B by:

$$J^*_{(u,\pi)}(\widetilde{y} \mid B) = \Sigma_{y\in Y}\eta(y \mid \widetilde{y})\,\Sigma_{x\in X}u(x)\,\Sigma_{t\in T}\mu_t\left(f^{*-1}_{(u,\pi)}(\widetilde{y} \mid B)(x)\right)\pi(t \mid (\widetilde{y}, y)).$$

Note that, for every $t \in T$ and $f \in F$ there must be an outcome $x \in X$ such that $\mu_t\left(f^{-1}(x)\right) > 0$. Otherwise the theory t is a-priori invalid.

Definition: An experiment $\widetilde{y}$ *is more informative than* $\widetilde{y}'$ *from the viewpoint of* (u, π) *at* $B \in 2^F$, denoted $\widetilde{y} \geqslant^{(u,\pi)}_B \widetilde{y}'$, if $J^*_{(u,\pi)}(\widetilde{y} \mid B) \geq J^*_{(u,\pi)}(\widetilde{y}' \mid B)$. An experiment $\widetilde{y}$ *is more informative than* $\widetilde{y}'$ *from the viewpoint of* (u, π), denoted $\widetilde{y} \geqslant^{(u,\pi)} \widetilde{y}'$, if it is more informative at B for all $B \in 2^F$.

Given $B \in 2^F$, the informational value-added of $\widetilde{y}$ over $\widetilde{y}'$ from the viewpoint of (u, π) is:

$$\Psi_{(u,\pi)}\left(\widetilde{y}, \widetilde{y}' \mid B\right) := J^*_{(u,\pi)}(\widetilde{y} \mid B) - J^*_{(u,\pi)}\left(\widetilde{y}' \mid B\right).$$

Because the set of feasible actions limits the opportunities to exploit the information generated by the experiments, the informational value-added depends on this set. For instance, if B is a singleton set then the informational value-added is zero.

Given (u, π) and $\widetilde{y} \in \mathcal{E}$ define the vector $v_{(u,\pi,\widetilde{y},B)} \in \mathbb{R}^{|Y|}$ by

$$v_{(u,\pi,\widetilde{y},B)}(y) = \Sigma_{x\in X}u(x)\,\Sigma_{t\in T}\mu_t\left(f^{*-1}_{(u,\pi)}(\widetilde{y} \mid B)(x)\right)\pi(t \mid (\widetilde{y}, y)),\ \forall y \in Y.$$

For given $\widetilde{y}' \in \mathcal{E}$ define the half-space $H^+(\widetilde{y}' \mid u, \pi, B) := \{\widetilde{y} \in \mathcal{E} \mid \langle \eta(\cdot \mid \widetilde{y}), v_{(u,\pi,\widetilde{y},B)}\rangle \geq \langle \eta(\cdot \mid \widetilde{y}'), v_{(u,\pi,\widetilde{y}',B)}\rangle\}$, where $\langle \cdot, \cdot \rangle$ denotes the inner product. Then we have the following:

Proposition 1: $\widetilde{y} \geqslant^{(u,\pi)}_B \widetilde{y}'$ *if and only if* $\widetilde{y} \in H^+(\widetilde{y}' \mid u, \pi, B)$ and $\widetilde{y} \geqslant^{(u,\pi)} \widetilde{y}'$ *if and only if* $\widetilde{y} \in H^+(\widetilde{y}' \mid u, \pi, B)$, *for all* $B \in 2^F$.

The proof of the proposition is immediate.

Let $\mathcal{M}$ denote the set of all $|\ Y\ | \times |\ Y\ |$ Markov matrices and let $K := \{\widetilde{y} \in \mathcal{E} \mid I(\widetilde{y})Q = I(\widetilde{y}'),$ for some $Q \in \mathcal{M}\}$. By Blackwell's (1951) theorem, $K \subset H^+(\widetilde{y}' \mid u, \pi, B)$, for all $B \in 2^F$. It follows that Blackwell's characterization of an experiment being more informative is sufficient condition for $\widetilde{y} \geqslant^{(u,\pi)} \widetilde{y}'$. However, it is not necessary condition for either $\widetilde{y} \geqslant_B^{(u,\pi)} \widetilde{y}'$ or $\widetilde{y} \geqslant^{(u,\pi)} \widetilde{y}'$.

4.3 The value of information

The choice of an act in the first stage and contingent plan for the second stage may require trading off material payoff in the first stage in exchange for information that can be used to increase the payoff in the second stage. Accordingly costly experimentation is justified if the informational value added resulting from improved choice of act in the second stage exceeds its cost.

Denote by $c(\widetilde{y}) \geq 0$ monetary cost of the experiment $\widetilde{y}$. Then, the set of outcomes in the first period is given by $\{(y, c(\widetilde{y})) \mid y \in Supp(\widetilde{y})\}$. Assume that running no experiment in the first stage (equivalently, running a non-informative experiment $\widetilde{y}_{\varnothing}$) is costless (i.e., $c(\widetilde{y}_{\varnothing}) = 0$). Let u be a real-valued function on $\mathbb{R}$, representing the utility of the monetary cost of the experiments, normalize so that $u(0) = 0$. Then, by Theorem 1, we get that $\widetilde{y}$ is worthwhile form the viewpoint of (u, π) given the feasible set of acts, B, if and only if

$$J^*_{(u,\pi)}(\widetilde{y} \mid B) - u(c(\widetilde{y})) \geq J^*_{u,\pi}((\widetilde{y}_{\varnothing}) \mid B),$$

where $J^*_{(u,\pi)}((\widetilde{y}_{\varnothing}) \mid B) = \max_{f \in B} \Sigma_{x \in X} u(x) \Sigma_{t \in T} \mu_t \left(f^{-1}(x)\right) \pi(t)$. The monetary value of the information generated by the experiment $\widetilde{y}$ form the viewpoint of (u, π) given B is $c^*(u, \pi \mid B) = u^{-1}\left(J^*_{(u,\pi)}(\widetilde{y} \mid B) - J^*_{(u,\pi)}((\widetilde{y}_{\varnothing}) \mid B)\right)$. More generally, given $B \in 2^F$, an experiment $\widetilde{y}$ is preferred over another experiment, $\widetilde{y}'$, form the point of view of (u, π), if $J^*_{(u,\pi)}(\widetilde{y} \mid B) - J^*_{(u,\pi)}(\widetilde{y}' \mid B) \geq u(c(\widetilde{y})) - u(c(\widetilde{y}'))$.

5 Elicitation of the Subjective Probabilities

5.1 The elicitation problem

Incentive compatible mechanisms designed to elicit subjective probabilities on a state space have been studied for more than half a century. Pioneered by the works of Brier (1950) and Good (1952) these studies include Savage (1971), Grether (1981), Kadane and Winkler (1988), and Karni (2009).[9] A common feature of these elicitation schemes is the conditioning of the subject's reward on the events of interest. This conditioning requires that the

[9] For a recent comprehensive review, see Chambers and Lambert (2020).

occurrence of the events of interest be observable and verifiable. Because, in general, theories are neither observable nor verifiable, these mechanisms do not apply to the elicitation of a subject's prior beliefs about the truth of theories.

Prelec (2004), Chambers and Lambert (2015, 2020), and Karni (2020) proposed elicitation mechanisms designed to elicit subjective probabilities on events that are private information and, consequently, unverifiable. However, the working of these mechanisms hinges on the presumption that the subject discovers, for himself, the truth of the unobservable event of interest. Because the uncertainty about the truth of theories may not dissipate in the subject's own mind, these mechanisms too do not apply to the elicitation problem with which we are concerned.

I propose next a new, indirect, scheme designed to elicit the subjective probabilities representing the subject's degree of belief in the truth of the theories and examine the conditions under which it yields the desired outcome.

5.2 The elicitation mechanism

The proposed, indirect, scheme for the elicitation of a subject's subjective degrees of belief in the truth of the theories invokes the observability of the signals of experiments. To describe the mechanism assume, provisionally, that there is an experiment $\widetilde{y} \in \mathcal{E}$, whose support, $Supp\left(\widetilde{y}\right)$, has cardinality that is at least as great as that of the set of theories, T , (i.e.,$\mid Supp\left(\widetilde{y}\right) \mid \geq \mid T \mid$). Let $\Upsilon = \left(Y_1, ..., Y_{|T|}\right)$ be a partition the set Y.[10]

Since the signals are observable and verifiable, it is possible to apply one of the existing schemes (e.g., Karni [2009]) to elicit the subject's subjective probabilities of the cells of the partition, $P\left(Y_i\right), i = 1, ..., \mid T \mid$ and let $P\left(\widetilde{y}\right) := \left(P(Y_1), ..., P(Y_{|T|})\right)$. By Theorem 2, for all $Y_i \in \Upsilon$, $P\left(Y_i\right) = \Sigma_{t \in T} \mu_t \left(\cup_{y \in Y_i} \widetilde{y}^{-1}\left(y\right)\right) \pi\left(t\right)$.

For each $Y_i \in \Upsilon$ and $t \in T$ define $\xi_t\left(Y_i\right) = \mu_t \left(\cup_{y \in Y_i} \widetilde{y}^{-1}\left(y\right)\right)$. Let $\boldsymbol{\pi} := \left(\pi\left(t_1\right), ..., \pi\left(t_{|T|}\right)\right)$, then $A\boldsymbol{\pi}^{\tau} = \left(P(Y_1), ..., P(Y_{|T|-1}), 1\right)^{\tau}$, where the superscript τ is the transpose and A is

[10] If $\mid Supp\left(\widetilde{y}\right) \mid = \mid T \mid$ then cells of the partition $\left(y_1, ..., y_{|T|}\right)$, are singlton sets, each containing an element of the support of $\widetilde{y}$.

the $| T | \times | T |$ matrix given by:

$$A = \begin{bmatrix} \xi_{t_1}(Y_1) & \cdot & \cdot & \xi_{t_{|T|}}(Y_1) \\ \cdot & & & \cdot \\ \xi_{t_1}\left(Y_{|T|-1}\right) & \cdot & \cdot & \xi_{t_{|T|}}\left(Y_{|T|-1}\right) \\ 1 & 1 & 1 & 1 \end{bmatrix}. \tag{1}$$

The following proposition is immediate.

Proposition 2: *The probability distribution* $\boldsymbol{\pi}$ *on* T *exists and is unique if and only if there is an experiment* $\widetilde{y} \in \mathcal{E}$ *and a partition of the support of* $\widetilde{y}$ *such that the corresponding matrix* A *is nonsingular.*

The elicitation scheme of P provides a procedure for indirect elicitation of the subjective probabilities on T as a solution of the linear system of equations given by $A\boldsymbol{\pi}^{\tau} = \left(P(Y_1), ..., P(Y_{|T|-1}), 1\right)^{\tau}$.

If no single experiment has support that includes larger number of observations than there are theories (that is, the case in which, for no $\widetilde{y} \in \mathcal{E}$, $| \; Supp\,(\widetilde{y}) \; | \geq | \; T \; |$) but $| \cup_{\widetilde{y} \in \mathcal{E}} Supp\,(\widetilde{y}) | \geq T$, it is possible to apply a modified version of the mechanism described above. To simplify the exposition, without loss of generality, suppose that there are K experiments, $\{\widetilde{y}_1, ..., \widetilde{y}_K\}$. Let $s_k :=| \; Supp\,(\widetilde{y}_k) \; |$, $k = 1, ..., K$, and suppose that $\Sigma_{k=1}^{K} s_k =| \; T \; |$.[11] Denote by $\mu_t\left(\widetilde{y}_k^{-1}(y_{k,j})\right)$ the probability, according to the theory t, of observing the signal $y_{k,j}$ if the experiment $\widetilde{y}_k$ is implemented.

The mechanism requires that the experiments $\{\widetilde{y}_1, ..., \widetilde{y}_K\}$ are implements and, for each experiment the vector $P\,(\widetilde{y}_k) := \left(P_{\widetilde{y}_k}(y_{k,1}), ..., P_{\widetilde{y}_k}(y_{k,s_k})\right)$ is elicited using one of the standard procedures. Consider the system of equations $\hat{A}\boldsymbol{\pi}^{\tau} = \hat{P}^{\tau}$, where

$$\hat{P} := \left(P_{\widetilde{y}_1}(y_{1,1}), ..., P_{\widetilde{y}_1}(y_{1,\,s_1-1}), ..., P_{\widetilde{y}_K}(y_{K,1}), ..., P_{\widetilde{y}_K}(y_{K,s_K-1}), 1\right)$$

[11] If $\Sigma_{k=1}^{K} s_k >| \; T \; |$ then, for some experiments, create partitions as necessary so that the probability vector defined below have the dimention $| \; T \; |$.

and

$$\hat{A} = \begin{bmatrix} \mu_{t_1}\left(\widetilde{y}_1^{-1}\left(y_{1,1}\right)\right) & \cdot\ \cdot\ \cdot & \mu_{t_{|T|}}\left(\widetilde{y}_1^{-1}\left(y_{1,1}\right)\right) \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \mu_{t_1}\left(\widetilde{y}_1^{-1}\left(y_{1,s_1-1}\right)\right) & \cdot\ \cdot\ \cdot & \mu_{t_{|T|}}\left(\widetilde{y}_1^{-1}\left(y_{1,s_1-1}\right)\right) \\ \mu_{t_1}\left(\widetilde{y}_2^{-1}\left(y_{2,1}\right)\right) & \cdot\ \cdot\ \cdot & \mu_{t_{|T|}}\left(\widetilde{y}_2^{-1}\left(y_{2,1}\right)\right) \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \mu_{t_1}\left(\widetilde{y}_2^{-1}\left(y_{2,s_2-1}\right)\right) & \cdot\ \cdot\ \cdot & \mu_{t_{|T|}}\left(\widetilde{y}_2^{-1}\left(y_{2,s_2-1}\right)\right) \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \mu_{t_1}\left(\widetilde{y}_K^{-1}\left(y_{K,s_K-1}\right)\right) & \cdot\ \cdot\ \cdot & \mu_{t_{|T|}}\left(\widetilde{y}_K^{-1}\left(y_{K,s_K-1}\right)\right) \\ 1 & 1\ 1\ 1 & 1 \end{bmatrix}$$

Then, the subjective probability $\boldsymbol{\pi}$ on T exists and is unique if and only if the matrix $\hat{A}$ is nonsingular. It is given by the solution to the system of equations $\hat{A}\boldsymbol{\pi}^{\tau} = \hat{P}^{\tau}$.

6 Discussion

6.1 States and consequences

Perhaps the simplest way to explain the distinction between the analytical framework proposed here to that of Savage (1954) is by invoking the model of Anscombe and Aumann (1963). In the Anscombe-Aumann model states are the outcomes of a horse race and acts are an assignment to each state a roulette lottery (that is, a simple objective probability distribution on an arbitrary set of prizes). A decision maker's subjective probabilities of the outcomes of a horse race are derived from her preferences over the set of acts. By contrast, in this paper, the probabilities of the outcomes of a horse race are forecasts of "horse-race theories" and the decision maker's subjective probabilities depict his degree of belief in the truth of these theories.

A shortcoming of Savage's analytical framework is the that it may confound the notion of states and consequences. This is exemplified in several scenarios described by Aumann and in a correspondence with Savage.[12] One of these scenarios depicts a man who loves his wife to the point that his life without her would be "less 'worth living.'" The wife falls ill and, to survive, she must undergo a routine but dangerous operation. The husband is offered a choice between betting $100 on his wife's survival and on the outcome of a flip of

[12] The correspondence is reproduced in Drèze (1987) and in the collected works of Aumann (2000).

a fair coin. Even supposing that the husband believes that his wife has an even chance of surviving the operation, according to Aumann, he still rather bet on her survival, because, betting on the outcome of a coin flip, the husband may win but winning \$100 if she does not survive is "somehow worthless."

In his response, Savage admits that the difficulty Aumann identifies is indeed serious. In defense of his model, Savage writes, "The theory of personal probability and utility is, as I see it, a sort of framework into which I hope to fit a large class of decision problems. In this process, a certain amount of pushing, pulling, and departure from common sense may be acceptable and even advisable.... To some–perhaps to you–it will seem grotesque if I say that I should not mind being hung so long as it be done without damage to my health or reputation, but I think it desirable to adopt such language so that the danger of being hung can be contemplated in this framework." (Drèze 1987, p. 78).

Cast in terms of the present model, the issue raised by Aumann can be addressed without having to resort to the convoluted reasoning of Savage. To begin with, in terms of the model of this paper, the results of the surgery are outcomes, not states. The way the situation is depicted does not leave room for alternative diagnoses, or hypotheses regarding the causes of the illness. However, this does not rule out the realistic possibility of modifying the probabilities of the outcomes by choosing the hospital or the surgeon to perform the operation. Supposing that surgeons that performed the operation frequently and hospitals in which such operations are routine improve chance of success, alternative theories may weigh these factors differently and, accordingly, assign the operation at a specific hospital by a particular surgeon different chances of success. Ultimately, the husband belief about his wife's chance of survival depends on his subjective beliefs regarding the validity of the different theories. Seen in this way, a choice of where to perform the surgery and by whom, may be regarded as a choice of a lottery with two outcomes, life, L, and death, D. Presumably, the decision makers should have no difficulty choosing among such imaginable acts if they were feasible.

Let w denotes the husband's wealth, than choosing a hospital and a surgeon to perform the operation, the husband's subjective expected utility is: $u(L,w)\,\Sigma_{t\in T}\mu_t\left(f^{-1}(L)\right)\pi(t)+u(D,w)\,\Sigma_{t\in T}\mu_t\left(f^{-1}(D)\right)\pi(t)$. The probability of the wife's survival, $\Sigma_{t\in T}\mu_t\left(f^{-1}(L)\right)\pi(t)$, depends on the husband's belief, quantified by π, regarding the likely validity of the theories about the importance of the experience of the hospital and surgeon.

A bet on the survival of the wife is an act, denoted $x_L y$, that pays x dollars in the event, $f^{-1}(L)$, that the wife survived, and y dollars otherwise, where $x > 0 > y$. Such a bet yields the expected utility payoff

$$u(L, w + x) \Sigma_{t \in T} \mu_t \left(f^{-1}(L)\right) \pi(t) + u(D, w + y) \Sigma_{t \in T} \mu_t \left(f^{-1}(D)\right) \pi(t).$$

By contrast, a bet on heads on a flip of a fair coin is the act $g = x_H y$, that pay x dollars in the event heads up, $g^{-1}(H)$ and y dollars in the complementary event, $g^{-1}(T)$. The probabilities of these events are determined by the theory about the parameter, $\lambda \in [0, 1]$, that represent the "bias" of the coin. Let $\bar{\lambda} := \int_0^1 \lambda \mu_\lambda \left(g^{-1}(H)\right) d\pi(\lambda)$, denote the expectations of τ according to the husband's subjective beliefs about the coin's bias. Then the subjective belief that the coin is fair means that $\bar{\lambda} = 1/2$ Supposing that the outcomes of the surgery and the coin flip are stochastically independent, the expected utility of betting on the heads is:

$$\begin{aligned}&\left[u(L, w + x) \Sigma_{t \in T} \mu_t \left(f^{-1}(L)\right) \pi(t) + u(D, w + x) \Sigma_{t \in T} \mu_t \left(f^{-1}(D)\right) \pi(t)\right]/2 + \\ &\left[u(L, w + y) \Sigma_{t \in T} \mu_t \left(f^{-1}(L)\right) \pi(t) + u(D, w + y) \Sigma_{t \in T} \mu_t \left(f^{-1}(D)\right) \pi(t)\right]/2.\end{aligned}$$

According to Aumann's story $\Sigma_{t \in T} \mu_t \left(f^{-1}(L)\right) \pi(t) = 1/2$. Hence, the expected utility of betting on the survival of the wife is $[u(L, w + x) + u(D, w + y)]/2$ and that of betting heads up on the coin flip is:

$$\{[u(L, w + x) + u(D, w + x)]/2 + [u(L, w + y) + u(D, w + y)]/2\}/2.$$

If $u(L, w + x) > [u(L, w + x) + u(D, w + x)]/2$ and $u(D, w + y) > [u(L, w + y) + u(D, w + y)]/2$, then the husband's strict preference for betting on his wife's survival is consistent with his beliefs. More importantly, there is no confounding of states and consequences. The outcomes of the surgery and the payoffs of the bets are clearly consequences. The underlying states (that is, elements of Ω) representing factors that impact the outcomes but are not accountable by the theories are implicit. The utility functions that figure in the representation are state-independent.

A theory, as defined in this work, is an abstract idea that, if correct, resolves the uncertainty associated with the outcomes of all acts up to unaccountable factors. According to this interpretation, $(t, \omega) \in T \times \Omega$ may be regarded as a state in the framework of Savage (1954) and $t \in T$ may be regarded as a state in the Anscombe-Aumann (1963) model. There is, however, an important difference between theories and Anscombe-Aumann states. Being an abstract idea, theories are not susceptible to be confound with outcomes, as is the case in some important applications, (e.g., health and life insurance). In these cases the utility functions are state-dependent and, as a result, the subjective probabilities of the states are not uniquely defined. By contrast, being abstractions, theories do not impact the decision makers well-being and, hence, the utility functions that figure in the representation of the preference relation are theory-independent. Furthermore, predicting the outcomes of alternative acts on the basis of some underlying laws that capture the regularity of the relation between acts and outcomes seems to correspond to the way we think. It also conforms well with the scientific method according to which general laws are parsimonious and efficient, in the sense of demanding less effort, way of describing the environment relevant to a decision problem, compared to describing it in all its details.

6.2 Related literature

Several ingredients of the analytical framework of this paper appear in Karni (2011). Specifically, the set of contingent plans and set of outcomes in the present analytical framework correspond, respectively, to the sets of strategies and the set of effects in Karni (2011). Signals are also an ingredients of both models. However, whereas in Karni (2011) the signals are exogenous to the decision making process, in the present paper the acquisition of signals is done by experimentation and, as such, is an endogenous aspect of the decision making process. More importantly, unlike the probabilities on the states (that is, the mappings from the set of strategies to the set of consequences) that the quantify the decision maker's beliefs about the occurrence of one time events, the probabilities of the theories in the present paper depict the decision maker's beliefs about the processes that determines these events.

The issue of experimentation as an information acquisition procedure that precedes the decision is discussed, rather informally, in Savage (1954). According to Savage observations generated by experiments allow the decision maker to determine, in advance, which event in a partition of the state space contains the true state and, consequently, choose the act that yields the highest expected utility conditional on that event. In terms of the present model, the events of the partition are the inverse images of the signals under the experiments, whose probability is predicted by the underlying theories. The choice of experiment is itself a decision problem, corresponding to the choice of the first-stage act in the present model, amalgamating the value of information and the cost of the experiment. What is described here as experiment is what Savage refers to as free observations. Observations that are not free correspond to first-stage acts. The upshot of this brief discussion is that this paper presents a theory of experiments that may, with appropriate reinterpretation, formalizes Savage's ideas.

Hyogo (2007) proposes a different decision theoretical model of experimentation whose focal point is subjective interpretation of relation between experiments and the distribution of signals. Decisions in Hyogo's model span two periods. In the first period the decision maker is supposed to choose an action and a subset of Anscombe-Aumann acts that is referred to as menu. The action generates a signal which is used by the decision maker to update her beliefs about the likelihoods of the states. In the second period, the decision maker chooses an Anscombe-Aumann act from the chosen menu. An experiment is a pair (Y, l), where $l : S \times A \to \Delta Y$ is a function, A is the set of actions, S is the set of states of the world, and ΔY is the set of distributions on a set, Y, of signals. The main objective Hyogo's model is "... to make the pair (Y, l), in addition to the prior, subjective." (Hyogo (2007), p. 317).

Hyogo's approach is fundamentally different from that of this paper in several important respects. To begin with, the modeling and definition of experiments. The analogue of states of the world in Hyogo's model are theories and that of actions are random variables on an abstract measure space taking their values in a signal space. However, unlike in Hyogo's model, the mapping of theories-experiment pairs to distribution of signals and the set of signals itself are objectively given. This is because, by definition, a theory generates predictions of the outcomes of experiments. Consequently, the objective of Hyogo's analysis has no counterpart in the present study in which focus is on the subjective degrees of belief of the decision maker in the truth of the theories. The different objectives require distinct analytical frameworks. Thus, in Hyogo's model elements of choice set in the first period are pairs, consisting of an action and a menu of acts, and that of the second period are acts from the menu that was selected in the first period. In the present model the elements of the choice set consist of experiment and plans of choosing acts contingent on the experiment-generated signals. Finally, the preference structures and their representations of the two models are different, reflecting the distinct objectives and analytical frameworks.

7 Proofs

7.1 Proof of theorem 1

The following claim is implied by Wakker (1988) Theorem 4.4.

Claim 1: *Let $\succcurlyeq$ on $\mathbb{C}$ be preference relation and assume the two components of $\mathbb{C}$ are essential then $\succcurlyeq$ is a continuous weak order satisfying component independence and the Thomsen condition if and only if it has continuously additive representation. Moreover, the representation is unique up to cardinal, unit-comparable, transformation.*[13]

By Claim 1, the sub-preference relations $\succcurlyeq_1$and $\succcurlyeq_2$ are well-defined and are represented by utility functions U_1 and U_2, respectively. Consider the sub-preference relation $\succcurlyeq_1$ on F. Claim 2 below is the Anscombe and Aumann (1963) representation of $\succcurlyeq_1$.[14]

Claim 2: *A preference relation $\succcurlyeq_1$ on F is a nontrivial, continuous, weak order satisfying independence and monotonicity if and only if there is non-constant, affine, function $\hat{U}_1 : \Delta X \to \mathbb{R}$, and a probability distribution π on T such that, for all $f, f' \in F$,*

$$f \succcurlyeq_1 f' \Leftrightarrow \Sigma_{t \in T} \left[\hat{U}_1 \left(f \left(t \right) \right) - \hat{U}_1 \left(f' \left(t \right) \right) \right] \pi \left(t \right) \geq 0.$$

Moreover, U_1 is unique up to positive affine transformation and π is unique.

For each $x \in X$, the probability of the outcome x under f if the theory t is true is $\mu_t \left(f_t^{-1} \left(x \right) \right)$. By the affinity of U_1, there exist continuous, real-valued, function u_1 on X such that $\hat{U}_1 \left(f \left(t \right) \right) = \Sigma_{x \in X} u_1 \left(x \right) \mu_t \left(f_t^{-1} \left(x \right) \right)$. Hence,

$$U_1 \left(f \right) = \Sigma_{t \in T} \hat{U}_1 \left(f \left(t \right) \right) \pi \left(t \right) = \Sigma_{x \in X} u_1 \left(x \right) \Sigma_{t \in T} \mu_t \left(f_t^{-1} \left(x \right) \right) \pi \left(t \right). \tag{2}$$

Consider next the sub-preference relation $\succcurlyeq_2$on $\mathcal{Z}$.

Claim 3: *A preference relation $\succcurlyeq_2$ on $\mathcal{Z}$ is a nontrivial continuous weak order satisfying independence if and only if there is a set of non-constant, affine, functions $\{\hat{U}_2 \left(\left(f, x \right), \cdot \right) : F \to \mathbb{R} \mid \left(f, x \right) \in F \times X \}$, such that, for all $\zeta, \zeta' \in \mathcal{Z}$ and $f \in F$,*

$$\zeta \succcurlyeq_2 \zeta' \Leftrightarrow \Sigma_{(f,x) \in F \times X} \hat{U}_2 \left(\left(f, x \right), \zeta \left(f, x \right) \right) \geq \Sigma_{(f,x) \in F \times X} \hat{U}_2 \left(\left(f, x \right), \zeta' \left(f, x \right) \right).$$

Moreover, for each observation $\left(f, x \right) \in F \times X$, $\hat{U}_2 \left(\left(f, x \right), \cdot \right)$, $\left(f, x \right) \in F \times X$, are unique up to cardinal, unit-comparable, transformation.

[13] Replacing the Thomsen condition with the hexagon condition, the same result follows from Wakker (1989) Theorem III.4.1.

[14] For a proof of Claim 2 see Kreps (1988).

Proof of Claim 3. The necessity is immediate. The proof of sufficiency is as follows:

By Claim 1, for all $\zeta, \zeta' \in \mathcal{Z}$, $\zeta \succcurlyeq_2 \zeta'$ if and only if $U_2(\zeta) \geq U_2(\zeta')$. Nontriviality implies that U_2 is non-constant. Independence implies that $U_2(\alpha\zeta + (1-\alpha)\zeta') = \alpha U_2(\zeta) + (1-\alpha) U_2(\zeta')$ and that U_2 is unique up to positive linear transformation.

Fix $\zeta^* \in \mathcal{Z}$ and let $| F \times X |= n$. For any $\zeta \in \mathcal{Z}$ define $\zeta^{(f,x)} = \zeta^*_{-(f,x)}\zeta(f,x)$, $(f,x) \in F \times X$. Then, by definition,

$$\frac{1}{n}\zeta + \frac{n-1}{n}\zeta^* = \frac{1}{n}\Sigma_{(f,x)\in F\times X}\zeta^{(f,x)}.$$

By the affinity of U_2,

$$\frac{1}{n}U_2(\zeta) + \frac{n-1}{n}U_2(\zeta^*) = \frac{1}{n}\Sigma_{(f,x)\in F\times X}U_2\left(\zeta^{(f,x)}\right).$$

Define a function $\hat{U}_2((f,x),\cdot) : F \to \mathbb{R}$ by:

$$\hat{U}_2((f,x),\zeta(f,x)) = U_2\left(\zeta^*_{-(f,x)}\zeta(f,x)\right) - \frac{n-1}{n}U_2(\zeta^*).$$

Then, for all $\zeta \in \mathcal{Z}$ we have:

$$\hat{U}_2((f,x),\zeta(f,x)) = U_2\left(\zeta^{(f,x)}\right) - \frac{n-1}{n}U_2(\zeta^*).$$

Thus,

$$\frac{1}{n}\Sigma_{(f,x)\in F\times X}\hat{U}_2((f,x),\zeta(f,x)) = \frac{1}{n}\Sigma_{(f,x)\in F\times X}U_2\left(\zeta^{(f,x)}\right) - \frac{n-1}{n}U_2(\zeta^*).$$

Hence,

$$U_2(\zeta) = \Sigma_{(f,x)\in F\times X}\hat{U}_2((f,x),\zeta(f,x)).$$

The uniqueness part follows form the uniqueness of U_2. ▲

The next claim dissects the functional form of U_2.

Claim 4: *A preference relation $\succcurlyeq_2$ on $\mathcal{Z}$ is a nontrivial continuous weak order satisfying independence and monotonicity if and only if there is a set of non-constant functions, $\{u_2((f,x),\cdot) : X \to \mathbb{R} \mid (f,x) \in F \times X\}$, such that, for all $\zeta, \zeta' \in \mathcal{Z}$ and $f \in F$, $\zeta \succcurlyeq_2 \zeta'$ if and only if*

$$\Sigma_{t\in T}\Sigma_{(f,x)\in F\times X}\left[u_2((f,x),x')\,\mu_t\left(\zeta_t^{-1}(f,x)(x')\right) - u_2((f,x),x)\,\mu_t\left(\left(\zeta'_t\right)^{-1}(f,x)(x')\right)\right]\pi(t \mid (f,x)).$$

Moreover, $u_2((f,x),\cdot)$, $(f,x) \in F \times X$, are unique up to cardinal unit-comparable transformation.

Proof of Claim 4. The necessity is immediate. The proof of sufficiency is as follows:

By Claim 3, $\succcurlyeq_2$ *on* $\mathcal{Z}$ has a representation $\zeta \mapsto \Sigma_{(f,x)\in F\times X}\hat{U}_2\left(\left(f,x\right),\zeta\left(f,x\right)\right)$, where $\zeta\left(f,x\right) \in F \subset \left(\Delta X\right)^T$, and $\hat{U}_2\left(\left(f,x\right),\cdot\right)$ is affine. By the same argument as in the proof of Claim 3, for all $(f,x) \in F \times X$,

$$\hat{U}_2\left(\left(f,x\right),\zeta\left(f,x\right)\right) = \Sigma_{t\in T}\bar{U}_2\left(\left(f,x\right),\zeta\left(f,x\right)\left(t\right),t\right),$$

where $\zeta\left(f,x\right)\left(t\right) = \mu_t\left(\zeta^{-1}\left(f,x\right)\right) \in \Delta X$ and $\bar{U}_2\left(\left(f,x\right),\cdot,t\right) : \Delta X \to \mathbb{R}$ are affine functions.[15]

By monotonicity, for every $f', f'' \in F$, $\bar{U}_2\left(\left(f,x\right),f',t\right) \geq \bar{U}_2\left(\left(f,x\right),f'',t\right)$, if and only if $\bar{U}_2\left(\left(f,x\right),f',t'\right) \geq \bar{U}_2\left(\left(f,x\right),f'',t'\right)$, for all $t,t' \in T$. Thus, by the affinity of $\bar{U}_2\left(\left(f,x\right),\cdot,t\right)$, for all $t,t' \in T$, $\bar{U}_2\left(\left(f,x\right),\cdot,t\right)$ and $\bar{U}_2\left(\left(f,x\right),\cdot,t'\right)$ are positive linear transformations of one another.

Fix t_0 and define $V_2\left(\left(f,x\right),\cdot\right) = \bar{U}_2\left(\left(f,x\right),\cdot,t_0\right)$. Thus, for each $t \in T$,

$$\bar{U}_2\left(\left(f,x\right),\cdot,t\right) = c\left(t;\left(f,x\right)\right)V_2\left(\left(f,x\right),\cdot\right) + a\left(t;\left(f,x\right)\right),\ c\left(t;\left(f,x\right)\right) > 0,$$

Let $\pi\left(t \mid f,x\right) := c\left(t;\left(f,x\right)\right)/\Sigma_{t'\in T}c\left(t';\left(f,x\right)\right)$ and $A\left(t;\left(f,x\right)\right) = a\left(t;\left(f,x\right)\right)/\Sigma_{t'\in T}c\left(t';\left(f,x\right)\right)$ to obtain

$$\bar{U}_2\left(\left(f,x\right),\cdot,t\right) = \pi\left(t \mid f,x\right)V_2\left(\left(f,x\right),\cdot\right) + A\left(t;\left(f,x\right)\right).$$

By the affinity of $V_2\left(\left(f,x\right),\cdot\right)$, there exist real-valued functions $u_2\left(\left(f,x\right),\cdot\right)$ on x such that $V_2\left(\left(f,x\right),\hat{f}\right) = \Sigma_{x'\in X}u_2\left(\left(f,x\right),x'\right)\mu\left(\hat{f}^{-1}\left(x'\right)\right)$, for all $\hat{f} \in F$. Combining these arguments we get the representation

$$\zeta \mapsto U_2\left(\zeta\right) = \Sigma_{t\in T}\Sigma_{x'\in X}u_2\left(\left(f,x\right),x'\right)\mu_t\left(\zeta^{-1}\left(f,x\right)\left(x'\right)\right)\pi\left(t \mid \left(f,x\right)\right). \tag{3}$$

The uniqueness of the representation follows from the uniqueness in Claims 1 and 3. ▲

Combining these results we get that $\succcurlyeq$ on $\mathbb{C}$ is nontrivial, continuous, weak order satisfying component independence, the Thomsen condition, independence, and monotonicity if and only if it is representable by

$$\left(f,\zeta\right) \mapsto U\left(f,\zeta\right) = U_1\left(f\right) + U_2\left(\zeta\right),$$

where, by (2)

$$U_1\left(f\right) = \Sigma_{x\in X}u_1\left(x\right)\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\pi\left(t\right),$$

[15] Recall that $\zeta\left(f,x\right)\left(t\right)\left(x'\right) = \mu_t\left(\zeta^{-1}\left(f,x\right)\left(x'\right)\right)$, for all $x' \in X$.

and, by (3),

$$U_2(\zeta) = \Sigma_{t \in T} \Sigma_{x' \in X} u_2\left((f, x), x'\right) \mu_t\left(\zeta^{-1}(f, x)\left(x'\right)\right) \pi(t \mid (f, x)).$$

The uniqueness of the utilities and probabilities are immediate implications of the uniqueness properties of the representations in Claims 1, 3 and 4. ■

7.2 Proof of theorem 2

Proof. By Claim 1 in the proof of Theorem 1, a preference relation $\succcurlyeq$ on $\mathcal{C}$ is nontrivial, continuous, weak order satisfying component independence and the Thomsen condition if and only if it admits continuous additive representation $(\widetilde{y}, \zeta) \mapsto U_1(\widetilde{y}) + U_2(\zeta)$. Experimental neutrality holds if and only if, for every constant strategy, $\zeta(\widetilde{y}, y) = f$, for all $y \in Y$, $U_1(\widetilde{y}) + U_2(f) = U_1(\widetilde{y}') + U_2(f)$, for all $\widetilde{y}, \widetilde{y}' \in \mathcal{E}$ and $f \in F$. Thus, U_1 is a constant function and, for all $(\widetilde{y}, \zeta), \left(\widetilde{y}', \zeta'\right) \in \mathcal{C}$, $(\widetilde{y}, \zeta) \succcurlyeq \left(\widetilde{y}', \zeta'\right)$ if and only if $U_2(\zeta) \geq U_2\left(\zeta'\right)$.

Moreover, by Claim 4 in the proof of Theorem 1, $\succcurlyeq_2$ on $\mathcal{Y}$ is a continuous weak-order satisfying independence and monotonicity if and only if there is a set of non-constant, affine, functions, $\{u_2((\widetilde{y}, y), \cdot) : X \to \mathbb{R} \mid (\widetilde{y}, y) \in \mathcal{E} \times Y\}$ such that, for all $\zeta \in \mathcal{Y}$,

$$U_2(\zeta) = \Sigma_{t \in T} \Sigma_{(\widetilde{y}, y) \in \mathcal{E} \times Y} u_2((\widetilde{y}, y), x) \mu_t\left(\zeta^{-1}(\widetilde{y}, y)(x)\right) \pi(t \mid (\widetilde{y}, y)).$$

But, for all $\zeta \in \mathcal{Y}$ and $f, g \in F$

$$U_2\left(\zeta_{-(\widetilde{y}, y)} f\right) - U_2\left(\zeta_{-(\widetilde{y}, y)} g\right) = \Sigma_{x \in X} u_2((\widetilde{y}, y), x)\left[p_f(\cdot \mid \widetilde{y}, y) - p_g(\cdot \mid \widetilde{y}, y)\right].$$

Hence, by observation independence, for all $f, g, f', g' \in F$ and $(\widetilde{y}, y), (\widetilde{y}', y') \in \mathcal{E} \times Y$, $p_f(\cdot \mid \widetilde{y}, y) = p_{f'}(\cdot \mid (\widetilde{y}', y'))$ and $p_g(\cdot \mid \widetilde{y}, y) = p_{g'}(\cdot \mid \widetilde{y}', y')$ imply that

$$U_2\left(\zeta_{-(\widetilde{y}, y)} f\right) - U_2\left(\zeta_{-(\widetilde{y}, y)} g\right) = \Sigma_{x \in X} u_2((\widetilde{y}, y), x)\left[p_f(x \mid \widetilde{y}, y) - p_g(x \mid \widetilde{y}, y)\right] \geq 0$$

if and only if

$$U_2\left(\zeta_{-(\widetilde{y}', y')} f'\right) - U_2\left(\zeta_{-(\widetilde{y}', y')} g'\right) = \Sigma_{x \in X} u_2\left(\left(\widetilde{y}', y'\right), x\right)\left[p_{f'}\left(x \mid \left(\widetilde{y}', y'\right)\right) - p_{g'}\left(x \mid \widetilde{y}', y'\right)\right] \geq 0.$$

Thus, the functions $u_2((\widetilde{y}, y), \cdot)$, $(\widetilde{y}, y) \in \mathcal{E} \times Y$, rank the lotteries ΔX in the same way.

Fix $(\widetilde{y}_0, y_0) \in \mathcal{E} \times Y$ and define a function $u : X \to \mathbb{R}$ by $u(x) = u_2((\widetilde{y}_0, y_o), x)$, for all $x \in X$. Then, by the uniqueness of the expected utility representation, for all $(\widetilde{y}, y) \in \mathcal{E} \times Y$,

$u_2\left(\left(\widetilde{y},y\right),x\right)=b\left(\widetilde{y},y\right)u\left(x\right)+a\left(\widetilde{y},y\right)$, $b\left(\widetilde{y},y\right)>0$. Let $\eta\left(y\mid\widetilde{y}\right):=b\left(\widetilde{y},\hat{y}\right)/\Sigma_{\hat{y}\in Y}b\left(\widetilde{y},\hat{y}\right)$, for all $y\in Y$. Then, for all $\zeta\in\mathcal{Y}$,

$$U_2\left(\zeta\right)=\Sigma_{x\in X}u\left(x\right)\Sigma_{y\in Y}\Sigma_{t\in T}\mu_t\left(\zeta_t^{-1}\left(\widetilde{y},y\right)\left(x\right)\right)\pi\left(t\mid\widetilde{y},y\right)\eta\left(y\mid\widetilde{y}\right).$$

Non-triviality implies that U_2 and, hence, u, are non-constant.

The uniqueness is an implication of the uniqueness of Theorem 1. ■

References

[1] Anscombe, Francis J. and Robert J. Aumann (1963) "A Definition of Subjective Probability," *Annals of Mathematical Statistics* 43, 199–205.

[2] Blackwell, David (1951) "The Comparison of Experiments," in J. Neyman (ed.) Proceedings of the Second Berkley Symposium on Mathematical Statistics and Probability (University of California, Berkely), 93-102.

[3] Brier, Glenn, W. (1950) "Verification of Forecasts Expressed in Terms of Probabilities" *Monthly Weather Review*, 78, 1-3.

[4] Chambers, Christopher P., Lambert, Nicolas S.(2015) "Introduction to Dynamic Belief Elicitation," *ACM SIGecom Exchanges* 14, 80-83 .

[5] Chambers, Christopher P., Lambert, Nicolas S. (2020) "Dynamic Belief Elicitation," *Econometrica* (forthcoming).

[6] Drèze, Jacques H. (1987) *Decision Theory with Moral Hazard and State-Dependent Preferences*, in Drèze, J. H. Essays on Economic Decisions Under Uncertainty. Cambridge University Press, Cambridge.

[7] Good, Irving J. (1952) "Rational Decisions," *Journal of Royal Statatistical Society,* Series B (Methodological), 14, 107-114.

[8] Grether, David M. (1981) "Financial Incentive Effects and Individual Decision Making," Cal Tech, Social Science Working Paper 401 .

[9] Hyogo, Kazuya (2007) "A Subjective Model of Experimentation," *Journal of Economic Theory*, 133, 316-330.

[10] Kadane, J. B., Winkler R. L. (1988) "Separating Probability Elicitation from Utility," *Journal of American Statatistical Association* 83, 357-363.

[11] Karni, Edi (2009) "A Mechanism for Eliciting Probabilities." *Econometrica*, 77, 603-606.

[12] Karni, Edi (2011) "Subjective Probabilities on a State Space," *American Economic Journal: Microeconomics* 3, 172-185.

[13] Karni, Edi (2020) "A Mechanism for the Elicitation of Second-Order Beliefs and Subjective Information Structures," *Economic Theory*, 69, 217-232.

[14] Kerps, David (1988) *Notes On the Theory of Choice* Westview Press, Boulder and London.

[15] Prelec, Drazen (2004) "A Bayesian Truth Serum for Subjective Data," *Science* 306, 462-466.

[16] Ramsey, Frank P. (1926), "Truth and Probability." (Published in Braithwaite, R. B. and F. Plumpton, *The Foundation of Mathematics and Other Logical Essays*. London: Routledge and Kegan, 1931).

[17] Savage, Leonard J. (1954) *The Foundations of Statistics.* New York: John Wiley.

[18] Savage, Leonard. J. (1971) "Elicitation of Personal Probabilities and Expectations," *Journal of American Statatistical Association* 66, 783-801 .

[19] Wakker, Peter (1988) "The Algebraic versus the Topological Approach to Additive Representations." *Journal of Mathematical Psychology*, 421-435 (1988)

[20] Wakker, Peter (1988) *Additive Representatios of Preferences.* Kluwer Academic Publishers. Dordrecht.