Decision-making under partial information: a smooth approach

Illia Pasichnichenko*
Alfred-Weber-Institute for Economics, Heidelberg University

July 14, 2019

Abstract

Belief functions, or normalized totally monotone capacities, are used in decision modelling to represent partial information and subjective judgements. In the paper, we propose a model of preferences among belief functions, which allows for smooth indifference curves and better separation of tastes and beliefs, than the Choquet integral. In our model, a generalised average is used instead of maximums and minimums. Thus, familiar methods from decision analysis under risk can be applied for studying decision maker's tastes.

1 Introduction

Belief functions introduced by Dempster [5] and Shafer [14] provide a convenient method for modelling partial or imprecise information as well as subjective judgements. For example, suppose that there is a probability $m(B)$ of receiving a message that the true state belongs to B. Then the probability of A containing the true state is at least $f(A) = \sum_{B \subseteq A} m(B)$. Here f is a set function possessing weaker properties than a probability measure. Indeed, f may not be additive, i.e. $f(A_1 \cup A_2) > f(A_1) + f(A_2)$ for some disjoint A_1 and A_2. It is possible, when there exists B, such that $B \subseteq A_1 \cup A_2$ and $m(B) > 0$, but B is a subset of neither A_1 nor A_2. Function f is called a belief function or a lower probability.

Consider an example of partial information, the Ellsberg paradox [6]. Here the decision maker gets different prizes depending on the color of the ball randomly drawn from an urn. The urn contains 30 red balls and 60 black and yellow balls, the later in unknown proportion. Possible bets and corresponding prizes are given by the table.

	R	B	Y
a	\$100	0	0
b	0	\$100	0
a'	\$100	0	\$100
b'	0	\$100	\$100

In this example, there are three possible sets of outcomes, namely $\{0\}$, $\{100\}$, and $\{0, 100\}$. When the decision maker evaluates bet a, he has complete information about chances to win. This is reflected in that all probability is assigned to singletons, namely

*Funding for this research has been provided by the Alexander von Humboldt Foundation.

$m_a(100) = 1/3$,[1] $m_a(0) = 2/3$, and $m_a(\{0, 100\}) = 0$. In this case, the corresponding belief function f_a is a probability measure. To the contrary, for bet b there is a chance $m_b(0) = 1/3$ of receiving nothing (when the drawn ball is red) and a chance $m_b(\{0, 100\}) = 2/3$ of getting 0 or \$100. The later is an example of partial information: probability 2/3 can not be divided between 0 and \$100, because the proportion of black to yellow balls is unknown. Corresponding belief function f_b is not additive. We summarize the chances in the next two tables.

	0	100	{0,100}
m_a	2/3	1/3	0
m_b	1/3	0	2/3
$m_{a'}$	0	1/3	2/3
$m_{b'}$	1/3	2/3	0

	0	100	{0,100}
f_a	2/3	1/3	1
f_b	1/3	0	1
$f_{a'}$	0	1/3	1
$f_{b'}$	1/3	2/3	1

While in this example we used belief functions to model partial information, they can also represent subjective judgements about possibilities of different outcomes. Indeed, belief functions are possibly better known in economics as a special case of nonadditive probabilities, or normalized capacities, that can be derived from decision maker's preferences over acts and reflect subjective beliefs about the unknown.

Such decision models also explain how nonadditive probabilities are used by the decision maker in evaluating decisions. For example, a model axiomatized by Schmeidler [13] is based on the idea of integration with respect to a non-additive set function. In particular, applying such (Choquet) integral for evaluation of bet a we get[2]

$$V(a) = m_a(0)u(0) + m_a(100)u(100) + m_a(\{0, 100\}) \min\{u(0), u(100)\}, \tag{1}$$

where u is a vNM utility function. Assuming $u(0) = 0$ and $u(100) = 1$, we get $V(a) = 1/3$, the expected utility of a. However, the value of bet b is only 0, because most of the probability mass, 2/3, is multiplied by $\min\{u(0), u(100)\} = 0$. This shows a pessimistic attitude incorporated in such evaluation. That is, if with probability $m(B)$ the true outcome is in B, then the decision maker acts as if with probability $m(B)$ he obtains the worst outcome in B. If the decision maker actually values b higher, we would have to admit that he has subjective judgement $m_b(100) > 0$. If at the same time he reveals that his lower estimate of observing black ball is still 0, we would have a contradiction. This happens because such form of evaluation is not flexible enough to encompass explicitly different attitudes other than extreme pessimism, thus pushing attitudes inside beliefs.

The problem is solved in the literature (see, for example, [8]) by replacing $m(B)u_*$ by $m(B)v(u_*, u^*)$, where u_* and u^* are the minimal and the maximal utility levels in B respectively. For example, when $v(u_*, u^*) = \alpha u_* + (1 - \alpha)u^*$, then $0 \le \alpha \le 1$ is called the pessimism index. An $\alpha < 1$ allows for both $V(b) > 0$ and $m_b(100) = 0$.

On the other hand, it seems possible that intermediate outcomes also influence the evaluation of B. Suppose that some black and yellow balls are replaced by the equal number of green balls. In case of choosing b and drawing a green ball, the decision maker wins $0 < x < 100$. According to the models from the previous paragraph, the value of b does not change, when x is changing. Indeed, assuming increasing u and omitting zero terms,

$$V(b) = m_b(\{0, x, 100\})v(0, 1),$$

[1] Abusing notation, we write $m_a(100)$ instead $m_a(\{100\})$.

[2] Using the representation of the Choquet integral as an average of minimums (see [7]).

which does not depend on x. However, some individuals may find bet b more attractive, when the value of x is higher. Also the existence of intermediate prizes in real-world lotteries suggests that such prizes may have value for buyers. This paper aims at studying such attitude.

We assume that decision maker's evaluation of B is better captured by $m(B)\Phi(B)$, where $\Phi(B)$ is a quasi-arithmetic mean[3] of utilities of elements of B. In other words,

$$\Phi(B) = \phi^{-1}\left(\frac{1}{n_B}\sum_{y\in B}\phi[u(y)]\right),$$

where ϕ is an increasing function and n_B is the number of elements in B.

Note that under risk, i.e. when $m > 0$ only for singletons, the model reduces to expected utility, because $\Phi(\{y\}) = u(y)$.

Function ϕ represents decision maker's tastes for situations of complete ignorance limited to B. The more concave is ϕ, the closer $\Phi(B)$ approaches u_*, representing a pessimistic attitude when facing unknown chances. To the contrary, convex ϕ represents optimistic attitude, pushing $\Phi(B)$ closer to u^*. In fact, we use here a familiar mechanism from decision analysis under risk.

In particular, the value of bet b in the last example is

$$V(b) = m_b(\{0, x, 100\})\Phi(\{0, x, 100\}),$$

where

$$\Phi(\{0, x, 100\}) = \phi^{-1}\left[\frac{1}{3}\phi(0) + \frac{1}{3}\phi(u(x)) + \frac{1}{3}\phi(1)\right].$$

The more concave is ϕ, the closer $\Phi(\{0, x, 100\})$ is to 0. Also note that $\Phi(\{0, x, 100\})$ is monotone in x.

In the next section we derive the decision-making criterion from preferences among belief functions. That is, we assume belief functions to be given and do not derive them from preferences among acts. In particular, such framework is suitable for situations, in which the information is partial and objective. Thus, this paper also belongs to the literature on exogenous ambiguity (see, for example, [2]).

2 Model

Suppose X is the set of lotteries over some set of consequences. We refer to elements of X as outcomes. Let $\mathcal{A}$ be an algebra of subsets of X containing all finite subsets.

A mapping $f : \mathcal{A} \to [0, 1]$ is a belief function on $\mathcal{A}$, when $f(\emptyset) = 0$, $f(X) = 1$, and f is monotone at all orders (for details see [4]). For example, monotonicity at order 2 is convexity, i.e. $f(A_1 \cup A_2) \geq f(A_1) + f(A_2) - f(A_1 \cap A_2)$.

Let $\mathcal{F}$ be the set of belief functions on $\mathcal{A}$ concentrated on a finite subset, i.e. $f(\mathcal{D}_f) = 1$ for some finite $\mathcal{D}_f$. Note that $\mathcal{F}$ is a mixture space, where $\lambda f + (1 - \lambda)g$ is interpreted as a lottery over two belief functions. We assume decision maker's preference relation on $\mathcal{F}$.

Note that each belief function in $\mathcal{F}$ can be represented as a convex combination of elementary belief functions $e_B \in \mathcal{F}$, where $B \in \mathcal{A}$, $e_B(A) = 1$ if $A \supseteq B$ and $e_B(A) = 0$ otherwise. Indeed, each belief function $f \in \mathcal{F}$ possess a Moebius inverse that is a mapping

[3] Quasi-arithmetic mean (or Kolmogorov mean) was axiomatically characterized by Kolmogorov [11]. In economics, it was used, for example, for evaluation of decision trees [9].

$m : \mathcal{A} \to [0,1]$ such that $m(A) > 0$ only for a finite number of subsets A, $m(\emptyset) = 0$, $\sum_{B\in\mathcal{A}} m(B) = 1$, and $f(A) = \sum_{B\subseteq A} m(B)$. We can rewrite the last equation,

$$f(A) = \sum_{B\in\mathcal{A}} m(B) e_B(A),$$

or simply $f = \sum_{B\in\mathcal{A}} m(B)e_B$. If V is a linear utility function on $\mathcal{F}$, then

$$V(f) = \sum_{B\in\mathcal{A}} m(B)V(e_B). \tag{2}$$

An elementary belief function e_B represents the situation of complete ignorance limited to B. In other words, the decision maker is sure that the true outcome belongs to B, but nothing more. In order to specify $V(e_B)$ in equation (2), additional assumptions should be made about decision maker's evaluation of such situations. For example, Jaffray [8] assumes that $V(e_B)$ depends only on the worst and the best outcomes in B[4]. To the contrary, we suppose that $V(e_B)$ satisfies the following equation

$$V(e_B) = \phi^{-1}\left(\frac{1}{n_B}\sum_{x\in B}\phi[u(x)]\right)$$

where ϕ is a continuous increasing function defined on the values of vNM utility function u, and n_B is the number of elements in B. We make here a standard assumption of independence of risk attitudes and ambiguity attitudes. In other words, first decision maker's risk preferences result in a particular list of vNM utilities for B, which are then averaged in a subjective scale ϕ.

To derive such criterion from decision maker's preferences on $\mathcal{F}$, we make the following assumptions. First, we assume that decision maker's preferences on $\mathcal{F}$ satisfy axioms of linear utility theory. Second, we assume reduction of a compound lottery[5], i.e.

$$\lambda e_{\{x\}} + (1-\lambda)e_{\{y\}} \sim e_{\{\lambda x + (1-\lambda)y\}}$$

for each x and y in X. This allows us to define a linear utility function on X[6]. Third, we assume that set X is rich, i.e. for each $x \in X$ there are countably many distinct outcomes in X that are indifferent to x.

Finally, we introduce a set of assumptions concerning decision maker's preferences among e_B, $B \in \mathcal{A}$. In what follows, for A and B in $\mathcal{A}$ we write $A \succ B$ instead of $e_A \succ e_B$. If $x \in A$ and $y \in X$, we denote $A_x(y)$ the set constructed from A by replacing x by y.

Axiom 1 (Monotonicity). If $y \succcurlyeq x$, then $A_x(y) \succcurlyeq A$.

That is, replacing one of the possible outcomes by a weakly preferred one does not make the situation of complete ignorance worse.

Axiom 2 (Set Betweenness). If $A \succcurlyeq B$ and $A \cap B = \emptyset$, then $A \succcurlyeq A \cup B \succcurlyeq B$.

A failure of Set Betweenness, say $A \cup B \prec B$, implies that adding better outcomes to B makes it less attractive. This may be true, if the decision maker is averse to the size of the set of possible outcomes.

---

[4] According to the natural ordering of $x \in X$ induced by decision maker's preferences among $e_{\{x\}}$.

[5] Recall that X is the set of lotteries.

[6] If V is a linear utility function on $\mathcal{F}$, then u defined by $u(x) = V(e_{\{x\}})$ is a linear utility function on X.

Axiom 3 (Set Continuity). For any $x \in A$ and $y, z \in X$ sets $\{0 \leq \lambda \leq 1 : A_x(\lambda y + (1-\lambda)z) \succcurlyeq A\}$ and $\{0 \leq \lambda \leq 1 : A_x(\lambda y + (1-\lambda)z) \preccurlyeq A\}$ are closed.

We call $\tau_A \in X$ a risky equivalent of $A \in \mathcal{A}$, if $\{\tau_A\} \sim A$. In the proof of the theorem we show that for some λ the lottery $\lambda x_* + (1-\lambda)x^*$ is a risky equivalent of a set A, where x_* and x^* are the worst and the best outcomes in A. Thus, for a finite A a risky equivalent always exists. We use this in the following axiom.

Axiom 4 (Partitioning). If $x_1, \ldots, x_{2n} \in X$, then

$$\{x_1, \ldots, x_{2n}\} \sim \{\tau_{\{x_1,\ldots,x_n\}}, \tau_{\{x_{n+1},\ldots,x_{2n}\}}\}.$$

Partitioning implies that the evaluation of $\{x_1, \ldots, x_{2n}\}$ can be done in three steps. That is, the evaluation of $\{x_1, \ldots, x_n\}$, $\{x_{n+1}, \ldots, x_{2n}\}$, and the resulting two-element set of risky equivalents. Such reduction seems reasonable, especially when $2n$ is large. However, this assumption may be a subject of experimental testing.

Now we are ready to state an existence theorem.

Theorem 1. *There exists a utility function V on $\mathcal{F}$ such that*

$$V(f) = \sum_{B \in \mathcal{A}} m(B)\, \phi^{-1}\left(\frac{1}{n_B}\sum_{x \in B}\phi[u(x)]\right),$$

where u is a linear utility function on X and ϕ is a continuous strictly increasing function defined on the values of u. Such V is unique up to a positive linear transformation. Given that V is fixed, u is unique and ϕ is unique up to a positive linear transformation.

It is straightforward to show that the assumptions (with the exception of richness of X) are also necessary for the representation. The proof of the theorem is given in section 4.

3 Discussion

Let us return to the original Ellsberg paradox. In our model, the value of a bet a is given by the formula

$$V(a) = m_a(0)u(0) + m_a(100)u(100) + m_a(\{0, 100\})\phi^{-1}\left(\tfrac{1}{2}\phi[u(0)] + \tfrac{1}{2}\phi[u(100)]\right).$$

Assuming $u(0) = 0$ and $u(100) = 1$, as we did before, and also $\phi(0) = 0$ and $\phi(1) = 1$, we get

$$V(a) = m_a(100) + m_a(\{0, 100\})\phi^{-1}\left(\tfrac{1}{2}\right).$$

Values of different bets are given in the following table. For comparison, we also give values obtained from formula (1).

	Our model	Choquet integral
a	$\frac{1}{3}$	$\frac{1}{3}$
b	$\frac{2}{3}\phi^{-1}\left(\frac{1}{2}\right)$	0
a'	$\frac{1}{3} + \frac{2}{3}\phi^{-1}\left(\frac{1}{2}\right)$	$\frac{1}{3}$
b'	$\frac{2}{3}$	$\frac{2}{3}$

If ϕ is concave between 0 and 1, then $\phi^{-1}\left(\frac{1}{2}\right) < \frac{1}{2}$, implying $a \succ b$ and $a' \prec b'$. In the literature, such preferences are linked with ambiguity aversion. If ϕ is convex, then the opposite relations hold. Finally, when ϕ is linear, the decision maker is indifferent between a and b, and between a' and b'.

On the other hand, it is not possible to obtain preferences $a \succ b$ and $a' \succ b'$ just by varying ϕ and holding belief functions fixed. This happens because ϕ only influences the evaluation of the "ignorance" component represented by the probability mass $m(\{0, 100\})$. This component is same in b and a', so increase in value of b due to changing the shape of ϕ leads to the equal increase in value of a'. Thus, preferences $a \succ b$ and $a' \succ b'$ imply that decision maker's belief functions are different from ones suggested in introduction.

Our model can explain Machina's [12] reflection example, but not the 50:51 example.

Estimation of ϕ can be done similarly to estimation of the vNM utility function (see, for example, [3]). Indeed, fix $x_*, x^* \in X$ such that $x_* \prec x^*$, and put $\phi[u(x_*)] = 0$ and $\phi[u(x^*)] = 1$. Find a risky equivalent τ_1^1 of $\{x_*, x^*\}$ and put $\phi[u(\tau_1^1)] = \frac{1}{2}$. If τ_1^2 and τ_2^2 are risky equivalents of $\{x_*, \tau_1^1\}$ and $\{\tau_1^1, x^*\}$ respectively, then $\phi[u(\tau_1^2)] = \frac{1}{4}$ and $\phi[u(\tau_2^2)] = \frac{3}{4}$. Continuing in the same fashion, we obtain an estimate of ϕ between $u(x_*)$ and $u(x^*)$ with a possibility of doing consistency checks. In theory, risky equivalents $\tau_1^1, \tau_1^2, \tau_2^2, \ldots$ exist (see section 4).

If ϕ is smooth, then indifference curves in our model are also smooth. This rises comparison to [10]. Analogously, a theory of comparison of ambiguity attitudes can be developed basing on the curvature of ϕ.

4 Proof of the theorem

Let V be a linear utility function on $\mathcal{F}$ and u be defined by $u(x) = V(e_{\{x\}})$ for each $x \in X$. Denote by I the set of values of u. Since u is a linear utility function on X, I is an (open, closed, half-closed, finite or infinite) real interval of positive length (omitting the trivial case of complete indifference). Define $\Phi(B) = V(e_B)$. We only have to show that there exists a continuous strictly increasing function $\phi : I \to \mathbb{R}$ such that

$$\Phi(B) = \phi^{-1}\left(\frac{1}{n_B}\sum_{x \in B}\phi[u(x)]\right) \tag{3}$$

holds for any finite $B \in \mathcal{A}$. To do this, we first prove (3) for two-element sets and then generalize the result to arbitrary finite sets.

Note that $\Phi(\{x_1, x_2\})$ depends only on $u(x_1)$ and $u(x_2)$. Indeed, if $u(x_1) = u(y_1)$ and $u(x_2) = u(y_2)$, then $\{x_1, x_2\} \sim \{y_1, y_2\}$ by Monotonicity, therefore $\Phi(\{x_1, x_2\}) = \Phi(\{y_1, y_2\})$. Hence,

$$\Phi(\{x_1, x_2\}) = M(u(x_1), u(x_2))$$

for some function $M : I^2 \to \mathbb{R}$. In what follows we study properties of this function.

For each $r \in I$ by richness of X there exist $x \neq y$ such that $u(x) = u(y) = r$. Since $x \sim y$, Set Betweenness implies $\{x\} \sim \{x, y\}$. Then

$$M(r, r) = M(u(x), u(y)) = V(\{x, y\}) = u(x) = r.$$

Let $r_1, r_2, s \in I$ and $r_1 < r_2$. Take different $x_1, x_2, y \in X$ such that $u(x_1) = r_1$, $u(x_2) = r_2$ and $u(y) = s$. By Monotonicity $\{x_1, y\} \prec \{x_2, y\}$, thus $M(r_1, s) < M(r_2, s)$. Therefore, M is strictly increasing in both variables.

For the next step, we have to show first that for any finite $B \in \mathcal{A}$ there exists a risky equivalent τ_B. Indeed, if $x^*, x_* \in B$ and $x^* \succcurlyeq x \succcurlyeq x_*$ for all $x \in B$, then $\{x^*\} \succcurlyeq B \succcurlyeq \{x_*\}$ by Set Betweenness. We can find $0 \leq \lambda \leq 1$ such that $V(B) = \lambda u(x^*) + (1-\lambda)u(x_*)$, therefore $\tau_B = \lambda x^* + (1-\lambda)x_*$ is a risky equivalent of B.

Now we would like to prove that

$$M\left(M(r,s), M(t,k)\right) = M\left(M(r,t), M(s,k)\right) \tag{4}$$

for arbitrary $r, s, t, k \in I$. To do this, take different $x, y, z, w \in X$ such that $u(x) = r$, $u(y) = s$ etc. Since $M(r,s) = u\left(\tau_{\{x,y\}}\right)$, we have

$$M\left(M(r,s), M(t,k)\right) = M\left(u\left(\tau_{\{x,y\}}\right), u\left(\tau_{\{z,w\}}\right)\right).$$

Partitioning implies that $\{\tau_{\{x,y\}}, \tau_{\{z,w\}}\} \sim \{x,y,z,w\} \sim \{\tau_{\{x,z\}}, \tau_{\{y,w\}}\}$, which leads to

$$M\left(u\left(\tau_{\{x,y\}}\right), u\left(\tau_{\{z,w\}}\right)\right) = M\left(u\left(\tau_{\{x,z\}}\right), u\left(\tau_{\{y,w\}}\right)\right)$$

from which (4) follows.

The fact that M is continuous follows from Set Continuity.

So far we proved that M is continuous and strictly increasing in both variables, satisfies (4) and $M(r,r) = r$ for all $r \in I$. According to the theorem characterizing the quasi-arithmetic mean (see [1], p. 287), M satisfies these conditions if and only if there exists a continuous and strictly monotonic function $\phi : I \to \mathbb{R}$ with which

$$M(r,s) = \phi^{-1}\left(\tfrac{1}{2}\phi(r) + \tfrac{1}{2}\phi(s)\right) \tag{5}$$

holds for each $r, s \in I$. If ϕ is decreasing, then $-\phi$ is a continuous strictly increasing function satisfying (5). Thus, we proved (3) for two-element sets.

Now we extend (3) to an arbitrary finite $B \in \mathcal{A}$. Suppose that (3) is true for all sets with no more than n elements, $n \geq 2$, and prove (3) for a set $B = \{x_1, \ldots, x_{n+1}\}$. If $y_1, \ldots, y_{n-1}$ are different risky equivalents of B, then by Set Betweenness

$$B \sim \{x_1, \ldots, x_{n+1}, y_1, \ldots, y_{n-1}\}.$$

Note that there are $2n$ elements in the later set. By Partitioning we have

$$B \sim \left\{\tau_{\{x_1,\ldots,x_n\}}, \tau_{\{x_{n+1}, y_1 \ldots, y_{n-1}\}}\right\}$$

The later set has only two elements, so

$$\phi\left(\Phi\left(B\right)\right) = \tfrac{1}{2}\phi\left(u\left(\tau_{\{x_1,\ldots,x_n\}}\right)\right) + \tfrac{1}{2}\phi\left(u\left(\tau_{\{x_{n+1}, y_1 \ldots, y_{n-1}\}}\right)\right).$$

Since

$$\phi\left(u\left(\tau_{\{x_1,\ldots,x_n\}}\right)\right) = \phi\left(\Phi\left(\{x_1, \ldots, x_n\}\right)\right) = \frac{1}{n}\sum_{i=1}^{n}\phi\left(u(x_i)\right)$$

and similarly

$$\phi\left(u\left(\tau_{\{x_{n+1}, y_1 \ldots, y_{n-1}\}}\right)\right) = \frac{1}{n}\phi\left(u(x_{n+1})\right) + \frac{1}{n}\sum_{i=1}^{n-1}\phi\left(u(y_i)\right),$$

we have

$$\phi\left(\Phi\left(B\right)\right) = \frac{1}{2n}\sum_{i=1}^{n+1}\phi\left(u(x_i)\right) + \frac{n-1}{2n}\phi\left(\Phi(B)\right),$$

which implies

$$\phi\left(\Phi\left(B\right)\right) = \frac{1}{n+1}\sum_{i=1}^{n+1}\phi\left(u(x_i)\right).$$

The uniqueness results are straightforward. The theorem is proved.

References

[1] J Aczél and J Dhombres. *Functional equations in several variables.* Cambridge university press, 1989.

[2] D S Ahn. Ambiguity without a state space. *The Review of Economic Studies*, 75(1):3–28, 2008.

[3] G M Becker, M H Degroot, and J Marschak. Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232, 1964.

[4] A Chateauneuf and J Y Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17(3):263–283, 1989.

[5] A P Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38(2):325–339, 1967.

[6] D Ellsberg. Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, 75(4):643–669, 1961.

[7] I Gilboa and D Schmeidler. Additive representations of non-additive measures and the choquet integral. *Annals of Operations Research*, 52(1):43–65, 1994.

[8] J Y Jaffray. Linear utility theory for belief functions. *Operation Research Letters*, 8(2):107–112, 1989.

[9] S Ke. Boundedly rational backward induction. *Theoretical Economics*, 14(1):103–134, 2019.

[10] P Klibanoff, M Marinacci, and S Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892, 2005.

[11] A N Kolmogorov. On the notion of mean. In *Selected Works of A.N. Kolmogorov: Volume I: Mathematics and Mechanics (1991)*, pages 144–146. Springer Science & Business Media, 1930.

[12] M J Machina. Risk, ambiguity, and the rank-dependence axioms. *American Economic Review*, 99(1):385–92, 2009.

[13] D Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[14] G Shafer. *A Mathematical Theory of Evidence.* Princeton Univ. Press, 1976.