Decision making under uncertainty comprising complete ignorance and probability

Phan H. Giang

George Mason University, MS 1J3
4400 University Dr., Fairfax, VA 22030, USA
pgiang@gmu.edu

Abstract
This paper investigates a model of decision making under uncertainty comprising opposite epistemic states of complete ignorance and probability. In the first part, a new utility theory under complete ignorance is developed that combines Hurwicz-Arrow’s theory of decision under ignorance with Anscombe-Aumann’s idea of reversibility and monotonicity used to characterize subjective probability. The main result is a representation theorem for preference under ignorance by a particular one-parameter function – the $\tau$-anchor utility function. In the second part, we study decision making under uncertainty comprising an ignorant variable and a probabilistic variable. We show that even if the variables are independent, they are not reversible in the Anscombe-Aumann’s sense. This insight leads to the development of a new proposal for decision under uncertainty represented by a preference relation that satisfies the weak order and monotonicity assumptions but rejects the reversibility assumption. A distinctive feature of the new proposal is that the certainty equivalent of a mapping from the state space of uncertain variables to the prize space depends on the order in which the variables are revealed. Explicit modeling of the order of variables explains some of the puzzles in multiple-prior model and the models for decision making with Dempster-Shafer belief function.

Keywords: Decision making, ignorance, probability, order of variables

1. Introduction
Ignorance and probability are opposite states of knowledge. On the one hand, probability, according to the aleatory interpretation, is a result of
knowing all that can be reasonably known about a phenomenon so that its outcome can be modeled as a random event, quantitatively indistinguishable from the randomness of coin toss, roulette spin or radioactive decay. Ignorance, on the other hand, is a singular state of knowledge characterized by knowing nothing or having no reliable information about the phenomenon of interest. Under this extreme state of uncertainty, it is impossible to say that an event except tautology is strictly more likely than another event. We hold the view that the two extreme and opposite states of knowledge form the basis on which other epistemic states are “spanned”. This is the main motivation to investigate the models of decision making under uncertainty comprising both ignorance and probability. AI agents would not be truly intelligent without the capability to make decision in situations of ignorance.

Let’s start with a motivating example. An investor is considering at the end of 2014 a one-year investment instrument that matures on 1 January 2016. The return on the investment depends on two uncertain variables. The first source of uncertainty is the prospect of a political settlement in country A (e.g. Afghanistan) and the second source of uncertainty is the prospect of 2015 coffee crop in country B (e.g. Brazil). The 2015 coffee harvest, denoted by $C$, can be either bumper ($b$) or normal ($n$) or poor ($p$). On the one hand, from extensive historical data, the probability distribution of Brazilian coffee crop in 2015 is estimated to be $(0.46, 0.2, 0.34)$ where the numbers are the chances of having bumper, normal and poor crop respectively. On the other hand, the political settlement variable, denoted by $S$, is modeled with two possible values: peaceful settlement among fighting factions in 2015 ($s$) or lack thereof ($\sim s$). The experts whose advice the investor seeks on the political settlement question, offer contradictory opinions and cannot come to any agreement. This underlies the fact that nobody knows the true driving forces behind a political settlement in that region of the world. The returns on the investment are given in the following table.

<table>
<thead>
<tr>
<th>$S/C$</th>
<th>$b$</th>
<th>$n$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s$</td>
<td>-3.0%</td>
<td>2.0%</td>
<td>8.0%</td>
</tr>
<tr>
<td>$\sim s$</td>
<td>5.0%</td>
<td>1.5%</td>
<td>-4.0%</td>
</tr>
</tbody>
</table>

Table 1: The returns on investment.

For example, if there will be a bumper coffee crop and a political settlement then the investment has negative return of $-3\%$. The question for the investor is whether she should invest in the financial instrument or keep
her money in the bank that pays interest of 1.2%.

A partial solution for this example is given in section 2.4 and a complete solution is given in section 5.1.

This example illustrates how ignorance and probability can coexist in our understanding of real world situations. Ignorance can arise when data and information are scarce, unreliable and contradictory. A player in competitive games may find herself in the state of ignorance if she suspects the information she has about the opponent is intentionally misleading. In the risk assessment practice, ignorance refers to the situations where there is a high level of uncertainty on both the likelihood of events and the consequences of such events [3]. A key condition that differentiates ignorance from uncertainty is the absence of knowledge about the factors that influence the issues [15]. Another condition often associated with the term “ignorance” is the sample space ignorance (SSI) [25] when the decision maker has difficulty in determining the set of alternative states.

Historically, with the development of decision theory under risk in the late 1940s, economists such as Shackle [28], Hurwicz and Arrow [2] started pondering the question how an individual makes decisions if she cannot associate any probability distribution to possible consequences. Earlier, in 1920s, Knight [20] and Keynes [17] came to conclusion that probability theory cannot capture all relevant aspects of uncertainty. In particular, the principle of insufficient reason that produces a probability distribution under ignorance by attributing equal probabilities to alternatives is not always appropriate because the same real situation can be modeled in different ways, with different sets of alternatives. Recently, Gilboa et al argue that rationality requires a compromise between internal coherence and justification, similarly to compromises that appear in moral dilemmas [12]. Under that view, “it is more rational to admit that one does not have sufficient information to generate a prior than to pretend that one does”.

This paper is structured as follows. Before tackling the problem of decision making under uncertainty involving both ignorance and probability we consider two special cases when uncertainty is either pure probability or complete ignorance. For the former we adopt the standard expected utility. For the latter, in section 2, a new utility theory under ignorance is developed by pulling together Hurwicz-Arrow’s decision theory with Anscombe-Aumann’s idea of reversibility of independent variables. The main result is a representation theorem for preference under ignorance. In section 3 we consider the decision problem that involves both ignorance and probability. We obtain a
negative result showing that the reversibility between ignorance and probability is not possible even when they are independent. In section 4, a solution is proposed that calls for explicit modeling of the order of variables. Section 5 has examples and application. Section 6 contains the discussion of related literature. All proofs are relegated to the appendix.

Before going into technical presentation, we briefly describe the notation convention used in this paper. The basic objects are uncertain variables, denoted by upper case letters such as $I, R$. A variable has a domain denoted by $\Omega$ indexed by variable’s name. Each variable is accompanied by a measure of uncertainty that gives the variable its type. In this paper, two types are allowed: ignorant variable and probabilistic variable. $\mathcal{O}$ is the set of quantifiable prizes/rewards. $\mathcal{O}$ is assumed to be the real unit interval $\mathbb{R}_{0,1}$.

A mapping from the Cartesian product of the domains of variables to the set of prizes is an act. Acts are denoted by lower case letters such as $d, f, g$. The behavior of an individual decision maker is described by a preference relation $\succeq$ on acts. We use different notations for acts to signal the type of associated uncertainty. Under Hurwicz-Arrow’s assumptions for preference under ignorance, the state space on which an act is defined is not important, the act can be identified with its set of prizes. So, HA acts are denoted by set notation $\{x_1, x_2, \ldots\}$. The set of HA acts is denoted by $\mathcal{F}(\mathcal{O})$. For acts that are defined (sequentially) on two ignorant variables we use the notation of set of sets $\{\{x_1, x_2\}, \{y_1, y_2\}, \ldots\}$. The set of such acts is denoted by $\mathcal{F}(\mathcal{F}(\mathcal{O}))$. An act defined on a probabilistic variable is denoted by a list of pairs $(p_1 : x_1, p_2 : x_2, \ldots, p_n : x_n)$ where $p_i$ is the probability of getting reward $x_i$. Finally, acts that are defined on an ignorant variable and a probabilistic variable are denoted by a set of lists (in case the ignorant variable precedes the probabilistic one) or a list of sets (in case the order is reversed). In this notation, the investment instrument in Table 1 can be coded as follows:

$$\{(0.46 : x_{11}, 0.2 : x_{12}, 0.34 : x_{13}), (0.46 : x_{21}, 0.2 : x_{22}, 0.34 : x_{23})\}$$

in case the political settlement is known before the coffee harvest or

$$(0.46 : \{x_{11}, x_{21}\}, 0.2 : \{x_{12}, x_{22}\}, 0.34 : \{x_{13}, x_{23}\})$$

if the order is reversed. For example, $x_{11}$ is the normalized prize in the unit interval corresponding to the return of $-3\%$ and $x_{12}$ corresponds to $2\%$. Table 2 has the details of the conversion from investment returns to normalized prizes.
2. A utility theory under ignorance

Our theory of decision under ignorance\(^1\) is a result of the marriage between Hurwicz-Arrow’s theory [2] and Anscombe-Aumann’s theory [1].

The ground-breaking result of decision under ignorance was made in early 1950s by Hurwicz & Arrow (HA) [2]. Their basic construct is a choice operator that for each set of acts \(D\) returns a non-empty subset of optimal acts \(\hat{D} \subseteq D\). They postulated four rationality properties that the choice operator must satisfy and proved a theorem comprising two parts: (a) among the prizes of an act, only the best and the worst prizes matter and (b) if both extreme prizes of an act are higher than the corresponding values of another then the former is preferable to the latter.

Anscombe and Aumann’s (AA) objective in [1] was to define subjective probability (e.g., on a Horse race) on the basis of objective probability (e.g., on Roulette spin). Basic objects in AA theory are three stage acts. The uncertainty in the first stage and the last stage is characterized by objective probability (Roulette). It is assumed that probabilistic lotteries are evaluated by expected utility. The aim is to characterize the subjective uncertainty in the second stage (Horse) via internal consistency requirements. They accomplished that by “apply utility theory twice over, and then connect the two systems of preferences and utilities” ([1] p.201). Two critical assumptions in AA approach are Monotonicity in prizes and the Reversal of order between variables. In a tree representation of an act, the monotonicity requires that the replacement of a subtree of an act by a better subtree would not make it less preferable. The reversal of order assumption requires that the acts, obtained by shifting probability between stages, are preferentially indifferent. The rationale of this assumption is, as Anscombe and Aumann put it, “if the prize you receive is to be determined by both a horse race and the spin of a roulette wheel, then it is immaterial whether the wheel is spun before or after the race”. It is important to note that the force of this argument does not rely on the nature of variables (probabilistic or not) but from the fact that the variables are independent (in a broad sense that knowing one does not affect the belief about the other).

\(^1\)A preliminary result in this section was reported in [10].
2.1. Hurwicz-Arrow's theory of choice under ignorance

In a short paper [2], reprinted in 1977, Hurwicz & Arrow outlined the theory of decision under ignorance developed in early 1950s. In this section we review the basic results of HA theory in our setting. Consider a collection of variables \( \{I_1, I_2, \ldots \} \) whose domains are sets \( \Omega_{I_i} \) which are finite subsets of the set of natural numbers \( \mathbb{N} \). A decision or act defined on variable \( I_i \) is a mapping of the form \( f : \Omega_{I_i} \to O \). The set of prizes, \( O \), is assumed to be the real unit interval \( \mathbb{R}_{0,1} \). The domain of \( f \) is denoted by \( \Omega(f) \). A decision problem is a non-empty set of decisions on the same domain. Denote the set of acts defined on variable \( I \) by \( D_I \) and the set of all acts by \( D \).

In this review, the HA optimal operator construct, which maps each decision problem \( A \) to a subset of optimal acts \( \hat{A} \), is replaced by a preference relation \( \succeq \) on the set of acts \( D \). The correspondence between them is as follows: \( f \in \hat{A} \) means \( \forall g \in A, f \succeq g \).

\( \succeq \) in HA theory must satisfy four properties, labeled after letters A to D. Property A stipulates that \( \succeq \) is a weak order. Property B requires invariance under relabeling. Formally, decisions \( f_1, f_2 \) are isomorphic if there is an one-to-one mapping \( h : \Omega(f_1) \to \Omega(f_2) \) such that \( \forall s \in \Omega(f_1), f_1(s) = f_2(h(s)) \). Property B requires that isomorphic acts are preferentially indifferent. Property C requires invariance under deletion of duplicate states. Formally, \( f_2 \) is said to be derived from \( f_1 \) by deleting duplicate states if (1) \( \Omega(f_2) \subset \Omega(f_1) \), \( f_1 \) and \( f_2 \) are coincided on \( \Omega(f_2) \) and (2) for each \( w \in \Omega(f_1) - \Omega(f_2) \), there exists \( w' \in \Omega(f_2) \) such that \( f_1(w) = f_1(w') \). Property C requires that if \( f_2 \) is derived from \( f_1 \) then they are preferentially indifferent. Property D is called the weak dominance property. If \( f_1 \) and \( f_2 \) are acts on the same domain \( \Omega \), and \( \forall w \in \Omega, f_1(w) \geq f_2(w) \) then \( f_1 \succeq f_2 \).

While properties A and D are the standard assumptions of rational behavior, the essence of ignorance is captured by properties B and C. B requires that no state in the domain is more important than any other state while C requires that merging/splitting states in a domain does not make it better or worse. C is intended to address the difficulty by the decision maker to determine the "right" set of possible states. In HA theory, states do not matter, only consequences do. States in this sense are not necessarily the states of nature but a result of modeling the nature. For instance, one can simply assume that the states are natural numbers used to index the consequences.
of acts.\footnote{Much later in the context of statistical inference, Walley \cite{30} proposes two criteria that an ignorance belief must satisfy. The embedding principle requires that plausibility of an event $A$ should not depend on the sample space in which $A$ is embedded. The symmetry principle says that all elements in the sample space should be assigned the same plausibility. These two principles are reincarnations of properties B and C in HA theory.}

**Theorem 1 (Hurwicz-Arrow).** The necessary and sufficient condition for preference relation $\succeq$ on the set of acts $\mathcal{D}$ satisfies properties A through D is that there exists a weak ordering $\succeq^2$ on the space of ordered pairs of real numbers $\mathbb{Z}^2 = \{(a, b) \mid 0 \leq a \leq b \leq 1\}$ that satisfies the following properties: (1) if $a \geq a'$ and $b \geq b'$ then $\langle a, b \rangle \succeq^2 \langle a', b' \rangle$; (2) for acts $f, g$

$$f \succeq g \text{ if } \left( \min_{w \in \Omega_f} f(w), \max_{w \in \Omega_f} f(w) \right) \succeq^2 \left( \min_{w \in \Omega_g} g(w), \max_{w \in \Omega_g} g(w) \right). \quad (1)$$

The preference relation that satisfies the conditions of theorem 1 is called **HA preference relation**. The preferential comparison between two acts reduces to comparing its extreme values. Under ignorance the state space is not important, so an act can be identified with its prizes. The set of finite non-empty bags of $\mathcal{O}$ is denoted by $\mathcal{F}(\mathcal{O})$. Technically, a bag (aka multiset) can contain the same item multiple times and in this sense it is different from a set. Using bags instead of sets to denote acts under ignorance permits convenient notation because for mapping $f : \Omega \rightarrow \mathcal{O}$, $f(\Omega)$ - the collection of potential prizes by $f$ - is a bag of values. The results of the paper require no deep property of bags as mathematical objects. Symbols $\mathcal{D}$ and $\mathcal{F}(\mathcal{O})$ denote the set of acts under ignorance.

We make two technical assumptions which are not part of HA theory. When the set of prizes of an act is a singleton, it is a constant act. The preference among constant acts is assumed to be the arithmetic order.

**Assumption 1 (Constant acts - CA).** For $x, y \in \mathcal{O}$, $x \succeq y$ iff $x \geq y$.

**Lemma 1.** Suppose $\succeq$ is HA preference relation on $\mathcal{F}(\mathcal{O})$ and satisfies condition (CA) then for each $A \in \mathcal{F}(\mathcal{O})$ there is a unique $c \in \mathcal{O}$ such that $A \sim c$, furthermore, $\min(A) \leq c \leq \max(A)$. 
The value \( c \) in the lemma is called the **certainty equivalent** (CE) of \( A \). Function \( CE : F(O) \rightarrow O \) such that \( CE(A) \sim A \) is called the **CE operator**.

Viewing an act as a vector of prizes naturally leads to the concept of convergence of a sequence of acts. Suppose \((f_i)_{i=1}^\infty\) is a sequence of acts, in the vector form \( f_i = \{x_{i1}, x_{i2}, \ldots, x_{in}\} \). The sequence \((f_i)_{i=1}^\infty\) is said to converge to act \( f = \{x_1, x_2, \ldots, x_n\} \), \( \lim_{i \to \infty} f_i = f \), if \( \lim_{i \to \infty} x_{ij} = x_j \) for \( 1 \leq j \leq n \).

**Assumption 2 (Continuity of certainty equivalence operator - C).**
If \( \lim_{i \to \infty} f_i = f \) then \( \lim_{i \to \infty} CE(f_i) = CE(f) \).

This assumption simply says that a small change in the prizes of an act under ignorance would not lead to a jump in its CE. The following lemma is an immediate consequence.

**Lemma 2.** If \((f_i)_{i=1}^\infty\) converges to \( f \) and \( \forall i, f_i \succeq g \) then \( f \succeq g \).

Because \( f_i \succeq g \), \( CE(f_i) \succeq CE(g) \). So \( \lim_{i \to \infty} CE(f_i) \succeq CE(g) \). By (C), \( \lim_{i \to \infty} CE(f_i) = CE(f) \). Thus, \( CE(f) \succeq CE(g) \) or equivalently \( f \succeq g \).

The properties of the CE operator of a HA preference relation are summarized in the following lemma.

**Lemma 3.** Suppose \( \succeq \) is a HA preference relation on \( F(O) \) that satisfies properties (CA) and (C) then operator \( CE \) is well defined and satisfies:

1. **Unanimity.** For \( x \in O \), \( CE(x) = x \).
2. **Range.** \( \forall A \in F(O), CE(A) = CE(\{\min(A), \max(A)\}) \).
3. **Monotonicity.** If \( a \geq a' \) and \( b \geq b' \) then \( CE(\{a, b\}) \geq CE(\{a', b'\}) \).
4. **Continuity.** \( \lim_{x \to a} CE(\{x, b\}) = CE(\{a, b\}) \); \( \lim_{x \to b} CE(\{a, x\}) = CE(\{a, b\}) \).

### 2.2. Multiple variables of ignorance

Now let us consider multiple variables of ignorance. Suppose \( X \) and \( Y \) are ignorant variables with domains \( \Omega_X \) and \( \Omega_Y \). We assume that they are mutually independent in the sense that knowing realization of one variable does not change our belief (in this case it is ignorance) about the other. Their join state space \( \Omega_X \times \Omega_Y \) can be thought as the domain of a new variable \( I \) \( (\Omega_I = \Omega_X \times \Omega_Y) \). Naturally, the join variable \( I \) must be an ignorant variable because we know nothing about likelihood of each state in \( \Omega_X \) and \( \Omega_Y \) and even knowing the realization of one does not tell us anything about the other.
The presence of multiple variables opens a sequential perspective on acts. Let’s assume for convenience $|\Omega_X| = m$ and $|\Omega_Y| = n$. Consider the act defined by mapping $f(x_i, y_j) \mapsto z_{ij}$ where $x_i \in \Omega_X$, $y_j \in \Omega_Y$ and $z_{ij} \in \mathcal{O}$ for $1 \leq i \leq m$, $1 \leq j \leq n$. $f$ has two sequential two-stage forms corresponding to the orders of variable realization. If $X$ realizes before $Y$: $f_X = \{f_{x_1}, f_{x_2}, \ldots, f_{x_m}\}$ where $f_{x_i}$ is a mapping $\Omega_Y \rightarrow \mathcal{O}$ and $f_{x_i}(y_j) = z_{ij}$ i.e., $f_{x_i}$ is the set of possible prizes of decision $f$ once $X = x_i$ is known. Similarly, if $Y$ realizes before $X$, then $f_Y = \{f_{y_1}, f_{y_2}, \ldots, f_{y_n}\}$ where $f_{y_j}$ is a function $\Omega_X \rightarrow \mathcal{O}$ such that $f_{y_j}(x_i) = z_{ij}$. Two sequential versions can be written in the matrix form as follows:

\[
\begin{align*}
\{ z_{11}, z_{12}, \ldots, z_{1n} \} \\
\{ z_{21}, z_{22}, \ldots, z_{2n} \} \\
\quad \quad \quad \cdots \\
\{ z_{m1}, z_{m2}, \ldots, z_{mn} \} \\
\end{align*}
\quad \text{and} \quad
\begin{align*}
\{ z_{11}, z_{21}, \ldots, z_{m1} \} \\
\{ z_{12}, z_{22}, \ldots, z_{m2} \} \\
\quad \quad \quad \cdots \\
\{ z_{1n}, z_{2n}, \ldots, z_{mn} \}
\end{align*}
\] (2)

Technically, two-stage act is a finite bag of elements in $\mathcal{F}(\mathcal{O})$. The set of such acts is $\mathcal{F}(\mathcal{F}(\mathcal{O}))$. The preference between acts is now a relation on $\mathcal{F}(\mathcal{F}(\mathcal{O}))$. Clearly $\mathcal{F}(\mathcal{O}) \subset \mathcal{F}(\mathcal{F}(\mathcal{O}))$.

Now the idea of Anscombe-Aumann’s theory [1] comes into play. The basic objects in AA framework are three-stage acts. The uncertainty in the first stage and the third stage is characterized by probability (variable $R$ for roulette spin). The second stage represents variable $H$ (for horse race). The preference on (probabilistic) lotteries is governed by expected utility. They showed that uncertainty on $H$ is representable by a probability measure if the preference relation on acts satisfies the reversibility and the monotonicity assumptions. The monotonicity assumption is a weak dominance requirement. It says that the replacement of a subtree of an act by a preferable subtree yields an act that is at least as good as the original one. The reversal-of-order assumption requires that shifting the order of variables $H$ and $R$ results in an act that is indifferent to the original act. The rationale is that the independence of the variables renders their order immaterial.

In our setting, we have two ignorant variables $X$ and $Y$ instead of $H$ and $R$. The preference relation $\succeq$ on acts under ignorance is assumed to be a HA preference and satisfies the monotonicity and the reversal-of-order assumptions.

**Assumption 3 (Monotonicity under ignorance (MI)).** Suppose for $1 \leq i \leq m$, $z_i, z'_i \in \mathcal{F}(\mathcal{O})$ and $z_i \succeq z'_i$ then $\{z_1, z_2, \ldots, z_m\} \succeq \{z'_1, z'_2, \ldots, z'_m\}$. 


Assumption 4 (Reversibility under ignorance (RI)).

\[
\{z_{ij}\}_{i=1}^{m} \sim \begin{cases} 
\{z_{11}, z_{12}, \ldots, z_{1n}\} \\
\{z_{21}, z_{22}, \ldots, z_{2n}\} \\
\vdots \\
\{z_{m1}, z_{m2}, \ldots, z_{mn}\}
\end{cases} \sim \begin{cases} 
\{z_{11}, z_{21}, \ldots, z_{m1}\} \\
\{z_{12}, z_{22}, \ldots, z_{m2}\} \\
\vdots \\
\{z_{1n}, z_{2n}, \ldots, z_{mn}\}
\end{cases}
\tag{3}
\]

The reversal of order for ignorance requires indifference \( f \sim f_X \sim f_Y \).

Lemma 4 (Iterated certainty equivalence under ignorance). Suppose \( \succeq \) is a HA preference relation on \( \mathcal{F}(\mathcal{O}) \) that satisfies (C), (CA), (MI) and (RI) then for \( A_i \in \mathcal{F}(\mathcal{O}), 1 \leq i \leq m \)

\[
\mathcal{CE}(\bigcup_{i=1}^{m} A_i) = \mathcal{CE}(\{\mathcal{CE}(A_i)| 1 \leq i \leq m\}). \tag{4}
\]

Equation (4) is referred to as the \textit{iterated certainty equivalence} property (ICE).

2.3. Representation theorem

Lemma 3 shows that \( \mathcal{CE} \), a function on finite bags, is completely determined by its component defined on the set of pairs of numbers. Given the certainty equivalence operator \( \mathcal{CE} \), a two-place function \( \gamma : \mathcal{Z}^2 \to \mathbb{R}_0^1 \) where \( \mathcal{Z}^2 = \{(a, b)| 0 \leq a \leq b \leq 1\} \) can be defined as follows:

\[
\forall A \in \mathcal{F}(\mathcal{O}), \ \mathcal{CE}(A) = x \iff \gamma(\min(A), \max(A)) = x. \tag{5}
\]

Lemma 5. Suppose \( \mathcal{CE} : \mathcal{F}(\mathcal{O}) \to \mathcal{O} \) satisfies Unanimity, Range, Monotonicity, Continuity and Iterated certainty equivalence then \( \gamma \), defined via (5), is continuous on each variable and satisfies

1. If for \( x \in \mathbb{R}_0^1 \), \( \gamma(x, 1) = a > x \) then \( \forall y \in [x, a], \ \gamma(y, 1) = \gamma(y, 1) = a. \)
2. If for \( z \in \mathbb{R}_0^1 \), \( \gamma(0, z) = b < z \) then \( \forall y \in [b, z], \ \gamma(0, y) = \gamma(0, y) = b. \)

The lemma is about a "sticky" property of function \( \gamma \). If \( \gamma(x, 1) = a > x \) then an increase in the first parameter in the range between \( x \) and \( a \) would not change the value of the function. (2) is a dual property to (1).

Theorem 2. Suppose \( \gamma : \mathcal{Z}^2 \to \mathbb{R}_0^1 \) the two following statements are equivalent
(1) \( \gamma \) is continuous in each argument and satisfies (i) \( \gamma(x, x) = x \) for \( 0 \leq x \leq 1 \); (ii) \( \gamma(x, y) \geq \gamma(x', y') \) if \( x \geq x', y \geq y' \); (iii) \( \gamma(x, y) = \gamma(\gamma(x, x), \gamma(x, y)) = \gamma(\gamma(x, y), \gamma(y, y)) \) for \( 0 \leq x \leq y \leq 1 \).

(2) There exists a value \( \tau \in [0, 1] \) such that

\[
\gamma(x, y) = \begin{cases} 
  y & \text{if } y < \tau \\
  \tau & \text{if } x \leq \tau \leq y \\
  x & \text{if } x > \tau 
\end{cases}
\]  

(6)

Combining Lemma 5 and theorem 2 we have a representation theorem for decision under ignorance.

**Theorem 3 (Representation theorem).** Suppose that \( \succeq \) is a HA preference relation on \( \mathcal{F}(\mathcal{F}(\mathcal{O})) \), \( \succeq \) satisfies (C), (CA), (MI) and (RI) iff there exists a value \( \tau \in \mathbb{R}_{[0,1]} \) such that the certainty equivalence operator \( \mathcal{CE} \) of \( \succeq \) has the form: for \( A_i \in \mathcal{F}(\mathcal{O}) \), \( 1 \leq i \leq m \)

\[
\mathcal{CE}(A_i) = \begin{cases} 
  \max(A_i) & \text{if } \max(A_i) < \tau \\
  \tau & \text{if } \min(A_i) \leq \tau \leq \max(A_i) \\
  \min(A_i) & \text{if } \min(A_i) > \tau 
\end{cases}
\]  

(7)

\[
\mathcal{CE}(\bigcup_{i=1}^{m} A_i) = \mathcal{CE}([\mathcal{CE}(A_i)|1 \leq i \leq m]).
\]  

(8)

Min and max decision criteria are special cases of (7) with \( \tau = 0 \) and \( \tau = 1 \) respectively. On the other hand, eq. (7) excludes Hurwicz’s \( \alpha \)-criterion as well as the median rule. Besides zero and unity, there is no nontrivial \( \alpha \) that makes Hurwicz’s \( \alpha \)-criterion equivalent to (7).

The value \( \tau \) in (7) is called the characteristic value under ignorance because it characterizes the behavior under ignorance of an individual decision maker. The CE of a set of prizes under ignorance is a point in the interval from the minimum to the maximum of the prize set that minimizes the distance to the characteristic value. Imagine an ideal rubber cord with one end fixed to \( \tau \) and the other end is allowed to move freely within the interval covering the set of prizes. The CE is the point where equilibrium is attained. The function in (7) will be called the \( \tau \)-anchor utility function.

The origami-like surface in Fig. 1 is the plot of the \( \tau \)-utility function with \( \tau = 0.4 \). The XY coordinates are the extremal values of sets of prizes. The CE of the sets are read on Z axis. For example points (0.2, 0.5) or (0.5, 0.2)
on XY surface represent sets of prizes whose minimal element is 0.2 and the maximal element is 0.5. The CE in these cases are 0.4 ($\tau$).

To get an interpretation for the characteristic value one can use the equality $\tau = CE(\{0, 1\})$. $\tau$ is the value that the decision maker would give in exchange for the entire set of prizes [0, 1]. This is a situation of total ignorance. Not only the decision maker is ignorant about the likelihood of variable realization but also no prize in the set of possible prizes is excluded.

The value of $\tau$ can be used to compare the extent of the tolerance for ignorance between different individuals. For two individuals A and B with characteristic values $\tau_A$ and $\tau_B$, we say that A is (strictly) more tolerant for ignorance than B if $\tau_A$ is (strictly) greater than $\tau_B$.

We have a simple corollary on ranking individuals by their tolerance for ignorance.

**Corollary 1.**

(i) If individual A is more tolerant for ignorance than B then for any act $f$ under ignorance $CE_A(f) \geq CE_B(f)$ i.e., the certainty equivalent of $f$ for A is greater than that for B.

(ii) A is strictly more tolerant for ignorance than B iff there exists an act $f$ such that $CE_A(f) > CE_B(f)$. 

Figure 1: The $\tau$-anchor utility function.
In theorem 3, we have seen that characterization of preference under ignorance by one parameter $\tau$-anchor function is obtained essentially by combining Anscombe-Aumann’s requirements of monotonicity and reversibility (MI and RI) with Hurwicz-Arrow’s assumptions. Assumptions CA and C are just regularity conditions.

In literature, Nehring and Puppe (NP) [24] obtained a characterization satisfying (7) by a different route. They consider a complete preference $\succeq_{NP}$ that satisfies two axioms Continuity ($C_{NP}$) and “Strong Independence” (SI) as follows. ($C_{NP}$): Suppose $A, B \in \mathcal{F}(\Omega)$ and $(A_n)_{n=1}^{\infty}$ is a sequence of sets converging to $A$. If $\forall n, A_n \succeq_{NP} B$ then $A \succeq_{NP} B$. If $\forall n, B \succeq_{NP} A_n$ then $B \succeq_{NP} B$. If a series converges to a set $A$ and all the members of the series is preferable to another set $B$ then $A$ is preferable to $B$. The concept of convergence of sets is understood to be the convergence in Hausdorff space. Assumption $C_{NP}$ conveys the same idea but is technically stronger than the continuity assumption (C). The second assumption in NP derivation is (Strong Independence - SI): For $A, B \in \mathcal{F}(\Omega), x \notin A \cup B, A \succeq_{NP} B \Rightarrow A \cup x \succeq_{NP} B \cup x$. Adding the same new element to sets $A$ and $B$ does not change the preference direction between the sets. The idea of this assumption is similar to the monotonicity assumption (MI). In addition to having different assumptions, the key difference between NP approach and our approach is that in the latter sequential acts are allowed while in the former only one stage acts are considered. As a consequence, in case of multiple ignorant variables $X, Y$ with domains $\Omega_X, \Omega_Y$, NP approach needs to collapse them into one variable with domain $\Omega_X \times \Omega_Y$. In NP approach the Reversibility under ignorance assumption (RI) is tautological because all three expressions in (3) have the same set of prizes.

2.4. Example

Let’s illustrate the evaluation of the investment instrument after the coffee harvest is known so that the return on the investment depends only on the result of political settlement (ignorant variable). The first step is to convert the values of investment return into normalized prizes in the unit interval. We assume that in the universe of investment options considered by the investor the return ranges from $-4\%$ to $16\%$. The conversion is reported in the following table by linear transformation.

Suppose that investor’s behavior under ignorance is summarized by three rules: (1) optimism in losses i.e., she equates a set of potential losses (negative returns) to the best return in that set; (2) pessimism in gains (positive
returns) i.e., she equates a set of potential gains to the lowest gain in the set; and (3) equates a set of returns that include both gain and loss to zero return. This behavior can be characterized by $\tau$-anchor utility function with $\tau = 0.20$ which corresponds to the return of zero percent (Table 2).

Let’s assume that an individual has a concave utility function under risk $u(x) = \sqrt{x}$ (she holds risk averse attitude). Now we can evaluate the certainty equivalence of the investment after knowing the coffee harvest (Fig. 2). If the coffee harvest is bumper then the investment reduces to an act under ignorance $\{0.05, 0.45\}$ (prizes calculated from the returns of $-3\%$ and $5\%$). The CE is $0.20$. If the harvest is normal then the corresponding act is $\{0.28, 0.30\}$ that has CE of $0.28$. Finally, if the harvest is poor then the corresponding act is $\{0, 0.60\}$ that has CE of $0.20$. The final answer to investor’s main question is not ready at this point and will be given in section 5.1.

3. Decision making under uncertainty comprising ignorance and probability

In the previous section, we considered a utility theory for complete ignorance. However, decisions in the real world are rarely made under complete ignorance alone but most likely involve more than one variables of different types of uncertainty. In this section we consider decisions whose prizes depend on an ignorant variable ($I$) and a probabilistic or roulette variable ($R$). Their domains are denoted by $\Omega_I$ and $\Omega_R$ respectively. The state space is $\Omega = \Omega_I \times \Omega_R$. Again we make an assumption that $R$ and $I$ are independent.

<table>
<thead>
<tr>
<th>Return (%)</th>
<th>-4.00</th>
<th>-3.00</th>
<th>0.00</th>
<th>1.20</th>
<th>1.50</th>
<th>2.00</th>
<th>5.00</th>
<th>8.00</th>
<th>16.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized</td>
<td>0.00</td>
<td>0.05</td>
<td>0.20</td>
<td>0.26</td>
<td>0.28</td>
<td>0.30</td>
<td>0.45</td>
<td>0.60</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2: Conversion from return to normalized prize.
in the sense that knowing the realization of one variable does not add any information about the other.

A mapping of the form \( f : \Omega_I \times \Omega_R \rightarrow O \) is called two-dimensional (2D) act because the prizes of \( f \) depend on both \( I \) and \( R \). In special case when the prizes of an act depend on \( I \) or \( R \) alone, it is called one-dimensional (1D). For example, if \( \forall x \in \Omega_I, \forall y, y' \in \Omega_R, f(x, y) = f(x, y') \) then \( f \) is an one-dimensional \( I \)-act. One-dimensional \( R \)-act is defined accordingly and is also referred to as lottery. Constant acts have zero dimension. The set of all (2D) acts is denoted by \( D \). The sets of 1D acts are denoted by \( D_I \) and \( D_R \) respectively. We have inclusion chain \( O \subset D_I, D_R \subset D \).

We consider a preference relation \( \succeq \) on \( D \) that is a weak order (complete, reflexive and transitive). The projections of \( \succeq \) on \( D_I \) and \( D_R \) are denoted by \( \succeq_I \) and \( \succeq_R \) respectively. That is, \( f, g \in D_I, f \succeq_I g \iff f \succeq g \). We assume that the preference order on prizes is exactly the arithmetic order.

**Assumption 5 (Preferences on \( R \)-acts, \( I \)-acts (1DR)).** \( \succeq_R \) is represented by the linear utility and \( \succeq_I \) by the \( \tau \)-anchor utility.

Under the linear utility assumption, an \( R \)-act can be written as a vector of probability-prize pairs \( (p_1 : x_1, \ldots, p_n : x_n) \) enclosed by a pair of round brackets. In particular, an implication of the linear utility assumption is that for triple \( x, y, z \in O \) such that \( x \leq y \leq z \), there is probability \( p \) such that \( y \sim (p : x, 1 - p : z) \). We continue with the convention that \( I \)-acts are denoted by finite bags of prizes.

Act \( f : \Omega_I \times \Omega_R \rightarrow O \) has two sequential forms: \( f_I : \Omega_I \rightarrow D_R \) and \( f_R : \Omega_R \rightarrow D_I \). The subscript indicates the variable which is supposed to be realized first. For example, suppose \( \Omega_R = \{r_1, r_2\} \) and \( \Omega_I = \{a_1, a_2\} \). Fig. 3 presents two sequential representations of act \( f \) with \( f(r_1, a_1) = x_1, f(r_1, a_2) = x_2, f(r_2, a_1) = y_1 \) and \( f(r_2, a_2) = y_2 \). The corresponding in-line notations are \( (p_1 : \{x_1, x_2\}, p_2 : \{y_1, y_2\}) \) (for the tree on the left) and \( \{(p_1 : x_1, p_2 : y_1), (p_1 : x_2, p_2 : y_2)\} \) where \( p_i \) is the probability of \( r_i \). On the left, the intermediate prizes are ignorant acts \( \{x_1, x_2\} \) and \( \{y_1, y_2\} \). The intermediate prizes on the right are lotteries \( (p_1 : x_1, p_2 : y_1) \) and \( (p_1 : x_2, p_2 : y_2) \).

To relate the projections \( \succeq_I \) and \( \succeq_R \) to their origin \( \succeq \) we again follow AA approach. In particular, we examine the consequences of adopting the monotonicity and the reversal order assumptions between an ignorant variable and a probabilistic one.
**Assumption 6 (Monotonicity (MIR)).** Suppose for \( z_i, z'_i \in D_I \) for \( 1 \leq i \leq n \) and \( u_j, u'_j \in D_R \) for \( 1 \leq j \leq m \)

\[
\forall i, z_i \succeq_I z'_i \Rightarrow (p_1: z_1, p_2: z_2, \ldots, p_n: z_n) \succeq (p_1: z'_1, p_2: z'_2, \ldots, p_n: z'_n),
\]

\[
\forall j, u_j \succeq_R u'_j \Rightarrow \{u_1, u_2, \ldots, u_m\} \succeq \{u'_1, u'_2, \ldots, u'_m\}.
\]

**Assumption 7 (Reversal of order (RIR)).** For any probability vector \((p_i)_{i=1}^n\)

\[
\{(p_i: x_i^1)_{i=1}^n, (p_i: x_i^2)_{i=1}^n, \ldots, (p_i: x_i^k)_{i=1}^n\} \sim (p_i: \{x_i^1, x_i^2, \ldots, x_i^k\})_{i=1}^n.
\]

Or in matrix form:

\[
\begin{pmatrix}
(p_1: x_1^1, p_2: x_2^1, \ldots, p_n: x_n^1) \\
(p_1: x_1^2, p_2: x_2^2, \ldots, p_n: x_n^2) \\
\vdots \\
(p_1: x_1^k, p_2: x_2^k, \ldots, p_n: x_n^k)
\end{pmatrix} \sim
\begin{pmatrix}
p_1: \{x_1^1, x_1^2, \ldots, x_1^k\} \\
p_2: \{x_2^1, x_2^2, \ldots, x_2^k\} \\
\vdots \\
p_n: \{x_n^1, x_n^2, \ldots, x_n^k\}
\end{pmatrix}.
\]

This assumption in effect equalizes two sequential forms \( f_I \) and \( f_R \) derived from a mapping \( f : \Omega_I \times \Omega_R \to \mathcal{O} \). In other words, it requires that \( f_I \) and \( f_R \) have the same certainty equivalent that is equal to the CE of \( f \). If (RIR) is satisfied then we say that \( I \) and \( R \) are fully reversible. We would expect that because \( R \) and \( I \) are independent, the order in which they are revealed does not matter. Given the intuitive appeal of (RIR), it is a disappointment to realize that (RIR) is not consistent with (MIR) and assumption 5 (1DR).
Theorem 4 (Non-reversibility). If preference relation $\succeq$ satisfies assumptions 5 (1DR) and 6 (MIR) then it does not satisfy assumption 7 (RIR).

In Fig. 3, the violation of the reversibility means that you can find values $x_i, y_j$ for $i, j = 1, 2$ and probabilities for $r_1, r_2$ such that the certainty equivalence of the (sequential) act on the left is different from that on the right.

In theorem 4, non-reversibility is obtained under assumption that preference under ignorance is governed by $\tau$-anchor utility. Actually, we can still draw that conclusion with a weaker assumption: that there exists a set of prizes such that the CE under ignorance is strictly less (more) than the CE under uniform distribution. Ellsberg’s paradox [9] is an example of this condition.

Let’s consider the example in Fig. 3. Variables $I, R$ have domains $\Omega_I = \{a_1, a_2\}$ and $\Omega_R = \{r_1, r_2\}$, $I$ is ignorant and $R$ is a fair coin with the uniform distribution $Pr(r_1) = Pr(r_2) = 0.5$. $f$ is an act $f(a_1, r_1) = 0, f(a_1, r_2) = 1, f(a_2, r_1) = 1$ and $f(a_2, r_2) = 0$ (in Fig. 3 $x_1 = y_2 = 0$ and $x_2 = y_1 = 1$). Suppose that the utility function under risk is $u(x) = \sqrt{x}$. On the one hand, the tree on the right side (Fig. 3) represents situation when $I$ is revealed before the coin is flipped. Given $I = a_1$, $f$ becomes a lottery $f_{a_1}(r_1) = 0$ and $f_{a_1}(r_2) = 1$ and its expected utility is 0.5. The CE of $f_{a_1}$ is 0.25. Conditional act $f_{a_2}$ given $I = a_2$ again has expected utility 0.5 and CE of 0.25. Thus, viewed as function of $I$ alone, $f$ is a constant act of 0.5 (utile), hence, the CE of $f$ is 0.25 (dollar). On the other hand, the tree on the left side of Fig. 3 corresponds to ordering $R \triangleright I$. After the coin is flipped and before $I$ is revealed, given $R = r_1$ $f$ becomes an act under ignorance with a set of monetary rewards $\{0, 1\}$. Suppose by the assumption, the CE of $\{0, 1\}$ under ignorance be a value $c < 0.25$. If $R = r_2$, $f$ becomes act under ignorance with the same set of monetary rewards $\{0, 1\}$ which also has CE equal to $c$. So, as a function of $R$ alone $f$ is a constant act with CE equal to $c$ which is strictly less than 0.25. Thus, the reversibility is violated.

Theorem 4 basically says that assumptions 5, 6 and 7 cannot be satisfied at the same time. If we insist on the representation of one-dimensional probabilistic and ignorant acts by linear utility and $\tau$-anchor utility respectively then the monotonicity and the reversibility assumptions contradict each other. Consequently, we have to choose one at the expense of the other. In fact, we choose the monotonicity and reject the reversibility in the following proposal.
4. Order-of-variables factor

4.1. Ordering information as a meta-variable

The violation of the reversibility assumption (RIR) implies that two sequential versions of an act, \( f_I \) and \( f_R \), are preferentially different even if \( R \) and \( I \) are independent. To account for that difference, we note that the sequential acts differ in a fundamental aspect which is the order in which variable values are revealed. In Fig. 3, the tree on the left makes an assumption that the value of the roulette spin is known before the value of the ignorant variable is known. The tree on the right makes an opposite statement. We argue that this difference in order of variables is responsible for the preferential difference between \( f_I \) and \( f_R \). To capture this relevant information, we introduce a meta variable \( T \) whose domain \( \Omega_T \) is the set of possible variable orderings. For two variables \( I, R \) there are two possible orderings \( \{ I \prec R, R \prec I \} \) where \( X \prec Y \) means “\( X \) precedes \( Y \)”.

With the inclusion of the meta variable, we can associate an act with a three-stage tree in which \( T \) is represented in the first stage. Such a tree is called the complete tree for the act. For example, the tree in Fig. 4 is obtained by joining two trees in Fig. 3. Given the complete tree, the ordinary trees on variables \( I \) and \( R \) are obtained by pruning the impossible branch.

Naturally, we have to deal with question what to do if we don’t know for sure the order by which variables are realized. The solution is, unsurprisingly, to treat \( T \) as a uncertain variable. \( T \) is allowed to be either probabilistic or ignorant. For example, when the information about the order of variables is missing, \( T \) is treated as an ignorant variable.

An important implication of explicitly modeling the order of variables is that the independence assumption for \( I \) and \( R \) can be dropped. Given a variable ordering, the later variable is allowed to be dependent on the previous one. The (asymmetric) dependency is realized via conditionalization mechanism. Given \( I \prec R \), knowing that \( I = a_i \) leads to updating the information about \( R \). This new information is described by conditional probability distribution \( Pr(R|I = a_i) \). Given \( R \prec I \), knowing \( R = r_j \) can change the domain of \( I \) from \( \Omega_I \) to \( \Omega_{I_j} \).

4.2. Formalization

The formal framework includes two ordinary variables \( I, R \) with the domains \( \Omega_I, \Omega_R \) and a meta variable \( T \) whose domain is the set of orderings \( \Omega_T = \{ R \prec I, I \prec R \} \).
Given the unit interval presenting prizes $\mathcal{O}$, $\mathcal{F}(\mathcal{O})$ is the set of finite subset of $\mathcal{O}$ and $\mathcal{L}(\mathcal{O})$ is the set of lotteries or finite support probability distributions on $\mathcal{O}$. An act given ordering $R \triangleleft I$ is a distribution on $\mathcal{F}(\mathcal{O})$, $(p_1:A_1, p_2:A_2, \ldots)$, where $A_i \in \mathcal{F}(\mathcal{O})$. This type of acts is referred to as $R$-$I$ acts to differentiate them from acts under a reverse ordering $I \triangleleft R$. Given ordering $I \triangleleft R$, an $I$-$R$ act is a collection of lotteries, $\{L_a | a \in \Omega_I\}$, where $L_a \in \mathcal{L}(\mathcal{O})$. The constant acts are both $I$-$R$ and $R$-$I$ acts. The preference relation on $R$-$I$ acts is denoted by $\succeq_{ri}$ and the preference relation on $I$-$R$ acts is denoted by $\succeq_{ir}$. They are the conditional preference relations associated with orderings.

We assume two primitive preference relations: $\succeq_I$ on the set of finite sets of prizes $\mathcal{F}(\mathcal{O})$ and $\succeq_R$ on the set of lotteries $\mathcal{L}(\mathcal{O})$. We adopt assumptions 1 (section 2) and 5 and 6 (section 3). Assumption 1 requires that preference on the constant acts (prizes) is exactly the arithmetic order. Assumption 5 requires $\succeq_I$ and $\succeq_R$ are governed by the $\tau$-anchor utility and the linear utility respectively. Assumption 6 is about the monotonicity of $\succeq_{ri}$ and $\succeq_{ir}$ with respect to $\succeq_I$ and $\succeq_R$ respectively. Finally, all preference relations $\succeq_{ri}, \succeq_{ir}, \succeq_I$ and $\succeq_R$ are weak orders. Their intersection includes the segment
on the constant acts $\mathcal{O}$.

Note that the uncertainty structure of the decision problem is different under different variable orderings. Under $R \triangleleft I$, the uncertainty is described by a probability measure on $\Omega_R$ and ignorance measure on $\Omega_I$ while under $I \triangleleft R$, it is an ignorance measure on $\Omega_I$ and a collection of conditional probability measures $\{Pr_a(\Omega_R)\}_{a \in \Omega_I}$.

Mappings of the form $\Omega_I \times \Omega_R \rightarrow \mathcal{O}$ are called incomplete acts because they are defined without reference to any ordering. An incomplete act $f$ has two complete versions $f_{ri}$ and $f_{ir}$ one for each possible ordering. The complete versions of an incomplete act are not necessarily indifferent even if $I$ and $R$ are independent. In general, there is no rational ground to require indifference between complete versions of an act because the uncertainty under different orderings are not comparable. Although an $I$-$R$ act and an $R$-$I$ act cannot be not directly compared by $\succeq_{ir}$ or $\succeq_{ri}$, thanks to the dual nature of constant acts, they are indirectly comparable via their CE.

In this framework, an incomplete act is actually a set of complete acts and only complete acts can be preferentially evaluated. This framework rejects the “one-stage” view of an act in which variables $I$ and $R$ are collapsed into a single variable. Implicitly, the idea of combining ignorance with probability to create a uncertainty measure on the product space $\Omega_I \times \Omega_R$ is rejected.

When an incomplete act cannot be made into a completed one due to the lack of information about the order of variables, the situation is modeled explicitly by an ignorance measure over the set of complete acts. Formally, if $f, g$ are incomplete acts, preference relation $\succeq_{\emptyset}$ is defined for ignorance over $\Omega_T$ and $\succeq_p$ is defined for a probability measure over $\Omega_T$.

\[
\begin{align*}
f \succeq_{\emptyset} g & \iff \{CE(f_{ri}), CE(f_{ir})\} \succeq_I \{CE(g_{ri}), CE(g_{ir})\}, \quad (13) \\
f \succeq_p g & \iff (p;CE(f_{ri}), (1-p);CE(f_{ir})) \succeq_R (p;CE(g_{ri}) , (1-p);CE(g_{ir})) \quad (14)
\end{align*}
\]

where $p$ is the probability of $R \triangleleft I$.

The definitions of $\succeq_{\emptyset}$ and $\succeq_p$ use the CE of complete acts. The existence of the CE can be shown recursively. For $A \in \mathcal{F}(\mathcal{O})$ and $L \in \mathcal{L}(\mathcal{O})$ their CE exist because $\succeq_I$ and $\succeq_R$ are represented by the $\tau$-anchor utility and the linear utility (Assumption 5). Moreover, the CE are unique because of assumption 1. For an $R$-$I$ act $(p_1:A_1, p_2:A_2, \ldots)$ where $A_i \in \mathcal{F}(\mathcal{O})$, suppose that the CE of $A_i$ are $c_i \in \mathcal{O}$. By the monotonicity $(p_1:A_1, p_2:A_2, \ldots) \sim_{ri} (p_1 : c_1, p_2 : c_2, \ldots)$. Thus, the CE of the $R$-$I$ act equals the CE of lottery $(p_1:c_1, p_2:c_2, \ldots)$. Similarly, the CE of an $I$-$R$ act is reduces to the CE of a
subset of prizes. Thus, the definitions of $\succeq_\emptyset$ and $\succeq_p$ by eqs. (13) and (14) are well defined. Because $\succeq_I$ and $\succeq_R$ are a weak order, so are $\succeq_\emptyset$ and $\succeq_p$.

5. Application and Example

In this section, two traditional models of decision under uncertainty, namely, the multiple-prior model and the decision making with Dempster-Shafer belief function are analyzed from the perspective of decision with ignorance and probability. But first we illustrate our method with several numerical examples.

5.1. Investment under uncertainty comprising ignorance and probability

Let’s continue the calculation given in Section 2.4 for the investment whose return depends on Brazil’s coffee crop and Afghanistan’s political settlement in 2015. Recall that investor’s utility function under risk is $u(x) = \sqrt{x}$ and her characteristic value under ignorance is $\tau = 0.2$. Under the variable ordering $C \triangleright S$, after the coffee crop is known the certainty equivalents for ignorant acts given $C = b$, $C = n$ and $C = p$ are 0.20, 0.28 and 0.20 respectively. So before the crop is known the investment is a lottery $(0.46 : 0.2, 0.28, 0.34 : 0.2)$. The calculation of the CE of this lottery is given in Table 3. The expected utility 0.4636, hence, the CE is 0.2149 which is translated into return of 0.3%.

<table>
<thead>
<tr>
<th>$C$</th>
<th>Prob</th>
<th>CE</th>
<th>C</th>
<th>Util</th>
<th>Exp util</th>
<th>CE</th>
<th>Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b$</td>
<td>0.46</td>
<td>0.20</td>
<td>0.4472</td>
<td>0.4636</td>
<td>0.2149</td>
<td>0.30%</td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>0.20</td>
<td>0.28</td>
<td>0.5292</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>0.34</td>
<td>0.20</td>
<td>0.4472</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Evaluation of investment in case $C \triangleright S$.

In the case $S \prec C$ the calculation is given in Table 4. Given a political settlement, the investment becomes a lottery $(0.46 : 0.05, 0.2 : 0.3, 0.34 : 0.6)$. Without a political settlement, it is $(0.46 : 0.45, 0.2 : 0.28, 0.34 : 0)$. Their CE are 0.2264 and 0.1717 respectively. So before knowing the value of $S$ the investment is an act under ignorance {0.2264, 0.1717} that has CE equal $\tau = 0.2$ (translated to the return of zero percent).

Thus, the answer to investor’s main question is that compared with putting money in the bank (earning 1.2%) the investment is an inferior choice. For a what-if question what should the investor do if the bank pays only 0.2%,
the answer depends on additional information about the order of realization of variables. If the coffee harvest is known before the political settlement then the investor should choose the investment. If the coffee harvest is known after the political settlement then she should put money in the bank. If she has no information about the order of variables then she also should put money in the bank.

5.2. Scenarios based on Ellsberg’s urn

Since its publication in 1961, Ellsberg’s famous paradox [9] has served as the standard test of reasonability for many proposals for decision making under uncertainty. We follow the practice here. The original Ellsberg’s experiment is set as follows. A urn of 90 balls of three colors: red (r), white (w) and yellow (y). The proportion of red is $\frac{1}{3}$. The proportions of white and yellow are unknown. $\bar{r}$ (or $\sim r$) denotes “not red” i.e., white or yellow. The same way $\bar{w}, \bar{y}$ are defined. A ball is drawn from the urn. A bet on proposition $\gamma$ where $\gamma$ is a propositional formula constructed from literals $r, w, y$ or their negations, pays $1 if $\gamma$ holds and nothing otherwise. The finding that contradicts the prescription of expected utility theory is that a typical decision maker strictly prefers a bet on $r$ to a bet on $w$ (or $y$) but at the same time strictly prefers a bet on $w \lor y$ to a bet on $r \lor y$.

Raiffa [26] claimed that randomization (flipping a coin) can resolve Ellsberg’s paradox. Refuting Raiffa’s claim, Eichberger et al [8] show that the apparent inconsistency of the smooth model for ambiguity [19] can be explained by considering whether randomization is done before or after learning about the color of the ball. In the following we use this augmented framework of ball color and a fair coin to illustrate the key points of our approach. We keep the same assumption made earlier that the individual’s utility function is $u(x) = \sqrt{x}$ and the characteristic value under ignorance $\tau = 0.20$. Under

<table>
<thead>
<tr>
<th>$S$</th>
<th>$C$</th>
<th>Prob</th>
<th>Prize</th>
<th>Util</th>
<th>$S$</th>
<th>$C$</th>
<th>Prob</th>
<th>Prize</th>
<th>Util</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>b</td>
<td>0.46</td>
<td>0.05</td>
<td>0.2236</td>
<td>$\sim s$</td>
<td>b</td>
<td>0.46</td>
<td>0.45</td>
<td>0.6708</td>
</tr>
<tr>
<td>s</td>
<td>n</td>
<td>0.20</td>
<td>0.30</td>
<td>0.5477</td>
<td>$\sim s$</td>
<td>n</td>
<td>0.20</td>
<td>0.28</td>
<td>0.5292</td>
</tr>
<tr>
<td>s</td>
<td>p</td>
<td>0.34</td>
<td>0.60</td>
<td>0.7746</td>
<td>$\sim s$</td>
<td>p</td>
<td>0.34</td>
<td>0</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 4: Evaluation of investment in case $S \triangleleft C$.  

---

3We abuse the notation to identify a proposition with the bet on that proposition.
the uniform probability distribution and with the utility function \( u(x) \), the lottery (0.5:0, 0.5:1) has the CE of 0.25.

We consider several scenarios. In the first scenario, we make an additional assumption that after the ball is drawn, the bettor will be told whether the ball’s color is red or not. In our framework, the given information about the urn is modeled by two variables \( X, Y \). The domain of \( X \) has two propositions \{r, \bar{r}\}. The conditional domain of \( Y \) given \( X = r \) has only one state denoted by \( \top \) (on this domain ignorance and certainty are the same). \( Y \) conditional on \( X = \bar{r} \) is an ignorant variable of two states \{w, y\}. (Fig. 5). In our modeling, since the ignorance about proportions of white and yellow balls is conditional on \( \bar{r} \), the logic dictates that the ordering must be \( X \prec Y \).

For a bet on \( r \), the prizes are \( x_1 = 1, x_2 = x_3 = 0 \). If the drawn ball is red, the prize is 1. If the color is not red then the prize is 0. The expected utility is \( \frac{1}{3}\sqrt{1} + \frac{2}{3}\sqrt{0} = \frac{1}{3} \). The CE is 0.1111.

The bet on \( w \) has the following prizes \( x_1 = x_3 = 0 \) and \( x_2 = 1 \). The prize if the ball is red is 1. If the color is not red, the individual has set of prizes \{0, 1\} under ignorance. The CE of that act is the characteristic value \( \tau = 0.20 \). The expected utility is 0.2981 and the CE of entire act is 0.0889.

The bet on \((w \lor y)\) has the following prizes \( x_1 = 0, x_2 = x_3 = 1 \). The expected utility is 0.6667 and the CE is 0.4444.

The bet on \((r \lor y)\) has the following prizes \( x_1 = 1, x_3 = 1 \) and \( x_2 = 0 \). If the ball is not red the individual has set of prizes \{0, 1\} under ignorance. The CE of this ignorant act is 0.20. The expected utility is 0.6315 and the CE of entire bet is 0.3988. Thus, \( r \succ w \) and \((w \lor y) \succ (r \lor y)\). At basic level, our theory prescribes that the decision maker strictly prefers a bet on
In the second scenario, we consider an individual with a fair coin who makes bets with Ellsberg’s urn. In the first case when the coin is flipped before the ball is drawn and the individual makes bet on $w$ or $y$ based on the outcome of the flipping (if Head then bet on $w$, if Tail bet on $y$). In Fig. 6 (a), $x_1 = x_{31} = x_{22} = 0$ and $x_{21} = x_{32} = 1$. The certainty equivalents in both branches (labeled with $H$ and $T$) are the same (0.0889). So the CE of the entire bet with randomization is also 0.0889. This randomization does not help to remove the ambiguity and the bet on $w$ ($y$) is still inferior to the bet on $r$. Raiffa’s claim is refuted if the coin is flipped before drawing ball.

The third scenario is presented in Fig. 6 (b). After learning that the ball is not a red one ($\bar{r}$), she flips the coin and bet on the outcome of the randomization, namely, she places a bet on $w$ if Head and on $y$ if Tail. Strictly speaking, the individual has more information than a person who has to bet before the ball is drawn. However, even in this case, randomization does not help. The person is still facing ignorance about $w$, $y$. The subtree associated with $\sim r$ has CE of 0.20. The entire bet has the CE of 0.0889.
Raiffa’s claim is refuted again in the case the coin is flipped after the ball is drawn and the bettor is told if the ball is not red. In our modeling, there are two probabilistic variables involved: one has domain \{H, T\} the other has domain \{r, \bar{r}\}. The indifference between versions (a) and (b) is an example that probabilistic variables are fully reversible.

5.3. Analysis of multiple-prior model

We use Ellsberg’s example to illustrate how our model can clarify a puzzle generated by multiple priors (MP) model (aka maximin expected utility (MMEU) [11]). In MMEU, the uncertainty is assumed to be a set of probability functions \(\Delta(S)\) on space \(S\). For an individual with vNM utility function under risk \(u\), the expected utility of act \(f\) given probability distribution \(\mu\) is \(e_{\mu,f} = \int u(f(x))d\mu\) from which the conditional CE of \(f\) given \(\mu\) is calculated. The set of probability functions \(\Delta(S)\) induces the set of conditional CEs \(C_{f,\Delta}\). The (unconditional) CE of the act is calculated from the set of conditional CEs. If one holds a probabilistic belief \(p\) on \(\Delta(S)\), the solution is again found by the expected utility. If such a second order probability is not known, MMEU compares acts by their the minimal conditional CE i.e., \(f \succeq_{\text{MMEU}} g\) if \(\min(C_{f,\Delta}) \geq \min(C_{g,\Delta})\).

The MMEU model without the second order probability is a special case of our model with an ignorant variable \(I\) and a chance variable \(R\). Suppose \(\Delta(S) = \{p_{\alpha}\}_{\alpha \in D}\), the domain of \(I\) is the set of identifiers \(D\). The domain of \(R\) is the state space \(S\). The order of variables is \(I \prec R\). The dependency of \(R\) on \(I\) and the conditional probability \(Pr(R|I = \alpha) = p_{\alpha}\). So MMEU is obtained when \(\tau = 0\). This value of \(\tau\) indicates ignorance aversion to the extreme which MMEU is often criticized for. In our model we can relax that extreme attitude. Namely, we can compare acts by their CE under \(\tau > 0\).

Not only our model allows non-extreme uncertainty attitudes, it clarifies potential modeling missteps. In MP model, Ellsberg’s urn is represented by a set of probability distributions \(\Delta = \{(\frac{1}{3}, \frac{2}{3} \alpha, \frac{2}{3}(1 - \alpha))\}_{0 \leq \alpha \leq 1}\) where the components in probability vectors are the proportions of \(r, w\) and \(y\) respectively. Fig. 7 (a) is the tree version of a bet on \(w\) according to MP interpretation. Given a value \(\alpha\), the expected utility of betting on \(w\) is \(\frac{2}{3}\alpha\) so the CE conditional on \(\alpha\) is \((\frac{2}{3})^2 \alpha^2\). Thus the set of conditional certainty equivalents is an interval \(C_{f,\Delta} = [0, \frac{4}{9}]\). Because this interval includes \(\tau = 0.2\), the CE of this interval under ignorance is \(CE([0, 0.4444]) = 0.2\). This CE is
greater than the CE of betting on \( r \) (0.1111) and it implies a counter-intuitive prescription \( w \succ r \).

Of course, one can make the bet on \( w \) less preferable than the bet on \( r \) by assuming more ignorance aversion i.e., choosing \( \tau < 0.1111 \), in particular, setting \( \tau = 0 \) as in MMEU. But extreme aversion to uncertainty is not realistic and has many undesirable implications.

However, our theory suggests a more plausible explanation for the failure to predict preference for \( r \) over \( w \) with a reasonable degree of ignorance aversion. We argue that the modeling of the sequence between the ignorant and the chance variables in version (a) Fig. 7 is incorrect.

Let’s compare versions (a) and (b) in Fig. 7. The ignorant variable in (a) is realized before the chance variable (i.e., \( \alpha \) must be known before the subtree with \( r, w, y \) can be evaluated) while in (b) the ignorant variable is resolved after the chance variable (the ball is red or not). In version (b), given \( \alpha \) the expected utility of subtree is \( \alpha \). The CE is \( \alpha^2 \). So the set of conditional certainty equivalents \( \{\alpha^2|0 \leq \alpha \leq 1\} \) or the interval \( [0,1] \).

Because this interval includes \( \tau = 0.2 \), the CE under ignorance given \( \sim r \) is 0.2. The expected utility is \( \frac{2}{3} \sqrt{0.2} = 0.2981 \) and the CE of entire act is 0.0889 which by the way is the same as the CE in Fig. 5. So the model in version (b) correctly predicts preference \( r \succ w \).
Because ignorant and chance variables are not reversible, the models (a) and (b) have different certainty equivalents. Model (a) is incorrect because the ignorance about the proportions of white and yellow balls is substituted by the ignorance on the index of probability function, reversing the order between ignorant and chance variables. A lesson from this example is that correct modeling requires careful isolation of the ignorance component of uncertainty to preserve the natural order of variables.

5.4. Decision making with Dempster-Shafer belief function

Belief function theory, which admits different interpretations, offers an instructive case study of the influence of uncertainty representation on decision making. Dempster [7] considered a uncertainty structure obtained by a multiple-valued mapping $m$ from a space $\Theta$ (the proto space) which is equipped with a probability measure $p$, to the power set of $\Omega$, $m: \Theta \rightarrow 2^\Omega$. Consider acts that map $\Omega$ to the set of prizes $\mathcal{O}$. The situation described by this belief function has a natural representation in our framework. We have a roulette variable $X$ (with domain $\Theta$) and an ignorant variable $Y$ in order $X \prec Y$. The the ignorant variable is dependent on the roulette variable. The conditional domain of $Y$ given $X = \theta_i$ is the subset $m(\theta_i)$.

Consider an example (Fig. 8). Suppose $\Theta = \{\theta_1, \theta_2\}$, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. The probability distribution is $Pr(\theta_1) = 0.4$, $Pr(\theta_2) = 0.6$. The mapping $m(\theta_1) = \{\omega_1, \omega_2\}$ and $m(\theta_1) = \{\omega_2, \omega_3, \omega_4\}$. Act $f$ on $\Omega$ has $f(\omega_1) = 0$, $f(\omega_2) = 0.15$, $f(\omega_3) = 0.35$, $f(\omega_4) = 1$. 

Figure 8: Dempster’s view of belief function.

Because ignorant and chance variables are not reversible, the models (a) and (b) have different certainty equivalents. Model (a) is incorrect because the ignorance about the proportions of white and yellow balls is substituted by the ignorance on the index of probability function, reversing the order between ignorant and chance variables. A lesson from this example is that correct modeling requires careful isolation of the ignorance component of uncertainty to preserve the natural order of variables.

5.4. Decision making with Dempster-Shafer belief function

Belief function theory, which admits different interpretations, offers an instructive case study of the influence of uncertainty representation on decision making. Dempster [7] considered a uncertainty structure obtained by a multiple-valued mapping $m$ from a space $\Theta$ (the proto space) which is equipped with a probability measure $p$, to the power set of $\Omega$, $m: \Theta \rightarrow 2^\Omega$. Consider acts that map $\Omega$ to the set of prizes $\mathcal{O}$. The situation described by this belief function has a natural representation in our framework. We have a roulette variable $X$ (with domain $\Theta$) and an ignorant variable $Y$ in order $X \prec Y$. The the ignorant variable is dependent on the roulette variable. The conditional domain of $Y$ given $X = \theta_i$ is the subset $m(\theta_i)$.

Consider an example (Fig. 8). Suppose $\Theta = \{\theta_1, \theta_2\}$, $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. The probability distribution is $Pr(\theta_1) = 0.4$, $Pr(\theta_2) = 0.6$. The mapping $m(\theta_1) = \{\omega_1, \omega_2\}$ and $m(\theta_1) = \{\omega_2, \omega_3, \omega_4\}$. Act $f$ on $\Omega$ has $f(\omega_1) = 0$, $f(\omega_2) = 0.15$, $f(\omega_3) = 0.35$, $f(\omega_4) = 1$. 

Figure 8: Dempster’s view of belief function.
\( f(\omega_2) = 0.15, \ f(\omega_3) = 0.35 \) and \( f(\omega_4) = 1 \). The decision maker has vNM utility function \( u(x) = \sqrt{x} \). The characteristic value under ignorance \( \tau = 0.2 \).

In our model, the CE of set of prizes \( \{0, 0.15\} \) under ignorance is 0.15 because all the prizes are below the \( \tau = 0.2 \); the CE of set of prizes \( \{0.15, 0.35, 1\} \) under ignorance is 0.2 because \( \tau = 0.2 \) is in between the minimum and maximum prizes in the set. The expected utility of the lottery \( (0.4:0.15, 0.6:0.2) \) is 0.4232 and the CE is 0.1791.

A belief function also admits an interpretation by a set of probability measures. A multiple-valued mapping generates a set of probability measures on \( \Omega \). The belief measure on \( \Omega \) is defined from \( m \) as follows:

\[
Bel(A) = \sum_{B \subseteq A} m(B) \tag{15}
\]

It is well known (see for example [14]) that Bel induces a set of probability measures on \( \Omega \), \( P = \{\mu | \forall B \subseteq \Omega, \mu(B) \geq Bel(B)\} \) i.e., \( P \) is the core of a transferable utility cooperative game Bel.

The representation of a belief function by \( P \) naturally suggests modeling with an ignorant variable \( X \) and a probabilistic variable \( Y \) in the order \( X \sim Y \). \( X \) can be thought of as the identifier of the probability measure in \( P \). \( Y \) conditioned on \( X = i \) is described by probability measure \( \mu_i \in P \). For an individual with vNM utility function \( u \) and characteristic value \( \tau \) the CE of act \( f \) is

\[
CE(f) = CE(\{u^{-1}(u_f), u^{-1}(\bar{u}_f)\}) \tag{16}
\]

where \( CE \) is given in Eq. 7, \( u_f (\bar{u}_f) \) is the minimal (maximal) expected utility of \( f \) over \( P \), \( u^{-1} \) is inverse function of \( u \).

For the example in Fig. 8, we have \( u_f = 0.2323 \) and \( u^{-1}(u_f) = 0.054 \). \( \bar{u}_f = 0.7549 \) and \( u^{-1}(\bar{u}_f) = 0.5699 \). So the CE of \( f \) is 0.2.

Because measures Bel and \( m \) are recoverable from the core set of probability measures \( P \), it is often claimed that different representations carry the same information. The fact that decision making on the “equivalent” interpretations lead to different results remains unsettling. Our approach can explain the puzzle by the non-reversibility between ignorance and probability. The difference is due to the fact that in Dempster’s belief function, the ignorance follows the probability while in the set of probability measures the order has been reversed.
6. Related literature and discussion

Puzzling phenomena, similar to what is described in theorem 4, have been known for long time under various terms such as “sequential incoherence” [27], [18] or “dynamic inconsistency” [22]. Such violations of the expected utility, in particular Ellsberg’s and Allais’ paradoxes, have spurred the efforts to develop non-expected theories of decision under uncertainty. Discussing the failure of stochastic dominance under substitution of indifferent alternatives (aka “sequential incoherence”), Seidenfeld [27] shows that it is a result of the failure of Independence (Monotonicity) axiom. To account for such violations, one group of proposals in the literature calls for the rejection of the Independence postulate. Another group of proposal calls for the rejection of the Ordering postulate i.e., to abandon the requirement that the preference is a weak order.

Our proposal is the third alternative, it keeps both the Monotonicity and the Ordering but rejects the Reversibility. In our theory, the reversibility holds between variables of the same type (either probabilistic or ignorant) but does not hold between variables of different uncertainty types. We don’t subscribe to the terms “incoherence” or “inconsistency” because we don’t believe that the reversibility is a rational property that must be taken for granted, at least, not under any circumstance. The reversibility is a tautological consequence of adopting the traditional one-stage view of acts as mappings from a state space to a reward space. Only in the context of Anscombe-Aumann’s multiple-variable setting, its rationale can be discussed. AA have shown that assuming the monotonicity and the reversibility implies the probabilistic representation of uncertainty. Contra-positively, if the uncertainty is not probabilistically representable then the reversibility is violated.

The τ-anchor utility theory under ignorance (Section 2) is developed on theoretical arguments but it is compatible with the findings from research on behavioral, neurobiological aspects of decision in the condition of ignorance.

One of few works dedicated to find experimental evidence about human decision making under ignorance is done by Hogarth and Kunreuther [16]. They argued that many practical situations where individuals have to make decisions, lack the basic features of a gamble, namely, the consequences and probabilities. For example, an individual decides whether to buy a warranty for an electronic device without knowing the probability of its breakdown and the repair cost. They designed a series of experiments on human subjects to understand the difference between the behavior under probability and that
when no probability or cost are known. They found clear evidence that the behaviors when the subjects have and do not have probability are different. They found that under ignorance the subjects use two types of strategies to arrive at a choice. One strategy is to employ a “principle that resolved the choice conflict and was insensitive to the particular features of different options”. The authors expressed some degree of surprise at the finding and wrote “It is perhaps ironic that, under ignorance, when people should probably think harder when making decisions, they do not. In fact, they may be swayed by the availability of simple arguments that serve to resolve the conflicts of choice.” We disagree with this comment. We think that the observed behavior is perfectly rational. Under ignorance, there is no reliable information upon which one can try to think harder. Thinking harder in this case often means filling the void left by lack of hard evidence with personal analytic/judgmental assumptions which may have no correspondence to the reality. The casual practice of acting upon those assumptions as if they were facts is misleading and sometime dangerous. As it was argued in section 2.3, the \( \tau \)-anchor utility fits the description of this type of strategy. For example, \( \tau \) can be interpreted as the “peace of mind” level for a subject.

Pushkarskaya et al [25] examined the neurological evidence of decision under ignorance using fMRI scan. They examine two types of missing information: ambiguity (vague probabilities) and sample space ignorance (SSI). They found that different types of missing information activate distinct neural substrates in the brain. The data partially reject the traditional reductive view held by neuroscientists, according to which, individuals reduce SSI (ignorance) to uncertainty/ambiguity and then to risk. It is done by forming subjective belief about the partition of a sample space and after that, forming a subjective probability measure. They found that the reductive view is only compatible with ambiguity averse individuals but not with ambiguity-tolerant individuals. The key recommendation they make is that theories of decision making under uncertainty should include a parameter on individual tolerance for missing information. We note that the characteristic value, \( \tau \), can be used to express individual’s extent of tolerance for ignorance (Corollary 1).

The axiomatic approach to decision under ignorance has been discussed, among others, by Maskin [23], Nehring & Puppe [24], Landes [21]. A literature survey on the topic is given in [4]. Cohen and Jaffray (CJ) [5] described a system of axioms for rational behavior under complete ignorance. CJ theory assumes a state space \( \Omega \) and acts are mappings from the state space to the
set of prizes. The basic object in CJ theory is the strict preference relation \( P \) i.e. \( fPg \) is the notation for “\( f \) is strictly preferred to \( g \)”.

\( P \) is assumed to be asymmetric and transitive. From relation \( P \) two relations \( R \) and \( I \) are defined. \( fRg \) means not \( fPg \) (it is not the case that \( f \) is strictly preferred to \( g \)). \( fIg \) means not \( (fPg \lor gPf) \). Intuitively one can view \( R \) as (non strict) preference and \( I \) as indifference but the difference between relation \( R \) and \( \geq \) in HA theory is that \( R \) is not transitive and neither is \( I \) even as \( P \) is. Giving up the transitivity requirement for non-strict preference, CJ were able to add a weak dominance axiom. A relation \( D \) is defined between two acts \( f, g \):

\[
fDg \iff \forall w \in \Omega, f(w) \geq g(w) \text{ and } \exists w_0 \in \Omega, f(w_0) > g(w_0). \quad (17)
\]

That is, for all states \( f \) is as good or better than \( g \) and there is at least in one state \( f \) is strictly better than \( g \). The weak dominance axiom stipulates that weak dominance implies strict preference i.e., \( fDg \Rightarrow fPg \). In addition to that, CJ also introduced an axiom “increase” (Axiom 6) which is less intuitive. They define a class of “rational decision criteria” consisting of those that satisfy their system of axioms. The central result is that CJ rational decision criteria in a “first-order approximation, depend on the sole comparison between the extremal values of acts, the taking into account of weak dominance which is required of criteria, or of other interactions between acts bringing in events, which remains a possibility, can only have a second-order influence on choices” [5]. An example of CJ rational criterion is

\[
fPg \iff \begin{cases} (m_f + M_f > m_g + M_g) \text{ or } (m_f + M_f = m_g + M_g, M_f - m_f < M_g - m_g) \text{ or } (m_f + M_f = m_g + M_g, M_f - m_f = M_g - m_g, \min_{\omega'f} > \min_{\omega'g}) \end{cases} \quad (18)
\]

where \( m_f = \min_{\Omega} f, M_f = \max_{\Omega} f \) and \( \Omega' = \{w \in \Omega | f(w) \neq g(w)\} \). That is a type of lexicographic criterion. The first condition used to compare two acts is the sum of their min and max elements. If that condition does not resolve in a strict preference, the second condition used is the difference between the max and min elements. If the second condition does not resolve the comparison then the minimal elements among the states where \( f \) and \( g \) are different are used.

In [6], Congar and Maniquet (CM) investigated an axiomatic system for decision under ignorance. In this setting, an act is a vector of outcomes which are von Neumann-Morgenstern (vNM) utilities. Implicitly, it is assumed that the ignorant variable precedes the risk variable. CM consider
five axioms: quasi-transitivity (transitivity required only for the strict preference relation), Savage’s independence, Duplication (split/merge of states with the same outcomes do not change preference), Strong dominance (dominance holds for all permutations of outcomes) and Scale invariance (linear transformation of the utilities does not affect the preference). Only three decision criteria satisfy all those requirements. The protective criterion, ignoring the common outcomes of acts $u$ and $v$, compares the minimal elements among remaining outcomes. This criterion reflects extreme pessimism. The hazardous criterion is the dual version of protective criterion. Instead of comparing the minimal elements of $u$ and $v$ excluding the common part, it compares the maximal elements. Finally, neutral criterion is the conjunction of both protective and hazardous criteria.

All three criteria compare acts by restricting the attention to the states in which the outcomes are different (due to Savage’s axiom). This feature differentiates CM from Hurwicz-Arrow’s decision criterion (theorem 1). Consider acts $u$ and $v$ which have the same minimal and maximal outcomes $m, M$ but are different on other outcomes. Hurwicz-Arrow’s decision criterion would make $u$ and $v$ indifferent but $u$ and $v$ may not be indifferent according to the protective or the hazardous or the neutral criteria. It follows that the decision criterion determined by $\tau$-anchor utility, a stricter form of HA criterion, does not satisfies all the CM axioms. In particular, the scale invariance axiom requiring that adding a constant to outcomes or multiplying the outcomes with a constant do not change the preference between two acts, is violated. Among five axioms, the rationale for this axiom is far from convincing. An often voiced critique for decision criteria including the protective, hazardous or neutral or the original HA criterion is that they do not permit individualization of attitude toward uncertainty. CM argue that three attitudes towards uncertainty, namely pessimism, optimism and neutral, are implemented by the protective, hazardous and neutral decision criteria. It follows that if two individuals A and B who are both pessimistic (optimistic) have the same vNM utility function, then they must have identical preference under ignorance. Thus, it is impossible to express the idea that while both individuals are pessimistic (optimistic), one is less so than the other. The adoption of Savage’s independence axiom for ignorance is also problematic. Perhaps, the key distinction between decision making under risk and under uncertainty has to do with Savage’s independence (sure-thing principle). A cornerstone in the theory of subjective probability, the axiom has been conclusively shown in many studies beginning with Ellsberg’s ground breaking work [9], to be
violated in the case of uncertainty. Because ignorance is the extreme form of uncertainty it would need a truly compelling argument, which is not there, to justify the independence axiom in the case of ignorance.

A proposal to the problem of decision under ignorance by Gravel, Marchant and Sen is based on the principle of insufficient reason and the expected utility theory. They define a “completely uncertain” decision as the finite set of its consequences and an “ambiguous decision” as a finite set of possible probability distributions over a finite set of consequences. They apply the principle of insufficient reason to assign to every consequence (probability distribution) an equal probability and as comparing decisions on the basis of the expected utility of their consequences (probability distributions) for some utility function. This proposal avoids dealing with difficulties caused by ignorance altogether. A similar proposal was found in the works by Smets et al in the context of decision making with Dempster-Shafer belief function. This family of proposals does not satisfy a basic property of ignorance in Hurwicz-Arrow sense, namely, the invariance under splitting/merging states. For example, adding a small random noise to the outcomes of an act would have a dramatic effect on its utility because it changes the set of different outcomes and hence the probability distribution derived from the principle of insufficient reason.

An assumption made in Anscombe-Aumann’s framework and later adopted in many studies of decision under ignorance, for example and , is that the prizes of acts under uncertainty (ambiguity, ignorance) are von Neumann-Morgenstern utilities or lotteries. This assumption implicitly includes a probabilistic variable in addition to the ignorant variable. In our setting, the assumption is spelled out in two parts: (1) the ignorant variable precedes the roulette variable and (2) the preference on lotteries is represented by vNM utility. The situation is described by the tree on the right hand side in Fig. 3. Consider an act with outcomes in vNM utility where are utility levels and is the linear utility function mapping from monetary values to utility scale. If , then the decision maker will be rewarded with utility level which after all is just an abstract quantity. To get monetary reward, she has to play another round of a roulette gamble that has expected utility . Such gambles are many but all of them are indifferent to the decision maker. Clearly, the outcome-in-utility assumption excludes from consideration the situation in which the ignorance variable precedes the probabilistic one (the tree on the left hand side in Fig. 3).
7. Conclusion

This paper describes a decision model under uncertainty that combines two extreme states of knowledge: complete ignorance and probability. It is motivated by a working hypothesis that uncertainty in general is the result of mixing probability and ignorance. We develop a new utility theory for decision under ignorance by marrying Hurwicz-Arrow’s decision theory for ignorance with the ideas from Anscombe-Aumann’s theory of subjective probability. The main result is a representation theorem by $\tau$-anchor utility function. The theorem establishes the existence of a value $\tau$ that characterizes decision maker’s behavior under ignorance. The characteristic value can be used to compare the extent of tolerance for ignorance between individuals.

The key insight of the new approach to the decision making under uncertainty comprising ignorance and probability is to recognize that ignorant and probability variables are non-reversible in Anscombe-Aumann’s sense. Hence, the ordering of variables is critical and must be included in preferential evaluation of acts. Our approach explains away the puzzling phenomenon of sequential inconsistency in many decision models under uncertainty. Our approach includes maximin expected utility model and the decision model for Dempster’s belief functions as special cases.

Reference


making under complete ignorance. *Mathematical Social Sciences*, 59:15–

[7] A. Dempster. Upper and lower probability induced by multivalued map-

consistency. Technical Report 2441566, Social science research network,


[10] P. H. Giang. Dynamic consistency and decision making under vacuous
belief. In F. G. Cozman and A. Pfeffer, editors, *Uncertainty in Artifi-
cial Intelligence: Proceedings of the 27nd Conference (UAI–2011)*, pages


for decision making under ignorance or objective ambiguity. *Journal of

probability and belief as evidence. *Artificial Intelligence*, 54(3):275–318,


[16] R. M. Hogarth and H. Kunreuther. Decision making under ignorance:


Appendix: Proofs of statements

Proof of Lemma 1

For \( A \in \mathcal{F}(\mathcal{O}) \) define \( A^\dagger = \{ x \in \mathcal{O} | x \succeq A \} \) and \( A^\dagger = \{ x \in \mathcal{O} | A \succeq x \} \). Clearly, both \( A^\dagger \) and \( A^\dagger \) are non-empty (0 \( \in A^\dagger \) and 1 \( \in A^\dagger \)) and because of completeness of \( \succeq \), \( A^\dagger \cup A^\dagger = \mathcal{O} \). By (CA) property, there is a unique \( c \in \mathcal{O} \) such that \( c \sim A \). To see that \( \min(A) \leq c \leq \max(A) \), suppose the contrary that either \( c < \min(A) \) or \( c > \max(A) \). Choose a value \( z \in \mathcal{O} \) such that \( c < z < \min(A) \). On the one hand, \( \{\min(A), \max(A)\} \succeq z \) because the min and max of the constant act \( z \) are less than those of act \( A \). On the other hand, \( z \succ c \) by CA. That contradicts the fact that \( A \sim c \). ■

Proof of Lemma 4

Denote by \( c_i \) the CE of \( A_i \). Suppose \( n = \max_i |A_i| \) - the size of largest \( A_i \). By duplicating prizes in \( A_i \) if necessary, all \( A_i \) can be transformed into ignorant acts \( A'_i \) on the same domain of size \( n \). Because \( \succeq \) is HA preference, \( A_i \sim A'_i \). By (MI), \( \{A_1, A_2, \ldots A_m\} \sim \{A'_1, A'_2, \ldots A'_m\} \). By (RI) \( \cup_i A'_i \sim \{A'_1, A'_2, \ldots A'_m\} \). By the way that \( A'_i \) is constructed from \( A_i \), the sets of distinct elements in \( A'_i \) and \( A_i \) are the same. So, \( \cup_i A_i \sim \cup_i A'_i \). By transitivity, \( \cup_i A_i \sim \{A_1, A_2, \ldots A_m\} \). By (MI), \( \{c_1, c_2, \ldots c_m\} \sim \{A_1, A_2, \ldots A_m\} \). By transitivity, \( \cup_i A_i \sim \{c_1, c_2, \ldots c_m\} \). Thus, \( \mathcal{CE}(\cup_i A_i) = \mathcal{CE}(\{c_1, c_2, \ldots c_m\}) \) ■

Proof of Lemma 5

It follows from its definition (5) and properties of \( \mathcal{CE} \) that \( \gamma \) is continuous in each argument and satisfies the following properties:

(i) for \( 0 \leq x \leq 1 \), \( \gamma(x, x) = x \);
(ii) if \( x \geq x', y \geq y' \), then \( \gamma(x, y) \geq \gamma(x', y') \); and
(iii) for \( 0 \leq x \leq y \leq 1 \), \( \gamma(x, y) = \gamma(\gamma(x, x), \gamma(x, y)) = \gamma(\gamma(x, y), \gamma(y, y)) \).

To prove \( \gamma(x, y) = \gamma(\gamma(x, x), \gamma(x, y)) \). Consider two bags \( f = \{x, y\} \) and \( g = \{x, x, y\} \). Because \( f \) and \( g \) have exactly the same min and max elements \( (x \text{ and } y) \), \( f \sim g \) i.e., they have the same CE (says \( a \)). By definition of \( \gamma \), \( \gamma(x, y) = a \). By ICE property, taking \( A_1 = \{x, x\} \) and \( A_2 = \{x, y\} \), \( \mathcal{CE}(\{x, x, y\}) = \mathcal{CE}(\{\mathcal{CE}(\{x, x\}), \mathcal{CE}(\{x, y\})\}) = \mathcal{CE}(x, a) = a \). Hence, \( \gamma(x, a) = \gamma(\gamma(x, x), \gamma(x, y)) = a \). The second part of (iii) is similar.

By properties (i) and (iii), we have \( a = \gamma(x, 1) = \gamma(\gamma(x, x), \gamma(1, 1)) = \gamma(x, a) \) and \( a = \gamma(x, 1) = \gamma(\gamma(x, 1), \gamma(1, 1)) = \gamma(a, 1) \). It follows from property (ii) that for any \( y \) in the interval \([x, a] \), \( a = \gamma(x, a) \leq \gamma(y, 1) \leq \gamma(a, 1) = a \). The proof of (2) is similar. ■

37
Proof of Theorem 2

(1 ⇔ 2) It is necessary to show that the function $\gamma$ given by (6) is continuous on each argument and satisfy (i), (ii) and (iii). For a fixed value $x_0$, $\gamma(x_0, y)$ is continuous on $y$. In the case $x_0 \leq \tau$ and if $x_0 \leq b \leq \tau$, $\lim_{y \to b} \gamma(x_0, y) = b = \gamma(x_0, b)$. If $b \geq \tau$, $\lim_{y \to b} \gamma(x_0, y) = \tau = \gamma(x_0, b)$. In the case $x_0 \geq \tau$, $\lim_{y \to b} \gamma(x_0, y) = x_0 = \gamma(x_0, b)$. Similarly, it can be shown that $\gamma(x, y_0)$ is continuous on $x$ for any fixed $y_0$. The verification of (i) and (ii) is straightforward.

To show (iii). Set $\gamma(x, y) = z$. There are three cases of $z$ relative to $\tau$ that we have to consider: $z$ is less/equal/greater than $\tau$. First, suppose $z < \tau$. This is the case of the first line in (6), hence, it follows that $z = y < \tau$. From $\gamma(x, y) = x, \gamma(x, y) = y, \gamma(y, y) = y$, (iii) reduces to $z = \gamma(x, y) = y$ which is satisfied. In the second case, suppose $z = \tau$. This is the case of the second line in (6). So, we have $x \leq \tau = z < y$. (iii) reduces to $\tau = \gamma(x, \tau) = \gamma(\tau, y)$ which is satisfied. Finally, suppose $z > \tau$. This is the case of the third line of (6), hence, $\tau < z = x \leq y$. (iii) reduces to $x = \gamma(x, x) = \gamma(x, y)$ which is satisfied.

(1 ⇒ 2) Suppose $\gamma$ is a continuous function and satisfies (i), (ii) and (iii). Set $\tau = \gamma(0, 1)$. Assume $0 < \tau < 1$. By Lemma 5, $\gamma(0, \tau) = \gamma(\tau, 1) = \tau$. By (ii), if $x \leq \tau \leq y$ then $\tau \leq \gamma(0, \tau) \leq \gamma(x, y) \leq \gamma(\tau, 1) = \tau$. In other words, if $\tau$ is between arguments then $\gamma(x, y) = \tau$.

For the case when both arguments are on the same side of $\tau$, assume for a moment that $x, y \geq \tau$. We show by contradiction that $\gamma(x, y) = x$. Suppose, on the contrary, there exists $x_0, y_0 \geq \tau$ such that $\gamma(x_0, y_0) > x_0$. Because of (ii) $\gamma(x_0, 1) = x_1 \geq \gamma(x_0, y_0) > x_0$. By Lemma 5, $x$ in the interval $[x_0, x_1]$, $\gamma(x, 1) = x_1$. Consider set $A_{x_1} = \{x | x \geq \tau, \gamma(x, 1) = x_1\}$. This set includes interval $[x_0, x_1]$. Choose $\tau' = \inf A_{x_1}$. Clearly, $\tau' \leq \tau' \leq x_0 < x_1$. For any $x$ in the neighborhood of $\tau'$ and $\tau' < x \leq x_1$, by definition of $\tau'$, there is $\tau' < \tau' \leq \tau$ such that $x' \in A_{x_1}$. By Lemma 5, $\gamma(x, 1) = x_1$.

We show that for $x < \tau'$, $\gamma(x, 1) \leq \tau'$. Suppose the contrary, $\gamma(x', 1) = \tau'' > \tau'$ for some $x' < \tau'$. By Lemma 5, for any $x \in [x', \tau'']$, $\gamma(x, 1) = \tau''$. Clearly, the intersection $(\tau', x_1] \cap [x', \tau'']$ is not empty. For $x$ in the intersection, $\gamma(x, 1) = x_1$ and $\gamma(x, 1) = \tau''$. Hence $x_1 = \tau''$. Thus, for $x' < \tau'$ and $\gamma(x', 1) = x_1$. This contradicts the fact that $\tau'$ is the infimum of $A_{x_1}$.

So, $\gamma(x, 1) \leq \tau'$ for $x < \tau'$.

For a sequence $\{x_i\}$ that approaches $\tau'$ from below $\lim_{x_i \to \tau'}(\gamma(x_i, 1)) \leq \tau'$. But for a sequence $\{y_i\}$ that approaches $\tau'$ from above $\lim_{y_i \to \tau'} \gamma(y_i, 1) = x_1$. So, $\gamma(\cdot, 1)$ is not continuous at $\tau'$. This contradiction shows that $\gamma(x, y) = x$.
for $\tau \leq x \leq y$. Similarly, one can show that $0 \leq x \leq y \leq \tau$, $\gamma(x, y) = y$. ■

Proof of Theorem 3

($\Leftarrow$) The fact that if the certainty equivalence operator $CE$ satisfies (7) and (8) then assumptions (C), (CA), (MI) and (RI) are satisfied can be directly verified from definition.

($\Rightarrow$) We show that if $\succeq$ is a HA preference and satisfies (C), (CA), (MI) and (RI) then the certainty equivalence operator $CE$ satisfies (7) and (8). By lemma 4, if HA preference $\succeq$ satisfies (C), (CA), (MI) and (RI) then the CE operator has the iterated certainty equivalence property (ICE) and satisfies (8). By lemma 5 and theorem 2, $CE$ must satisfy (7).

Proof of Corollary 1

(i) If $\tau_A = \tau_B$ then $CE_A(f) = CE_B(f)$ for any $f$. So, we need consider the case of strict more tolerance for ignorance. Suppose $0 < \tau_B < \tau_A < 1$. Denote the minimum and maximum prizes of act $f$ by $\underline{f}$ and $\overline{f}$. There are in total 6 possible cases for the topological location of interval $[\underline{f}, \overline{f}]$ relative to interval $[\tau_B, \tau_A]$.

1. If $\underline{f} < \overline{f} \leq \tau_B < \tau_A$ then $CE_B(f) = CE_A(f) = \overline{f}$
2. If $\underline{f} < \tau_B < \overline{f} \leq \tau_A$ then $CE_B(f) = \tau_B < CE_A(f) = \overline{f}$
3. If $\underline{f} < \tau_B < \tau_A < \overline{f}$ then $CE_B(f) = \tau_B < CE_A(f) = \tau_A$
4. If $\tau_B < \underline{f} < \overline{f} \leq \tau_A$ then $CE_B(f) = \underline{f} < CE_A(f) = \overline{f}$
5. If $\tau_B < \underline{f} < \tau_A < \overline{f}$ then $CE_B(f) = \underline{f} < CE_A(f) = \tau_A$
6. If $\tau_B < \tau_A < \underline{f} < \overline{f}$ then $CE_B(f) = CE_A(f) = \underline{f}$.

(ii) If $\tau_B < \tau_A$ then an act $f$ as in case (2) of part (i) is sufficient to show that $CE_B(f) < CE_A(f)$. Conversely, suppose $CE_B(g) < CE_A(g)$ for some $g$ then $\tau_B < \tau_A$ follows by contradiction. Suppose $\tau_B \geq \tau_A$ then applying part (i), $CE_B(f) \geq CE_A(f)$ that is a contradiction. ■

Proof of Theorem 4

Let $u$ and $v$ be the utility functions representing $\succeq_R$ and $\succeq_I$, we have:

\[(p_i : x_i)_{i=1}^n \succeq_R (q_j : y_j)_{j=1}^m \iff \sum_{i=1}^n p_i u(x_i) \geq \sum_{j=1}^m q_j u(y_j).\]
\[\{x_1, x_2, \ldots, x_n\} \succeq_I \{x_1', x_2', \ldots, x_m'\} \iff v(\{x_1, x_2, \ldots, x_n\}) \geq v(\{x_1', x_2', \ldots, x_m'\}).\]
where

\[
\forall A \subseteq \mathcal{O}, \quad v(A) = \begin{cases} 
\max(A) & \text{if } \max(A) < \tau \\
\tau & \text{if } \min(A) \leq \tau \leq \max(A) \\
\min(A) & \text{if } \min(A) > \tau
\end{cases}
\]  

(19)

The proof is done by constructing a counter-example for each of three cases:

0 < \tau < 1, \tau = 0 \text{ and } \tau = 1. \text{ Also the counter example is constructed by choosing values } x_i^j \text{ and } p_i \text{ so that the left and right sides of (12) have different certainty equivalents.}

Suppose that 0 < \tau < 1. Consider (12) with k = 2, n = 2 and choose the following configuration of values 0 = x_1^1 < x_1^2 < \tau < x_2^1 < x_2^2 = 1. Thus, the left hand side of (12) becomes \((p:0,(1-p):x_2^2),(p:x_1^1,(1-p):1)\). For a large enough but strictly less than unity p, both lotteries \((p:0,(1-p):x_2^2)\) and \((p:x_1^1,(1-p):1)\) have CE that are (strictly) less than \tau. Between the lotteries, the latter is strictly preferable to the former because of monotonicity. So, the act under ignorance on the left hand side of (12) is indifferent to

\((p:x_1^2,(1-p):1)\).  

(20)

On the other hand, the right hand side of (12) becomes lottery \((p:\{0,x_1^2\},(1-p):\{x_2^1,1\})\). Because \(x_1^2 < \tau\), \(\{0,x_1^2\} \sim x_1^2\). Because \(\tau < x_2^1\), \(\{x_1^2,1\} \sim x_2^1\). So the right hand side lottery becomes

\((p:x_1^1,(1-p):x_2^2)\).  

(21)

Comparing (20) with (21) and taking into account the facts that \(x_2^1 < 1\) and \((1-p) > 0\) we see that the left hand side of (12) is strictly preferable to the right hand side.

Consider the second case, suppose \(\tau = 0\), \(v\) becomes min function. Choose the following configuration of values \(x_1^1 < x_1^2\) and \(x_2^1 < x_2^2\). The right hand side reduces to lottery \((p:x_1^1,(1-p):x_2^2)\) because \(\{x_1^1,x_1^2\} \sim x_1^1\) and \(\{x_1^2,x_2^2\} \sim x_2^2\). The left hand side of (12) becomes an ignorant act \((p:x_1^1,(1-p):x_1^2),(p:x_2^1,(1-p):x_2^2)\). We show that each of the lotteries inside the ignorant act is strictly preferable to the lottery on the right hand side of (12). For \(0 < p < 1\), we have \(pu(x_1^1) + (1-p)u(x_1^2) > pu(x_1^1) + (1-p)u(x_2^2)\) and \(pu(x_1^2) + (1-p)u(x_2^2) > pu(x_1^1) + (1-p)u(x_2^2)\). So, \(p:x_1^1,(1-p):x_2^2 \succ (p:x_1^1,(1-p):x_2^1)\) and \(p:x_1^1,(1-p):x_2^2 \succ (p:x_1^1,(1-p):x_2^1)\). Thus, \(\{(p:x_1^1,(1-p):x_2^1),(p:x_1^1,(1-p):x_2^2)\} \succ (p:x_1^1,(1-p):x_2^2)\). This is a violation of \((RIR)\). The proof for the case \(\tau = 1\) is similar.