# When Choices are Mistakes\*

Kirby Nielsen<sup>†</sup> and John Rehbeck<sup>‡</sup>

Draft: January 24, 2020

#### **Abstract**

Using a laboratory experiment, we identify whether decision makers consider it a mistake to violate canonical choice axioms. To do this, we incentivize subjects to report which of several axioms they want their decisions to satisfy. Then, subjects make lottery choices which might conflict with their stated axiom preferences. We give them the opportunity to re-evaluate their decisions when lotteries conflict with desired axioms. We find that a majority of individuals want to follow the axioms and revise their lottery choices to be consistent with them. We interpret this to mean that most axiom violations in our sample were mistakes.

KEYWORDS: mistakes; rationality; normative versus descriptive theory; procedural decision making

<sup>\*</sup>We thank Dan Benjamin, Doug Bernheim, Paul Feldman, Tzachi Gilboa, Paul J. Healy, Miles Kimball, Muriel Niederle, Ryan Oprea, Collin Raymond, Frank Schilbach, Joel Sobel, and Colin Sullivan for many helpful comments and suggestions. A previous version of this paper was circulated with the title "Are Axioms Normative? Eliciting Axiom Preferences and Resolving Conflicts with Lottery Choices." This study was approved by the Institutional Review Boards at Stanford University and The Ohio State University.

<sup>†</sup>kirbyn@stanford.edu; Department of Economics, Stanford University

<sup>&</sup>lt;sup>‡</sup>rehbeck.7@osu.edu; Department of Economics, The Ohio State University

In reversing my preference... I have corrected an error. There is, of course, an important sense in which preferences, being entirely subjective, cannot be in error; but in a different, more subtle sense they can be.

Leonard Savage (1954)

### I. INTRODUCTION

An enormous experimental literature—spanning at least six decades—has shown that individuals consistently violate the canonical axioms in decision theory. However, it is not clear from the literature whether these violations are intentional deviations from the axioms or are simply mistakes. While it is obvious that these violations demonstrate that the theory is not *descriptively* accurate, it could be the case that the theory is still *normatively* appealing to the decision makers. An individual who wants to follow an axiom might just find it difficult to make choices that are reliably consistent with the axiom. When an individual violates an axiom but would not have done so had they *known* they were violating the axiom, we call the violation a mistake. If violations of canonical axioms stem mainly from mistakes, then we can retain the normative content of the theory. However, if an individual violates the axioms because they do not *want* to follow them, then one should look for other "behavioral axioms" that the individual agrees with.

In this paper, we report results from the first incentivized experiment designed to detect mistakes when individuals violate axioms in the domain of risk. The difficulty in detecting a mistake is that it requires two pieces of information in addition to the choice data collected in standard experiments: knowledge of the axioms an individual wants to follow and observation of how individuals reconcile lottery choice and axioms when they conflict. By directly and incentive-compatibly eliciting these two missing pieces of information, we provide a general method to detect the "subtle sense" in which individuals make mistakes as mentioned by Savage (1954).

<sup>&</sup>lt;sup>1</sup>Examples include May (1954); MacCrimmon (1968); Tversky (1969); Slovic and Tversky (1974); Kahneman and Tversky (1979); Huber et al. (1982); Segal (1988); Loomes et al. (1991); Wedell (1991); Loomes et al. (1992); Camerer (1995); Birnbaum and Chavez (1997); Seidl (2002); Birnbaum and Martin (2003); Birnbaum and Schmidt (2008); Regenwetter et al. (2011); Birnbaum et al. (2016), among many others.

To accomplish this, our experiment has three main parts. First, we elicit the axioms a subject wants to follow by presenting the axioms as "decision rules" for making risky choices. In particular, subjects choose whether to have the rule select lotteries for them or whether they want to make the lottery choices on their own. Subjects then face standard binary decision problems in which the axioms are routinely violated. Our main focus is on subjects who want to follow an axiom but violate it when making lottery choices. To observe how subjects perceive this inconsistency, we give them the opportunity to re-evaluate their decisions when lottery choices conflict with the axioms they want to follow. If the axioms remain normatively appealing, then subjects should change their lottery choices to conform to the axioms. If an axiom does not remain normatively appealing in the face of a violation, then subjects should abandon the axiom as a decision rule. In short, we are able to identify mistakes whereby individuals reconcile axioms and lotteries to be consistent—separately from deliberate violations. Not only can we see when a mistake occurs, but we can see, from the subject's own perspective, whether the mistake was in the axiom choice or in the lottery choice.

We focus on six fundamental axioms in the domain of risk—independence of irrelevant alternatives, first order stochastic dominance, transitivity, independence, branch independence, and consistency, though we emphasize that our methods could be applied to any axioms. We find that subjects ex-ante want to follow these axioms at high rates—around 85% of subjects desire their choices to be consistent with each of the six axioms. However, as in previous experiments, subjects often violate these axioms in their lottery choices.<sup>2</sup> We give subjects the opportunity to reconcile their axiom and lottery choices when inconsistent. Subjects can make their choices consistent either by changing their lottery decisions or by declaring they no longer want their choices to obey the axiom. Subjects revise about 70% of inconsistent choices, and of these revisions, about 80% change their lottery choices to be consistent with the axioms. We interpret this as subjects treating the axioms as normative and viewing their lottery choices as mistakes.

Our results provide compelling evidence that subjects prefer to choose in accordance with many of the canonical axioms of decision theory. However, a major concern when eliciting whether a subject wants to follow a rule is the presence of experimenter demand effects. In particular, a subject in the experiment may believe that

<sup>&</sup>lt;sup>2</sup>As we discuss in section III, we intentionally chose questions to induce violations.

the experimenter is only offering rules which are helpful to them or that inconsistencies between rules and lotteries should be reconciled in favor of the rules. To prevent this confound from driving the results, we also present the subjects with *not axioms* ( $\neg axioms$ ). The  $\neg axioms$  were designed to be the "opposite" of our main axioms. For example, we present subjects with the rule  $\neg transitivity$ , which says, "If A is preferred to B and B is preferred to C, then C is preferred to A." We use the rates at which subjects select and reconcile these axioms as a benchmark for results on the main axioms of interest.

We find that subjects are much less likely to select the ¬axioms, doing so only about 10% of the time. This confirms that subjects are not simply following the axioms due to experimenter demand effect. Furthermore, when subjects reconcile choices that are inconsistent with the ¬axioms, they are more likely to revise the conflict by abandoning the ¬axiom rather than by altering their lottery choices to conform to the ¬axiom. In doing so, they reveal that following the ¬axiom was the mistake, rather than their lottery choices. This is in stark contrast to the axioms, where a majority of lottery choices are revised to be consistent with the axiom.

In addition to identifying mistakes, this paper contributes to a sparse literature on how individuals perceive axioms. There are two main branches of this literature, depending on whether the axioms are explicitly stated to subjects or not. In our paper, the axioms are explicitly stated and elicited as preferences. The closest papers with explicit axioms are MacCrimmon (1968), Slovic and Tversky (1974), and MacCrimmon and Larsson (1979). However, these studies do not incentivize the participants and do not elicit whether the individuals ex-ante want to follow the axiom. These studies also do not control for demand effects and even try to induce them in some cases. Thus, the results in our paper are the first to show, in an incentivized setting controlling for demand effects, that there is normative appeal to canonical axioms that are used to model risk preferences. In the branch of the literature where the axioms are not explicitly stated, but the choices are designed to elicit attitudes towards the axioms, the closest paper is Benjamin et al. (2019).<sup>3</sup> They also attempt to address demand effects and allow for reconciliation of choices. However, the questions in their paper were designed to be directly applicable to questions about life-cycle saving so the choices could not be incentivized. We describe these related studies in more detail and additional literature in Section V.

<sup>&</sup>lt;sup>3</sup>A version of this appeared earlier in Fontana (2015).

Finally, we view our paper partially as a methodological contribution and proof of concept that opens the door to a number of future research directions. Researchers can use our experimental paradigm to elicit the normative appeal of—and identify mistakes in implementing—axioms, strategies, social choice rules, and many other objects of interest. We purposefully chose simple axioms to study, but one could easily use a similar procedure to study more complicated axioms such as reduction of compound lotteries, the weak axiom of revealed preference, dynamic consistency, time stationarity, among many others. In strategic games, one could use this methodology to elicit whether individuals view obeying dominant strategies and best-responding to beliefs as normative principles for different games, even if they fail to do so successfully in practice. Researchers could also elicit attitudes toward fairness or aggregation rules in the domain of social preferences. In the case of impossibility theorems (e.g. Arrow, 1950), these methods could be used to identify which axioms are the most desirable to relax or abandon. In this way, we view the methods here as a paradigm that can be transplanted to inform other areas of economics.

### II. THEORETICAL FRAMEWORK

Before outlining the experimental design, we define the theoretical framework underlying the experiment. We presented all questions and axioms in the domain of non-negative monetary lotteries. We considered lotteries with US dollars as prizes, with potential outcomes in X = [0,30]. We represent the set of lotteries with prizes in X by  $\Delta(X)$ , with strict preferences  $\succ$  defined over  $\Delta(X)$ . We denote generic prizes in X by x,y,z, and denote generic lotteries in  $\Delta(X)$  by p,q,r,s. We represent the degenerate lottery giving x for sure as x. Lastly, for a set of lotteries, x, we denote the set of lotteries chosen from x as x. We write x to mean x to mean x to x

Throughout the experiment, we study six fundamental axioms:

- 1. Independence of Irrelevant Alternatives (IIA):  $p = C(\{p,q,r\}) \Rightarrow p = C(\{p,q\})$  IIA states that if a lottery p is chosen from the set of lotteries p, q and r, then it is also chosen from the subset p and q.
- 2. First Order Stochastic Dominance (FOSD):<sup>4</sup>  $\forall x$   $1 P(x) \ge 1 Q(x) \Rightarrow p > q$

<sup>&</sup>lt;sup>4</sup>Where P(x) and Q(x) are the cumulative distribution functions to x of p and q respectively. For example,  $P(x) = \sum_{y \le x} p(y)$  where p(y) if the probability of winning prize y.

FOSD states that if the probability of winning a prize greater than x is higher in p than in q, for all prizes, then p will be chosen over q.

- 3. Transitivity (TRANS): p > q and  $q > r \Rightarrow p > r$ TRANS states that if a lottery p is chosen over lottery q, and q is chosen over r, then p will be chosen over r.
- 4. Independence (IND):  $\forall \lambda \quad p > q \Rightarrow \lambda p + (1 \lambda)r > \lambda q + (1 \lambda)r$  IND states that if p is chosen over q, then the mixture of p with any lottery r will be chosen over the equivalent mixture of q with r.
- 5. Branch Independence (BRANCH):  $\lambda p + (1-\lambda)r > \lambda q + (1-\lambda)r \Rightarrow \lambda p + (1-\lambda)s > \lambda q + (1-\lambda)s$ BRANCH states that if the mixture of p and r is chosen over the mixture of q and r, then the preference will not change when r is swapped out for a different lottery, s.
- 6. Consistency (CONS):  $p > q \Rightarrow p > q$  CONS states that if p is chosen over q, then p always will be chosen over q.

In addition to these six main axioms, we included the "opposite" of each axiom (denoted as ¬axioms). The ¬axioms reverse the preference relation after the implication for each of the six main axioms. We use the ¬axioms to test for understanding, control for potential demand effects, and disentangle explanations for selecting different axioms.

Formally, we included the following six ¬axioms:

- 1. Not Independence of Irrelevant Alternatives ( $\neg$ IIA):  $p = C(\{p,q,r\}) \Rightarrow q = C(\{p,q\})$
- 2. Not First Order Stochastic Dominance (¬FOSD):  $\forall x \quad 1 P(x) \ge 1 Q(x) \Rightarrow q > p$
- 3. Not Transitivity ( $\neg$ TRANS): p > q and  $q > r \Rightarrow r > p$
- 4. Not Independence  $(\neg IND): \forall \lambda \quad p > q \Rightarrow \lambda q + (1 \lambda)r > \lambda p + (1 \lambda)r$
- 5. Not Branch Independence (¬BRANCH):  $\lambda p + (1 \lambda)r > \lambda q + (1 \lambda)r \Rightarrow \lambda q + (1 \lambda)s > \lambda p + (1 \lambda)s$
- 6. Not Consistency ( $\neg CONS$ ):  $p > q \Rightarrow q > p$

We also designed six meaningless *distractor rules*, which were over unrelated lotteries. For example, one distractor rule is  $p > q \Rightarrow r > s$  where the lotteries p, q, r,

and s are unrelated. This rule essentially implements a random choice. We used the distractor rules only as a buffer so that subjects were less likely to notice the relationships between the axioms and  $\neg$ axioms. The full list of the distractor rules is in the Supplemental Appendix. When we refer to the axioms,  $\neg$ axioms, or distractor rules as general choice objects, we refer to them as *rules*, which is the language used in the experimental instructions.

#### III. EXPERIMENTAL DESIGN

The experiment consisted of four separate blocks, presented to subjects in the order described below. Within each block, we presented questions in a pre-defined randomized order. In Block 1, we directly elicited preferences over rules. In Block 2, we had subjects choose a lottery from binary decision problems which, unbeknownst to them, were related to the rules. In Block 3, we elicited subjects' willingness to re-evaluate any inconsistencies between their rule choices and lottery choices. In Block 4, we provided reconciliation opportunities for almost all lottery choices that violated one of the axioms or ¬axioms a subject selected in Block 1. In the following subsections, we describe in detail the choices made in each block. Subjects were paid from exactly one randomly selected decision which could have been from any of the four blocks.

We describe the choices and payment mechanisms in detail below. To overview the payments, subjects could be paid for one of four possibilities: original rule choices (Block 1), original lottery choices (Block 2), revised rule choices (Block 4), or revised lottery choices (Block 4). The incentivization procedures are the same for original and revised rules, and are the same for original and revised lotteries. Rules are incentivized by applying them on a set of lotteries and paying subjects what the rule prescribes selecting. If individuals do not want to follow a rule, instead they make the lottery choices themselves. Original and revised lottery choices are incentivized in the standard manner, by paying subjects a realization from the lottery they selected. The choice of which question would be paid is based on random chance as well as

<sup>&</sup>lt;sup>5</sup>Due to complications in programming and incentivizing all of the questions and reconciled choices, we were unable to randomize question order across subjects. Therefore, while we can ensure subjects did not see all FOSD questions in a row, for example, all subjects saw questions in the same order.

 $<sup>^6</sup>$ As we explain below, we did not have subjects reconcile one violation of ¬TRANS, given that it was impossible for us to explain how to make a price-list completely intransitive.

<sup>&</sup>lt;sup>7</sup>All payment uncertainty was resolved using physical randomization devices, in particular two ten-sided dice.

subjects' decisions in Block 3. Subjects were paid at the very end of the experiment, regardless of which decision was selected for payment.

### III.A. Block 1: Rule Choices

The objective in Block 1 was to elicit whether subjects want their choices to satisfy canonical choice axioms. The first challenge lies in making the rules accessible and easy for subjects to understand, while still retaining their broad implications on choices. To address this, the rules for making choices were presented as a form of logic statements. We represented the choice objects in an abstract manner, denoting lotteries by colored circles (see Figure I). Subjects were told that the colored circles represent monetary lotteries but they did not know the exact lotteries associated with each rule. We inform subjects that the lotteries could have payoffs from \$0-\$30, with any probabilities from 0%-100%.

The second challenge is incentivizing the rule choice so that subjects select all of the rules they view as desirable and do not select any others. We did this by asking subjects to decide whether they prefer the rule to make choices for them or whether they would rather make the relevant choices themselves.<sup>8</sup> If the subject chose a rule to make their decisions and it was selected as the payoff-relevant decision, then we applied it to a set of lotteries where the rule has implications and the subject was paid a realization of the lottery prescribed by the rule. Since the rules often require choices as inputs, the subject would answer any relevant "input" questions, and the program would select a lottery as an "output" based on the rule. If the subject instead chose to make the choice themselves, then they would answer both the input and output questions themselves. For any rules not chosen for payment, the subject would never see the input or output lottery choices.

We give a detailed example to help understand the choices in Block 1. Figure I shows how we represent IIA to the subjects. The rule reads as follows: "From the options of black, grey, and white, you chose black. Therefore, from the options of black and grey, we would select black for you." We also included a written description of

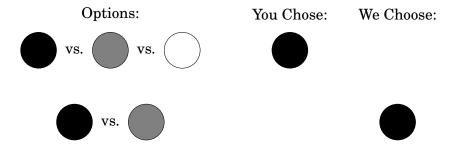
<sup>&</sup>lt;sup>8</sup>Subjects were not allowed to choose between these two options until at least 30 seconds had passed. This design feature encourages subjects to consider the rules carefully before deciding.

<sup>&</sup>lt;sup>9</sup>In the experiment, we used colors—red, orange, green, and blue—rather than black and white alternatives. Colorblind subjects had the opportunity to take the show-up payment instead of participating, but we did not have any subjects take this option.

<sup>&</sup>lt;sup>10</sup>For ¬IIA, the implication instead would be "Therefore, from the options of black and grey, we

the rule on the subjects' screens under the abstract depiction. 11

If this rule were chosen for payment and the subject chose the rule to make their decisions, then we would pay them according to the following procedure. First, there is a bank of three lotteries corresponding to "black," "grey," and "white." Subjects knew that these could be any possible lotteries with payments from \$0–\$30, but only saw the specific lotteries in the event this rule was chosen for payment. The subject would choose their most preferred of the three. This represents the "black" lottery. Following this choice, we would select "black" over "grey" and pay them the realization of the "black" lottery. If they did not select the rule to make their choice, they first would choose their most preferred of the three lotteries "black," "grey," and "white." Following this, we presented them with a decision between the lottery they chose ("black") versus one of the other two lotteries ("grey"). Here the subject would have been paid the realization of the lottery they chose in this second decision. <sup>12</sup> In the Supplemental Appendix, we show how we represent the other five axioms in rule format. Combining the axioms, ¬axioms, and distractor rules, subjects made eighteen decisions in Block 1.



**Figure I:** Rule representation of IIA

<u>Notes:</u> We represent axioms in this manner. Colored circles represent any possible lotteries with payoffs from \$0–\$30. Subjects choose whether to have this rule make choices for them or instead make choices on their own.

We incentivized the rules this way to encourage subjects to select exactly the ax-

would select grey for you."

<sup>&</sup>lt;sup>11</sup>We described the rule using fruits which match the colored circles. For example, we say "If you choose a blueberry over an apple and a grape, then we would choose a blueberry over an apple for you." Subjects are reminded, though, that the circles represent money and not consumption goods, and there are no complementarities or tastes involved in mixtures.

<sup>&</sup>lt;sup>12</sup>During payment, subjects were not told which rule was being implemented before they made the input choices. If we had told them which rule was being implemented, they could answer the input questions "opposite" their true preferences for the ¬axioms and still receive their truly-preferred alternative.

ioms which *always* should be true on the relevant domain of lotteries. Thus, subjects should select the rules that they want to describe all of their choices, regardless of the specific lotteries and payments. However, a worry is that subjects might be indifferent between selecting a rule and making the choice on their own. A subject who does select a rule reveals that they want to make decisions according to the rule, since it can be applied over any lotteries. The interpretation is less clear for subjects who do not select a rule. A subject who agrees with a rule but believes their choices will align with the rule anyway has no strict incentive to select the rule, aside from the time and effort cost of making choices. In Appendix Section B, we present results from another treatment where subjects had to pay a small cost, \$1, to make the choice on their own. This makes it strictly costly not to select a rule, eliminating this concern. We find that the rules are selected slightly more often in this treatment, responding to the incentives, but all qualitative results remain unchanged. This indicates that the vast majority of subjects who do not choose a rule aren't doing so out of indifference, but rather because they do not find it desirable.

### III.B. Block 2: Lottery Choices

Given that our main interest is in studying how individuals reconcile inconsistent choices, we selected lottery questions from previous papers which found violations of the axioms. We do not focus on the specifics of the lotteries here, but we picked questions to maximize axiom violations. Our intention is not to compare violations across axioms since the questions are not representative of the axiom more generally. A description of references and violations for each question can be found in Table I, along with the question numbering that we use to present the results. The full set of questions can be found in Appendix C.

The payment amounts span from \$0 to \$30, and lotteries range in expected value from \$1.40 to \$26. Within each question (with either two or three lotteries), the difference in expected value between the lotteries ranges from \$0 to \$6. On average, the expected value difference is just shy of \$1.75. Furthermore, these lottery questions are not the same as the lottery questions that incentivized Block 1 rule choices. This

<sup>&</sup>lt;sup>13</sup>Here, however, the interpretation is less clear for subjects who do select a rule. A subject who selects a rule does not necessarily indicate they *always* want to follow the rule. It could be that they want to follow the rule "most" of the time, so in expectation, they believe it is not worth paying \$1 to make choices on their own.

Question	Reference		
IIA1–IIA3	Huber et al. (1982) (Attraction Affect)		
IIA4	Huber et al. (1982) (Compromise Effect)		
FOSD1-FOSD4	Birnbaum and Martin (2003)		
TRANS1-TRANS2	Loomes et al. (1991) (Regret Theory)		
TRANS3	Brown and Healy (2018) (Multiple Switch Points)		
IND1	Birnbaum and Chavez (1997)		
IND2	Kahneman and Tversky (1979) (Certainty Effect)		
IND3	3 Jain and Nielsen (2019) (Reverse Certainty Effect)		
BRANCH1	Birnbaum and Chavez (1997)		
CONS1-CONS2	Brown and Healy (2018) (Near Indifference)		

**Table I:** Description of Questions

was to ensure that subjects do not update on the type of lotteries where the rules would apply. We briefly describe the source of the questions.

INDEPENDENCE OF IRRELEVANT ALTERNATIVES — We used four IIA questions similar to Huber et al. (1982). Three of the four questions targeted violations of IIA by adding a dominated lottery to a binary choice problem to "attract" the subject to the dominating alternative. We refer to these as IIA1, IIA2, and IIA3. The fourth question, IIA4, targeted a violation of IIA by adding a lottery to a binary choice problem that makes one of the initial two lotteries a "compromise" option.

TRANSITIVITY — We used two transitivity questions (TRANS1 and TRANS2) similar to Loomes et al. (1991) that were used to examine regret theory. In addition, we included six binary questions which together comprised a separated price list, as demonstrated in Brown and Healy (2018). These six questions involved binary comparisons between a risky lottery and a sure payment. The risky lottery was the same in all six questions while the sure payment varied. Multiple switch points on the price list constitute a violation of transitivity, which we refer to as TRANS3.

FIRST-ORDER STOCHASTIC DOMINANCE — We asked four binary questions to target violations of FOSD (FOSD1–FOSD4). All four questions followed the structure in Birnbaum and Martin (2003).

INDEPENDENCE — We included three questions that targeted violations of independence, including one question from Birnbaum and Chavez (1997) (IND1), one

 $<sup>^{14}</sup>$ For example, one question was a binary choice between lottery p and \$14.00. Another was the choice between p and \$14.50, another between p and \$15.00, etc. Presenting questions of this form is a common procedure used to elicit a certainty equivalent for lottery p.

question from (Kahneman and Tversky, 1979) demonstrating the certainty effect (IND2), and one question from Jain and Nielsen (2019) demonstrating the reverse certainty effect (IND3).

BRANCH INDEPENDENCE — We included one question targeting a violation of Branch Independence from Birnbaum and Chavez (1997) (BRANCH1).

CONSISTENCY — We included two questions to target violations of consistency in which we asked two binary decision problems that were each repeated twice (CONS1 and CONS2).<sup>15</sup> We chose the binary decision problems based on the questions that were nearest to indifference in Brown and Healy (2018).

We displayed the lotteries simply by reporting the probabilities and payoffs of each possible outcome. Subjects saw the lotteries on their screens as below and made their choices by selecting the rectangular button corresponding to their preferred option. Altogether, subjects make choices from thirty-three binary or trinary decision problems in Block 2.



**Figure II:** Representation of lotteries

We chose to standardize the presentation of the lotteries to ease subjects' understanding and to allow for clean comparisons across axioms. The drawback, however, is that our results are not directly comparable to the original papers due to differences in framing. For example, Loomes et al. (1991) presented lotteries in a table format that displayed probabilities as the length of a given cell. It is possible that these different presentation formats invoke different behavioral responses, so our experiment does not directly replicate previous papers. <sup>16</sup>

One issue when studying preferences over axioms is that if an individual views

<sup>&</sup>lt;sup>15</sup>The questions are not repeated back to back.

<sup>&</sup>lt;sup>16</sup>Indeed, we find fewer violations of transitivity than Loomes et al. (1991). The most likely explanation is that their presentation of the decision problem induced stronger feelings of regret, leading to more violations.

many axioms as normative, the axioms combined may have unexpected implications across questions. For example, continuity, completeness, transitivity, and independence together imply that there is an expected utility representation consistent across all questions. To ameliorate this concern, we chose lotteries so that the questions have no implications across multiple axioms. In particular, we did not use any of the same lotteries to target more than one axiom. The lotteries used in FOSD1, for example, are completely distinct from lotteries used in any other question. Therefore, there are no immediate implications of the axioms across questions, and each decision applies only to its targeted axiom.

We make no assumptions on preferences except for dominance in degenerate lotteries, i.e.  $\delta_x > \delta_y$  if and only if x > y.<sup>17</sup> In using the random problem selection payment mechanism, we also assume a form of monotonicity in the space of two-stage lotteries (Azrieli et al., 2019). Brown and Healy (2018) give evidence that this condition is met in a risky choice experiment similar to ours.

### III.C. Block 3: Willingness to Reconcile

To elicit willingness to reconcile inconsistent choices, we had subjects rank any of the six main axioms they selected against a \$1 outside option. For example, a subject who selected all six axioms would see seven boxes on their screen—one for each of the axioms, and one with an option that says "I would rather have \$1 than reconcile choices associated with any of the remaining rules." Subjects first select the axiom they would most want to reconcile should their choices violate it, or select the outside option if they would rather have \$1 than reconcile their choices. Then, subjects select the axiom they would next-most want to reconcile among the remaining axioms, and so on.

If the ranking were chosen for payment, we would randomly select two of the available axioms or the outside option and pay the subject according to the Block 4 reconciled choices from whichever axiom they ranked higher. For example, take a subject who ranks the axioms in the order FOSD > IIA > TRANS > CONS > Outside Option > BRANCH > IND. We would randomly select two options, say IIA and the Outside Option. The subject ranked IIA higher, so they would be paid based on their recon-

<sup>&</sup>lt;sup>17</sup>This assumption allows us to represent multiple switches on a price list as a violation of transitivity. It is also relevant for FOSD violations.

ciled choices in the IIA question, as described below.<sup>18</sup> Regardless of their ranking, a subject still faces all of the reconciliation decisions in Block 4; the ranking only impacts which would be paid.

Technically, the reconciled choices for the last-ranked axiom would not be incentivized in this procedure. The reconciled choices of last-ranked axiom would never be implemented since we implement the reconciled choices of whichever axiom is ranked higher. To ensure incentive compatibility, there's an independent chance that we would randomly select the reconciled choices to pay, as described below. This means that the reconciled choices are almost twice as likely to be paid as the original choices. Therefore, if anything, subjects should be more concerned with their reconciled choices than their original choices. This also helps encourage subjects to carefully consider the reconciliation opportunity.

The main purpose of the ranking is to see whether subjects strictly prefer to reevaluate their choices. This gives us only a coarse measure—whether individuals are willing to give up at least \$1. We felt this would be easier for subjects to understand than trying to elicit a finer willingness-to-pay for each axiom. We also use the reported rankings to look for any consistent patterns within rankings. <sup>19</sup> This gives us a finer measure of subjects' perceptions of the rules compared to the binary ranking in Block 1.<sup>20</sup>

### III.D. Block 4: Reconciliation

After subjects ranked all of the axioms, we presented them with every inconsistency between their lottery choices and selected axioms or ¬axioms. For example, a subject who selected IIA in Block 1 but violated IIA with their lottery choices in Block 2 saw these choices side by side on their screen. We color-highlighted the original choices that the subject made. We color-coded the choices to match the colors in the rule so the subject could better understand how the rule mapped onto their choices. We

<sup>&</sup>lt;sup>18</sup> If they had ranked the outside option higher, they would be paid for their original choices in the IIA question and would receive an extra \$1 bonus. If they did not violate IIA, they would be paid for their original choices and would receive a \$1 bonus, which ensures their ranking is not affected by their perceptions of which axioms they were more likely to have violated.

<sup>&</sup>lt;sup>19</sup>With this procedure, we cannot rule out that subjects are indifferent among the axioms or to the outside option. We look for systematic patterns in the rankings, but acknowledge this as a shortcoming of our ranking procedure. We also cannot rule out that rankings are driven by subjects' beliefs about the expected value differences of the associated lotteries.

 $<sup>^{20}</sup>$ Eliciting a ranking over axioms is also similar to the work of MacCrimmon and Larsson (1979).

also included a color-coded explanation of why choices violated the rule and how they could be reconciled. We provide a screenshot in the Appendix A and reproduce an example below in Figure III.

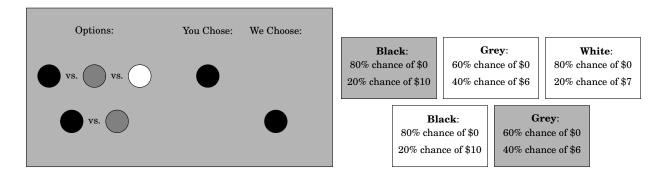


Figure III: Example of Reconciliation Screen

Notes: The options highlighted in grey indicate subjects' original choices in Blocks 1 and 2. For example, this subject selected IIA in Block 1, but chose "black over grey and white" in one question and chose "grey over black" in another question. Below this, subjects saw an explanation of why the rule would have selected something different than what they chose for themselves.

Subjects could change any of their choices, or could leave them as they were. We impressed upon subjects that they could change any of their lottery choices to match the rule, could unselect the rule, or could leave choices inconsistent. As a result, there was no default direction for any potential experimenter demand effect. Consistency in choices could arise equally by unselecting rules or by changing lottery choices.

When a reconciled choice was selected for payment, we paid subjects for the rule or lottery in the same manner as the original choices were paid in Blocks 1 and 2. That is, if a reconciled rule were chosen for payment and the subject kept the rule selected, then we would apply this rule on a set of lottery questions.<sup>21</sup> If instead the subject unselected the rule during the reconciliation opportunity, then they would make the lottery choice themselves. If reconciled lottery choices were selected for payment, then the individual would be paid the lottery ultimately chosen in the reconciliation stage. Furthermore, choices which were not violating a rule also could be paid again as reconciliation choices to maintain equal probability of all rules and lotteries being paid. In this case, we paid the subject based on their original rule or lottery choice.

<sup>&</sup>lt;sup>21</sup>We did not specify one way or another to subjects, but this set of lottery questions was different from the set of lottery questions incentivizing Block 1 rule choices.

Subjects reconciled choices inconsistent with each of the six axioms and the six ¬axioms. Subjects reconciled each violation independently. That is, a subject who selected IIA and violated it on two separate occasions had two separate opportunities to reconcile the violations, rather than reconciling all choices together. We did this to encourage subjects to analyze each choice in isolation. This allows us to see whether some types of violations are reconciled in favor of rules while others are reconciled in favor of lotteries. We also had subjects reconcile inconsistencies within their rule choices. For example, a subject who chose both IIA and ¬IIA in Block 1 would also see these rules side by side on their screen and choose which, if any, to keep selected. These decisions were incentivized in the same way as other revised rules.

The number of reconciliation opportunities varied per subject, based on number of violations and on number of axioms and ¬axioms selected. On average, subjects had 5.5 reconciliation opportunities. The number of the reconciliations ranged from zero to twenty-two.

We analyze data from 110 subjects, primarily undergraduate students at the Ohio State University where the sessions took place. We programmed the experiment using z-Tree (Fischbacher, 2007), and recruited subjects using ORSEE (Greiner, 2015). Sessions lasted about one hour, and subjects earned \$15 on average, including a \$7 show-up payment. Instructions are included in a Supplemental Appendix.

### IV. RESULTS

Figure IV shows the percentage of subjects who selected each axiom in Block 1. We break this down by whether a subject selected the axiom only, the axiom and the ¬axiom, only the ¬axiom, or neither. FOSD is most popular, selected by 90% of subjects overall. For the remaining axioms, 85% select Consistency, 83% select Transitivity, 83% select IIA, and 83% select Independence, and 82% select Branch Independence. The axioms are much more often selected than the ¬axioms. In

 $<sup>^{22}</sup>$ We did not have subjects reconcile  $\neg TRANS$  with the price list, as we could not explain how to make the price list completely intransitive. We did not have subjects reconcile the meaningless distractor rules given that there is no natural way to present the violating choices.

<sup>&</sup>lt;sup>23</sup>This also means that the reconciliation was not dynamic. That is, a subject who selected both IIA and ¬IIA in Block 1, and then violated IIA in Block 2, may have reconciled these choices to be consistent with IIA. In doing so, this might lead them to be inconsistent with ¬IIA! We did not present them this subsequent reconciliation. The reconciliation opportunities were fixed and determined by their choices in Blocks 1 and 2.

particular, 15% selected ¬Branch Independence, 14% selected ¬Independence, 12% selected ¬IIA, 11% selected ¬Transitivity, 10% selected ¬FOSD, and 5% selected ¬Consistency.

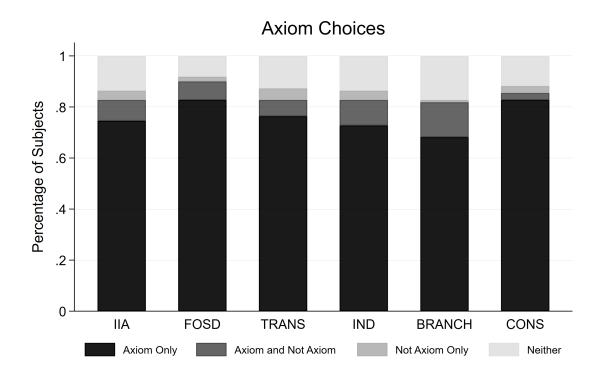


Figure IV: Percentage of Subjects Selecting Each Rule in Block 1

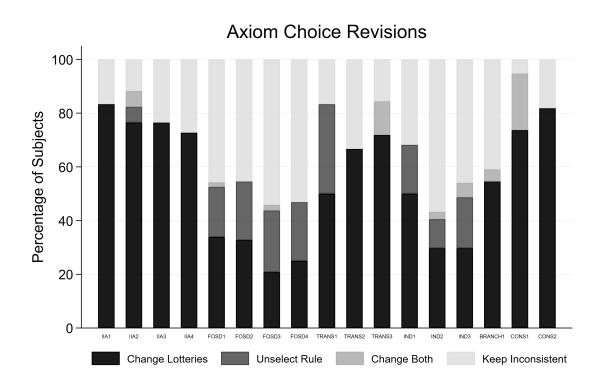
Interestingly, FOSD is the most popular axiom while ¬FOSD is among the least popular ¬axioms. Similarly, Branch Independence is the least popular axiom while ¬Branch Independence is the most popular ¬axiom. This might indicate that there are some patterns to how subjects perceive the axioms. FOSD is most "obviously" desirable, and therefore ¬FOSD is obviously not desirable. The opposite is true for Branch Independence. This suggests some features of axioms might be more or less compelling to individuals. It is also interesting that both FOSD and Branch Independence involve "mixing," so it is not that case that individuals are just averse to, or confused by, mixing. From our data, however, we cannot say whether this is driven by differences in understanding of the axioms or differences in preferences over the axioms.

Table IX in the Appendix shows the number of axioms and ¬axioms selected on an

individual level. 60% of subjects selected all six axioms, and 65% of subjects never selected a ¬axiom. Among individuals who ever select a ¬axiom, it is most common for individuals to select only one of them. Therefore, we have confidence that subjects understand the decision rules and incentivization procedure. In the Appendix, we report correlations of selecting the various axioms. All pairwise correlations are significantly positive, and we do not detect any obvious patterns.

Overall, we conclude that individuals do view these axioms as desirable methods for choosing lotteries since they are chosen at much higher rates than the ¬axioms. While these same individuals will go on to violate the axioms in their choices, they express a preference for their choices to satisfy these axioms as a general principle.

**Result 1.** Nearly all individuals reveal a preference for their choices to satisfy canonical choice axioms. These axioms are selected at much higher rates ( $\approx 85\%$ ) than their negations ( $\approx 10\%$ ).



**Figure V:** Percentage of Subjects Revising Choices in Block 4, Conditional on Selecting Axiom

Given that subjects prefer to satisfy these axioms, the question remains as to how

individuals respond when their choices do not. Among those who select the respective axiom, 85% of subjects violated FOSD, 75% violated Independence, 46% violated Consistency, 43% violated transitivity, 38% violated IIA, and 24% violated Branch Independence. Over 85% of subjects who violated an axiom indicated in Block 1 that they wanted to follow the axiom, so "wanting" to follow a rule does not ensure that a subject can or will follow the rule. Our main focus is whether these subjects reconcile their choices in Block 4 and whether they reconcile by changing their lottery choices or by abandoning the decision rule.

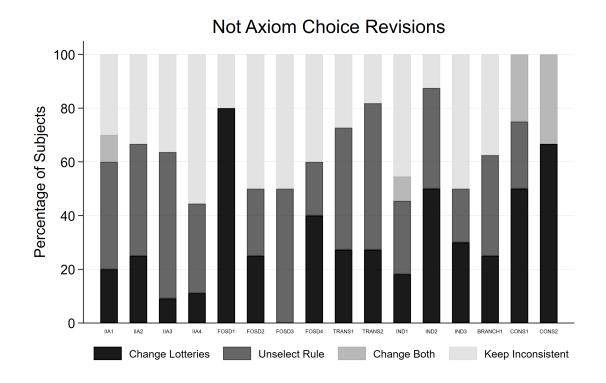
Figure V shows choices in the reconciliation block (Block 4), broken down by question. A non-trivial percentage of inconsistencies are not reconciled, which we discuss below. Of those who do change their choices, it is far more common for individuals to change their lottery choices to be consistent with the axiom than to unselect the axiom. Across all questions, 79% of revisions are in favor of the axioms, rather than the opposite. Naturally, this is significantly different from a 50-50 split in the direction of revisions (p = 0.0000).

Table II breaks this down by axiom. In column two, we report the percentage of individuals who changed any of their choices in the reconciliation. In columns three and four, we report the conditional percentages of revising the lottery versus revising the axiom, excluding those who revise both.<sup>24</sup>

		Conditional	Conditional
Axiom	% Revised Either	% Revise Lottery	% Revise axiom
Any (n=782)	63%	79%	21%
IIA (n=116)	76%	99%	1%
FOSD (n=359)	44%	66%	34%
TRANS (n=46)	80%	94%	6%
IND (n=159)	50%	78%	22%
BRANCH (n=22)	59%	100%	0%
CONS (n=80)	91%	100%	0%

**Table II:** Percentage of Violations Revised and Direction of Reconciliation Notes: The second column gives the percentage of violations that were reconciled. The third column reports the percentage of these reconciliations where subjects revised their lottery choices, and the last column reports the percentage of these reconciliations where subjects revised their rule choices.

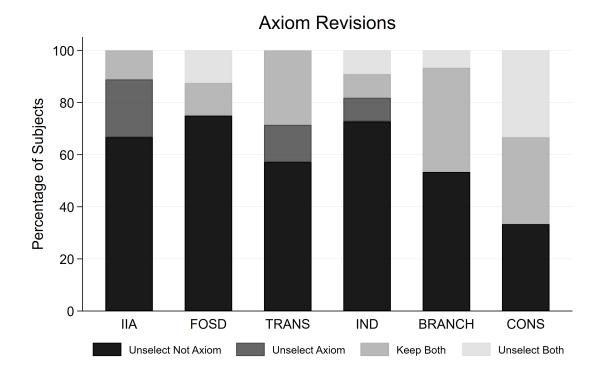
 $<sup>^{24}</sup>$ Note, the sample sizes vary widely as individuals violated some axioms more than others, and some axioms had more questions than others. For example, there were four FOSD questions, and 85% of subjects violated FOSD at least once. On the other hand, there was only one Branch Independence question and 24% of subjects violated the axiom.



**Figure VI:** Percentage of Subjects Revising Choices in Block 4, Conditional on Selecting Opposite Axiom

We use the reconciliation of the ¬axioms as a benchmark for the direction of reconciliation. We might worry that individuals change their choices to match the rules as a form of experimenter demand effect. We tried to eliminate this in the design—since subjects could make their choices consistent either by changing their lottery choices or by un-selecting the rule, there is no obvious direction for the demand effect. However, the description of the rule violation informed subjects of how they could make their lottery choices consistent with the rule, so changes in lottery choices could reflect subjects simply "following instructions" rather than a real desire to choose consistently with the axiom. We will see that this is not the case, as individuals are less likely to reconcile their choices in favor of the ¬axioms.

Figure VI shows the same breakdown of revised choices conditional on subjects selecting the ¬axioms. Across all questions, 44% of revisions are in favor of the ¬axioms, not significantly different from a 50–50 split (p = 0.311). This is in significant contrast to the 79% of instances where subjects reconcile in favor of the main axioms (Wilcoxon Ranksum p = 0.0000).



**Figure VII:** Percentage of Subjects Revising Choices in Block 4, Conditional on Selecting axiom and ¬axiom

One might still worry, however, that individuals who select the  $\neg$ axioms are systematically different from those who select the axioms. For instance, they might not understand the rule representations or payment incentives. Differences in reconciliation patterns could be driven by these factors, but we have no reason to believe this type of noise would be systematic. We can rule this out further by looking at individuals who select both an axiom and corresponding  $\neg$ axiom. Conducting the same analysis as above, restricted to this subsample, we find the results unchanged. When reconciling violations of the axioms, individuals revise their choices 69% of the time, and 65% of revisions follow the axiom. When reconciling violations of the  $\neg$ axioms, these individuals revise their choices 72% of the time, and only 36% of revisions follow the  $\neg$ axiom, which is significantly lower than for the main axioms (p = 0.0149).

Furthermore, Figure VII shows the rule reconciliation pattern for these individuals. That is, we look to see how individuals reconcile their rule choices when faced with both the axiom and ¬axiom. Within that sample, we find individuals still favor the main axioms. Among individuals who un-select one of the rules, over 89%

of them un-select the ¬axiom. That is, when individuals are faced with two decision rules that prescribe opposite choices, they realize this and abandon the less-sensible rule.

We conclude that a majority individuals who wanted to follow the axioms but violated them made a mistake in their lottery choices. There is a baseline tendency to keep choices inconsistent (as we will discuss below, in Result 4). However, among reconciliations, the axioms are usually followed.

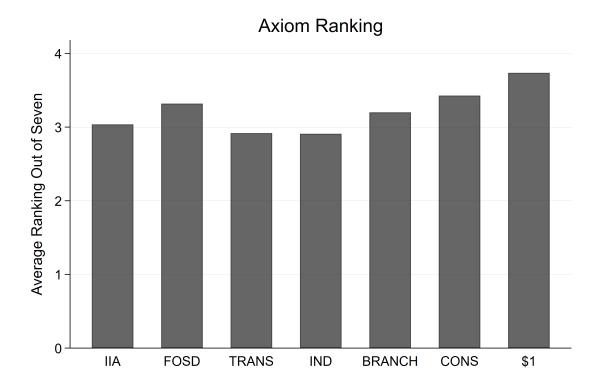
**Result 2.** Individuals violating canonical axioms often change their choices to be consistent with the axiom ( $\approx 79\%$  of revisions). Individuals violating  $\neg$ axioms are much less likely to do so ( $\approx 44\%$  of revisions).

Finally, we look at how subjects ranked the axioms in Block 3. We find that, for each of the six axioms, about 60% of subjects are willing to give up at least \$1 to re-evaluate their inconsistent choices. We find that 66% of subjects rank Branch Independence before the \$1 outside option, and this is similar at 62% for Independence, 62% for FOSD, 62% for IIA, 60% for Transitivity, and 56% for Consistency. The pairwise comparison between Consistency and Branch Independence is significant (p = 0.041) but all other pairwise comparisons show insignificant differences in willingness to pay (Wilcoxon Sign Rank tests, p > 0.134). On an individual level, 85% of subjects are willing to give up \$1 to re-evaluate choices corresponding to at least one axiom.

Looking at the average ranking of each axiom, again we find very few differences. Figure VIII presents the average ranking for each of the alternatives. Transitivity and Independence are ranked lowest (most preferred), while the outside option is the least preferred, on average. There are some significant differences across axioms, but all axioms are preferred to the outside option on average. Therefore, it seems that subjects do not have any systematic preferences among the axioms they wish the follow.

**Result 3.** A majority of subjects ( $\approx 85\%$ ) are willing to pay for the opportunity to reconcile their choices when choices conflict with an axiom they want to follow.

Finally, Figures V, VII, and VI show that about 30% of subjects keep their choices inconsistent across these revision opportunities. Inconsistencies with the axioms are revised 63% of the time and inconsistencies with the ¬axioms are reconciled 67% of



**Figure VIII:** Average Ranking of Each Axiom Notes: Lower ranking corresponds to the rule being more preferred.

the time (p = 0.37). This is true even among individuals who have a strict preference for re-evaluating their decisions. While this might seem odd at first blush, there are a few reasons why individuals might keep their choices inconsistent. The most obvious to us is simple effort cost. Subjects have already thought about these decisions and chosen what they prefer. Revising choices is costly in terms of time and cognitive effort, and individuals may view the cost as too high.

Again, we view this as a baseline in our experimental design. Across the axioms, ¬axioms, and rule reconciliations, there is a latent minority of choices which go unreconciled. This could result from a number of factors, and our experiment is not designed to disentangle these, but we see clear patterns among those who do reconcile their choices. It would be interesting for future research to better understand more about the choices kept inconsistent.

**Result 4.** Individuals keep their choices inconsistent in about one third of all reconciliation opportunities.

While we could look at comparative statics of the reconciliations, for example whether revised choices become more or less risk averse, our experiment is not designed for these questions. We chose the lottery questions in order to maximize violations of the axioms, and therefore the questions are in no sense representative of the violations and revisions we might see more generally. However, we believe our methodology could be very useful in answering these types of questions in future research. As a step in this direction, see Benjamin et al. (2019) who study risky investment decisions before and after reconciliation opportunities.

### V. LITERATURE REVIEW

Naturally, our paper relates directly to the large literature on violations of decision theory axioms. Many papers, in addition to those cited in Section III, have found violations of the axioms we study. We confirm these violations, but our experiment allows us to identify a majority of these violations as mistakes. Individuals *do* want their choices to satisfy these axioms, and, when given the opportunity for reflection and revision, individuals often prefer to change their choices to be consistent with the axiom.

We are not aware of any papers that incentive-compatibly elicit the axioms an individual wants to follow. Even without incentives, we are not aware of any papers that elicit axioms and allow for reconciliation with violating choices, which is necessary to detect whether an individual made a mistake. There are three related papers in economics which come closest to our motivation: MacCrimmon (1968), Slovic and Tversky (1974), and MacCrimmon and Larsson (1979). Since our work builds off these pioneering studies, we describe them in detail.

MacCrimmon (1968) asked questions to thirty-eight business executives designed to induce potential violations of five postulates of decision theory. After making decisions, the postulates of choice were discussed with the executives and the executives were allowed to change their choices during the discussion. Within the experiment, most executives were either consistent with the postulates of decision theory or changed their choices in favor of the postulates after some discussion, consistent with our results. While an ambitious study, the unstructured nature of the discussion, lack of incentives, and specialized subject pool are limitations to external validity of the results. Furthermore, the postulates were presented to subjects by the

experimenter in the reconciliation opportunity. A main contribution of our experiment is that subjects themselves choose both the axioms and the lotteries, so the reconciliation is less prone to experimenter demand effects and other outside motivations.

In a related study, Slovic and Tversky (1974) ask unincentivized lottery questions related to the Allais and Ellsberg paradoxes to undergraduate students. After making their decisions, subjects were presented with "advice" in the form of explained arguments for and against the independence axiom and the sure thing principle. Subjects were told that the arguments came from prominent decision theorists. After reading the arguments, subjects were asked to reconsider their original lottery choices. Overall, they find few subjects changing their decisions to be consistent with the advice. In addition to being incentivized, a contribution of our design is to elicit individuals' preferences over these axioms, rather than presenting them as advice or arguments from others. This allows individuals to see the inconsistencies within their own preferences, rather than deciding whether to follow the logic from an outsider. This could be one explanation for why our subjects are more likely to revise their choices compared to Slovic and Tversky's.

MacCrimmon and Larsson (1979) asked graduate students to indicate their level of agreement with 20 decision rules related to the axioms of expected utility theory. Additionally, they had subjects make unincentivized lottery choices and indicate any decision procedures involved in making the choices. They find that individuals view the decision rules favorably but often violate them. In this experiment, there is no reconciliation stage. We design our experiment to carefully incentivize these choices and to study how individuals reconcile discrepancies between their choice of lotteries and decision rules.

A number of contemporaneous papers study reconciled choices from disparate motivations. Benjamin et al. (2019) develop a survey procedure to reduce framing effects when measuring risk aversion, in order to get better risk measurements to inform retirement saving policies. They have subjects make hypothetical life-cycle savings choices which are related to one another according to various consistency principles. They ask subjects to reconsider their conflicting choices and rate whether the situations are sufficiently different to warrant choosing different options. Breig and Feldman (2019) ask subjects a number of risky budget set decisions, each repeated twice. Subjects are given the opportunity to revise their choices with both choices

shown on the screen. Individuals often revise, and revised choices are slightly closer to rationality according to standard revealed preference measures. In this study, the decisions are most closely related to our Consistency axiom. Crosetto and Gaudeul (2019) study the asymmetric dominance effect and find that the effect disappears after subjects are given time and incentives to revise their choices.

Though very different environments, our paper is also related to the literature studying strategies in repeated games. Romero and Rosokha (2018), Cason and Mui (2019), and Dal Bó and Fréchette (2019), among others, allow subjects to design comprehensive strategies in indefinitely repeated prisoner's dilemma games, rather than choosing actions each period. While our subjects do not design their own "axioms" to follow, our paper can be thought of as a similar *procedural* experiment where subjects choose rules to implement decisions for them. This is similar to the distinction between substantive and procedural rationality as outlined in Simon (1976), who calls for economists to "become interested in the procedures—the rational processes—that economic actors use to cope with uncertainty" (Simon, 1976, p.81). Halevy and Mayraz (2020) take a step in that direction, allowing subjects to design procedures to carry out their investment decisions. They find that subjects prefer using procedures to making decisions on their own, which is similar to our subjects' preference for following the axioms.

### VI. DISCUSSION

We present the first incentivized experimental evidence supporting the view that canonical choice axioms have normative content and that violations of axioms represent mistakes. In directly eliciting preferences over axioms, we find that individuals view them as rules that they want their choices to follow. When lottery choices conflict with stated axiom preferences, individuals change their choices to be consistent with the axiom, rather than inferring from their choices that the axiom is not desirable.

Our experiment takes a step towards identifying individuals' choices as "preferences" versus "mistakes," but also highlights the difficulties in doing so. The evidence suggests that most subjects do view these axioms as desirable and change choices accordingly, leading us to interpret their inconsistent lottery choices as mistakes. However, a substantial minority of individuals do not change their choices despite

wanting to follow the axiom. In this case, it is not obvious how to declare either the axiom or lottery decisions as preferences or mistakes in these cases. However, these situations might not be surprising. For example, Gilboa et al. (2009) argue that it is natural to encounter situations where a preference for a decision rule conflicts with preferences over a single decision problem. Sometimes individuals resolve a conflict by adhering to the rule (axiom) and other times by adhering to their decisions, but neither needs to be abandoned in general. Subjects in our experiment who conflict in their rules and choices demonstrate that these cases occur in practice.

This experiment also demonstrates the difficulty of making welfare statements from choices. For example, whether or not a subject reconciles an inconsistency often depends on the question that is asked. One case of this occurs for the three questions related to independence. For question IND1, subjects are more likely to revise their choices to satisfy independence than IND2 or IND3. Thus, using independence to evaluate welfare may depend on the type of question asked. Other issues of measuring welfare are brought up in Bernheim and Rangel (2009) and focus on the difficulty of making welfare statements in the presence of framing. In contrast, all of our questions are framed the same way so we demonstrate issues of making welfare statements even holding the framing of questions fixed.

We believe our results provide implications for both theoretical models and experimental design, and we discuss these both in turn. We also see a number of open directions for future research, and we outline these explicitly.

### VI.A. Implications for Theory

Our results suggest a role for economic theory to model preferences over axioms alongside modeling the choices individuals make. Our experiment elicits two revealed preferences—one over axioms and one over lotteries. These preference relations do not always align, resulting in violations of the axioms. However, results suggest that, in these cases, the preference over axioms supersedes the lottery preference. Our results suggest that individuals do have preferences over axioms directly, so it might prove fruitful to incorporate these preferences into theoretical models. This would provide structure to exploring the interaction between axiom preferences and choices. While there is little theoretical work that explicitly models different types preferences that are related, one example is Gilboa et al. (2010) which models

the relation between objective and subjective preferences.

Our results also contribute to an interesting discussion on the role of decision theory, outlined in Gilboa (2010), who writes:

"We are equipped with the phenomenally elegant classical decision theory and faced with the outpour of experimental evidence á la Kahneman and Tversky, showing that each and every axiom fails in carefully designed laboratory experiments. What should we do in face of these violations? One approach is to incorporate them into our descriptive theories, to make the latter more accurate. This is, to a large extent, the road taken by behavioral economics. Another approach is to go out and preach our classical theories, that is, to use them as normative ones... In other words, we can either bring the theory closer to reality (making the theory a better descriptive one) or bring reality closer to the theory (preaching the theory as a normative one). Which should we choose?"

Our results call for the latter, and suggest that individuals already view the classical theory as normative in many instances. For a majority of individuals, violating canonical axioms is revealed irrational by their own choices. We help individuals make better decisions, according to their own preferences, when we assist them in satisfying these axioms. This is not to diminish the role of descriptive theories, but to draw attention to the different roles that descriptive and normative theories may play.

### VI.B. Implications for Experiments

Given that we find individuals making mistakes, it's unclear how elicited preferences would change after individuals are given the opportunity to "correct" their choices. For example, are revised choices systematically more/less risk averse than original choices? As mentioned above, we chose specific questions that would result in violations of the axioms, so our experiment was not designed to answer these questions. However, we can speculate on the implications. For example, one of our questions (TRANS3) comprised a separated price list. Price lists are common tools for experimental elicitation, and there is no general consensus as to whether experimenters should enforce a single switching point (see discussion in Brown and Healy (2018)). We find that over 70% of subjects who exhibit intransitivites in the form of multiple switching points change their choices to be transitive. This suggests that enforcing a single switching point might actually help subjects express their underlying desire

for transitivity.

In this vein, it might be the case that there are other features of the environment or experimental design that cause axiom preferences and choice preferences to align. For example, experimental interfaces could notify subjects when they violate consistency or independence of irrelevant alternatives. If we want to elicit subjects' "rational" preferences, or the choices that they do not regret, this might require more study into the structure and design of experiments.

More generally, our results suggest caution in designing and interpreting experimental tests of axioms. While it is possible to design choices to induce violations of nearly any axiom, researchers (ourselves included) should think carefully about whether these violations are revealing meaningful preferences or whether we are simply setting "traps" for subjects to fall into. If we "trick" subjects into making choices they regret, this might not be the kind of evidence we want to use to design economic models.

### VI.C. Directions for Future Research

We view our experiment as one in a new line of experiments in procedural choice. We see many interesting directions in which to take this agenda, and we outline a few below.

In our experiment, people tend to follow "rules" over following choices. It would be interesting to understand more about when and where this is true. It is also interesting to what aspects of the environment (e.g. framing) alters an individual's perceptions of decision rules. In a somewhat-related study, Oprea (2019) analyzes aspects of the decision environment that make rules more complex to implement. It would be interesting to understand more about how these measures of complexity interact with the questions we answer in our paper. For example, what features make axioms more complex to understand? Are more complex axioms less appealing? Does the complexity of the environment (here, relatively simple lottery choices) affect the rules one wishes to implement in that environment? More generally, we believe it fruitful to study when and why it is difficult for people to implement the principles they feel should guide their choices.

As we referenced in the Introduction, there are many other environments in which our methodology could prove useful. For example, researchers often have hypotheses about competing heuristics, and could use our methodology to elicit the desirability of these heuristics directly. In short, we could elicit what individuals "want to be" or think they "should be" doing, in addition (or instead of) eliciting what they actually do. Naturally these methodologies are complementary, and we believe this to be a fruitful avenue for future research.

#### REFERENCES

- Kenneth Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, pages 328–346, 1950.
- Yaron Azrieli, Christopher P Chambers, and Paul J Healy. Incentives in experiments with objective lotteries. *Experimental Economics*, pages 1–29, 2019.
- Daniel Benjamin, Mark Fontana, and Miles Kimball. Reconsidering risk aversion, Oct 2019.
- B. Douglas Bernheim and Antonio Rangel. Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, 124(1):51–104, 2009.
- Michael H Birnbaum and Alfredo Chavez. Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and human decision Processes*, 71(2):161–194, 1997.
- Michael H Birnbaum and Teresa Martin. Generalization across people, procedures, and predictions: Violations of stochastic dominance and coalescing. *Emerging perspectives on decision research*, pages 84–107, 2003.
- Michael H Birnbaum and Ulrich Schmidt. An experimental investigation of violations of transitivity in choice under uncertainty. *Journal of Risk and Uncertainty*, 37(1):77, 2008.
- Michael H Birnbaum, Daniel Navarro-Martinez, Christoph Ungemach, Neil Stewart, Edika G Quispe-Torreblanca, et al. Risky decision making: Testing for violations of transitivity predicted by an editing mechanism. *Judgment and Decision Making*, 11(1):75–91, 2016.
- Zachary Breig and Paul Feldman. Risky mistakes and revisions, Oct 2019.
- Alexander L Brown and Paul J Healy. Separated decisions. *European Economic Review*, 101:20–34, 2018.
- Colin Camerer. Individual decision making. *Handbook of experimental economics*, 1995.

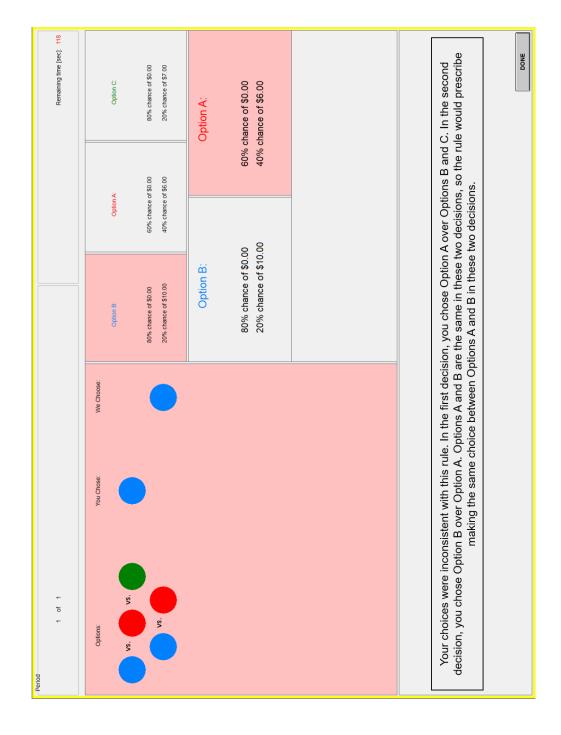
- Timothy N. Cason and Vai-Lam Mui. Individual versus group choices of repeated game strategies: A strategy method approach. *Games and Economic Behavior*, 114:128–145, 2019.
- Paolo Crosetto and Alexia Gaudeul. The rise and fall of asymmetric dominance. *Presentation*, 2019.
- Petro Dal Bó and Guillaume Fréchette. Strategy choice in the infinitely repeated prisoners' dilemma. *American Economic Review*, forthcoming, 2019.
- Urs Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, 2007.
- Mark Alan Fontana. *Mitigating Risk: Smartphone Notifications, Adaptive Surveying, and Genetics*. PhD thesis, University of Michigan, 2015.
- Itzhak Gilboa. Questions in decision theory. *Annual Review of Economics*, 2:1–19, 2010.
- Itzhak Gilboa, Andrew Postlewaite, and David Schmeidler. Is it always rational to satisfy savage's axioms? *Economics & Philosophy*, 25(3):285–296, 2009.
- Itzhak Gilboa, Fabio Maccheroni, Massimo Marinacci, and David Schmeidler. Objective and subjective rationality in a multiple prior model. *Econometrica*, 78(2): 755–770, 2010.
- Ben Greiner. Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125, 2015.
- Yoram Halevy and Guy Mayraz. Modes of rationality: Act versus rule-based decisions. *Working Paper*, 2020.
- Joel Huber, John W Payne, and Christopher Puto. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, 9(1):90–98, 1982.
- Ritesh Jain and Kirby Nielsen. A systematic test of the independence axiom near certainty. *Working Paper*, 2019.

- Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.
- Graham Loomes, Chris Starmer, and Robert Sugden. Observing violations of transitivity by experimental methods. *Econometrica: Journal of the Econometric Society*, pages 425–439, 1991.
- Graham Loomes, Chris Starmer, and Robert Sugden. Are preferences monotonic? testing some predictions of regret theory. *Economica*, pages 17–33, 1992.
- Kenneth R MacCrimmon. Descriptive and normative implications of the decision-theory postulates. In *Risk and uncertainty*, pages 3–32. Springer, 1968.
- Kenneth R. MacCrimmon and Stig Larsson. Utility theory: Axioms versus 'paradoxes'. In *Expected Utility Hypotheses and the Allais Paradox*, pages 333–409. 1979.
- Kenneth O May. Intransitivity, utility, and the aggregation of preference patterns. *Econometrica: Journal of the Econometric Society*, pages 1–13, 1954.
- Ryan Oprea. What is complex? Working Paper, 2019.
- Michel Regenwetter, Jason Dana, and Clintin P Davis-Stober. Transitivity of preferences. *Psychological review*, 118(1):42, 2011.
- Julian Romero and Yaroslav Rosokha. Constructing strategies in the indefinitely repeated prisoner's dilemma game. *European Economic Review*, 104:185–219, 2018.
- Leonard J Savage. The foundations of statistics. Courier Corporation, 1954.
- Uzi Segal. Does the preference reversal phenomenon necessarily contradict the independence axiom? *The American Economic Review*, 78(1):233–236, 1988.
- Christian Seidl. Preference reversal. *Journal of Economic Surveys*, 16(5):621–655, 2002.
- Herbert A. Simon. From substantive to procedural rationality. 25 Years of Economic Theory, pages 65–86, 1976.
- Paul Slovic and Amos Tversky. Who accepts savage's axiom? *Behavioral science*, 19 (6):368–373, 1974.

Amos Tversky. Intransitivity of preferences. Psychological review, 76(1):31, 1969.

Douglas H Wedell. Distinguishing among models of contextually induced preference reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4):767, 1991.

## A. SCREENSHOTS



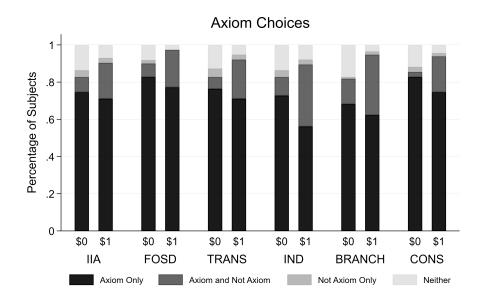


Figure IX: Axiom Selection Rates Across Treatments

#### B. STRICT COST OF DECIDING

As we described in Section III, we ran an additional treatment where subjects had to pay a strictly positive cost, \$1, to make decisions on their own rather than following a rule. Figure IX shows the difference in rule selection rates across treatments. For each of the axioms, individuals are more likely to select both the axiom and the  $\neg$ axiom in the \$1 cost treatment. The Fisher's Exact test confirms significant differences across treatments (p = 0.000).

Figure X shows the choice revisions in the \$1 cost treatment. 51% of inconsistencies are revised, which is significantly lower than the 63% in our main treatment (Fisher's Exact, p = 0.000). However, among reconciled choices, it is still the case that individuals reconcile in favor of the axiom. 91% of reconciliations change lottery choices, which is significantly higher than the 79% in the main treatment (p = 0.000). We find a similar story for the ¬axiom revisions. 47% of inconsistencies are revised, which is significantly lower than the 77% in the main treatment (p = 0.000). Among the revised choices, 56% are revised in favor of the axiom, which is slightly higher than the 44% in the main treatment (p = 0.095).

Overall, we find the same qualitative results in both treatments: Individuals find the axioms to be normatively appealing and revise inconsistencies in favor of the ax-

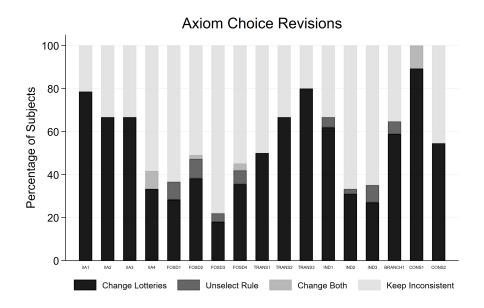


Figure X: Axiom Reconciliation Choices for \$1 Cost Treatment

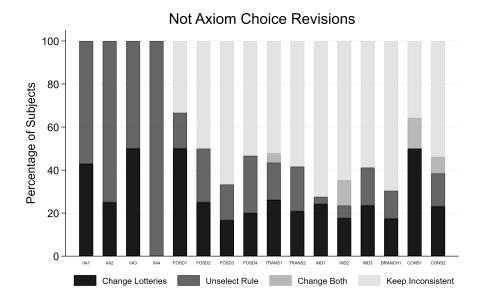


Figure XI: ¬Axiom Reconciliation Choices for \$1 Cost Treatment

ioms. The treatment differences do give some interesting insights into how subjects perceive these axioms. Individuals are more willing to follow rules when it saves them \$1, which is intuitive. They are also less likely to revise inconsistent choices. This suggests that individuals in the \$0 cost treatment view selecting the axiom as indicating that it should be "always" true, while individuals in the \$1 cost treatment view it as something that should be "mostly" true.

In Tables XII and XIII, we report the percentage of violations and direction of reconciliation for each question.

# C. LOTTERIES

**Table III:** IIA Questions

Question		${\rm Lottery}\ B$	
IIA1	60% chance of \$0	80% chance of \$0	80% chance of \$0
	40% chance of \$6	20% chance of \$10	20% chance of \$7
IIA2	60% chance of \$0	80% chance of \$0	85% chance of \$0
	40% chance of \$6	20% chance of \$10	15% chance of \$10
IIA3	60% chance of \$0	80% chance of \$0	85% chance of \$0
	40% chance of \$6	20% chance of \$10	15% chance of \$7
IIA4	60% chance of \$0	80% chance of \$0	70% chance of \$0
	40% chance of \$6	20% chance of \$10	30% chance of \$8

**Table IV:** FOSD Questions (B FOSD A)

Question		Lottery $B$
FOSD1	10% chance of \$1.25 5% chance of \$9 85% chance of \$9.75	5% chance of \$1.25 5% chance of \$1.50 90% chance of \$9.75
FOSD2	10% chance of \$2 5% chance of \$16 85% chance of \$19	5% chance of \$2 5% chance of \$3 90% chance of \$19
FOSD3	21% chance of \$1 18% chance of \$10.25 61% chance of \$11	1% chance of \$1 19% chance of \$2 80% chance of \$11
FOSD4	21% chance of \$0.50 18% chance of \$13 61% chance of \$16	1% chance of \$0.50 19% chance of \$4 80% chance of \$16

Table V: Transitivity Questions

Question		${\rm Lottery} B$	
TRANS1	30% chance of \$6	30% chance of \$0.50	30% chance of \$8
	30% chance of \$6	30% chance of \$11	30% chance of \$8
	40% chance of \$20	40% chance of \$11	40% chance of \$8
TRANS2	45% chance of \$7.50	45% chance of \$1.25	45% chance of \$9
	25% chance of \$7.50	25% chance of \$10.50	25% chance of \$9
	30% chance of \$19	30% chance of \$10.50	30% chance of \$9

# Table VI: Price List Transitivity Questions

Question	A	B	C	D	E	F	G
	25% chance of \$5	100% chance					
TRANS3	75% chance of \$25	of \$14	of \$14.50	of \$15	of \$15.50	of \$16	of \$16.50

# Table VII: Independence Questions

Question		Lottery $B$	Lottery  C	$\operatorname{Lottery} D$
BRANCH1	80% chance of \$0 10% chance of \$2 10% chance of \$12	80% chance of \$0 10% chance of \$8 10% chance of \$9	10% chance of \$2 10% chance of \$12 80% chance of \$15	10% chance of \$8 10% chance of \$9 80% chance of \$15
IND1	80% chance of \$0 10% chance of \$6 10% chance of \$11	80% chance of \$0 10% chance of \$8 10% chance of \$9	40% chance of \$0 30% chance of \$6 30% chance of \$11	40% chance of \$0 30% chance of \$8 30% chance of \$9
IND2	100% chance of \$10	20% chance of \$0 80% chance of \$13.50	75% chance of \$0 25% chance of \$10	80% chance of \$0 20% chance of \$13.50
IND3	20% chance of \$10 80% chance of \$30	100% chance of \$20	24% chance of \$10 8% chance of \$20 68% chance of \$30	8% chance of \$10 84% chance of \$20 8% chance of \$30

# **Table VIII:** Consistency Questions

Question	${\rm Lottery}\ A$	Lottery B
CONS1	50% chance of \$3 50% chance of \$15	25% chance of \$5 75% chance of \$12
CONS2	50% chance of \$5 50% chance of \$10	30% chance of \$0 70% chance of \$15

# D. ADDITIONAL RESULTS

Number of ¬Axioms Selected								
#Axioms								
Selected	0	1	<b>2</b>	3	4	5	6	Total
0	2.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.9%	3.6%
1	0.9%	0.0%	0.9%	0.0%	0.0%	0.0%	0.0%	1.8%
2	2.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.7%
3	3.6%	0.0%	0.0%	0.0%	1.8%	0.9%	0.0%	6.4%
4	3.6%	2.7%	1.8%	0.0%	0.0%	0.0%	0.0%	8.2%
5	11.8%	2.7%	1.8%	0.0%	0.9%	0.0%	0.0%	17.3%
6	39.1%	17.3%	0.9%	1.8%	0.0%	0.0%	0.9%	60.0%
Total	64.5%	22.7%	5.5%	1.8%	2.7%	0.9%	1.8%	100.0%

**Table IX:** Number of Axioms and ¬Axioms Selected

IΙΑ

0.427	TRANS				
0.409	0.489	FOSD			
0.427	0.300	0.569	IND		
0.470	0.283	0.393	0.283	BRANCH	
0.425	0.425	0.464	0.425	0.340	CONS

Figure XII: Pairwise Correlations in Axiom Choices for Main Treatment

Question	# of Violators	# Keep Inconsistent	# Revise Axiom	# Revise Lotteries	# Revise Both
IIA1	18	3	0	15	0
IIA2	17	2	1	13	1
IIA3	17	4	0	13	0
IIA4	11	3	0	8	0
FOSD1	59	27	11	20	1
FOSD2	55	25	12	18	0
FOSD3	48	26	11	10	1
FOSD4	32	17	7	8	0
TRANS1	6	1	2	3	0
TRANS2	3	1	0	2	0
TRANS3	32	5	0	23	4
IND1	22	7	4	11	0
IND2	37	21	4	11	1
IND3	37	17	7	11	2
BRANCH1	22	9	0	12	1
CONS1	19	1	0	14	4
CONS2	33	6	0	27	0

**Table X:** Number of Violations and Revisions Per Question, Conditional on Selecting Axiom

Question	# of Violators	# Keep Inconsistent	# Revise Axiom	# Revise Lotteries	# Revise Both
¬IIA1	10	3	4	2	1
$\neg IIA2$	12	4	5	3	0
$\neg IIA3$	11	4	6	1	0
¬IIA4	9	5	3	1	0
$\neg FOSD1$	5	1	0	4	0
$\neg FOSD2$	4	2	1	1	0
$\neg FOSD3$	2	1	1	0	0
$\neg FOSD4$	5	2	1	2	0
$\neg TRANS1$	11	3	5	3	0
$\neg TRANS2$	11	2	6	3	0
$\neg IND1$	11	5	3	2	1
$\neg \text{IND2}$	8	1	3	4	0
$\neg IND3$	10	5	2	3	0
$\neg BRANCH1$	8	3	3	2	0
$\neg \text{CONS1}$	4	0	1	2	1
¬CONS2	3	0	0	2	1

**Table XI:** Number of Violations and Revisions Per Question, Conditional on Selecting  $\neg$ axiom

Question	# of Violators	# Keep Inconsistent	# Revise Axiom	# Revise Lotteries	# Revise Both
IIA1	14	3	0	11	0
IIA2	21	7	0	14	0
IIA3	18	6	0	12	0
IIA4	12	7	0	4	1
FOSD1	60	38	5	17	0
FOSD2	55	28	5	21	1
FOSD3	50	39	2	9	0
FOSD4	31	17	2	11	1
TRANS1	<b>2</b>	1	0	1	0
TRANS2	3	1	0	2	0
TRANS3	35	7	0	28	0
IND1	21	7	1	13	0
IND2	42	28	1	13	0
IND3	37	24	3	10	0
BRANCH1	17	6	1	10	0
CONS1	28	0	0	25	3
CONS2	22	10	0	12	0

**Table XII:** Number of Violations and Revisions Per Question, Conditional on Selecting and Violating Axiom, in \$1 Cost to Decide Treatment

Question	# of Violators	# Keep Inconsistent	# Revise Axiom	# Revise Lotteries	# Revise Both
¬IIA1	7	0	4	3	0
$\neg IIA2$	4	0	3	1	0
$\neg IIA3$	6	0	3	3	0
¬IIA4	3	0	3	0	0
$\neg FOSD1$	6	2	1	3	0
$\neg FOSD2$	4	4	2	2	0
$\neg FOSD3$	12	8	<b>2</b>	2	0
$\neg FOSD4$	15	8	4	3	0
$\neg TRANS1$	23	12	4	6	1
$\neg TRANS2$	24	14	5	5	0
$\neg IND1$	29	21	1	7	0
$\neg \text{IND2}$	17	11	1	3	2
$\neg IND3$	17	10	3	4	0
$\neg BRANCH1$	23	16	3	4	0
$\neg \text{CONS1}$	14	5	0	7	2
¬CONS2	13	7	2	4	1

**Table XIII:** Number of Violations and Revisions Per Question, Conditional on Selecting and Violating ¬axiom, in \$1 Cost to Decide Treatment

# E. SUPPLEMENTAL APPENDIX

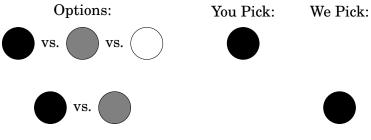


Figure XIII: IIA Rule

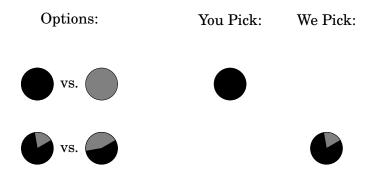


Figure XIV: First Order Stochastic Dominance Rule

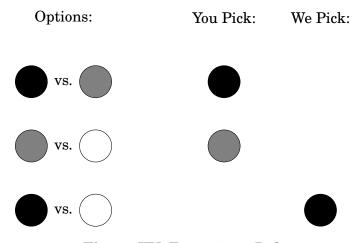


Figure XV: Transitivity Rule

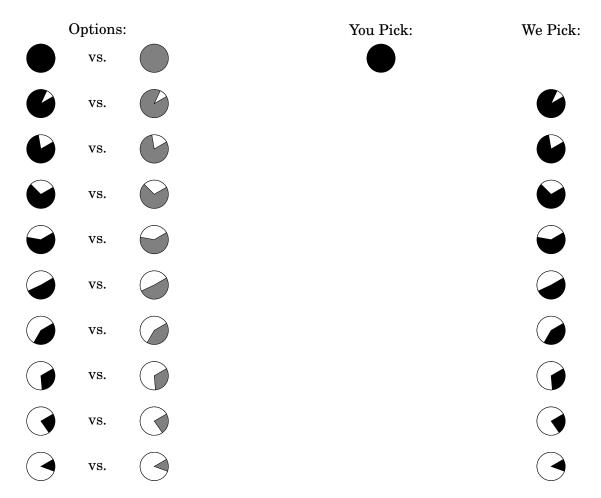


Figure XVI: Independence Rule

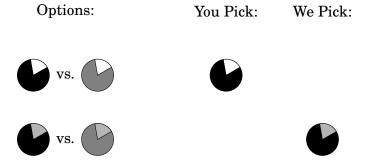


Figure XVII: Branch Independence Rule

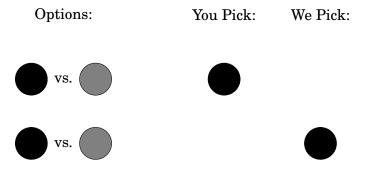


Figure XVIII: Consistency Rule

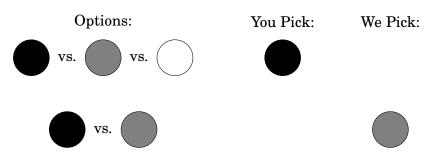


Figure XIX: Not IIA Rule

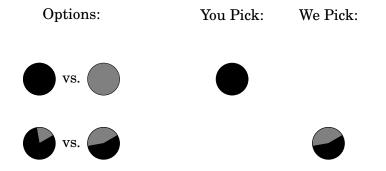


Figure XX: Not First Order Stochastic Dominance Rule

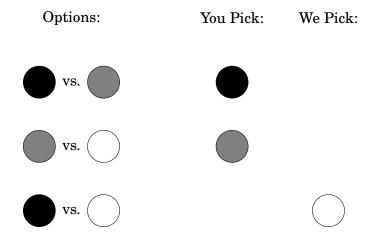


Figure XXI: Not Transitivity Rule

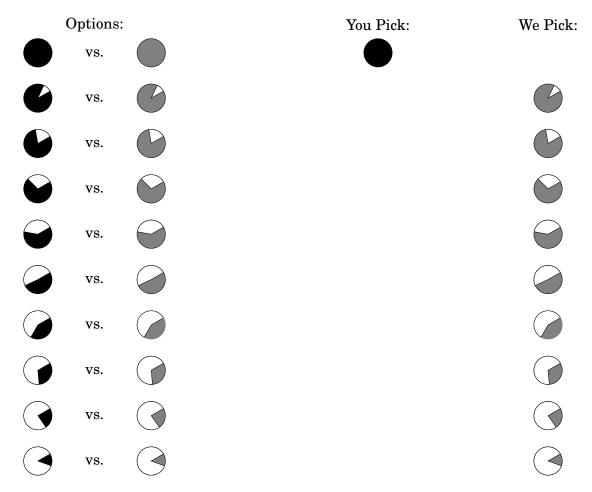
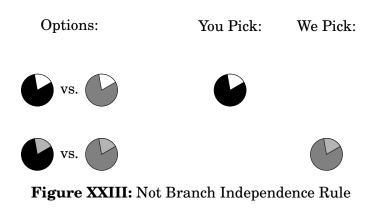


Figure XXII: Not Independence Rule



Options: You Pick: We Pick:

vs.

vs.

Figure XXIV: Not Consistency Rule

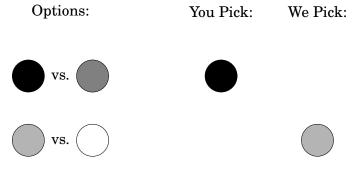


Figure XXV: Distractor Rule #1

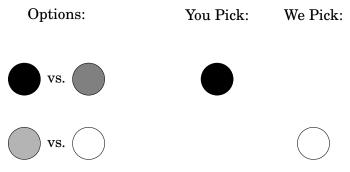
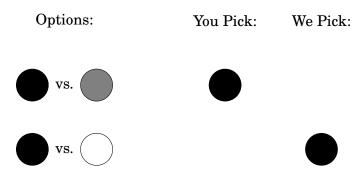


Figure XXVI: Distractor Rule #2



 $\textbf{Figure XXVII:} \ Distractor \ Rule \ \#3$ 

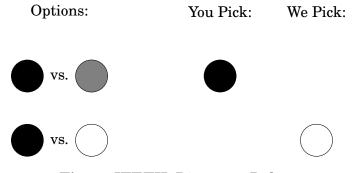
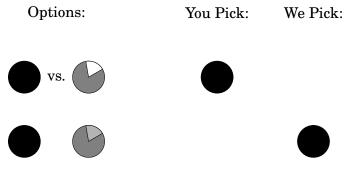


Figure XXVIII: Distractor Rule #4



**Figure XXIX:** Distractor Rule #5

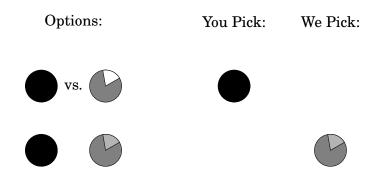


Figure XXX: Distractor Rule #6

#### **INSTRUCTIONS**

This is an experiment in the economics of decision making. Stanford University and the Ohio State University have provided the funds for this research. Feel free to ask questions while we go over the instructions. Please do not speak with any other participants during the experiment and please put away your cell phones and anything that you might have brought with you.

This experiment has three different parts, and each part has many decisions. These instructions are for the first part. You will be paid based on your choice in exactly ONE randomly selected decision from the entire experiment. Each of your decisions is equally likely to be paid. In addition, you will receive \$7 for completing the experiment.

For each part, we will explain how you would be paid from a decision in that part. The decision chosen to determine your payment will be shown at the end of the experiment, and we will roll dice at the front of the room to determine which decision it will be.

#### **RULES**

In this part of the experiment, you will be presented with various decision "rules." These rules are abstract ways to represent choices, and we are interested in how people perceive them. Your task is to evaluate various rules and decide whether you would want to implement choices using the rule or make choices yourself. If you select your preferred rules carefully, will implement your preferred choices from the rules and you won't have to make these choices on your own.

The rules will use colors to represent **positive money payments.** These payments could be "lotteries" or sure payments, which we'll describe below. The possible payments range from \$0 to \$30. You will choose whether you want the rule to make decisions for you or whether you want to make the choices yourself. Remember, when you evaluate these rules, you should only be considering money payments, and only positive amounts (meaning you can't lose money).

A "rule" is like an algorithm that observes your initial choices and then figures out what to pick later based on those initial choices. You can choose to use various rules to make your choices, or you can choose to make each later choice on your own.

For example, here's one possible rule:

Options:	You Choose:	We Choose:
vs.		
vs.		

This rule establishes that, when given the choice between "black" and "white," you chose "black." Then, the rule says that if you were given the choice between "black" and "grey," we would choose black for you.

The colors don't have any inherent meaning (e.g. grey is not "in between" black and white). They can represent any possible money amount or lottery. As a result, you probably wouldn't want to select this rule. If you like black over white, whatever they may be, this doesn't necessarily mean you'd like black over grey, since grey can represent any possible payment. Grey might be \$30 and black might be \$5!

You want to select rules that would be <u>always</u> best for you, regardless of what the colors stand for. It's worthwhile to make the choice on your own if the rule would not always be best for you, like in the example above. This way, if "grey" were \$30, you could choose grey over black. If instead grey were \$0, you'd be able to choose black.

#### **MIXTURES**

Some of the rules will involve "mixtures." Here's an example.

Options:	You Choose:	We Choose:
vs.		
vs.		

This rule says that if you prefer black over grey, then you would prefer a mix of black and white to the same mix of grey and white.

You can think of the mixtures like a spinner wheel. If the spinner wheel lands in the "white" region, you'd receive white. If it lands in the black (or grey) region, you'd receive black (or grey). Essentially, this rule would say that if you like black over grey, then you'd also like black-mixed-with-white over grey-mixed-with-white, regardless of what black, white, and grey are.

For example, imagine black and grey are both lotteries (remember, they can be any lottery), and imagine white is \$10. If the spinner lands in the white region (which happens with probability ¾), you would receive \$10. If it lands in the black or grey region (which happens with probability ¾), you would receive the original black or grey lottery.

### Here's another example:

Options:	You Choose:	We Choose:
vs.		
vs.		

Many people would find this to be a desirable rule. If you like black more than grey, you might prefer the mix of black-and-grey to grey. Because there's a chance you can get black, which you prefer to grey.

# **TYPES OF QUESTIONS**

In the experiment, the colored circles will represent various lotteries. These lotteries can have any payments from \$0 to \$30, and can have any probabilities. For example, maybe one lottery is a 100% chance of \$20. Another lottery might be a 40% chance of \$5, 30% chance of \$11, and 30% chance of \$16.

You won't know what the exact lotteries are while you're making your decisions about the rules, but the lotteries are sufficiently varied. Some might be very risky, some are safe, some involve low payments, some high, etc.

Therefore, when choosing your rules, you'll want to select only rules which are always best for you. By selecting rules which are always best for you, you'll get the option which is best for you without having to make the decision on your own. By not selecting rules which aren't always best for you, you'll ensure that you have the opportunity to pick your most preferred option in the cases where the rule wouldn't be good for you.

#### **PAYMENT**

Along with all decisions you make in this experiment, each of your rule decisions is equally likely to be paid, and we will roll dice at the end of the experiment to determine which decision will be paid. If one of the rules is the randomly selected question to determine your payment, here's how we would pay you.

Each rule corresponds to some specific lottery questions. If you selected to make the choice on your own, you will be shown the lottery question(s) and you will make the choice on your own.

If you selected to implement your choices using the rule, we will make the choice for you based on what the rule says.

You will be paid based on the lottery chosen, either by you or by the rule.

For example, imagine you are being paid based on the following rule:

Options:	You Choose:	We Choose:
vs.		
+ vs.		+

There would be two lotteries corresponding to "black" and "grey." Remember, these lotteries could be anything, but let's say the two lotteries were

- 1. 100% chance of \$12
- 2. 75% chance of \$20, 25% chance of \$10

Let's say you prefer option 2 (75% chance of \$20, 25% chance of \$10), so think of this as "black" and the other (100% chance of \$12) as "grey." In this case, you preferred black over grey, as the rule assumes. Then, the rule says we will pay you "black & white" instead of "grey." The white lottery also could be anything, say it's 100% chance of \$5. Then we would pay you "black & white," or 75% chance of \$20, 25% chance of \$10, as well as 100% chance of \$5, instead of "grey," which would have been 100% chance of \$12.

If you hadn't selected this rule, you would make the choice of black + white vs. grey on your own.

## **HOW TO ANSWER**

There are no "right" or "wrong" answers in these questions. We are interested in people's preferences over these rules, so we simply want to know which rules you like and which you don't.

Since we want to know people's true preferences, the payment is set up such that you'll be paid your "favorite" thing if you answer truthfully. Therefore, you have no incentive to answer differently from what you really think. If you think the rule should describe your choices, you should select it (and then you'll save the cost of making the decision on your own). If you think there are situations where the rule would not give you your favorite option, you should not select it.

#### **LOTTERIES**

In this part of experiment, you will be making choices between "lotteries." A lottery specifies the chance of receiving certain payoffs. In this experiment, the possible payoffs will range from \$0 to \$30. The chance of each payoff can be anything from 0% to 100%.

For example, one lottery could give you an 80% chance of \$18, 10% chance of \$7, and a 10% chance of \$4. Another lottery might give a 100% chance of \$13. There are many different possible lotteries.

In each decision, you will see two or three lotteries on your screen. Your screen will display the written values for the payoffs and probabilities in a table.

This is an example of what that could look like:

Option A:	Option B:
40% chance of \$5	50% chance of \$3
30% chance of \$7	50% chance of \$25
30% chance of \$19	

To visualize these lotteries, imagine rolling two 10-sided dice to generate a number from 1 to 100. In Option A, the first 40 numbers pay \$5, the next 30 numbers pay \$7, and the last 30 numbers pay \$19. In Option B the first half of the numbers pay \$3 and the other half pay \$25. If this is the question to determine your payment, we would roll the dice and pay you according to which number is rolled.

Your task is simply to choose the lottery you prefer. The computer will record your choice and then will bring you to the next decision.

## **PAYMENT**

You will be making 33 decisions. Each decision will be presented on a different screen. Along with all decisions you make in this experiment, each of your lottery decisions is equally likely to be paid, and we will roll dice at the end of the experiment to determine which decision will be paid.

For example, in the randomly selected problem, imagine you chose the lottery which gives

30% chance of \$0

50% chance of \$11.50

20% chance of \$17

If we roll a number 1-30, you'd receive \$0. There are 30 out of 100 possible numbers between 1 and 30, so this corresponds to a 30% chance of \$0.

If we roll a number 31—80, you'd receive \$11.50. There are 50 out of 100 possible numbers between 31 and 80, so this corresponds to a 50% chance of \$11.50.

If we roll a number 81—100, you'd receive \$17. There are 20 out of 100 possible numbers between 81 and 100, so this corresponds to a 20% chance of \$17.

### **HOW TO ANSWER**

Simply choose the lottery you prefer! In any given decision, if it's the decision to determine your payment, we will give you the lottery you selected. So you should select the lottery that you'd rather have determine your payment.

There are no "right" or "wrong" answers in these questions. We are interested in people's preferences over these lotteries, so we simply want to know which lotteries you prefer.

These lotteries are not the same questions as the lotteries which would determine your payment from your "rule" decisions.

#### **RULES AND CHOICES**

In the first part of the experiment, you indicated which "rules" you liked and which ones you did not like. In the second part of the experiment, you made choices over lotteries. It's possible that you selected a rule in Part 1 which could have been applied to make one of your choices in Part 2.

In this final part of the experiment, you will have the opportunity to reevaluate your choices. We will show you if you ever selected a rule and then made lottery choices which were not what the rule would have chosen for you. This way, you can analyze the rule and the choices together. You might decide that you want to change your lottery choices to align with what the rule would have chosen for you. Or you might decide that the rule isn't always true of your choices, so you can "un-select" it. Or, you can leave your choices as they were.

Here's how that will work.

#### **RANKING OF THE RULES**

First, for some of the rules you selected in Part 1, we will ask you to "rank" them in order of how much you would like to revise your choices if they conflict with the rule. An example is shown on the board. Essentially, we are asking you to rank the rules according to how important you think they are. The rule you think is most important would be ranked first, the next most important would be ranked second, etc.

At the bottom of the screen, there is also an option which says "I would rather have \$1 than revise my choices for any of the remaining rules." This is an alternative which can enter into your ranking, too.

For example, let's say you selected six rules. In order of importance, you rank them

Rule 4, Rule 1, Rule 2, Rule 3, Rule 6, Rule 5.

Now suppose that you'd be willing to pay at least \$1 to revise choices that conflict with rules 4, 1, 2, and 3, but you'd rather have an extra \$1 than revise choices that conflict with rules 6 and 5.

Then, you would report a ranking of

Rule 4, Rule 1, Rule 2, Rule 3, \$1, Rule 6, Rule 5.

If this ranking question is randomly chosen for payment at the end of the experiment, we will randomly select two of the things you ranked. We would pay you according to whichever you ranked *higher*. If that is one of the rules, we would pay you for your revised choices in that rule (more on this below). If that is the \$1, you would receive a \$1 bonus and would be paid for your original choices.

Your decision will be implemented only IF your choices are not in agreement with the rule. If your choices *are* in agreement with the rule, we will just implement your original decisions and you still get a \$1 bonus. This means that you don't need to worry about whether you think your choices were in agreement with the rule or not. If they were, your decision in these questions won't matter (since we'll

just implement your original choices). So you should report your ranking as if your choices conflict with the rules.

#### **REVISING YOUR CHOICES**

Regardless of your ranking, we will show you your rule/lottery decisions and you can choose to revise them or not. These decisions actually could be paid if you are paid for your revised choices, as described above.

To give you the opportunity to revise your choices, we will show you the rule on your screen and we'll show you your choices which are inconsistent with that rule. An example is shown on the board.

The choices you made in Parts 1 and 2 will be shown in red. Below that, you'll see an explanation of why the rule would have selected something different than what you chose. To change your choice of the rule, just click the box containing the rule. It will change from red to grey and vice versa. To change your lottery decisions, just click the box of the option you prefer instead. It will turn red, also. You can change any of your choices as many or as few times as you wish. Your final choices will be those in red.

You can revise your choices in any way you wish. For example, your choices being inconsistent with the rule might indicate that you don't like the rule after all and it is not a rule which is always best for you. In that case, you could choose to "un-select" the rule. On the other hand, you might decide that the rule does give you things which are always best for you, and therefore you want to change your choices to align with the rule. Finally, you might decide that you like the rule but also like your choices, even though they're inconsistent. In this case, you can leave all your choices as they are.

In other words, you can change any of your choices or keep them the same. You are under no obligation to make your choices agree with the rules, or vice versa.

#### **PAYMENT**

If one of your revised choices is selected for payment, we would pay you in the same way that we described in Parts 1 and 2. That is, if a revised rule were selected for payment, we would pay you just like how we described for the rules in Part 1. If you selected the rule *in your revised choices*, we would use the rule to make decisions for you. If you did not select the rule, you would make the decision for yourself.

If a revised lottery decision were selected for payment, we would pay you the lottery you selected *in your revised choices*. We would roll dice to generate a number between 1—100 to determine which payoff you receive.

### **HOW TO ANSWER**

There are no "right" or "wrong" answers in these questions. We are interested in people's preferences over these rules and choices, so we want to know which options you prefer.

Since we want to know people's true preferences, the payment is set up such that you'll be paid your "favorite" thing if you answer truthfully. Therefore, you have no incentive to answer different from what you really think. If you think the rule should describe your choices, you should select it. If you think there are situations where it should not describe your choices, you should not. If you like one lottery over the other, you should choose that.

Your revised rule selections or your revised lottery choices could be selected to determine your final payment. If this is the case, we will pay you based on your revised rule or lottery choice, as described in Parts 1 and 2. Therefore, since your revised choices might be selected to determine your payment, you should choose the options you most prefer.

You are also under no obligation to make your lottery choices and rules agree. And, if you do want to change your decisions, you can change either the rule or the lottery choices depending on your preferences.