

Chapter 20

Decision Making Under Ignorance

Phan H. Giang

Abstract Ignorance is an extreme form of uncertainty. In most narrow and technical sense, it means inability to assign a meaningful probability to the phenomena of interest. In more general sense, the state of ignorance is the result of the absence of knowledge about structural factors that influence the issues, the lack of reliable information or inability to completely determine the space of alternatives and consequences. We argue that the practice of casually papering over the ignorance with subjective judgments and analytic assumptions can have serious consequences. This chapter provides a structured survey (and necessarily selective) of significant ideas and proposals for decision making under ignorance, from the ground breaking work by Hurwicz and Arrow to the latest result of τ -anchor utility theory. A careful analysis and isolation of ignorance in the system of knowledge about a subject or a problem is of particular importance in the context of risk assessment and risk management.

20.1 Introduction

In the article *Rumsfeld's Knowns and Unknowns: The Intellectual History of a Quip* appeared in *The Atlantic* magazine on March 2014, journalist David Graham (2014) tells a fascinating story about the memorable quip that Donald Rumsfeld, who served as Secretary of Defense in George W. Bush's administration, is remembered for:

As we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns - the ones we don't know we don't know.

The statement was a reply to a question raised by reporter Jim Miklaszewski of *NBC News* in an official news briefing by US Department of Defense in February 2002,

P.H. Giang (✉)
George Mason University, MS 1J3, 4400 University Dr., Fairfax, VA 22030, USA
e-mail: pgiang@gmu.edu

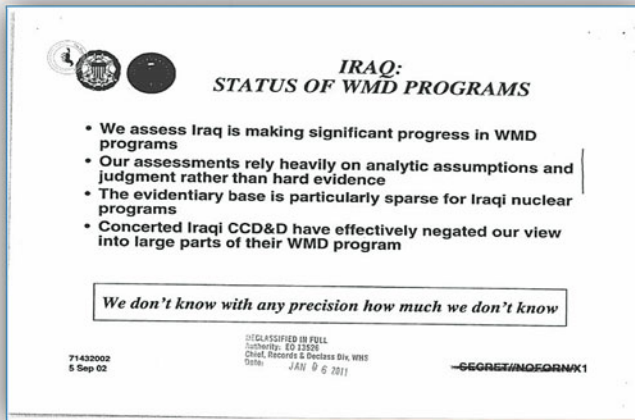


Fig. 20.1 Rumsfeld's memo to Gen. Myers

demanding the evidence for the claims that Iraq's dictator Saddam Hussein had weapons of mass destruction (WMD) and was willing to transfer them to terrorist networks.¹

It was a difficult question for Rumsfeld to answer and, to give him credit, he was smart and in this case at least, truthful. We now know his true state of knowledge from a recently declassified secret memo he sent to the Chairman of the Joint Chiefs of Staff, Gen. Richard Myers, dated September 5, 2002.² In that memo, Rumsfeld included a slide (Fig. 20.1) from the top secret presentation prepared on his specific request by the Chief of Military Intelligence. The presentation summarized the information that US Government knew about the status of WMD programs [of Iraq]. The salient points of the slide were that the assessments about significant progress of Iraq's WMD program relied mostly on *analytic assumptions* and *judgments* (emphasis added) not the hard evidence; the evidentiary base for the assessments was sparse and finally, the concerted efforts of Iraq's government to hide their intention effectively blocked US intelligence view into WMD program. The slide concluded "We don't know with any precision how much we don't know." The state of knowledge that Mr. Rumsfeld described in such a convoluted way, fits the technical definition of ignorance.

We all know how the war that was sold to the public mostly on supposed WMD dangers, ended. Admittedly, it was quite possible that if the WMD were not an issue, something else could be used to sell the war. Everyone can draw a lesson from this

¹For the context of the quip, the entire transcript of the news briefing can be found here: <http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636>.

²http://library.rumsfeld.com/doclib/20110202_20020909toMyersreWMD.pdf.

tragic history. One lesson that we can draw from this story is that substituting so called analytic and judgmental assumptions for the hard evidence may have adverse, potentially disastrous, consequences.

Most applications in science, engineering, medicine, business, and military express uncertainty in the language of the probability theory. Probability estimation is based on data and/or human opinions. Accurate estimation of probability requires a well-defined state space and availability of reliable and relevant data in sufficient quantity. Reliability and relevancy are essential for data to be useful. Subjective or personal probability is estimated on the basis observed or hypothetical behavior of individuals. An essential condition for estimation of subjective probability is that the behaviors must be internally consistent. Without sufficient data or behavioral observations, probability can be arrived by invoking the philosophical principle of insufficient reason (or its close cousin the principle of maximum entropy) that assigns equal probability to alternative states. An important but casually dismissed implication of using probability is that an individual commits herself to very strong epistemic assumption according to which anything that can be learned about a phenomenon is known to the individual. In that state of knowledge, the outcomes are random events which, from the individual's standpoint, are quantitatively exchangeable with the outcomes of coin tossing, roulette spinning, or radioactive decay.

However, in many practical situations, the real rationale behind the adoption of probability is convenience rather than justification. Almost a century ago, Knight (1921) and Keynes (1921) came to the conclusion that probability theory is not a universally appropriate language of uncertainty. In particular, the principle of insufficient reason is problematic because the same real situation can be modeled in different ways resulting in different sets of alternatives. The rationality of imposing the internal consistency, the critical condition for the estimation of subjective probability, is also debatable. Gilboa, Postlewaite, and Schmeidler argue that rationality requires a compromise between internal coherence and justification, the same way that the compromises are forced on moral dilemmas (Gilboa et al. 2009). When the requirements for internal coherence and justification contradict each other, they wrote "it is more rational to admit that one does not have sufficient information to generate a prior than to pretend that one does." More mundane challenges in probability estimation that everyone can bear witness to are routine unavailability of the data necessary for estimation, the cost of collecting data is too expensive for one's budget, the relevant data are scarce or conflicting, the data that you can trust are irrelevant and finally in competitive games, the information about the opposing party can be intentionally misleading.

The opposite epistemic state of knowing everything is the state of ignorance. This is an extreme form of uncertainty when an individual has no reliable information about the phenomenon of interest and therefore is not able to produce, in any meaningful way, a probability distribution.

In the context of risk assessment and risk management (Aven and Steen 2010), the term ignorance is used to describe the situations characterized by the lack of knowledge about the subject, a poor basis for probability assignments (data are

scarce, unreliable, conflicting, etc.) and inability of the decision makers to fully determine all possible consequences. A more structural view held by the health and safety executive (HSE) in the UK (Health and Safety Executive 2001) highlights the key difference between ordinary uncertainty and ignorance is the knowledge about the factors that influence the issues. Uncertainty refers to a state of knowledge in which influencing factors are known, but the likelihood of consequences or effects cannot be precisely described. Ignorance, on the other hand, refers to the lack of knowledge about the factors influencing an issue. Another condition often associated with the term ignorance is the sample space ignorance (SSI) (Pushkarskaya et al. 2010) where the decision maker has difficulty in determining the set of possible alternative states. Daily situations that can be identified with ignorance occur without much attention perhaps because the stakes involved are quite low. For example, you are deciding whether or not to buy warranty for a piece of electronics made by a new to the market manufacturer but the information necessary to set up an optimization problem such as the rate of failure and the cost of fixing it is not known. Practically, in this situation you have to make decision under ignorance (Hogarth and Kunreuther 1995).

It is necessary to emphasize that, strictly speaking, the situations of pure or complete ignorance are rare. The rationale for studying decision making under ignorance is not based on the argument that pure or complete ignorance occurs frequently in practice. The reason that makes study of decision making under ignorance necessary is that an adequate description of uncertainty for most practical problems almost always includes some form of the singular state of knowledge. If anyone still dismisses the relevance of ignorance in the age of Big data, Rumsfeld's episode should be enough to refute that argument. Instead of making the ignorance disappear, modeling efforts should focus on separating the parts of the problem for which reliable data are available from those where reliable evidence is absent.

This chapter is structured as follows. Section 20.2 offers a review of literature on decision under ignorance. In Sect. 20.3, a new utility theory under ignorance is developed by imposing a condition for the certainty equivalent operator on Hurwicz–Arrow's decision theory under ignorance. Finally, Sect. 20.4 contains the discussion and examples.

20.2 A Brief Review of Decision Under Ignorance

Before going into technical presentation, we list the basic set of notations used in this paper. Our framework includes *variables* which are denoted by upper case letters I, X, Y , and so on. A variable (X) has a *domain* (Ω_X). A *decision* or *act* involves one or more variables. It is a mapping from the state space formed by the domains of its variables to the set of prizes. The set of prizes denoted by \mathcal{O} . Acts are denoted by lower case letters such as d, f, g . The decision maker's behavior is described by preference relation \succeq on acts. In this chapter \succeq is assumed to be a weak order with some exceptions which will be made explicitly.

With the development of formal decision theory under risk in the 1940s spurred by pioneering work of von Neumann and Morgenstern (1953), economists like Shackle (1949), Hurwicz and Arrow (1977) started pondering the question how an individual makes decision if she cannot associate any probability distribution to consequences of an act.

20.2.1 Hurwicz–Arrow’s Theory of Choice Under Ignorance

A short paper (Arrow and Hurwicz 1977), reprinted in 1977, outlines the theory of decision under ignorance developed by Hurwicz and Arrow in early 1950s. The result was first obtained by Hurwicz and then improved by Arrow. The axiomatic approach they employed laid the foundation for later studies of decision making under ignorance.

We review the basic results of HA theory in our setting. Consider a collection of variables $\{I_1, I_2, \dots\}$ whose domains are sets Ω_{I_i} . A *decision* or *act* defined on variable I_i is a mapping $f : \Omega_{I_i} \rightarrow \mathcal{O}$ where \mathcal{O} is the (normalized) *prize space*. The domain of f and is denoted by $\Omega(f)$. A *decision problem* is a non-empty set of decisions that have the same domain. Denote the set of acts defined on variable I by \mathcal{D}_I and the set of all acts by \mathcal{D} . Further we make technical assumptions: Ω_{I_i} are finite subsets of the set of natural numbers \mathbb{N} and \mathcal{O} is the (normalized) *prize space* which is assumed to be the real unit interval $\mathbb{R}_{0,1}$.

In the HA paper, an optimal operator $\hat{\cdot}$ that maps each decision problem A to a subset of optimal acts \hat{A} . In this review, the optimal operator construct is replaced by a preference relation \succeq on the set of acts \mathcal{D} . We assume that \succeq is a weak order, i.e., it is reflexive, transitive, and complete. Formally, for any $f, g, h \in \mathcal{D}$ $f \succeq g$, if $f \succeq g$ and $g \succeq h$ then $f \succeq h$; either $f \succeq g$ or $g \succeq f$. From preference relation \succeq , a strict preference (\succ) and indifference (\sim) are defined as follows. $f \succ g$ means $f \succeq g$ and $g \not\succeq f$. $f \sim g$ means $f \succeq g$ and $g \succeq f$. For act f and prize $x \in \mathcal{O}$ if $f \sim x$ then x is called *certainty equivalent* of f .

The correspondence between HA operator $\hat{\cdot}$ and \succeq is as follows: For a decision problem A , $f \in \hat{A} \Leftrightarrow \forall g \in A, f \succeq g$. From the definition it follows that if $f, g \in \hat{A}$ then $f \sim g$.

The lasting impact of HA theory is due to the axioms that elegantly capture the essence of the notion of ignorance. Formally, HA theory presupposes that \succeq must satisfy four axioms, originally named as the properties A–D as follows. Property A requires that \succeq is a weak order. Property D or the weak dominance property, stipulates that if f_1 and f_2 are acts on the same domain $\Omega(f)$, and $\forall w \in \Omega(f), f_1(w) \geq f_2(w)$ then $f_1 \succeq f_2$. Properties A and D are standard axioms of preference which are not specific for preference under ignorance but hold for preference under risk and uncertainty. Properties B and C on the other hand are specific requirements that make sense only in the case of ignorance.

Property B is called the invariance under relabeling. Formally, if acts f_1, f_2 are isomorphic in the sense that there is an one-to-one mapping $h : \Omega(f_1) \rightarrow \Omega(f_2)$ such that $\forall s \in \Omega(f_1), f_1(s) = f_2(h(s))$ then f_1 and f_2 are indifferent $f_1 \sim f_2$.

The acts f_1, f_2 in (B) have domains of the same cardinality. An act can be viewed as a vector of its prizes. Thanks to the one-to-one mapping h , vector f_2 can be seen as obtained from vector f_1 by a permutation. Property B requires that the (preferential) valuation of a vector does not depend on the positions of its values. The states in the domain of Ω_{f_i} are treated symmetrically. No state is considered more likely than another. If the uncertainty is described a probability function only the uniform distribution among all distributions satisfies this symmetry property.

Property C, invariance under deletion of duplicate states, is a requirement unique for the state of ignorance. Suppose f_1, f_2 are two decisions, f_2 is said to be *derived from* f_1 by deleting duplicate states if (1) $\Omega(f_2) \subset \Omega(f_1)$ and f_1 and f_2 are coincide 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 f_1 and f_2 are indifferent. Viewed as vectors of prizes, this property requires that if vector f_1 has two equal components $f_1(w_i) = f_1(w_j)$ then one of the component can be deleted. The newly obtained vector f_2 is indifferent to the original vector f_1 . The term “duplicate states” refers to the states that f -map to the same prize.

This property addresses one of the epistemic challenges that create the state of ignorance, namely, the inability to determine and justify the complete set of alternative states [for example, the SSI (Pushkarskaya et al. 2010)]. Let us consider a simple example.

Example 1. Suppose that you are standing in front of two identical urns that contain 100 balls each. You are told that the first urn may include balls of two different sizes (small and big) but no information about their proportions. You need to decide how much to pay for a bet that pays \$0 if a ball drawn from the urn is small and \$1 if the ball is big. In the second urn, you have the same information about the sizes and proportions as in the first urn. But on top of it, you are told that small balls have only one color—white, while big balls can be painted in one of three colors: red, green, or blue. Again, you do not know the proportions of balls by colors. You have to decide how much you are willing to pay for a bet that pays \$0 if the ball is small-white, and \$1 if the ball is big-red or big-blue or big-green.

Let us formally describe the urns. For the first one, the domain of the variable “size” is {small, big} and the bet is $f_1(\text{small}) = 0$ and $f_1(\text{big}) = 1$. For the second one, the state space has elements {small.white, big.red, big.green, big.blue} and the bet is $f_2(\text{small.white}) = 0, f_2(\text{big.red}) = 1, f_2(\text{big.green}) = 1, f_2(\text{big.blue}) = 1$. Property C requires that bets f_1 and f_2 are indifferent, that is, you would pay the equal prices for f_1 and f_2 . It is quite easy to see the reasonableness of the indifference in this case because the color is clearly irrelevant to bet f_2 .

To make color variable relevant, let us modify f_2 into bet f'_2 as follows $f'_2(\text{small.white}) = 0, f'_2(\text{big.red}) = 0, f'_2(\text{big.green}) = 1, f'_2(\text{big.blue}) = 1$. Property C still requires that $f_1 \sim f'_2$. Thus, under property C, the price you pay for a bet depends on the *set of distinct prizes*, not the vector of prizes where a prize

can appear multiple times. The rationale of this axiom is that if it is difficult for an individual to determine and justify the space of alternative states then the evaluation should not depend on the notational form tied to a particular state space. If you are ignorant about the proportions of balls by sizes, colors then it does not matter if you model the act using only size variable or color variable or both. One can say that the domain of f_2 is obtained from the domain of f_1 by splitting states or vice versa domain of f_1 is obtained from merging states in the domain of f_2 (big into big.red, big.blue, big.green). ■

Property C is unique for ignorance. No probability distribution has this property because the uniformity of a distribution is destroyed when states are splitted/merged. Thus, properties B and C formulated by Hurwicz and Arrow capture the essence of ignorance. Much later in the context of statistical inference, Walley (1996) proposes two criteria that a 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. The main result of HA theory is a theorem that characterizes the preference relation that satisfies properties A through D.

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 $\mathcal{Z}^2 = \{\langle a, b \rangle \mid 0 \leq a \leq b \leq 1\}$ 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 of the same domain Ω_I ,*

$$f \succeq g \text{ if } \left\langle \min_w f(w), \max_w f(w) \right\rangle \succeq^2 \left\langle \min_w g(w), \max_w g(w) \right\rangle. \quad (20.1)$$

Let us call the preference relation that satisfies the conditions of HA theorem *HA preference relation*. Few implications can be drawn from HA theorem. First, the preferential comparison between two acts reduces to comparing its extreme values. Intermediate prize values do not matter. Under ignorance the state space is not important, therefore, acts can be identified with their prizes. The order of prizes is not important either. The set of finite non-empty *bags*³ of \mathcal{O} is denoted by $\mathcal{F}(\mathcal{O})$. Thus, two symbols denote the same thing $\mathcal{D} \equiv \mathcal{F}(\mathcal{O})$. From now on, we have a flexibility to denote, at our convenience, an act under ignorance either by a vector or a set of its prizes. Finally, HA theorem does not fully describe the preference relation \succeq . It does not specify the preference between two acts when the comparisons of their minimal elements and maximal elements point to opposite directions.

Facing the same informational ignorance and the same set of prizes, the prices that different individuals would be willing to pay for the bet are different. Intuitively, a pessimistic person would pay more attention to the negative side, i.e., the worst

³A bag is different from a set in the sense that a bag can contain the same item multiple times while all the elements in a set are distinct.

prize in the set while an optimistic person would look more into the positive side, i.e., the best prize in the set. Hurwicz proposed to quantify that attitude by a parameter α to control the weights attached to the worst and best prizes in a set. According to Hurwicz’s α -rule, for acts $f_1, f_2 \in \mathcal{F}(\mathcal{O})$

$$f_1 \succeq_{\alpha} f_2 \text{ iff } \alpha \min(f_1) + (1 - \alpha) \max(f_1) \geq \alpha \min(f_2) + (1 - \alpha) \max(f_2). \quad (20.2)$$

That is, comparing two acts, an individual must calculate the combinations of the worst and the best prizes with α coefficient for each act and then compares the calculated values. For example, acts that have the same combined value are indifferent. In particular, the value $\alpha \min(f) + (1 - \alpha) \max(f)$ is the certainty equivalent of act f .

Hurwicz’s rule or criterion belongs to the family of preferences sanctioned by HA theorem. Moreover, it is intuitively appealing because α parameter allows a smooth variation of the degree of pessimism in different individuals. An extremely pessimistic individual who focuses exclusively on the worst possibility has $\alpha = 1$ while on the opposite end, an extremely optimistic individual who focuses on the best prizes has $\alpha = 0$. Despite its popularity and appeal, Hurwicz’s rule suffers from a drawback—it is sequentially inconsistent. The following example clarifies the problem.

Example 2. Consider a urn of 100 balls. A ball has two characteristics: size (small/large) and color (black/white). The composition of the balls is not known (on both size and color). A bet is offered whose rewards depend on both size and color as follows. $f(\text{small.black}) = 1, f(\text{small.white}) = 0.4, f(\text{large.black}) = 0.7,$ and $f(\text{large.white}) = 0$. Figure 20.2 presents two ways of viewing the bet.

Suppose an individual has $\alpha = 0.6$ (slightly pessimistic). First, she can reason as follows (Fig. 20.2a). Suppose the size of the ball is small, she has a bet on color only. The set of rewards in this case is $\{1, 0.4\}$. Using Hurwicz’s rule the set is

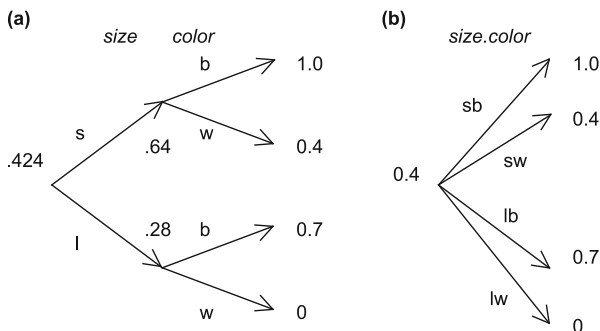


Fig. 20.2 Hurwicz’s rule: sequential inconsistency illustration: (a) “size” is followed by “color” and (b) “size” and “color” are merged

indifferent to $0.64 = 0.4 * 0.6 + 0.4 * 1$. With similar reasoning, if the ball turns out to be large ball then she gets a set of prizes $\{0, 0.7\}$ which is indifferent to 0.28. Under the ignorance about the proportions of small and large balls, the set of rewards $\{0.28, 0.64\}$ is indifferent to 0.424.

The individual wants to verify the value she calculated using a different line of reasoning. Because she has no information about both size and color, she lumps two variables into one called “size.color”. Under this view, the set of rewards is $\{1, 0.7, 0.4, 0\}$ (Fig. 20.2b). Using $\alpha = 0.6$ she finds the set of rewards is indifferent to 0.4.

The difference is puzzling because no new information is added in case (b). The difference between (a) and (b) is due to purely subjective view of the individual. ■

The notion of sequential consistency is closely related to an important identity in probability theory—the law of iterated expectation. Basically, if X and Y are random variables then

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]. \quad (20.3)$$

The unconditional expectation of X equals the expectation of the conditional expectation of X given Y . In particular, this law allows a divide-and-conquer strategy to compute $\mathbb{E}[X]$. One can divide the state space of X into several subclasses and create a variable Y to be the class ID, i.e., $\mathbb{E}[X|Y = j]$ is the mean of class j . Compute the mean for each class and then compute the mean of those conditional expectations. From a conceptual point of view, the law is assuring because no matter how the state space of X is divided the final value of expectation remains the same.

As the example shows, Hurwicz’s criterion does not have this property. That is, the value of the certainty equivalent of a set of prizes calculated by Hurwicz’s rule (denoted by $\mathcal{CE}_\alpha(A)$) depends on the way the set is partitioned.

Formally, suppose A is a set of prizes and $\{B_i | 1 \leq i \leq m\}$ is a partitions of A , i.e., $A = \cup_{i=1}^m B_i$ and $B_i \cap B_{i'} = \emptyset$ for $i \neq i'$. Equation (20.4) below is referred to as the *law of iterated certainty equivalence*. The law says that the certainty equivalent of a set of prizes is equal to the certainty equivalent of a set of certainty equivalents calculated for each subset of the partition. In other words, the certainty equivalent is invariant of the partition of the set.

$$\mathcal{CE}(A) = \mathcal{CE}(\{\mathcal{CE}(B_i) | 1 \leq i \leq m\}). \quad (20.4)$$

In general, Hurwicz’s rule does not satisfy that law. But there are two exceptions: for $\alpha = 0$ and $\alpha = 1$. The fact that (20.4) holds if $\mathcal{CE}_0(A) = \max(A)$ and $\mathcal{CE}_1(A) = \min(A)$ is not difficult to prove. Later we will show a sufficient statement that (20.4) holds only if $\alpha = 0$ or $\alpha = 1$.

20.2.2 Decision Under Ignorance: Alternatives to HA Theory

One of complaints about HA theory is its reliance only on two extremal values and ignore the intermediate values in the set even the result is derived from reasonable axioms. There are many attempts to relax the HA axioms to account for non-extremal values. We review representative works and ideas in the following.

Cohen and Jaffray (CJ) (1980) described a system of axioms for rational behavior under complete ignorance. CJ theory assumes a state space Ω 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 $(fPg \vee gPf)$. Intuitively one can view R as (non-strict) preference and I as indifference but the difference between R and \succeq 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 as follows:

$$fDg \Leftrightarrow \forall w \in \Omega, f(w) \geq g(w) \quad \text{and} \quad \exists w_0 \in \Omega, f(w_0) > g(w_0). \tag{20.5}$$

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., $fDg \Rightarrow fPg$. This axiom implies that states are non-null. 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” (Cohen and Jaffray 1980). An example of CJ rational criterion is

$$fPg \Leftrightarrow \begin{cases} (m_f + M_f > m_g + M_g) \text{ or} \\ (m_f + M_f = m_g + M_g, M_f - m_f < M_g - m_g) \text{ or} \\ (m_f + M_f = m_g + M_g, M_f - m_f = M_g - m_g \\ \min_{\Omega'} f > \min_{\Omega'} g \end{cases}, \tag{20.6}$$

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 Congar and Maniquet (2010), Congar and Maniquet (CM) investigated axiomatic system for a rational decision under ignorance which is defined as “no available information regarding plausible probability distributions over the possible outcomes.” An act is a vector of outcomes which are von Neumann–Morgenstern utilities. CM’s assumption that the outcome is vNM utility makes it clear that the ignorant variable precedes the risk variable in their model.

CM theory includes 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). The quasi-transitivity property is inspired by Cohen–Jaffray work reviewed earlier. Only three decision criteria satisfy these requirements. The first criterion, 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 second decision criterion, *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, the third criterion, *neutral* criterion, is the conjunction of both protective and hazardous criteria. A prominent property is that all three criteria compare acts by restricting the attention to the states in which the outcomes are different. This feature, due to Savage’s axiom, is also present in the example of CJ rational behavior as in Eq. (20.6). The Savage’s independence axiom differentiates CM theory 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. 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. 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 toward uncertainty, namely pessimism, optimism, and neutral, are implemented by the protective, hazardous, and neutral decision criteria. However, if individual A and individual B who are both pessimistic (optimistic) and have the same vNM utility function and then they have identical preference under ignorance. It is impossible to express the idea that while both 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 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 (Ellsberg 1961), to be violated in decision under uncertainty. Because ignorance is the extreme form of uncertainty it would require truly compelling arguments, which are not there, to convince the validity of the independence axiom in the case of ignorance.

A proposal to the problem of decision under ignorance by Gravel et al. (2012) is based on the principle of insufficient reason and 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 would then 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 (2005) in the context of decision making with Dempster–Shafer belief functions. 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.

Viewing an act under ignorance as a set of prizes naturally suggests using a familiar descriptive statistic as the certainty equivalent of the set. Among three basic statistics, mean, mode, and median, the calculations of mean and mode rely on the frequency information which is not known in the state of ignorance and therefore excluded. Nitzan and Pattanaik (1984) describes a system that characterizes the median criterion according to which the certainty equivalent of an act under ignorance is equal to the median of its set of prizes. The most important property of the median statistic is its insensitivity to the outliers, i.e., the extreme (max and min) values of a set can change dramatically without affecting its median value. In this sense, the median criterion is radically different from the prescription of HA theory. The most obvious problem with the median criterion is its violation of HA property C. The certainty equivalent of an act is dependent on the way state space is modeled or merging/splitting of states.

The axiomatic approach to decision under ignorance has been discussed, among others, by Maskin (1979), Nehring and Puppe (1996), and more recently Puppe and Schlag (2009). A literature survey on the topic by Barbera, Bossert, and Pattanaik is given in Barbera et al. (2004).

20.3 A Utility Theory Under Ignorance

We have seen in the previous section the approaches to decision under ignorance that reject some of the assumptions in HA theory. In this section we present a theory that accepts all the HA assumptions and further imposes the law of iterated certainty equivalence that Hurwicz’s α -rule, the most famous criterion in the family of criteria sanctioned by HA theory, violates. The result was first reported in Giang (2011) where the proofs for the propositions in this section are found.

We assume a preference relation \succeq which is a weak order and satisfies HA axioms A–D (Sect. 20.2.1). We make two technical assumptions which are not part of HA theory. First, when the prize set of an act is a singleton (mapping all states into a single prize) we have a constant act. In such a case the uncertainty (ignorance included) about the states does not matter. We assume that the preference among constant acts is exactly the arithmetic order.

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

For $z \in \mathcal{F}(\mathcal{O})$ define $z^\uparrow = \{x \in \mathcal{O} \mid x \succeq z\}$ and $z^\downarrow = \{x \in \mathcal{O} \mid z \succeq x\}$. Clearly, both z^\downarrow and z^\uparrow are non-empty ($0 \in z^\downarrow$ and $1 \in z^\uparrow$) and because of completeness of \succeq , $z^\downarrow \cup z^\uparrow = \mathcal{O}$. By (CA) property, there is a unique $x \in \mathcal{O}$ such that $x \sim z$. Thus, for each act $A \in \mathcal{F}(\mathcal{O})$ there is a unique prize $c \in \mathcal{O}$ which is the *certainty equivalent* of A . It follows from HA Theorem 1 that $\min(A) \leq c \leq \max(A)$. To see that $\min(A) \leq \mathcal{CE}(A) \leq \max(A)$, suppose the contrary that either $\mathcal{CE}(A) < \min(A)$ or $\mathcal{CE}(A) > \max(A)$. Choose a value $z \in \mathcal{O}$ such that $\mathcal{CE}(A) < z < \min(A)$. 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, $z \succ \mathcal{CE}(A)$ by CA. That contradicts the fact that $A \sim \mathcal{CE}(A)$. \mathcal{CE} will be referred to as *certainty equivalent operator*. Because a HA preference relation \succeq completely determines (and is completely determined by) its certainty equivalent operator, we can interchangeably discuss the properties of \succeq and \mathcal{CE} .

The second technical assumption is about continuity of preference relation (equivalently of ce operator). Viewing acts as vectors 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 $f_i = \{x_{i1}, x_{i2}, \dots, x_{in}\}$. The sequence $(f_i)_{i=1}^\infty$ is said to converge to act $f = \{x_1, x_2, \dots, x_n\}$ (notation $\lim_{i \rightarrow \infty} f_i = f$) if $\lim_{i \rightarrow \infty} x_{ij} = x_j$ for $1 \leq j \leq n$.

Assumption 2 (Continuity of Certainty Equivalence Operator—C). If $\lim_{i \rightarrow \infty} f_i = f$ then $\lim_{i \rightarrow \infty} \mathcal{CE}(f_i) = \mathcal{CE}(f)$.

This assumption simply says that a small change in the prizes of an act under ignorance would not lead to a jump in the certainty equivalent. If $(f_i)_{i=1}^\infty$ converges to f and every member of that series is preferred to g then f is preferred to g . Because $f_i \succeq g$, $\mathcal{CE}(f_i) \geq \mathcal{CE}(g)$. So $\lim_{i \rightarrow \infty} \mathcal{CE}(f_i) \geq \mathcal{CE}(g)$. By (C), $\lim_{i \rightarrow \infty} \mathcal{CE}(f_i) = \mathcal{CE}(f)$. Thus, $\mathcal{CE}(f) \geq \mathcal{CE}(g)$ or equivalently $f \succeq g$.

The following lemma summarizes the properties of \mathcal{CE} operator.

Lemma 1. Suppose \succeq is a HA preference relation on $\mathcal{F}(\mathcal{O})$ that satisfies properties (CA) and (C) then operator \mathcal{CE} is well defined and satisfies:

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

The ce operator \mathcal{CE} is a function that maps from $\mathcal{F}(\mathcal{O})$, the set of finite subsets of reals in the unit interval, to the unit interval \mathcal{O} . Because \mathcal{CE} is a certainty

equivalent operator for HA preference, only the maximal and minimal elements of its argument matter. So we can define a two-placed function $\gamma : \mathcal{Z}^2 \rightarrow \mathbb{R}_0^1$ where $\mathcal{Z}^2 = \{\langle a, b \rangle \mid 0 \leq a \leq b \leq 1\}$ that carries all information of \mathcal{CE} .

$$\forall A \in \mathcal{F}(\mathcal{O}), \mathcal{CE}(A) = x \Leftrightarrow \gamma(\min(A), \max(A)) = x. \tag{20.7}$$

The following lemma details the properties of γ .

Lemma 2. *Suppose $\mathcal{CE} : \mathcal{F}(\mathcal{O}) \rightarrow \mathcal{O}$ satisfies Unanimity, Range, Monotonicity, Continuity and Iterated certainty equivalence and γ is defined via (20.7) then*

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

It is easy to verify that γ is continuous in each argument and satisfies: (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 $0 \leq x \leq y \leq 1, \gamma(x, y) = \gamma(\gamma(x, x), \gamma(x, y)) = \gamma(\gamma(x, y), \gamma(y, y))$. By properties (i) and (iii), we have $a = \gamma(x, 1) = \gamma(\gamma(x, x), \gamma(x, 1)) = \gamma(x, a)$ 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 $[x, a], a = \gamma(x, a) \leq \gamma(y, 1) \leq \gamma(a, 1) = a$. Symmetrically, (2) can be proved.

It turns out that γ function corresponding to a HA preference relation that satisfies the law of iterated certainty equivalence must have a special functional form.

Lemma 3. *Suppose $\gamma : \mathcal{Z}^2 \rightarrow \mathbb{R}_0^1$ two following statements are equivalent*

- (1) *γ 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*

$$\gamma(x, y) = \begin{cases} y & \text{if } y \leq \tau \\ \tau & \text{if } x \leq \tau \leq y \\ x & \text{if } x \geq \tau \end{cases} \tag{20.8}$$

Combining Lemmas 2 and 3 leads to a representation theorem.

Theorem 2. *A HA preference relation \succeq on $\mathcal{F}(\mathcal{F}(\mathcal{O}))$ that satisfies (C), (CA) and the law of iterated certainty equivalence iff there exists a value $\tau \in \mathbb{R}_0^1$ such that the certainty equivalent operator \mathcal{CE} of \succeq has the form: for $A_i \in \mathcal{F}(\mathcal{O}), 1 \leq i \leq m$*

$$\mathcal{CE}(A_i) = \begin{cases} \max(A_i) & \text{if } \max(A_i) \leq \tau \\ \tau & \text{if } \min(A_i) \leq \tau \leq \max(A_i), \\ \min(A_i) & \text{if } \min(A_i) \geq \tau \end{cases} \tag{20.9}$$

$$\mathcal{CE}(\cup_{i=1}^m A_i) = \mathcal{CE}(\{\mathcal{CE}(A_i) \mid 1 \leq i \leq m\}). \tag{20.10}$$

While Eq. (20.10) is the law of iterated certainty equivalence, Eq. (20.9) describes the special form that ce operator must have. The central role is the value τ . The behavior of \mathcal{CE} operator has three cases depending on the position of τ relative to the set (actually its minimal and maximal values). If the entire set of prizes lies above τ then \mathcal{CE} behaves like $\min()$ function. If the entire set of prizes lies below τ then \mathcal{CE} behaves like $\max()$ function. If τ is in between the minimal and maximal elements of the set then the entire set of prizes is indifferent to τ . For that reason, τ in (20.9) is called the *characteristic value* of a preference relation. Since the behavior of an individual decision maker under ignorance is described by a preference relation, τ can also be viewed as the characteristic value of the individual.

Another interesting property of Eq. (20.9) is that the ce of a set of prizes is the value between the \min and \max elements of the set that minimizes the distance to the characteristic value. For this reason, the function in (20.9) will be called the τ -*anchor* utility function. Imagine an ideal rubber cord, one end of it is fixed to the anchor τ . The other end of the cord is fixed to a movable point which is allowed to move within interval $[m_A, M_A]$ where $m_A = \min(A)$ and $M_A = \max(A)$. The ce of a set is the point where equilibrium is attained.

To get an interpretation for the characteristic value one can use the equality $\tau = \mathcal{CE}(\{0, 1\})$. τ 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 the consequences. No prize in the set of possible prizes is excluded.

We can see how the value of τ can be used to classify individual's qualitative attitude toward ignorance. As ignorance is an extreme form of uncertainty one can reasonably argue that an uncertainty averse (seeking) individual must demonstrate the ignorance averse (seeking) attitude. Suppose a "probabilistically sophisticated" individual has a vNM utility function under risk $u(x)$. Facing the prize set $\{0, 1\}$ with no probability distribution, she would adopt the uniform distribution. In effect, she converts the act under ignorance to a fair coin lottery ($H:0, T:1$) that brings reward of \$0 if the coin lands *Head* and \$1 if the coin lands *Tail*. The ce of this lottery is $c_u^{un} = u^{-1}(0.5 * u(0) + 0.5 * u(1))$. The probabilistically sophisticated person is by definition uncertainty neutral (Epstein 1999). A uncertainty averse (seeking) attitude implies that the certainty equivalent under ignorance is less (more) than the certainty equivalent under the uniform probability. This leads to the following classification of attitude toward ignorance. For an individual with utility function under risk u , her attitude toward ignorance is averse (seeking) if $\tau \leq (\geq) c_u^{un}$. For example, if an individual is risk neutral with vNM utility function $u(x) = x$, then $c_u^{un} = 0.5$. $\tau = 0.4 < 0.5$ is classified as ignorance averse. For another individual with vNM utility function $u(x) = x^{0.5}$ (risk averse), hence, $c_u^{un} = 0.25$. $\tau = 0.4 > 0.25$ is classified as ignorance seeking.

For two individuals with the same utility under risk, the comparison of their characteristic values tells their relative attitude toward ignorance, i.e., $\tau_1 < \tau_2$ means that individual number 1 is more averse toward ignorance than the individual number 2.

Finally, we note that min and max decision criteria are special cases of (20.9) with $\tau = 0$ and $\tau = 1$, respectively. On the other hand, (20.9) excludes Hurwicz’s α -criterion as well as the median rule. There is no nontrivial $0 < \alpha < 1$ that makes Hurwicz’s α -criterion equivalent to (20.9).

There is experimental evidence to support the model of decision making under ignorance based on to the τ -anchor utility. In one of few works dedicated specifically to find evidence about human decision making under ignorance, Hogarth and Kunreuther (1995) designed series of experiments on human subjects to understand the difference between decision making under ignorance and that when probabilities and consequences are known. They found that under ignorance, a significant number of subjects would switch to “meta-strategy” and use some higher-order rules to solve the choice problem. An example of such high order rules is “buy warranty to for the peace of mind” or “I would regret not buying the warranty should a breakdown occur.” Hogarth and Kunreuther observed that “An important feature of a meta-strategy is that, although the use of the strategy is triggered by the stimulus encountered, it is not responsive to detailed features of the stimulus.” This observation fits well with τ -anchor utility model. In our model, τ represents the “peace of mind” level of an individual. Although the certainty equivalence of a decision under τ -anchor utility depends on its prizes, it is quite stable when the prizes vary.

In Fig. 20.3 is the plot of the τ -utility function with $\tau = 0.4$. The coordinates on XY surface 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) 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 (τ). The origami-like surface is continuous but not differentiable at the folding lines.

In literature, a result that closest to (20.9) was established by Nehring and Puppe (1996). They derived the functional form from two axioms: a continuity axiom (C) and a strong independence axiom (SI) and other technical assumptions. Strong independence condition requires that adding a new prize that was not included in either of two sets of prizes would not change the preference between the sets.

$$\forall A, B \in \mathcal{F}(\mathcal{O}), x \in \mathcal{O}, x \notin A, x \notin B, \text{ if } A \succeq_{NP} B \Rightarrow A \cup x \succeq_{NP} B \cup x. \quad (20.11)$$

SI condition is often contrasted with the condition named Independence (I) which requires that preference between two sets made by adding two new prizes to a set is the same as the preference between added prizes. $A \in \mathcal{F}(\mathcal{O}), x, y \in \mathcal{O}$ and $x, y \notin A$, if $x \geq y$ then $A \cup \{x\} \succeq A \cup \{y\}$. (SI) is stronger than (I). For example both Hurwicz’s α -criterion with $\alpha \neq 0, 1$ and the median rule satisfy (I) but not (SI). Theorem 2, on the other hand, is based on imposing the law of iterated certainty equivalence on HA preference. Taken together, Nehring–Puppe’s result and Theorem 2 provide independent supports for the conceptual foundation of τ -anchor utility function.

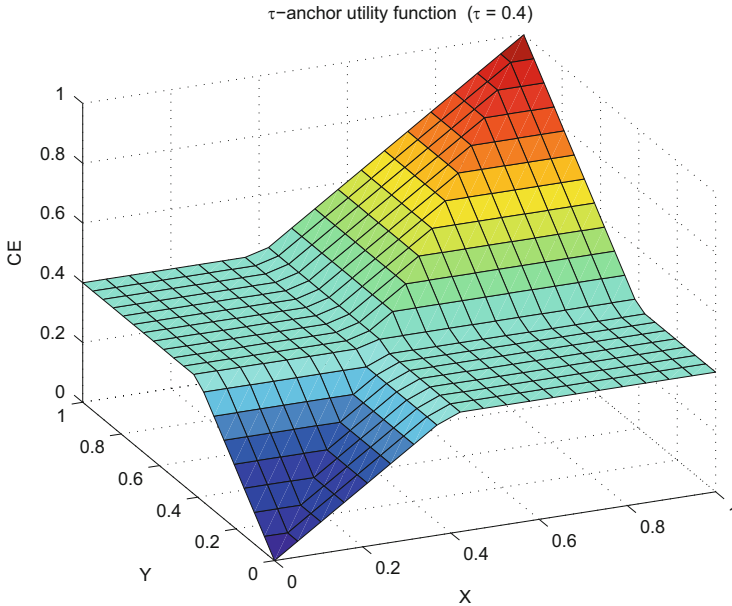


Fig. 20.3 τ -Anchor utility function

20.4 Related Literature and Discussion

On this chapter the topic of decision in the condition of ignorance has been examined mostly from a mathematical and economic point of view. Understanding of the topic would not be possible without contributions from researchers in behavioral and neurobiological fields. Hogarth and Kunreuther (1995) argued that many practical situations where individuals have to make decision lack the basic features of a gamble, namely, the outcomes and probabilities. For example, an individual decides whether to buy warranty for an electronic device without knowing probability of its breakdown and the repair cost. They designed a series of experiments on human subjects to understand the difference between behavior under probability and that when no probability or cost are known. They found clear evidence for difference of behaviors when the subjects have and do not have probability (ignorance). They found that under ignorance the subjects use two types of strategies to arrive at a choice. One of those is using 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 hold a different opinion. We think that the observed behavior is perfectly rational. Under ignorance, there is no reliable information as input for one to think harder. Thinking harder in this case often

means filling the void left by lack of hard evidence with personal analytic or judgmental assumptions which may be false. The casual practice of acting upon those assumptions as if they were facts is misleading and sometime dangerous. As it was argued in the previous section, the τ -anchor utility fits the description of this type of strategy. For example, τ can be interpreted as “peace of mind” level for a subject.

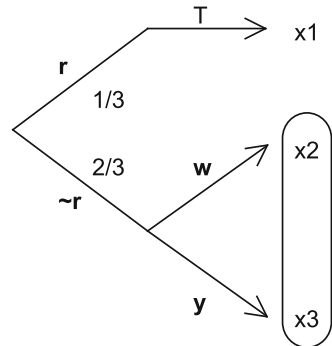
Pushkarskaya et al. (2010) examined the neurological evidence of decision under ignorance using fMRI scan. They consider two types of missing information (MI): ambiguity (vague probabilities) and SSI. They found that different types of MI activate distinct neural substrates in the brain. A popular view held by neuroscientists, the reductive viewpoint, suggests that individuals reduce SSI (extreme uncertainty) to ambiguity and then to risk by forming subjective beliefs, about the partition of a sample space and then about the corresponding probability. However, the prediction of the reductive view holds only for ambiguity averse individuals and not for ambiguity-tolerant individuals. They concluded with a key suggestion that theories of decision making under uncertainty should include individual tolerance for missing information. The characteristic value that plays a central role of in the utility theory in Sect. 20.3 provides an answer to this challenge.

Finally, we illustrate with an example based on Ellsberg’s paradox (Ellsberg 1961) on how ignorance and probability can be used together.

Example 3 (Ellsberg’s Urn). 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 α where α can be r, w, y or their negations, pays \$1 if α holds and nothing otherwise.

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. 20.4).

Fig. 20.4 Ellsberg’s urn



Let assume that an individual has a risk averse (concave) utility function $u(x) = \sqrt{x}$ and characteristic value under ignorance $\tau = 0.20$. Note that under the uniform distribution and with utility function $u(x) = \sqrt{x}$, the lottery (0.5:0, 0.5:1) has the certainty equivalent of 0.25. So for a uncertainty averse individual, her characteristic value under ignorance must be $\tau < 0.25$. For a bet on red, 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 $1/3 * \sqrt{1} + 2/3 * \sqrt{0} = 1/3$. The certainty equivalent 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 0. If the ball 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 $(w \vee y)$ 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 $(r \vee y)$ 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 $(w \vee y) \succ (r \vee y)$. ■

20.5 Conclusion

Our goal in this chapter is to convince readers that the state of ignorance does occur in the real world, almost always within a bigger problem. A careful analysis and isolation of ignorance in the system of knowledge about a subject or a problem is of particular importance in the context of risk assessment and risk management. We argue that the practice of casually papering over the ignorance with subjective judgments and analytic assumptions can have serious consequences. We provide a structured survey (and necessarily selective) of significant ideas and proposals for decision making under ignorance, from the ground breaking work by Hurwicz and Arrow to the latest result of τ -anchor utility theory.

References

- Arrow KJ, Hurwicz L (1977) An optimality criterion for decision making under ignorance. In: Arrow KJ, Hurwicz L (eds) Studies in resource allocation processes. Cambridge University Press, Cambridge
- Aven T, Steen R (2010) The concept of ignorance in a risk assessment and risk management context. Reliab Eng Syst Saf 95:1117–1122
- Barbera S, Bossert W, Pattanaik PK (2004) Ranking sets of objects. In: Barbera S, Hammond P, Seidl C (eds) Handbook of utility theory: extensions, vol 2. Springer, Berlin, pp 893–978
- Cohen M, Jaffray J-Y (1980) Rational behavior under complete ignorance. Econometrica 48(5):1281–1299

- Congar R, Maniquet F (2010) A trichotomy of attitudes for decision-making under complete ignorance. *Math Soc Sci* 59:15–25
- Ellsberg D (1961) Risk, ambiguity and the Savage's axioms. *Q J Econ* 75(4):643–669
- Epstein LG (1999) A definition of uncertainty aversion. *Rev Econ Stud* 66:579–608
- Giang PH (2011) Dynamic consistency and decision making under vacuous belief. In: Cozman FG, Pfeiffer A (eds) *Uncertainty in artificial intelligence: proceedings of the 27nd conference (UAI–2011)*. The Association for Uncertainty in Artificial Intelligence Press, pp 230–237
- Gilboa I, Postlewaite A, Schmeidler D (2009) Is it always rational to satisfy Savage's axioms? *Econ Philos* 25:285–296
- Graham DA (2014) Rumsfeld's knowns and unknowns: the intellectual history of a quip. *Atl Mag*. <http://www.theatlantic.com/politics/archive/2014/03/rumsfelds-knowns-and-unknowns-the-intellectual-history-of-a-quip/359719/>
- Gravel N, Marchant T, Sen A (2012) Uniform expected utility criteria for decision making under ignorance or objective ambiguity. *J Math Psychol* 56:297–315
- Health and Safety Executive (UK) (2001) *Reducing risks, protecting people: HSE's decision-making process*. HSE Books, London
- Hogarth RM, Kunreuther H (1995) Decision making under ignorance: arguing with yourself. *J Risk Uncertain* 10:15–36
- Keynes JM (1921) *A treatise on probability*. Macmillan, London
- Knight FH (1921) *Risk, uncertainty and profit*. Houghton Mifflin, Boston
- Maskin E (1979) Decision-making under ignorance with implications for social choice. *Theor Decis* 11:319–337
- Nehring K, Puppe C (1996) Continuous extensions of an order on a set to the power set. *J Econ Theory* 68:456–479
- Nitzan S, Pattanaik PK (1984) Median-based extensions of an ordering over a set to the power set: an axiomatic characterization. *J Econ Theory* 34:252–261
- Puppe C, Schlag KH (2009) Choice under complete uncertainty when outcome spaces are state dependent. *Theor Decis* 66:1–16
- Pushkarskaya H, Liu X, Smithson M, Joseph JE (2010) Beyond risk and ambiguity: deciding under ignorance. *Cogn Affect Behav Neurosci* 10(3):382–391
- Shackle GLS (1949) *Expectations in economics*. Cambridge University Press, Cambridge
- Smets P (2005) Decision making in the TBM: the necessity of the pignistic transformation. *Int J Approx Reason* 38(2):133–147
- von Neumann J, Morgenstern O (1953) *Theory of games and economic behavior*, 3rd edn. Princeton University Press, Princeton, NJ
- Walley P (1996) Inferences from multinomial data: learning about a bag of marbles. *J R Stat Soc Ser B* 58(1):3–57