CONSISTENCY AND HETEROGENEITY IN CONSIDERATION AND CHOICE*

KEATON ELLIS[♭], EMEL FILIZ-OZBAY[†], AND ERKUT Y. OZBAY[§]

Abstract. We empirically investigate the consideration and choice behaviors of individuals under uncertainty. Our experiment elicits these functions by repeatedly questioning the decision makers in a rich lottery domain and, hence, allows them to reveal their stochastic or deterministic consideration and choice. The subjects consider more options on larger menus. 93% of the decision-makers have stochastic considerations on at least one menu and the randomization is more frequent when similar options are jointly presented. Most subjects' consideration data violates monotonicity (Cattaneo et al., 2020.) About 24% have (almost) deterministic choice behavior and they are all consistent with the Weak Axiom of Revealed Preferences. Our tests for stochastic choice properties provide limited support for Regularity, but the majority satisfies Weak Binary Regularity, Strong Stochastic Transitivity, and Independence (Filiz-Ozbay and Masatlioglu, 2023, Gul et al., 2014.)

Date: July 21, 2024.

*We thank Marina Agranov, Paul J. Healy, Efe Ok, Paola Manzini, Pietro Ortoleva, and the audiences of several seminar presentations for helpful comments and discussions. This study was reviewed by the Institutional Review Board at the University of Maryland, and the experiments are funded by the University of Maryland.

[♭]UC-Berkeley, The Simons Institute, 314 Melvin Calvin Laboratory, Berkeley, CA 94720. Email: khkellis@berkeley.edu.

[†]University of Maryland, 3114 Tydings Hall, Department of Economics, College Park, MD 20742. Email: efozbay@umd.edu.

[§]University of Maryland, 3114 Tydings Hall, Department of Economics, College Park, MD 20742. E-mail: ozbay@umd.edu.

1. Introduction

Although the standard choice theory assumes that a decision maker fully evaluates all the options and picks the best according to some preference maximization (Kreps, 1988), the choice options are often initially vaguely described, and the decision maker may want to commit to a shortlist to collect more information on those selected alternatives before making a better-informed decision. For example, a recruiting department may first decide on a few applicants to interview and interact with further and then choose a candidate to make an offer. There are models acknowledging the two-stage idea by describing choice as an optimization on a consideration set induced by a behavioral limitation.[1] The empirical studies also documented limited consideration as a factor affecting choice.[2] In this paper, we collect unique data on the consideration and choice behavior of the decision makers. In our setup, the options are ambiguously described lotteries initially and a decision maker first commits to consider a short list of options to learn the associated risks with the understanding that the final choice will be one of the considered options. Our data is rich enough to test well-known properties studied for deterministic and stochastic consideration and choice functions at the individual level.

Recent developments in choice theory generalizes the classical one by acknowledging not only the possibility of limited consideration but also the tendency to act stochastically either at the consideration stage (Manzini and Mariotti, 2014, Brady and Rehbeck, 2016, Cattaneo et al., 2020, 2023) or choice stage (building on the random utility model of McFadden, 1977, McFadden and Richter, 1990), or both (Kashaev and Aguiar, 2022). The empirical studies confirmed the stochastic property of choice with lab and field data.[3] We aim to empirically test the axioms imposed by those appealing models of limited (stochastic) consideration and (stochastic) choice. There are two obstacles to doing that: (i) observing the decision-makers' consideration function on overlapping menus without altering the information they collect while considering, (ii) repeating the same consideration and choice problems several times

[1]The behavioral limitations include limited attention (Masatlioglu et al., 2012, Lleras et al., 2017), shortlisting (Manzini and Mariotti, 2007), rationalization (Cherepanov et al., 2013), and categorization (Mariotti and Manzini, 2012).

[2]Reutskaja et al. [2011] provides experimental evidence against the full consideration assumption using eye tracking in a search problem; Draganska and Klapper [2011] and Honka and Chintagunta [2017] use survey data to show limited consideration.

[3]Hey [2001] asked subjects to make 100 choices in five sessions. Nobody had consistent preferences throughout the entire experiment. For stochastic behavior in choice, see also Agranov and Ortoleva [2017], McCausland et al. [2020], Rubinstein [2002], Chen and Corter [2006].

to allow for (possibly) stochastic behavior for all the menus of interest without boring the subjects. Our unique design addresses both of these challenges and, to the best of our knowledge, ours is the first individual-level data set with the consideration and choice functions defined on the whole domain being observed. Based on this data, we study the descriptive properties of consideration and choice functions, document heterogeneity in complying with the well-known axioms of interest (such as monotonicity, independence, stochastic transitivity, etc.), and detect certain menu-dependent trends in consideration and choice behavior. Some trends that we observe in the consideration data, such as more randomization between similar-looking options, are new to the consideration literature and should inspire future theoretical models.

We have a two-part design motivated by two-stage models, i.e., consideration followed by choice. The first part collects consideration functions of the subjects fully and the second part collects their choice functions. Our design allows us to analyze Part 1 data without referring to the choice data in Part 2 (or vice versa) and we can test consideration axioms such as monotonicity and attention filter or choice axioms such as regularity, stochastic transitivity, and independence, separately. Since we observe these functions fully (on every menu in the domain rather than only those menus that are needed for testing particular axioms), this data can be utilized to test various existing theories as well as future ones to be developed.

Our consideration elicitation method in Part 1 is new to the literature because previous empirical work testing the two-stage models analyzes consideration indirectly using choice data by committing to a consideration model and studying the implications of limited consideration on choice (e.g., Aguiar et al., 2023, Abaluck and Adams-Prassl, 2021). To the best of our knowledge, ours is the first experiment eliciting the consideration function directly and fully.[4] Moreover, we allow the observed consideration function to be stochastic as we repeat each menu by employing the strategy method (this prevents us from altering the subjects' knowledge about options.)

Each problem in Part 1 presents a menu of ambiguous lotteries (prizes are known but the probabilities are unknown) and subjects indicate what options they want to consider for

[4]Other empirical studies on consideration are either within the optimal search framework (Reutskaja et al., 2011 and Thomas et al., 2021 utilize eye-tracking technology and Honka and Chintagunta, 2017 collects search history of consumers to elicit consideration in search problems), or based on the recall survey of the consumers (Draganska and Klapper, 2011). These are not suitable methods for our purpose of eliciting full consideration function in the two-stage setup at the individual level without altering the decision makers' knowledge.

each menu without receiving any feedback on the probabilities between problems. Without feedback, the information content of an ambiguous lottery (the number of possible prizes and the associated payoffs) remains constant throughout the experiment. The subjects consider options only after we fully elicit their consideration function. At the end of Part 1, a subject is presented with one randomly selected problem and is asked to perform a real effort counting task to fully learn the probabilities of the ambiguous options that she committed to consider earlier for this problem. Upon learning the probability information of the lotteries, she chooses one of the considered lotteries and is paid based on the outcome of this lottery. During the consideration elicitation part, the subjects know that what they indicated to consider on a menu will limit their options to choose at the payment stage of Part 1. Hence, on the one hand, they may want to consider all the options to make better-informed decisions at the payment stage, but considering an option requires performing a counting task perfectly, and the subjects may want to consider a limited set of options to minimize the counting task if they do not like that task or they are afraid of making mistakes, which is severely punished with zero payments.

Part 2 elicits the choice functions of the subjects. The same four lotteries of Part 1 are used in Part 2, but this time they are presented as standard risky lotteries where the prizes and their associated probabilities are known. The subjects indicate what they want to pick from each non-singleton subset of the grand set, and each menu is presented multiple times to allow for stochastic choice behavior.[5] After eliciting their possibly stochastic choice functions, we randomly select one choice problem, and pay the subjects based on the outcome of their choice in that decision problem. The subjects know the payment rule from the beginning of the experiment so they understand that what they indicated to choose in Part 2 will determine their earnings.

Results from Part 1 (Consideration) show that most subjects act stochastically rather than deterministically. They have limited consideration on every type of menu and this behavior is persistent throughout the experiment. They consider more alternatives on larger menus indicating that they understand the trade-off between making a better-informed decision and the cost of investigating more options. The subjects who had deterministic consideration

[5]Agranov and Ortoleva [2017] and McCausland et al. [2020] also repeat the choice problems and document stochastic choice. Earlier studies repeating choice problems in experiments include Balakrishnan et al. [2022] and Hey and Orme [1994]. Our design is closest to McCausland et al. [2020] but we have more granularity in choice frequencies as we repeat the menus 10 times rather 6.

(i.e., always consider the same options on a menu when repeated) satisfy the Attention Filter property (Masatlioglu et al., 2012) and Competition Filter property (Lleras et al., 2017). We designed our ambiguous options so that two of them had a wide range of prizes (i.e., high maximum and low minimum prizes) and two of them had a narrow range (i.e., low maximum and high minimum prizes). Hence, each option had a similar (or somewhat duplicate) option in the domain. The randomization in consideration is more likely if similar options are presented together and these options are treated similarly by the subjects. We have limited support for the Monotonic Consideration axiom (Cattaneo et al., 2020) although this is the weakest stochastic consideration axiom offered by the literature. We discuss possible ways to further weaken this axiom.

In Part 2 (Choice), most subjects acted stochastically on at least some menus. Stochastic choice is more common if the menu contains similar options. Our deterministic subjects ($\sim 24\%$) satisfy the Weak Axiom of Revealed Preference. The others acted stochastically depending on the menu. We test some well-known axioms of stochastic choice theory for those subjects and observe heterogeneity. Similar to McCausland et al. [2020], we have limited support for the Regularity axiom. However, the majority of subjects satisfy the Weak Regularity, Weak Binary Regularity, Strong Stochastic Transitivity, (Weak) Independence, and Elimination of Duplicates properties (Echenique et al., 2011, Filiz-Ozbay and Masatlioglu, 2023, Gul et al., 2014.)

We can also interpret our choice data in the style of Balakrishnan et al. [2022] and Ok and Tserenjigmid [2023], which interpret stochastic choice functions as noisy implementations of choice correspondences: options above a *ex ante* selected threshold relative choice frequency λ are included in the choice correspondence, and options below are classified as unwanted noise and excluded from the choice correspondence. Under this interpretation, we test the Weak Axiom of Revealed Non-Inferiority (Eliaz and Ok, 2006a) and find that the vast majority of subjects satisfy the axiom regardless of the threshold λ.

Understanding how limited consideration impacts the final choice is an important problem for choice architectures. To address this issue, we generate an implicit choice data set for each subject by taking into account their elicited consideration function. More precisely, on each option set we combine the observed consideration function with the observed choice on the considered set and generate a derived choice function. Then we compare the actual choice with the derived one to see if some desirable alternatives (those that would be chosen if considered) are eliminated in the consideration stage. We find that about half of the subjects' choices were significantly impacted by their limited considerations.

In sum, our contribution has three folds: (i) we offer a method to observe consideration and choice functions at the individual level, (ii) we test stochastic properties of these functions which have been the focus of recent literature, (iii) we relate the consideration and choice behavior at the individual level that should help with welfare improvements through menu design. Moreover, our findings highlight the type of deviations observed from the well-known theories. The need for modeling a consideration stage as an optimization problem with bounded rationality is striking.

In what follows, we explain the design in Section 2 and in Section 3 we introduce the basic notation for the relevant theoretical models and axioms that will be studied. Section 4 first analyzes the consideration data (Part 1) followed by the choice data (Part 2). Each part is first analyzed descriptively and then by testing the relevant axioms of the literature. Then, we analyze a hypothetical choice function based on the observed μ and π for each subject. Section 5 suggests an alternative axiom for consideration data and finds majority support in our sample. Section 6 concludes. Additional results are reported in Appendix D. The instructions of the experiment can be found in Appendix A.

2. Design of the Experiment

The experiment consists of two main parts followed by a demographics questionnaire on risk and ambiguity preferences, age, gender, and education level. We elicit what subjects want to consider on a set of options in Part I and what they want to choose from an option set in Part II. The same decision problems are presented multiple times to allow subjects to reveal their stochastic consideration or choice if they prefer to do so. The instructions are in Appendix A.

Throughout the experiment, the subjects make decisions involving four main options with uncertainty.[6] Each option is a lottery but the subjects only see its possible outcomes (not the probabilities attached) in Part I. In Part II, the subjects see the lotteries fully (the payoffs and the corresponding probabilities.) The lotteries used are in Table 1. Each

[6]There is a fifth option, E, which is used to detect inattentive subjects to eliminate noise in the data.

lottery is presented as a box filled with 100 colored balls. We tell subjects that the color of a randomly picked ball from the box they select eventually will determine their earnings in the experiment. Note that Lotteries A and B have relatively larger ranges of prizes than C and D. E is the option we introduced for attention check throughout the experiment. Its highest outcome is lower than the lowest outcome of any other lottery and therefore it is dominated under any distribution.

In both parts of the experiment, subjects are presented a sequence of decision problems. We use all the subsets of $\{A, B, C, D\}$ of size two or greater to construct 11 decision problems (six binary, four trinary, one quaternary option sets) and repeat each one of those ten times to allow for stochastic behavior. This creates 110 decisions to make for each part. In addition to that, the option set $\{C, E\}$ is presented five times as an understanding check in Part I. Since lottery C clearly dominates E (even without knowing the probabilities of outcomes of C and E) we expected that any subject, who prefers more money to less and understands the experiment, shouldn't consider E alone in Part I. The order of the 110 decisions is randomized at the subject level. For Part I, the five repetitions of $\{C, E\}$ are evenly distributed in that sequence. In total, 115 consideration decisions and 110 choice decisions were made in part Parts I and II, respectively. In addition to those, at the end of Part I a randomly selected decision problem is picked to actually perform the consideration task. Below, the procedural details of Parts I and II are explained, separately.

TABLE 1. The lotteries used in the Experiment

Lottery	Payoffs			Probabilities		
A	20	128	360	50%	30%	20%
B	40	120	320	51%	29%	20%
C	80	84	160	20%	24%	56%
D	72	136	140	25%	50%	25%
E	4	8	16	33%	33%	34%

Part I. At the beginning of Part I, we tell subjects that when they see a box in this part, the colors of the individual balls are not specified, but they will know the possible colors and their corresponding prizes. Hence, they do not know how likely it is to pick a certain color in a box. We tell them that they will see a sequence of decision problems and in each one there will be a set of options for them to consider. In order to make subjects gain some experience with the difficulty of the consideration task that they will face in this experiment, we first provide a sample problem with different colors and payoffs than what we used in the actual experiment. Counting the color content of the presented boxes in this dry round should give the subjects an idea of how costly or pleasant the consideration task is, hence, they may have a better understanding of the trade-off considering more or less options.

Figure 1 presents an example of a consideration problem with two options. This figure is copied from one of the actual rounds we used in the experiment. Note that the content of the boxes are not revealed but the possible prizes of each box are announced. The subject needs to choose what to consider, i.e., which boxes to investigate further to learn the content. They need to click on at least one box and they may click as many boxes as they want to learn the content. This determines what they want to consider for that problem. They decide what to consider for all 115 problems without feedback. This is an application of a strategy method. To incentivize Part I, one of the 115 problems is randomly selected for actual consideration after all the problems of this part are completed. The subject sees the color combinations of the boxes they had chosen to consider for that problem. She needs to count the color content of the boxes they consider, correctly. Subjects are told that if they make a mistake in counting, then they will not be paid for Part I. So if a subject wants to avoid counting several boxes in this stage, they should choose to consider less boxes in Part I; if they do not mind counting and believe they can perform the task perfectly, it is best to consider all the boxes to learn the probability of each prize and make informed decision at this stage. After considering the boxes of their choice, then they choose one of the considered boxes and the realization of a random draw from that box determines the subject's payoff in this part. Subjects will learn whether they will be paid for this choice at the end of the experiment. This is to avoid any wealth effect impacting Part II performances.

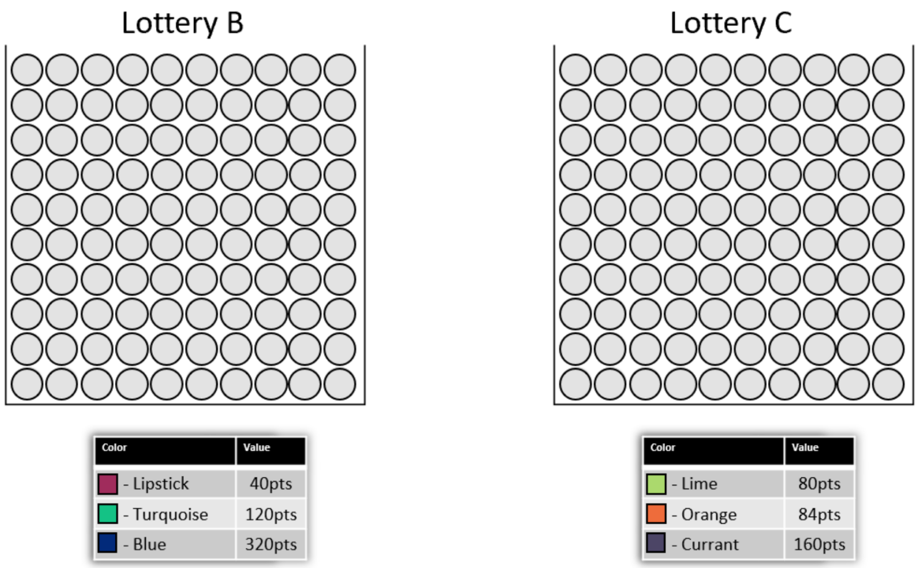


FIGURE 1. A sample consideration problem with two options.

There are three important design details that we cautiously implemented:

First, note that we incentivize the consideration decisions in Part I by making subjects to eventually choose from what they considered on a randomly selected decision problem. This is our method to elicit (possibly stochastic) consideration data in an incentivized way.

Second, no feedback is given between decision problems in Part I. Otherwise, if a subject considers a box in one problem and immediately learns its content, they would make an informed decision on whether to consider that box the next time it appears in a different problem. Then the decision problems would not be independent from each other because there would be some learning. Since the consideration models that we will test are static, we needed to shut down such possibilities of learning. Our procedure gives us stochastic consideration data at the individual level.

Third, even though there were 115 consideration problems where subjects had to reveal what they would consider on that problem, they performed this task only on one randomly selected problem. This serves two purposes: simplicity and elimination of learning. Note that counting the color content of boxes with 100 balls is a hard and boring task. Therefore, we can only ask subjects to perform this a limited number of times. Moreover, counting the content of a box once is enough to learn it, so if we made the subjects perform the same consideration multiple times, they would not count it every time as they would remember the numbers. No feedback feature of the design guarantees independent observations on each decision problem.

Part II. This part is similar to Part I in terms of how 110 decision problems are generated. This time, a decision problem presents the available options by both showing the color content of the boxes and the corresponding prizes (see Figure 2.) Hence, each option is presented as an objective lottery. The subjects need to choose a single option for each problem and after completion of this part, they are paid based on one randomly selected decision problem of Part II and paid based on the realization of that lottery.

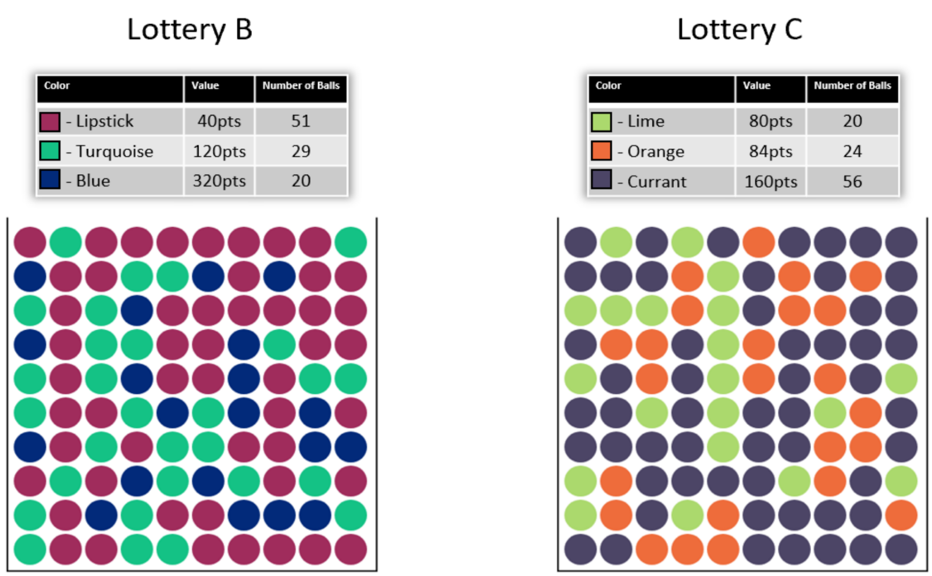


FIGURE 2. A sample choice problem with two options.

After Part II, a questionnaire takes place. The first page of the questionnaire elicits risk preferences using a standard multiple-price list method. Subjects will consider a bag of red and blue balls. There are 10 balls - 5 red and 5 blue - in the bag and drawing red ball pays 60 points while a blue ball pays zero. Then, subjects answer 11 binary decisions. Each decision asks subjects whether they prefer to draw a ball from the bag or whether they prefer to take a sure amount, which starts at 15 and increases in increments of two. One offer will be randomly chosen for payment. The second page of the questionnaire elicits ambiguity preferences and closely mimics the same elicitation process as risk preference elicitation, but subjects are told that the bag of balls used for the section has an unknown proportion of red and blue balls. Subjects then choose which color to bet on, in contrast to the previous page's set choice of red.[7] Finally, a third page asks subjects questions on age, gender, and highest level of schooling achieved. Subjects can choose a "Prefer not to answer" option in the third page.

The experiment was conducted on Prolific with 402 subjects in September and October 2022. 201 and 189 specified their gender as male and female, respectively, and 12 of them did not specify their gender or did not identify as either male or female. The median age range was 28-37 and the median subject had a bachelor's degree. The details of the demographics of the subject pool can be found in Tables 6, 7, and 8 in Appendix B. We implemented an attention test to make sure that the subjects were engaged in our experiments fully: Those who never considered only dominated alternative E in the decision problem $\{C, E\}$ are labeled as inattentive and excluded from the analysis. Hence, the results are based on 315 subjects.[8] The average hourly payment was \$9.21 and it took approximately 27 minutes on average for a subject to finish the experiment.

3. Theoretical Framework

In Part 1, we sample from each subject's consideration function. Let $X = \{A, B, C, D\}$ be the grand set of options. For each $S \subseteq X$, a subject's consideration function that is sampled in Part I is

$$\mu(. \mid S) : 2^S \setminus \varnothing \longrightarrow [0,1] \text{ such that } \sum_{\substack{T \subseteq S \\ T \neq \varnothing}} \mu(T \mid S) = 1$$

We say that a subject has a deterministic consideration function if for each $S \subseteq X$, there is a unique $T \subseteq S$ such that $\mu(T \mid S) = 1$ and for all $T' \subseteq S$ such that $T' \neq T$, $\mu(T \mid S) = 0$. Otherwise, we say the subject has a stochastic consideration function. Note that a subject

[7]This is standard in ambiguity attitude elicitation experiments and it aims to assure subjects that the experimenter does not have any color bias.

[8]In addition to that we may filter those who did not complete the dry run of the consideration task successfully and that would drop the number of subjects to 249. The analysis does not qualitatively change with or without these subjects and they do not change the demographic composition of the subject pool. Nevertheless, we choose to include those since the dry round is not incentivized.

with a deterministic consideration would consider the same options every time the same decision problem appears in the experiment.

It is useful to introduce an implicitly derived attention frequency function of an alternative given an option set because it will be referenced by some of the existing models that we will test later. Given $\mu(. \mid S)$ (which is observed in Part 1), one can calculate how likely it is for the subject that an option x attracts attention in the decision problem of S by adding up the probabilities of each subset containing x being considered. Cattaneo et al. [2023] defines the attention frequency of an alternative in an option set S as follows: Given the consideration function $\mu(. \mid S)$, the attention attracted by option $x \in S$ is defined by

$$\phi_\mu(x \mid S) = \sum_{\substack{T \subseteq S \\ T \ni x}} \mu(T \mid S) \tag{1}$$

For example, when presented with set $\{x, y\}$ if a subject's consideration function is $\mu(\{x\} \mid \{x, y\}) = \mu(\{x, y\} \mid \{x, y\}) = 0.5$, then we say that x attracted attention by probability 1, because it is contained in every subset that has positive consideration weight.

For each subject, we denote the induced attention frequency function by $\phi(. \mid S)$ rather than $\phi_\mu(. \mid S)$ to simplify the notation. Since each subject has only one observed μ function, this simplification should not lead to any confusion.

Next, we introduce notation for the data observed in Part II of the experiment. For each $S \subseteq X$, the choice function of a subject in Part II is

$$\pi(\cdot \mid S) : S \longrightarrow [0, 1] \text{ such that } \sum_{x \in S} \pi(x \mid S) = 1$$

We say that subject i has a deterministic choice function, if for each $S \subseteq X$ there exists a unique $x \in S$ such that $\pi(x \mid S) = 1$ and for all other $y \in S$ such that $y \neq x$, $\pi(y \mid S) = 0$. Otherwise, we say the subject has a stochastic choice function. Part II data provides us a sampling of π for every subject, as we elicit how frequently a subject chooses each option presented in all ten appearances of a decision problem. Note that a subject with a deterministic choice would choose the same option every time the same decision problem is presented.

4. Results

We will present first the consideration results followed by the choice results based on Parts I and II of the experiment, respectively. For each case the descriptive will be provided to identify certain trends in the data and then the axiom testing will come. Recall from the earlier discussion on how we filtered noise, we are reporting results for 315 subjects who passed the attention check and never considered option E alone.

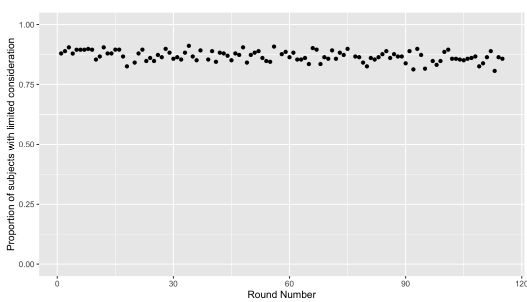
4.1. **Descriptive Analysis of Consideration.** Note that the behavior in Part I may deviate from the standard full consideration hypothesis in two ways: A subject may consistently consider a unique subset of an option set or she may consider a different subset when the option set is presented repeatedly. The former behavior would be a limited deterministic consideration and the later one would be a stochastic consideration. The following two subsections address these behaviors.

4.1.1. *Full vs Limited Consideration.* Recall that in Part I, each decision problem presents a set of 2-4 boxes (lotteries), and asks subjects on which non-empty subset of these boxes they want to consider, i.e., count the color content. On one hand, a subject may want to learn the color content of *more* boxes so that they can make an informed lottery choice after learning the probability of each corresponding prize. On the other hand, they may want to count the balls in *fewer* boxes as counting may not be their favorite activity or they may want to minimize the possibility of mistakes in counting (note that the subjects are not paid for Part I if they make a mistake in this part of the experiment.) It turns out to be that the latter concern is indeed valid as subjects did not choose to consider all the presented options in Part I. Out of 315 subjects, only 7 subjects exhibit full consideration in all questions and full consideration is rare for any menu size (see Table 2, column 2 for the menu sized of 2-4.)

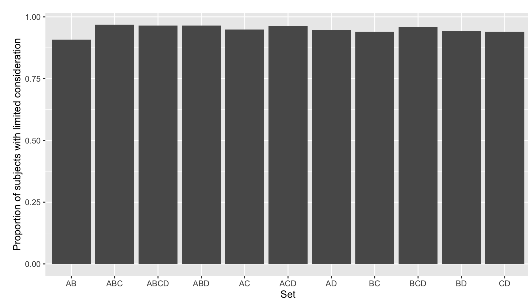
Table 2. Number of subjects with full and limited consideration

	Full Consideration	Limited Consideration
All Data	7	308
$\lvert S\rvert = 2$	8	307
$\lvert S\rvert = 3$	9	306
$\lvert S\rvert = 4$	11	304

The limited consideration results are robust. Figure 3a reports that about 87% of the subjects had limited consideration in each one of the 115 rounds. Figure 3b provides further evidence that limited consideration applies to any type of menu. Almost all the subjects had limited consideration for each of the 11 menus they saw.



(A) Proportion of limited consideration subjects by round.

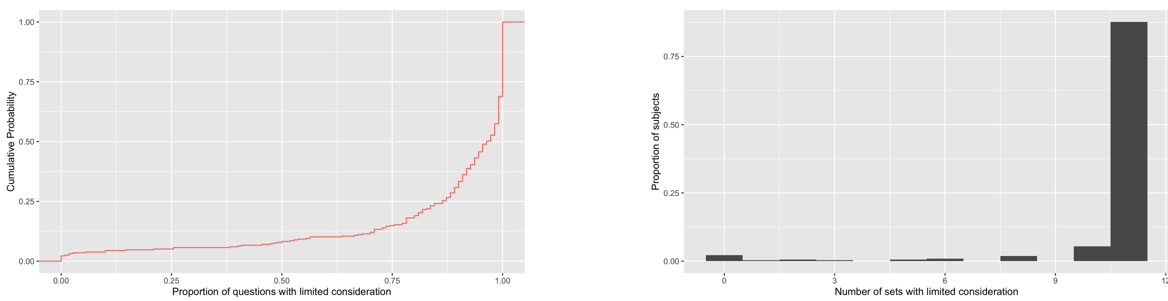


(B) Fraction of subjects with limited consideration by set.

Figure 3. Robustness of Limited Consideration

Figure 4a shows the distribution of proportions of questions with limited consideration by subjects. In other words, it calculates how many subjects considered a certain percentage of 110 questions in a limited fashion. Note that 98 subjects (31.1%) had limited consideration every time they faced a menu. These subjects are located on the far right of the cdf in Figure 4a where there is a big jump. 268 subjects (85.1%) had limited consideration on at least 75% of the questions out of 110 consideration questions they answered. In sum, limited consideration is robustly exhibited by almost every subject on every type of menu throughout the experiment. We may perform this analysis at the menu level as well. Figure 4b reports the percentages of subjects who considered a certain number of menus in a limited fashion.

Note that 276 subjects (87.6%) had limited consideration on all 11 menus in Figure 4b. The 7 subjects who fully considered every menu (as reported in Table 2) appear on the very left of the distribution in this figure.



(A) Distribution of proportion of questions with limited consideration, by subject.

(B) Distribution of subjects by the number of sets that have limited consideration.

FIGURE 4. Robustness of Limited Consideration

The number of alternatives that are considered increases by the menu size. The average size of the considered set is 1.18, 1.37, and 1.59 for menus with size 2, 3, and 4, respectively. As it can be seen in Figure 5, this behavior is robust in every round of the experiment and more alternatives are considered on larger menus throughout the session. The blue, orange, and gray correspond to data coming from problems with two, three, and four options presented, respectively.

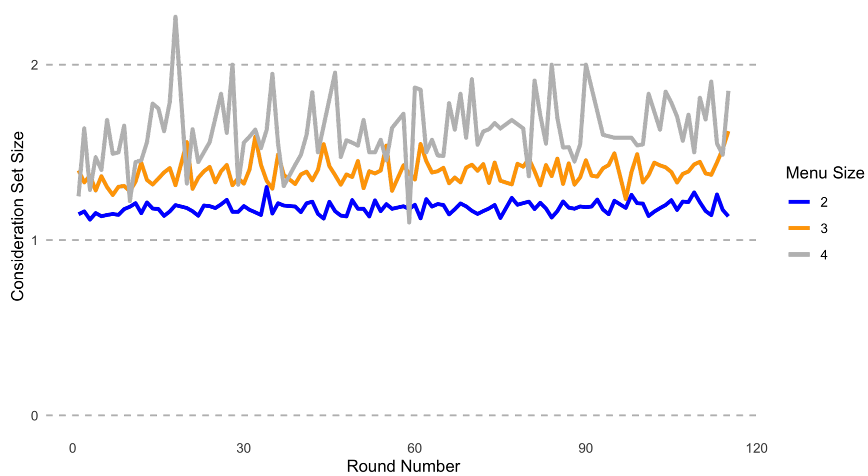


FIGURE 5. Consideration set size by round and total set size

Result 1: The first set of observations we made on limited consideration is listed below.

i. Participants do not consider all of the alternatives (Figure 5);
ii. The number of options that are considered increases with the menu size (Figure 5.)

4.1.2. *Stochastic vs. Deterministic Consideration.* The limited consideration of the subjects mostly occurs in stochastic fashion. Recall that 7 out of 315 subjects had full consideration all the time and that behavior is trivially deterministic. Among the remaining 308 subjects only 16 acted deterministically in their limited consideration, i.e., they considered the same sub-menu in every repetition. Table 3 reports the number of subjects who considered deterministacally by the menu size as well. The small numbers of deterministic considerations indicate that the majority had stochastic considerations.

Table 3. Number of subjects with stochastic and deterministic consideration

	Stochastic	Deterministic Full	Deterministic Limited
All Data	292	7	16
$\lvert S\rvert = 2$	283	8	24
$\lvert S\rvert = 3$	262	9	44
$\lvert S\rvert = 4$	227	11	77

Unlike the behavior of limited consideration, stochasticity of the consideration depends on the menu. Only about 18% acts stochastically on every menu. Figure 6 shows the distribution of subjects who have stochastic consideration by the number of menus. The set dependence of stochasticity in consideration can be further seen in Figure 7. We see that more subjects act stochastically on some sets than the others.

Notably, among the binary menus the subjects behaved more stochastically on menus $\{A, B\}$ and $\{C, D\}$. Recall that the options in these menus are similar, i.e., A and B are high range and C and D are low range options. The subjects tend to randomize when they face with similar options.

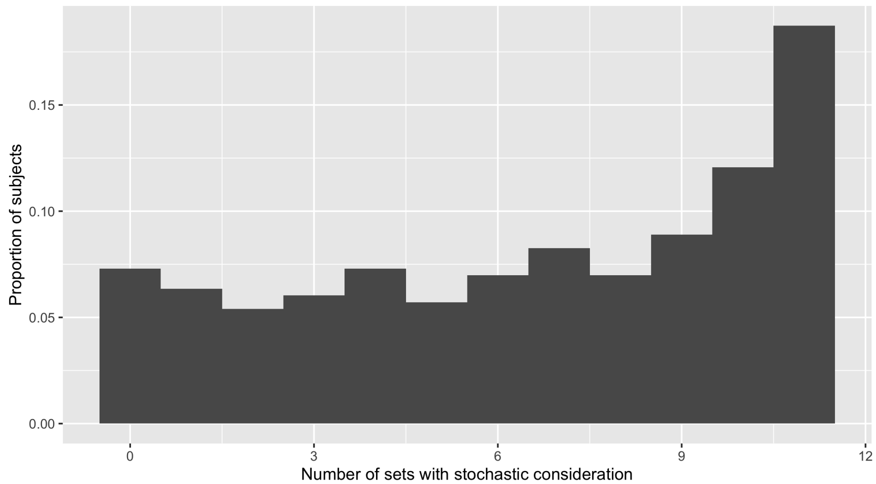


FIGURE 6. Distribution of subjects by the number of menus that have stochastic consideration.

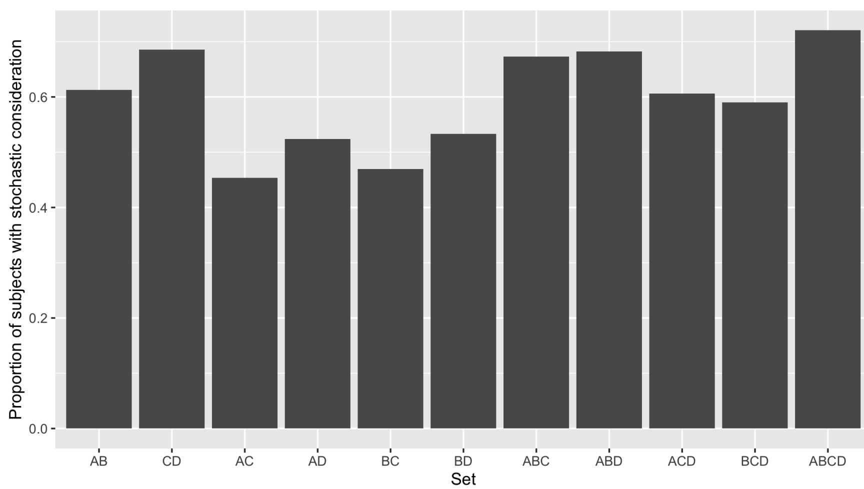


FIGURE 7. Fraction of all subjects with stochastic consideration by set.

There are some robust behavior on what subjects prefer to consider on different menus. Since it is easier to graph the behavior on binaries, we explain the data on binary menus as summarized by Figure 8. Each triangle represents the consideration data on a binary menu, such as $\{A, B\}$, $\{C, D\}$, etc.. Each vertex of a triangle is a possible consideration set. The size of a red circle on a triangle shows the number of subjects who have that type of consideration function. The subjects with deterministic considerations are on the vertices and those with stochastic considerations are either on the edges or in the interior of a triangle. Closer a red circle to a vertex means the frequency of considering the sub-menu on that vertex is higher. For example, if a subject always considered only A on menu $\{A, B\}$, she is denoted on the left bottom vertex of the first triangle in Figure 8. If she considered A alone half of the time and together with B half of the time, then her data is denoted on the mid-point of the bottom edge of this triangle.

We observe three robust behavior in Figure 8. First, the right bottom corner of each triangle has little data indicating that discrete full consideration is rare in all of these four representative menus. This observation is a piece of further evidence on the extensive limited consideration that we observed in Figure 3b.

Second, the behavior on menus in the first row is similar to each other and the same is true for the second row of the figure. The first-row menus include similar options and the second-row menus include different options in the sense of outcome ranges. Note that the deterministic consideration is more pronounced when non-similar options are presented, as the masses on the left bottom vertices of all the triangles in the second row are larger.

Third, the majority of those who act deterministically on menus with non-similar options (the second row in the figure) have preferences for considering the higher range option. For example, the mass of subjects deterministically considering A on menu $\{A, C\}$ is much higher than those considering C deterministically. This contrasting tendency for high range option when it is compared with a low range one is robust, as can be seen in all four menus in the second row of the figure. This means that the subjects who want to consider only one option would like to investigate the option with extreme outcomes hoping for a jackpot.

At the consideration stage, the range of ambiguous lotteries is one way to identify options so it is expected that consideration behavior is robustly sensitive to this characteristic. We intentionally designed the experiment by making A and B as a high range and C and D as low range lotteries. Figure 8 provided the first evidence for the fact that the subjects indeed responded to similar options similarly on binary menus. We may also establish this by checking the consideration behavior on tripletons. Since it is not possible to plot consideration function triangles for tripletons, as we did for the binary menus, we can instead show this result by checking the attention frequencies of each alternative on tripletons. Figure 9 reports the relevant attention frequencies.[9] Note that the attention frequencies of A and B are similar

[9]The attention frequency of an option on a menu is calculated using the definition of ϕ given in (1).

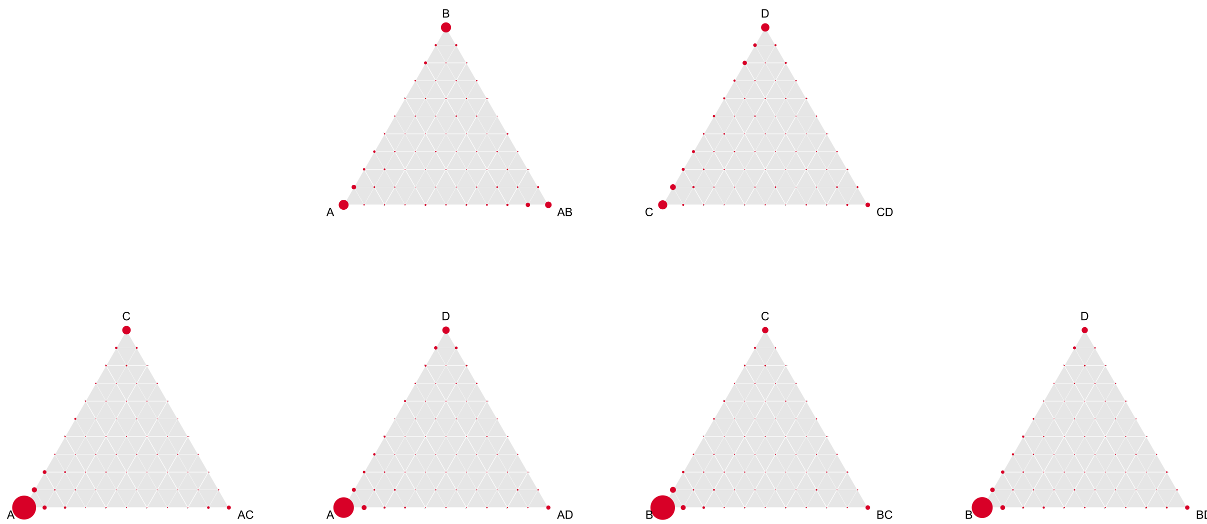
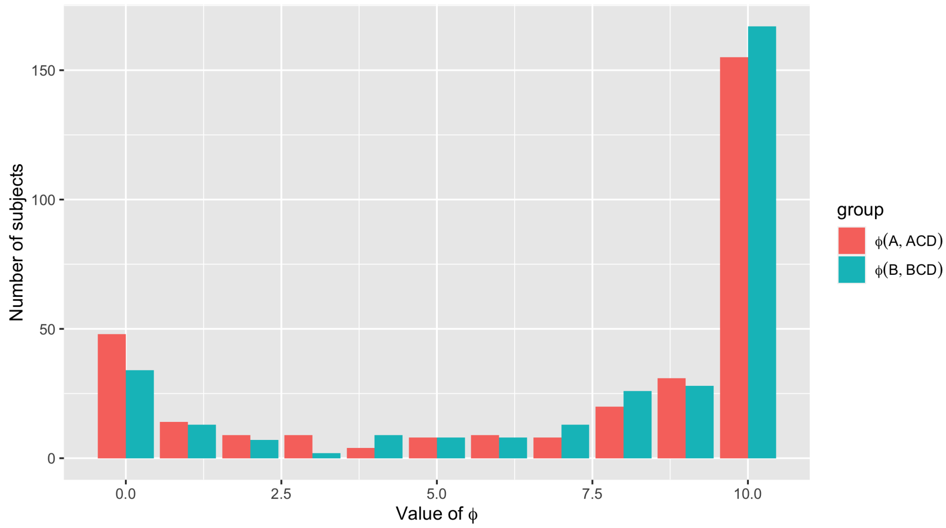
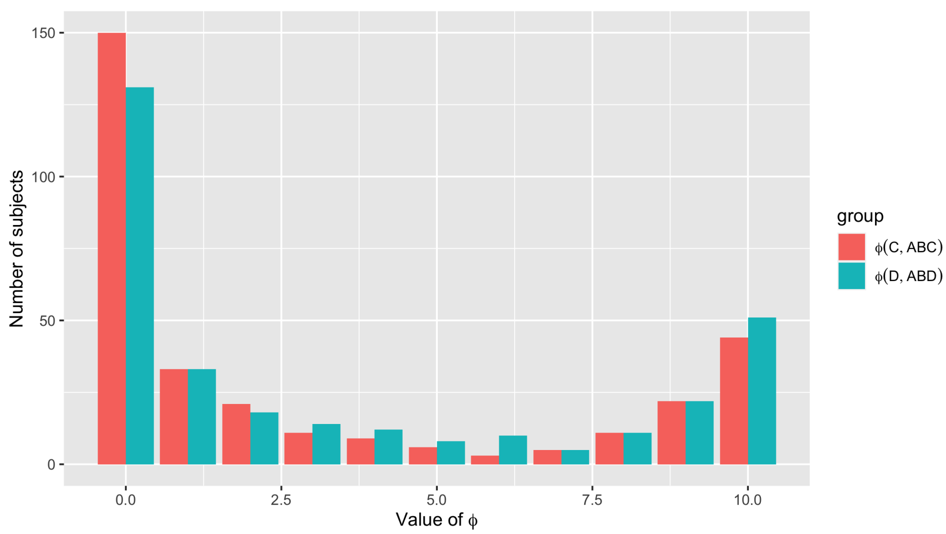


FIGURE 8. Consideration behavior on binary menus. A coordinate in the triangle simplex shows the frequency with which the single options or the binary option are considered. Vertices are deterministic behavior.

on menus $\{A, C, D\}$ and $\{B, C, D\}$, respectively. In other words, adding C and D to a menu affects the appeal of A and B similarly. Also, the attention frequencies of C and D are similar on menus $\{A, B, C\}$ and $\{A, B, D\}$, respectively. Hence, adding A and B to a menu affects the appeal of C and D similarly.



(A) Attention frequencies of A and B in the presence of both C and D.



(B) Attention frequencies of C and D in the presence of both A and B.

FIGURE 9. Attention frequencies of similar options are similar in tripletons

Result 2: Stochastic Consideration results are listed below.

i. Almost all subjects considered stochastically on at least some menus (Table 3 and Figure 6);

ii. Stochastic consideration is more likely when similar options are presented together (Figures 7 and 8);

iii. Similar options are treated similarly when each is presented together with a set of other options (Figures 8 and 9a and 9b.)

4.2. **Testing Consideration Axioms.** Only 7.3% of the subjects (23 out of 315 subjects) had deterministic considerations, i.e. they considered the same set of options in every repetition of each decision problem. Out of those 23 subjects, 7 subjects had full consideration, the rest had a non-trivial deterministic consideration function. In addition to those 23 subjects, if we consider those who deviated from being deterministic in only one out of 110 questions answered (i.e., behaved almost deterministically) as well, we have 37 subjects (11.7%) with almost deterministic consideration functions.[10] All of these subjects satisfied the Attention Filter and the Competition Filter properties.[11]

Since most subjects chose what to consider stochastically, our axiom testing exercise will focus on models that allow for stochastic consideration sets. Cattaneo et al. [2020] and Cattaneo et al. [2023] provide two relatively weaker axioms of consideration. Both are monotonicity properties. One of them is the monotonicity of the consideration function, μ, that we sample in Part I and the other one is the monotonicity of the attention frequency, ϕ, derived from μ as we defined in (1).

Cattaneo et al. [2020] propose a general model of random attention where consideration sets compete for attention, hence, the probability of considering a set in a menu increases when the menu gets smaller. This is captured by the monotonic consideration rule, μ, stated below.[12]

Monotonic consideration. For any T, S, $x \in S$ such that $T \subset S \setminus \{x\} \subseteq X$, $\mu(T \mid S) \leq \mu(T \mid S \setminus \{x\})$

[10]Among those almost deterministic ones, only 1 subject behaved almost full consideration and the rest behaved almost deterministic with limited consideration.

[11]Let $\Gamma(S) := \{T \mid \mu(T \mid S) = 1\}$ denote what is considered on a menu given a deterministic μ function. The Attention Filter property of Masatlioglu et al. [2012] requires that "$\Gamma(S) = \Gamma(S \setminus \{x\})$ if $x \notin \Gamma(S)$" and the Competition Filter property of Lleras et al. [2017] requires that "if for all $x \in S \subset T$ and $x \in \Gamma(T)$ then $x \in \Gamma(S)$".

[12]Cattaneo et al. [2020] call this property as monotonic attention as they name function μ as the attention function.

Cattaneo et al. [2023] impose a similar monotonicity requirement on attention frequency, ϕ, instead.

Consideration overload. For any T, S, x such that $x \in T \subseteq S \subseteq X$, $\phi(x \mid S) \leq \phi(x \mid T)$.

Table 4 reports the subjects who fail to reject these two monotonicity properties. Tests are performed using the estimation techniques adopted from Cattaneo et al. [2020] at 95% confidence levels.[13] Note that all of the 37 subjects with almost deterministic consideration satisfied the Consideration Filter property, hence, they trivially satisfy Monotonic Consideration as it is the stochastic version of Competition Filter. Since 49 subjects satisfy Monotonic Consideration in Table 4 only 4% of the subjects satisfy Monotonic Consideration non-trivially.[14]

Although we are not aware of any other research testing these axioms, the limited support for them in our data is not completely unexpected. We will revisit this discussion in Section 5, but for now, one should note that these consideration axioms are similar to the regularity condition (see subsection 4.4 for its definition) which is well-known in the stochastic choice domain. The regularity in choice is challenged both theoretically and empirically, hence, it is somewhat expected to see limited support for its analogous properties in the consideration domain.

TABLE 4. The percentage (number) of subjects who are consistent with models of stochastic consideration.

Property Name	% (#) satisfying
Monotonic Consideration Cattaneo et al. [2020]	15.6% (49)
Consideration Overload Cattaneo et al. [2023]	28.6% (90)

4.3. **Descriptive Analysis of Choice.** In this subsection, we analyze the choice data following the same structure that we used for the analysis of the consideration data. We first

[13] Additional econometric detail is provided in Appendix C.

[14] The independent attention model (Manzini and Mariotti, 2014) and the logit consideration model (Brady and Rehbeck, 2016) are special cases of monotonic consideration and hence they cannot be satisfied by our mostly non-monotonic subjects who were acting stochastically.

start with the descriptive analysis of Part II data followed by reporting the results from axiom testing.

4.3.1. *Stochastic vs. Deterministic Choice.* Deterministic behavior in choice decisions is much more frequent than it is for consideration decisions. Out of 315 subjects, 47 subjects exhibit deterministic choice in all questions. There are an additional 27 subjects who are almost deterministic, i.e., only in 1 out of 110 questions they act stochastically. Hence, the stochastic behavior is less pronounced in choice problems than in consideration problems. We are not aware of any other results from the literature documenting this observation. Moreover, the subjects who choose stochastically do not act that way on all the menus (see Figure 10.) Recall that Figure 6 was a counterpart of Figure 10 and notably, the stochastic consideration summarized in Figure 6 was much more skewed to the left than that in Figure 10 indicating further evidence for a higher rate of stochasticity in the consideration domain than the choice domain.

In Figure 11, we document the frequency of subjects with stochastic choice by menu, and again the randomness in choice increases when the menu is larger as the bars in this figure tend to be taller for the menus of sizes 3 and 4.

Result 3: Descriptive observations on choice behavior are listed below.

i. 74 subjects choose almost deterministically;
ii. Stochastic choice is less likely than stochastic consideration;

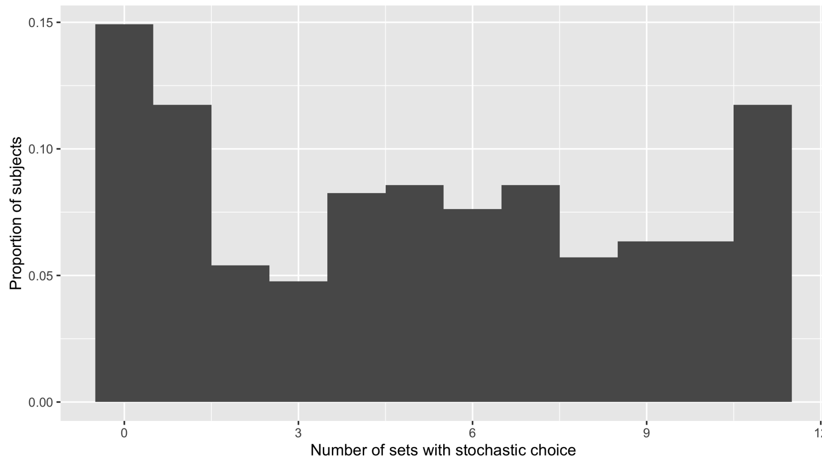


FIGURE 10. Distribution of subjects by the number of sets that have stochastic choice.

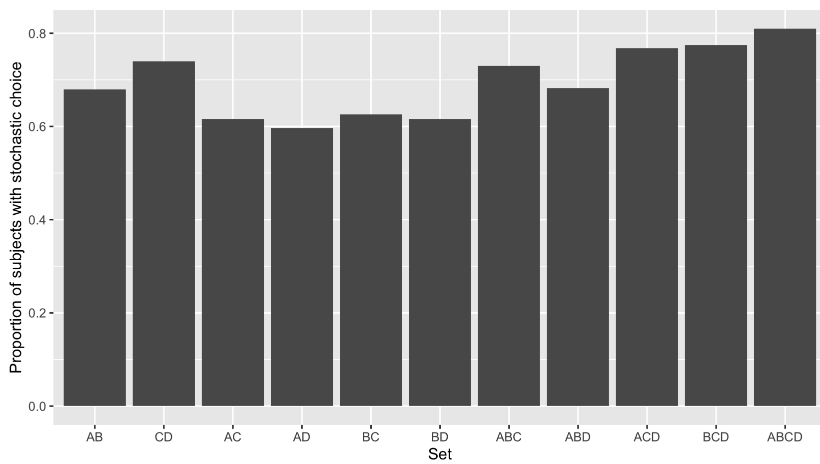


FIGURE 11. Fraction of subjects with stochastic choice by set.

4.4. **Testing Choice Axioms.** Recall that 47 subjects chose fully deterministically and 27 chose almost deterministically (by deviating from a deterministic choice behavior in only 1 out of 110 questions). For these subjects, it is meaningful to check the Weak Axiom of Revealed Preferences (WARP) as the main axiom of rationality in standard choice theory.[15] Strikingly, all of these 74 subjects satisfy WARP.

[15]WARP, stated for choice correspondences $c : 2^X \to 2^X$, states that if for all $S \subseteq X$ and $y \in S$, if there exists an $x \in S$ such that $y \in c(T)$ for some $T \subseteq X$ with $x \in T$, then $y \in c(S)$.

Next, we analyze the data for other well-known stochastic choice axioms. We focus on three classes of axioms: versions of regularity (how menu size affects the choice probabilities), versions of transitivity (stochastic generalizations of the standard transitivity axiom), and versions of independence (how relative choice probabilities depend on the menu.) We first define these axioms below and then report the subjects whose behavior cannot reject the corresponding axiom.

Random Utility theories are offered by the literature to explain stochastic choice behavior. The regularity is a standard axiom for such models. The regularity requires that the chance of choosing an option cannot decrease when another option is removed from the menu.

Regularity. For any $x \neq y \in S$, $\pi(x \mid S) \leq \pi(x \mid S \setminus \{y\})$.

Several empirical studies documented robust violations of regularity due to the asymmetrical dominance effects or attraction effects (see Rieskamp et al., 2006, Section 5 for a review.) Filiz-Ozbay and Masatlioglu [2023] weaken the regularity condition by requiring it to be satisfied by some binary subset of a menu rather than with respect to every subset. In other words, the probability of choosing an option on a rich menu cannot exceed its probability of being chosen in all binary comparisons of that option within the menu. This condition allows for regularity violations but does not allow them to happen with respect to every binary comparison.

Weak Binary Regularity. For any $S \subseteq X$ and $x \in S$, $\pi(x \mid S) \leq \max_{y \in S \setminus \{x\}} \{\pi(x \mid \{x, y\})\}$.

Echenique and Saito [2019] provides the following weak regularity property which applies to the options that have zero weight in the stochastic choice. It requires that if an option y is selected with zero chance against option x then y must be selected with zero chance on any menu that contains x. This axiom is relevant for our design, as we observe several participant never choosing some options.

Weak Regularity. For any $x, y \in X$ and $S \subseteq X$ such that $x, y \in S$, $\pi(y \mid \{x, y\}) = 0 \Rightarrow \pi(y \mid S) = 0$.

The transitivity axiom in deterministic theories has been a cornerstone because it implies utility representation (Samuelson, 1953). The Strong Stochastic Transitivity is proposed as a natural extension of the deterministic version of transitivity and it is implied by all fixed utility theories such as Luce's (Rieskamp et al., 2006). Many empirical studies found violations of this property (see Mellers and Biagini, 1994, for a review.) The Weak Stochastic Transitivity and Moderate Stochastic Transitivity are proposed as the weaker versions of it (He and Natenzon, 2024). The moderate stochastic transitivity is often attributed to Chipman [1958, 1960] and it characterizes the moderate utility model (He and Natenzon, 2024.) Earlier research found that the violations of these weaker notions of stochastic transitivity are rare –consistent with our data as we will report shortly (Rieskamp et al., 2006.)

Weak Stochastic Transitivity: $\min\{\pi(x|\{x,y\}),\pi(y|\{y,z\})\} \geq 0.5$ implies $\pi(x|\{x,z\}) \geq 0.5$.

Moderate Stochastic Transitivity: $\min\{\pi(x|\{x,y\}),\pi(y|\{y,z\})\} \geq 0.5$ implies $\pi(x|\{x,z\}) \geq \min\{\pi(x|\{x,y\}),\pi(y|\{y,z\})\}$.

Strong Stochastic Transitivity: $\min\{\pi(x|\{x,y\}),\pi(y|\{y,z\})\} \geq 0.5$ implies $\pi(x|\{x,z\}) \geq \max\{\pi(x|\{x,y\}),\pi(y|\{y,z\})\}$.

Another well-studied principle in stochastic choice domains is the independence of irrelevant alternatives. This principle relies on the idea that the relative probabilities of choosing two options should not vary when each is compared with respect to options in menu Q or R.[16] Below we state a version of the Independence property studied by Gul et al. [2014] and for that we first need to define the probability of a set Y being chosen in menu Z by adding the probability of each element in Y being chosen in Z, i.e., $\psi(Y \mid Z) = \sum_{y\in Y}\pi(y \mid Z)$.

Independence: For any non-empty Q,R,S,T, $\psi(S \mid S\cup Q) \geq \psi(T \mid T\cup Q)$ implies $\psi(S \mid S\cup R) \geq \psi(T \mid T\cup R)$ if $Q,R \in 2^X \setminus \varnothing$ and $(S\cup T)\cap(Q\cup R) = \varnothing$.

Gul et al. [2014] shows that Independence in rich domains is equivalent to the Luce model for stochastic choice. Since the Luce model has unintuitive implications when there are some similarities between options[17], they weaken the Independence property by formally introducing a notion of duplicates and checking Independence for non-duplicates.

[16] Luce [1959] provides a stronger version of the independence notion which requires not only the order of probabilities of the two options but also the ratios of choice probabilities to be independent of what else is in the menu.

[17] Debreu [1960] intuitively described a decision problem for a violation of the independence principle where the introduction of an option diminishes the likelihood of choice for other similar options in the menu without affecting the weights on the nonsimilar ones.

Duplicates: S and T are duplicates if for any Q such that $Q \cap (S \cup T) = \varnothing$, $\pi(x \mid Q \cup S) = \pi(x \mid Q \cup T)$ for any $x \in Q$. Duplicates T and S are denoted by $T \sim S$.

Since the notion of duplicates is introduced for identifying similarities in the stochastic choice domain, the upcoming analysis first focuses on the 241 subjects with non-deterministic choice (i.e., excluding 74 almost deterministic subjects satisfying WARP.) Only 12 of these subjects have no duplicates. For most subjects A and B are duplicates followed closely by C and D (by 166 and 153 subjects, respectively.) This is inline with our choice of these lotteries because the prospects of A and B are similar both in terms of outcomes and probabilities associated to them; C and D are only similar in terms of the range of outcomes but not in terms of probabilities.

The next two axioms, Weak independence and Elimination of Duplicates (Gul et al., 2014), require versions of independence property to hold for non-overlapping comparisons.[18]

Weak Independence: For any non-empty Q, R, S, T such that $(S \cup T) \perp (Q \cup R)$, $\psi(S \mid S \cup Q) \geq \psi(T \mid T \cup Q)$ implies $\psi(S \mid S \cup R) \geq \psi(T \mid T \cup R)$.

Elimination of Duplicates. For any $T \sim S' \subseteq S \perp Q$ and $x \in Q$, $\pi(x \mid S \cup Q) = \pi(x \mid T \cup S \cup Q)$.

Elimination of duplicates indicates that the duplicates are treated equally by the decision maker. More precisely, if T and S' are duplicates then adding T to a menu that contains S' does not affect the probability of an option being chosen from a non-overlapping set Q. In our experiment, for example, if a subject viewed options A and B as duplicates, then likelihood of choosing C from $\{B, C\}$ and $\{A, B, C\}$ should be the same in order to satisfy this property.

Table 5 starts with our earlier observation of 74 subjects who almost always chose deterministically satisfying WARP. Since all the other stochastic choice axioms of interest imply WARP for the deterministic behavior, these 74 people will trivially satisfy the remaining stochastic axioms in Table 5. To highlight this point, the second column of the table reports the results for each stochastic axiom as the sum of those who fail to reject the axiom in a non-trivial way (by acting stochastically) and those who act consistent with WARP.

[18] Two sets are called non-overlapping if all their duplicates are pairwise disjoint, i.e. T and S are called non-overlapping if $T \sim T'$, $S \sim S'$ implies $T' \cap S' = \varnothing$. Non overlapping sets T and S are denoted by $T \perp S$.

Table 5. The percentage (number) of subjects consistent with models of deterministic or stochastic choice.

Choice Property	% (#) satisfying
WARP - (Almost) Deterministic	**23.5% (74)**
Regularity	**7.9%**+23.5% (**25**+74)
Weak Binary Regularity	**57.8%**+23.5% (**182**+74)
Weak Regularity	**41.0%**+23.5% (**129**+74)
Weak Stochastic Transitivity	**76.2 %**+23.5% (**240**+74)
Moderate Stochastic Transitivity	**73.3 %**+23.5% (**231**+74)
Strong Stochastic Transitivity	**28.9 %**+23.5% (**91**+74)
Independence	**27.0 %**+23.5% (**85**+74)
Weak Independence	**26.3%**+23.5% (**83**+74)
Elimination of Duplicates	**73.7%**+23.5% (**232**+74)

Tests are performed using the estimation techniques adopted from Cattaneo et al. [2020] at 95% confidence levels. Deterministic or almost deterministic choices that satisfy WARP (74 subjects in the first row) also satisfy the other stochastic properties trivially.

Similar to the earlier studies, only 7.9% satisfy the Regularity in a non-trivial way in addition to those who satisfy it because they act consistent with WARP. Each of the two weaker versions of this property explains the behavior of the majority. Again consistent with the literature only 28.9% fails to reject the Strong Stochastic Transitivity in a non-trivial way. This number increases to 73.3% for the moderate version. Weakening it further to the Weak Stochastic Transitivity does not improve the results much (adds only 3 percentage points to the pool of subjects who acted consistent with the rule.)

We are not aware of any other study empirically testing Independence and Weak Independence. We find that 27% of the subjects act consistent with the Independence but weakening it by checking it only for non-overlapping comparisons does not improve the results.[19] On the contrary, almost all the subjects acted consistent with the Elimination of Duplicates either trivially or non-trivially.

While we only focused on the most well-known axioms in the classes of regularity, transitivity, and independence in Table 5, one may extend this analysis to any stochastic choice axiom of interest since we sample the stochastic choice function, π, on every menu, i.e. the whole domain. The correlations between different choice properties that are analyzed in this section are reported in Table 9 in Appendix D.

4.5. **Consistency of choice with incomplete preferences.** We observe the choices of individuals through repeated observations in Part II data. So far we have interpreted that data as sampling from an underlying stochastic choice function. Alternatively, Balakrishnan et al. [2022] builds choice correspondences using such data, infers the underlying (possibly incomplete) binary relation, and studies its properties. In this subsection, we will depart from the standard stochastic choice interpretation of the frequently observed choice behavior, and treat our data in the spirit of Balakrishnan et al. [2022].

This is a two-part exercise to test for consistency with potentially incomplete preference relations. First, we utilize the notion of Fishburn correspondences from Balakrishnan et al. [2022] to convert our stochastic choice data to a choice correspondence. Based on $\lambda \in [0,1]$ that is prespecified by the analyst, a choice correspondence $c : 2^X \to 2^X$ is generated from a probability distribution of choices over sets, π, as:

$$c(S) = \left\{ x \in S \mid \pi(x \mid S) \geq \lambda \cdot \max_{y \in S} \pi(y \mid S) \right\}$$

Smaller values of λ results in a more permissible conversion; in the extreme case of $\lambda = 0$, the choice correspondence is $c(S) = S$. In the other extreme case of $\lambda = 1$, the choice correspondence consists solely of alternatives with maximal choice probability.

Second, after generating a choice correspondence, $c(S)$, for a given λ, we evaluate the consistency of this correspondence through the Weak Axiom of Revealed Non-Inferiority

[19]The fact that slightly more subjects satisfy Independence than Weak Independence appears to be a limitation of testing methodology. See Appendix C for more details.

(WARNI) from Eliaz and Ok [2006b]. In our environment, using our notation, the axiom reads as:

Weak Axiom of Revealed Non-Inferiority (WARNI): For any $S \subseteq X$ and $y \in S$, if *for every* $x \in c(S)$ there exists a $T \subseteq X$ with $y \in c(T)$ and $x \in T$, then $y \in c(S)$.

Intuitively, the condition relaxes WARP from "revealed preferred to" to "not revealed inferior to"; if there is no $x \in S$ such that y is never chosen when x is present for all $T \supseteq \{x, y\}$, then the implication is that y is $\succeq$-maximal for some $\succeq$, and thus must be chosen in S.[20] However, unlike WARP, the $\succeq$ used to rationalize choice data may be incomplete, and thus multiple alternatives may be $\succeq$-maximal, generating a choice correspondence. Theorem 2 from Eliaz and Ok [2006b] states that the axiom is necessary and sufficient to have a unique, but not necessarily complete, regular preference relation $\succeq$ on X such that $c = \max(\cdot, \succeq)$.

We conduct the exercise for $\lambda = 0.1, 0.2, \ldots, 1$.[21] Figure 12 presents the proportion of subjects that satisfy WARNI for each λ level. Overall, we observe a large majority of subjects satisfying WARNI for all λ levels.[22] For $\lambda = 1$, 249 (79.0%) of subjects satisfy WARNI. This proportion is monotonically decreasing in λ, with 312 (99%) of subjects satisfying WARNI when $\lambda = 0.1$. Hence, while only a quarter of subjects are (almost) deterministic and satisfy WARP, the necessary and sufficient condition for complete preference maximization, the vast majority of subjects are consistent with the concept of incomplete preference maximization.

[20] The set of all $\succeq$-maximal elements from a set S is defined as $\{x \in S : y \succ x \text{ for no } y \in S\}$. An element in the set is said to be $\succeq$-*maximal*.

[21] $\lambda = 0$ trivially satisfies WARNI, since a perfectly incomplete preference relation rationalizes choices.

[22] Note that $\lambda = 1$ does not necessarily test for WARP in cases where $\arg\max_{y \in S} \pi(y \mid S)$ is not unique.

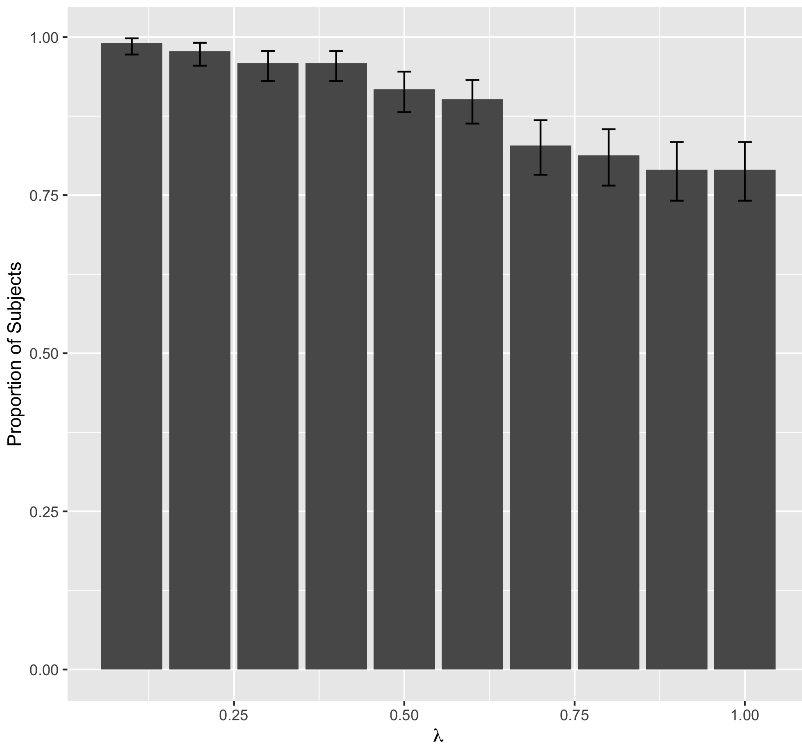


FIGURE 12. Percentage of 315 subjects satisfying WARNI, by λ.

4.6. **Combining Part I and Part II data sets.** The existing two-stage stochastic models, where the decision maker chooses from what is considered, are built primarily on the choice data. For example, Manzini and Mariotti [2012] or Cattaneo et al. [2020] assume the choice data as the only observables of the model and construct the underlying consideration or attention functions for the first stage based on their modeling assumptions. This approach became standard because the consideration data is typically unobservable. Our data is free from that limitation because we observe what each subject considered and what she chose on every menu. In the following analysis, we will analyze what properties the choice behavior of an individual would satisfy if we forced them to implement the two-stage procedure based on their actual consideration and choice.

For this purpose, we first create a hypothetical choice function for each subject, π^*, based on the observed μ and π of that subject. This is interpreted as how the subject would behave if she first considered according to her μ function and then chose from what she considered according to her π function:

$$\pi^* = (x \mid S) = \sum_{\substack{T \subset S \\ x \in T}} \mu(T \mid S)\pi(x \mid T)$$

π^* would be the ideal data to test the characterization axioms of the two-stage models of the literature. If we perform that analysis, we see that for 289 subjects (91.7%), their constructed π^* is consistent with both monotonic consideration and consideration overload. According to those models, all these subjects' choice data can be represented by a two-stage procedure where the consideration stage satisfies the relevant monotonicity property. Then, how can one explain the low support for monotonic consideration and consideration overload in our Part I data as reported in Table 4 (15.6% and 28.6%, respectively), while the choice data implied by that consideration function being perfectly in line with the corresponding model? We argue that this exercise does not falsify those models necessarily because those models do not rule out a consideration behavior generating the same choice while violating monotonicity because the representation is not unique. They only claim the existence of a monotonic consideration function that is consistent with the data. Our exercise with π^* argues that the seemingly weak properties of unobservable parts (consideration parts) of these models may not be satisfied by the actual function as it is the case with our Part I data.

Additionally, we can compare the observed choice behavior, π, in Part I and implied choice behavior by the imposed two-stage choice, π^* to understand the impact of consideration stage. Adding a consideration stage to the choice problem changes the choice behavior non-trivially. For 15 subjects, π^* and π are identical. In total, for 146 subjects we fail to reject the hypothesis that $\pi = \pi^*$ at the 5% level.[23]

All the axioms of choice in Table 5 can be tested for π^*, as well. We report this analysis in Table 10 in Appendix D. As expected π^* satisfies WARP for less subjects than π did due to more stochasticity imposed by the two stages in the construction of π^*.

[23]To test whether $\pi = \pi^*$, a Wilcoxon two-sample paired signed-rank test is performed. An initial check determines whether the distributions are exactly the same - if so, no test is run and subjects are classified as having $\pi = \pi^*$. The remaining subjects undergo the test.

5. Revisiting consideration data

We have noted earlier that the monotonicity property for stochastic consideration was analogous to the regularity property in stochastic choice. Regularity in choice is often violated (including our data) and monotonicity of consideration (the Consideration Overload property) was also violated by the majority of our subjects. To fit the data better in the choice framework, models such as the nested logit and the attribute rule are offered. The idea of these models is to split the set of alternatives into classes based on shared attributes of options and the characterization axioms are checked only within class (Kovach and Tserenjigmid, 2022 for the axiomatization of nested logit, and Gul et al., 2014 for attribute rule.) In this section, we will apply the classification idea of the choice theory to our consideration data, and offer a new property of consideration overload which is only checked with respect to classes.

The theory side of consideration literature is more limited than the choice literature. So we do not have much guidance from the existing theories on how to generalize the consideration overload property. Nevertheless, the classification idea of choice is natural in our setting, as we intentionally designed our ambiguous lotteries so that some options might be classified more similar to each other than the other options (recall that A and B are the high range ambiguous lotteries and C and D are the low range ambiguous lotteries by design.) Hence, we may naturally classify the options in the consideration problem based on option similarity and check the consideration overload with respect to the similarity classes. Nevertheless, this might be ad-hoc because even though we designed those pairs as similar, the subjects may not have viewed them as similar. To address this, below we formally define "Consideration Duplicates" which motivated from the notion of "duplicates" in choice theory by Gul et al. [2014] that we used in Subsection 4.4.

Consideration Duplicates. Alternatives x and y are considered duplicates for consideration function μ if for all sets S such that $\{x, y\} \cap S = \varnothing$, and for all $W \subseteq S$, $\mu(W \mid x \cup S) = \mu(W \mid y \cup S)$.

This notion means that the probability of considering W in a menu is the same when either of the two duplicates is added to the menu. Hence, the impact of two consideration duplicates on considering set W are the same. For example, for a subject who views A and B as duplicates, her likelihood of considering C when it is presented together with A versus B must be the same, i.e., we must have $\mu(\{C\} \mid \{A, C\}) = \mu(\{C\} \mid \{B, C\})$. 129 subjects

(41%) viewed A and B as duplicates and 80 subjects (25%) viewed C and D as duplicates at the consideration stage.

Next, we offer a generalization of the Consideration Overload property offered only for the removal of duplicates rather than arbitrary subsets.

Consideration overload due to duplicates: For any T, S, x such that $x \in T \subseteq S \subseteq X$ and x is a duplicate to any $y \in S \setminus T$, $\phi(x \mid S) \leq \phi(x \mid T)$.

Note that if a subject does not view any alternatives as duplicates, the property is automatically satisfied. Hence, for half of the subjects (167 subjects), the condition is trivially satisfied. When we check the property for those who viewed A and B or C and D (or both) as duplicates (148 subjects), 88% of them satisfy Consideration overload due to duplicates. The support for the weaker version of the Consideration Overload is similar to the impact of the duplicate notion on Independence property of choice in Table 5.

6. Conclusion

We design and implement an experiment to fully observe consideration and choice functions over a set of four alternatives. We generate multiple observations per choice set, providing us with an estimate of potentially stochastic consideration and choice. Hence, our data set can be used to evaluate any existing and future axioms of choice and/or consideration. Overall, we find results consistent with the literature on previously tested choice axioms such as regularity and stochastic transitivity, and test additional choice axioms that have not been previously tested, such as weak independence. For consideration, we find limited support for monotonicity, both on the consideration function and on functions of alternative-level consideration frequency. We suggest a generalization of these axioms in the spirit of duplicates implemented in choice theory.

References

Abaluck, J. and A. Adams-Prassl (2021): "What Do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses," *The Quarterly Journal of Economics*, 136, 1611–1663.

Agranov, M. and P. Ortoleva (2017): "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, 125, 40–68.

Aguiar, V. H., M. J. Boccardi, N. Kashaev, and J. Kim (2023): "Random Utility and Limited Consideration," *Quantitative Economics*, 14, 71–116.

Balakrishnan, N., E. A. Ok, and P. Ortoleva (2022): "Inference of Choice Correspondences," *Working Paper, Princeton University.*

Brady, R. L. and J. Rehbeck (2016): "Menu-Dependent Stochastic Feasibility," *Econometrica*, 84, 1203–1223.

Cattaneo, M. D., P. Cheung, X. Ma, and Y. Masatlioglu (2023): "Attention Overload," .

Cattaneo, M. D., X. Ma, Y. Masatlioglu, and E. Suleymanov (2020): "A Random Attention Model," *Journal of Political Economy*, 128, 2796–2836.

Chen, Y.-J. and J. E. Corter (2006): "When mixed options are preferred in multiple-trial decisions," *Journal of Behavioral Decision Making*, 19, 17–42.

Cherepanov, V., T. Feddersen, and A. Sandroni (2013): "Rationalization," *Theoretical Economics*, 8, 775–800.

Chipman, J. S. (1958): "Stochastic Choice and Subjective-Probability," *Econometrica*, 26, 613–613.

——— (1960): "Stochastic choice and subjective probability," in *Decisions, Values, and Groups, Vol. I*, ed. by D. Willner, Pergamon Press, 70–95.

Debreu, G. (1960): "Review of R. D. Luce, Individual Choice Behavior: A Theoretical Analysis," *American Economic Review*, 50, 186–188.

Draganska, M. and D. Klapper (2011): "Choice set heterogeneity and the role of advertising: An analysis with micro and macro data," *Journal of Marketing Research*, 48, 653–669.

Echenique, F., S. Lee, and M. Shum (2011): "The money pump as a measure of revealed preference violations," *Journal of Political Economy*, 119, 1201–1223.

Echenique, F. and K. Saito (2019): "General Luce model," *Economic Theory*, 68, 811–826.

Eliaz, K. and E. A. Ok (2006a): "Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences," *Games and Economic Behavior*, 56, 61–86.

——— (2006b): "Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences," *Games and economic behavior*, 56, 61–86.

Filiz-Ozbay, E. and Y. Masatlioglu (2023): "Progressive random choice," *Journal of Political Economy*, 131.

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," *Econometrica*, 82, 1873–1912.

He, J. and P. Natenzon (2024): "Moderate utility," *American Economic Review: Insights*, 6, 176–195.

Hey, J. D. (2001): "Does Repetition Improve Consistency?" *Experimental Economics*, 4, 5–54.

Hey, J. D. and C. Orme (1994): "Investigating generalizations of expected utility theory using experimental data," *Econometrica: Journal of the Econometric Society*, 1291–1326.

Honka, E. and P. Chintagunta (2017): "Simultaneous or Sequential? Search Strategies in the US Auto Insurance industry," *Marketing Science*, 36, 21–42.

Kashaev, N. and V. H. Aguiar (2022): "A random attention and utility model," *Journal of Economic Theory*, 204, 105487.

Kovach, M. and G. Tserenjigmid (2022): "Behavioral foundations of nested stochastic choice and nested logit," *Journal of Political Economy*, 130, 2411–2461.

Kreps, D. M. (1988): *Notes on the Theory of Choice*, Underground Classics in Economics, Westview Press.

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay (2017): "When More Is Less: Limited Consideration," *Journal of Economic Theory*, 170, 70–85.

Luce, R. D. (1959): *Individual choice behavior*, Wiley, New York.

Manzini, P. and M. Mariotti (2007): "Sequentially Rationalizable Choice," *American Economic Review*, 97, 1824–1839.

——— (2012): "Categorize then choose: Boundedly rational choice and welfare," *Journal of the European Economic Association*, 10, 1141–1165.

——— (2014): "Stochastic Choice and Consideration Sets," *Econometrica*, 82, 1153–1176.

Mariotti, M. and P. Manzini (2012): "Choice by lexicographic semiorders," *Theoretical Economics*, 7.

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012): "Revealed Attention," *American Economic Review*, 102, 2183–2205.

McCausland, W. J., C. Davis-Stober, A. Marley, S. Park, and N. Brown (2020): "Testing the random utility hypothesis directly," *The Economic Journal*, 130, 183–207.

McFadden, D. (1977): "Modelling the Choice of Residential Location," *Cowles Foundation Discussion Paper*.

McFadden, D. and M. K. Richter (1990): "Stochastic rationality and revealed stochastic preference," *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz, Westview Press: Boulder, CO*, 161–186.

Mellers, B. A. and K. Biagini (1994): "Similarity and Choice," *Psychological Review*, 101, 505–518.

Ok, E. A. and G. Tserenjigmid (2023): "Measuring Stochastic Rationality," *arXiv preprint arXiv:2303.08202*.

Reutskaja, E., R. Nagel, C. F. Camerer, and A. Rangel (2011): "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study," *American Economic Review*, 101, 900–926.

Rieskamp, J., J. R. Busemeyer, and B. A. Mellers (2006): "Extending the Bounds of Rationality: Evidence and Theories of Preferential Choice," *Journal of Economic Literature*, 44, 631–661.

Rubinstein, A. (2002): "Irrational diversification in multiple decision problems," *European Economic Review*, 46, 1369–1378.

Samuelson, P. (1953): *Foundations of economic analysis*, Harvard University Press.

Thomas, A. W., F. Molter, and I. Krajbich (2021): "Uncovering the computational mechanisms underlying many-alternative choice," *Elife*, 10, e57012.

Appendices

A. Instructions

Section I

In this Section, you will be presented a collection of bins. Your task is to select which bins' content to examine and to choose one among from the examined bins. The bin you choose at the end will determine your payoff.

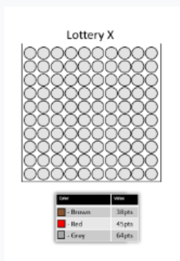
What is a bin?

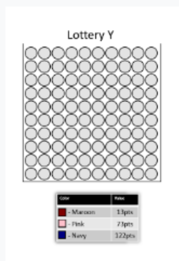
A bin corresponds to a lottery with monetary prizes. There are 100 balls in each bin and the color of a randomly drawn ball from a bin determines the prize. You will know the prize associated to each color in a bin, but in order to learn number of balls in each color in a bin, you need to examine the bin further.

For illustration of your task, we show you below two lotteries corresponding to two bins. Note that each bin has 100 balls with unidentified color compositions.

The first bin consists of Brown, Red and Grey balls. The corresponding lottery X pays you based on the color of a randomly drawn ball from this bin. If a randomly drawn ball from this bin is Brown, Lottery X pays 38 points. If it is Red, it pays 45 points. If it is Grey, it pays 64 points.

The second bin, corresponding to lottery Y, pays 73 points if it is Pink, 122 points if it is Navy, and 13 points if it is Maroon.





Choosing bins to examine

Since you don't know how many balls there are in each color in a bin, at this stage you cannot know how likely it is to draw a certain color. You may learn the exact color composition of a bin, by choosing to examine that bin.

YOUR TASK IS TO CHOOSE WHICH BINS YOU WANT TO EXAMINE. You may examine any number of bins you want!

The content of a bin that you selected for further examination will be revealed to you. For example, the figure below shows you the content of the first bin above corresponding to lottery X.

Color	Number of Balls
Red	
Brown	
Grey	

You need to count each color in a bin you selected, and type those numbers correctly in the corresponding box. For example, there are 21 Brown, 33 Red, and 46 Grey balls in the bin of lottery X. So you have to enter these numbers correctly for the corresponding colors.

What happens in a round?

In this experiment there are four bins: A, B, C, and D. In each round, you will be presented with a subset of bins A, B, C, and D. You will know the corresponding prizes on each bin but you will not know the color combinations unless you choose to examine them.

From the set of presented bins, choose the ones you want to examine and learn the color content.

In each round you will make such decisions on a set of bins.

There are 115 rounds with similar decision problems. Once 115 rounds are completed, one of the rounds will be randomly selected and you will examine the bins that you chose to examine for that round.

Remember that on the bins you examine, you need to count the colored balls and enter the number of balls in each color correctly. If you make any mistakes, your payment in this section will be zero.

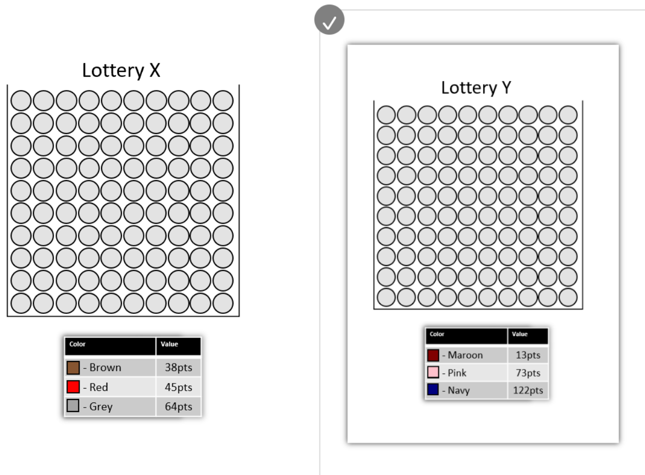
If you enter the number of balls in each color correctly, then you will be asked to choose one of the bins you examined for that round and your payment for this section will be the payoff of the randomly drawn ball from the bin you chose.

Next

Figure 13. Part I instructions

Sample question 1 out of 3:

Please see the available lotteries with the corresponding prizes below and click on the ones you want to examine further to understand the color content. You may choose as many or as few as you want to examine later. Note that if this round is selected at the end of the sample questions, you need to count the balls in all the bins you selected to examine in this round correctly.



Next

FIGURE 14. Section 1 sample problem

Section 2

In this section you are dealing with the same bins A, B, C, and D as in Section 1. This time you **will know** the color composition of each bin. Therefore, you will know the likelihood of drawing a ball in certain color from any bin.

In each round we will show you a subset of bins A, B, C, and D. Your task is to choose exactly one bin to draw a ball from. For example, if you are offered bins {A, B, C} and if you choose bin A, this means you want a ball to be randomly drawn from A for this round and be paid according to that draw.

There are 110 rounds in this section. In each round you have to pick exactly one bin to base your payment on.

Once you finish all the rounds of the section, one round will be randomly drawn, and we will draw a random ball from the bin you selected for that round.

Next

FIGURE 15. Part II instructions

B. Demographics

Table 6. Gender identity frequency

Gender	Frequency
Male	201
Female	189
Other	9
Did not respond	3

Table 7. Age bracket frequency

Age bracket	Frequency
18-27	129
28-37	140
38-47	76
48-57	35
58+	21
Did not answer	1

Table 8. Educational attainment frequency

Highest education level	Frequency
Less than high school diploma	4
High school diploma or GED	35
Some college, but no degree	95
Associate's Degree (for example: AA, AS)	36
Bachelor's Degree (for example: BA, BBA, and BS)	159
Master's Degree (for example: MA, MS, and MEng)	52
Professional Degree (for example: MD, DDS, JD)	13
Doctorate (for example: PhD, EdD)	5
Did not answer	3

C. Econometric Tests

In order to test consideration and choice axioms, we closely follow the framework of Cattaneo et al. [2020], whereby each axiom is tested by using an appropriate set of linear inequalites. The procedure is similar for all axioms, we will only explain testing the monotonic consideration.

First, let

$$\mu = [\mu(\cdot \mid \{A, B, C, D\}), \mu(\cdot \mid \{A, B, C\}), \dots, \mu(\cdot \mid \{C, D\})]'$$

where $\mu(\cdot \mid S)$ represent the vector of probabilities for menu S.

Second, construct a maxtrix R_{MC} such that each cell of the matrix is based on the inequalities of the monotonic consideration axiom: for each set S, for each alternative $x \in S$, for each $T \in S - x$, $\mu(T \mid S) \leq \mu(T \mid S - x)$. Similar to the construction in Cattaneo et al. [2020], a "1" is entered in the column corresponding to $\mu(T \mid S)$, and a "-1" is entered in the column corresponding to $\mu(T \mid S - x)$. All other entries are "0". For example, consider three alternatives x, y, and z. In this case, μ is the column vector

$$\mu = [\mu(x \mid xyz), \mu(y \mid xyz), \mu(z \mid xyz), \mu(xy \mid xyz), \mu(xz \mid xyz), \mu(yz \mid xyz), \mu(xyz \mid xyz)$$

$$\mu(x \mid xy), \mu(y \mid xy), \mu(xy \mid xy), \mu(x \mid xz), \mu(z \mid xz), \mu(xz \mid xz), \mu(y \mid yz), \mu(z \mid yz), \mu(yz \mid yz)]'$$

R_{MC} is constructed from all the inequalities:

S	x	T	x|xyz	y|xyz	z|xyz	xy|xyz	xz|xyz	yz|xyz	xyz|xyz	x|xy	y|xy	xy|xy	x|xz	z|xz	xz|xz	y|yz	z|yz	yz|yz
xyz	x	y	0	1	0	0	0	0	0	0	0	0	0	0	0	-1	0	0
xyz	x	z	0	0	1	0	0	0	0	0	0	0	0	0	0	0	-1	0
xyz	x	yz	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	-1
xyz	y	x	1	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0
xyz	y	z	0	0	1	0	0	0	0	0	0	0	0	-1	0	0	0	0
xyz	y	xz	0	0	0	0	1	0	0	0	0	0	0	0	-1	0	0	0
xyz	z	x	1	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
xyz	z	y	0	1	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
xyz	z	xy	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0

Consideration sets are assumed to be generated by a true underlying μ, with observations forming an empirical distributions $\hat{\mu}$. The null hypothesis is that the generating process μ satisfies the condition $R_{MC}\mu \leq 0$. The following tests the alternative hypothesis that $R_{MC}\mu \nleq 0$.

Following Cattaneo et al. [2020], the standard deviation is $\sigma_\mu = \sqrt{\text{diag}(R_{MC}\Omega_\mu R'_{MC})}$, where Ω_μ is block diagonal, with blocks given by $\{1/11\}\Omega_{\mu,S}$ and $\Omega_{\mu,S} = \text{diag}(\mu_S) - \mu_S\mu'_S$. The estimate of standard deviation, $\hat{\sigma}$, is the estimate of σ_μ using observed data: $\hat{\sigma} = \sqrt{\text{diag}(R_{MC}\hat{\Omega}R'_{MC})}$. Ω_μ is estimated as $\hat{\Omega}$ by plugging in the empirical estimate $\hat{\mu}$.

The test statistic is $T = \sqrt{N} * \max\{R_{MC}\hat{\mu} \oslash \hat{\sigma}, 0\}$, where $\oslash$ represents elementwise division (aka Hadamard division) and the maximum is the largest element of $R_{MC}\hat{\mu} \oslash \hat{\sigma}$ if it is positive, or zero otherwise. Intuitively, the smaller the test statistic value is, the closer $\hat{\mu}$ is to satisfying all inequalities.

The critical value of the test statistic is calculated by randomly sampling a normal distribution plus an additional estimate for moment conditions. By adding and subtracting $\sqrt{N} * R_{MC}\mu$, T becomes $T = \max\{(R_{MC}\sqrt{N}(\hat{\mu} - \mu) + \sqrt{N}R_{MC}\mu) \oslash \hat{\sigma}, 0\}$. The first term is approximately normal with variance Ω_μ, whereas under the null hypothesis the second term is bounded above by zero, but the "less conservative" estimate used in Cattaneo et al. [2020] estimates it using $\frac{1}{\sqrt{\log N}}(R_{MC}\hat{\mu} \oslash \hat{\sigma})_-$, where $(a)_- = a \circ \mathbb{I}(a \leq 0)$ and $\circ$ is elementwise multiplication (aka Hadamard product).

Thus the simulated values are $T^* = \sqrt{N} \cdot \max\{(R_{MC}z^*) \oslash \hat{\sigma} + \frac{1}{\sqrt{\log N}}(R_{MC}\hat{\mu} \oslash \hat{\sigma})_-, 0\}$, where z^* is simulated from the normal distribution $N(0, \hat{\Omega}/N)$. The critical value comes from sampling T^* 1,000 times, and the $(1-\alpha)\%$ percentile of simulated values becomes the critical value: $c_\alpha = \inf\left\{t \mid \frac{1}{1000}\sum_{m=1}^{1000}\mathbb{I}(T^* \leq t) \geq 1 - \alpha\right\}$. The null hypothesis is rejected if $T > c_\alpha$. α is set to be 0.05.

C.1. **Nested Axioms.** The test statistic distribution $\mathcal{T}^*$ is generated using a 28-dimensional normally distributed random variable z with mean 0 and standard deviation $\hat{\Omega}/N$, where N is the sample size of 110 and $\hat{\Omega}$ is generated using the distribution π. The distribution multiplies the inequality matrix R with z, conducts additional operations to obtain

$$\mathcal{T}^* = \sqrt{N} \cdot \max\{(R_{MC}z^*) \oslash \hat{\sigma} + \frac{1}{\sqrt{\log N}}(R_{MC}\hat{\pi} \oslash \hat{\sigma})_-, 0\}$$

. The null hypothesis of axiom satisfaction is rejected if the test statistic $\mathcal{T}$ is too large relative to the distribution $\mathcal{T}^*$.

The max function of the above expression is sensitive to the amount of inequalities there are, since having more inequalities means that more components of the variable z are used in calculating the test statistic distribution. In particular, if the $\mathcal{T}$ **statistic is the same** between a nested axiom and a nesting axiom, there is a natural tendency for the case with *more* inequalities (the nesting axiom) to be shifted to the right, and thus the axiom would be satisfied more frequently.

D. Additional Results

Table 9. Number of subjects that satisfy two choice axioms.

	Almost WARP	Regularity	Weak Regularity	Weak Binary Regularity	Weak Stochastic Transitivity	Moderate Stochastic Transitivity	Strong Stochastic Transitivity	Independence	Weak Independence	Elimination of Duplicates	Total
Almost WARP	74	74	74	74	74	74	74	74	74	74	74
Regularity	74	99	99	99	99	97	89	95	95	99	99
Weak Regularity	74	99	203	172	202	195	118	129	127	201	203
Weak Binary Regularity	74	99	172	256	255	251	142	151	149	250	256
Weak Stochastic Transitivity	74	99	202	255	314	305	165	159	157	305	314
Moderate Stochastic Transitivity	74	97	195	251	305	305	165	159	157	296	305
Strong Stochastic Transitivity	74	89	118	142	165	165	165	120	119	160	165
Independence	74	95	129	151	159	159	120	159	157	155	159
Weak Independence	74	95	127	149	157	157	119	157	157	153	157
Elimination of Duplicates	74	99	201	250	305	296	160	155	153	306	306
Total	74	99	203	256	314	305	165	159	157	306	315

Table 10. The percentage (number) of subjects consistent with models of deterministic or stochastic two-stage choice π^*.

Property Name	% (#) satisfying
WARP - Almost Deterministic	**14.3% (45)**
Regularity	**8.3%**+14.3% (**26**+45)
Weak Regularity	**41.0%**+14.3% (**129**+45)
Weak Binary Regularity	**42.2%**+14.3% (**133**+45)
Weak Stochastic Transitivity	**84.8 %**+14.3% (**267**+45)
Moderate Stochastic Transitivity	**81.0 %**+14.3% (**255**+45)
Strong Stochastic Transitivity	**53.3 %**+14.3% (**168**+45)
Independence	**24.1 %**+14.3% (**76**+45)
Weak Independence	**24.1%**+14.3% (**76**+45)
Elimination of Duplicates	**77.5%**+14.3% (**244**+45)

Tests are performed using the estimation techniques adopted from Cattaneo et al. [2020] at 95% confidence levels. Deterministic or almost deterministic choices which satisfy WARP (45 subjects in the first row) also satisfy the other stochastic properties trivially.