Journal of Economic Behavior and Organization 211 (2023) 251–269

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Random network consideration: Theory and experiment☆

Ian Chadd

Rensselaer Polytechnic Institute, Department of Economics, 110 8th St., Troy NY 12180, USA

ARTICLE INFO

Article history:
Received 24 January 2022
Revised 15 February 2023
Accepted 19 February 2023

JEL classification:
D83
D90

Keywords:
Limited consideration
Random attention
Product network

ABSTRACT

In many settings, it is natural to think of limited consideration exhibiting spillovers: attention paid to a particular alternative may "spill over" to another alternative based on shared characteristics, complementarities, features of the choice environment, shared advertising campaigns, product bundling, etc. Limited consideration of this form gives rise to new methods of revealing preferences and attention. Using a novel laboratory experiment, I test the attention properties of a deterministic Network Consideration model proposed in previous work and find a plethora of violations thereof, even at the individual level. I then propose a stochastic generalization, Random Network Consideration, and analyze its properties regarding the formation of consideration sets. When applied to the laboratory data, I find greater consistency with the general Random Network Consideration model. These results reveal that intuitively appealing features of theories of limited attention may indeed be too stringent in practice, an insight which should guide future theoretical research.

1. Introduction

The modern consumer is confronted with an abundance of options in many decision making contexts. For example, a search on Amazon for a TV between 50 and 59 in. in size yields 419 results; no consumer could be reasonably expected to meaningfully consider and evaluate all 419 of these options.[1] In some instances, limited consideration of many options is mandated. For example, students in NYC choose from among more than 500 different high school programs in a centralized procedure where they can submit a ranking *only* over 12 programs.[2] How decision makers (DMs) construct consideration sets from larger available sets is presently unknown in many contexts. One possibility is that initial consideration of some available option "spills over" into consideration for related options in the available set. For example, a worker who is offered a good job in San Francisco may search for other jobs based in San Francisco to which to apply. A shopper with a coupon for a brand of sandwich bread may further consider other brands located nearby in the grocery store aisle. Initial interest in a particular political candidate may lead to consideration of other candidates who have similar policy positions.

These attention spillovers could result from a number of channels, say, complementarities between distinct goods, shared advertising campaigns, bundling, and other mechanisms. The important point here is that, contrary to more general methods

☆ Subject incentives for the experiment contained herein are supported by National Science Foundation Doctoral Dissertation Research Improvement Grant [Award ID 1757253]. I thank Erkut Ozbay, Emel Filiz-Ozbay, Yusufcan Masatlioglu, and the participants of the Microeconomics Seminar at University of Maryland for helpful comments and advice. I also thank Amanda Statland for invaluable administrative guidance and support. Finally, this version benefited immensely from the insightful comments of the editor and two anonymous reviewers.

E-mail address: chaddi@rpi.edu

[1] Website accessed September 6, 2019.

[2] See Abdulkadiroğlu et al. (2005).

https://doi.org/10.1016/j.jebo.2023.02.020
0167-2681/

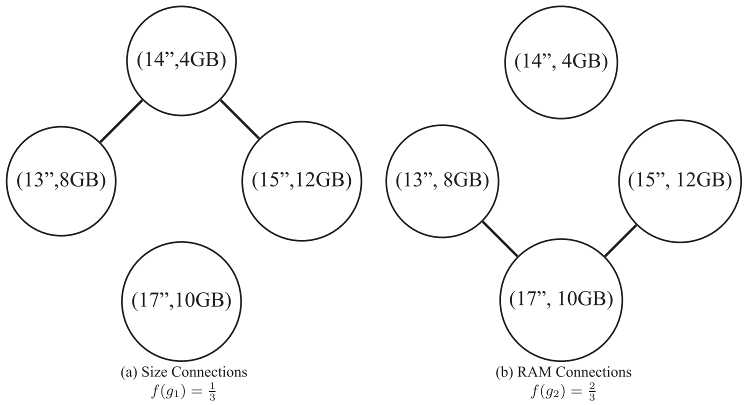

Fig. 1. Example random graph with four laptops Notes: each node represents a laptop uniquely identified by its screen size characteristic and available RAM (e.g. (13”, 8GB) represents a laptop with a 13 inch screen and 8 gigabits of RAM). In graph g_1, connections are determined by the size attribute using the criteria “x is connected to y if they differ in screen size by no more than 1 inch.” In graph g_2, connections are determined by the RAM attribute using the criteria “x is connected to y if they differ in available RAM by no more than 2 gigabits.”.

of consideration, these attention spillovers form consideration sets in a structured, precise, and consistent manner. This structure gives rise to new methods of revealing attention and/or preferences.

Indeed, there is evidence from the marketing body of literature to suggest that DMs exhibit such attention spillovers. Shapiro (2018) shows that Direct-To-Consumer advertising exhibits positive spillovers in the case of pharmaceutical anti-depressants: sales of a given drug increase by about 1.6% in response to the advertisement of a rival drug. Sahni (2016) provides experimental evidence that suggests that these positive spillovers are indeed attention-based by studying the response to online advertising in the restaurant market. Advertising a particular restaurant online can increase sales leads to a competing restaurant by around 4%.[3] Finally, Lewis and Nguyen (2015) show that online advertisements can lead to an increase in online searches for *competitors'* brands by up to 23%.

These marketing studies on attention spillovers have been focused on brand or product *categories*: the advertisement of a particular good has an effect on consideration of all goods in the same category as that which is advertised. However, two stylized facts suggest that this implicit modeling restriction may be too strong: i) Shapiro (2018) finds a variety of advertising elasticities between goods even in the same defined “category,” and ii) Sahni (2016) finds differential effects of rival advertising based on features of the firm (e.g. firm age, aggregate review scores, etc.). A general model of attention spillovers would then need to represent such spillovers as operating on a *network of connections* between options, with this category-specific treatment as a special case.

In this work, I present the results of an experiment designed to test the consideration set properties of several nested models of network consideration. To my knowledge, this is the first experimental study of a decision (i.e., non-strategic) environment with a network structure. First, the deterministic special case, studied previously by Masatlioglu and Suleymanov (2021) and called “Network Consideration” (NC hereafter), is leveraged to structure the parameterization of the laboratory experiment. NC also serves as a deterministic baseline model of attention against which to test the elicited data. The consideration set properties of NC are quite strong and I find evidence that attention, even at the subject level, is not consistent with NC in the observed data. In light of the pervasiveness of violations of NC, I suggest a more general stochastic model, which I'll call Random Network Consideration (RNC hereafter). This model shares several features with NC. First, it exhibits limited consideration, whereby the DM only considers a subset of the available set of alternatives. It also possesses a form of status quo bias where the status quo or “starting point” of the DM determines the set of alternatives that are reachable according to the random network structure. However, RNC utilizes a random network structure, as opposed to a deterministic network, as is assumed in NC. The generalization to a random network structure allows for more general consideration set mappings while still maintaining the desired dependence on a network of attention connections between options. To preview how this model works, consider the following example:

Suppose a shopper is considering the purchase of a new laptop. For simplicity, let a given laptop be fully described by two attributes: screen size and available RAM. There are four laptops available, each with a different screen size and RAM attribute: the available set is $\{(13'', 8GB), (14'', 4GB), (15'', 12GB), (17'', 10GB)\}$. In Fig. 1, these are connected to one another in a random graph structure. Let's first consider the graph on the left, g_1. This graph represents a consideration process wherein attention “spills over” from one laptop to another based on similarity in screen size: laptop x is connected to laptop y if they differ in size by no more than 1 inch. A shopper who starts their search by considering laptop $(13'', 8GB)$ would therefore also consider the14// and15// laptops, but never the17// laptop. The other graph, g_2, represents attention spillovers based on similar RAM attributes: laptop x is connected to laptop y if they differ by less than 4 gigabits of RAM.

[3] Sales leads are defined in Sahni (2016) as the consumer searching for the restaurant's phone number, which is observable in his dataset.

Thus, this same shopper who starts at the 8GB laptop will end up considering the 10GB and 12GB laptops, but never the 4GB laptop.

The crux of RNC lies in the fact that a given shopper might follow graph g_1 **or** g_2 based on the realization of some random draw: the shopper follows size connections with probability $\frac{1}{3}$ and RAM connections with probability $\frac{2}{3}$. All other possible networks of consideration occur with probability 0. Consideration (attention) under RNC is therefore *random* where, in our example, the shopper will consider the set $\{(13'', 8GB), (14'', 4GB), (15'', 12GB)\}$ with probability $\frac{1}{3}$ and $\{(13'', 8GB), (15'', 12GB), (17'', 10GB)\}$ with probability $\frac{2}{3}$ if they always start their search at $(13'', 8GB)$. The NC deterministic model is a special case of RNC where $f(g) = 1$ for some network g, and $f(g') = 0$ for all other networks g'.

Given that RNC is simply a random attention generalization of NC, it should come as no surprise that it is more consistent with the experimental data than NC, which assumes that attention is deterministic. What then is the use of this empirical exercise? First, I find that consistency with RNC consideration set properties is asymmetric. Across three classes of properties (characterized by features exhibiting monotonicity, symmetry, and connectedness, as explained in Sections 4 and 5), I find that the RNC generalization of two of these classes (symmetry and connectedness) performs significantly better than the third (monotonicity). Stated differently, while the NC property describing how consideration sets should change as the set of available options grows performs poorly, even a random attention generalization of this same property does not adequately fit the experimental data. This observation could help guide theorists attempting to model attention allocation in environments with a great deal of structure: properties that may be intuitively appealing could easily be too stringent for real-world DMs.

Second, I also present exploratory analysis to report trends in the formation of consideration sets (where both NC and RNC are silent) and the determinants of the optimality of choice. These observations may again prove useful to theorists or applied researchers interested in further developing models of attention or empirical investigations of data assuming limited attention.

This paper proceeds as follows. Related literature, both theoretical and experimental studies, is reviewed in Section 2. The experimental design and results of tests of NC are presented in Sections 3 and 4. Section 5 presents the RNC model which is tested in Section 6. The exploratory and complementary analysis mentioned above is presented in Section 7. Section 8 concludes.

2. Related literature

2.1. Experiments

The experiment contained herein is most closely connected to a growing body of literature in economics on experimental investigations of limited attention. Firstly, this experiment elicits data regarding consideration sets in a manner complementary to earlier work. Reutskaja et al. (2011) rely on eye-tracking technology to infer the consideration and search behavior of subjects.[4] Caplin et al. (2011) elicit *choice process data* as defined previously in Caplin and Dean (2011). Instead of directly observing consideration through eye-tracking technology, Caplin et al. (2011) incentivize the revelation of the path of present-best options at each point in time during which the subject is evaluating a set of options. Geng (2016) studies the impact of a status quo on attention allocation as measured by decision and consideration time. Finally, Gabaix et al. (2006) use the MouseLab coding language to investigate subject attention in a setting with attribute-level information regarding available options.

Several studies of attention and information acquisition have been devoted to testing, estimating, or informing theoretical models. Dean and Neligh (2019) present a set of experiments regarding the rational inattention model of Sims (2003, 2006), generalized in Caplin and Dean (2015), where they document consistency with a generalized model beyond the Shannon case. Chadd et al. (2021) show that the presentation of irrelevant information can affect the consideration set in a manner not predicted by extant models of limited attention. The aim of this experiment is similar to these previous studies in that it seeks to determine consistency with a model of network consideration formation.

The experimental body of literature on networks is often focused on environments where the nodes on the network are optimizing agents and not feasible options to be considered by a central DM. A number of studies exist of network games, where agents are connected to one another via a network structure (see Charness et al. (2014) for a canonical example and Choi et al. (2016) for a thorough survey of such experiments through 2015). In a similar vein, more recent studies have been focused on dynamic network formation, in which agents enter a network sequentially and choose to connect themselves to a subset of extant nodes (agents) in the network. Neligh (2020) shows that entrants to a network "vie for dominance" by connecting to many extant nodes in a manner consistent with forward-looking behavior.

While the experiments above on network formation and network games are at least nominally related to the experiment contained herein insofar as they are explorations of "networks" in economic settings, their connection to the current experiment ends there. All of the above are game-theoretic explorations of behavior in network structures, whether they be exogenously determined or endogenously determined in equilibrium. RNC is a decision theoretic model and involve no strategic interaction between multiple agents.

[4] See Orquin and Loose (2013) for a review of eye-tracking studies in decision making.

2.2. Theory

RNC is closely related to several models of path-dependent attention and choice. Masatlioglu and Suleymanov (2021) present a model of Network Choice (NC) where attention spills over between options in a given deterministic network. Kovach and Suleymanov (2021) present a Reference-Dependent Random Attention Model that bears similarities to RNC in that stochastic consideration sets, and subsequent choice, are dependent on an initial reference (i.e., starting point). Kovach and Suleymanov (2021) combines and supersedes Kovach (2018) and Suleymanov (2018). It should be noted, however, that RNC does not satisfy a central axiom of the Reference-Dependent Random Attention Model, Dominant Alternative, and is thus not nested within it.

Several other models of limited consideration are based on stochastically determined consideration set mappings. Manzini and Mariotti (2014) first explore consideration sets that are stochastically determined. In contrast to RNC, the model of Manzini and Mariotti (2014) focuses on consideration of individual options in the feasible set where each feasible option is considered with some fixed probability, allowing for violations of independence of irrelevant alternatives, as required earlier by Luce (1959). Brady and Rehbeck (2016) generalize Manzini and Mariotti (2014) by allowing the consideration probability of a given option to depend on the available set of options. Caplin et al. (2018) show that rational inattention with Shannon entropy implies the formation of random consideration sets. In more recent work, Cattaneo et al. (2020) present a "Random Attention Model" (RAM hereafter). They apply a monotonicity condition on attention rules of the following form:

For any $a \in S - T$, $G(T \mid S) \leq G(T \mid S - a)$. where $G(T \mid S)$ is the probability that the consideration set is T when S is available.[5]

RNC also shares features with a number of models that exhibit status quo bias. Note that, in accordance with the distribution over networks of options, a change in the starting point may change both consideration probabilities and, subsequently, choice in RNC. In this way, RNC exhibits a form of status quo bias akin to that explored in Masatlioglu et al. (2005), Masatlioglu and Ok (2013), and Dean et al. (2017). However, in the models presented in Masatlioglu et al. (2005) and Masatlioglu and Ok (2013), the status quo affects what is considered by the DM according to whether the status quo dominates an option, with only undominated options being considered. The status quo rules out consideration of certain options more generally in Dean et al. (2017). In contrast, the status quo (or starting point) of RNC simply affects which networks of connections may feasibly be followed in the DM's search - an assumption that is independent of preferences and which is undefined in the absence of a status quo (starting point).

Finally, at the intersection of both theory and empirics, Abaluck and Adams-Prassl (2021) develop econometric methods for demand estimation in the presence of consumer limited consideration. They consider models both with consideration independence in the style of Manzini and Mariotti (2014) and dependence of a particular form. In the dependent model, which they dub "Default-Specific Consideration" (DSC), consumers are endowed with a default option and are either "asleep" (i.e., consume the default) or "wake-up" and consider all available options. In a laboratory experiment they demonstrate that they can recover preferences in the presence of limited consideration using only choice data.

3. Experiment

In order to test the deterministic NC model, I construct a laboratory environment with several goals. First, the environment must mimic a setting where options are linked to one another via a network. Second, choices must be incentivized. Finally, I err on the side of creating a restrictive environment in order to test the NC model where it is most likely to succeed. That is, if NC fails in this context, it is not likely to succeed in a real world analogue with fewer restrictions.

3.1. General environment

A total of 107 undergraduate subjects at the Experimental Economics Laboratory at University of Maryland, College Park participated in this experiment across eight sessions[6]. On average, subjects earned \$23.63 for approximately 90 minutes of time spent in the lab.[7]

It is helpful to consider the experiment from the perspective of a subject. The subject faces 31 extended problems, each defined by a starting point x and a set of available options S. For each extended decision problem (x, S), the subject's task is to select the option with the highest value among the ones they consider. The subject's payoff for that problem is simply the value of that chosen option, converted to cash.

Similar in spirit to the example presented in the introduction, each option in the experiment is described by four separate attributes: Shape, Pattern, Size, and Number. The value of an option is simply the sum of the value of its attributes,

[5] Manzini and Mariotti (2014), Brady and Rehbeck (2016), Kovach and Suleymanov (2021), Caplin et al. (2018), and Cattaneo et al. (2020) are not the only examples of random attention models. See Cattaneo et al. (2020) for a full review of random attention models and their connection to RAM, of which RNC is a starting-point contingent special case.

[6] See Table A1 for subject demographics.

[7] A single pilot session was conducted to test the experimental interface and receive feedback regarding clarity of the instructions. These subjects were excluded in all analysis.

Table 1
Option Example.

Option X			
Shape	Pattern	Size	Number
		SMALL	5

denominated in Experimental Currency Units (ECUs). Each attribute can take on one of 5 values, from 1 ECU to 5 ECU, resulting in 625 possible options, with values ranging from 4 to 20 ECU.[8] For clarity, consider the option displayed in Table 1.

Option X in Table 1 is described by 4 attributes: Square, Two-Bar Pattern, Small, and 5. These are worth 2 ECU, 3 ECU, 2 ECU, and 5 ECU, respectively. Then the value of Option X is 12 ECU ($= 2 + 3 + 2 + 5$). Deciding which option has the highest value in any extended decision problem is thus non-trivial, since it requires i) associating an attribute with its value per the payoff table provided in the instructions and ii) calculating the resultant option value from the sum of its attribute values.

At the start of each extended decision problem, the subject is first shown attributes for a single option (i.e., the starting point) and no other available option. In order to navigate to information for another option, the subject can utilize two lists on their screen: i) a list of “Linked Options” and ii) a list of “Options Already Viewed.” The list of “Options Already Viewed” simply lists the options within the available set for which the subject has already viewed attribute information (defined as having navigated previously to the option information page for that option). To see information for an option other than the one currently displayed, the subject clicks on the option label in one of these lists and then clicks a box labeled “View the Selected Option.”

The list of “Linked Options” displays a list of options that are said to be “linked” to the currently displayed option. An option is said to be “linked” to another if the two share two or more attributes. Thus, for Option X in Table 1, if another option had the Shape attribute “Square” and the Number attribute “5,” it would be included in this list of linked options. An option that only shared one attribute, but no more, with Option X would not be included in this list. It is through this method that the design induces an exogenous network structure on the set of available options.

This system of “linking” options to one another was chosen for two reasons. First, in order to mimic a real-world environment where NC may be an appropriate model, the experiment necessitated an exogenous network of some form. Second, this particular exogenous network structure was chosen over a more conservative alternative in order to avoid potential subject confusion or experimenter demand effects. In an alternative design where “links” between options were agnostic of option attributes, subjects may ask themselves why providing the network structure is necessary in the first place. This could lead to confusion or the appearance of deception. The chosen network structure is both easy to understand and mimics real world scenarios where we might believe that NC is the correct model for individual choice.

It is through this navigation process that I argue the current design properly incentivizes revelation of the consideration set for each extended decision problem. In effect, navigating from one option to another “uncovers” hidden information in the extended decision problem regarding the attribute information for each option. I view the design used herein as complementary to approaches incorporating MouseLab, eye-tracking, and choice-process elicitation procedures discussed in Section 2.1.

In the baseline version of this experiment, “linked” options were displayed in a list without any additional information regarding these options. For robustness, a variation of this display method was used for half of the sessions. In this variation, the full list of “linked” options was split into four lists, one for each attribute. The option linked to the currently displayed option was then displayed in the lists for the attributes that it shared with the currently displayed option. The goal of using this variation is to determine whether consistency with NC was dependent on arguably minor features of the laboratory environment. Wherever relevant, I use “Baseline” to refer to the original contextless display and “Context” to refer to those observations that came from the variant with more context provided as to the source of the link between options.

Each extended decision problem has a time limit of 75 seconds,[9] and the subject can choose to stop viewing information at any time prior by clicking a “Stop” button located at the bottom of the interface. If this is done, the subject may not view any additional information for options and may not further alter their provisional choice. Stopping the extended decision problem does not allow the subject to immediately move on to the next extended decision problem, however; they must wait for the entirety of the 75 allotted seconds to pass before moving on. This design was chosen to disincentivize haphazard choices on the part of the subject in the interest of finishing the experiment early.

At the end of the experiment, subjects filled out a post-experiment questionnaire.[10] One extended decision problem is chosen at random for each subject and they are paid based their earnings for that problem only. Subjects do not know

[8] A full table of attribute values was included in the instructions and can be found in the Online Appendix.

[9] Specifically, subjects could view options and make changes to their provisional choice for the entirety of the 75 second period. At the end of each period, a “decision timeg was chosen randomly from a uniform distribution from 2 to 75 seconds and the provisional choice held by the subject at that decision time was treated as the chosen option for that period. This procedure incentivizes the collection of so-called *choice process data* in the style of Caplin et al. (2011), though the investigation of this choice process data is beyond the scope of the current project.

[10] Subjects were asked for age; gender; self-reported ACT, SAT, and GPA scores; native language; and major of study. They were also given the opportunity to explain their decisions and indicate whether they felt they sufficiently understood the instructions to the experiment.

which extended decision problem will be chosen when making decisions, so they are incentivized to treat each decision as if it is the one for which they are paid.

3.2. Data generation process

For each extended decision problem, both the set of available options and the starting point were chosen intentionally to create explicit tests of properties of consideration sets in NC. This design was chosen to ensure that there would be a sufficient number of tests of each consideration set property. One alternative design would have randomized the extended decision problems presented to subjects. With four attributes, each taking on one of five different values, the grand set of alternatives is of size 625, with $2^{625}-1$ unique non-empty subsets. With such a large dataset over which to randomize, it would be highly unlikely that the final dataset would end up with a sufficient number of tests of the properties of NC using a reasonable number of laboratory subjects. Extended decision problems were thus chosen such that the observations gleaned from each would constitute, at minimum, one test of some axiom of NC. I provide additional information about how these extended decision problems were constructed in Section 4 when each consideration set property of NC is tested.[11]

4. Results: NC

Given that the focus of this project is on consideration set formation in the presence of a network structure, I focus in this section on analyzing consistency of the consideration set data elicited in the experiment with properties of the NC consideration sets. That is, I ask whether consideration sets satisfy Upward Monotonicity, Symmetry, and Path Connectedness.[12] I find that all three properties are violated with great frequency, both in the aggregate and at the subject level (where possible). The prevalence of violations does not in general depend on whether the subject participated in the Baseline or Context treatments.[13]

4.1. Upward monotonicity

For some extended decision problem (x, S), let $\Gamma_x(S)$ be the set of all options in S which are considered when x is the starting point. Then the NC property of Upward Monotonicity is as follows:

A1 **Upward Monotonicity:** $\Gamma_x(T) \subseteq \Gamma_x(S)$ for all $T \subseteq S$

In essence, this property describes the process of aggregation across nested extended decision problems. Under the deterministic NC model, if the DM faces (x, T), they will consider all of those options which are "reachable" from the starting point x and are also in T.[14] Then when the DM is confronted with (x, S) for $T \subseteq S$, all those options which were reachable from x and in T remain reachable under (x, S) (i.e., nothing about the underlying connections between options has been changed) and should therefore still be considered under (x, S).

In order to test this property in the lab, I define five extended decision problems that are "nested" within one another. Let (x, A_i) be one of these five extended decision problems where the option x is the starting point and set A_i is the available set. Each A_i was then chosen such that $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_5$. The starting point x was chosen such that $x \in A_1$. A violation of Upward Monotonicity would then be an observation of a realized consideration set such that $\Gamma_x(A_i) \not\subseteq \Gamma_x(A_{i+1})$. These same five available sets are presented to subjects twice, but with two different starting points (in order to test Symmetry, as explained in the following subsection). The experimental dataset therefore includes 20 tests of Upward Monotonicity per subject (2140 tests in total).

In the aggregate, 79.8% of these observations were inconsistent with Upward Monotonicity. Moreover, an analysis of the CDF of the proportion of Upward Monotonicity violations (MV) per subject, displayed in Fig. 2, reveals that nearly 50% of subjects had more than approximately 80% of their observations in violation of Upward Monotonicity. No subjects had fewer than 20% of their observations in violation of Upward Monotonicity. Taken together, these results suggest that Upward Monotonicity may be too strong an assumption on how consideration sets are formed in the presence of an exogenous network, even at the individual level.

Given the prevalence of violations of Upward Monotonicity, I ask what characteristics of extended decision problems are correlated with these violations. First, note that Upward Monotonicity does not take into account the "distance" between the

[11] While testing these properties ends up only requiring 17 unique extended decision problems, an additional 14 (for a total of 31 for each subject) were constructed to test NC choice axioms, along with the choice axioms of a related model contained in Suleymanov (2018) (now superseded by Kovach and Suleymanov (2021)). Because the focus of this paper is on consideration set formation only, these choice axiom tests are excluded.

[12] Both NC and RNC also require that the DM choose a $\succ$-maximal option, for some transitive and antisymmetric preference relation $\succ$, from the consideration set once it is formed. Since the focus of this project is on consideration set formation, tests of choice optimality are excluded from the main results. Nevertheless, I explore the prevelance and determinants of optimal choice in exploratory analysis presented in Section 7.

[13] Analysis of these treatment effects is therefore relegated to the Online Appendix.

[14] Masatlioglu and Suleymanov (2021) say that y is "reachable" from x if there exists a path from x to y in the network of connected options. I use the same terminology here.

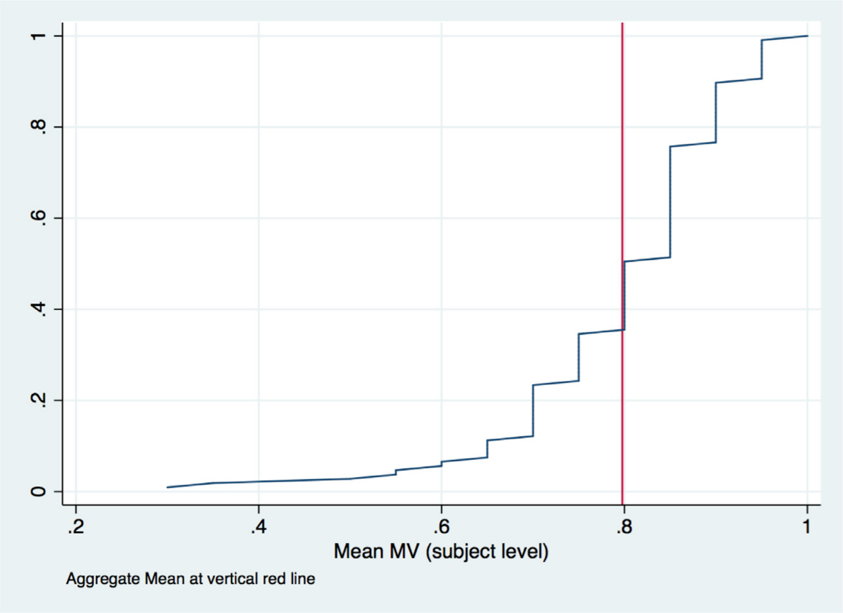

Fig. 2. Cumulative Distribution of Mean Upward Monotonicity Violations (MV) by Subject.

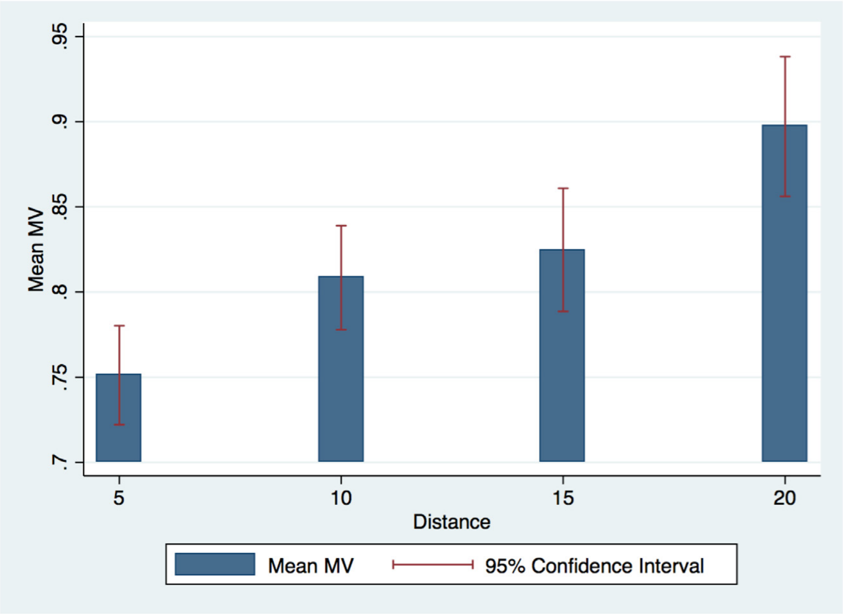

Fig. 3. Upward Monotonicity Violations (MV) by Distance ($|A_j| - |A_i|$, for $A_i \subseteq A_j$).

sizes of the relevant available sets: search is exhaustive within the sub-network connected to the starting point. Empirically, it may matter whether the relevant available sets differ greatly in their size.

Fig. 3 displays the frequency of violations for each distance between relevant available set sizes. For extended decision problems (x, A_i), for $A_i \subseteq A_j$, I define the Distance between A_i and A_j to be equal to $|A_j| - |A_i|$. On the whole, I find that as Distance increases, so too does the likelihood of observing a violation of Upward Monotonicity (though there is considerable overlap in 95% confidence intervals for these categories).

Table 2
NC Symmetry Subject Level Summary Statistics.

	Mean	SD	Min	Max	N
Hypothesis N	3.748	(1.237)	0	5	107
Symmetric N	0.589	(0.672)	0	3	107
Violation N	3.159	(1.175)	0	5	107

4.2. Symmetry

The Symmetry property of NC consideration sets is primarily concerned with the "undirectedness" of how options are connected: if option y can be reached when the starting point is x, then option x should be reached when the starting point is y. This is the result of the assumption that networks under NC are *undirected*. This has a clear implication for how consideration sets should compare across the same available set, but given distinct starting points:

A2 **Symmetry:** If $y \in \Gamma_x(S)$, $\Gamma_x(S) = \Gamma_y(S)$ for all S.

To test Symmetry, I present to subjects the five A_i available sets constructed to test Upward Monotonicity two times: once with x as the starting point and once with y as the starting point. A violation of Symmetry would therefore appear as $y \in \Gamma_x(S)$, but $\Gamma_x(S) \neq \Gamma_y(S)$. This results in ten possible tests of Symmetry for each subject: for each $(x, A_i) - (y, A_i)$ pair of extended decision problems, we can write two statements of Symmetry to be tested in the data:

$$\Gamma_x(A_i) = \Gamma_y(A_i) \text{if } y \in \Gamma_x(A_i) \tag{1}$$

$$\Gamma_y(A_i) = \Gamma_x(A_i) \text{if } x \in \Gamma_y(A_i) \tag{2}$$

These two conditions are clearly interrelated. If both the hypothesis and implication of condition 1 are satisfied for some observation, then so will those of 2, and vice versa. However, if the hypothesis is not satisfied in one, it is possible that the other test may fail. In order to rule out double-counting successes (and failures), in all of the following I exclude tests of condition 2 unless condition 1 is satisfied only trivially (i.e., $y \notin \Gamma_x(A_i)$). I thus only include a maximum of five tests of Symmetry per subject.

Out of a possible maximum of 535 tests of Symmetry, 401 were such that the hypothesis was satisfied. In the aggregate, 84.29% of these observations violated Symmetry (Wilcoxon signed-rank $p < .001$ for $H_0 : \mu = 0$). Table 2 presents summary statistics for subject-level data on the number of tests per subject (Hypothesis) and violations of Symmetry at the subject level. Of a total of 5 possible tests of Symmetry per subject, subjects satisfied a hypothesis of conditions 1 or 2 for 3.75, on average. Notably, the maximum number of symmetric observations for a given subject is 3 (out of 5 tests). Furthermore, from Fig. 4 we can see that nearly 50% of all subjects violated Symmetry in each observation where the hypothesis was satisfied.

4.3. Path connectedness

The final property of NC to be tested concerns the impact of an option that uniquely provides a connection between two other options. This property, Path Connectedness, essentially states that the revelation that some option y is required to make z reachable from x should also reveal i) that y is reachable from x in the absence of z and ii) that z is reachable from y in the absence of x. Formally, Path Connectedness is written as follows:

A3 **Path Connectedness:** If $z \in \Gamma_x(S)$ and $z \notin \Gamma_x(S \setminus y)$, then $y \in \Gamma_x(S \setminus z)$ and $z \in \Gamma_y(S \setminus x)$

This property is best understood through a simple example. In Fig. 5, clearly $z \in \Gamma_x(\{x, y, z\})$, but $z \notin \Gamma_x(\{x, z\})$; the only connection between x and z passes through y. This tells us two pieces of information. First, y must then be connected to x, independent of z. Similarly, y must then be connected to z, independent of x. So, we can then say i) $y \in \Gamma_x(\{x, y\})$ and ii) $z \in \Gamma_y(\{y, z\})$, as stated in the property above.

Path Connectedness involves four separate extended decision problems: (x, S), $(x, S \setminus y)$, $(x, S \setminus z)$, and $(y, S \setminus x)$. Furthermore, note that the hypothesis will be endogenously determined by consideration data: it may be the case that a subject makes choices in both (x, S) and $(x, S \setminus y)$ and considers some option z in both problems, thus violating the hypothesis. In order to increase the probability that there are sufficiently many tests, two separate options for y and z in the above are presented to each subject holding x and S fixed, resulting in seven extended decision problems constructed to test Path Connectedness.

The dataset includes 396 possible tests of Path Connectedness.[15] In the aggregate, only about 24.5% of possible tests were such that the hypothesis of Path Connectedness was satisfied, making for 97 total tests used. Of these 97 tests, roughly 46.4% were consistent with Path Connectedness, as reported in Table 3.

[15] By design, there were 4 tests for each of 107 subjects. In one session, a program bug in the experimental software caused the loss of two of these tests for 16 subjects (32 total tests excluded).

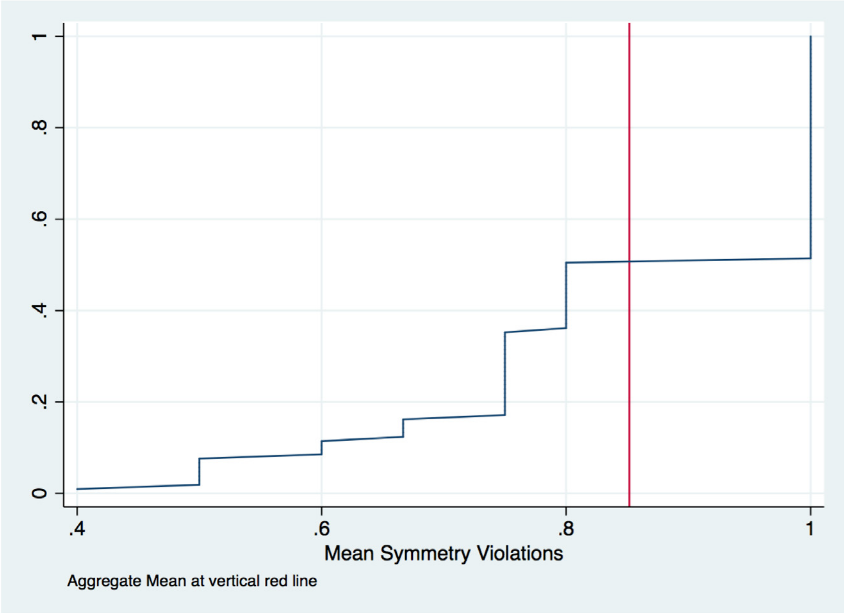

Fig. 4. Cumulative Distribution of Mean Symmetry Violations per Subject.

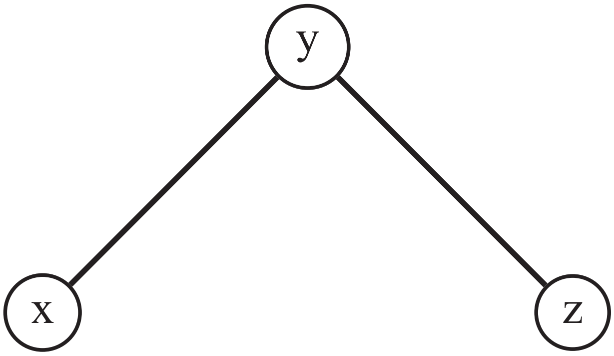

Fig. 5. Example graph with three options.

Table 3
Aggregate Test of NC Path Connectedness.

	Path Connectedness
Mean	0.464
Std Error	0.051
N	97

Wilcoxon signed-rank $p < .001$ for aggregate test of $H_0 : \mu = 1$.

These results, taken together, are largely inconsistent with the consideration set properties of NC. I summarize this as follows:

Result 1. Empirical consideration sets are largely inconsistent with the deterministic NC model:

- 79.8% of tests violate Upward Monotonicity.
- 84.3% of tests violate Symmetry.
- 53.6% of tests violate Path Connectedness.

5. A random attention generalization

Given that the experimental data is largely inconsistent with the deterministic NC model, I propose a stochastic generalization to be tested against the same dataset. I propose a generalized Random Network Consideration (RNC) model and discuss necessary properties that this model imposes on stochastic consideration set mappings. These properties are directly related to the deterministic properties of NC. I should note that this theory is meant only to provide a suitable modeling alternative to NC that may be consistent with the experimental dataset. Masatlioglu and Suleymanov (2021) fully characterize choice, providing both necessary and sufficient conditions for choice to be represented by NC. Because the focus of this paper is on consideration set formation only, such an axiomatic characterization of RNC is beyond the scope of the current work. I focus instead on necessary conditions of consideration set mappings to be tested against the experimental data.[16]

5.1. Random network consideration

Let X be a finite set of alternatives with $\Omega = 2^X \setminus \emptyset$ as the set of all non-empty subsets of X. I consider random networks on X. To that end, let $g = [X, E, \psi]$ be a network consisting of a set of nodes (alternatives) X, edges E, and an incidence function ψ such that $\psi(i, j) \in \{0, 1\}$ indicates whether nodes i and j have an edge between them ($\psi(i, j) = 1$) or not ($\psi(i, j) = 0$). I restrict ψ to be such that $\psi(i, i) = 1$ for all $i \in X$ and $\psi(i, j) = \psi(j, i)$. In other words, I restrict attention in this model to undirected networks on X, as in NC. In slight abuse of notation, I'll use g_{ij} to refer to the value of $\psi(i, j)$ associated with network g, such that $g_{ij} = 1$ indicates that nodes i and j have an edge between them under g (and vice versa for $g_{ij} = 0$). Let $\mathcal{G}$ be the set of all possible networks on X.

To consider subsets $S \in \Omega$, I must restrict attention to only those nodes that are in S. Abusing notation slightly, let $g[S]$ be a *node-induced sub-network* of g such that $g[S] = [X, E[S], \psi[S]]$ where $E[S]$ is simply the edge-set E minus all edges that involve any nodes $i \notin S$ and $\psi[S]$ is defined as follows:

$$\psi[S](i, j) = \begin{cases} \psi(i, j) & \{i, j\} \subseteq S \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Given a network g, i and j are **connected** under g if there exists an $i - j$ path in g. That is, they are connected if there exists a sequence $(x_0, x_1, \ldots x_n)$ with $x_0 = i$ and $x_n = j$ where $g_{x_k, x_{k+1}} = 1$ for all x_k and x_{k+1}. Using this terminology, the definition of what it means for a given subset of nodes to be connected under some network g directly follows:

Definition 1. A network $g \in \mathcal{G}$ is said to be **T - Connected** for some set $T \in \Omega$ if

1. t and t' are connected under g for all $t, t' \in T$ with $t \neq t'$
2. t and t' are not connected under g for each $t \in T$ and $t' \in X \setminus T$

In other words, a network g is $T-$Connected if all of the elements in T are connected to one another under g and *no* element of T is connected to a path that leads out of T.

Given an extended problem (x, S) consisting of an alternative set S and starting point $x \in S$, the DM forms a consideration set stochastically. In contrast to NC, RNC utilizes a *random network* structure to form consideration sets. Presented with a set of alternatives and a starting point, the DM will follow a single network of connections between alternatives that is realized from some distribution over the set of all possible networks. To this end, let $f(g)$ denote the probability that network g occurs.

If the DM follows this random network consideration procedure, consideration probabilities will have a specific structure and will be explicitly related to the distribution f and network of connections g. To this end, I define a *random network consideration set mapping*, denoted G_x, where $G_x(T \mid S)$ gives the probability that T is the consideration set when the set S is available and x is the starting point. For ease of notation, let $\mathcal{G}_T^S = \{g \in \mathcal{G} \mid g[S] \text{is T-Connected}\}$. The definition of a *random network consideration set mapping* is as follows:

Definition 2. Given a distribution f on $\mathcal{G}$, a **random network consideration set mapping** is a function $G_x : \Omega \times \Omega \to [0, 1]$ such that the following is true:

$$G_x(T \mid S) = \begin{cases} \sum_{g \in \mathcal{G}_T^S} f(g) & \text{if } \{x\} \subseteq T \subseteq S \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

A random network consideration set mapping is best understood as being constructed according to a sequential process. First, given S, restrict attention of networks to $g[S]$. This is done to include those networks in $G_x(T \mid S)$ that are *not* T-Connected only due to some element $t' \in X \setminus S$. Second, among all $g[S] \in \mathcal{G}$, consider those that are T-Connected, further restricting attention to $\mathcal{G}_T^S \subseteq \mathcal{G}$. Finally, given these networks that connect set T under available set S, the probability that T is the consideration set is simply the sum of the probabilities of each network occurring.

[16] Nevertheless, an interested reader may consult the Online Appendix for additional features of RNC beyond consideration set mapping necessary properties. These include revealed preference analysis and connections to the general Random Attention Model of Cattaneo et al. (2020).

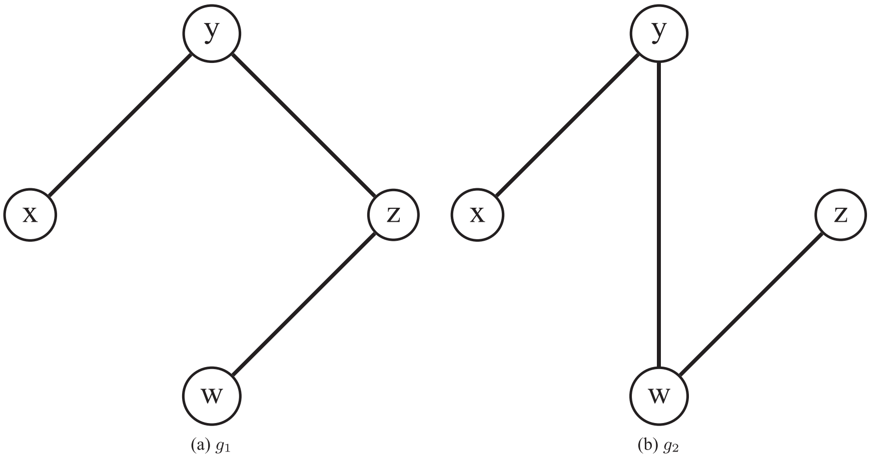

Fig. 6. An RNC Upward Monotonicity Example.

5.2. RNC necessary properties

For the purposes of this project, we focus on the properties of random network consideration set mappings themselves, exploring the means through which consideration sets are formed. We first look at a natural implication of the definition of T-Connectedness for some network $g[S]$. Consider both $g[S]$ and $g[S \cup \{a\}]$, for $a \notin S$. A network $g[S]$ that is T-Connected for some $T \subseteq S$ may or may not remain T-Connected under $S \cup \{a\}$: a may or may not connect to some $t \in T$. What is certain, however, is that all of the elements in t remained connected to one another when this new element is added. This is formally stated in the following Lemma.

Lemma 1. *For any g such that $g[S]$ is T-Connected for some $T \subseteq S$, $g[S']$ is T'-Connected for some unique T' such that $T \subseteq T'$ and $T' \subseteq S'$, for all $S \subseteq S'$. Equivalently, $\mathcal{G}_T^S \subseteq \bigcup_{T' \subseteq S' : T \subseteq T'} \mathcal{G}_{T'}^{S'}$ for all $S \subseteq S'$.*

Proof. The proof is straightforward and comes from the definitions of $g[S]$ and T-Connectedness. Recall that $g[S] = g - \sum_{\{i,j\} \subseteq X \setminus S} g_{ij}$. If $g[S]$ is T-Connected, by definition, there exists a $t - t'$ path in $g[S]$ for all $t, t' \in T$, $t \neq t'$. Each of these paths survives in $g[S']$ for some $S' \supseteq S$, since $g[S'] = g[S] + \sum_{i \in S'} \sum_{j \in S' \setminus S} g_{ij}$. Therefore, each $t, t' \in T$ is connected under $g[S']$.

Let T' be the largest set of nodes in S' such that each $t, t' \in T'$ is connected under $g[S']$ and $T' \supseteq T$. Clearly $T' \neq \emptyset$, since T itself is connected under $g[S']$ by the above logic. Then $g[S']$ is T'-Connected.

To show that T' is unique, suppose first that it isn't and consider $T'' \subseteq S'$, but $T'' \neq T'$. Note that $T \subset T''$, by construction, so either i) $\exists t'' \in T''$ such that $t'' \notin T'$ or ii) $T'' \subseteq T'$. Suppose it is case (i), then T' was not the largest set of nodes in S' such that each $t, t' \in T'$ is connected under $g[S']$, since t'' is connected to some $t \in T$ (by $T \subseteq T''$) and $T' \cup \{t''\} \supseteq T'$, a contradiction. Next, suppose it is case (ii), then $g[S']$ is not T''-Connected, since $\exists t'' \in T' \setminus T''$ that is connected to some $t \in T \subseteq T''$ by construction, which is a contradiction. Therefore, T' is unique. □

This leads us to our first characteristic of random network consideration set mappings:

B1 **RNC Upward Monotonicity** For each $x \in T \subseteq S \subseteq S'$, the following must be true:

$$\sum_{T' \subseteq S : T \subseteq T'} G_x(T' \mid S) \leq \sum_{T'' \subseteq S' : T \subseteq T''} G_x(T'' \mid S')$$

RNC Upward Monotonicity is a direct result of Lemma 1. If a network is T-Connected under set S (i.e. it is part of the sum that makes up $G_x(T \mid S)$), then by Lemma 1, that same network is T''-Connected for some $T'' \subseteq S'$ with $S \subseteq S'$. That same network will appear as part of the sum that makes up some (unique) $G_x(T' \mid S')$. This logic then applies for all $T' \supseteq T$. In short, if a network is included on the left-hand side of B1, it will show up on the right-hand side as well. To see why this expression does not hold with equality, consider the example network in Fig. 6:

Then when $S = \{x, y, z\}$ is available, the set $T = \{x, y, z\}$ is the consideration set with probability $f(g_1)$ when the starting point is x. However, when $S' = \{w, x, y, z\}$ is available, the probability that some $T'' \supseteq \{x, y, z\}$ is the consideration set is $f(g_1) + f(g_2)$. Under g_2, node z is connected to x and y *through* node w. When w is removed, z cannot connect to x or y under g_2, resulting in $f(g_2)$ being included on the right-hand side of B1, but not the left hand side when $T = \{x, y, z\}$. Notice that, in this example, had we considered $T = \{x, y\}$ for $S = \{x, y, z\}$ and $S' = \{w, x, y, z\}$, the expression would have held with equality.

This property is clearly a stochastic generalization of the Upward Monotonicity property of Masatlioglu and Suleymanov (2021). When we restrict attention to degenerate distributions over networks such that $f(g) \in \{0, 1\}$, B1 is equivalent to A1. This, along with the relationship between other properties of RNC and NC, will be further explored in the proof to Proposition 1.

I now consider the effect of changing the starting point on consideration set probabilities. When switching from x to y while both are in T, we reveal a fundamental characteristic of random network consideration:

B2 **RNC Symmetry:** $G_x(T|S) = G_y(T|S)$ for all $\{x, y\} \subseteq T \subseteq S$

This comes from the straightforward observation that $\mathcal{G}_T^S$ does not depend on the starting point and will be the same for all $t \in T$. We are thus summing over the same set of networks, resulting in the same consideration probabilities for each $t \in T$ as the starting point.[17]

In a similar fashion to the RNC Upward Monotonicity property above, RNC Symmetry is a generalization of NC Symmetry in the deterministic case. In the Symmetry property for NC, the inclusion of y in $\Gamma_x(S)$ indicates that y is connected to x when S is available. When we change the starting point to y, this connection remains. Contrary to B2, Symmetry in NC restricts the DM to follow the same sub-network of consideration on S in both extended decision problems (x, S) and (y, S), such that not only x, but the entirety of $\Gamma_x(S)$ must be considered in (y, S), since we know that y and x exist on the same sub-network. In RNC, it is not required that the *same* sub-network be followed by the DM to construct the consideration set in both extended decision problems. The only requirement is that the probability of a given sub-network occurring does not depend on the starting point, conditional on the starting points being included in that sub-network.

Finally, we explore what this random network structure implies about the connectedness of certain options. Consider the following: there exists some $T \subseteq S$ with $z \in T$ where $G_x(T \mid S) > 0$. This then implies that there exists some network g that connects x to z when S is available which occurs with positive probability. Now consider the removal of some element $y \in S$. If there still exists some $T' \subseteq S \setminus \{y\}$ with $z \in T'$ and $G_x(T' \mid S \setminus \{y\}) > 0$, we do not learn anything additional about how x, y, and z are connected, other than that y is not required for there to exist a path between z and x. However, if there exists no such T', we learn that y is *required* for there to be a path between x and z under g. The same will be true for any $T \subseteq S$ such that $\{x, y, z\} \subseteq T$, implying that we can express this relationship using simply the consideration probabilities of each option under S. This leads us to our final necessary property of random network consideration set mappings:

B3 **RNC Path Connectedness** Define $\gamma_x(z \mid S) \equiv \sum_{T \subseteq S: z \in T} G_x(T \mid S)$ as the probability that option z is considered when the available set is S and the starting point is x. Then RNC Path Connectedness requires that if $\gamma_x(z \mid S) > \gamma_x(z \mid S \setminus \{y\})$, then the following must hold:
(a) $\gamma_x(y \mid S \setminus \{z\}) > 0$
(b) $\gamma_y(z \mid S \setminus \{x\}) > 0$

The hypothesis, as mentioned previously, reveals that y is required to establish a connection between z and x under some network g[S]. In other words, all such paths that have x and z as terminal nodes must include y as an intermediate node. Then, we can break up one such path into its $x - y$ and $y - z$ sub-paths. The $x - y$ sub-path survives when z is removed, which means y is considered with some positive probability when z is removed (the first implication of the above). Similarly, the $y - z$ sub-path survives when x is removed, so z is considered with some positive probability when x is removed and the starting point is y.

Again, RNC Path Connectedness is a stochastic generalization of Path Connectedness in the deterministic NC model. Thus far, I have claimed that each of B1 - B3 are stochastic generalizations of consideration set properties of NC. The following Proposition captures this notion.

Proposition 1. *If $G_x(T \mid S)$ is a random network consideration set mapping such that i) $G_x(T \mid S) \in \{0, 1\}$ for all $x \in T \subseteq S$ and ii) G_x satisfies B1 - B3, then $\Gamma_x(S)$ satisfies A1 - A3 where $\Gamma_x(S) = T$ for $G_x(T \mid S) = 1$.*

Proof. Let $\Gamma_x(T \mid S)$ be a random consideration set mapping such that i) $\Gamma_x(T \mid S) \in \{0, 1\}$ for all $x \in T \subseteq S$ and ii) Γ_x satisfies B1 - B3. Let $\Gamma_x(S) = T$ for $\Gamma_x(T \mid S) = 1$

A1 Consider $\Gamma_x(S) = T$. Since G_x satisfies B1, $G_x(T \mid S) = 1 \leq \sum_{T'' \subseteq S' : T \subseteq T''} G_x(T'' \mid S')$ for $x \in T \subseteq S$ and $S' \supseteq S$. By the definition of $G_x(T'' \mid S')$, which requires that $G_x(T'' \mid S') = 1$ for some unique $T'' \subseteq S'$, there exists some T''' such that $G_x(T''' \mid S') = 1$. This, together with $\sum_{T'' \subseteq S' : T \subseteq T''} G_x(T'' \mid S') \geq 1 = G_x(T \mid S)$, ensures that $T \subseteq T'''$. Then $G_x(S) = T \subseteq T''' = \Gamma_x(S')$ and Γ_x satisfies A1.

A2 Consider $\Gamma_x(S)$ and $\Gamma_y(S)$, assuming $x, y \in S$ and that $y \in \Gamma_x(S)$. Then $\exists T, T'$ where $G_x(T \mid S) = 1 = G_y(T' \mid S)$. Since $y \in \Gamma_x(S)$, $y \in T$ and $T = T'$, since by B2 $y \in T$ implies $G_x(T \mid S) = G_y(T \mid S) = 1$. Then $\Gamma_x(S) = \Gamma_y(S)$ and Γ_x satisfies A2.

[17] It should be clear from the definition of G_x that exploring the relationship between $G_x(T|S)$ and $G_y(T|S)$ for $\{x, y\} \not\subseteq T$ is trivial since $G_x(T|S) = 0$ for $x \notin T$.

Table 4
Aggregate Test of RNC Monotonicity.

	All	NT
Mean	0.479	0.693
Std Error	0.005	0.005
N	8927	6169

Wilcoxon signed-rank $p < .001$ for aggregate test of $H_0 : \mu = 0$ for both All and NT.
Mann-Whitney $p < .001$ for $H_0 : \mu_{All} = \mu_{NT}$.
NT results exclude observations where $\sum_{T' \subseteq S : T \subseteq T'} G_x(T' \mid S) = 0$.

A3 Let $z \in \Gamma_x(S)$ and $z \notin \Gamma_x(S \setminus \{y\})$. Then $z \in T$ for T such that $G_x(T \mid S) = 1 > \sum_{T \subseteq S : z \in T} G_x(T \mid S \setminus \{y\}) = 0$ (otherwise there exists some $T' \supseteq \{z\}$ where $G_x(T' \mid S \setminus \{y\}) = 1$ and $z \in \Gamma_x(S \setminus \{y\})$ by the definition of Γ_x). Since G_x satisfies B3, we therefore have that $\exists T''$ with $y \in T''$ such that $G_x(T'' \mid S \setminus \{z\}) > 0$ and $\exists T'''$ with $z \in T'''$ such that $G_y(T'' \mid S \setminus \{x\}) > 0$. Then, by definition of $\Gamma_x(S)$, $y \in \Gamma_x(S \setminus \{z\})$ and $z \in \Gamma_y(S \setminus \{x\})$. Therefore, Γ_x satisfies A3. □

Finally, it should be clear at this point that RNC consideration set mappings necessarily exhibit all of the above properties. Proposition 2 captures this idea.

Proposition 2. *If G_x is a random network consideration set mapping, then it satisfies RNC Symmetry, RNC Upward Monotonicity, and RNC Path Connectedness.*

Proof. Consider some random network consideration set mapping G_x.

RNC Symmetry: Consider any $T \subset S$ with $\{x, y\} \subseteq T$ and $x \neq y$. If $G_x(T \mid S) > 0$, then there exists some network $g[S]$ that is T-Connected. Since $y \in T$, $g[S]$ will also be included in $G_y(T \mid S)$. Therefore, $G_x(T \mid S) \leq G_y(T \mid S)$. $G_y(T \mid S) \leq G_x(T \mid S)$ by the same logic. Finally, if $G_x(T \mid S) = 0$, then there are no $g \in \mathcal{G}$ such that $g[S]$ is T-Connected. This will hold regardless of the starting point in T, so $G_y(T \mid S) = 0$ as well. Thus, RNC Symmetry holds.

RNC Upward Monotonicity: Given Lemma 1, the proof is trivial. With $\mathcal{G}_T^S \subseteq \bigcup_{T' \subseteq S' : T \subseteq T'} \mathcal{G}_{T'}^{S'}$ for all $S \subseteq S'$, the statement follows directly from the definition of $G_x(T \mid S)$.

RNC Path Connectedness: For some $\{x, y, z\} \subseteq S$, let $\gamma_x(z \mid S) > \gamma_x(z \mid S \setminus \{y\})$. Then there exists some $g[S] \in \mathcal{G}_T^S$ where $f(g[S]) > 0$; $g[S]$ is T-Connected for some $T \supseteq \{z\}$; and $g[S \setminus \{y\}]$ is not T'-connected for any $T' \supseteq \{z\}$ and $T' \subseteq S \setminus \{y\}$. Then every path that connects x to z under $g[S]$ must include y as an intermediate node. To see why this is the case, consider some $x - z$ path in $g[S]$ that does *not* include y as an intermediate node. When y is removed from S, this path remains, since y was not on this path under $g[S]$ and $g[S \setminus \{y\}]$ would be T'-Connected for some $T' \supseteq \{z\}$, a contradiction. Since there exists an $x - y - z$ path in $g[S]$, we can consider each sub-path independently.

Consider the $x - y$ sub-path. When z is removed from S, this path survives and $g[S \setminus \{z\}]$ is T''-Connected for some $T'' \supseteq \{x, y\}$. Then $\gamma_x(y \mid S \setminus \{z\}) > 0$.

By similar logic for the y- z sub-path, $\gamma_y(z \mid S \setminus \{x\}) > 0$ and G_x satisfies RNC Path Connectedness. □

6. Results: Random network consideration

In this section, tests of the more general stochastic properties of RNC are conducted. For each test, in order to generate consideration probabilities, observations are aggregated over all subjects, treating each observation as if it came from a representative subject who encountered each problem multiple times. Thus, for a given extended decision problem (x, S) and for some consideration set $T \subseteq S$, $\Gamma_x(T \mid S)$ was set to be equal to the frequency of consideration set T observed in the full data set, conditional on the extended decision problem being (x, S).

6.1. RNC monotonicity

For each T observed with strictly positive probability, RNC Monotoncity is constructed by comparing the sum of probabilities over supersets of T when S was presented $[\sum_{T' \subseteq S : T \subseteq T'} \Gamma_x(T' \mid S)]$ to the sum of probabilities over supersets of T for some presented superset of S $[\sum_{T'' \subseteq S' : T \subseteq T''} \Gamma_x(T'' \mid S')]$, offering a direct test of the RNC Monotonicity property. Table 4 presents the aggregate mean violations of RNC Monotonicity. Many consideration sets T are feasible for a given (x, S) extended decision problem, in that they are such that $T \subseteq S$, but they do not occur with positive probability. Then RNC Monotonicity will be satisfied trivially for these sets. While these observations are technically consistent with RNC Monotonicity, they are excluded in the column labeled "NT", for "Non-Trivial," in Table 4. In the aggregate, 47.9% of all observations result in a violation of RNC Monotonicity, compared to 79.8% when testing against the NC model. Even when only considering "Non-Trivial" observations, the rate of Monotonicity violations is lower under RNC than under NC at 69.3% (Wilcoxon signed-rank

Table 5
RNC Symmetry: Aggregate.

	All
Choose	
Starting Point	-0.0862
	(0.0848)
Observations	38,082

Standard errors in parentheses.
Odds ratios from conditional logit regression specifications.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6
RNC Symmetry Regressions.

	N = 5	N = 10	N = 15	N = 20	N = 25
Choose					
Starting Point	-0.252	-0.0155	-0.155	-0.0216	0.134
	(0.168)	(0.176)	(0.186)	(0.208)	(0.232)
Observations	776	5100	13,448	11,088	7670

Standard errors in parentheses.
Odds ratios from conditional logit regression specifications.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

$p < .01$ for $H_0 : \mu = 0.798$). Though consistency with Monotonicity appears to improve under the RNC generalization, we thus still see a non-trivial number of violations. This suggests that consideration set properties based on this strict version of monotonicity may, in general, be too stringent.

6.2. RNC symmetry

Consistency with RNC Symmetry is tested using two methodologies: first, by conditional logit regression estimation and second, by individual difference-in-means tests for each pair of consideration set and available set.

Initially, each observation used for the conditional logistic regression specifications consists of a subject, an extended decision problem (i.e., a starting point that is an element of $\{x, y\}$ and an available set A_i), and the set of options considered by the subject (i.e., the observed consideration set). A test of RNC Symmetry consists of a test of whether the probability of a given consideration set being observed is dependent on the starting point in $\{x, y\}$, given that the consideration set contains both starting points. To this end, for each of the five available sets used for the Symmetry extended decision problems, I treat each observed unique consideration set that includes both starting points as an "available consideration set" that the subject may then "choose". Thus, a case for the purposes of these conditional logistic regressions is defined as a subject-available set pair, with the unique consideration sets which include both starting points observed in the data for that available set *across all subjects* constituting the "available consideration sets" from which this subject can "choose." Note that, for a given case, each of these consideration sets is offered to the subject as an available consideration set *twice*: once for each extended decision problem that utilized the available set for this case. This is done to allow for the possibility that the same consideration set was chosen by an individual subject in *both* extended decision problems that utilize the available set for this case. Thus, in these conditional logit specifications, the dependent variable Choose indicates whether the "available consideration set" was "chosen" for an individual case. The lone dependent variable, Starting Point, is a binary variable that takes the value 1 when the starting point for the observation is y and 0 otherwise.

Of the 553 distinct consideration sets observed for the 10 extended decision problems constructed to test the Symmetry property, 235 were such that $\{x, y\} \subseteq T$. Results at the aggregate level are displayed in Table 5. When we aggregate over all possible available sets ($N \in \{5, 10, 15, 20, 25\}$), we see that there is no relationship between the starting point and whether a consideration set is "chosen." This result is robust to whether a separate conditional logit regression is run on each available set individually, as can be see in Table 6, indicating that the size of the available set does not affect consistency with RNC Symmetry. There are thus broad indications that consideration set formation is consistent with RNC Symmetry.

Each consideration set observed in the data individually can be examined directly by constructing $\Gamma_x(T \mid S)$ and $\Gamma_y(T \mid S)$ for each $T \subseteq S$ combination. If the frequency of T conditional on S being available is significantly different between Γ_x and Γ_y, this is a violation of RNC Symmetry. We thus conduct a Wilcoxon sign-rank test on each of the Y consideration sets where $\{x, y\} \subset T$. Note, however, that if a given T is only ever chosen once in the entire sample, a sign-rank test will result in an insignificant difference by starting point.[18] A relatively conservative approach is used here, where only those consideration sets that occur with non-trivial frequency, defined as "having occurred more than once across all 10 symmetry

[18] Consideration set-available set pairs have 214 observations, 107 for each starting point. If the consideration set only occurs once, the sign-rank test will result in $p > .10$.

Table 7
RNC Path Connectedness.

	Case 1		Case 2		Case 3		Case 4	
Panel A: Hypothesis								
$\gamma_x(z \mid \cdot)$ Difference	0.062		0.598		-0.066		0.084	
p-value	0.146		0.000		0.933		0.001	
Panel B: Implication								
	$\gamma_x(y \mid \cdot)$	$\gamma_y(z \mid \cdot)$	$\gamma_x(y \mid \cdot)$	$\gamma_y(z \mid \cdot)$	$\gamma_x(y \mid \cdot)$	$\gamma_y(z \mid \cdot)$	$\gamma_x(y \mid \cdot)$	$\gamma_y(z \mid \cdot)$
Mean	0.407	0.429	0.925	0.561	0.290	0.121	0.879	0.178
Std	0.494	0.497	0.264	0.498	0.456	0.328	0.328	0.037
N	91		107		91		107	

$\gamma_x(z \mid \cdot)$ Difference is equal to $\gamma_x(z \mid S) - \gamma_x(z \mid S \setminus \{y\})$.
$\gamma_x(y \mid \cdot)$ refers to $\gamma_x(y \mid S \setminus \{z\})$; $\gamma_y(z \mid \cdot)$ refers to $\gamma_y(z \mid S \setminus \{x\})$.
p-values for Hypothesis come from one-sided t-tests.

extended decision problems," are included in this analysis. Of the 235 distinct consideration sets observed in these extended problems that satisfy $\{x, y, \} \in T$, 183 occurred only once. The remaining consideration sets are then used to construct $\Gamma_x(T \mid S)$ and $\Gamma_y(T \mid S)$ for each S such that T appeared at least once across (x, S) and (y, S). This resulted in 69 separate tests of $T \subset S$ pairs.[19] Of these 69 tests, only 5 (7.23%) resulted in a statistically significant rejection of $H_0 : \Gamma_x(T \mid S) = \Gamma_y(T \mid S)$. Support for RNC Symmetry is therefore robust, being independent of the test method (conditional logit vs. rank-sum at the consideration set-available set pair level).

6.3. RNC path connectedness

Recall that, in the construction of the extended decision problems used to test Path Connectedness, four different cases resulted from varying the option used in each extended decision problem. Tests of RNC Path Connectedness therefore begin with the consideration of the hypothesis that $\gamma_x(z \mid S) > \gamma_x(z \mid S \setminus \{y\})$ in each case, respectively. For each case, after confirming that the hypothesis is true, RNC Path Connectedness then implies a simple test of whether $\gamma_x(y \mid S \setminus \{z\}) > 0$ and $\gamma_y(z \mid S \setminus \{x\}) > 0$.

Table 7 presents the results of these aggregate tests separately for each case. Panel A displays the difference in probabilities of consideration of option z for each relevant available set in the case. All but Case 3 are generally consistent with the hypothesis of RNC Path Connectedness in that this difference is positive (though statistically insignificant for Case 1). Panel B displays the relevant consideration probabilities of y and z for each portion of the implication of this property. We can see that in every case (even for Case 3 where the implication was not satisfied), these probabilities are positive. On the whole, this is consistent with RNC Path Connectedness.

Result 2. Aggregate consideration set frequencies exhibit mixed consistentcy with RNC:

- 47.9% of all observations violate RNC Upward Monotonicity.
- Fewer than 8% of all consideration sets observed in the aggregate data violate RNC Symmetry.
- Aggregate results are wholly consistent with RNC Path Connectedness.

7. Complementary and exploratory analysis

In this section, I explore i) additional empirical trends in the consideration set data and ii) choice optimality. None of the additional analysis using consideration set data is based on theoretical predictions from NC and RNC, but should serve as useful observations for the development of new models of attention allocation. Analysis of choice optimality is based on a shared choice rule used by NC and RNC, but is complementary to the main analysis herein since it does not directly deal with attention allocation.

7.1. Trends in consideration

Subjects consider more options i) over time and ii) when more are available.

Table 8 presents two OLS regression specifications with the size of the consideration set ($|\Gamma_x(S)|$) as the dependent variable.[20]

I replicate a version of the results contained in Reutskaja et al. (2011), that additional available options lead to more options being considered by the DM (the coefficient estimate of N, the size of the available set, being positive in both specifications). Additionally, there is evidence of learning: subjects consider more in later periods in the experiment.

[19] Note that this is greater than $235 - 183 = 52$. This is due to the fact that several consideration sets were observed under multiple available sets.

[20] In these model specifications, $\hat{\mu}^{GPA}$ is a within-sample normalization of self-reported GPA (see Cohen et al., 1999; Filiz-Ozbay et al., 2018; Chadd et al., 2021) and, for subject i, is as follows: $\hat{\mu}_i^{GPA} = \frac{\mu_i^{GPA} - \mu_{\min}^{GPA}}{\mu_{\max}^{GPA} - \mu_{\min}^{GPA}}$ where $\mu_{\max}^{GPA}$ ($\mu_{\min}^{GPA}$) is the maximum (minimum) value of GPA in the sample.

Table 8
Determinants of Consideration Set Size.

	Model 1	Model 2
N	0.217***	0.218***
	(0.011)	(0.011)
Period	0.0271***	0.0264***
	(0.008)	(0.008)
Female		0.146
		(0.466)
$\hat{\mu}_{GPA}$		2.131
		(1.418)
Constant	4.241***	2.687**
	(0.234)	(1.106)
Observations	25,811	25,811

Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

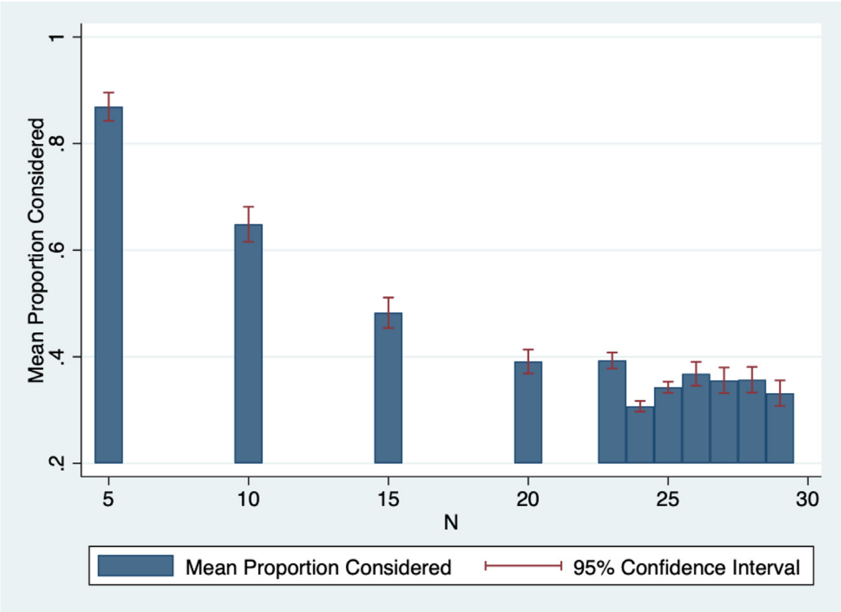

Fig. 7. Proportion Considered by N.

Subjects consider a smaller proportion *of available options when more are available.*

Fig. 7 displays the average proportion of available options considered by the size of the available set N. This runs from nearly exhaustive search for small sets (roughly 87% for N = 5; median = 100%) to very little search for large sets (roughly 33% for N = 29; median = 31%).

7.2. Choice optimality

Both NC and RNC share a common choice rule: once the consideration set is formed, the DM chooses a $\succ$-maximal option for some transitive and asymmetric preference relation $\succ$. I therefore say that a subject chose "optimally relative to the NC/RNC benchmark" if the chosen option had the highest possible ECU payoff *among those options that the subject considered*. Note that this, in general, is not equivalent to standard mistake rate analysis conducted in Caplin et al. (2011) and Chadd et al. (2021), for example. For completeness, I also test for "choice optimality relative to the full attention benchmark," meaning that the chosen option had the highest possible ECU payoff among *all available options.*

Subjects choose more optimally i) in later periods, ii) when they consider fewer options, and iii) when fewer options are available, relative to the NC/RNC benchmark.

Subjects chose optimally relative to the NC/RNC benchmark in 85.675% of problems (Wilcox $p < .001$ for $H_0 : \mu = 1$). Given that sub-optimal choice is non-trivial, I further investigate determinants through several logistic regression specifications. In Table 9, the dependent variable in models 1 and 2 is a binary variable that takes the value 1 if the subject chose chose optimally relative to the NC/RNC benchmark and 0 otherwise. In both model specifications, marginal effects from a

Table 9
Determinants of Optimal Choice.

	NC/RNC		Full Attention	
	Model 1	Model 2	Model 3	Model 4
Context	-0.016	-0.016	0.059*	0.065**
	(0.031)	(0.031)	(0.033)	(0.032)
Period	0.002**	0.002**	-0.001**	-0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
CS_N^{ϵ}	-0.013***	-0.013***		
	(0.003)	(0.003)		
N	-0.010***	-0.010***	-0.027***	-0.027***
	(0.001)	(0.001)	(0.002)	(0.002)
Female		-0.036		-0.063**
		(0.027)		(0.032)
$\hat{\mu}_{GPA}$		0.036		0.159*
		(0.080)		(0.089)
Observations	3276	3276	3276	3276

Standard errors in parentheses.
Marginal effects from logit regression specifications.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

logistic regression are reported, along with robust standard errors clustered at the subject level. Context is a binary variable used to indicate whether the observation came from the Context display. CS_N^{ϵ} is the residual generated from Model 1 in Table 8, an OLS regression of CS_N onto Period and N. These residuals are the portion of CS_N left unexplained by N and Period, and they are used in both models to estimate the effect of consideration set size on choice optimality separately from the effects of N.

From Table 9, we can see that choice optimality improves with learning and diminishes with the size of the consideration set. Additionally, the size of the available set matters; for each additional element added to the set of available options, the probability that the chosen element will be consideration set optimal decreases by approximately 1 percentage point. This may hint at some attention fatigue-induced perception issues when it comes to the actual evaluation of considered options.

Subjects choose more optimally i) in smaller available sets, ii) in earlier periods, and iii) when context is provided regarding the network of connections, relative to the full attention benchmark.

Relative to the full attention benchmark (i.e., where the DM considers all available options), subjects choose optimally in 61.86% of problems (Wilcox $p < .001$ for $H_0 : \mu = 1$). In Table 9, the dependent variable in models 3 and 4 is a binary variable that takes the value 1 if the subject chose chose optimally relative to the full attention benchmark and 0 otherwise. Subjects choose the highest valued option less frequently in i) larger available sets and ii) over time. However, the relationship between choice optimality and the size of the available set does not appear to be monotonic: the frequency of choosing the highest valued option is highest in the smallest sets (N= 5) at 94.34% and lowest in sets of medium size (N= 23) at 29.06%. Note that the largest sets made available to subjects were of size N= 29 with 48.6% of observations at the highest valued available option. Providing additional information on network of connections also appears to matter: the estimated coefficient on Context is positive and significant in both models 3 and 4. This suggests that subjects who are provided with information as to why an option is connected to another are better able to include more highly-valued options in their consideration set, increasing overall choice optimality relative to the full attention benchmark.

8. Conclusion

In this work, I explore the consistency of two models of network consideration with experimentally gleaned attention data. Notably, the consideration data that I observe is inconsistent with deterministic network consideration, even at the individual subject level. I generalize these network consideration properties by allowing attention to evolve according to random networks. This random network consideration is much more consistent with the experimental data, though violations remain. In particular, it seems as if Monotonicity assumptions governing network consideration may be too stringent, even when attention is allowed to be random. While roughly 80% of observations were inconsistent with the deterministic form of Monotonicity, still 47.9% of the analogous tests of the random form of Monotonicity were inconsistent. This is a particularly troubling finding given the intuitive appeal of Monotonicity in the context of a network of options.

There are several limitations to this study that are worth discussing. First, the experimental design included a time constraint (75 seconds) for each decision problem - a seemingly artificial constraint given that search could be exhaustive in the field. However, time constraints in decision-theoretic experiments are common (e.g., Caplin et al., 2011; Chadd et al., 2021) and often viewed as one way of inducing limited attention in the lab: indeed, consumers in the field will have time constraints on their attention, albeit often endogenously chosen. Nevertheless, it may be that the extent of violations of consideration set properties is at least partly determined by the use of these time constraints. There are several hints that this is not the case. For example, the median number of options considered when the available set has 20 or more options is 8 (75% of such consideration sets are of size 6 or higher), suggesting that the search capacity of most subjects was weakly higher. Now consider Upward Monotonicity: 8 of the 20 sets *per subject* were such that the smaller available set had only 5 options, well within this search capacity for the allotted time. Yet 71.5% of observations involving this smallest set led to a violation of Upward Monotonicity; more starkly, 49.1% of observations where the smaller set was of size 5 and the larger set was of size 10 led to a violation. Still, it remains an interesting question whether an endogenous time constraint or self-directed network search in the field would lead to more or fewer violations of network consideration properties.

Additionally, the attention model developed herein is agnostic as to *how* the network of attention is formed. The development of an explicit attribute-based network model of attention could prove useful in applied settings. Finally, the consideration set properties of both NC and RNC are based on the assumption that the underlying networks are *undirected*. In many field settings, it seems quite feasible that attention may spill over from option x to y, but not the other way around. It would thus be worthwhile to consider attention allocation over a directed network.

Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Additional Results

Table A1
Demographic Information.

	Age	SAT	ACT	GPA	Female
mean	20.682	1810.833	30.114	3.363	0.449
sd	1.552	324.116	3.947	0.439	0.500
min	18	1100	20	2	0
max	27	2360	36	4	1
count	107	84	35	107	107

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jebo.2023.02.020.

References

Abaluck, J., Adams-Prassl, A., 2021. What do consumers consider before they choose? identification from asymmetric demand responses. Q. J. Econ. 136 (3), 1611–1663.
Abdulkadiroğlu, A., Pathak, P.A., Roth, A.E., 2005. The new york city high school match. Am. Econ. Rev. 95 (2), 364–367.
Brady, R.L., Rehbeck, J., 2016. Menu-dependent stochastic feasibility. Econometrica 84 (3), 1203–1223.
Caplin, A., Dean, M., 2011. Search, choice, and revealed preference. Theor. Econ. 6 (1), 19–48.
Caplin, A., Dean, M., 2015. Revealed preference, rational inattention, and costly information acquisition. Am. Econ. Rev. 105 (7), 2183–2203.
Caplin, A., Dean, M., Leahy, J., 2018. Rational inattention, optimal consideration sets, and stochastic choice. Rev. Econ. Stud. 86 (3), 1061–1094.
Caplin, A., Dean, M., Martin, D., 2011. Search and satisficing. Am. Econ. Rev. 101 (7), 2899–2922.
Cattaneo, M.D., Ma, X., Masatlioglu, Y., Suleymanov, E., 2020. A random attention model. J. Polit. Econ. 128 (7), 2796–2836.
Chadd, I., Filiz-Ozbay, E., Ozbay, E.Y., 2021. The relevance of irrelevant information. Exp. Econ. 24 (3), 985–1018.
Charness, G., Feri, F., Meléndez-Jiménez, M.A., Sutter, M., 2014. Experimental games on networks: underpinnings of behavior and equilibrium selection. Econometrica 82 (5), 1615–1670.
Choi, S., Gallo, E., Kariv, S., 2016. Networks in the Laboratory. In: Bramoullé, Y., Galeotti, A., Rogers, B. (Eds.), Oxford Handbook of the Economics of Networks. Oxford University Press.
Cohen, P., Cohen, J., Aiken, L.S., West, S.G., 1999. The problem of units and the circumstance for POMP. Multivariate Behav. Res. 34 (3), 315–346.
Dean, M., Kıbrıs, Ö., Masatlioglu, Y., 2017. Limited attention and status quo bias. J. Econ. Theory 169, 93–127.
Dean, M., Neligh, N., 2019. Experimental tests of rational inattention. Working Paper.
Filiz-Ozbay, E., Ham, J.C., Kagel, J.H., Ozbay, E.Y., 2018. The role of cognitive ability and personality traits for men and women in gift exchange outcomes. Exp. Econ. 21, 650–672.
Gabaix, X., Laibson, D., Moloche, G., Weinberg, S., 2006. Costly information acquisition: experimental analysis of a boundedly rational model. Am. Econ. Rev. 96 (4), 1043–1068.
Geng, S., 2016. Decision time, consideration time, and status quo bias. Econ. Inq. 54 (1), 433–449.
Kovach, M., 2018. Thinking inside the box: status quo bias and stochastic consideration. Working Paper.
Kovach, M., Suleymanov, E., 2021. Reference dependence and random attention. Working Paper.
Lewis, R., Nguyen, D., 2015. Display advertisings competitive spillovers to consumer search. Quant. Market. Econ. 13 (2), 93–115.
Luce, R.D., 1959. Individual Choice Behavior: A Theoretical Analysis. John Wiley & Sons, Inc..
Manzini, P., Mariotti, M., 2014. Stochastic choice and consideration sets. Econometrica 82 (3), 1153–1176.
Masatlioglu, Y., Ok, E.A., 2013. A canonical model of choice with initial endowments. Rev. Econ. Stud. 81 (2), 851–883.

Masatlioglu, Y., Ok, E.A., et al., 2005. Rational choice with status quo bias. J. Econ. Theory 121 (1), 1–29.
Masatlioglu, Y., Suleymanov, E., 2021. Decision making within a product network. Econ. Theory 71 (1), 185–209.
Neligh, N., 2020. Vying for dominance: an experiment in dynamic network formation. J. Econ. Behav. Org. 178, 719–739.
Orquin, J.L., Loose, S.M., 2013. Attention and choice: a review on eye movements in decision making. Acta Psychol. (Amst.) 144 (1), 190–206.
Reutskaja, E., Nagel, R., Camerer, C.F., Rangel, A., 2011. Search dynamics in consumer choice under time pressure: an eye-tracking study. Am. Econ. Rev. 101 (2), 900–926.
Sahni, N.S., 2016. Advertising spillovers: evidence from online field experiments and implications for returns on advertising. J. Market. Res. 53 (4), 459–478.
Shapiro, B.T., 2018. Positive spillovers and free riding in advertising of prescription pharmaceuticals: the case of antidepressants. J. Polit. Econ. 126 (1), 381–437.
Sims, C.A., 2003. Implications of rational inattention. J. Monet. Econ. 50 (3), 665–690.
Sims, C.A., 2006. Rational inattention: beyond the linear-quadratic case. Am. Econ. Rev. 96 (2), 158–163.
Suleymanov, E., 2018. Stochastic attention and search. Working Paper.