A Theory-Based Decision Model

Edi Karni[*,†]
Johns Hopkins University

June 4, 2021

Abstract

The paper proposes a theory-based model of decision making under uncertainty whose main premise is that predictions of the outcomes of acts are derived from theories. Realized act-outcome pairs provide information on the basis of which decision-makers update their beliefs regarding the validity of the theories. Consequently, acts are, simultaneously, initiatives that have material consequences and information generating experiments. Pure experiments (that is, information generating acts of no direct material consequences), are characterized and the value of information they generate is defined. An incentive-compatible mechanism is introduced, by which the beliefs that 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 predicting the consequences of alternative actions. The second is the curiosity-driven desire to the explore, by means of observation and experimentation, the physical, biological and social environments.

In the context of decision making under uncertainty, theories, a term that I use interchangeably with hypotheses or propositions, are general laws that decision makers use, explicitly or implicitly, to make sense of alternative courses of action and predict their outcomes. Popper (2002) distinguishes between two kinds of predictions, prophecy, refers to forecasting an event which one can do nothing to prevent (e.g., the coming of a typhoon) and technological predictions which intimate the steps one may take to achieve certain outcomes.[1] Theories of decision making under uncertainty based on the analytical framework of Savage (1954) presume that decision makers predict the results of their actions by assigning consequences to events (i.e., subsets of state space) and attributing a subjective probabilities to these events. In other words, in these theories decision makers choice behavior is governed by *unspecified* combination of prophetic and technological predictions.

The main premise of this paper is that, in many situations involving decision making under uncertainty, decision makers invoke *formal theories*, or hypotheses, to predict the outcomes of the alternative courses of action and choose among feasible actions on the basis of these predictions. Put differently, this works assumes that decision makers' assessment

[1] According to Popper "The distinction between these two sorts of predictions roughly coincides with the lesser or greater importance of the part played by designed experiment, as opposed to mere patient observation, in the science concerned. The typical experimental sciences are capable of making technological predictions, while those employing mainly non-experimental observations produce prophecies." [Popper (2002) pp. 38-39].

of the courses of action in the choice set is based on the predictions of theories rather than intuition alone. In situations that admit competing theories that yield distinct prophecies and technological predictions, decision makers weight these predictions by their beliefs in the validity of these theories.[2] Observations, in this context, consist of action-outcome pairs and are used to update the decision makers' beliefs. Theories are modified or discarded when they are falsified by observations. In other words, a theory is falsified when an action results in an outcome that is outside the set of outcomes that, according to the theory, are feasible given the action. Observations that are consistent with the existing theories induce the updating of the beliefs of the decision makers in the truth of the theories.

It is common to distinguish between acts and experiments. Acts are commonly perceived as initiatives that have material consequences and no informational content, while experiments are perceived as information-generating devices of no 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 kind of experiments. Experiments may have material implications (e.g., cost) and are, therefore, kind of acts. In 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 that applies to a large class of decision problems, I propose a context-dependent approach to modeling choice behavior. According to this approach, it is the context determines the relevant theories. For instance, in the context of medical decisions, the relevant theories predict the probable states of health (i.e., outcomes) that would result from alternative treatments (i.e., acts). In the context of financial decision, the relevant 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 balls 50 of which are red and 50 are black. Balls are drawn sequentially, at random, and their colors observed. A decision maker can place a bet on the event A, (e.g., all the balls are of the same color), or on any other color combination. The decision maker entertains two hypotheses regarding the random process generating

[2]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.

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 various events. If, for example, two balls are drawn then according to hypothesis I, the event A and its complement, A^c, are equally probable; according to hypothesis II, the probability of event A is 49/99 and that of A^c is 50/99. In this scenario, the stochastic process takes place behind a "veil of ignorance" and decision makers get to observe only the colors of the balls. Acts are sequences of random draws of balls and outcomes are the resulting color combinations.

A focal issue in the theory of decision making under uncertainty is the existence and elicitation of decision makers' subjective probabilities of events, such as A and A^c. In this example, these probabilities are induced by underlying (prior) beliefs regarding the validity of the two hypotheses. Suppose that these beliefs are quantifiable by probabilities and let p and $1-p$ denote the probabilities of hypotheses I and II, respectively.

Consider betting on the color combinations of the balls. Following Borel (1924) and Ramsey (1927) the acceptable odds in such bets allow an observer to infer a decision maker's beliefs regarding the likely realization of 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 recover from the betting odds 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 $[99p + (1-p)\,98]\,/198 = 0.496$.[3]

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 and cannot be ascertained except by actually employing them. Suppose, for simplicity, that the employer entertains alternative hypotheses regarding the relationship between employees' productivity and their education

[3] Applying Bayes' rule to obtain the posterior beliefs regarding the validity of hypoteses I and II, it is possible 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.

levels. 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 job vacancies. After having observed the productivity of the employees, the employer updates her beliefs 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 provides information regarding the possible underlying afflictions 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. Competing theories regarding the effect of such policies on unemployment vary from the Phillips curve model that predicts persistent negative correlation between the rate of inflation and the rate of unemployment, to the long-run Phillips curve model that predicts that a higher rate of inflation may reduce the rate of unemployment temporarily, however, 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 models and, based on these beliefs, implements a policy. Once the effects of the policy are observed and analyzed, the central bank updates its beliefs regarding the validity of the alternative models and invokes its posterior beliefs the next time it is called upon to intervene.

In these examples, the hypotheses do not necessarily assign the available courses of action 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 hypotheses.

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, observations, and decision models

Theories depict the causal relationships between acts and outcomes. Consequently, theories constitute the basis for predicting the likelihoods of the outcomes that follow from each act. To formalize this idea I introduce two primitives a set F of *acts* and a finite set, $X = \{x_1, ..., x_n\}$, of *outcomes.* In general, theoretical predictions of the outcomes of acts include a stochastic component, reflecting the fact that factors not accounted for be a theory may play a role in determining the outcomes. Formally, let Ω be the unit interval and $\mathcal{B}$ the Borel sigma-algebra on Ω. Then the set of acts F takes the formal meaning of $\mathcal{B}$–measurable functions on Ω taking values in X. I denote by ΔX the simplex in $\mathbb{R}^n$ whose elements represent probability distributions on X.

A *theory* is a mapping $t : F \to \Delta X$ defined by $t(f)(x) = \mu_t\left(f^{-1}(x)\right)$, for all $x \in X$, where μ_t is a probability measure on $(\Omega, \mathcal{B})$. In other words, $t(f)$ is the prediction of theory t of the likelihoods of the outcomes of the act f. Let $T = \{t_1, ..., t_m\}$ denote the (finite) set of theories.

For each $f \in F$ and $t \in T$, define $f(t)(x) = \mu_t\left(f^{-1}(x)\right)$, for all $x \in X$. Under this definition F is identified with $(\Delta X)^T$.[4] For all $f, f' \in F$ and $\alpha \in [0,1]$ define $(\alpha f + (1-\alpha) f') \in F$, by:

$$\left(\alpha f + (1-\alpha) f'\right)(t)(x) = \alpha \mu_t\left(f^{-1}(x)\right) + (1-\alpha)\mu_t\left(f'^{-1}(x)\right), \forall (t,x) \in T \times X.$$

Thus, F is a convex set in $\mathbb{R}^{|T\times X|}$.

Observations are act-outcome pairs, $(f, x) \in F \times X$. For every $f \in F$ and $t \in T$, the support of $f(t)$ is the set $X(t,f) := \{x \in X \mid \mu_t\left(f^{-1}(x)\right) > 0\}$. An observation, (f, x) is *consistent* with theory t if $x \in X(t, f)$. If (f, x) is inconsistent with a theory t, (that is, $x \notin X(t, f)$) then the theory t is said to be *falsified* by the observation (f, x).

A *decision model* is a set $\{F, X, (\Omega, \mathcal{B}, \mu_t)_{t\in T}\}$.

The analytical framework includes two layers of randomness. The first layer, modeled by $(\Omega, \mathcal{B}, \mu_t)$, formalizes the randomness inherent in the predictions due to presence of

[4] Alternatively, one may think of f as a $n \times m$ left-stochastic matrix.

factors not explicitly accounted for by the theories. The second layer concerns the decision maker's subjective uncertainty regarding the validity, or 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 dynamic processes. In the first stage, the decision maker chooses an act, $f \in F$, and obtains an outcome, $x \in X$. In the second stage the decision maker chooses the subsequent act contingent on the observations. Formally, let $\mathcal{Z} := \{\zeta : F \times X \to F\}$ the set of mappings representing choices of the second stage acts contingent on the first-stage observations. For each $f \in F$ let $\mathcal{Z}(f) := \{\zeta(f) : X \to F \mid \zeta \in \mathcal{Z}\}$.

The *choice set* is $\mathbb{C} := \{(f, \zeta(f)) \in F \times \mathcal{Z}(f)\}$, whose generic element consists of an act $f \in F$ and a plan for choosing a second-stage act $\zeta(f, x) \in F$, $x \in X$. For all $(f, \zeta(f)), (f', \zeta'(f')) \in \mathbb{C}$ and $\alpha \in [0, 1]$, define the convex operation

$$\alpha(f, \zeta(f)) + (1-\alpha)(f', \zeta'(f')) = (\alpha f + (1-\alpha) f', \alpha\zeta(f) + (1-\alpha)\zeta'(f')),$$

where

$$\alpha\zeta(f) + (1-\alpha)\zeta'(f') := (\alpha\zeta(f, x) + (1-\alpha)\zeta'(f', x))_{x \in X},$$

then $\mathbb{C}$ is a convex set in a linear space.

A *preference relation* $\succcurlyeq$ on $\mathbb{C}$ is a binary relation that has the following interpretation. For all $(f, \zeta(f)), (f', \zeta'(f')) \in \mathbb{C}$, $(f, \zeta(f)) \succcurlyeq (f', \zeta'(f'))$ means that choosing the act f in the first stage followed by the implementation of the contingent plan $\zeta(f)$ in the second is at least as preferred as choosing the act f' in the first stage followed by the implementation of the contingent plan $\zeta'(f')$ in the second. The strict preference relation, $\succ$, and the indifference relation, $\sim$, are the asymmetric and symmetric parts of $\succcurlyeq$, respectively. A preference relation, $\succcurlyeq$, is *nontrivial* if the corresponding strict preference relation is nonempty. I assume throughout that the preference relations being considered are nontrivial. The two components of $\mathbb{C}$ are *essential* if $\neg((f, \zeta(f)) \sim (f, \zeta'(f)), \forall \zeta(f), \zeta'(f) \in F^X)$ and $\neg(f, \zeta(f)) \sim (f', \zeta'(f')), \forall f, f' \in F$ and $\zeta(f), \zeta'(f') \in F^X$ such that $\zeta(f) = \zeta'(f')$.

3 Preference Relations: Structures and Representation

3.1 The axiomatic structure

The first two axioms are standard and require no elaboration.

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

(A.2) (**Archimedean**) For each $(f,\zeta(f)), (f',\zeta'(f')), (f'',\zeta''(f'')) \in \mathbb{C}$ such that $(f,\zeta(f)) \succ (f',\zeta'(f')) \succ (f'',\zeta''(f''))$ there are $\alpha,\beta \in (0,1)$ such that $\alpha(f,\zeta(f)) + (1-\alpha)(f'',\zeta''(f'')) \succ (f',\zeta'(f')) \succ \beta(f,\zeta(f)) + (1-\beta)(f'',\zeta''(f''))$.

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

(A.3) (**Independence**) For all $(f,\zeta(f)), (f',\zeta'(f')), (f'',\zeta''(f'')) \in \mathbb{C}$ and $\alpha \in (0,1]$, $(f,\zeta(f)) \succcurlyeq (f',\zeta'(f'))$ if and only if $\alpha(f,\zeta(f)) + (1-\alpha)(f'',\zeta''(f'')) \succcurlyeq \alpha(f',\zeta'(f')) + (1-\alpha)(f'',\zeta''(f''))$.

In particular, for all $(f,\zeta(f)), (f,\zeta'(f)), (f,\zeta''(f)) \in \mathbb{C}$ and $\alpha \in (0,1]$, $(f,\zeta(f)) \succcurlyeq (f,\zeta'(f))$ if and only if $(f,\alpha\zeta(f)+(1-\alpha)\zeta''(f)) \succcurlyeq (f,\alpha\zeta'(f)+(1-\alpha)\zeta''(f))$ and for all $(f,\zeta(f)), (f',\zeta'(f')), (f'',\zeta''(f'')) \in \mathbb{C}$ such that $\zeta(f) = \zeta'(f') = \zeta''(f'') = \hat{\zeta}$ and $\alpha \in (0,1]$, $(f,\hat{\zeta}) \succcurlyeq (f',\hat{\zeta})$ if and only if $(\alpha f + (1-\alpha) f'', \hat{\zeta}) \succcurlyeq (\alpha f' + (1-\alpha) f'', \hat{\zeta})$. Thus, independence holds for each component separably.

The next axiom asserts that theories, being abstract ideas, inform the decision making process by predicting the outcomes of acts but do not impact the decision maker's well-being directly. To state this formally, I introduce the following additional notations: Given $f, f' \in F$, let $(f_{-t}f') \in F$ be defined by: $(f_{-t}f')(t') = f(t')$ if $t' \in T\backslash\{t\}$ and $(f_{-t}f')(t') = f'(t)$ if $t' = t$. Similarly, given $\zeta(f), \zeta'(f) \in \mathcal{Z}(f)$, let $\zeta(f)_{-t}\zeta'(f) \in \mathcal{Z}(f)$ be defined by $(\zeta(f)_{-t}\zeta'(f))(t') = \zeta(f)$ if $t' \in T\backslash\{t\}$ and $(\zeta(f)_{-t}\zeta'(f))(t') = \zeta'(f)(t')$ if $t' = t$. A theory t is *ex-ante irrelevant* if $(f_{-t}f', \zeta(f_{-t}f')) \sim ((f_{-t}f''), \zeta'(f_{-t}f''))$, for all $f, f', f'' \in F$ and $\zeta(f_{-t}f'), \zeta'(f_{-t}f'')$ such that $\zeta(f_{-t}f', x) = \zeta'(f_{-t}f'', x)$, for all $x \in X$. A theory, t, is *ex-post irrelevant* if it is ex-ante irrelevant or if $(f,x) \notin X(t,f)$ and is *ex-post relevant* if $(f,x) \in X(t,f)$.[5]

[5] In the theory of decision making under uncertainty the same condition applied to events indicates that the decision maker believes that the event to be null. In the present context it means that the decision maker believes that the theory is invalid and, consequently, insofar as the evaluation of the acts and contingent plans is concerned, is irrelevant.

(A.4) **(Theory-independence)** (a) For all $f, f', f'' \in F$, $\zeta, \zeta' \in \mathcal{Z}$ and ex-ante relevant $t, t' \in T$, if $\zeta(f_{-t}f') = \zeta'(f_{-t}f'')$ and $\zeta(f_{-t'}f') = \zeta'(f_{-t'}f'')$ then $(f_{-t}f', \zeta(f_{-t}f')) \succcurlyeq \left(f_{-t}f'', \zeta'(f_{-t}f'')\right)$ if and only if $(f_{-t'}f, \zeta(f_{-t'}f')) \succcurlyeq \left(f_{-t'}f'', \zeta'(f_{-t'}f'')\right)$. (b) For all $f \in F$ and $\zeta, \zeta' \in \mathcal{Z}$, and ex-post relevant $t, t' \in T$, $\left(f, \zeta(f,x)_{-t}\zeta'(f,x)\right) \succcurlyeq \left(f, \zeta(f,x)_{-t}\zeta'(f,x)\right)$ if and only if $\left(f, \zeta(f,x)_{-t'}\zeta'(f,x)\right) \succcurlyeq \left(f, \zeta(f,x)_{-t'}\zeta'(f,x)\right)$, for all $x \in X$.

3.2 Representation

The first result characterizes the existence and uniqueness of subjective expected utility representation of $\succcurlyeq$ on $\mathbb{C}$. The subjectivity is due to decision makers' idiosyncratic valuations of the sequences of outcomes and their personal beliefs regarding the validity of the theories.

Theorem 1: *A preference relation $\succcurlyeq$ on $\mathbb{C}$ is nontrivial, Archimedean, weak order satisfying independence and theory-independence if and only if there is nonconstant function $u : X \times X \to \mathbb{R}$, and a probability distribution μ on T such that, for all $(f, \zeta(f))$,$(f', \zeta'(f')) \in \mathbb{C}$,*

$$(f, \zeta(f)) \succcurlyeq (f', \zeta'(f')) \Leftrightarrow U(f, \zeta(f)) \geq U(f', \zeta'(f')),$$

where

$$U(f, \zeta(f)) := \Sigma_{x \in X}\left[\Sigma_{x' \in X} u(x, x')\, \Sigma_{t \in T}\mu_t\left(\zeta(f,x)^{-1}(x')\right)\mu(t \mid f, x)\right]\Sigma_{t \in T}\mu_t\left(f^{-1}(x)\right)\mu(t).$$

Moreover, u is unique up to positive affine transformation and μ is unique.

The representation in Theorem 1 may be reformulated as follows: Let

$$\Pr(x' \mid \zeta(f,x)) := \Sigma_{t \in T}\mu_t\left(\zeta(f,x)^{-1}(x')\right)\mu(t \mid f, x), \ \forall (f,x) \in F \times X$$

and

$$\Pr(x \mid f) := \Sigma_{t \in T}\mu_t\left(f^{-1}(x)\right)\mu(t).$$

Then the joint probability distribution on $X \times X$ induced by $(f, \zeta(f))$ is:

$$\Pr\left(x, x' \mid (f, \zeta(f))\right) = \Pr\left(x' \mid \zeta(f,x)\right)\Pr(x \mid f).$$

The representation may be expressed as follows:

$$(f, \zeta(f)) \mapsto \Sigma_{x \in X}\Sigma_{x' \in X} u(x, x') \Pr(x, x' \mid f, \zeta(f).$$

Dissecting the representation we observe that the subjective probabilities of the theories in the second-stage decisions are conditional on the observations generated by the first-stage decision. The decision maker is *Bayesian* if the second-stage conditional probabilities are $\mu\left(t \mid f,x\right)=\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right)/\Sigma_{t'\in T}\mu_{t'}\left(f^{-1}\left(x\right)\right)\mu\left(t'\right)$, for all $(f,x)\in F\times X$ and $t\in T$. In particular, $\mu(t)=0$ if and only if t is ex-ante irrelevant and $\mu\left(t \mid f,x\right)=0$ if and only if t is either ex-ante or ex-poste irrelevant.

In general, decisions involve trade-offs between material benefits and information acquisition. The model described here allows for an act to be chosen that foregoes imminent material benefits if it generates information that improves the subsequent choices.

4 Experiments

4.1 Preferences and representation

Experiments are acts whose outcomes are dubbed *signals*. Experiments are *free* if they are devoid of material implications (that is, they have no direct impact insofar as the decision maker's well-being is concerned) and whose sole significance resides in their informational content. Formally, experiments are random variables, $\widetilde{y}$, 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 Z:$=\{\zeta:\mathcal{E}\times Y\rightarrow F\}$ be sets of mappings representing strategies of choosing acts contingent on the observations.[6] Let $\mathcal{Z}\left(\widetilde{y}\right):=\{\zeta\left(\widetilde{y}\right):X\rightarrow F\mid\zeta\in\mathcal{Z}\}$, then the choice set is $\mathcal{C}:=\{(\widetilde{y},\zeta\left(\widetilde{y}\right))\in F\times\mathcal{Z}\left(\widetilde{y}\right)\}$, whose generic element, $\left(\widetilde{y},\zeta\left(\widetilde{y}\right)\right)$, 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 premise 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 next axiom formalizes the idea that neither the experiment itself nor its signals affect the decision-maker well-being except through the update of his beliefs about the

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

likely outcomes of the second-stage acts. To formalize this idea it is necessary to separate the informational effects of experiments from potential signal effects on the decision maker's well-being. To achieve this separation, I introduce the notion of uninformative experiments. Formally, an experiment $\widetilde{y} \in \mathcal{E}$ is *uninformative* if the distribution on the set of signals is theory-independent. For example, a single random draw from the urn described in the introduction has the same probability of resulting in a red ball (or black ball) under the two hypotheses. Hence, insofar as the two hypotheses are concerned, one draw is an uninformative experiment. Another example is the education signals mentioned in the introduction. Consider the experiment of hiring an employee who was deprived of the opportunity of acquiring education. Observing the productivity of this employee is uninformative since its distribution is the same under both hypotheses (i.e., it is the prior distribution of productivity in the population). Formally an experiment $\widetilde{y}$ is uninformative if $\mu_t\left(\widetilde{y}^{-1}(y)\right) = \mu_{t'.}\left(\widetilde{y}^{-1}(y)\right)$, for all $y \in Y$ and $t, t' \in T$. The next axiom captures that idea that experiments are valuable only in as much as they are informative. It asserts that if the experiment is uninformative then decision makers are indifferent among strategies involving exchangeability of acts across signals.

(A.4) (**Signals exchangeability)** For every uninformative experiment $\widetilde{y}$, for all $y, y' \in Y$ and $f, f' \in F$, $\left(\widetilde{y}, \zeta(\widetilde{y})_{-y} f_{-y'} f'\right) \sim \left(\widetilde{y}, \zeta(\widetilde{y})_{-y} f'_{-y'} f\right)$.

The next Theorem characterizes the representation of preference ranking of experiments.

Theorem 2: *Let $\succcurlyeq$ on $\mathcal{C}$ be preference relation then $\succcurlyeq$ is nontrivial, Archimedean, weak order satisfying independence, theory-independence, and signals exchangeability if and only if there is a non-constant function $u : X \to \mathbb{R}$, and probability distribution μ on T such that* for all $(\widetilde{y}, \zeta(\widetilde{y})), \left(\widetilde{y}', \zeta'(\widetilde{y}')\right) \in \mathcal{C}$,

$$(\widetilde{y}, \zeta(\widetilde{y})) \succcurlyeq \left(\widetilde{y}', \zeta'(\widetilde{y}')\right) \Leftrightarrow U(\widetilde{y}, \zeta(\widetilde{y})) \geq U\left(\widetilde{y}', \zeta'(\widetilde{y}')\right),$$

where, for all $(\widetilde{y}, \zeta(\widetilde{y})) \in \mathcal{C}$,

$$U(\widetilde{y}, \zeta(\widetilde{y})) = \Sigma_{y \in Y}\left[\Sigma_{x \in X} u(x) \Sigma_{t \in T} \mu_t\left(\zeta(\widetilde{y}, y)^{-1}(x)\right) \mu(t \mid \widetilde{y}, y)\right] \Sigma_{t \in T} \mu_t\left(\widetilde{y}^{-1}(y)\right) \mu(t).$$

Moreover, u is unique up to positive affine transformation and μ is unique.

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. To compare experiments in terms of their values we need to define their informational advantage they confer.

Given a nonempty set $B \subset F$ of *feasible* acts, the value of an experiment, $\widetilde{y}$, is the maximal expected utility of a contingent plan $\zeta(\widetilde{y}) \in B^Y$. Formally, given a utility function $u : X \to \mathbb{R}$ and a subjective prior distribution μ on T, define a value function $J_{(u,\mu)} : \mathcal{E} \times 2^F \to \mathbb{R}$ by:

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

Then $J_{(u,\mu)}(\widetilde{y}, B)$ 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 contingent plan $\zeta^*(\widetilde{y} \mid B)$ that is observation-wise optimal given the B, and the information acquired. Formally, given $\widetilde{y} \in \mathcal{E}$, for every $y \in Y$, $\zeta^*(\widetilde{y}, y \mid B) \in \arg\max_B \Sigma_{x \in X} u(x)\, \Sigma_{t \in T} \mu_t \left(f^{-1}(x)\right) \mu(t \mid \widetilde{y}, y)$.

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

$$V^*_{(u,\mu)}(\widetilde{y} \mid B) = \Sigma_{y \in Y}[\Sigma_{x \in X} u(x)\, \Sigma_{t \in T} \mu_t \left(\zeta^*_{(u,\mu)}(\widetilde{y}, y \mid B)^{-1}(x)\right) \mu(t \mid (\widetilde{y}, y))] \Sigma_{t \in T} \mu_t \left(\widetilde{y}^{-1}(y)\right) \mu(t).$$

Let $\mathcal{P} := \{p \in [0,1]^{|Y|} \mid \Sigma_{y \in Y} p(y) = 1\}$ and $\mathcal{U} := \mathbb{R}^X$.

Definitions: An experiment $\widetilde{y}$ *is more informative than* an experiment $\widetilde{y}'$ *from the viewpoint of* $(u, \mu) \in \mathcal{U} \times \mathcal{P}$ if $V^*_{(u,\mu)}(\widetilde{y} \mid B) \geq V^*_{(u,\mu)}(\widetilde{y}' \mid B)$ for all nonempty $B \in 2^F$. An experiment $\widetilde{y}$ *is more informative than* an experiment $\widetilde{y}'$ if it is more informative from the viewpoint of (u, μ), for all $(u, \mu) \in \mathcal{U} \times \mathcal{P}$.

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 *information structure,* whose (y, t) entry, $\mu_t\left(\widetilde{y}^{-1}(y)\right)$, is the probability that the signal y is generated by the experiment $\widetilde{y}$ conditional on theory t being true. Let $\mathcal{M}$ denote the set of all $\mid Y \mid \times \mid Y \mid$ Markov matrices. By Blackwell's (1951) theorem an experiment $\widetilde{y}$ is more informative than $\widetilde{y}'$ if and only if there is $M \in \mathcal{M}$ such that $I(\widetilde{y}') = MI(\widetilde{y})$. Thus, Blackwell's characterization of an experiment $\widetilde{y}$ is more informative than $\widetilde{y}'$ is sufficient but not necessary for condition for $\widetilde{y}$ to more informative than $\widetilde{y}'$ from the viewpoint of (u, μ).

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

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 benefits in the first stage in exchange for valuable information. Accordingly *costly experimentation* is justified if the informational value added 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 stage of the decision process is $\{(y, c(\widetilde{y})) \mid y \in Y\}$. Not experimenting is informationally equivalent to conducting a non-informative experiment, denoted by $\widetilde{y}_\varnothing \in \mathcal{E}$. Assume that not experimenting is costless (i.e., $c(\widetilde{y}_\varnothing) = 0$). Then, letting

$$V^*_{(u,\mu)}(\widetilde{y} \mid B, c(\widetilde{y})) := \Sigma_{y \in Y} \Sigma_{x \in X} u(-c(\widetilde{y}), x) \left[\Sigma_{t \in T} \mu_t \left(\zeta^*_{(u,\mu)}(\widetilde{y}, y \mid B)^{-1}(x)\right) \mu(t \mid \widetilde{y}, y)\right] \Sigma_{t \in T} \mu_t \left(\widetilde{y}^{-1}(y)\right) \mu(t$$

By Theorem 1, we get that $\widetilde{y}$ is worthwhile form the viewpoint of (u, μ) given the feasible set of acts, B, if and only if

$$V^*_{(u,\mu)}(\widetilde{y} \mid B, c(\widetilde{y})) \geq V^*_{(u,\mu)}(\widetilde{y}_\varnothing \mid B),$$

where $V^*_{(u,\mu)}(\widetilde{y}_\varnothing \mid B) = \max_{f \in B} \Sigma_{x \in X} u(x) \Sigma_{t \in T} \mu_t \left(f^{-1}(x)\right) \mu(t)$. The monetary value of the information produced by an experiment is given by the solution $c_{(u,\mu)}(\widetilde{y} \mid B)$ of the equation $V^*_{(u,\mu)}(\widetilde{y} \mid B, c) = V^*_{(u,\mu)}(\widetilde{y}_\varnothing \mid B)$.

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).[7] A common feature of these elicitation schemes is the condition-

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

ing of the subject's reward on the events of interest. This conditioning requires that the 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, 2021), 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, incentive compatible 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. The proposed scheme invokes the observability of the signals of experiments.

5.2 The elicitation mechanism

To describe the mechanism assume, provisionally, that there is an experiment $\widetilde{y} \in \mathcal{E}$, whose support, (that is, the set $S(\widetilde{y}) := \{y \in Y \mid \Sigma_{t \in T} \mu_t \left(\widetilde{y}^{-1}(y)\right) \mu(t) > 0\}$) has cardinality that is at least as great as that of the set of theories, T (i.e.,$\mid S(\widetilde{y}) \mid \geq m$). Let $\Upsilon(\widetilde{y}) = (Y_1, ..., Y_m)$ be a partition the set Y, and denote by Υ_m the set of all $m-$elements partitions of Y.[8]

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, $P(Y_i), i = 1, ..., m$, and let $P(\widetilde{y}) := (P(Y_1), ..., P(Y_m))$. By Theorem 2, for all $Y_i \in \Upsilon$, $P(Y_i) = \cup_{y \in Y_i} \Sigma_{t \in T} \mu_t \left(\widetilde{y}^{-1}(y)\right) \mu(t)$.

For each $Y_i \in \Upsilon(\widetilde{y})$ and $t \in T$ define $\xi_t(Y_i) = \cup_{y \in Y_i} \mu_t \left(\widetilde{y}^{-1}(y)\right)$. Let $\boldsymbol{\mu}(\Upsilon(\widetilde{y})) := (\mu(t_1 \mid \Upsilon(\widetilde{y})), ..., \mu(t_m \mid \Upsilon(\widetilde{y})))$, then $A\boldsymbol{\mu}^{\tau}(\Upsilon(\widetilde{y})) = (P(Y_1), ..., P(Y_{m-1}), 1)^{\tau}$, where the

[8] If $\mid S(\widetilde{y}) \mid= m$ then cells of the partition $(y_1, ..., y_m)$, are singlton sets, each containing an element of the support of $\widetilde{y}$.

superscript τ is the transpose and A is the $m \times m$ matrix given by:

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

The following proposition is immediate:

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

Assume that, for every $\Upsilon \in \Upsilon_m$, $\boldsymbol{\mu}(\Upsilon(\widetilde{y}))$ exists. A prior probability distribution $\boldsymbol{\mu}$ on T is said to *represesnt a decision maker's beliefs* if and only if $\boldsymbol{\mu}(\Upsilon(\widetilde{y})) = \boldsymbol{\mu}(\Upsilon'(\widetilde{y})) = \boldsymbol{\mu}(\Upsilon(\widetilde{y}')) = \boldsymbol{\mu}(\Upsilon'(\widetilde{y}')) = \boldsymbol{\mu}$ for all $\Upsilon(\widetilde{y}), \Upsilon'(\widetilde{y}), \Upsilon(\widetilde{y}'), \Upsilon'(\widetilde{y}') \in \Upsilon_m$ and $\widetilde{y}, \widetilde{y}' \in \mathcal{E}$.

Note that this elicitation mechanism may not work if it is applied to general acts rather than to experiments. The reason is that the in general, the valuations of the payoffs in all the known schemes designed elicit the probabilities $P(f) := (P(X_1), ..., P(X_m))$ of a partition of the set of outcomes X may not be outcome-independent. Therefore, the elicitation mechanism may fail to produce reliable values of the probabilities of the elements of the partition.[9]

If no single experiment has support that includes larger number of observations than there are theories (that is, the case in which, for no $\widetilde{y} \in \mathcal{E}$, $| S(\widetilde{y}) | \geq m$) but $| \cup_{\widetilde{y} \in \mathcal{E}} S(\widetilde{y}) | \geq m$, it is possible to apply a modified version of the mechanism described above. To simplify the exposition, without loss of generality, let there be K experiments, $\{\widetilde{y}_1, ..., \widetilde{y}_K\}$. Let $s_k :=| S(\widetilde{y}_k) |$, $k = 1, ..., K$, and suppose that $\Sigma_{k=1}^K s_k = m$.[10] 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 $P(\widetilde{y}_k) := (P(y_{k,1}), ..., P(y_{k,s_k}))$ is elicited using one of the standard procedures. Consider the system of equations $\hat{A}\boldsymbol{\pi}^\tau = \hat{P}^\tau$, where

$$\hat{P} := (P(y_{1,1}), ..., P(y_{1,\, s_1-1}), ..., P(y_{K,1}), ..., P(y_{K,s_K-1}), 1)$$

[9] This is a version of the familiar state-dependent problem with the known elicitations procedures.

[10] 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 $\boldsymbol{\mu}$ 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 $\hat{A}\boldsymbol{\mu}^{\tau} = \hat{P}^{\tau}$.

6 Discussion

6.1 States and consequences

Perhaps the simplest way to explain the distinction between the traditional analytical framework employed in the theories of decision making under uncertainty and the one proposed here is by reviewing the similarities and differences with the model of Anscombe and Aumann (1963). The apparent similarity is due to the fact that because the theory-dependent consequences of acts are distributions on X (that is, $\mu_t\left(f^{-1}(\cdot)\right) \in \Delta X$, for all $(t, f) \in T \times F$) it may seem as if the decision model in this paper is analogous to that of Anscombe and Aumann with theories replacing the states. This conclusion is wrong for several reasons. Most importantly, in the Anscombe-Aumann model states are *directly observable* while theories are not. In other words, the Anscombe-Aumann states, whether they are the outcomes of actions, as in the case of running a horse race, or naturally occurring, as in the case of the weather, are verifiable independently of the acts.[11] By contrast, the truth of theories is not directly observable and their validity may not be verified *except through the acts or experiments*. Furthermore, because the same observations may be consistent with distinct theories, in general, observing act-outcome pairs is not sufficient to identify, indirectly, which of the underlying theories is true. Another

[11] This is a necessary condition for the payoffs of the acts to be effectuated.

difference has to do with the interpretation of the subjective probabilities. Unlike the Anscombe-Aumann model in which a decision makers' entertain beliefs, represented by subjective probabilities, on the states (i.e., the outcomes of a horse race or the weather), in model of this paper, the probabilities of the outcomes of a horse race are objective forecasts of "horse-race theories" or weather-forecasts models. By contrast, decision makers' entertain beliefs, represented by subjective probabilities, in the truth of the alternative "horse-race theories" or weather-forecasts models. Finally, according to Anscombe and Aumann, running a horse race is an action that determines the relevant state space while, in the present model, it is an experiment and its possible outcomes (i.e., the Anscombe-Aumann states) are signals that inform about the "horse-race theories" on the basis of which the likely outcomes are forecasted.

An important shortcoming of the traditional analytical framework is that the notion of states and consequences may be confounded. This problem was 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 a fair coin. According to Aumann, even supposing that the husband believes that his wife has an even chance of surviving the operation, he still rather bet on her survival. This is because winning the bet on the outcome of a coin flip, the husband the reward in the event that the wife 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

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

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 the diagnosis is correct. This does not rule out, however, the realistic possibility of choosing the hospital or the surgeon to perform the operation. In terms of the present model, these options correspond to distinct acts. There are alternative hypotheses regarding the effect on the probability of success of the surgeon performing the operation and the hospital in which it is performed. In particular, surgeons that preform the operation frequently and hospitals in which such operations are routine have better 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 decision makers is capable choosing among such feasible acts. We can now analyze the husbands behavior using the present model. 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:

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

The probability of the wife's survival, $\Sigma_{t\in T}\mu_t\left(f^{-1}\left(L\right)\right)\mu\left(t\right)$, depends on the husband's belief, quantified by μ, regarding the likely validity of the theories and f the choice of a surgeon and hospital in which to perform the operation.

Given the act f, let $\beta\left(f,L\right)=z_{f^{-1}(L)}z'$, where $z>0>z'$ denote a bet on the survival of the wife conditional on f being implemented. Thus, the bet $\beta\left(f,L\right)$ pays off z dollars in the event, $f^{-1}\left(L\right)$, that the wife survived, and z' dollars otherwise. Such a bet yields the expected utility payoff

$$u\left(L,w+z\right)\Pr\left(L\right)+u\left(D,w+z'\right)\left(1-\Pr\left(L\right)\right),$$

where $\Pr\left(L\right)=\Sigma_{t\in T}\mu_t\left(f^{-1}\left(L\right)\right)\mu\left(t\right)$.

Let g be the act of flipping a fair coin. Then a $\beta\left(g,H\right)=z_{g^{-1}(H)}z'$ denotes the bet on heads up on a flip of a fair coin. Thus, $\beta\left(g,H\right)$ pay off z dollars in the event heads up, $g^{-1}\left(H\right)$ and z' dollars in the complementary event, $g^{-1}\left(T\right)$. Assuming, that the husband believes that the coin is fair means, the probability of $H=1/2$. Supposing that the outcomes of the surgery and the coin flip are stochastically independent, the expected utility of betting on the heads up is:

$$\begin{aligned}&[u(L, w+z)\Pr(L) + u(D, w+z)(1-\Pr(L))]/2 + \\ &\left[u\left(L, w+z'\right)\Pr(L) + u\left(D, w+z'\right)(1-\Pr(L))\right]/2.\end{aligned}$$

In Aumann's example the choice of f is such that $\Pr(L) = 1/2$. Hence, the expected utility of betting on heads up in the coin flip is:

$$\left[u(L, w+z) + u(D, w+z) + u\left(L, w+z'\right) + u\left(D, w+z'\right)\right]/4,$$

and the expected utility of betting on the survival of the wife is $[u(L, w+z) + u(D, w+z')]/2$. According to the scenario described by Aumann, presumably $u(L, w+z) > [u(L, w+z) + u(D, w+z)]/2$ and $u(D, w+z') > [u(L, w+z') + u(D, w+z')]/2$. Hence, 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 true, 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 analogous to a state in the Anscombe-Aumann (1963) model. There are, however, important differences between theories and Anscombe-Aumann states. First, being an abstract idea, theories, unlike states, are not susceptible to be confounded with outcomes. More specifically, Anscombe-Aumann states (e.g., the results of a horse race or states of health) are outcomes in the theory-based model of this paper. Second, unlike states which are verifiable and, therefore, it is possible to bet on them, in general the truth of theories may not be verifiable. Finally, predicting the outcomes of acts on the basis of laws that capture the regularity of the relation between acts and outcomes seems to conform to the way we think and with the scientific method, according to which general laws are parsimonious and efficient mean (i.e., demanding less effort) of describing the environment relevant to a decision problem.

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 this paper the acquisition information is done by experimentation and, as such, the production of signals 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 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 determine these events.

The issue of experimentation as an information acquisition process 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. Savage's notion of experiment and the one advanced here seem to correspond to Popper's (2002) distinction between the term experiment that is synonymous with *an action whose outcome is uncertain* as opposed to experiment which is a terms that denotes *a means of acquiring knowledge, by comparing results obtained with results expected.*

The choice of experiment is itself a decision problem, that amalgamates 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 the 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 $l : S \times A \rightarrow \Delta Y$ 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

(Sufficiency) By the von Neumann-Morgenstern theorem $\succcurlyeq$ is an Archimedean weak order satisfying independence if and only if there is an affine, real-valued, function U on $\mathbb{C}$ such that, for all $(f, \zeta(f)), (f', \zeta'(f')) \in \mathbb{C}$,

$$(f, \zeta(f)) \succcurlyeq (f', \zeta'(f')) \Leftrightarrow U(f, \zeta(f)) \geq U(f', \zeta'(f')).$$

Fix $f^* \in F$ and, for any $f \in F$ and $\zeta, \zeta', \zeta'' \in \mathcal{Z}$ such that $\zeta(f) = \zeta'(f^*) = \zeta''(f^*_{-t} f)$, (that is, $\zeta(f,x)=\zeta'(f^*,x)=\zeta''\left(f^*_{-t}f,x\right)$ for all $x\in X$) by definition,

$$\frac{1}{m}\left(f,\zeta\left(f\right)\right)+\frac{m-1}{m}\left(f^*,\zeta'\left(f^*\right)\right)=\frac{1}{m}\Sigma_{t\in T}\left(f^*_{-t}f,\zeta''\left(f^*_{-t}f\right)\right).$$

By the affinity of U,

$$\frac{1}{m}U\left(f,\zeta\left(f\right)\right)+\frac{m-1}{m}U\left(f^*,\zeta^*\left(f^*\right)\right)=\frac{1}{m}\Sigma_{t\in T}U\left(f^*_{-t}f,\zeta''\left(f^*_{-t}f\right)\right).$$

For each $t\in T$ let $\zeta\left(f\right)\left(t\right):=\left(\zeta\left(f,x_1\right)\left(t\right),...,\zeta\left(f,x_n\right)\left(t\right)\right)\in\left(\Delta X\right)^n$ and define a function $W:T\times\Delta X\times\left(\Delta X\right)^n\to\mathbb{R}$ by:

$$W\left(t,p,\vec{p})\right)=U\left(\left(f^*_{-t}f\right),\zeta''\left(f^*_{-t}f\right)\right)-\frac{m-1}{m}U\left(f^*\left(t\right),\zeta'\left(f^*\right)\right),$$

where $p=\left(f^*_{-t}f\right)\left(t\right)$ and $\vec{p}=\zeta''\left(f^*_{-t}f\right)\left(t\right)=\zeta'\left(f^*\right)\left(t\right)$. Thus,

$$\frac{1}{m}\Sigma_{t\in T}W\left(t,f(t),\zeta\left(f\right)\left(t\right)\right)=\frac{1}{m}\Sigma_{t\in T}U\left(\left(f^*_{-t}f\right)\left(t\right),\zeta''\left(f^*_{-t}f\right)\left(t\right)\right)-\frac{m-1}{m}U\left(f^*\left(t\right),\zeta'\left(f^*\right)\left(t\right)\right).$$

Hence,

$$U\left(f,\zeta\left(f\right)\right)=\Sigma_{t\in T}W\left(t,f\left(t\right),\zeta\left(f\right)\left(t\right)\right).$$

By the affinity of U, $W\left(t,\cdot,\cdot\right)$ is affine in its second and third arguments.

By theory-independence, for all ex-ante relevant $t,t'\in T$ and $f,f'\in F$ and all $\zeta,\zeta',\zeta'',\zeta'''\in\mathcal{Z}$, $\zeta\left(f^*_{-t}f'\right)=\zeta'\left(f^*_{-t}f'\right)$ and $\zeta''\left(f^*_{-t'}f\right)=\zeta'''\left(f^*_{-t'}f'\right)$,

$$\left(f^*_{-t}f,\zeta\left(f^*_{-t}f\right)\right)\succcurlyeq\left(f^*_{-t}f',\zeta'\left(f^*_{-t}f'\right)\right)$$

if and only if

$$\left(f^*_{-t'}f,\zeta''\left(f^*_{-t'}f\right)\right)\succcurlyeq\left(f^*_{-t}f',\zeta'''\left(f^*_{-t'}f'\right)\right).$$

Then, by the additivity of U across theories, for all ex-ante relevant $t,t'\in T$,

$$W\left(t,f\left(t\right),\zeta\left(f\right)\left(t\right)\right)\geq W\left(t,f'\left(t\right),\zeta'\left(f'\right)\left(t\right)\right)$$

if and only if

$$W\left(t',f\left(t'\right),\zeta''\left(f\right)\left(t'\right)\right)\geq W\left(t',f'\left(t'\right),\zeta'''\left(f'\right)\left(t'\right)\right).$$

Define $w\left(\cdot,\cdot\right):\Delta X\times\left(\Delta X\right)^m\to\mathbb{R}$ by $w\left(\cdot,\cdot\right)=W\left(t_1,\cdot,\cdot\right)$.

Then, the uniqueness of the affine utility representation on ΔX, implies that, for every $\vec{p} \in (\Delta X)^n$,

$$W(t, \cdot, \vec{p}) = b_t w(\cdot, \vec{p}) + a_t,$$

where $b_t > 0$, for all ex-ante relevant $t \in T$ and $b_t = 0$ for all ex-ante irrelevant $t \in T$. By nontriviality, $\Sigma_{t' \in T} b_{t'} > 0$. Define $\mu(t) = b_t / \Sigma_{t' \in T} b_{t'}$, for all $t \in T$, then

$$U(f, \zeta(f)) = (\Sigma_{t' \in T} b_{t'}) \Sigma_{t \in T} \mu(t) w(f(t), \zeta(f)(t)) + \Sigma_{t \in T} a_t.$$

Thus,

$$U(f, \zeta(f)) \geq U(f', \zeta'(f'))$$

if and only if

$$\Sigma_{t \in T} \mu(t) w(f(t), \zeta(f)(t)) \geq \Sigma_{t \in T} \mu(t) w\left(f'(t), \zeta'(f')(t)\right). \tag{2}$$

Define $\delta_x \in F$ by $\Sigma_{t \in T} \mu_t\left(\delta_x^{-1}(x')\right) \mu(t) = 1$ if $x' = x$ and $\Sigma_{t \in T} \mu_t\left(\delta_x^{-1}(x')\right) \mu(t) = 0$, otherwise. Define a function $\hat{U} : X \times F \to \mathbb{R}$ by $\hat{U}\left(x, \zeta\left(\delta_x^{-1}\right)\right) = w\left(\delta_x, \zeta\left(\delta_x^{-1}\right)\right)$, for all $\zeta \in \mathcal{Z}$. By the affinity of w and induction on the size the support (see Kreps (1988) p. 50) we have

$$w(f(t), \zeta(f)(t)) = \Sigma_{x \in X} \hat{U}\left(x, \zeta\left(\delta_x^{-1}, x\right)\right) \mu_t\left(f^{-1}(x)\right).$$

Hence, by (2),

$$U(f, \zeta(f)) \geq U(f', \zeta'(f'))$$

if and only if

$$\Sigma_{x \in X} \hat{U}(x, \zeta(f, x)) \Sigma_{t \in T} \mu_t\left(f^{-1}(x)\right) \mu(t) \geq \Sigma_{x \in X} \hat{U}\left(x, \zeta'(f', x)\right) \Sigma_{t \in T} \mu_t\left(f'^{-1}(x)\right) \mu(t). \tag{3}$$

Consider next the function $\hat{U}(x, \zeta(f', x))$. Fix $\zeta^* \in \mathcal{Z}$ and $(f, x) \in F \times X$ define $\left(\zeta^*(f, x)_{-t} \zeta(f, x)\right)(t') = \zeta^*(f, x)(t')$ if $t' \in T \backslash \{t\}$ and $\left(\zeta^*(f, x)_{-t} \zeta(f, x)\right)(t') = \zeta(f, x)(t')$, otherwise. Then, by the same argument as above

$$\frac{1}{m}(f, \zeta(f, x)) + \frac{m-1}{m}(f, \zeta^*(f, x)) = \frac{1}{m} \Sigma_{t \in T}(f, \zeta^*(f, x)_{-t} \zeta(f, x)).$$

Thus, by (3) and the affinity of $\hat{U}(x, \cdot)$,

$$\Sigma_{x\in X}\left[\frac{1}{m}\hat{U}\left(x,\zeta\left(f,x\right)\right)+\frac{m-1}{m}\hat{U}\left(x,\zeta\left(f,x\right)\right)\right]\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right)$$
$$= \frac{1}{m}\Sigma_{t'\in T}\left[\Sigma_{x\in X}\hat{U}\left(x,\zeta^*\left(f,x\right)_{-t'}\zeta(f,x)\right)\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right)\right].$$

Define a function $H : T \times X \times \Delta X \to \mathbb{R}$ by:

$$H\left(t,x,p\right)=\hat{U}\left(x,\zeta^*\left(f,x\right)_{-t'}\zeta(f,x)\right)-\frac{m-1}{m}\left(f,\zeta^*\left(f,x\right)\right).$$

Then, by the same argument as above,

$$\hat{U}\left(x,\zeta\left(f,x\right)\right)=\Sigma_{t\in T}H\left(t,x,\zeta\left(f,x\right)\left(t\right)\right). \tag{4}$$

For all ex-post relevant $t,t'\in T$ and $\zeta,\zeta'\in\mathcal{Z}$, by theory independence, for every given $(f,x)\in F\times X$,

$$\left(f,\zeta^*\left(f,x\right)_{-t}\zeta\left(f,x\right)\right)\succcurlyeq\left(f,\zeta^*\left(f,x\right)_{-t}\zeta'\left(f,x\right)\right)$$

if and only if,

$$\left(f,\zeta^*\left(f,x\right)_{-t'}\zeta\left(f,x\right)\right)\succcurlyeq\left(f,\zeta^*\left(f,x\right)_{-t'}\zeta'\left(f,x\right)\right).$$

Hence, by (4),

$$\Sigma_{x\in X}H\left(t,x,\zeta\left(f,x\right)\left(t\right)\right)\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right)\geq\Sigma_{x\in X}H\left(t,x,\zeta'\left(f,x\right)\left(t\right)\right)\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right)$$

if and only if

$$\Sigma_{x\in X}H\left(t',x,\zeta\left(f,x\right)\left(t\right)\right)\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right)\geq\Sigma_{x\in X}H\left(t',x,\zeta'\left(f,x\right)\left(t\right)\right)\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right).$$

Let $h\left(x,p\right):=H\left(t_1,x,p\right)$ for all $(x,p)\in X\times\Delta X$, then, by the uniqueness of the affine representation, for all $t\in T$,

$$H\left(t,x,\zeta\left(f,x\right)\left(t\right)\right)=\hat{b}_t\left(f,x\right)h\left(x,\zeta\left(f,x\right)\left(t\right)\right)+\hat{a}_t\left(f,x\right)$$

By nontriviality, $\Sigma_{t\in T}\hat{b}_t\left(f,x\right)>0$. Let $\mu\left(t\mid f,x\right):=\hat{b}_t\left(f,x\right)/\Sigma_{t'\in T}\hat{b}_{t'}\left(f,x\right)$. Define $u : X\times X\to\mathbb{R}$ by $u\left(x,x'\right)=h\left(x,\zeta\left(\delta_{x'},x'\right)\right)$. Then, by the affinity of $h\left(x,\cdot\right)$, we have $h\left(x,\zeta\left(f,x\right)\left(t\right)\right)=\Sigma_{x'\in X}u\left(x,x'\right)\mu_t\left(\zeta\left(f,x\right)^{-1}\left(x'\right)\right)$ and, by (4),

$$\hat{U}\left(x,\zeta\left(f,x\right)\right)=\Sigma_{t'\in T}\hat{b}_{t'}\left(f,x\right)\Sigma_{x'\in X}u\left(x,x'\right)\Sigma_{t\in T}\mu_t\left(\zeta\left(f,x\right)^{-1}\left(x'\right)\right)\mu\left(t\mid f,x\right)+\Sigma_{t\in T}\mu_t\hat{a}_t\left(f,x\right) \tag{5}$$

Combining (3) and (5) we obtain:

$$U\left(f,\zeta\left(f\right)\right)\geq U(f',\zeta'\left(f'\right))$$

if and only if

$$\begin{aligned} & \Sigma_{x\in X}\left[\Sigma_{x'\in X}u\left(x,x'\right)\Sigma_{t\in T}\mu_t\left(\zeta\left(f,x\right)^{-1}\left(x'\right)\right)\mu\left(t\mid f,x\right)\right]\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\right) \\ \geq\; & \Sigma_{x\in X}\left[\Sigma_{x'\in X}u\left(x,x'\right)\Sigma_{t\in T}\mu_t\left(\zeta'\left(f',x\right)^{-1}\left(x'\right)\right)\mu\left(t\mid f,x\right)\right]\Sigma_{t\in T}\mu_t\left(f'^{-1}\left(x\right)\right)\mu\left(t\right). \end{aligned}$$

(Necessity) The necessity of weak order, Archimedean and independence forllow from the von Neumann-Morgenstern theorem. The necessity of theory-independece is immediate.

The uniqueness part follows form the uniqueness of U. ■

7.2 Proof of theorem 2

Proof. By Theorem 1, a preference relation $\succcurlyeq$ on $\mathcal{C}$ is nontrivial, continuous weak order satisfying independence and theory- independence if and only if it admits the representation: $\left(\widetilde{y},\zeta\left(\widetilde{y}\right)\right)\mapsto U\left(\widetilde{y},\zeta\left(\widetilde{y}\right)\right)$, where

$$U\left(\widetilde{y},\zeta\left(\widetilde{y}\right)\right)=\Sigma_{y\in Y}\left[\Sigma_{x'\in X}u\left(y,x\right)\Sigma_{t\in T}\mu_t\left(\zeta\left(\widetilde{y},y\right)^{-1}\left(x'\right)\right)\mu\left(t\mid\widetilde{y},y\right)\right]\Sigma_{t\in T}\mu_t\left(\widetilde{y}^{-1}\left(y\right)\right)\mu\left(t\right).$$

Let $\widetilde{y}$ be uninformative then, $\mu_t\left(\widetilde{y}^{-1}\left(y\right)\right)=\mu_t\left(\widetilde{y}^{-1}\left(y'\right)\right)$, for all $y,y'\in T$. Hence, $\mu\left(t\mid\widetilde{y},y\right)=\mu\left(t\mid\widetilde{y},y'\right)$, for all $y,y'\in T$, where $\mu\left(t\mid\widetilde{y},y\right)=\mu_t\left(\widetilde{y}^{-1}\left(y\right)\right)/\Sigma_{t'\in T}\mu_{t'}\left(\widetilde{y}^{-1}\left(y\right)\right)$. Let $\Pr\left(x\mid f,\widetilde{y},y\right):=\Sigma_{t\in T}\mu_t\left(f^{-1}\left(x\right)\right)\mu\left(t\mid\widetilde{y},y\right)$, for all $x\in X$, then, by signal exchangeability,

$$U\left(\widetilde{y},\zeta\left(\widetilde{y}\right)_{-y}f_{-y'}f'\right)=\Sigma_{x\in X}u\left(y,x\right)\Pr\left(x\mid f,\widetilde{y},y\right)+\Sigma_{x\in X}u\left(y',x\right)\Pr\left(x\mid f',\widetilde{y},y'\right)=$$

$$\Sigma_{x\in X}u(y,x)\Pr\left(x\mid f',\widetilde{y},y\right)+\Sigma_{x\in X}u\left(y',x\right)\Pr\left(x\mid f,\widetilde{y},y'\right)=U\left(\widetilde{y},\zeta\left(\widetilde{y}\right)_{-y}f'_{-y'}f\right).$$

But $\mu_t\left(f^{-1}\left(\cdot\right)\right)\neq\mu_t\left(f'^{-1}\left(\cdot\right)\right)$ for some $f,f'\in F$ imply that

$$\Sigma_{x\in X}u\left(x\right)\Pr\left(x\mid f,\widetilde{y},y\right)\neq\Sigma_{x\in X}u\left(x\right)\Pr\left(x\mid f',\widetilde{y},y'\right).$$

Hence, the equalities above imply that $u(y, x) = u(y', x) = u(x)$, for all $x \in X$. Therefore,

$$U(\widetilde{y}, \zeta(\widetilde{y})) = \Sigma_{y \in Y} \left[\Sigma_{x \in X} u(x) \Sigma_{t \in T} \mu_t \left(\zeta(\widetilde{y}, y)^{-1}(x) \right) \mu(t \mid \widetilde{y}, y) \right] \Sigma_{t \in T} \mu_t \left(\widetilde{y}^{-1}(y) \right) \mu(t).$$

To show the necessity of signal exchangeability is strightforward.

The uniqueness follows from 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. (2021) "Dynamic Belief Elicitation," *Econometrica*, 89, 375–414..

[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] Popper, Karl (2002) *The Poverty of Historicism.* Routledge, London and New York.

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

[17] 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).

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

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

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

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