A Procedural Model of Complexity Under Risk[*†]

En Hua Hu
University of Toronto

October 30, 2023

Abstract

I consider a decision-maker who uses rules to simplify lotteries in order to compare them. I characterize expected utility in this setting and highlight its complexity requirements which a purely axiomatic characterization overlooks. I relax these requirements to characterize two models of complexity aversion: outcome support size cost and entropy cost models. I consider an additional aspect of complexity: decision-makers find it easier to evaluate a lottery when outcomes are close in value. To capture this, I characterize a third model of complexity aversion. Here the DM first partitions together outcomes that are close in value and then evaluates the lottery along with the complexity of the partition. This representation offers a measure of complexity that is not restricted to the probability and support size but also accounts for the cardinal values of the outcomes. I also compare empirically the models and find support for partition complexity.

keywords: complexity aversion, procedural choice, non-expected utility

JEL codes: D8, D91

[*]I am indebted to Yoram Halevy and Colin Stewart for their excellent supervision and guidance. I also thank Alexis Ghersengorin, Itzhak Gilboa, Sonia Mun, Kirby Nielsen, Ryan Oprea, Jean-Marc Tallon, Martin Vaeth, and the participants at the PSE-SciencePo PhD seminars, IESEG brown bag, and Toronto PhD reading group for discussion and feedback.

[†]This is a very early draft version, so I greatly appreciate feedback on all aspects and especially potential mistakes at enhua.hu@mail.utoronto.ca

1 Introduction

There is now sufficient evidence to conclude that decision-makers often do not behave according to expected utility theory[1]. Under the axiomatic decision theoretic paradigm, this can only be reconciled by relaxing the theory's axioms. However, the axioms of expected utility, even though violated, carry with them strong intuition and appeal. Certainly, as a student who first learned of mixture independence, I, and no doubt many others, accepted it as truth, and certainly, von Neumann and Morgenstern aimed to pick believable axioms. This conflict is further evidenced in experimental works where subjects state agreement with the axioms only to go on to violate them when decisions become too complex[2]. In conjunction, it is also well documented that subjects prefer simple lotteries to complex ones (Huck and Weizsäcker, 1999; Kovářík et al., 2016). The goal of this work is to study the relationship between complexity and non-expected utility behavior via a procedural model. In particular, I model violations of expected utility which are due to complexity concerns, rather than a conscious rejection of the axioms.

The paper studies decision-making under risk where the objects of choice are lotteries with monetary outcomes. I offer a procedural model where a decision-maker uses rules to simplify lotteries before comparing them. The key rule, *simplification*, allows the decision-maker to merge two outcomes of the lottery, along with their probabilities, into a single outcome. For example, the lottery $(\$4, 20\%; \$5, 20\%; \$10, 60\%)$ may be simplified to $(\$4.5, 40\%; \$10, 60\%)$. I first characterize the expected utility in this model and highlight the theory's complexity assumptions. In particular, the theory assumes that simplification is a costless operation and does not discriminate between different types of simplifications. Relaxing the first assumption and stipulating that simplifications may be costly allows me to obtain the following model of complexity aversion:

- Support Size Cost: there exists a continuous and strictly increasing utility function u and a weakly increasing cost function C such that for any lottery p and q:

$$p \succeq q \Leftrightarrow \sum_{x \in supp(p)} u(x)p(x) - C(|supp(p)|) \geq \sum_{x \in supp(q)} u(x)q(x) - C(|supp(q)|)$$

I then further consider the possibility that the decision-maker may discriminate simplifications by the proportion of the two outcomes that were merged. This yields the following representation:

- Entropy Cost: there exists a continuous and strictly increasing utility function u and $\gamma \in [0, \infty)$ such that for any lottery p and q,

$$p \succeq q \Leftrightarrow \sum_{x \in supp(p)} u(x)p(x) - \gamma\mathcal{H}(p) \geq \sum_{x \in supp(q)} u(x)q(x) - \gamma\mathcal{H}(q)$$

Where $\mathcal{H}$ denotes the Shannon entropy of a lottery.

[1] Tversky (1969); Kahneman and Tversky (1979)

[2] See Nielsen and Rehbeck (2021); MacCrimmon and Larsson (1979); Moskowitz (1974); Slovic and Tversky (1974); MacCrimmon (1968)

While these two models have recently been characterized in purely axiomatic frameworks by Puri (2020) and Mononen (2021), respectively, my characterization in a procedural model highlights how the DM's attitude towards rules reveals what she considers complex. As rules are used to reduce complexity, what the decision-maker seeks to change with rules is precisely what she considers complex. Finally, I consider an additional aspect of complexity that the previous models do not capture: decision-makers find it easier to evaluate a lottery when outcomes are close in value. For instance, a decision-maker may find it easy to evaluate a lottery with outcomes of $5.99 and $6.00 with equal probability, but a lottery of $1,000 and $50 with equal probability may be much more difficult just by virtue of how different the outcomes are. With this in mind, I characterize the following model, which I call the partition complexity model:

- Partition Cost: there exists a continuous and strictly increasing utility function u, and for every lottery p and q, there exist partitions of outcomes of p, q denoted by P_p, P_q such that the cells of each partition are contained in mutually disjoint intervals and:

$$p \succeq q \Leftrightarrow \sum_{x \in supp(p)} u(x)p(x) - \mathcal{C}(|supp(P_p)|) \geq \sum_{x \in supp(q)} u(x)q(x) - \mathcal{C}(|supp(P_q)|)$$

  Where $\mathcal{C}$ is the cost function of the partition which is weakly increasing in the support size of the partition.

This representation nests the support size cost representation and additionally takes into account that decision-makers may first decide to merge outcomes that they consider close in value together and then evaluate the resulting lottery as well as its complexity afterward. From the perspective of rules, the model is characterized by distinguishing certain simplifications as easy, and others as hard. Then I consider DM who considers easy simplifications costless and always perform them before hard simplifications. Therefore again, the decision-maker's attitude towards rules reveals her measure of complexity.

This work contributes to the intersection of three fields: decision under risk, complexity aversion, and procedural choice. For decisions under risk and complexity aversion, models have been offered by Puri (2020) and Mononen (2021), my contribution here is to 1) show how these models have a procedural foundation and how they may be obtained from an intuitive modification of expected utility and 2) to provide a partition model which offers a richer measure of complexity. Decision-makers have been shown in various experiments to 1) display complexity aversion (Huck and Weizsäcker, 1999; Kovářík et al., 2016) and 2) display more violations of EU when choices are complex (Agranov and Ortoleva, 2017; Nielsen and Rehbeck, 2021). My paper offers a procedural foundation for some of these behaviors. For complexity aversion and procedural choice, it has been shown in recent experimental papers that individuals frequently use heuristics and rules to make decisions and that these may be costly (Oprea, 2020; Halevy and Mayraz, 2021). My paper makes a theoretical contribution and provides the insight that the ways rules are used may reveal what the decision-maker perceives as complex. Finally, the paper takes inspiration from Rubinstein (1988) where a procedural rule based on similarity is explored in parallel with EU violation under risk. I also take a procedural approach and whereas Rubinstein (1988) claims the decision maker eliminates similar outcomes in a comparison between lotteries, I consider the merging of similar outcomes within a lottery.

The remainder of the paper proceeds as follows. In section II, I introduce the theoretical framework, as well as the *standard procedural model* (SPM) which characterizes expected utility. In section III, I show that relaxing the complexity assumptions of the SPM yields the two models of complexity aversion. In section IV, I consider a decision-maker who partitions outcomes by their cardinal values and obtains the partition complexity model's characterization. Section V considers a cross-validation exercise on choice data between the models explored in this paper and show shows support for partition complexity.

2 Framework and Rules

2.1 Framework and Notation

Let $\mathcal{X} = (m^-, m^+)$ denote the set of monetary outcomes, where $m^-, m^+ \in \mathbb{R} \cup \{-\infty, \infty\}$, members of $\mathcal{X}$ are denoted by x, y, z. The set of *simple lotteries* are probability distributions on $\mathcal{X}$ with finite support, and the set of simple lotteries is denoted by $\mathcal{L}$ with members denoted by p, q, r. Denote also by δ_x the degenerate lottery which yields x with probability 1.

A standard notation is the mixture operation where $r = p\alpha q$ denotes the lottery which is obtained as $r = \alpha p + (1-\alpha)q$. In this paper, I introduce an additional notation, I denote by $r \doteq p\alpha q$ if $supp(p) \cap supp(q) = \emptyset$ and $r = p\alpha q$. Therefore, $\doteq$ implies that r is a mixture of p and q, two lotteries that share no common outcome.

2.2 Rules

My model of decision making takes the DM as characterized by *rules* which determine her *choices*. Rules are asymmetric binary relations on lotteries or pairs of lotteries. These are mental processes that the DM uses to simplify the evaluation of a lottery as well as the comparison between pairs of lotteries. Her choices are characterized by a standard preference relation over lotteries. My model relates explicitly to the rules and choices to produce different representations.

I begin by introducing the two rules that are central to the various models in this paper. The first is *cancellation*, which states that if two lotteries share a common outcome, and this outcome occurs with the same probability in both lotteries, then the DM is able to spot this and cancel out this common outcome.

Definition 1. *Cancellation*

$(p, q)C(p', q')$ *if* $\exists \alpha, x,\ p \doteq p'\alpha x, q \doteq q'\alpha x$

In other words, p, q both have exactly $1-\alpha$ proportion of the outcome x, so the DM can cancel it out to obtain p', q'. Therefore, cancellation is a rule that the DM uses to simplify the comparison of lotteries. I highlight here that she can only cancel out the entire proportion of an outcome, therefore cancellation is strictly weaker than mixture independence over outcomes which would be defined with $=$ instead of $\doteq$.

The next rule instead is used to simplify the evaluation of a lottery. This rule, *simplification*, takes a binary support lottery, $r \doteq x\alpha y$, and reduces it to a sure outcome z. Then in any lottery p where r is a distinct sublottery, $p \doteq p'\beta r$, the DM can use simplification to replace the outcomes of r by z to obtain $q = p'\beta z$ and we say pSq to mean p simplifies to q. Therefore to define the set of simplifications, I define some properties of simplifications on binary support lotteries, and simplification on more complex lotteries is defined as simplifying a binary sub-lottery.

Definition 2. *Simplification*

Let p, q be any binary support lotteries then S is such that the following holds:

- *Existence: pSx holds for some $x \in \mathcal{X}$.*
- *S-Monotonicity: If p strictly FOSD q and pSx_p, qSx_q then $x_p > x_q$.*
- *Continuity: $\forall x > y > z$, $\exists \alpha$ such that $(x\alpha z)Sy$.*

Let p be any lottery, then pSq if $p \doteq p'\alpha(x\beta y)$, $q = p'\alpha z$ and $(x\beta y)Sz$.
Finally, S is a simplification if the following also holds:

- *Consistency: For any lottery p, suppose p becomes x and y after two sequences of simplifications, then $x = y$.*

I note, just like for cancellation, that simplification is fairly restrictive in the sense that $p \doteq p'\alpha(x\beta y)$ is required to simplify $x\beta y$, namely, it is not possible to simplify away only a portion of x in p and leave the rest in p. I impose four conditions on simplification. The first three are regarding binary support lotteries. Existence claims that any binary support lottery can be simplified. Monotonicity further regulates the value of simplification over binary support lotteries by first-order stochastic dominance. Continuity states that if $x > y > z$ then there is some mixture of x and y which simplifies to z. These three are standard assumptions and frequently appear in the literature in different formulations. The fourth assumption is a tractability and normative assumption, namely, it states that each lottery can only be reduced to a unique outcome no matter the order of simplifications. It is necessary for tractability as otherwise a lottery may have two values but also it is normative as otherwise a lottery may be strictly preferred to itself.

One final notation is $(p, q) \to (p', q')$ which denotes that (p, q) becomes (p', q') after some sequence of application of rules, these could be the simplification of individual lotteries: $(p, q) \to (p', q)$ where pSp'. Or cancellation on the pair of lotteries $(p, q) \to (p', q')$ where $(p, q)C(p', q')$.

3 Standard Procedural Model and EU Representation

Given the two rules, I define the *standard procedural model*:

The Standard Procedural Model

1. Rules:
   - Simplification.
   - Cancellation.
2. Choice:
   - L-Monotonicity: $\forall x, y \in \mathcal{X}$, $x > y$ implies $\delta_x \succ \delta_y$.
   - O-Deduction: $\forall p_1, p_2$, $p_1 \succeq p_2$ if $(p_1, p_2) \to (p_1', p_2')$ and $p_1' \succeq p_2'$.

Under the standard procedural model (SPM), the DM first evaluates degenerate lotteries by monotonicity. For a choice between more complicated lotteries, the DM uses simplifications and cancellation to simplify the decision. This process is called an *O-deduction* (O for omniscient).

The first result states that if choices are generated with simplification and cancellation in such a way, then the DM is necessarily an expected utility maximizer. Similarly, by picking the appropriate simplifications, it is easy to see that any EU DM can be generated via some SPM.

Theorem 1. *Expected Utility Representation*

The following are equivalent:

- $\succeq$ *arises out of an SPM.*
- $\succeq$ *has an EU representation with continuous and strictly increasing utility.*

Proof: Appendix

This result highlights the complexity assumption of the EU model. In particular, O-Deduction assumes that the lengths of the sequence of application of rules do not matter. This further implies that each simplification and cancellation are equally costly, as they have zero costs. As I show in the next section, different forms of *deductions*, which put different restrictions on the type or length of simplification allowed, give rise to different representations of complexity aversion.

The reader may believe instead that O-Deduction is not too strong but rather instead the rules are too permissive. I offer now three additional characterizations of the expected utility representation with more restrictive rules.

Definition 3. *Adjacent Simplifications*

Let p be a lottery with outcomes $x_1 > x_2 > ... > x_n$, we say that $pS^a p'$ if pSp' and $p \doteq q\alpha(x_i \beta x_{i+1})$ and $p' = q\alpha z$.

S^a then denotes adjacent simplification. As the name suggests, an adjacent simplification is a simplification of adjacent outcomes. Consider the lottery $p = (10, 0.1; 8, 0.1; 6, 0.4; 4, 0.4)$, then a DM who is allowed to perform simplification may merge $4, 0.4$ and $10, 0.1$ together. Whereas a DM who is limited to adjacent simplifications can only merge $10, 0.1$ with $8, 0.1$. Therefore she simplifies first outcomes which are close in payoff terms.

Similarly, I weaken cancellation as follows by *replacement*:

Definition 4. *Replacement*
$(p,q)R(p',q')$ *if* $\exists \alpha, x, y,\ p \doteq p^*\alpha x, q \doteq q^*\alpha x$ *and* $p' \doteq p^*\alpha y$, $q' \doteq q^*\alpha y$.

Note that replacement is a weakening of cancellation. In particular, cancellation may be too strong in the sense that it supposes decision-makers can derive correctly the conditional probability after canceling out outcomes. With replacement, this issue does not arise and further replacement is strictly weaker as it says nothing about the relationship between (p,q) and (p',q') when $(p,q)C(p',q')$ holds. I now introduce three models with weaker rules than the SPM.

Definition 5. *Three Models with Restrictions on Rules*

- *Adjacent PM: Replace simplification with adjacent simplification in the SMP.*
- *Replacement PM: Replace simplification with adjacent simplification and replace cancellation with replacement in SMP.*
- *No Cancellation PM: Remove cancellation from SMP.*

My next result states that these models, with restricted types of simplification and removal/weakening of cancellation, are still equivalent to an expected utility representation. Hence, suggesting that the core complexity assumption is on deduction.

Theorem 2. *Generalized Expected Utility Representation*
The following are equivalent:

- $\succeq$ *arises out of an SPM.*
- $\succeq$ *arises out of an Adjacent PM.*
- $\succeq$ *arises out of a Replacement PM.*
- $\succeq$ *arises out of a No Cancellation PM.*
- $\succeq$ *has an EU representation with continuous and strictly increasing utility.*

Proof: Appendix

Before moving on to relaxing O-Deduction, I highlight another perspective on my results. So far I have considered preference as arising out of choice between lotteries. In experiments and real-world decisions, the DM often chooses a certainty equivalent, for instance, how much to pay for insurance or to play roulette. Therefore, one question is whether a DM who evaluates lotteries by certainty equivalents via these procedures also must satisfy expected utility. Formally speaking, let S be a simplification relation. Suppose a DM evaluates each lottery p the same as the certainty equivalent $CE(p)$ obtained through applications of S. Then a corollary of Theorem 2 is that such a DM is necessarily using expected utility to evaluate each lottery.

4 Complexity Aversion

In this section, I first consider two models of complexity aversion and show how they arise out of restrictions of *O-Deduction*.

I consider two models: the *support size cost* representation of Puri (2020) and the *entropy cost* representation of Mononen (2021) . I start with the support size cost representation:

Definition 6. *Support Size Cost Representation*
$\succeq$ has a support size cost representation if:

$$p \succeq q \Leftrightarrow \sum_{x \in supp(p)} u(x)p(x) - C(|supp(p)|) \geq \sum_{x \in supp(q)} u(x)q(x) - C(|supp(q)|)$$

Where u is a continuous, strictly increasing utility function and C is a non-decreasing cost function.

Therefore, the model posits that the DM evaluates a lottery by its expected utility but may penalize it for being too complex. Where complexity is taken to be the outcome support size.

First, I consider a modification of O-Deduction. Let S-Deduction be defined as follows:

Definition 7. *S-Deduction*
S-Deduction: $\forall p, q$, $p \succeq q$ if $(p, q) \to (p', q')$ after n_p, n_q many simplifications or replacements such that:

- $n_p \leq n_q$
- $p' \succeq q'$

Therefore S-Deduction differs from O-Deduction in two ways. First, the number of applications of simplification matters. The DM prefers to have to perform fewer simplifications. Therefore complexity aversion is introduced directly as an aversion to having to perform difficult simplifications. Second, cancellation is replaced by replacement, but as we know from Theorem 2, this is not what drives non-EU behavior. Finally, note that, unlike O-deduction, S-deduction is more limited and does not offer a complete preference relation over lotteries.

Given this, I define a second procedural model:

The Outcome Procedural Model

1. Rules:
   - Adjacent Simplification.
   - Replacement.
2. Choice:
   - L-Monotonicity: $\forall x, y \in \mathcal{X}$, $x > y$ implies $\delta_x \succ \delta_y$.
   - S-Deduction.
   - Weak order: $\succeq$ is complete and transitive.

As mentioned, S-deduction does not guarantee completeness, so the model is closed off directly by assuming that $\succeq$ is a weak order. Note then that the only difference between this model and the Replacement PM model which is equivalent to expected utility is that O-deduction is replaced by S-deduction, as for both cases $\succeq$ is a weak order. This slight change in terms of the decision-maker's attitudes towards applying rules generates exactly the support size cost representation as my next theorem indicates.

Theorem 3. *Support Cost Size Representation*

The following are equivalent:

- *$\succeq$ arises out of an OPM.*
- *$\succeq$ has a support size cost representation.*

Proof: Appendix

It is intuitive to see how this model translates to the support size cost representation. By definition, an adjacent simplification reduces the number of outcomes by one. In addition, the DM only cares about the number of simplifications, the less the merrier. Therefore it is unsurprising that a support size cost arises out of the outcome procedural model (OPM).

I note that while cancellation and replacement are equivalent under O-Deduction, adding it to the S-deduction implies that the cost function is a multiplicative function of the support size or that the cost of each additional outcome is higher than the utility of the best outcome.

Corollary 1. *Cancellation Support Size Cost*

Suppose cancellation occurs instead of replacement in the definition of S-deduction and the Outcome Procedural Model. Then $\succeq$ has a support size cost representation with cost $C(p) = \lambda|supp(p)|$ with $\lambda \geq 0$ or $C(n) - C(n-1) > u(m^+)$ for all n.

One might ask whether lottery support size is "the" measure of complexity in this procedural framework. After all, simplification reduces the number of outcomes, and as I pointed out, complexity aversion is defined as an aversion to the application of rules (in this case reducing outcome number). I show next that such is not the case. In particular, S-deduction makes only restrictions on the number but not the type of simplifications. One can indeed obtain different costs with another restriction on the type of simplifications allowed. To do this I define another model of complexity cost, axiomatized recently by Mononen (2021):

Definition 8. *Entropy Cost Representation*
$\succeq$ *has an entropy cost representation if:*

$$p \succeq q \Leftrightarrow \sum_{x \in supp(p)} u(x)p(x) - \gamma \mathcal{H}(p) \geq \sum_{x \in supp(q)} u(x)q(x) - \gamma \mathcal{H}(q)$$

Where u is a continuous, strictly increasing utility function, $\mathcal{H}$ is the Shannon entropy and $\gamma \in [0, \infty)$ a cost parameter.

Therefore this model also discounts the complexity of a lottery, but by its entropy instead of its outcome size. To characterize this representation in my framework, two definitions are necessary:

Definition 9. *Mixture Continuity*
$\succeq$ *is mixture continuous if $\forall p, q, r$, the sets $\{\alpha \in [0,1] \mid p\alpha q \succeq r\}$ and $\{\alpha \in [0,1] \mid p\alpha q \preceq r\}$ are closed.*

Mixture continuity implies that if $\succeq$ is represented by some value function, then it is continuous in probabilities. Note this fails for the support size cost representation as when α goes to 0 or 1, the outcome support size may change, leading to a non-continuous change in value.

Definition 10. *Parallel Simplifications*
Let (p, q) become $(p^* \alpha z_p, q^* \alpha z_q)$ by applying one simplification each to p and q. We say this pair of simplification is parallel if $\exists \beta$ such that $p \doteq p^* \alpha (x_p \beta y_p)$ and $q \doteq q^* \alpha (x_q \beta y_q)$.

Consider the simplification of $p\alpha(x_p \beta y_p)$ to $p \alpha z_p$, this process simplifies away an α proportion of a β mixture of two outcomes. A parallel pair is precisely a pair of simplifications such that they both simplify away the same α proportion which has the same β mixture. In other words, a pair of simplifications is parallel whenever the same probabilities are simplified away. The intuition is that a DM may be able to spot that some outcomes have the same probabilities and choose to simplify simultaneously these outcomes.

With these two definitions, I introduce E-deduction which characterizes the entropy cost model. Similar to S-deduction, it imposes assumptions on how the DM values the complexity of rules, and here the assumption is that whenever two simplifications are parallel, they have the same complexity value.

Definition 11. *E-Deduction*
$\forall p, q$, $p \succeq q$ if either one holds:

- *p can be obtained from q through some sequence of adjacent simplifications.*

- $(p, q) \to (p', q')$ *through a sequence of parallel pairs of simplifications or cancellations and* $p' \succeq q'$.

Therefore E-deduction has two clauses, the first one is exactly complexity aversion to simplifications. The second states that the DM can apply sequences of simplifications as long as they are "obvious" in the sense that they are pairs of parallel simplifications. With these definitions, I introduce the entropy procedural model (EPM):

The Entropy Procedural Model

1. Rules:
   - Adjacent Simplification.
   - Cancellation.
2. Choice:
   - L-Monotonicity: $\forall x, y \in \mathcal{X}$, $x > y$ implies $\delta_x \succ \delta_y$.
   - E-Deduction.
   - Weak order: $\succeq$ is complete and transitive.
   - Mixture continuity.

The next result states that if a DM uses adjacent simplification and cancellation à la E-deduction to derive a preference that is a weak order, monotonic, and mixture continuous, then it has an entropy cost representation. As well as the reverse:

Theorem 4. *Entropy Cost Representation*
The following are equivalent:

- $\succeq$ *arises out of an EPM.*
- $\succeq$ *has an entropy cost representation.*

Proof: Appendix

Therefore this result illustrates that outcome size is not "the" measure of complexity in this model. Rather, with different types of deductions, the DM will have different measures of complexity.

Lastly, I comment on cancellation. Note that the two models differ also in whether cancellation holds. On the surface, cancellation seems to be a natural rule and should be included in deductions, i.e. $(p, q) \to (p', q')$ via a cancellation then the preferences should match. Indeed, we have this intuition that cancellation is "obvious" and that DMs perform such a rule. This is exactly what the entropy cost model and expected utility models do, but the support size cost representation as well as some other non-EU models do not require this. This type of deduction implies that a cancellation applied to a pair of lotteries changes the complexity of both lotteries by an equal amount. In the support size cost model, a change of 100 outcomes to 99 outcomes is allowed to have a different effect than that of 2 outcomes to 1 outcome.

5 Partition Complexity

The last two models' approaches to complexity have a common point in that complexity is independent of the cardinal value of outcomes. To elaborate on this, the support size cost model uses as measure the number of outcomes, whereas the entropy cost model uses entropy, both disregarding the values of the outcomes and focusing only on an outcome's contribution to the composition of the lottery. This approach misses many behaviors relevant to complexity. For instance, it may be easy for a DM to figure out the certainty equivalent for a lottery yielding \$4 and \$4.5 with equal probability, but much harder for a lottery yielding \$7 and \$100 with equal probability. To illustrate this further, consider lotteries p and q below. The lotteries have the same number of outcomes as well as entropy, and thus it is permissible in both models to replace the common outcome $(\$1, 20\%)$ by $(\$5.2, 20\%)$ which yields p', q'. Furthermore, both theories claim that the four lotteries have equal complexity.

p		q	
\$1.00	**20%**	**\$1.00**	**20%**
\$1.25	20%	\$5.00	20%
\$1.50	20%	\$5.10	20%
\$9.00	20%	\$5.25	20%
\$9.50	20%	\$5.50	20%

p'		q'	
\$5.20	**20%**	**\$5.20**	**20%**
\$1.25	20%	\$5.00	20%
\$1.50	20%	\$5.10	20%
\$9.00	20%	\$5.25	20%
\$9.50	20%	\$5.50	20%

Nevertheless, one gets the impression that q' has become much easier to evaluate and to give a certainty equivalent for because now all values are within \$0.5 of each other. On the contrary, one gets the impression that p' is at least as complex as p was. The outcomes of p can be categorized as low and high, but now in p', there are plausibly three categories of outcomes, rendering the lottery potentially more complex.

In this section, I consider a DM who evaluates a lottery in two steps. In the first step, she groups outcomes that are "similar" enough in value, creating a partition of outcomes. Then, in the second step, she evaluates the lottery by its value minus some cost on the complexity of the partition. For such a DM, it is then indeed the case that q is less complex than q' as long as the DM considers $\$1$ to belong to a different partition than the rest. And indeed p is at least as complex as p as it has up to three natural cells in its partition whereas p only has two. I define the representation as follows:

Definition 12. *Partition Complexity*

We say $\succeq$ has a partition complexity representation if, for every lottery p and q, there exists partitions of outcomes of p, q denoted by P_p, P_q such that the cells of each partition are contained in mutually disjoint *intervals and:*

$$p \succeq q \Leftrightarrow \sum_{x \in supp(p)} u(x)p(x) - \mathcal{C}(|supp(P_p)|) \geq \sum_{x \in supp(q)} u(x)q(x) - \mathcal{C}(|supp(P_q)|)$$

Where C *is the cost function of the partition which is weakly increasing in the support size of the partition.*

Therefore for each lottery, the decision-maker partitions its outcome such that each cell is contained in its own interval. Note here that we do not impose strong conditions on what counts as "close in value". Rather, it could be different depending on the lottery at hand. For instance for a lottery with outcomes all of which are within $1 of each other, close may be $0.25, however for a lottery with amounts in the millions, close may be in the tens of thousands.

I introduce two conditions first and then introduce the deduction rules.

Definition 13. *Fixed Simplifications*
We say a decision-maker uses fixed simplifications if for each lottery p *she uses a fixed sequence of simplifications each time.*

Definition 14. *Efficient Simplifications*
For each lottery, the set of simplifications is divided into easy and hard simplifications, we say a sequence of simplifications is efficient for this lottery if the easy simplifications occur before hard simplifications.

First, unlike previous models where the deductions were allowed to be order independent, here the decision-maker fixes some order to make her simplifications. Note here that it does not impact the final product as the consistency condition of simplification implies that any order yields the same unique outcome. However, the order of application is important to ensure that the same partition arises for each lottery during each comparison. Second, I impose that, for each lottery, the decision-maker finds certain simplifications to be easier than others, and that she always performs the easy ones first.

Given these two conditions, I define the deduction rule:

Definition 15. *SP-Deduction*
SP-Deduction: $\forall p, q$, $p \succeq q$ *if* $(p, q) \to (p', q')$ *is a fixed and efficient sequence which contain* n_p, n_q *many hard simplifications and:*

- $n_p \leq n_q$
- $p' \succeq q'$

SP-Deduction is therefore the same as S-Deduction with the only difference that only hard simplifications are costly as well as the fact that we no longer consider the replacement rule.

The Partition Procedural Model

1. Rules:

   - Adjacent+Efficient+Fixed Simplifications.

2. Choice:

   - L-Monotonicity: $\forall x, y \in \mathcal{X}$, $x > y$ implies $\delta_x \succ \delta_y$.
   - SP-Deduction.
   - Weak order: $\succeq$ is complete and transitive.

Therefore the Partition Procedural Model (PPM)'s main difference compared to the OPM is that the decision-maker distinguishes between easy and hard simplifications as well as only finds the hard simplifications costly to perform. Then the idea is that she performs first the easy simplifications, and given that they must be adjacent, she really is partitioning outcomes. Once there are no more easy simplifications, she has fully partitioned the lottery's outcomes. Then when comparing lotteries, she must make hard simplifications, so instead, she now discounts the number of cells in the partition as each hard simplification merges two cells of a partition.

Theorem 5. *Partition Complexity Representation*

The following are equivalent:

- $\succeq$ *arises out of a PPM.*

- $\succeq$ *has a partition complexity representation.*

Proof: Appendix

This result highlights that distinguishing between easy and hard simplifications leads to a partition representation. I note that the partition is endogenous because what counts as easy or hard may differ from lottery to lottery. Note again the theme of the paper, the rules reveal the complexity, here the attitudes towards different simplifications end up showing us what the decision-maker finds complex: the cells of partitions rather than just outcomes.

Suppose instead that a universal categorization of easy and hard were applied for all lotteries, then the partition and utility will have to satisfy certain compatibility conditions which I plan on elaborating in a future draft.

6 Empirical Tests

I now consider a simple empirical analysis of the models considered so far. I obtained data from CPC-18[3] and compared the different models via a cross-validation task. CPC-18 consists of around

[3]See `http://www.cpc-18.com`.

200 distinct choices between two lotteries, I remove lotteries with correlation and ambiguity, which leaves 171 choices. Subjects choose between these lotteries repeatedly and each choice is made 25 times by each subject. I use the aggregate choice probabilities[4] to test between the different models.

For each model, I estimate its parameters by assuming a random CARA utility model with a Gumbel error. For example, the entropy cost model gives:

$$P(p_1|\{p_1,p_2\}) = \frac{\exp(EU(p_1) - \gamma\mathcal{H}(p_1))}{\exp(EU(p_1) - \gamma\mathcal{H}(p_1)) + \exp(EU(p_2) - \gamma\mathcal{H}(p_2))}$$

Here EU is estimated using a CARA utility: $EU(p) = \mathbb{E}_p[\frac{x^{1-\gamma}}{1-\gamma}]$. For each of the three models and EU, I estimate γ along with complexity aversion parameters by minimizing the MSE of a training sample, then using these estimated parameters, I compute the MSE of the testing sample. A few details are necessary for the outcome support size and partition complexity model. For the outcome support size model, as the sample of lotteries has at most ten outcomes, I estimate ten cost parameters, one for each support size. Similarly, for the partition complexity, the same approach is taken, in addition I construct the partition by picking a number $\delta > 0$ such that the partition is taken to be the one with the smallest support size given that each cell only contains outcomes within δ of each other. I note δ is also a parameter which is estimated. Note further that I am therefore imposing a restriction on the partition complexity model as δ is a uniformity assumption on the size of cells of partitions whereas the theory makes no such assumption.

Given these procedures, I selected a training sample of size 30 and a testing sample of size 141. These samples are randomly drawn 200 times and Table 1 presents the results of the cross-validation task. I report here the mean squared error and its standard deviation times 100.

Table 1: Cross Validation Task

	Expected Utility		Entropy Cost		Support Size Cost		Partition Complexity	
	MSE Train	MSE Test	MSE Train	MSE Test	MSE Train	MSE Test	MSE Train	MSE Test
Mean $\times 100$	6.78	6.83	6.77	6.87	6.72	6.84	6.27	6.66
SD $\times 100$	(0.29)	(1.37)	(0.29)	(1.37)	(0.28)	(1.33)	(0.30)	(1.36)

The results first reveal that the entropy cost and support size cost models do not differ significantly from EU for both the MSEs of the training sample and testing sample. This may be partially due to the data structure and the choice of the utility function. Fudenburg and Puri (2022) shows that the support size cost model is very good at predicting *certainty equivalent* when paired with cumulative prospect theory. Here I predict the *aggregate choice probabilities* instead and the data structure is also different from theirs. Nonetheless, a paired t-test shows that both the training and testing sample MSEs of the partition complexity model are statistically significantly lower than all three other models, with $p < 0.001$. For example, the mean (trial by trial) difference of

[4] Available at `http://www.cpc-18.com/data`.

MSE$\times 100$ between the support size cost model and partition complexity model is $0.46/0.17$ for the training/testing sample and this difference has an STD$\times 100$ of $0.13/0.51$.

7 Conclusion

This paper explores the relationship between rules, complexity, and non-expected utility behavior. I introduce a procedural choice model where the decision-maker's attitude towards different rules is used to derive representations of her choice. The key channel is as follows: the rules a decision-maker uses, and how she uses them, reveals what she finds complex and therefore gives a measure of complexity. For an expected utility decision-maker, she uses simplification without discriminating between different simplifications or incurring any costs. Thus my method highlights a complexity assumption of expected utility when framed in a procedural manner. When the decision-maker starts to discriminate against the rules and distinguish between different types of rules, I obtain different representations where the difference is precisely in the complexity aversion cost. Finally, a crude empirical analysis shows some support for the partition complexity model.

The paper has chosen decision-making under risk, where objects are lotteries. This is an explicit choice for two reasons. First, there is already some literature in this area: theoretical models as well as experimental findings to build upon. Second, the goal was not to focus on the complexity of events or outcomes. Therefore, lotteries are the perfect setting as both the outcome, monetary values, and event, probability, are well understood and not sources of complexity. This allows the side-stepping of these issues and focuses on the complexity of evaluating a lottery procedurally. I fully believe, however, that a promising direction is to study the complexity of events, in say a Savage framework, and see whether these are related to ambiguity and stochastic behavior.

References

Agranov, M. and P. Ortoleva (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy 125*(1), 40–68.

Fudenburg, D. and I. Puri (2022). Evaluating and extending theories of choice under risk. *Available at* `https://economics.mit.edu/files/21287`. Working Paper.

Halevy, Y. and G. Mayraz (2021). Identifying rule-based rationality. Working Paper.

Huck, S. and G. Weizsäcker (1999). Risk, complexity, and deviations from expected-value maximization: Results of a lottery choice experiment. *Journal of Economic Psychology 20*(6), 699–715.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 363–391.

Kovářík, J., D. Levin, and T. Wang (2016). Ellsberg paradox: Ambiguity and complexity aversions compared. *Journal of Risk and Uncertainty 52*(1), 47–64.

MacCrimmon, K. R. (1968). Descriptive and normative implications of the decision-theory postulates. In *Risk and uncertainty*, pp. 3–32. Springer.

MacCrimmon, K. R. and S. Larsson (1979). Utility theory: Axioms versus 'paradoxes'. In *Expected utility hypotheses and the Allais paradox*, pp. 333–409. Springer.

Mononen, L. (2021). On preference for simplicity and probability weighting. Working Paper.

Moskowitz, H. (1974). Effects of problem representation and feedback on rational behavior in allais and morlat-type problems. *Decision Sciences 5*(2), 225–242.

Nielsen, K. and J. Rehbeck (2021). When choices are mistakes. *Available at SSRN 3481381*. Working Paper.

Oprea, R. (2020). What makes a rule complex? *American economic review 110*(12), 3913–51.

Puri, I. (2020). Preference for simplicity. *Available at SSRN 3253494*. Working Paper.

Rubinstein, A. (1988). Similarity and decision-making under risk (is there a utility theory resolution to the allais paradox?). *Journal of economic theory 46*(1), 145–153.

Slovic, P. and A. Tversky (1974). Who accepts savage's axiom? *Behavioral science 19*(6), 368–373.

Tversky, A. (1969). Intransitivity of preferences. *Psychological review 76*(1), 31.

Appendix

A Proofs

Theorem 1

I show Theorem 1 in two steps. First, via the Mixture Space Theorem I show that there is an expected utility representation. Second, I extend this expected utility representation to be continuous and strictly increasing.

Step 1: Recall the Mixture Space Theorem says that a weak order that satisfies continuity and independence must have an expected utility representation (not necessarily continuous).

Weak Order. From the consistency requirement, we know each p has a unique certainty equivalent from any sequence of simplifications. Therefore lotteries can be compared via this certainty equivalent and $\succeq$ must be complete and transitive.

Independence. Recall independence states $p \succeq q$ implies $p\alpha r \succeq q\alpha r$. Now note if $supp(p) \cap supp(r) = \emptyset$ and $supp(q) \cap supp(r) = \emptyset$ then the proof is done by cancellation. To circumvent this, for each $x \in supp(r) \cap [supp(p) \cup supp(q)]$, we can find x_1, x_2 such that $x_1 \alpha x_2 S x$ and $x_1, x_2 \notin supp(r) \cap [supp(p) \cup supp(q)]$. We can also pick x_1, x_2 such that they are distinct from other $y \in supp(r) \cap [supp(p) \cup supp(q)]$ and their respective y_1, y_2 by the continuity property of Simplification. Then we transformed r to r' such that $supp(r') \cap [supp(p) \cup supp(q)] = \emptyset$ and note by construction $p\alpha r' \to p\alpha r$ as well as $q\alpha r' \to q\alpha r$, so $p\alpha r' \sim p\alpha r$ and $q\alpha r' \sim q\alpha r$. But on the pair $p\alpha r', q\alpha r'$ we can apply cancellation and obtain the desired result.

Continuity. Recall the type of continuity required is $p \succ q \succ r$ then $\exists \alpha, \beta$ such that $p\alpha r \succ q \succ p\beta r$. Again let us define r' such that $supp(p) \cap supp(r') = \emptyset$ while $p\alpha r' \to p\alpha r$. Then let $x_p > x_q > x_r$ be the respective certainty equivalents arising from the lotteries. Note we can also pick r' such that no outcome of r' arises in the sequence of deduction $p \to x_p$ which implies $p\alpha r' \to x_p \alpha x_r$ is possible. Then we need show only $x_p \alpha x_r > x_q > x_p \beta x_r$ holds for some α, β. Note that must be true by S-monotonicity and Existence.

Step 2: Now that we have this utility function, we wish to show it is continuous. Suppose not and a discontinuity exists at x. Then consider some $x_{\min} < x < x_{\max}$ and now consider $f : [x_{\min}, x_{\max}] \to [0, 1]$ such that $f(y) = \alpha$ if and only if $(x_{\max} \alpha x_{\min}) S y$. Note of course that f is injective, onto, and strictly increasing, so necessarily continuous. Then the value at $u(x)$ is clearly continuous as it is a continuous function of $f(x)$.

Finally the strictly increasingness of u falls out naturally from L-monotonicity and the construction of u.

To show the other direction, suffice to pick the appropriate simplifications and note that the expected utility form arises naturally.

Theorem 2

Fix some simplification relation S, and let $\succeq_S, \succeq_A, \succeq_R, \succeq_{NC}$ be preferences arising from the SPM, Adjacent PM, Replacement PM and No Cancellation PM.

Note first that the proof of Theorem 1 still holds when we are restricted to adjacent simplifications. The only tweak necessary is in **Step 1** the proof of independence and continuity requires us to pick x_1, x_2 arbitrarily close to x to guarantee that $p\alpha r' \to p\alpha r$. This implies that $\succeq_S = \succeq_A$.

Now I show that adjacent simplification and replacement imply cancellation. This implies an expected utility representation as the proof of Theorem 1 then follows. Let $p \succeq q$ we want to show that $p\alpha x \succeq q\alpha x$ whenever x is an outcome that does not appear in p or q. Note first that if $x \notin (\min\{x \in supp(p) \cup supp(q)\}, \max\{x \in supp(p) \cup supp(q)\})$, then the result is immediate by first deriving the certainty equivalents of p and q via adjacent simplifications and then applying monotonicity. Now suppose $x \in (\min\{x \in supp(p) \cup supp(q)\}, \max\{x \in supp(p) \cup supp(q)\})$ then the goal is to show that $p\alpha x \sim \delta_p \alpha x$, where $p \to \delta_p$ and similarly for q. Note again that if $x \notin (\min\{x \in supp(p)\}, \max\{x \in supp(p)\})$ then $p\alpha x \sim \delta_p \alpha x$ is immediate. Suppose not and let $p\alpha x \to (\overline{p}, \alpha_1; x, 1-\alpha; \underline{p}, \alpha_2)$. Where $\overline{p}$ and $\underline{p}$ are outcomes obtained after applying adjacent simplifications to all outcomes larger and smaller than x. Then we know $\overline{p}\frac{\alpha_1}{\alpha_1+\alpha_2}\underline{p} \to \delta_p$ but cannot apply this here as it is not an adjacent simplification with x in between. Note that $m_- < \underline{p}$ and $m^+ > \overline{p}$ then clearly $\exists z$ such that $\delta_p \alpha z \sim (\overline{p}, \alpha_1; z, 1-\alpha; \underline{p}, \alpha_2)$ by picking z small or large enough that the two ps can be simplified adjacently, then applying replacement we get the desired result that $\succeq_S = \succeq_R$

Thirdly, note that simplifications imply cancellation. To see this, let $p \doteq p'\alpha x$ and $q \doteq q'\alpha x$, then cancellation is $p' \succeq q'$ implies $p \succeq q$. To see this, use simplification to simplify $p'\alpha x \to x_p \alpha x$ and $q'\alpha x \to x_q \alpha x$. Then note $p' \succeq q'$ implies $x_p > x_q$, finally S-monotonicity implies $x_p \alpha x \succeq x_q \alpha x$ which is precisely $p \succeq q$, as the latter becomes the former via a set of simplifications. Therefore this implies that simplification alone is sufficient for the EU representation and $\succeq_S = \succeq_{NC}$.

Theorem 3

To show this result, consider $\succeq^*$ defined as follows:

- L-Monotonicity: $\forall x, y \in \mathcal{X}$, $x > y$ implies $\delta_x \succ^* \delta_y$.
- S*-deduction: $\forall p, q$, $p \succeq^* q$ if there exists adjacent simplifications and replacements such that $p \to p'$ and $q \to q'$ and $p' \succeq^* q'$.

Therefore $\succeq^*$ is the same as $\succeq$ except we remove the restrictions on the number of simplifications. We note by **Theorem 2** that $\succeq^*$ arises out of a Replacement PM and thus has an expected utility representation. Let us denote the utility function by u.

Now we show that $\succeq^* = \succeq$ when restricted to sets of lotteries with the same support size. That is, if $|supp(p)| = |supp(q)|$ then $p \succeq^* q \Leftrightarrow p \succeq q$. To see this, suffice to notice that one way to compare p, q under S-deduction is to find their certainty equivalent. In which case $n_p = n_q$ and

$x_p \geq x_q$ implies $p \succeq q$. It remains to note that the same procedure works for S*-deduction

Therefore, each set of lotteries that have the same support size have $\succeq$ represented by u. Now by transitivity and completeness of $\succeq$, I show there is a cost per outcome size and construct it inductively.

If possible, pick two lotteries $p_1 \sim p_2$ such that $p_1 = \delta_p$ and $|supp(p_2)| = 2$. Then define $c(2) - c(1) = u(p_1) - u(p_2)$. We know this cost is positive as by S-deduction we have $u(p_1) \geq u(p_2)$. To see that this is well defined for all lotteries of support size 1 and 2, note that transitivity gives it to us. If there were two such costs, then transitivity gives a simple contradiction. If such a pair of lottery is not possible, then set $c(2) = u(m^+) - u(m^-)$, some arbitrarily high enough cost. We then construct the rest of the cost function similarly.

To show the reverse direction is trivial. As the representation easily satisfies the rules of OPM.

To show Corollary 1, note simply that if cancellation occurs instead of replacement, then removing 1 outcome from each lottery preserves their comparison, so the same amount of "cost" must have been removed. Or it is preserved because each support size addition has a cost higher than $u(m^+)$.

Theorem 4

To show this result, consider the following definitions. Let p and q be two lotteries that have the same *co-probability* whenever there is a bijective function f of outcomes of p to outcomes of q such that $p(x) = q(f(x))$. In other words, for every outcome of p with some probability $p(x)$ there is an outcome of q with the same probability. Another final way to understand is that p and q have the same probability but potentially different outcomes. Consider now a few properties defined on $\succeq$, for more discussion, see Mononen (2021).

Definition 16. *Sequential continuity with respect to co-probability*

Let $p_n \to p$ be a sequence of lotteries with the same co-probability and $\forall n$, $p_n \succeq (\preceq) q$ then $p \succeq (\preceq) q$.

Definition 17. *Same outcome independence*

Let $p \succeq q$, $\forall x, p', q'$ such that $p' \doteq p\alpha x$ and $q' \doteq q\alpha x$ we have $p' \succeq q'$.

Definition 18. *Co-Probability FOSD*

Let p, q have the same co-probability, if p FOSD q then $p \succeq q$.

From Mononen (2021), it is shown that $\succeq$ is a weak order, mixture continuous, sequentially continuous with respect to co-probability, same outcome independence, and co-probability FOSD if and only if it has an entropy cost representation. Our proof verifies the axioms.

Weak order and mixture continuity are directly assumed so they hold. Similarly, the same outcome independence is exactly cancellation, so it holds as well.

To show the remaining two, consider $\succeq^*$ defined as follows:

- L-Monotonicity: $\forall x, y \in \mathcal{X}$, $x > y$ implies $\delta_x \succ^* \delta_y$.
- E*-deduction: $\forall p, q$, $p \succeq^* q$ if there exists adjacent simplifications and cancellations such that $p \to p'$ and $q \to q'$ and $p' \succeq^* q'$.

Therefore $\succeq^*$ is the same as $\succeq$ except we remove the restrictions on the number and type of simplifications. We note by **Theorem 2** that $\succeq^*$ arises out of an Adjacent PM and thus has an expected utility representation. Let us denote the utility function by u.

I show that $\succeq = \succeq^*$ when restricted to sets of probability with the same co-probability. Again, this is easy to see as the definition of E-deduction and E*-deduction yield the same comparison of such pairs (as long as S is fixed, which it is). This implies that Co-Probability FOSD and sequential continuity holds. This implies $\succeq$ has an entropy cost representation.

To show the reverse direction is trivial. Given some utility function, first, pick the set of simplifications that correspond to it. Then for any two lotteries with the same co-probability, we know the choice from EPM is the same as that of the entropy cost model. Now for some choice of γ we can simply complete the preference ordering such that two lotteries from different co-probability are ranked according to the entropy cost model.

Theorem 5

First let $p \to p^s$ denote that p^s is obtained from p after applying all the easy simplification of the fixed sequence assigned to p. Then note that $p \to p^s$ and $q \to q^s$ implies $p \succeq q \Leftrightarrow p^s \succeq q^s$. This is held by SP-deduction.

Second, consider $\succeq^*$ obtained as follows:

- L-Monotonicity: $\forall x, y \in \mathcal{X}$, $x > y$ implies $\delta_x \succ^* \delta_y$.
- SP*-deduction: $\forall p, q$, $p \succeq^* q$ if there exists adjacent and fixed simplifications such that $p \to p'$ and $q \to q'$ and $p' \succeq^* q'$.

I show that $\succeq^*$ has an expected utility representation. Note that S is consistent so it is order independent. Let $CE(p) \in \mathcal{X}$ be the "certainty equivalent" of p obtained through the adjacent and fixed simplification. Note of course that the same $CE(p)$ would be obtained through any simplification as well. So $\succeq^* = \succeq_{NC}$ given some fixed S, which proves this claim.

Now consider any two p, q, and p^s, q^s such that p^s, q^s have the same number of hard outcomes, it is clear then that $p^s \succeq q^s$ if and only if $p^s \succeq^* q^s$. So there is the same utility function u such that whenever p and q are such that $supp(p^s) = supp(q^s)$, we have that these two are ranked by expected utility on u.

Then we can construct recursively the cost function in a manner analogous to Theorem 3. Finally, to see that these must form partitions that are over mutually exclusive intervals, suffice to note that adjacent simplifications necessarily lead to such intervals.

To show the sufficiency part is easier. First, given the utility function, we pick S such that it matches. Second, we classify easy and hard simplifications such that the correct partitions arise, note that it is simple to do as we can classify all simplifications within the cell to be easy and any outside of it to be hard. Third, we simply complete the preference between p^s, q^s with different support sizes to match our desired cost function.